applied sciences

Article

Lane-Changing Recognition of Urban Expressway Exit Using Natural Driving Data

Lei Zhao [1], Ting Xu [1,*], Zhishun Zhang [1] and Yanjun Hao [2]

[1] College of Transportation Engineering, Chang'an University, The Middle Section of the Second Ring Road, Xi'an 710064, Shaanxi, China
[2] Institute of Intelligent Traffic in Shanxi Province Co., Ltd., Taiyuan 30053, Shanxi, China
* Correspondence: xuting@chd.edu.cn; Tel.: +86-183-9213-0710

Abstract: The traffic environment at the exit of the urban expressway is complex, and vehicle lane-changing behavior occurs frequently, making it prone to traffic conflict and congestion. To study the traffic conditions at the exit of the urban expressway and improve the road operation capacity, this paper analyzes the characteristics of lane-changing behaviors at the exit, adds driving style into the influencing factors of lane-changing, and recognizes one's lane-changing intention based on driving data. A UAV (unmanned aerial vehicle) is used to collect the natural driving track data of the urban expressway diverge area, the track segments of vehicle lane-changing that meet the standards are extracted, and 374 lane-changing segments are obtained. K-means++ is used to cluster the driving style of the lane-changing segments which is grouped into three clusters, corresponding to "ordinary", "radical", and "conservative". Through the random forest model used to identify and predict driving style, the accuracy reaches 93%. Considering the characteristics of a single time point and the characteristics of the historical time window, XGBoost, LightGBM, and the Stacking fusion model are established to recognize one's lane-changing intention. The results show that the models can well recognize the lane-changing intention of drivers. The Stacking fusion model has the highest accuracy, while the LightGBM model takes less time; the model considering the characteristics of the historical time window performs better than the other one, which can better improve the prediction accuracy of lane-changing behavior.

Keywords: traffic safety; urban expressway exit; driving style; data driving; lane-changing intention; machine learning

Citation: Zhao, L.; Xu, T.; Zhang, Z.; Hao, Y. Lane-Changing Recognition of Urban Expressway Exit Using Natural Driving Data. *Appl. Sci.* **2022**, *12*, 9762. https://doi.org/10.3390/app12199762

Academic Editors: Yuchuan Du and Yu Shen

Received: 31 August 2022
Accepted: 26 September 2022
Published: 28 September 2022

1. Introduction

The urban expressway is an important segment of the urban roadway system. The diverging behavior of vehicles often occurs in the exit area of the urban expressway. Comparing the traffic behavior of the mainline, acceleration or deceleration, and lane-changing behaviors are frequent in the exit area, which will result in a higher risk of traffic crashes.

It is stated by the National Highway Traffic Safety Administration that 27% of accidents were caused by vehicle lane-changing [1]. Traffic safety data in recent years also shows that 23.91% of traffic accidents are caused by lane-changing [2]. Therefore, if the intentions of surrounding vehicles can be recognized in advance, traffic accidents could be avoided to some extent. The urban expressway exits have more short off-ramps than the freeway and large volumes of vehicles make the situation even worse. It is very important to accurately recognize the lane-changing behavior of vehicles to reduce the incidence of road traffic accidents.

Based on the vehicle trajectory data collected in the urban expressway diversion area, this paper extracts the characteristics that represent the driver's style and vehicle driving intention and establishes a machine learning model to identify the vehicle lane-changing intention in time, which can improve driving safety and feasibility to a certain extent.

1.1. Literature Review

Lane-changing is a common operation for drivers. It has been proposed for many years to study traffic safety by recognizing lane-changing behaviors; many studies focus on this field to test new theories, enhance the accuracy of lane-changing recognition, and reduce traffic crashes. Gipps connected the decisions made before lane-changing when analyzing the decisions made in the process of lane-changing, which makes the overall logic of the model more complete, and established one of the most classic vehicle lane-changing decision models [3]. However, the model is too idealistic to be applied in practice. The relevant scholars have optimized and improved the classical model from various angles and proposed new models, such as the MITSIM model, SITRAS model, CORISM model, etc. [4–6].

With the development of science and technology, researchers have applied new methods to lane-changing behavior analysis, including SVM, the LSTM model, the hidden Markov model, etc. Zyner and Worrall et al., proposed a long short-term memory (LSTM) network model to identify the driver's intention when the vehicle enters the intersection. The model inputs the vehicle position, heading angle, speed, and other parameters for learning and training to achieve a good recognition effect [7]. Phillips et al., collected traffic data at intersections and built an intention recognition model based on the LSTM network to predict a left turn, right turn, and straight ahead intentions [8]. Kim et al., proposed a new preprocessing algorithm for the advanced driver assistance system to improve the accuracy of identifying the driver's intention to change lanes [9]. The verification results on the driving simulator show that the recognition accuracy can be improved by combining the neural network model with the support vector machine model. Guo used the bi-directional long and short-term memory network based on the attention mechanism (AT-BiLSTM) to establish a lane-changing intention model, which improved the prediction accuracy [10]. Song Xiaoling et al., constructed the revenue function to reflect the interaction between vehicles and introduced the attention mechanism and conditional random field to build an LSTM model to identify the vehicle lane-changing intention [11]. Through comparison, it is found that the performance of this model is better than that of the SVM and hidden Markov model.

At the same time, some researchers have developed various methods to improve the effect of identifying vehicle lane-changing behavior. Zhang Mingfang et al., established a combined HMM and Bayesian model to identify the discrete behavior of intersections, including common lane-changing, emergency lane-changing, and turning [12]. The training model parameters include the steering angle, angular velocity, and point cloud data. Li Keqiang et al., proposed another new algorithm, which combines HMM and Bayesian filtering to identify the changed behavior of the left and right lanes [13]. Yi et al., used a Bayesian classifier and decision tree theory to predict vehicle lane-changing probability based on NGSIM data [14].

Deep learning technology has an excellent performance in many research applications [15,16]; some studies on lane-changing behavior also use deep learning technology for analysis and research [17]. Based on the real vehicle data (NGSIM) from the next-generation simulation data, Dou Yangliu et al., proposed a vehicle lane-changing prediction model by using the characteristics of speed difference, vehicle spacing, and position, combined with SVM and the artificial neural network (ANN) [18]. Peng Jinshuan et al., developed a back propagation neural network (BPNN) model for predicting vehicle lane-changing intention [19]. The model accurately predicted vehicle lane-changing behavior at least 1.5 s in advance by analyzing the drivers' visual data, operation data, vehicle motion data, and traffic environment. Xu Ting et al., established a two-layer convolutional neural network model (CNN) to identify the drivers' lane-changing behavior [20]. Mammer et al., only used a small number of manually labeled samples to train the convolutional neural network (CNN) for identifying lane-changing intention [21].

Although various machine learning models are widely used in lane-changing recognition, there are still some shortcomings. For example, due to the correlation between input variables, the Bayesian network model has a great impact on lane-changing recognition results [22]; the artificial neural network has many parameters and a slow operation speed; the decision tree is prone to overfitting; and the support vector machine algorithm is better than the above machine learning algorithm, but it is not easy to determine the optimal penalty factor and kernel function parameters that affect the model [14]. Therefore, how to find the optimal support vector machine parameters to improve the accuracy of model recognition is a research hotspot. The traditional support vector machine uses the grid search algorithm to determine the appropriate support vector machine parameters, but the algorithm has the problems of complex calculation and being time consuming [23]. Salvucci et al., proposed a driver model for detecting vehicle lane-changing behavior by comparing simulated driving behavior with actual driving behavior [24]. Although many studies have classified specific lane-changing behaviors and extracted the characterization parameters, these works are mainly focused on the research of traffic flow and are mostly used for microscopic traffic simulation rather than vehicle lane-changing recognition [25].

1.2. Problem Description and Formulation

During the actual lane-changing, the driver may wait to change lanes when approaching the lane line, or give up the lane-changing operation according to the surrounding traffic conditions. Driving style has a direct impact on the intention of changing lanes; drivers with different styles often have different driving habits. Most of the previously mentioned studies did not comprehensively consider the relationship between driver style and lane-changing intention recognition but instead studied the impact of one of them on traffic safety alone. In light of this, it is interesting and necessary to use the data-driven method and introduce the consideration of the driver's style in lane-changing recognition research.

At present, the rise of automatic driving technology has driven the research of vehicle lane-changing behavior. The recognition of autonomous vehicle lane-changing behavior is mainly realized by machine vision technology and radar. Road image algorithms can help autonomous vehicles to recognize and judge. The most reliable way to train an automatic driving algorithm is to carry out a road test, but the road environment is single, so it is difficult to train the CNN network of the system comprehensively [26,27].

However, the image techniques make it available for data-driven methods. In this paper, the UAV is used to collect the vehicle trajectory data at the exit of the urban expressway and vehicle lane-changing segments are extracted by the lane-changing rules. On this basis, the driving styles are clustered, and a model is established to predict and recognize driver styles. In the end, a machine learning model is established to recognize the intention of vehicle lane-changing. This can be applied to automatic driving vehicles to identify dangerous lane-changing and promote the development and maturity of the advanced driving assistance system (ADAS). In this way, a lane-changing collision risk can be identified by vehicles and the driving track of autonomous vehicles can be optimized to improve driving safety.

Driving style directly affects the driver's lane-changing behavior, but traditional lane-changing recognition research does not consider this factor. This paper makes some assumptions as driving style is introduced in the research, which can effectively improve the accuracy and integrity of lane-changing recognition strategies, but we also make several new contributions. The recognition of lane-changing behavior at the exit of the urban expressway perfects the research in this field and improves driving safety. Three machine learning models are established to recognize the intention of vehicle lane-changing in time by mining the effective information of vehicle trajectory data and comparing the results of a single time point and different time windows, which improves the accuracy and feasibility of the research.

The remainder of the paper is organized as follows: the details of the data and methods are presented in Section 2, including the data collection and the trajectory data extraction. Driving style clustering and its classification are explained in Section 3, including the clustering characteristics, method, and classification model features. The lane-changing recognition model application and a discussion of the modeling results are presented in Section 4. The last is the conclusion section which summarizes the study and recommends future research directions.

