International Journal of Engineering Research And Management (IJERM)
ISSN: 2349- 2058, Volume-12, Issue-06, June 2025

Vehicle Forced Lane Change Stage Recognition Based On Trajectory Features

Jianyun Shi, Bowen Ying

Abstract— **Aimed at the problem of vehicle lane change phase recognition affected by the vehicle trajectory characteristics in the merging area of expressways, this study conducted a multi-dimensional analysis of vehicle trajectory characteristics, proposed a force model between vehicles based on the Lennard-Jones potential energy function, and constructed a vehicle lane change phase recognition model based on Particle Swarm Optimization eXtreme Gradient Boosting (PSO-XGBoost).By analyzing the vehicle trajectory data in the merging area of Mengshan Avenue in Linyi City, it was found that the lane-changing points of ramp vehicles followed a Gaussian distribution in space, the lane-changing conditions in the middle of the acceleration lane were better than those at both ends, and most drivers chose to change lanes in the middle. Most vehicles were in the speed range of 45-75 km/h when changing lanes, with 55-60 km/h being the optimal speed range. When driving to the end of the acceleration lane, drivers adopted aggressive driving strategies to force lane changes, resulting in increased risk levels.The intermolecular force model was introduced, and a force model between vehicles based on the Lennard-Jones potential energy function was constructed. The interaction between front and rear vehicles was quantified, and the correlation mechanism between vehicle dynamic game behavior and lane-changing decisions was revealed. On this basis, a vehicle phase recognition model based on PSO-XGBoost was developed, achieving accurate recognition of vehicle lane changes in each phase. The results showed that the classification accuracy of the model was 0.92, and the area under the ROC curve was 0.987. Compared with the LightGBM (LGB), Support Vector Machine (SVM), and Random Forest (RF) , the accuracy improved by 0.03, 0.06, and 0.04, respectively, demonstrating the model's excellent recognition capability.The characteristics of forced lane changing are analyzed based on vehicle trajectory, and the Lennard-Jones potential energy function is introduced to quantify the interaction strength between vehicles. A recognition model of vehicle lane changing phase based on PSO-XGBoost is established. The model is optimized by particle swarm optimization algorithm, which improves the accuracy of the model, and provides a reference for the study of forced lane changing in the merging area.**

Manuscript received July, 2025

Jianyun Shi · School of Transportation Engineering,Dalian Jiaotong University, Dalian, China

Bowen Ying · School of Transportation Engineering,Dalian Jiaotong University, Dalian, China

***Index Terms*—Intelligent transport; Lane change phase identification; XGBoost; Forced lane change**

I. INTRODUCTION

Accurate identification of vehicle lane changing stage is the key link of vehicle driving safety evaluation. The control strategy and index performance of each stage directly affect the safety efficiency of lane changing behavior. However, the existing studies still have obvious deficiencies in the division of lane changing stages. On the one hand, the influencing factors of vehicle lane changing are not clear, and the setting of dividing nodes in each stage is lack of reasonable basis; On the other hand, relying on manual division leads to the reduction of recognition accuracy and limited efficiency. These problems seriously restrict the effectiveness and reliability of lane change safety assessment.

In the study of vehicle lane changing behavior, researchers analyzed the influencing factors of lane changing behavior based on traffic flow theory. Majid Sarvi[1]-[3] conducted an in-depth analysis of the traffic flow characteristics of the freeway merging area under congestion, and built a driver behavior model of the merging area based on empirical research. Song Jia[4] proposed an early lane change model in the pre merge area for the lane change behavior of vehicles in heterogeneous traffic flow environment. Wang Siqi[5] discussed the visual characteristics of drivers in the merging area of high-density interchange groups, and found that the complex environment would increase the driving load. Wang Jiaqi[6] established the minimum spacing model by analyzing the driving characteristics of the driver merging area, and put forward the recommended minimum spacing value of the continuous merging area.

In the research of lane change phase recognition based on model training, Syama[7] established a joint model combining support vector machine and random forest to identify vehicle lane change intention, which improved the accuracy of lane change recognition. Ma Yanfeng[8] took NGSIM data set as a sample to study the execution and adjustment behavior of forced lane change process. Xu Bing[9] analyzed the relative speed and collision time between the target vehicle and other vehicles, and established a lane change decision model based on gradient lifting decision tree.

To sum up, although great progress has been made in the study of vehicle lane change, some studies have problems such as unclear lane change stage division and low stage recognition accuracy. In view of this, the study focuses on the forced lane changing behavior of ramp vehicles in the traffic flow interaction environment in the merging area, analyzes the influence mechanism of vehicle lane changing based on vehicle trajectory, constructs the force model between vehicles based on Lennard-Jones potential function, quantifies the interaction strength between vehicles, and constructs the machine learning model based on PSO-XGBoost to realize the accurate identification of ramp vehicle lane changing stage.

