Intelligent Vehicle Lane Change Risk Prediction Model Based on Deep Learning and Interactive Features

Lingyi Meng

School of Transportation, Shandong University of Science and Technology, Qingdao, Shandong 266590, China

Abstract: ***With the development of autonomous driving technology, predicting the risk of vehicle lane changing has become one of the key tasks to improve the safety and performance of intelligent vehicles. Traditional lane change risk prediction methods often overlook the complex interaction characteristics between lane changing vehicles and the surrounding environment, resulting in insufficient prediction accuracy. This article proposes an intelligent vehicle lane change risk prediction method based on interactive features. Firstly, based on the high-D dataset, the trajectory data of surrounding interactive vehicles is extracted. The parking distance index is used to establish a risk assessment index, and the risk level is divided through clustering algorithm. Finally, the LSTM model was used for risk prediction, and the experimental results showed that the LSTM model with added interactive features had better prediction performance.***

Keywords: Interactive features; Intelligent vehicles; Risk prediction for lane changing; LSTM model.

1. INTRODUCTION

With the continuous increase in the number of motor vehicles, the transportation system is becoming increasingly complex, and the problem of traffic congestion is worsening. To enhance safety and alleviate traffic pressure, China actively promotes the intelligentization and networking of automobiles. However, the coexistence of autonomous and manually driven vehicles in mixed traffic environments increases traffic complexity and poses challenges to the safe operation of autonomous vehicles. The lane changing behavior involves longitudinal and lateral movements, as well as interactions with surrounding vehicles, which increases driving complexity. Identifying potential risk lane changing situations in advance is crucial for assessing traffic risks and preventing accidents.

Jin (2025) proposed an attention-based temporal convolutional network combined with reinforcement learning (RL) to predict delays and optimize inventory, offering a data-driven approach to logistical challenges [1]. Similarly, Wang and Liang (2025) applied RL with graph neural networks and self-attention mechanisms to enhance supply chain route optimization, further highlighting AI's role in operational efficiency [6]. Beyond logistics, Shen et al. (2025) developed an AI system using long short-term memory (LSTM) networks to manage anesthetic dosing in cancer surgery, showcasing clinical applicability [2]. In cybersecurity, Xu et al. (2025) investigated adversarial machine learning, analyzing attack strategies and defensive frameworks to safeguard systems [3]. Autonomous systems also benefit from AI, as seen in Wang et al. (2025)'s end-to-end autonomous driving model, which integrates multimodal data for real-time decision-making [4]. Computer vision techniques, such as Zhou et al. (2024)'s ResNet-50 and weakly supervised CNN model, have improved garbage recognition for sustainable urban development [5]. Meanwhile, Xie et al. (2025) introduced RTop-K, an ultra-fast GPU-accelerated algorithm for top-k selection, optimizing neural network inference [7]. In healthcare and behavioral research, Lin et al. (2025) demonstrated that AI-driven physical exercise monitoring enhances executive function in children with ADHD [8], while Peng et al. (2025) explored how aerobic exercise intensity influences cognition and sleep, underscoring AI's utility in health analytics [9]. Lastly, Tang and Zhao (2025) employed neural networks to analyze aging populations' impact on real estate dynamics, linking demographic trends to economic models [10]. Collectively, these studies reflect AI's transformative potential, from technical innovations (e.g., [1, 3, 7]) to societal applications (e.g., [2, 8, 10]), while emphasizing interdisciplinary collaboration. Traditional methods mostly rely on vehicle dynamics models and rule-based algorithms to estimate lane change risk by analyzing the vehicle's position, velocity, acceleration, and other characteristics. For example, some studies determine whether a vehicle is at risk of collision by setting a safe distance threshold. However, these methods ignore the interactive effects between vehicles during lane changing and are difficult to cope with complex traffic environments.