2. Materials and Methods

This study mainly focuses on the exit diversion area of the urban expressway. To collect more natural driving track data within the effective time, the exit of the urban expressway at the South Third Ring Road of Xi'an was selected in this study. It is a two-way six-lane expressway; the lane width is 3.75 m and the design speed is 100 km/h.

2.1. Data Collection

For the exit, the traffic conditions during the morning peak hours (8:30–9:30 am) and afternoon peak hours (5:30–6:30 pm) on a weekday were collected by the UAV and were stabilized at about 200 m to ensure an adequate viewing height to cover the whole functional area (as shown in Figure 1). To meet the requirements of needed for the UAV, the weather is sunny with a high visibility and no wind, which ensures that the UAV can capture the video in a stable state.

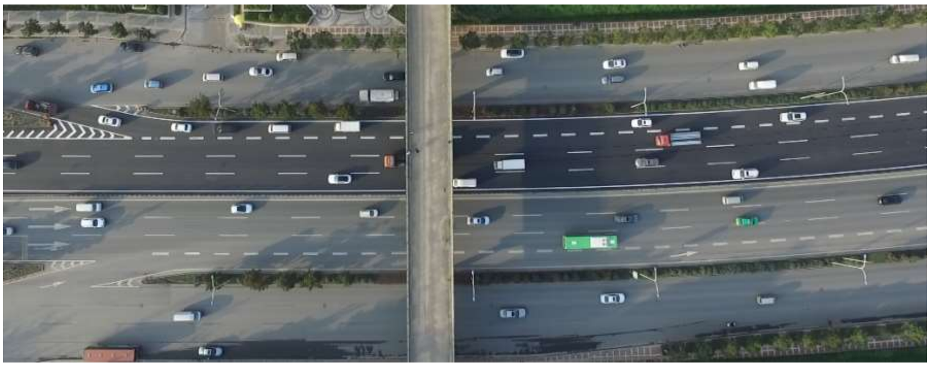
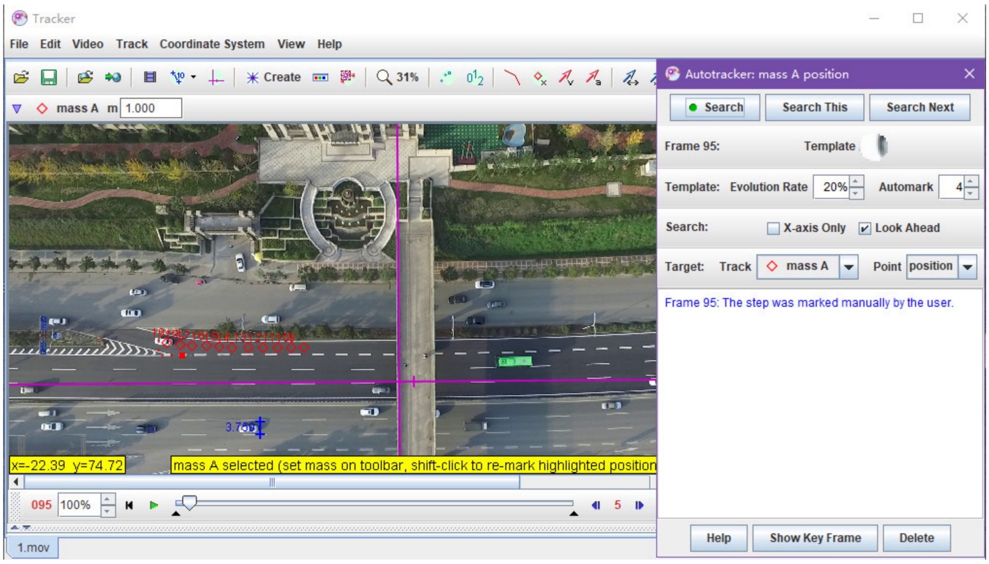


Figure 1. Range of aerial photography of UAV.

2.2. Trajectory Data Extraction

As a video analysis software, Tracker was first widely used in the physical field. In recent years, with its advantages of its simple operation, high accuracy, and powerful functions, Tracker has been widely known by the public and has also been widely used in the transportation field. The tracker software is used to extract the vehicle trajectory data. By marking the coordinate system and tracking the target automatically or manually, the coordinates, speed, and other information of each vehicle at any time can be obtained. The steps to extract the natural driving trajectory data based on the tracker are as follows.

First, the coordinate system and reference distance calibration were determined. In this study, the location of the overpass in the video is used as the coordinate origin to establish a two-dimensional coordinate system, in which the X-axis coincides with the lane line in the diversion area, and the Y-axis is perpendicular to the lane line direction. In addition, to obtain accurate data, it is necessary to take the fixed length and static target in the video as the reference. The lane width of the survey site was 3.75 m. Therefore, the distance value in the coordinate system is calibrated with the lane width as the reference. The calibration diagram is shown in Figure 2. For the convenience of the following description, the lanes are numbered uniformly. From the deceleration lane inward, the lanes are numbered as 1, 2, 3, and 4.


Tracker
File Edit Video Track Coordinate System View Help
Create
31%
Autotracker: mass A position
Search
Search This
Search Next
Help
Show Key Frame
Delete
1.mov

mass A m 1.000

Frame 95: Template

Template: Evolution Rate 20% Automark 4

Search: □ X-axis Only ☑ Look Ahead

Target: Track mass A Point position

Frame 95: The step was marked manually by the user.

x=-22.39 y=74.72

mass A selected (set mass on toolbar, shift-click to re-mark highlighted position

095 100% 5

Figure 2. Coordinate calibration.

Second, target tracking and trajectory extraction were carried out. Each vehicle was numbered, a tracking point at the front of the vehicle was created, and the vehicle was tracked in real-time through the automatic tracking function of Tracker software, as shown in the red circle in Figure 2. The red diamond point in the figure indicates the position of the vehicle at each moment. Because of the large error caused by inaccurate identification in automatic tracking, the speed and acceleration change curve trend of each vehicle will be checked after the automatic tracking is completed, and the vehicle with a large error will be manually marked for tracking to reduce the error and finally obtain complete natural driving track data.

Finally, 1482 pieces of complete natural driving track data were obtained, which contained 322 tracks of heavy vehicles and 1160 tracks of cars. At the same time, the position and other information of each vehicle at any time can be obtained. The relationship between each index value is shown in Figure 3. On this basis, the vehicle speed, acceleration, and other related indicators can be calculated, mainly including the frame number (time frame), vehicle ID, vehicle type class, vehicle lane number (lane id), vehicle X-axis speed v_x, vehicle Y-axis speed v_y, vehicle X-axis acceleration a_x, vehicle Y-axis acceleration a_y, vehicle X-axis coordinate value X, and the Y-axis coordinate value Y. Table 1 is the extracted track data of a certain part of the vehicle.

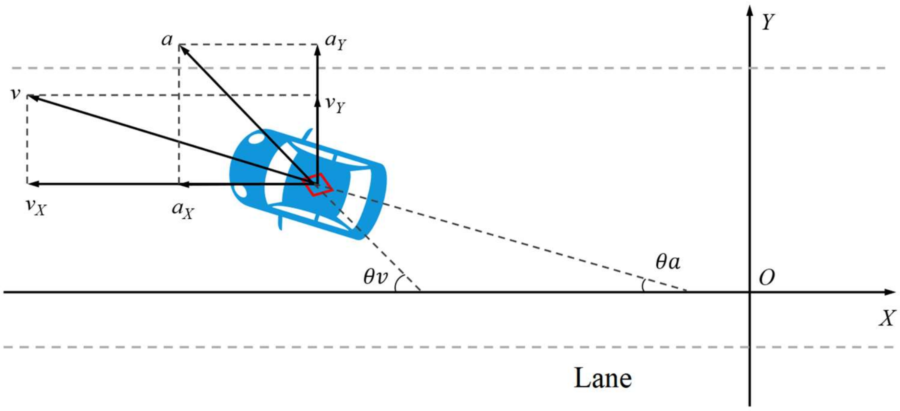


Figure 3. Vehicle track parameters.

Table 1. Vehicle running track data (part).

Frame ID	ID	X (m)	Y (m)	v_x (m/s)	v_y (m/s)	a_x (m/s^2)	a_y (m/s^2)	Lane Id	Class
101	32	−54.12	15.13	−30.71	1.04	0.27	0.18	3	Truck
102	32	−52.87	15.08	−30.69	1.05	0.27	0.16	3	Truck
103	32	−51.62	15.04	−30.68	1.06	0.26	0.15	3	Truck
104	32	−50.39	14.99	−30.67	1.06	0.26	0.13	3	Truck
105	32	−49.18	14.95	−30.66	1.07	0.26	0.12	3	Truck
106	32	−47.95	14.91	−30.65	1.07	0.26	0.1	3	Truck
107	32	−46.72	14.86	−30.64	1.07	0.26	0.09	4	Truck
108	32	−45.52	14.82	−30.63	1.07	0.25	0.08	4	Truck
…	…	…	…	…	…	…	…	…	…

2.2.1. Lane-Changing Segment Extraction Criteria

Lane-changing refers to the behavior of the driver driving to the adjacent lane to meet the needs during the driving process. The complete lane-changing process is shown in Figure 4, which can be divided into discretionary lane-changing and compulsive lane-changing.