II. PROBLEM AND SCENARIO DESCRIPTION

Lane changing behavior is a continuous process[10]. Combined with the driving habits of actual drivers, the lane changing process of vehicles can be divided into three stages, as shown in Figure 1. The first stage is the vehicle following stage, in which the vehicle keeps running stably in the current state; The second stage is the lane change decision stage, in which the driver will judge whether to change lanes according to the surrounding traffic flow information[11]; The third stage is the lane change execution stage. The driver will adjust the driving state to change the lane to the target lane, so as to drive stably on the target lane.

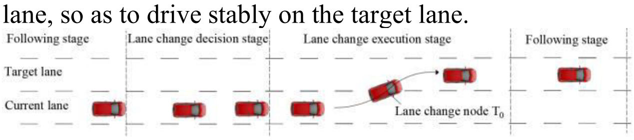

Fig.1 Stage division of vehicle lane changing behavior

The lane changing behavior of ramp vehicles is different from other lane changing behaviors. Because the vehicles on ramp need to merge with the main line vehicles through the acceleration lane[12], but the length of the acceleration lane is limited, the vehicles must complete lane changing before reaching the end of the acceleration lane, which makes them perform forced lane changing[13]. Due to its special lane changing mode, the forced lane changing behavior of vehicles will be studied.

Taking the merging area of Mengshan Avenue Expressway in Linyi City as the research object, the length of the merging area is 120m. The lane in the merging area is four lanes, of which the inner three lanes are the main lane, and the outer single lane is the acceleration lane connecting with the ramp. The UAV is used to capture the vehicle, and the tracker software is used to extract the video track, and the data is post processed.

Fig.2 Top view of the merging area of Mengshan Avenue

III. ANALYSIS ON CHARACTERISTICS OF FORCED LANE CHANGE OF VEHICLES

A. Space distribution

The forward direction of vehicles in the merging area of Mengshan Avenue in Linyi city is set as longitudinal, and the length of acceleration lane is 120m. The road section is divided into 12 sections in 10m units, and the frequency distribution histogram of lane changing behavior is drawn by crossing the lane line as a node.

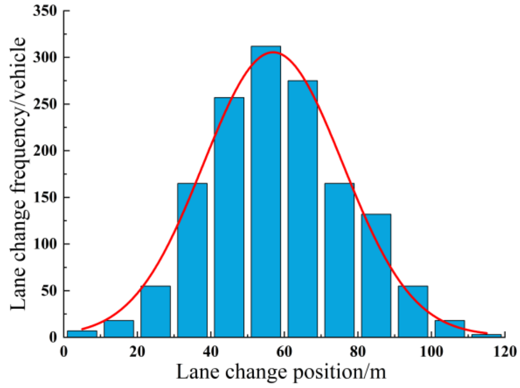

Fig.3 Spatial distribution of lane change

The traffic flow of ramp vehicles is 1468veh/h. From the frequency distribution histogram, it can be seen that about 90% of vehicles choose to change lanes within 30-90m from the starting point of merging; Before the vehicle enters the acceleration lane, due to the barrier and height difference between the ramp and the main road, the driver can not accurately observe the driving situation of the vehicles on the main road, so after entering the acceleration lane, he can not directly change lanes, but chooses to continue driving, and only after confirming the safety of the lane changing environment can he change lanes. When most vehicles change lanes in the middle of the acceleration lane, a few vehicles drive to the end of the acceleration lane. Due to the length limit of the acceleration lane, vehicles have to change lanes. By analyzing the spatial distribution characteristics of lane changing nodes, it can be concluded that the spatial distribution of lane changing nodes conforms to the Gaussian distribution.

B. Vehicle spacing

When the vehicle is changing lanes, affected by the surrounding environment, the driver must consider whether the merging conditions of the vehicle support the vehicle to change lanes safely[14]. During driving, the vehicle is not affected by the front and rear vehicles, including vehicles in the lane where the vehicle is located and vehicles in the target lane[15]. The acceptable spacing of vehicles is related to drivers' habits, road characteristics, vehicle status and other factors. The difference in driving habits makes drivers' judgment of acceptable spacing different, and the diversity of vehicle types and road conditions further aggravates this uncertainty[16]. In order to reduce the interference of uncertain factors, the study will conduct a comprehensive analysis of the collected vehicle trajectories, and draw the density distribution diagram of the spacing between lane changing vehicles and surrounding vehicles with the vehicle crossing the lane line as the node, as shown in Figure 4.