2. ANALYSIS OF VEHICLE CHARACTERISTICS IN SURROUNDING INTERACTIONS

2.1 Impact of Surrounding Interactive Vehicles

2.1.1 Impact of vehicles in front of the original lane

During the deviation phase of free lane changing, the lane changing vehicle gradually leaves the original lane. If the longitudinal distance from the vehicle in front of the original lane is not accurately evaluated, or if the vehicle speed is adjusted improperly, it may lead to a rear end collision. When the lane changing vehicle accelerates too fast or fails to slow down in time, the safety distance from the vehicle in front of the original lane is shortened, and the risk of collision increases. This type of collision usually occurs in situations where the vehicle's dynamic changes are significant, reflecting the requirements for longitudinal distance and speed control during the deviation phase.

2.1.2 Impact of vehicles in front of the target lane

During the cross lane phase, the lane changing vehicle must maintain a safe distance from the vehicle in front of it when entering the target lane. When the lane changing vehicle fails to accurately determine the speed or dynamic changes of the preceding vehicle, or fails to adjust its own speed in a timely manner, a rear end collision may occur. If the vehicle in front of the target lane suddenly decelerates and the lane changing vehicle has insufficient response, the risk of collision will significantly increase. Matching the distance and speed between the lane changing vehicle and the vehicle in front of the target lane is an important factor in avoiding such collisions.

2.1.3 Impact of vehicles behind the target lane

During the cross lane and return to normal phases of free lane changing, the lane changing vehicle needs to coordinate with the vehicle behind the target lane to complete merging. If the lane changing vehicle fails to accurately assess the speed and distance of the following vehicle in the target lane, it may interfere with the following vehicle when entering the lane, which may result in the following vehicle being unable to avoid in a timely manner, leading to rear or side collisions. This type of collision usually stems from insufficient perception of the dynamic changes of the vehicle behind the target lane during lane changing, as well as deviations in the vehicle's path planning.

2.2 Data Extraction

This article uses the German highway large-scale natural driving trajectory dataset highD for research, which has sufficient vehicle interaction data. The specific steps for extracting data are as follows:

(1) Exclude non automotive vehicle trajectories from the high-D dataset;

(2) Determine the lane change timestamp by analyzing the lane ID changes of adjacent frames in the LaneID column;

(3) Starting from the lane change timestamp, search forward and backward for timestamps with four consecutive frames of lateral displacement less than 0.03m to determine the start and end time of the lane change;

(4) Extract trajectory data of vehicles around the lane changing vehicle.

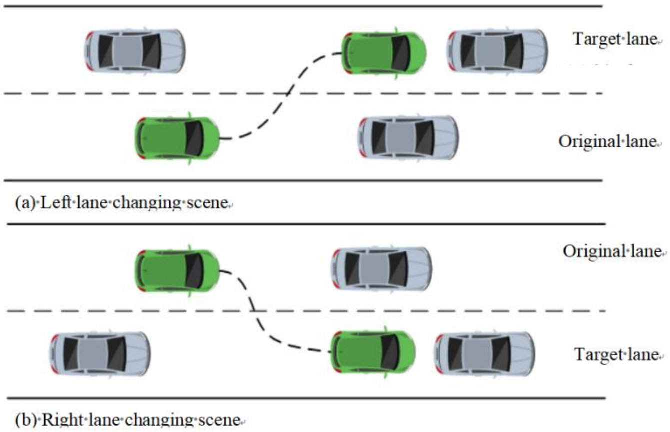

Figure 1: Lane changing scene extraction

3. RISK PREDICTION OF LANE CHANGING CONSIDERING INTERACTIVE FEATURES

3.1 Classification of Lane Changing Risk Levels Based on Cluster Analysis

This article establishes an evaluation index for interactive conflicts during lane changing based on the Parking Distance Index (SDI). Firstly, calculate the stopping distance index between lane changing vehicles and surrounding interacting vehicles, and then use the fuzzy c-means clustering algorithm to achieve unsupervised risk level labeling. Risk level classification clustering center results:

(1) Security level

The clustering center is (97.84123.85119.68), indicating that under this risk level, the SDI values of the three groups are all in a high range, indicating that the safe distance between the changing lane vehicles and surrounding vehicles is large, the speed difference is small, and the acceleration changes are stable. Therefore, the risk of lane changing is relatively low, the relative position and dynamic behavior between vehicles are relatively stable, the safety of lane changing vehicles is high, and the collision risk is low.