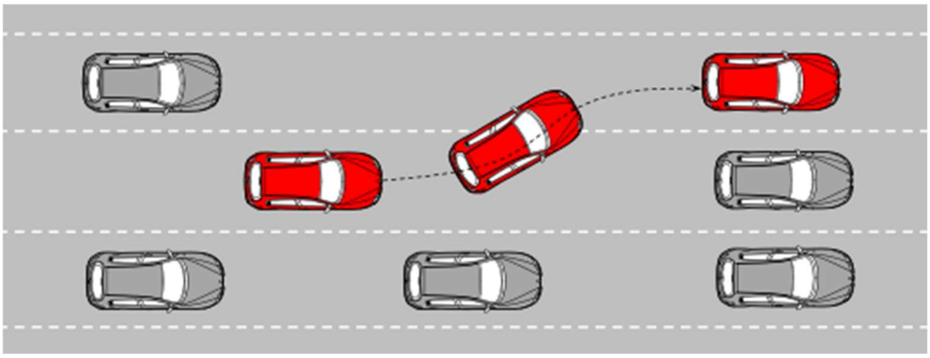
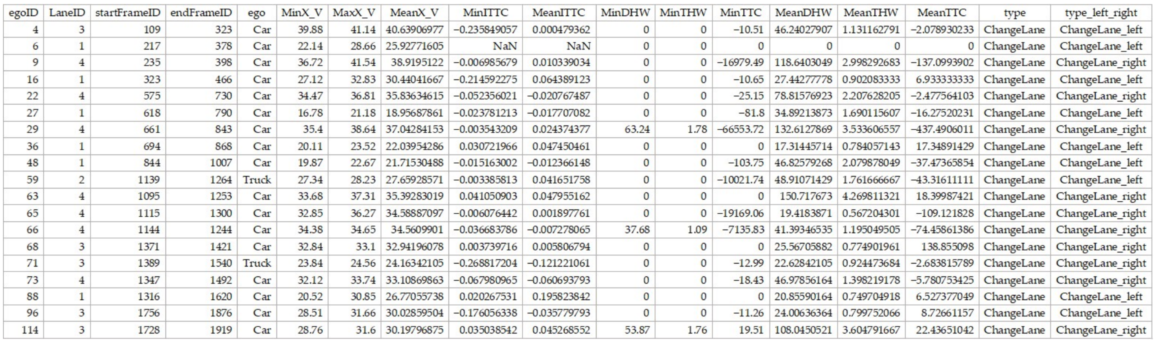


Figure 4. Schematic diagram of vehicle lane-changing.

Compulsive lane-changing mostly occurs in the ramp junction area, construction area, and when an accident occurs in front of the current lane. This paper does not distinguish between discretionary lane-changing and compulsive lane-changing, but mainly studies the driver's discretionary lane-changing behavior.

Before extracting the lane-changing segment, the time when the vehicle crosses the lane marking is taken as the key frame, and the starting and ending points of lane-changing are found forward and backward, respectively. In this paper, the longitudinal moving distance of the vehicle within 10 frames is less than 0.1 m, which is the basis for the start and the end of lane-changing, and lane-changing segments are extracted based on this logic. Some lane-changing results and the extracted partial lane-changing data are shown in Figure 5 below.



egoID	LaneID	startFrameID	endFrameID	ego	MinX_V	MaxX_V	MeanX_V	MinITTC	MeanITTC	MinDHW	MinTHW	MinTTC	MeanDHW	MeanTHW	MeanTTC	type	type_left_right
4	3	109	323	Car	39.88	41.14	40.63906977	−0.235849057	0.000479362	0	0	−10.51	46.24027907	1.131162791	−2.078930233	ChangeLane	ChangeLane_left
6	1	217	378	Car	22.14	28.66	25.92771605	NaN	NaN	0	0	0	0	0	0	ChangeLane	ChangeLane_left
9	4	235	398	Car	36.72	41.54	38.9195122	−0.006985679	0.010339034	0	0	−16979.49	118.6403049	2.998292683	−137.0993902	ChangeLane	ChangeLane_right
16	1	323	466	Car	27.12	32.83	30.44041667	−0.214592275	0.064389123	0	0	−10.65	27.44277778	0.902083333	6.933333333	ChangeLane	ChangeLane_left
22	4	575	730	Car	34.47	36.81	35.83634615	−0.052356021	−0.020767487	0	0	−25.15	78.81576923	2.207628205	−2.477564103	ChangeLane	ChangeLane_right
27	1	618	790	Car	16.78	21.18	18.95687861	−0.023781213	−0.017707082	0	0	−81.8	34.89213873	1.690115607	−16.27520231	ChangeLane	ChangeLane_left
29	4	661	843	Car	35.4	38.64	37.04284153	−0.003543209	0.024374377	63.24	1.78	−66553.72	132.6127869	3.533606557	−437.4906011	ChangeLane	ChangeLane_right
36	1	694	868	Car	20.11	23.52	22.03954286	0.030721966	0.047450461	0	0	0	17.31445714	0.784057143	17.34891429	ChangeLane	ChangeLane_left
48	1	844	1007	Car	19.87	22.67	21.71530488	−0.015163002	−0.012366148	0	0	−103.75	46.82579268	2.079878049	−37.47365854	ChangeLane	ChangeLane_left
59	2	1139	1264	Truck	27.34	28.23	27.65928571	−0.003385813	0.041651758	0	0	−10021.74	48.91071429	1.761666667	−43.31611111	ChangeLane	ChangeLane_left
63	4	1095	1253	Car	33.68	37.31	35.39283019	0.041050903	0.047955162	0	0	0	150.717673	4.269811321	18.39987421	ChangeLane	ChangeLane_right
65	4	1115	1300	Car	32.85	36.27	34.58887097	−0.006076442	0.001897761	0	0	−19169.06	19.4183871	0.567204301	−109.121828	ChangeLane	ChangeLane_right
66	4	1144	1244	Car	34.38	34.65	34.5609901	−0.036683786	−0.007278065	37.68	1.09	−7135.83	41.39346535	1.195049505	−74.45861386	ChangeLane	ChangeLane_right
68	3	1371	1421	Car	32.84	33.1	32.94196078	0.003739716	0.005806794	0	0	0	25.56705882	0.774901961	138.855098	ChangeLane	ChangeLane_right
71	3	1389	1540	Truck	23.84	24.56	24.16342105	−0.268817204	−0.121221061	0	0	−12.99	22.62842105	0.924473684	−2.683815789	ChangeLane	ChangeLane_right
73	4	1347	1492	Car	32.12	33.74	33.10869863	−0.067980965	−0.060693793	0	0	−18.43	46.97856164	1.398219178	−5.780753425	ChangeLane	ChangeLane_right
88	1	1316	1620	Car	20.52	30.85	26.77055738	0.020267531	0.195823842	0	0	0	20.85590164	0.749704918	6.527377049	ChangeLane	ChangeLane_left
96	3	1756	1876	Car	28.51	31.66	30.02859504	−0.176056338	−0.035779793	0	0	−11.26	24.00636364	0.799752066	8.72661157	ChangeLane	ChangeLane_left
114	3	1728	1919	Car	28.76	31.6	30.19796875	0.035038542	0.045268552	53.87	1.76	19.51	108.0450521	3.604791667	22.43651042	ChangeLane	ChangeLane_right

Figure 5. The extracted partial lane-changing data.

2.2.2. Fragment Extraction Results and Analysis

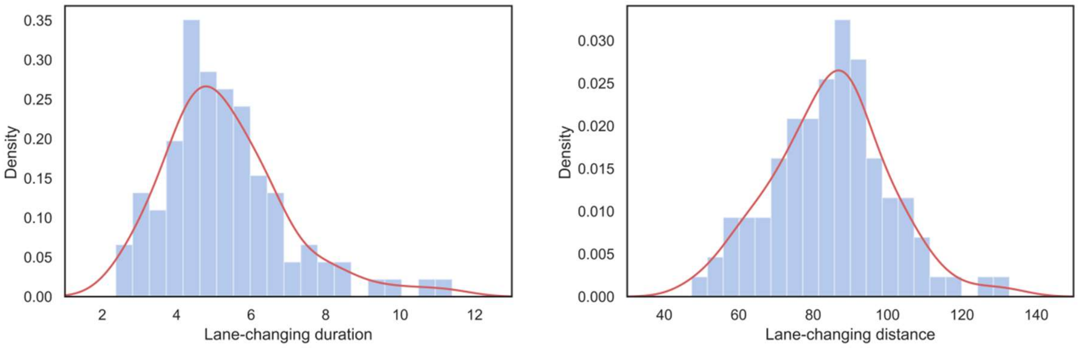
The lane-changing segment extraction was realized based on python. In this process, 374 lane-changing segments were obtained, and the duration of all lane-change segments is 1959 s. The statistical results of the lane-changing segments are shown in Table 2. It can be seen that Lane 2 has the most lane-changing segments, and the lane-changing segments of trucks are mainly concentrated in Lane 2 and 3.

Table 2. List of lane-changing segment extraction results.

Lane ID	Car			Truck		
	Number	Duration (s)	Mean Velocity (m/s)	Number	Duration (s)	Mean Velocity (m/s)
1	25	121	28.01	4	34	24.33
2	157	905	28.59	25	134	24.99
3	81	357	29.62	19	108	25.32
4	65	341	33.50	8	59	26.30
Total	328	1624		46	335	

2.2.3. Characteristic Analysis of Lane-Changing Segment

According to the analysis of the lane-changing characteristics, the density of lane-changing duration and distance are shown in Figure 6. It can be seen that the duration of the lane-changing segments reaches a peak near 4.5 s, and the duration of most lane-changing segments is between 4 s and 6 s; the lane-changing distance is mainly distributed between 70 m and 100 m, which reaches a peak near 90 m.



(**a**) Lane-changing Duration density (**b**) Lane-changing Distance density

Figure 6. Duration and distance density map of lane-changing segments.

2.3. Driving Style Clustering

To cluster the driving style, it is first necessary to obtain the features that can represent the driving style. In this paper, the main extracted features include the distance headway (DHW), time headway (THW), time to collision (TTC), and the inverse of TTC (ITTC). The DHW represents the distance between the front and rear vehicles. The THW represents the time difference between the front and rear vehicles passing through the same place; it can be calculated by dividing the DHW by the following vehicle speed. The TTC indicates the time required for the collision if two vehicles continue to collide at the current speed and on the same path; it can be calculated by dividing the DHW by the speed difference between two vehicles. The minimum, maximum, average, and standard deviation of these indicators can be calculated as features and 12 features were obtained, as shown in Table 3.