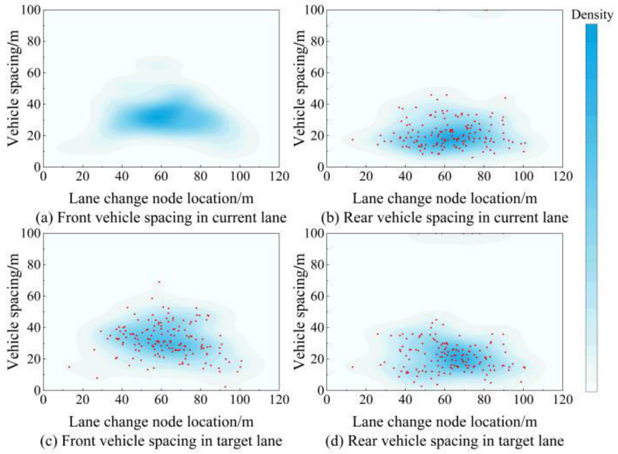

Fig.4 Distribution of vehicle spacing

From the vehicle spacing diagram, it can be found that the front vehicle spacing is mostly within 50m, and the rear vehicle spacing is mostly within 40m. With the change of lane changing nodes, the distribution of vehicle spacing presents the characteristics of diffusion from the middle to the surrounding. The vehicle spacing value in the middle of the acceleration lane is higher than that on both sides, which indicates that the driving environment in the middle of the merging area is more suitable for vehicle lane changing conditions. The vehicles in front of the acceleration lane enter the ramp with high vehicle density, resulting in a small time interval for a few vehicles to change lanes; When the vehicle continues to drive to the middle of the acceleration lane, the vehicle density is low, the vehicle spacing is large, and the lane changing conditions are ideal. Most vehicles choose to change lanes in the middle of the acceleration lane; As a small number of vehicles gradually drive to the end of the acceleration lane, due to the limitation of the length of the acceleration lane, the rigid demand for lane changing behavior can not be met, and drivers will face psychological pressure, tend to accept smaller vehicle spacing, and force lane changing when approaching the end of the acceleration lane.

Comparing the front vehicle spacing and the rear vehicle spacing, it is found that the rear vehicle spacing is less than the front vehicle spacing as a whole. This is because front the vehicle spacing is mainly controlled by the vehicles changing lanes. The vehicles need to accelerate when changing lanes, and the acceleration space needs to be reserved. The distance divergence between the rear workshops is small, because the distance between the rear vehicles is mainly controlled by the rear vehicles. When the rear vehicles are in the following state, the distance between the vehicles remains stable, resulting in the distance divergence between the rear workshops is less than that of the front vehicles.

C. Vehicle speed

Through the extraction of vehicle trajectory information, the speed characteristics of vehicles can be obtained, and the speed of lane changing nodes can be analyzed. It can be seen from Figure 5 that the speed of vehicles at lane changing nodes shows a significant centralized distribution characteristics, and its speed value is mainly distributed in the area of 45-75 km/h, accounting for about 90%. Among them, the speed distribution reaches the peak in the range of 55-60 km/h, accounting for 25.5% of the total sample. This statistical result shows that the speed range of 45-75 km/h can meet the needs of most drivers' lane changing behavior, and 55-60 km/h is the optimal lane changing speed range.

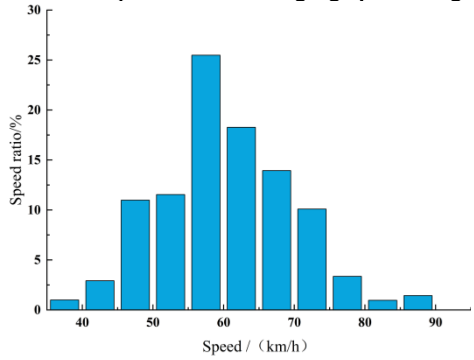

Fig.5 Speed distribution diagram of lane changing nodes

D. TTC

In order to explore the safety of vehicle lane changing, the time to collision(TTC) of vehicle lane changing is used as the index for risk analysis. TTC is defined as the time required for the vehicle to collide with the vehicle in front in the current motion state, and its calculation is shown in equation 2.

$$T=\begin{cases}\dfrac{\Delta l}{\Delta v} & v_{rear}>v_{front}\\ +\infty & v_{rear}\le v_{front}\end{cases} \tag{1}$$

Where T is the collision to time, Δl is the relative distance and Δv is the relative speed.

The study takes the range where the lateral distance between the vehicle and the lane line before and after the lane change node is less than 1m as the analysis interval, and uses TTC as the risk assessment index. According to the TTC, the lane changing process is divided into four risk levels: (0,1] high risk, (1,3] medium risk, (3,5] low risk and (5,+∞) safety state. The trajectory analysis in Figure 6 shows that the high-risk lane changing behavior is mainly concentrated in the second half of the merging area. When the vehicle drives to the end of the acceleration lane for lane changing, the safety level of the vehicle is almost high risk. When the lane changing position of the vehicle is reduced, the risk level will be reduced. The reason for this phenomenon is that as the vehicle approaches the end of the acceleration lane, the remaining driving distance continues to shorten, and drivers tend to adopt aggressive driving strategies to forcibly change lanes when the distance between vehicles is small, resulting in a decrease in TTC and a corresponding increase in collision risk.