(2) Low risk level

The clustering center is (93.02, 48.76, 41.11), and the SDI value of this level is significantly lower than the safety level, indicating that during the lane change process, although the safety distance is relatively large, the speed and acceleration differences between vehicles have increased, and there may be a brief state of tension. At this point, although a certain level of safety can still be maintained, the behavior of surrounding vehicles and the relative position of lane changing vehicles may pose a certain degree of risk, and vigilance should be maintained.

(3) Medium risk level

The clustering centers are (39.58, 32.07, 38.00), and the SDI value is low, which means that the distance between lane changing vehicles and surrounding vehicles becomes smaller, the speed difference increases, and the acceleration fluctuation is larger. In this case, vehicles changing lanes may face complex traffic conditions, with a certain risk of collision during the lane changing process. The dynamic interaction between vehicles is more intense, resulting in lower safety.

(4) High risk level

The clustering centers are (24.45, 22.42, 32.07), and all three SDI values are in a low range, indicating that the distance between lane changing vehicles and surrounding vehicles is extremely small, with significant differences in speed and acceleration, and the interaction between vehicles is very tense. In this situation, vehicles changing lanes are highly likely to face collisions or other significant risks, and lane changing operations should be carried out with great caution and emergency avoidance measures should be taken quickly.

3.2 Determination of Network Model Structure and Parameter Setting

The model selected a three-layer neural network structure model, including a data input layer LSTM hidden layer and data output layer. Parameter settings:

(1) Learning rate: 0.01;

(2) Batch_size: 32;

(3) Dropout: 0.5;

(4) Number of LSTM neurons: 64;

(5) LSTM input_dim :24;

(6) LSTM hidden_dim: 64;

(7) LSTM output_dim: 4;

(8) Loss function: categorical_crossentropy;

(9) Epoch: 200.

3.3 Feature Selection

When using LSTM neural network for lane change decision prediction, the states of the vehicle and surrounding vehicles need to be considered, so interactive features are included in the risk prediction model. The model feature parameters are as follows:

(1) Self driving features:

Lateral position, longitudinal position, lateral velocity, longitudinal velocity, lateral acceleration, longitudinal acceleration.

(2) Interactive features:

Vehicle in front of the original lane: lateral position, longitudinal position, lateral velocity, longitudinal velocity, lateral acceleration, longitudinal acceleration;

Target lane preceding vehicle: lateral position, longitudinal position, lateral velocity, longitudinal velocity, lateral acceleration, longitudinal acceleration;

Target lane rear vehicle: lateral position, longitudinal position, lateral velocity, longitudinal velocity, lateral acceleration, longitudinal acceleration.

4. RESULTS AND ANALYSIS

This article uses natural driving trajectory data as the training raw data, dividing the data into an 80% training set and a 20% validation set. Select accuracy, precision, recall, and F1 score as evaluation metrics for the model.

Table 1: Model Evaluation Indicators

Input features	Accuracy	Precision	Recall	F1 score
Self driving characteristics	79.47	82.10	79.47	78.60
Self driving features+interactive features	82.65	83.03	82.65	82.52

In order to verify the performance of the lane change risk prediction model based on LSTM established in this article, the feature values of the self vehicle features plus interactive features and the feature values without interactive features were input separately. According to the lane change risk prediction results in Table 9, it can be seen that when interactive features are added to the input features, the accuracy of the LSTM model is 82.65%, while the accuracy of the model without interactive features is 79347%. This indicates that the LSTM model considering interactive features proposed in this paper has better performance in lane change risk prediction.