Through analysis of the above indicators, it was found that some indicators are highly correlated, as shown in Figure 7, which will increase the training time and affect the performance of the model to a certain extent. Therefore, it is necessary to reduce the dimension of the primal data.

Table 3. List of driving style clustering characteristics.

Symbols	Unit	Meaning
TTC_{min}	s	The minimum value of TTC
TTC_{mean}	s	The average value of TTC
TTC_{std}	s	Variance of TTC
$ITTC_{min}$		The minimum value of ITTC
$ITTC_{mean}$		The average value of ITTC
$ITTC_{std}$		Variance of ITTC
THW_{min}	s	The minimum value of THW
THW_{mean}	s	The average value of THW
THW_{std}	s	Variance of THW
DHW_{min}	m	The minimum value of DHW
DHW_{mean}	m	The average value of DHW
DHW_{std}	m	Variance of DHW

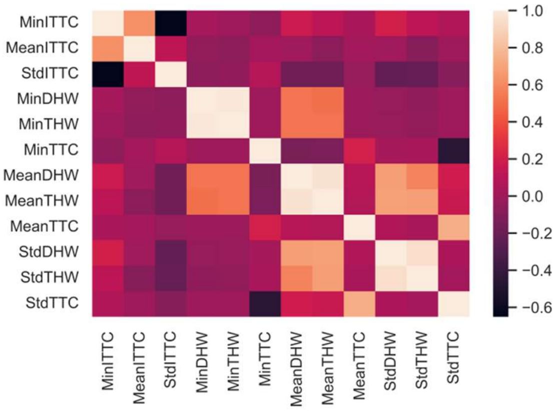


Figure 7. Correlation of driving style characteristics.

This paper normalizes the features and reduces the dimension by PCA. Feature normalization and PCA dimensionality reduction based on Python can obtain the cumulative contribution rate of the principal components, as shown in Figure 8. It can be seen that the first four principal components have been able to represent more than 85% of the information of the original features. Therefore, this paper selects the first four principal components for subsequent analysis.

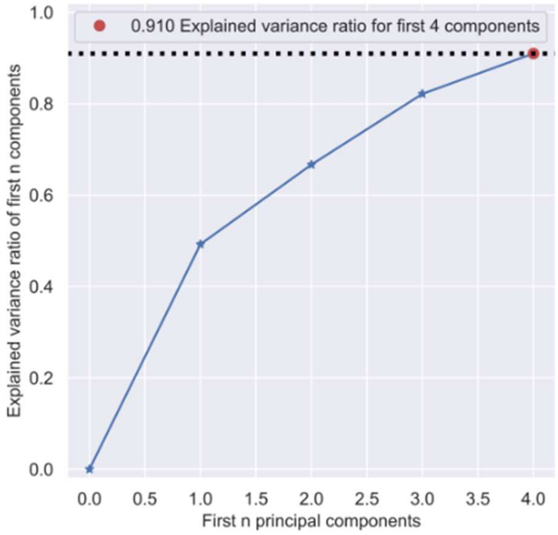
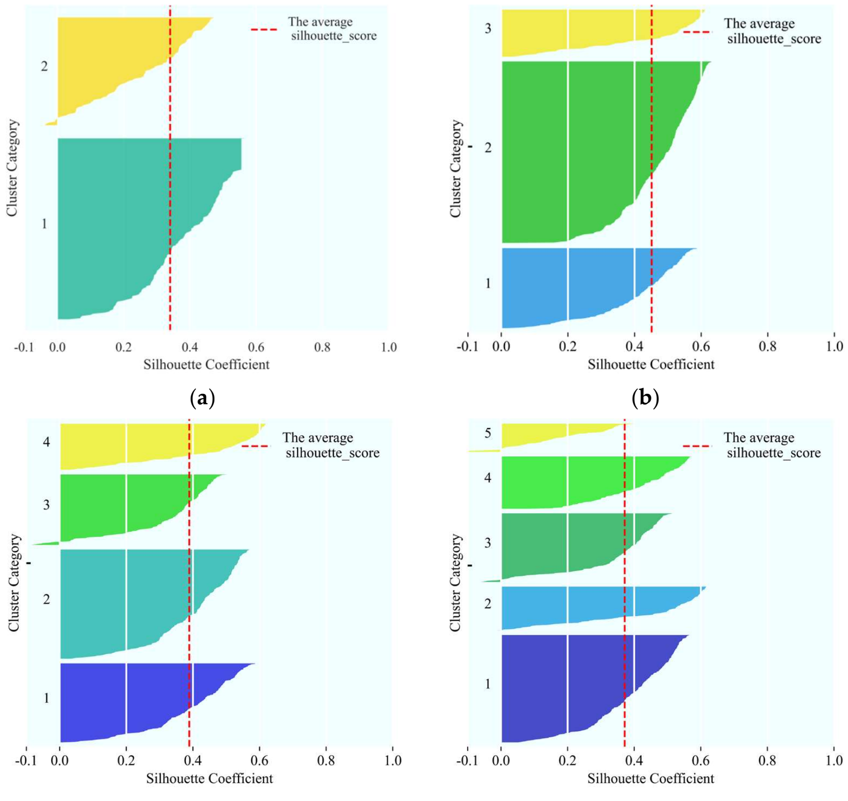


Figure 8. The cumulative contribution rate of PCA.

The K-means algorithm is the most commonly used unsupervised clustering algorithm, which divides a group of samples into several clusters without intersections [28]. An important step in K-means is to place the initial centroid. Theoretically, K-means must converge after a long time of operation, but the sum of squares in the cluster may converge to the local minimum. Whether it can converge to the real minimum depends largely on the initialization of the centroid. If the initial centroid is placed at different positions, the clustering results are likely to be inconsistent. A good centroid selection can make K-means avoid more calculations and make the algorithm converge stably and faster. The traditional K-means algorithm uses a random method to extract the samples from the sample points as the initial centroid, which has certain limitations. The improved k-means++ algorithm based on the traditional K-means algorithm can make the initial centroids far away from each other, to guide more reliable results than random initialization. When the k-means++ algorithm selects K cluster centers, it selects the first cluster center by a random method [29]. When selecting the next cluster center, it will give priority to the samples farther away from the selected cluster center. This cycle will continue until all cluster centers are selected. Although the improved principle of the k-means++ algorithm is simple, it is more effective than the traditional K-means algorithm.

In this paper, the k-means++ algorithm is used for driving styles cluster analysis. The data involved in clustering is the vehicle driving data after the dimension reduction (DHW, THW, TTC, ITTC) when the number of clusters is two, three, four, and five, respectively, which can obtain the silhouette coefficient results, as shown in Figure 9. The red line in each figure represents the average silhouette coefficient of the current cluster. In the clustering results, the silhouette coefficients of the samples on the right side of the red line are higher than the average silhouette coefficient, which plays a positive role in the average silhouette coefficient of the model. It can be seen that negative values of the silhouette coefficient occur under different clustering results, but the proportion of negative values of the silhouette coefficient is smaller when the cluster number is three.



(**a**)

(**b**)

(**c**)

(**d**)

Figure 9. Comparison of silhouette coefficients under different cluster numbers. (**a**) The number of clusters is 2, (**b**) the number of clusters is 3, (**c**) the number of clusters is 4, (**d**) the number of clusters is 5.

The silhouette coefficients (SC) and Calinski–Harabasz (CH) scores of the different clustering results are shown in Figure 10, it can be seen that the average SC are 0.340, 0.451, 0.389, and 0.371 and the CH score is 143.5, 573.3, 594.3, and 537.7 when the cluster number is two, three, four, and five clusters, respectively. Although the CH score is higher when the cluster number is four, the SC score is relatively low at this time, while the SC and CH scores are high when the cluster number is three, so it is more reasonable to set the number of clusters to three.

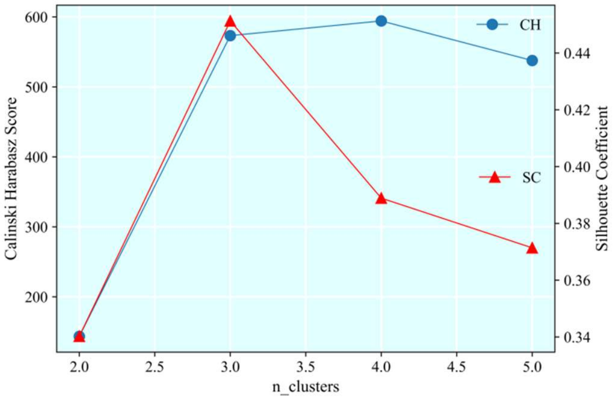


Figure 10. Silhouette coefficients and Calinski–Harabasz Score.

On this basis, the index values of each cluster of driving behavior are analyzed and the density map of lane-changing time is shown in Figure 11. It can be seen that the three types of driving behaviors obtained from the k-means++ clustering have significant differences, cluster 1, cluster 2, and cluster 3 can represent the "conservative", "ordinary", and "aggressive" driver styles. The average time required for lane-changing under the various driving styles is 4.88 s, 6.65 s, and 8.41 s, respectively. Conservative drivers are the most cautious. They will change lanes when the traffic environment risks are extremely low. The operation of lane-changing is relatively stable and smooth, so the lane-changing time is the longest; on the contrary, the operation of aggressive drivers is slightly aggressive, so the lane change time is the shortest. There are obvious differences in the safety of various driving behaviors. The three types of driving styles will be used as features for the following research on vehicle lane change intention recognition.

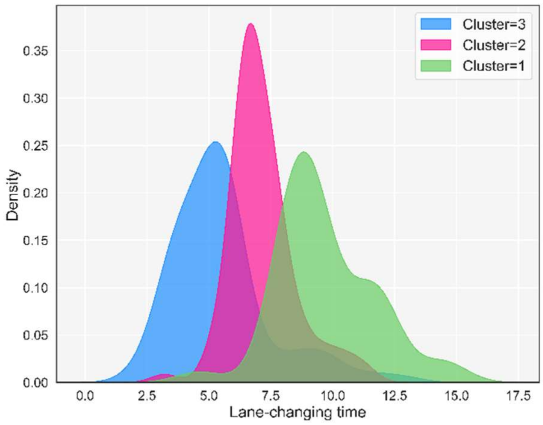


Figure 11. The density map of lane-changing time.