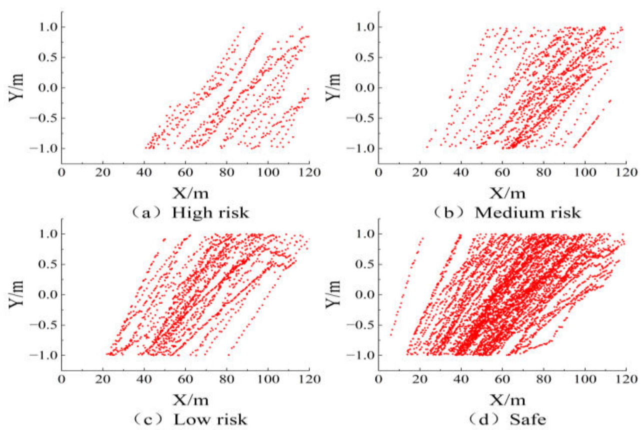

Fig.6 Risk distribution of lane changing

IV. CONSTRUCTION OF DATA SET

A. Determination of starting and ending points of lane change

There has been no unified standard for determining the starting and ending points of lane changing[17]. However, starting from the actual data of this study, because the upstream of the ramp is a curved section, the vehicle can not drive parallel to the lane line after entering the merging area, and has been moving towards the inner side of the merging area in the direction perpendicular to the lane line. The study will judge the starting and ending points of vehicle lane change based on the distance between the midpoint of the vehicle and the lane line. As shown in Figure 7, the two times when the distance between the target vehicle center and the lane line is half of the vehicle width are respectively defined as the beginning and end of lane change[18].

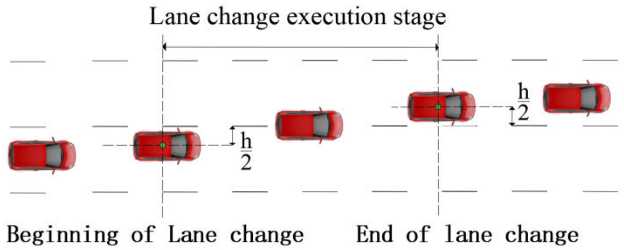

Fig.7 Risk distribution of lane changing

B. Determination of lane change decision point

The interaction between vehicles on the road is similar to the interaction between molecules. The force between molecules will be in the state of attraction, repulsion and equilibrium with the difference of molecular spacing. The interaction between vehicles will be affected by speed, spacing and other factors, and will also be in the state of attraction, repulsion and equilibrium[19]. When the distance between vehicles is large, the vehicles behind will accelerate and gradually reduce the distance between the vehicles in front, corresponding to the attraction state between molecules; When the distance between vehicles is small, the vehicles behind will control the distance within a reasonable range by controlling the speed, which is similar to the repulsive state between molecules; When the distance between vehicles is at a safe distance, the vehicle will drive in a relatively stable state, which corresponds to the equilibrium state between molecules.

The theory is based on the Lennard-Jones potential used in the molecular dynamics model, the potential function and potential field distribution are as follows:

$$U_{6,12}(r)=4\varepsilon\left[\left(\frac{r_0}{r}\right)^{12}-\left(\frac{r_0}{r}\right)^{6}\right] \qquad (2)$$

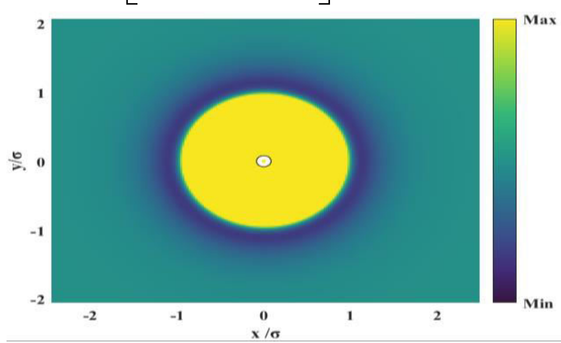

Fig.8 Molecular potential field distribution

However, the gravitational and repulsive terms in Lennard-Jones potential are suitable for the interaction of microscopic particles in fluid mechanics, and can not accurately describe the vehicle following behavior in macro traffic scenes. Therefore, the potential function needs to be properly improved.