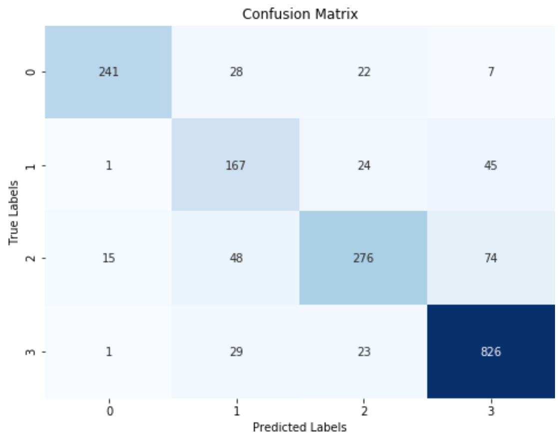

Figure 2: Confusion matrix without added interactive features

In Figure 2, label 0 represents high risk, label 1 represents medium risk, label 2 represents low risk, and label 3 represents safety.

From the figure, it can be seen that the model can better utilize time series features and interaction features after adding interaction features, significantly improving classification accuracy, especially in high-risk categories where it performs the best, correctly identifying 241 samples, significantly higher than the LSTM model without interaction features. The LSTM model can better combine time series features and interaction features, effectively improving the recognition ability of high-risk samples.

5. CONCLUSION

This article studies the risk prediction problem of intelligent vehicle lane changing. Firstly, the impact of surrounding interactive vehicles was analyzed, and lane changing vehicles and trajectory data of surrounding vehicles were screened; Then, based on SDI, interactive risk assessment indicators are constructed, and risk levels are classified through fuzzy c-means clustering; Finally, the LSTM model was used for risk prediction, and the interaction features were input as feature values. The experimental results showed that the model with added interaction features had higher prediction accuracy than the model without added interaction features.

REFERENCES

[1] Jin, T. (2025). Attention-Based Temporal Convolutional Networks and Reinforcement Learning for Supply Chain Delay Prediction and Inventory Optimization.

[2] Shen, Z., Wang, Y., Hu, K., Wang, Z., & Lin, S. (2025). Exploration of Clinical Application of AI System Incorporating LSTM Algorithm for Management of Anesthetic Dose in Cancer Surgery. Journal of Theory and Practice in Clinical Sciences, 2, 17-28.

[3] Xu, J., Wang, Y., Chen, H., & Shen, Z. (2025). Adversarial Machine Learning in Cybersecurity: Attacks and Defenses. International Journal of Management Science Research, 8(2), 26-33.

[4] Wang, Y., Shen, Z., Hu, K., Yang, J., & Li, C. (2025). AI End-to-End Autonomous Driving.

[5] Zhou, Y., Wang, Z., Zheng, S., Zhou, L., Dai, L., Luo, H., ... & Sui, M. (2024). Optimization of automated garbage recognition model based on resnet-50 and weakly supervised cnn for sustainable urban development. Alexandria Engineering Journal, 108, 415-427.

[6] Wang, Y., & Liang, X. (2025). Application of Reinforcement Learning Methods Combining Graph Neural Networks and Self-Attention Mechanisms in Supply Chain Route Optimization. Sensors, 25(3), 955.

[7] Xie, X., Luo, Y., Peng, H., & Ding, C. RTop-K: Ultra-Fast Row-Wise Top-K Selection for Neural Network Acceleration on GPUs. In The Thirteenth International Conference on Learning Representations.

[8] Lin, L., Li, N., & Zhao, S. (2025). The effect of intelligent monitoring of physical exercise on executive function in children with ADHD. Alexandria Engineering Journal, 122, 355-363.
[9] Peng, Y., Zhang, G., & Pang, H. (2025). Impact of Short-Duration Aerobic Exercise Intensity on Executive Function and Sleep. arXiv preprint arXiv:2503.09077.
[10] Tang, Y., & Zhao, S. (2025). Research on Relationship Between Aging Population Distribution and Real Estate Market Dynamics based on Neural Networks.