2.4. Driving Style Classification

The premise of establishing the model to predict the driver style is to extract the driving characteristics of the target vehicle and the surrounding vehicles, including the speed, acceleration, distance, and other indicators. The specific indicators are shown in Table 4. The surrounding vehicles mainly refer to the front vehicle (FV) and the rear vehicle (RV).

Table 4. Driving style classification model features.

Symbols	Unit	Characteristic Meaning
$xVel^t_{ego}$	m/s	Target vehicle speed in the x-direction at time t
$yVel^t_{ego}$	m/s	Target vehicle speed in the x-direction at time t
$xAcc^t_{ego}$	m/s^2	Acceleration of target vehicle in the x-direction at time t
$yAcc^t_{ego}$	m/s^2	Acceleration of target vehicle in the x-direction at time t
$laneID^t_{ego}$		Lane of target vehicle at time t
$Class_{ego}$		Type of target vehicle, car or truck
$leftLaneDis^t_{ego}$	m	Distance from target vehicle to left lane line at time t
$rightLaneDis^t_{ego}$	m	Distance from target vehicle to right lane line at time t
$\Delta x^t_{vehicle}$	m	Distance difference between surrounding vehicle and target vehicle in the x-direction at time t
$\Delta y^t_{vehicle}$	m	Distance difference between surrounding vehicle and target vehicle in the y-direction at time t
$\Delta xVel^t_{vehicle}$	m/s	The speed difference between the surrounding vehicle and the target vehicle in the x-direction at time t
$\Delta yVel^t_{vehicle}$	m/s	The speed difference between the surrounding vehicle and the target vehicle in the y-direction at time t
$\Delta xAcc^t_{vehicle}$	m/s^2	Acceleration difference between surrounding vehicle and target vehicle in the x-direction at time t
$\Delta yAcc^t_{vehicle}$	m/s^2	Acceleration difference between surrounding vehicle and target vehicle in the y-direction at time t
$Class^t_{vehicle}$		The type of vehicles around at time t, car or truck

It will cause “overfitting” or “underfitting” when the model has too many or too few features. Appropriate features can reduce the complexity of the model and improve training accuracy. The optimization method (RFECV) based on recursive feature elimination (REF) is adopted to screen important features in this paper. First, all features are modeled and sorted according to their importance. After deleting one or several features with the lowest importance, the modeling and analysis are conducted again. The cycle is repeated until the importance of all the features is sorted. The final feature importance ranking is shown in Table 5.

Table 5. RFECV Feature filtering.

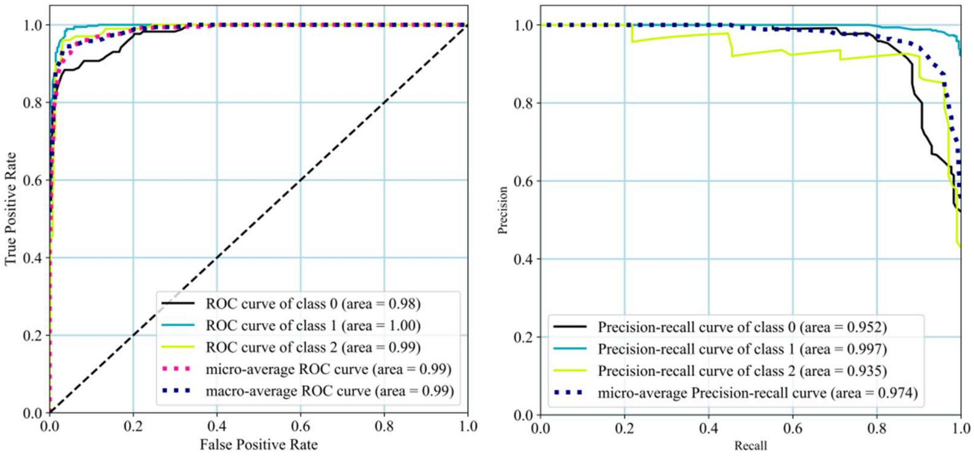
Symbols	Characteristic Meaning	Importance Value	Importance Ranking
Δx^t_{FV}*-median*	The median of the distance difference between the target vehicle and the front vehicle in the x-direction	0.2875	1
$\Delta xAcc^t_{FV}$*-mean*	Mean value of acceleration difference in the x-direction between target vehicle and front vehicle	0.2546	2
$xVel^t_{ego}$*-absolute_sum_of_changes*	Sum of absolute values of continuous changes in the speed sequence of the target vehicle in the x-direction	0.1978	3
$xVel^t_{ego}$*-median*	Speed median of the target vehicle in the x-direction	0.1385	4
Δy^t_{FV}*-mean*	Mean value of speed difference in the y-direction between target vehicle and front vehicle	0.1187	5
$yAcc^t_{ego}$*-mean_change*	Absolute value’s mean value of continuous change value of speed in the y-direction of the target vehicle	0.0946	6
$Class^t_{FV}$	Front vehicle type	0.0694	7

The random forest algorithm can complete the task of classification and regression, and the default parameters have a well performance on most data sets. The random forest algorithm is adopted to classify and predict the results after the k-means++ clustering in this paper; 70% of the data is divided into the training set and 30% into the test set. The final prediction results of the trained model are shown in Table 6. Cluster 1, cluster 2, and cluster 3 can represent the “conservative”, “ordinary”, and “aggressive” driver styles, of which the corresponding labels in the classification model are class 0, class 1, and class 2. The model has the highest prediction accuracy and recall for the aggressive driving style. That is, the model can identify the aggressive driving style well. The main reason is that the aggressive sample data is more recognizable than the other two types of feature distribution. It is also more practical to accurately identify the aggressive style in practical applications.

Table 6. Driving style model prediction results.

RF	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Macro Avg**	**Weighted Avg**	**Accuracy**
	Conservative	**Ordinary**	**Aggressive**			
	Class 0	**Class 1**	**Class 2**			
Precision	0.8378	0.9761	0.8803	0.8983	0.9327	0.9306
Recall	0.9394	0.9874	0.9435	0.9268	0.9306	
F1	0.8857	0.9617	0.9570	0.9115	0.9312	

The ROC and PR curves of the random forest model are shown in Figure 12. It can be seen that the area under the ROC and PR curves is greater than 0.9, almost approaching 1, which indicates that the model can distinguish three types of samples well.



(**a**) (**b**)

Figure 12. ROC and PR curves. (**a**) ROC curve, (**b**) PR curve.

3. Lane-Changing Recognition Model

Machine-learning models including XGBoost (extreme gradient boosting), LightGBM (light gradient boosting machine), and the Stacking fusion model are often used by scholars for recognition and prediction research. Compared with the traditional gradient lifting algorithm (GBDT), XGBoost is faster. At present, the algorithm has an excellent performance in classification and regression analysis. LightGBM avoids some shortcomings of XGBoost, such as the high requirements for computing devices and the large consumption of hardware facilities, and the training speed of the LightGBM model is faster, consumes less memory, and has a higher accuracy. As the best method to improve the effectiveness of the machine learning model, the Stacking fusion model can further improve the prediction accuracy based on the traditional integrated model which uses multi-fold cross-validation, so the results of the model are more stable, but the process of stacking generally takes a long time. This paper uses three different methods to identify lane-changing behavior and selects a suitable model for the exit of the urban expressway.

To improve the accuracy of the lane-changing intention recognition model, lane-changing recognition is analyzed from two aspects: a single time point in the lane-changing window and the whole historical segment of the lane-changing window.

The former mainly considers the characteristics of a time in the lane-changing window, and the latter considers the characteristics of each period in the whole lane-changing window. For each moment, the extracted features mainly include the horizontal and vertical speed and acceleration of the target vehicle, the distance from the self-vehicle to the left and right lane lines, the relative speed, relative acceleration, and relative distance of the target vehicle and the surrounding vehicles, as well as the type of self-vehicle and the surrounding vehicles and the driver's driving style. For any vehicle at each time, it must belong to the set {TV, FV, RV, LFV, LPV, LRV, RFV, RPV, RRV}. TV represents the target vehicle, FV represents the vehicle in front of the target vehicle, RV represents the vehicle behind the target vehicle, LFV represents the vehicle in front of the target vehicle on the left, LPV represents the vehicle adjacent to the target vehicle on the left, LRV represents the vehicle behind the target vehicle on the left, RFV represents the vehicle in front of the target vehicle on the right, RPV represents the vehicle adjacent to the target vehicle on the right, and RRV represents the vehicle behind the target vehicle on the right.

The relative positions of the target vehicle and the surrounding vehicles are shown in Figure 13. If the vehicle at a corresponding position does not exist, its corresponding characteristics are set to null. Finally, the characteristic meanings and symbols of the features extracted at each time point are the same as those in Table 4. In addition, the driver styles of the target vehicle and surrounding vehicles are added. This value is the label of the driving style cluster. The new features are shown in Table 7. Zero, one, and two represent "conservative", "ordinary", and "aggressive" driving styles, respectively.

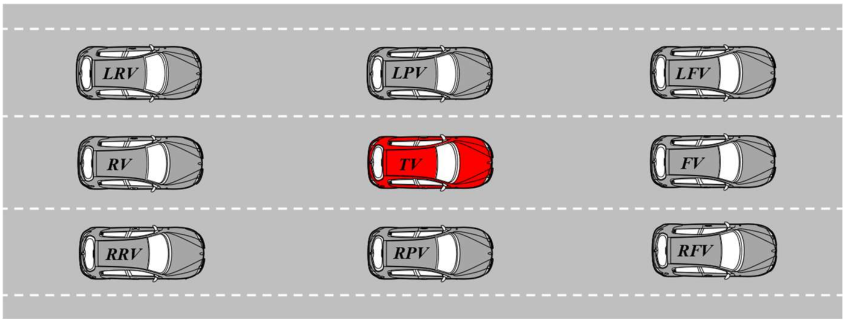


Figure 13. The relative position of vehicles.