The improved potential function is as follows:

$$U_V=\frac{\varepsilon}{v-u}\left[u\left(\frac{X_r}{L}\right)^{v}-v\left(\frac{X_r}{L}\right)^{u}\right] \qquad (3)$$

Where X_r is the vehicle balance distance and L is the actual vehicle distance. Use the actual driving data for parameter calibration[20], v=1.648, u=0.713, ε=48.494×e^3.

For the definition of X_r, researchers believe that the safety distance is related to the speed and speed difference, The X_r can be calculated as:

$$X_r=(t_0-c_v v_r)v_n \qquad (4)$$

Where t_0 and c_v are constant coefficients, which are taken as 1.5 and 0.05 respectively. v_r is the relative speed of the front and rear vehicles, and v_n is the speed of the following vehicle. The displacement derivative of the interaction potential between vehicles is used to obtain the force acting on the rear vehicle:

$$f_V=\frac{vu\cdot\varepsilon}{v-u}\left(\frac{X_r^u}{L^{u+1}}-\frac{X_r^v}{L^{v+1}}\right) \qquad (5)$$

The interaction force of vehicles can be used to quantify the interaction between vehicles. When f_V>0, it shows the gravitational effect, which makes the target vehicle keep following or accelerating; When f_V<0, it is repulsive, and the front vehicle has a suppression effect on the target vehicle. When the vehicle is subjected to the maximum suppression effect of the vehicle ahead, that is, when f_V is the minimum value, the target vehicle has the intention to change lanes because it can not maintain the current driving state.

Considering the dynamic characteristics of vehicles at high speed, the lane change decision point is defined as the time corresponding to the minimum value of f_V within 1-3s before the actual lane change operation. At this time, the rejection effect imposed by the vehicle in front is the most significant, marking the critical trigger point of the driver's lane change intention.

C. Construction of samples

The data used in the study are from the merging area of Mengshan Avenue Expressway in Linyi City, with a length of 120m. Through UAV aerial photography and trajectory extraction, a total of 153 groups of 18126 sequence data can be obtained. At the same time, the data are divided into training set and test set according to the ratio of 4:1.

Through the above analysis of the characteristics of vehicle forced lane change, the factors that have a great impact on vehicle forced lane change will be selected as the model input features.The main research object is the vehicle that has changed lanes, and the surrounding vehicles are the front and rear vehicles in its current lane and the front and rear vehicles in the target lane, mainly including the following characteristics:

Table1 Input feature variables

Feature variable number	Characteristic variable
x_1	Target vehicle forward distance
x_2	Distance from lane markings
x_3	Front vehicle spacing in current lane
x_4	Rear vehicle spacing in current lane
x_5	Front vehicle spacing in target lane
x_6	Rear vehicle spacing in target lane
x_7	The speed of the target vehicle
x_8	Front vehicle speed in current lane
x_9	Rear vehicle speed in current lane
x_{10}	Front vehicle speed in target lane
x_{11}	Rear vehicle speed in target lane
x_{12}	Type of target vehicle
x_{13}	Type of front vehicle in current lane
x_{14}	Type of rear vehicle in current lane
x_{15}	Type of front vehicle in target lane
x_{16}	Type of rear vehicle in target lane

The vehicle type is divided into small vehicle and large vehicle, which are represented by 0 and 1 respectively.

Table2 Output characteristic variables

Feature variable number	Characteristic variable
0	Following stage
1	Lane change decision stage
2	Lane change execution stage

Through the division of vehicle phases, the output characteristics of vehicles can be divided into vehicle following phase, decision-making phase and execution phase, which are represented by 0, 1 and 2 respectively.

V. RECOGNITION MODEL OF VEHICLE LANE CHANGING PHASE BASED ON PSO XGBOOST

A. Construction of XGBoost algorithm

XGBoost is fully known as the extreme gradient lifting model, which was proposed by Chen Tianqi in 2016. It is based on the improvement of the gradient lifting decision tree and is widely used in ensemble learning. Compared with the traditional GBDT, XGBoost has made a number of key improvements, which make XGBoost show excellent performance and generalization ability in machine learning.The specific steps of xgboost algorithm are as follows.

The prediction function expression is:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F \tag{6}$$

Where $\hat{y}_i$ is the prediction result of sample x_i; K is the total number of decision trees; $f_k(x_i)$ is the prediction result of the k tree.

The expression of the objective function is:

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{7}$$

Where: $l(y_i, \hat{y}_i)$ is the prediction error of sample x_i; $\Omega(f_k)$ is the regularization term of the k tree.

