Multi-Feature-Based Integrated Vehicle Trajectory Prediction with Driving Intention Recognition

Shuobo Wang [1, *], Shijie Liu [2]

[1] School of Mechanical and Precision Instrument Engineering, Xi'an University of Technology, Xi'an, China, 710048

[2] Department of Mechatronics and Vehicle Engineering, Taiyuan University, Taiyuan, China, 030032

* Corresponding Author Email: shuobo0629@163.com

Abstract. Vehicle trajectory prediction and driving intention recognition are essential technologies for improving safety and optimizing traffic efficiency in autonomous driving systems. Although traditional LSTM models are effective in trajectory prediction, their complexity and computational requirements impede practical implementation efficiency. To overcome this challenge, this study proposes a hybrid model, MTF-GRU, which integrates multi-feature fusion. Initially, the datasets are preprocessed in this study through denoising, feature extraction, and timing extraction to capture vehicle information from single and fused multi-features. Subsequently, a GRU encoding-decoding model is developed. The encoder processes the feature data to generate context vectors, while the decoder employs a combination of recursive and teaching-driven input modes. Furthermore, a teaching rate control mechanism is integrated to dynamically convert context vectors into future trajectories. The proposed model is validated using the NGSIM datasets, demonstrating superior prediction performance with multi-feature inputs outperforming single features by reducing the average endpoint displacement error by 20.5%. Our model also achieves improved accuracy rates, particularly excelling in long-term predictions with an endpoint displacement error of only 2.31 meters at 5 seconds. Moreover, the overall accuracy rate for lane change intention recognition reaches 91.3%. The model's computational efficiency supports practical deployment in real-time autonomous systems, while future efforts will integrate multi-modal sensor data to enhance adaptability in complex urban scenarios and extreme conditions. Further validation will extend to diverse traffic environments and edge computing platforms to optimize real-world robustness.

Keywords: Intelligent Connected Vehicles, Multi Feature Fusion, Vehicle Trajectory Prediction, Driving Intention Recognition.

1. Introduction

Within the exponential advancement of intelligent transportation ecosystems, contemporary intelligent connected vehicles (ICVs) harness cutting-edge multisensor architectures to acquire high-fidelity environmental cognition while orchestrating real-time multimodal data symbiosis across vehicular-infrastructure networks. This technological confluence facilitates unprecedented optimization of traffic management paradigms through enhanced operational efficiency and robust safety protocols. Central to this innovation matrix reside two critical enablers:Vehicle Trajectory Prediction and Driving Intention Recognition. These dual technological competencies not only elevate autonomous navigation safety thresholds and optimize traffic hydrodynamic efficiency, but also establish the computational bedrock for next-generation driver-assistance architectures and autonomous decision-making ecosystems. Through accurate vehicle trajectory prediction frameworks, ICVs attain heightened environmental perceptivity to navigate dynamic traffic scenarios with predictive responsiveness [1]. Simultaneously, driving intention recognition algorithms facilitate anticipatory modeling of agent behaviors, enabling recursive optimization of decision-making heuristics [2]. This synergistic integration drives substantive enhancements in both transportation network throughput and collision mitigation efficacy. The transformative potential inherent in these technological frontiers has precipitated intensified scholarly investigation, positioning this discipline as a critical catalyst in redefining intelligent mobility paradigms.

Early vehicle trajectory prediction methodologies primarily relied on physical modeling approaches. These methods conceptualized vehicles as physical entities obeying kinematic and dynamic laws to construct motion prediction frameworks. Xie et al. developed an optimized trajectory selection framework incorporating behavioral recognition and curvature constraints, achieving a computational latency of 0.103 milliseconds [3]. Comparatively, Feng et al.proposed an uncertainty-aware prediction model using second-order Markov chains, which exhibited 40% higher accuracy than first-order models [4].