Table 7. Lane-changing intention recognition feature.

Symbol	Characteristic Meaning
Styleego	Target vehicle driver style, 0, 1, or 2
$Style^t_{vehicle}$	The driver's style of surrounding vehicles at time t, 0, 1, or 2

3.1. Lane-Changing Intention Recognition Model Considering Single-Time Point Characteristics

To analyze the effect of the three models on lane-changing recognition, the starting point time of lane-changing is taken as the modeling time, and the instantaneous characteristic values of left lane-changing and right lane-changing track segments are extracted at this time. All sample data were divided into the training set and test set, in which the training set accounts for 70% and the test set accounts for 30%. On this basis, XGBoost, LightGBM, and the Stacking fusion model were trained and tested respectively.

The final prediction results of the three models are shown in Table 8, respectively. It can be seen that XGBoost and LightGBM have the same accuracy in lane-changing intention recognition. However, the XGBoost algorithm has a longer training time and better fitting effect on various data. The Stacking fusion model also has a good performance in lane-changing intention recognition, the recognition accuracy of lane-changing to the right is even close to 100%, but the comprehensive realization of the fusion model is less improved than that of a single model. On the one hand, it may be that the two integrated algorithm models of XGBoost and LightGBM have a strong enough prediction ability, so it is difficult to improve the accuracy to a new level by modeling skills alone. At the same time, the amount of sample data is too small, so the model cannot learn a lot of effective information, which limits the upper limit of the model to a certain extent.

It can be seen that the Stacking fusion model can well recognize the driving intention. The model has a high recognition accuracy for the driving behavior of left lane-changing and right lane-changing; the classifications of the latter are correct basically, but the recognition accuracy of left lane-changing is lower than the right. On the one hand, the difference between the two kinds of characteristic information is not very obvious. On the other hand, this kind of sample accounts for the least proportion in the original sample, which results in insufficient learning of this behavior.

To study the accuracy of the model for lane-changing intention recognition at different single time points, the lane-changing intention time window is set to a fixed length, which is set to 3 s in this paper. The prediction accuracy, model training duration, and other information of each time point model are recorded respectively. The final statistics of some modeling results are shown in Table 9.

Table 8. Prediction results of lane-changing intention recognition with single-time feature.

Evaluating Indicator		XGBoost	LightGBM	Stacking
Turn right	Precision	0.9583	0.9592	0.9669
	Recall	1.0000	0.9692	0.9851
	F1	0.9926	0.9692	0.9925
Turn left	Precision	0.9175	0.9227	0.9377
	Recall	0.9091	0.8667	0.9091
	F1	0.9231	0.9123	0.9375
Macro avg	Precision	0.9685	0.9640	0.9780
	Recall	0.9640	0.9398	0.9618
	F1	0.9661	0.9510	0.9696
Weighted avg	Precision	0.9766	0.9630	0.9770
	Recall	0.9769	0.9630	0.9769
	F1	0.9767	0.9626	0.9767
Accuracy		0.9569	0.9530	0.9632

Table 9. Comparison of lane-changing intention recognition accuracy at different times.

Timestamp (s)	XGB_Score	LGB_Score	Stacking_Score	LGB_Time	XGB_Time
−2.96	0.7951	0.8076	0.7826	0.1640	0.6962
−2.72	0.7693	0.7932	0.8043	0.1320	0.7007
−2.48	0.8171	0.8183	0.7779	0.1460	0.7270
−2.24	0.7811	0.7811	0.7758	0.1625	0.7827
−2	0.8580	0.8598	0.8631	0.2981	1.4621
−1.76	0.8276	0.8331	0.8214	0.2976	1.4220
−1.52	0.8371	0.8508	0.8275	0.2959	1.3836
−1.28	0.8547	0.8550	0.8404	0.2866	1.4134
−1.04	0.8620	0.8768	0.8752	0.3161	1.3544
−0.8	0.8775	0.9112	0.9049	0.2851	1.4724
−0.56	0.9326	0.9404	0.9200	0.2641	1.4393
−0.32	0.9395	0.9332	0.9373	0.2641	1.6104
−0.04	0.9523	0.9582	0.9568	0.2431	1.4414
0.32	0.9539	0.9652	0.9609	0.2389	1.2929
0.56	0.9622	0.9708	0.9638	0.3358	1.4483
0.8	0.9718	0.9708	0.9680	0.2291	1.2190
1.04	0.9677	0.9736	0.9749	0.1910	1.2101
1.28	0.9730	0.9777	0.9791	0.1970	1.1799
1.52	0.9775	0.9791	0.9833	0.1840	1.1535
1.76	0.9747	0.9833	0.9861	0.2030	1.2070
2	0.9774	0.9818	0.9860	0.1995	1.0757
2.24	0.9775	0.9832	0.9887	0.2090	1.0410
2.48	0.9802	0.9818	0.9873	0.1900	1.0212
2.72	0.9789	0.9858	0.9830	0.1826	1.0933
2.96	0.9814	0.9843	0.9857	0.1851	1.0522

It can be seen that the prediction accuracy of the three models has been able to maintain an accuracy above 80% in the first two seconds of the lane-changing. However, 2 s before the start of the lane-changing, the prediction accuracy decreased significantly and the accuracy is basically below 80%. The prediction results of the three models are similar, but the training time of LightGBM is less than XGBoost, which is also an important reason why LightGBM has been widely used in the engineering field in recent years.

The curves of the lane-changing intention recognition accuracy of the three models at different times within 3 s before and after the start of the lane-changing are shown in Figure 14; the XGB_score, LGB_score, and Stacking_score, respectively, represent the accuracy of the three models. The negative timestamp indicates the time before lane-changing, which means how much time is left before the lane-changing occurs.

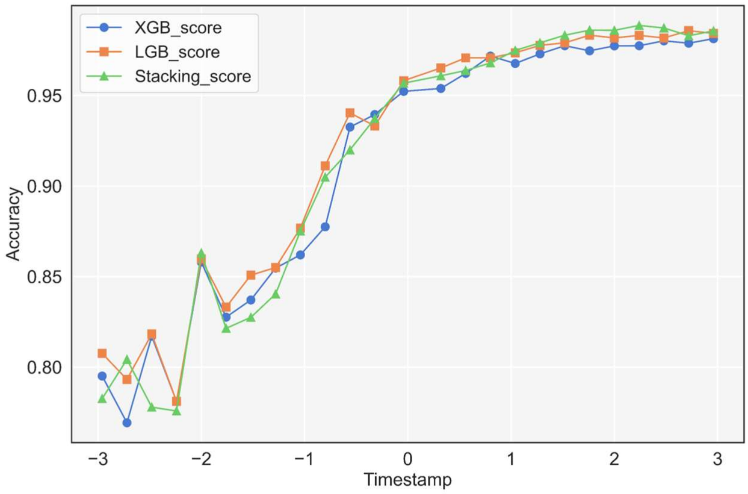


Figure 14. Comparison diagram of lane-changing intention recognition accuracy of each model at different times.

The result shows that the second before the starting time point of lane-changing is an important dividing point. The prediction accuracy of each model before this point drops to a low level, and they fluctuate greatly. This is because as the modeling time is too early, the vehicle may not show obvious behavior characteristics that can represent the intention of the lane-changing, so the recognition accuracy decreases. After the start of the lane-changing, the three model bases tend to be stable, and the prediction accuracy reaches the maximum value. In general, the difference between the recognition results is small, but the recognition accuracy of the Stacking fusion model is better than a single machine learning model, and the performance of the overall model is better considering the single-time point characteristics.

3.2. Lane-Changing Intention Recognition Model Considering Historical Time Window Characteristics

The above models only consider the characteristics of a certain time point in the lane-changing window. However, the process of the vehicle during the whole lane-changing stage, from the generation of the lane-changing intention to the end of the lane-changing, has a certain memory effect, and the vehicle state at the current moment will affect the state at the next moment. Thus, it is not comprehensive to consider the characteristics of a certain time point alone. Inspired by this, considering the characteristics of each time, the lane-changing window is set to different time thresholds. For each lane-changing window with a fixed length of time, all the characteristics in the whole lane-changing window are extracted.

The maximum time length of the lane-changing window was set as 3 s and decreased in steps of 0.12 s to get a total of 24 lane-changing window lengths. Three models were built separately and tested at different lane-changing window lengths. The curves of the accuracy of the three models are drawn in Figure 15 to intuitively show the impact of different lane-changing window lengths on the accuracy of lane-changing recognition models. It can be seen that there is little difference in the prediction accuracy of different lane-changing window lengths. However, the accuracy of the three models is at a high level when the lane-changing window length is about 2.2 s, and the LightGBM model and Stacking fusion model are higher than XGBoost. Whether the length of the lane-changing window is too long or too short, the prediction accuracy is reduced. Therefore, the length of the lane-changing window is determined as 2.2 s for subsequent research and analysis.

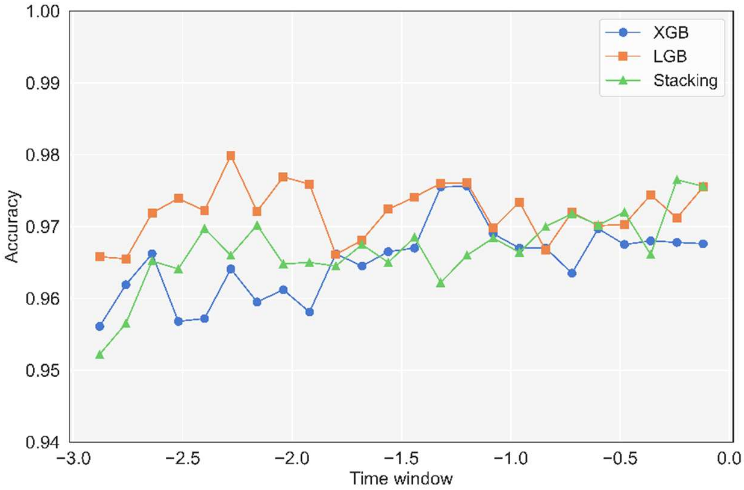


Figure 15. Accuracy comparison of different lane-changing time windows.