In the XGBoost algorithm, gradient boosting strategy is iteratively applied to continuously learn and retain the existing model $\hat{y}_i^{(t-1)}$. After iterations through the t-th and i-th samples, a regression tree $f_k(x_i)$ is generated, enabling the predicted values to progressively approach actual values. Through these iterations, the objective function can be expressed as:

$$Obj^{(t)} = \sum_{i=1}^{n} \left(y_i - \left(\hat{y}_i^{(t-1)} + f_i(x_i) \right) \right)^2 \tag{8}$$

The approximate objective function is obtained by expanding the objective function with the second-order Taylor formula：

$$Obj^{(t)} \approx \sum_{i=1}^{n} \left[\left(y_i - \hat{y}^{(t-1)} \right)^2 + 2\left(y_i - \hat{y}_i^{(t-1)} \right) f_t(x_i) - h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{9}$$

When the constant term is removed, the objective function is only related to the first and second derivatives of the error function. At this time, the objective function can be expressed as:

$$Obj^{(t)} = \sum_{j=1}^{T} \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \tag{10}$$

When w_j takes the minimum value, the objective function is the minimum, and finally the objective function is obtained:

$$Obj = -\frac{1}{2} \sum_{j=1}^{T} \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \tag{11}$$

B. PSO hyperparameter optimization

In the process of model training, the model can run efficiently by controlling the parameters. The research will call XGBoost model directly from the Scikit-learn framework, so the XGBoost classification model will be controlled by the parameters under the Scikit-learn framework.

The following are the common parameters of the model:

N_estimators: the number of iterations, the study will add the early stop method to dynamically select the best number of iterations.

Learning_rate: learning rate, value range [0.01,1].

Max_depth: maximum tree depth, value range [1,10].

Min_child_weight: sum of minimum sample weights of leaf nodes, value range [1,10].

Subsample: proportion of random samples, value range [0.5,1].

Colsample_bytree: random sampling feature ratio, value range [0.5,1].
Gamma: minimum loss reduction value required for splitting nodes, value range [0,5].
Reg\u lambda: controls the L2 regularization intensity of leaf node weights, with a value range of [1,10].
Objective: learning objectives, the research is a multi classification problem, so multi:softmax is selected.
The research will adopt the particle swarm optimization model to optimize the super parameters. The particle swarm optimization algorithm is derived from the foraging behavior of birds. The principle of particle swarm optimization algorithm is that the behavior of particles is affected by their own experience and group experience, and gradually approaches the optimal solution by adjusting their own speed and position. The iterative formula of its speed and position is as follows.

$$v_{id}^{t+1} = w \times v_{id}^{t} + c_1 r_1 \left(p_{id,pbest}^{t} - x_{id}^{t}\right) + c_2 r_2 \left(p_{id,gbest}^{t} - x_{id}^{t}\right) \tag{12}$$

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t+1} \tag{13}$$

Where v is the particle speed; w is inertia weight; c_1 and c_2 denote individual learning factor and social learning factor respectively; r_1 and r_2 are random numbers between [0,1] respectively; p_{pbest} and p_{gest} are the individual optimal solution and the overall optimal solution obtained after several iterations respectively; x is the position vector of the particle; i is the particle serial number, i=1,2,3...,N; d represents particle dimension, d=1,2,3...,D.
After super parameter optimization, the following parameters can be obtained: N_estimators=336, Learning_rate=0.01,Max_depth=3,Min_child_weight=4.292 , Subsample=0.665 , Colsample_bytree=0.630 , Gamma=0 , Reg_lambda=7.989.

C. Training results

According to the training results of the model, the study will select different evaluation indexes to analyze the performance of PSO-XGBoost model. The performance evaluation indexes of the model include confusion matrix analysis, subject working characteristic curve, accuracy rate, recall rate and F1 score, so as to comprehensively quantify the classification performance of different algorithms. The performance of the four classifiers on the test set is comprehensively evaluated through comparative experiments, using Light Gradient Boosting Machine (LGB), Support Vector Machine (SVM) and Random Forest (RF) as comparative models.
Each row of the confusion matrix represents the real category, and the sum of each row represents the real quantity under the category; Each column represents the forecast category, and the sum in each column represents the forecast quantity under this category. The quantity in each column represents the actual value predicted as the quantity of the category. Combined with the confusion matrix in Figure 9 and the analysis of evaluation index table 3, the overall accuracy of the constructed PSO-XGBoost model is 0.92, and its recall rate and F1 score are 0.92, and the three indicators are more than 0.9, which shows that the model can accurately identify the following stage, lane change decision-making stage and lane change execution stage.