The complex and highly dynamic operational environments of ICVs impose stringent demands on prediction accuracy, rendering conventional modeling approaches insufficient for real-world deployment. In contrast, data-driven methodologies have emerged as promising solutions due to their superior capacity to learn intricate patterns. By leveraging sophisticated non-linear architectures, these approaches effectively capture the complex spatiotemporal features and latent regularities of vehicular motion, thereby enhancing model generalization and predictive performance. In research on such methods, Jin et al. employed a driving-centric perspective to design a long short-term memory (LSTM)-based framework for extracting features from historical trajectories [5]. Ji et al. proposed a joint LSTM architecture for driving intention recognition and trajectory prediction, which enhanced interpretability of complex interactive behaviors [6]. Meng et al. introduced a hybrid system combining continuous hidden Markov models for lane-changing intention recognition with an advanced LSTM model for trajectory prediction, demonstrating significant gains in both prediction precision and reliability [7]. Fang et al. incorporated a mixed teacher forcing decoding mechanism into LSTM networks, improving their ability to learn real-world trajectory data [8]. Li et al. developed a bidirectional graph-LSTM architecture that incorporates interactive features from surrounding vehicles and historical motion patterns of target vehicle [9].

In current research, mainstream approaches to trajectory prediction and lane change intention recognition predominantly employ LSTM-based learning frameworks. However, LSTM models exhibit high parameter complexity and computational costs, limiting their practical implementation efficiency. To address these challenges and enhance prediction accuracy and computational efficiency in complex dynamic traffic environments, this study proposes a Mixed Teaching Forcing Gated Recurrent Unit(MTF-GRU) hybrid model for integrated trajectory prediction and driving intention recognition. This framework aims to develop a more comprehensive intelligent driving perception system by leveraging multi-feature fusion strategies. The MTF-GRU model adopts structural principles from hybrid teaching frameworks, replacing traditional LSTM architectures with GRU networks, which offer reduced computational overhead. Furthermore, this study examines the modeling impact of fusing surrounding environmental information specifically related to bicycles. It investigates how multi-feature fusion affects trajectory prediction and the recognition of lane change intentions. Experimental results demonstrate that the MTF-GRU model achieves comparable prediction accuracy to state-of-the-art methods while significantly improving computational efficiency. These advancements provide strong technical support for trajectory forecasting and driving decision-making in ICVs.

2. Introduction to the Model Principle

The trajectory prediction model proposed in this study employs a Gated Recurrent Unit (GRU), a specialized variant of recurrent neural networks (RNN) optimized for temporal data processing. GRU leverage update and reset gates to dynamically regulate information flow across sequences, addressing key limitations of traditional RNN architectures. While preserving RNN's capacity to handle sequential dependencies, GRU mitigate gradient vanishing and explosion phenomena in long-sequence scenarios, significantly improving long-term dependency modeling. The update gate modulates the integration of historical hidden states with current candidate states, whereas the reset gate governs the influence of prior states on new inputs. This synergistic gating mechanism enables simultaneous extraction of local temporal patterns and global contextual features. Furthermore, GRU exhibit structural simplicity and parameter efficiency compared to LSTM networks, achieving comparable predictive performance with reduced computational overhead. These advantages have established GRU as a cornerstone architecture for time-series forecasting, natural language processing, and speech recognition applications. The model's robustness and computational efficiency position it as a critical framework for analyzing complex temporal relationships in trajectory prediction tasks. The model framework diagram is presented as shown in Figure 1.