Models were established to recognize the lane-changing intention at each time point within 2.2 s before the starting time point of the lane-changing. The step size was taken as 0.12 s and the prediction results of 18 different time points were obtained. The specific results of each model are shown in Table 10. It can be seen that the accuracy of the three models has less difference, the Stacking fusion model has the best comprehensive effect. With the regression of the prediction time, the accuracy of the three models decreases to a certain extent.

Table 10. Comparison of lane-changing intention recognition accuracy at different times.

Timestamp (s)	XGB_Score	LGB_Score	Stacking_Score	LGB_Time	XGB_Time
−2.20	0.8015	0.7996	0.8151	0.1722	1.2171
−2.08	0.8256	0.8202	0.8384	0.1722	1.0249
−1.96	0.8417	0.8495	0.8512	0.2009	1.0903
−1.84	0.8652	0.8616	0.8756	0.1968	1.0947
−1.72	0.8711	0.8811	0.8882	0.1722	1.0527
−1.60	0.8984	0.8945	0.8976	0.1886	1.1133
−1.48	0.9101	0.9052	0.9207	0.2091	1.1519
−1.36	0.9155	0.9158	0.9255	0.1722	1.1194
−1.24	0.9161	0.9201	0.9275	0.1805	1.1395
−1.12	0.9215	0.9252	0.9326	0.2132	1.1007
−1.10	0.9315	0.9300	0.9355	0.1722	1.0914
−0.88	0.9412	0.9312	0.9425	0.1970	1.1558
−0.76	0.9296	0.9458	0.9414	0.2130	1.1148
−0.64	0.9408	0.9443	0.9505	0.2170	1.1344
−0.52	0.9325	0.9478	0.9472	0.2150	1.1094
−0.4	0.9483	0.9586	0.9579	0.2713	1.1771
−0.28	0.9536	0.9642	0.9723	0.1980	1.1288
−0.16	0.9586	0.9744	0.9791	0.1784	1.1243

The curve of the prediction accuracy of the three models is drawn as shown in Figure 16. The result shows that the accuracy decreases continuously with the regression of the prediction time, especially from 1.5 s before the start time point of lane-changing, where the decline speed increases. This is because as the prediction time moves back, the characteristics of the lane-changing behavior are not obvious, so the learning effect of the model on lane-changing is poor.

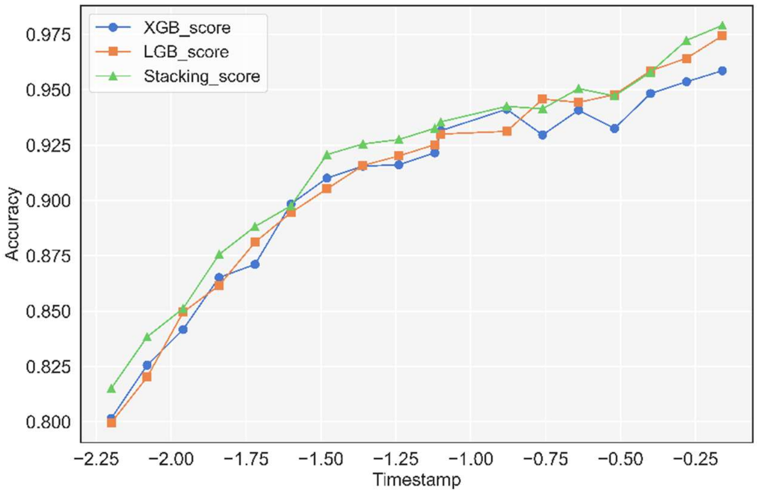


Figure 16. Comparison of accuracy of lane-changing intention recognition at different times.

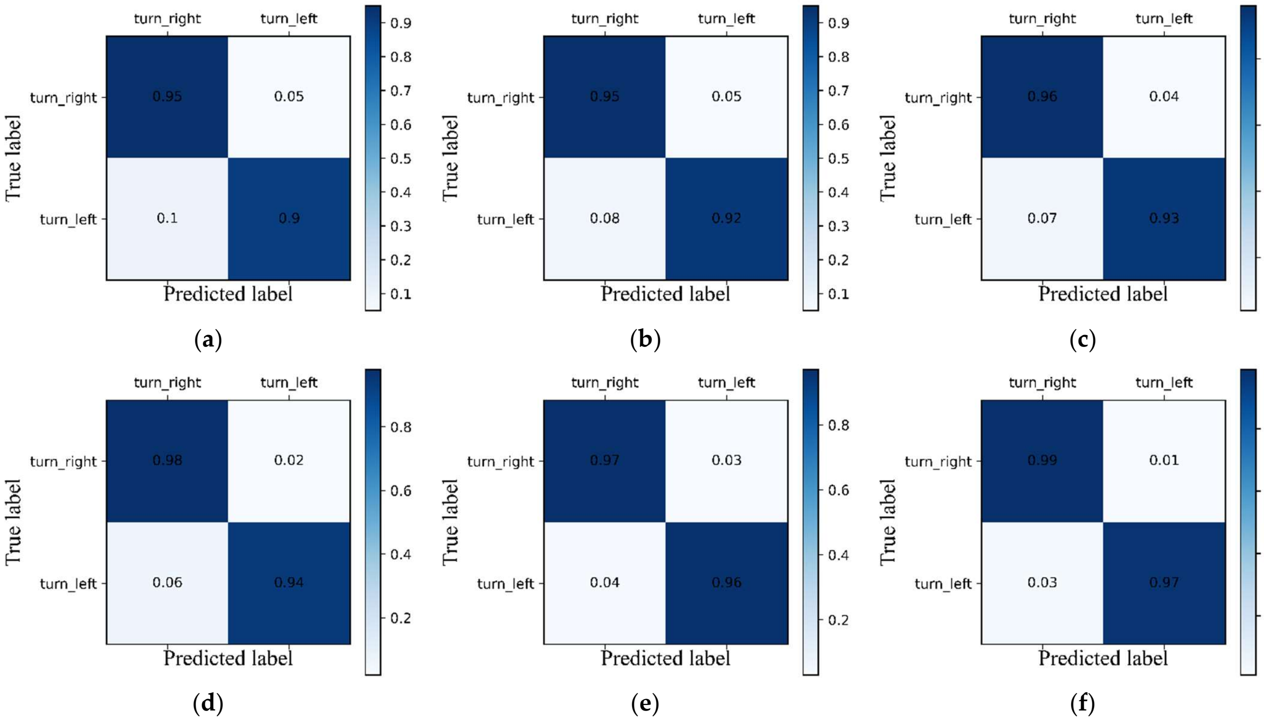
4. Evaluation and Discussion Section

The accuracy of the traditional lane-changing recognition methods (SVM, LSTM) is maintained at about 93% [7,8,11]. To compare with the model in this paper, the confusion matrix is drawn as shown in Figure 17. XGB_ history, LGB_ history, and Stacking_ history, respectively, represent the accuracy of the three models of considering the historical time window, which means for each lane-changing window with a fixed length of time, all the characteristics in the whole lane-changing window are extracted. XGB_ moment, LGB_ moment, and Stacking_ moment, respectively, represent the accuracy of the three models of considering single-time point characteristics, which means the model only considers the characteristics of a certain time point in the lane-changing window.

It can be seen that the Stacking fusion model can well identify the driving intention, whether considering a single-time point or the historical time window; the accuracy of the XGBoost and LightGBM models for the lane-changing intention recognition is similar. The recognition accuracy of the model for changing lanes to the right is at a high level, which can identify most samples correctly, but the recognition accuracy of the model for changing lanes to the left is lower. This may be because the left lane-changing sample accounts for the least proportion in the original sample, which results in the model's insufficient learning of this kind of driving behavior.

By comparing the lane-changing intention recognition models that only consider the single time point feature and consider the historical time window feature, it can be found that the three models have good effects. In general, the prediction accuracy of the Stacking fusion model is the highest and the training speed of the LightGBM model is the fastest. The accuracy of all models under the two types is drawn in Figure 18. The curve with the suffix "history" represents the model considering the characteristics of the historical time window, and the curve with the suffix "moment" is the model considering only the characteristics of a single time point. It can be found that the difference between the prediction results based on the two feature information is small when it is close to the start time of lane-changing, but the difference increases significantly with the regression of the prediction time.

In addition, it can be seen that the prediction accuracy of the lane-changing intention recognition model considering the characteristics of the historical time window is higher, which can guarantee an accuracy of more than 90% after 1.6 s before the start time point of lane-changing, while the accuracy of more than 90% can be guaranteed after 0.6 s before the start time point of lane-changing for the model considering only the characteristics of a single time point. At the same time, the model considering the historical time window feature can still ensure an accuracy of more than 82% at 2.2 s before the lane-changing time point, but the accuracy of the model considering only the feature of a single time point is about 78%, which indicates that the model considering the historical time window feature has a better comprehensive performance in lane-changing intention recognition. In addition, the prediction accuracy of the model that only considers the characteristics at a single time point has a sudden change at 2 s before the time point of lane-changing, which is likely caused by the instability of the characteristic information at this time.



(**a**) (**b**) (**c**)

(**d**) (**e**) (**f**)

Figure 17. Confusion matrix. (**a**) XGBoost_moment, (**b**) LightGBM_moment, (**c**) Stacking fusion model_moment, (**d**) XGBoost_history, (**e**) LightGBM_history, (**f**) Stacking fusion model_history.