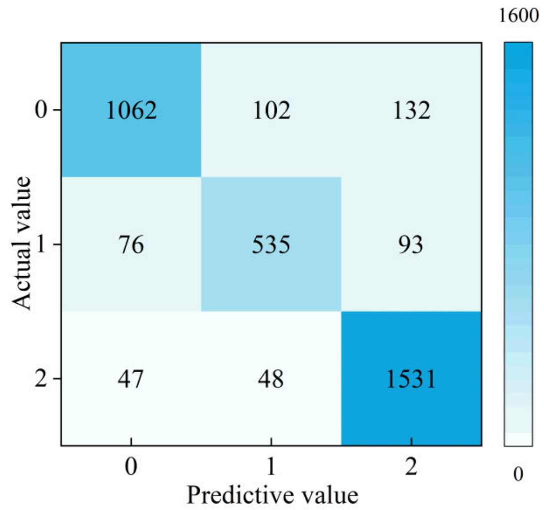

Fig.9 Confusion matrix diagram

Table3 Model performance evaluation indicators

Model	Evaluation	Stage			
		Following	Decision	Execute	Total
XGBoost	Precision	0.90	0.82	0.97	0.92
	Recall	0.90	0.87	0.95	0.92
	F1-score	0.90	0.84	0.96	0.92

To further verify the performance of the model, the recognition results of the model are compared with those of LGB, SVM and RF models. Table 4 shows the evaluation indexes of each model. Compared with LGB, SVM and RF models, the accuracy of PSO-XGBoost model is improved by 0.03, 0.06 and 0.04 respectively, which shows that the introduction of xgboost optimized by particle swarm optimization can better adjust the parameters of the training model, so that the model can identify each stage of vehicle lane changing behavior with high-precision standards.

Table4 Model performance evaluation indicators

Model	Evaluation	Stage			
		Following	Decision	Execute	Total
LGB	precision	0.84	0.85	0.96	0.89
	recall	0.92	0.74	0.93	0.89
	f1-score	0.88	0.79	0.95	0.89
SVM	precision	0.90	0.78	0.87	0.86
	recall	0.82	0.76	0.94	0.86
	f1-score	0.86	0.77	0.91	0.86
RF	precision	0.94	0.70	0.91	0.88
	recall	0.80	0.80	0.96	0.87
	f1-score	0.87	0.74	0.94	0.87

Figure 10 shows the ROC curve of each model. The ROC curve visually shows the discrimination ability of the classification model by drawing the corresponding relationship between the true positive rate and the false positive rate under different classification thresholds. Among them, the morphological characteristics and AUC value of ROC curve are important performance evaluation indicators: the closer the curve is to the upper left corner, the higher the AUC value, indicating that the better the prediction performance of the model.
In the ROC curve comparison, the ROC curve of xgboost model is the closest to the upper left corner, and the AUC value is the closest to 1 in the quantitative comparison. This result further highlights the superior performance of XGBoost algorithm in vehicle lane change phase recognition.

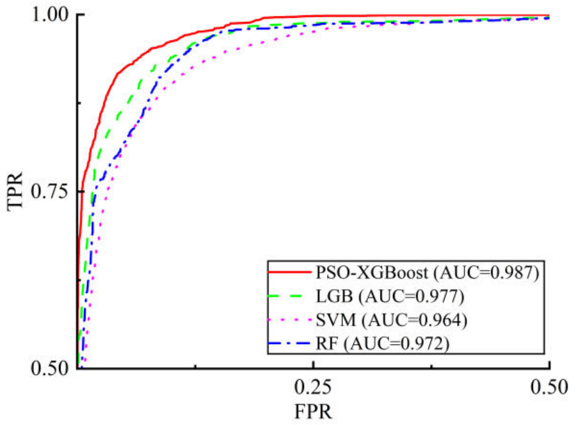

Fig.10 ROC curve diagram

VI. CONCLUSION

(1)The lane changing behavior of ramp vehicles presents Gaussian distribution in space. The distance between vehicles in the middle of the acceleration lane is greater than that at both ends, and the lane changing environment is better than that at both ends. Most drivers will choose to change lanes in the middle. By quantifying the risk of vehicle lane change, it is found that when the vehicle reaches the end of the acceleration lane, the driver will take aggressive driving strategies to force lane change, resulting in an increase in the risk level.

(2)The intermolecular force model is innovatively introduced to construct the force model between vehicles based on the Lennard Jones potential energy function. The force expression between vehicles obtained by improving the potential energy function can effectively express the attraction and repulsion effect between vehicles, quantify the strength of the interaction between vehicles, reflect the relationship between the dynamic game behavior of vehicles and the lane change decision, and provide a theoretical basis for the decision-making point of lane change.

(3)A lane change phase recognition model based on PSO xgboost is constructed. The parameters of the limit gradient lifting model are optimized by particle swarm optimization algorithm to improve the overall performance of the model. The overall classification accuracy of the model is 0.92, which is 0.03, 0.06 and 0.04 higher than that of LGB, SVM and RF models, respectively. This shows that the constructed model has excellent recognition performance after particle swarm optimization.