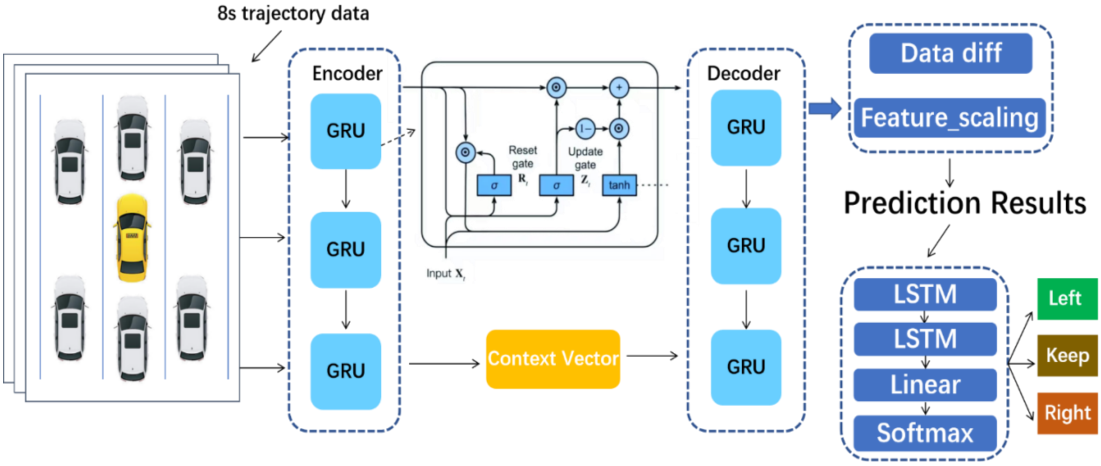

Figure 1. Model framework diagram

2.1. GRU Encoder

To develop a robust trajectory prediction model, this study extracted and fused features from the target vehicle, surrounding vehicles, and road environment using the filtered US101 datasets. At each time step *t*, six surrounding vehicles were selected to comprehensively assess their influence on the target vehicle: those positioned directly in front and behind, as well as vehicles in the left-front, left-rear, right-front, and right-rear quadrants relative to the target vehicle, as shown in Table. 1.

Table 1. Features description table of input data

Parameter	Description	Calculation Formula
x^t, y^t	Lateral & Longitudinal Coordinates of Target Vehicle	/
vx^t, vy^t	Lateral & Longitudinal Instantaneous Velocity of Target Vehicle	$vx^t = \frac{x^{t+1} - x^t}{t}, vy^t = \frac{y^{t+1} - y^t}{t}$
ax^t, ay^t	Lateral & Longitudinal Instantaneous Acceleration of Target Vehicle	$ax^t = \frac{vx^{t+1} - vx^t}{t}, ay^t = \frac{vy^{t+1} - vy^t}{t}$
X_i^t, Y_i^t	Relative Lateral & Longitudinal Coordinates of Surrounding Vehicle *i*	$X_i^t = x_i^t - x^t, Y_i^t = y_i^t - y^t$
VX_i^t, VY_i^t	Instantaneous Velocity of Surrounding Vehicle *i*	$VX_i^t = \frac{x_i^t - x^t}{t}, VY_i^t = \frac{y_i^t - y^t}{t}$
AX_i^t, AY_i^t	Instantaneous Acceleration of Surrounding Vehicle *i*	$AX_i^t = \frac{VX_i^{t+1} - VX_i^t}{t}, AY_i^t = \frac{VY_i^{t+1} - VY_i^t}{t}$
R_{left}, R_{right}	Left/Right Lane Presence Flags of Target Vehicle	/

To enhance model training efficiency and stability, this study implemented a two-stage preprocessing protocol on the fused datasets. First, feature scaling was applied using min-max normalization to constrain all feature values within the range [−1, 1]. This standardization mitigates dimensional disparities among variables, ensuring balanced parameter updates during optimization. Subsequently, first-order differencing was employed to compute temporal gradients, thereby capturing dynamic trends in feature evolution. This step reduces temporal redundancy in sequential data while sharpening the model's sensitivity to trajectory variations. The mathematical formulation for these operations is:

$$X_{input} = diff\left(\frac{2x-(x_{\max}+x_{\min})}{x_{\max}-x_{\min}}\right) \tag{1}$$

2.2. GRU Decoder

The decoder input is derived from the encoder's output context feature vector. During decoding, two standard operational modes are commonly utilized: Teacher-forcing mode, which employs ground-truth values alongside the current hidden state as inputs for the subsequent timestep. While this approach accelerates convergence, it risks overfitting due to dependency on labeled data. Recursive mode, which iteratively uses model-predicted values and the current hidden state for future timesteps[10]. Though this reduces overfitting, cumulative prediction errors degrade long-term accuracy. To address these limitations, this study proposes a mixed teacher-forcing strategy that dynamically interleaves both modes. A teaching rate parameter($\alpha = 0.5$)governs the selection process: before each timestep prediction, a random value η is sampled uniformly from [0,1]. If $\eta > \alpha$, Recursive mode is activated; otherwise, teacher-forcing is applied. This stochastic switching mechanism enhances the model's exposure to real trajectory data while curbing error propagation and overfitting, as shown in Figure. 2.