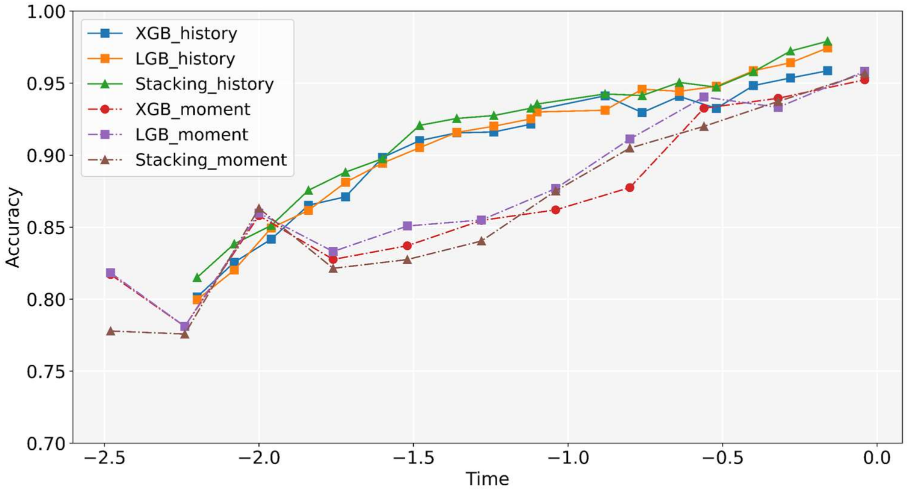


Figure 18. Comparison of lane-changing intention recognition accuracy based on a single time and historical time window.

5. Conclusions

The main conclusions of this study are as follows:

1. In this paper, the safety evaluation indexes including TTC, ITTC, THW, and DHW were calculated by lane-changing data. When k-means++ is used to group the driving style into three clusters, the silhouette coefficient of the evaluation index of the model reaches its height. The clustering result is selected as the driving style label value of the track segment, and the features are further extracted to establish the random forest models to identify the driving style. The results show that the accuracy of the random forest model is 93%, thus it is able to identify driving style well. Because different driving styles directly affect the driving characteristic information before the lane-changing, compared with the traditional lane-changing recognition model, this paper adds the driving style variable to the parameters of the lane-changing recognition model to improve the recognition accuracy of the model.

2. XBGoost, LightGBM, and the Stacking fusion model are established to recognize the lane-changing intention. At present, most models only consider the characteristics of a single time point. This study compares the lane-changing intention recognition models that only consider the characteristics of a single time point with those that consider the characteristics of the historical time window. The results show that the model which considers the characteristics of the historical time window can still achieve a recognition rate of more than 85% at 2 s before the starting time point, and the accuracy is higher than that of the model which only considers the characteristics of a single time point. At the same time, all three models can better identify the lane-changing intention, but the Stacking fusion model has the highest comprehensive accuracy, and the training speed of LightGBM is much faster than XGBoost, so it should be gradually applied more and more widely in the practical engineering field.

This paper only analyzes and studies the vehicle driving track in the urban expressway diversion area. The data source is relatively simple, but the vehicle driving area will involve various types of roads and traffic environments. In the future, it is necessary to conduct a comprehensive study on the driver's style, vehicle lane-changing intention recognition, and driving track prediction in different traffic environments.

Author Contributions: Conceptualization, L.Z. and Y.H.; methodology, L.Z. and T.X.; software, L.Z. and Y.H.; validation, L.Z. and Z.Z.; data curation, L.Z., Z.Z. and T.X.; writing—original draft preparation, L.Z.; writing—review and editing, L.Z., Y.H. and T.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant U1664264 and Grant No. 51878066, Funds for Central Universities and Colleges of Chang'an University (No. 300102229201 and No. 300102220204), the Major scientific and technological innovation projects of Shandong Province under Grant No. 2019JZZY020904, and Xi'an scientific and technological projects under Grant No. 2019218514GXRC021CG022-GXYD21.5.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. National Highway Traffic Safety Administration. *Traffic Safety Facts 2015: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System*; National Highway Traffic Safety Administration: Washington, DC, USA, 2017.
2. Chen, T.; Shi, X.; Wong, Y.D. A lane-changing risk profile analysis method based on time-series clustering. *Phys. A Stat. Mech. Its Appl.* **2021**, *565*, 125567. [CrossRef]
3. Gipps, P.G. A model for the structure of lane-changing decision. *Transp. Res. Part B Methodol.* **1986**, *20*, 403–414. [CrossRef]

4. Ahmed, K.I. Modeling Drivers' Acceleration and Lane Changing Behavior. Ph.D. Dissertation, Department of Civil and Environmental Engineering, MIT, Cambridge, MA, USA, 1999.
5. Hidas, P. Modelling lane changing and merging in microscopic traffic simulation. *Transp. Res. Part C Emerg. Technol.* **2002**, *10*, 351–371. [CrossRef]
6. Ratrout, N.T.; Rahman, S.M.; Box, K. A comparative analysis of currently used microscopic and macroscopic traffic simulation software. *Arab. J. Sci. Eng.* **2009**, *34*, 121–133.
7. Zyner, A.; Worrall, S.; Ward, J.; Nebot, J. Long short term memory for driver intent prediction. In Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV), Los Angeles, CA, USA, 11–14 June 2017; pp. 1484–1489.
8. Phillips, D.J.; Wheeler, T.A.; Kochenderfer, M.J. Generalizable intention prediction of human drivers at intersections. In Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV), Los Angeles, CA, USA, 11–14 June 2017; pp. 1665–1670.
9. Kim, I.H.; Bong, J.H.; Park, J.; Park, S. Prediction of driver's intention of lane change by augmenting sensor information using machine learning techniques. *Sensors* **2017**, *17*, 1350. [CrossRef] [PubMed]
10. Guo, Y.; Zhang, H.; Wang, C.; Sun, Q.; Li, W. Driver lane change intention recognition in the connected environment. *Phys. A Stat. Mech. Its Appl.* **2021**, *575*, 126057. [CrossRef]
11. Song, X.; Zeng, Y.; Cao, H.; Li, M.; Yi, B. Lane changing intention recognition method based on an LSTM Network. *China J. Highw. Transp.* **2021**, *34*, 236–245.
12. Zhang, M.; Fu, R.; Morris, D.D.; Wang, C. A Framework for Turning Behavior Classification at Intersections Using 3D LIDAR. *IEEE Trans. Veh. Technol.* **2019**, *68*, 7431–7442. [CrossRef]
13. Li, K.; Wang, X.; Xu, Y.; Wang, J. Lane changing intention recognition based on speech recognition models. *Transp. Res. Part C Emerg. Technol.* **2016**, *69*, 497–514. [CrossRef]
14. Yi, H.; Edara, P.; Sun, C. Modeling Mandatory Lane Changing Using Bayes Classifier and Decision Trees. *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 647–655. [CrossRef]
15. He, P.; Wu, A.; Huang, X.; Rangarajan, A.; Ranka, S. Machine Learning-Based Highway Truck Commodity Classification Using Logo Data. *Appl. Sci.* **2022**, *12*, 2075. [CrossRef]
16. Sultonov, F.; Park, J.H.; Yun, S.; Lim, D.-W.; Kang, J.-M. Mixer U-Net: An Improved Automatic Road Extraction from UAV Imagery. *Appl. Sci.* **2022**, *12*, 1953. [CrossRef]
17. Zyner, A.; Worrall, S.; Nebot, E. Naturalistic driver intention and path prediction using recurrent neural networks. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 1584–1594. [CrossRef]
18. Dou, Y.; Yan, F.; Feng, D. Lane changing prediction at highway lane drops using support vector machine and artificial neural network classifiers. In Proceedings of the 2016 IEEE International Conference on Advanced Intelligent Mechatronics (AIM), Banff, AB, Canada, 12–15 July 2016; pp. 901–906.
19. Peng, J.; Guo, Y.; Fu, R.; Yuan, W.; Wang, C. Multi-parameter prediction of drivers' lane-changing behaviour with neural network model. *Appl. Ergon.* **2015**, *50*, 207–217. [CrossRef] [PubMed]
20. Xu, T.; Zhang, Z.; Wu, X.; Qi, L.; Han, Y. Recognition of lane-changing behaviour with machine learning methods at freeway off-ramps. *Phys. A Stat. Mech. Its Appl.* **2021**, *567*, 125691. [CrossRef]
21. Mammeri, A.; Zhao, Y.; Boukerche, A.; Siddiqui, A.J.; Pekilis, B. Design of a semi-supervised learning strategy based on convolutional neural network for vehicle maneuver classification. In Proceedings of the 2019 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE), Ottawa, ON, Canada, 16–18 October 2019; pp. 65–70.
22. Zhao, S.; Ke, T.; Liu, P. Decision Model of Vehicle Lane-changing Based on Bayesian Network. *J. Chongqing Jiaotong Univ. Nat. Sci.* **2020**, *39*, 130–137+144.
23. Zhu, L.L.; Liu, L.; Zhao, X.P.; Yang, D. Driver Behavior Recognition Based on Support Vector Machine. *J. Transp. Syst. Eng. Inf. Technol.* **2017**, *17*, 91–97.
24. Salvucci, D.D.; Mandalia, H.M.; Kuge, N.; Yamamura, T. Lane-change detection using a computational driver model. *Hum. Factors* **2007**, *49*, 532–542. [CrossRef] [PubMed]
25. Rehder, T.; Koenig, A.; Goehl, M.; Louis, L.; Schramm, D. Lane change intention awareness for assisted and automated driving on highways. *IEEE Trans. Intell. Veh.* **2019**, *4*, 265–276. [CrossRef]
26. Ahmad, I.; Pothuganti, K. Design & implementation of real time autonomous car by using image processing & IoT. In Proceedings of the 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 20–22 August 2020; pp. 107–113.
27. Widaa, A.H.A.; Talha, W.A. Design of Fuzzy-based autonomous car control system. In Proceedings of the 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), Khartoum, Sudan, 16–18 January 2017; pp. 1–7.
28. Van dana Kaur, M. An Improved K-Means Based Text Document Clustering Using Artificial Bee Colony with Support Vector Machine. JETIR. 2021. Available online: www.jetir.org (accessed on 25 September 2022).
29. Kapoor, A.; Singhal, A. A comparative study of K-Means, K-Means++ and Fuzzy C-Means clustering algorithms. In Proceedings of the 2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT), Ghaziabad, India, 9–10 February 2017; pp. 1–6.