The research adopts the research method of combining actual data analysis and model building, and completes the identification of forced lane change of vehicles in the merging area of expressways, which provides a scientific basis for traffic control in the merging area under the intelligent network environment, and also provides an important reference for the construction of traffic safety evaluation system. In the follow-up study, the data dimension of the model can be further expanded by combining the driver's psychological characteristics and physiological indicators, so as to improve the adaptability and robustness of the model in different traffic scenarios.

REFERENCES

[1] Sarvi M ,Kuwahara M ,Ceder A .Observing freeway ramp merging phenomena in congested traffic[J].Journal of Advanced Transportation,2007,41(2):145-170.
[2] Sarvi M, Kuwahara M. Using ITS to improve the capacity of freeway merging sections by transferring freight vehicles.[J].IEEE transactions on intelligent transportation system,2008,9(4):580-588.
[3] Sarvi M, Kuwahara M. Microsimulation of freeway ramp merging processes under congested traffic conditions[J].IEEE transactions on intelligent transportation systems,2007,8(3):470-479.
[4] SONG Jia.Research on Vehicle Lane-changing behavior in freeway merging areas in heterogeneous traffic flow environment[D].Jilin University,2023.
[5] WANG Siqi.Study on the visual characteristics of drivers in merging areas of high-density interchange groups on expressways[D].Chongqing Jiaotong University,2023.
[6] WANG Jiaqi,HUANG Jieling,GUO Jiangang,et al.Study on minimum space of continuous merging zone of urban river crossing bridges[J].Journal of Dalian Jiaotong University,2024,45(03):31-37.
[7] Syama R., Mala C. An ensemble model for lane change intention inference for autonomous driving[C]. 2022 International Conference on Computing, Communication, Security and Intelligent Systems (IC3SIS). IEEE, 2022: 1-6.
[8] MA Yanfeng.Research on vehicles' mandatory lane change behavior in execution and post-adjustment period on multi-lane freeway based on ngsim trajectory data[D].Chang'an University,2020.
[9] XU Bing,Liu Xiao,Wang Ziyang,et al.Fusion decision model for vehicle lane change with gradient boosting decision tree[J].Journal of Zhejiang University (Engineering Science) ,2019,53(06):1171-1181.
[10] SHI Jianyun,HUANG Huiyuan.Optimization of vehicle routing for hazardous material transportation in time-varying risk environments[J].Journal of Safety and Environment,2024,24(11):4416-4423.
[11] Wen P .Minimum-time lane changing problem of vehicle handling inverse dynamics based on adaptive mesh refinement and collocation optimization method[J].Journal of Vibroengineering,2023,25(6):1198-1216.
[12] Matthew V ,Long R C .Fault tolerance analysis of an adaptive neuro-fuzzy inference system for mandatory lane changing decisions in automated driving[J].International Journal of Transportation Science and Technology,2023,12(2):594-605.
[13] DING Yang,ZHONG Hong,QIAN Yahui, et al.Lane-Change collision avoidance control for automated vehicles with control barrier functions[J].International Journal of Automotive Technology,2023,24(3):739-748.
[14] Rizehvandi A ,Azadi S ,Eichberger A .Enhancing highway driving: high automated vehicle decision making in a complex multi-body simulation environment[J].Modelling,2024,5(3):951-968.
[15] HAO Xiaohui,XIA Yuanqing,YANG Hongjiu, et al.Typical motion-based modelling and tracking for vehicle targets in linear road segment[J].International Journal of Systems Science,2024,55(5):833-843.
[16] XIA Taokai ,CHEN Hui ,SU Shaoka .Risk field enhanced game theoretic model for interpretable and consistent lane-changing decision makings[J].SAE International Journal of Advances and Current Practices in Mobility,2024,7(2):827-837.
[17] Mohamadreza H ,Aryan H ,Robert K .A lane-changing behaviour algorithm for connected and autonomous vehicles to facilitate emergency vehicle movements on urban arterials in microsimulation[J].Journal of Simulation,2024,18(3):413-427.

[18] Matthew V ,Long R C .Comparative evaluation of adaptive fuzzy inference system and adaptive neuro-fuzzy inference system for mandatory lane changing decisions on freeways[J].Journal of Intelligent Transportation Systems,2022,26(6):746-760.
[19] SU Xiaozhi,MA Yanfeng,LI Rui, et al.Placement distance of exit advance guide sign on an eight-lane expressway considering lane changing behaviour in China[J].Promet - Traffic&Transportation,2022,34(2):285-295.
[20] JIA Yanfeng,QU Dayi,ZHAO Zixu,et al.Car-following Decision-making and model for connected and autonomous vehicles based on safety potential field[J].Journal of Transportation Systems Engineering and Information Technology,2022,22(01):85-97.