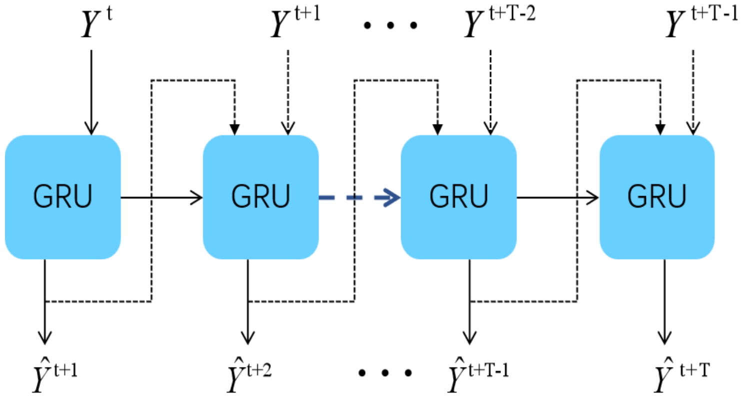

Figure 2. Mixed Teaching Force Mode

The trajectory data generated by the decoder module is subsequently transmitted to the data processing module, where it undergoes feature differential processing and feature scaling operations to ultimately produce trajectory predictions for the subsequent 5-second temporal window. Specifically, the feature differential operation captures temporal variations in motion dynamics through discrete differentiation of feature vectors between consecutive time steps, while the feature scaling mechanism employs parameterized transformation layers to dynamically adjust feature magnitudes. This dual-processing architecture enables adaptive feature representation optimization, thereby enhancing the model's capability to maintain prediction accuracy across diverse driving scenarios characterized by varying kinematic patterns and environmental conditions.

2.3. Driving Intention Recognition Model

The proposed model employs a dual-layer Long Short-Term Memory (LSTM) architecture, utilizing a 44-dimensional input vector that integrates both 3-second historical trajectory data and 5-second predictive trajectory data. Each LSTM layer maintains a hidden state dimension of 128, with the final temporal step output being directed to a fully-connected layer for driving intention classification. A Softmax activation function is implemented to perform probability normalization across three distinct driving intention categories. During the training phase, the model optimization is achieved through a weighted cross-entropy loss function, while the Adam optimizer is synergistically combined with a learning rate scheduler to enhance convergence efficiency and model robustness.

3. Data Preparation

The NGSIM datasets, developed by the Federal Highway Administration (FHWA), is an open-source datasets designed for vehicle trajectory prediction research. It captures the spatiotemporal trajectories of motorcycles, passenger vehicles, and heavy-duty vehicles on specific roadway segments. The datasets includes essential kinematic parameters such as longitudinal and lateral positions, velocity, acceleration (as temporal derivatives of velocity), and inter-vehicular spacing. For instance, the US-101 corridor datasets covers a 640-meter detection zone, with data collected during morning peak hours (07:50-08:35) at a 0.1-second sampling interval. Detailed descriptions of key data variables and their definitions are systematically presented in tabular form for research purposes, as shown in Table. 2.

Table 2. The raw data of NGSIM

Field	Description	Unit
Vehicle_ID	/	/
Local_X	Longitudinal Position	ft
Local_Y	Lateral Position	ft
v_Class	Vehicle Class	1-motorcycle; 2-small car; 3-large car
v_Vel	Vehicle Speed	ft/s
v_Acc	Vehicle Acceleration	ft/s^2
Lane_ID	/	/

An adaptive denoising method based on wavelet transform is proposed in this study to address track dithering resulting from sensor noise and environmental disturbances. Traditional linear filters are ineffective in handling non-stationary noise; however, wavelet transform can accurately distinguish noise from genuine motion patterns due to its time-frequency localization capabilities. Initially, the original trajectory coordinate sequence is standardized using Z-score normalization to eliminate dimensional variations. Subsequently, the velocity sequence undergoes decomposition through a 5-layer discrete wavelet transformation employing the sym11 wavelet basis. Finally, the denoised trajectory is reconstructed through inverse wavelet transformation. Visual inspection confirms a notable reduction in high-frequency noise energy, thereby enhancing the robustness of the input for the time series model, as shown in Figure. 3.

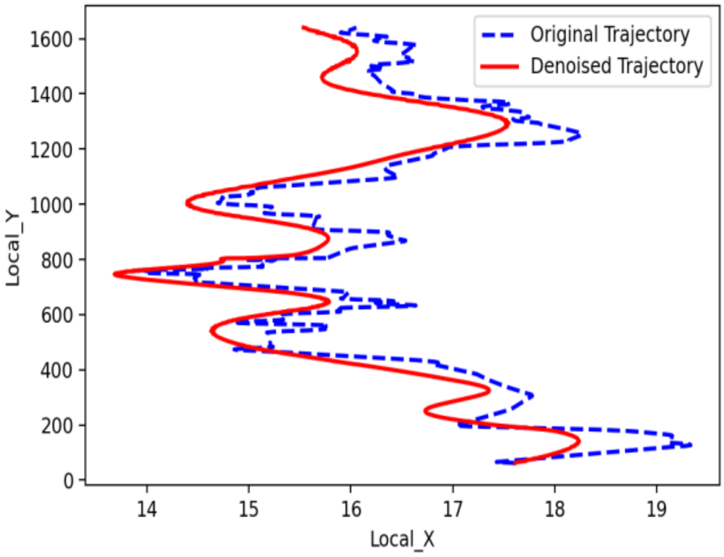

Figure 3. Trajectory Comparison

For time series analysis, 80 frames (8 seconds) of data were extracted using a sliding window algorithm. A 5-second future trajectory prediction was conducted based on 3 seconds of historical trajectory data. Consequently, the driving intention label for each track was assigned every 3 seconds. Upon labeling the driving intentions in the datasets, a notable discrepancy in the sample sizes across different driving intention categories (such as left lane change, lane keeping, and right lane change) was observed. This imbalanced distribution can compromise the model's generalization capability and prediction performance. To address this issue, the study employs an undersampling strategy to equalize the sample sizes of the left lane change and lane keeping categories with that of the right lane change category, which has the fewest samples. This approach aims to rectify category imbalances and mitigate model bias stemming from such disparities, as shown in Figure. 4 Additionally, the datasets is partitioned into training, testing, and validation sets following a 7:1:2 ratio.

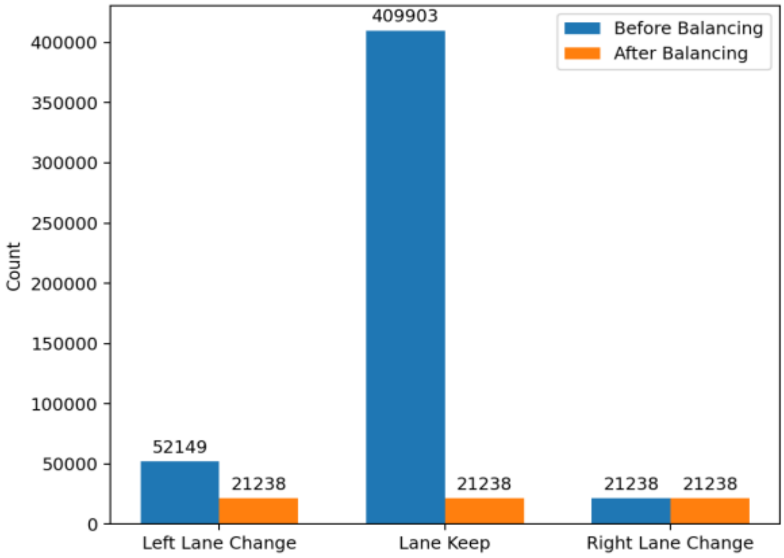

Figure 4. Number of Samples Before and After Balancing

4. Analysis

4.1. Analysis of single Feature and multiple feature input selection

The MTF-GRU network for trajectory prediction outputs continuous values, leading to the adoption of root mean square error (RMSE) as the evaluation metric.

$$RMSE = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2} \tag{2}$$

The experiment utilizes the PyTorch deep learning platform on a hardware setup comprising Windows 11 OS, an Intel Core i7-13650HX processor (3.20GHz), 16GB system memory, and NVIDIA GeForce RT4060 GPU for training. The network architecture includes an encoder-decoder with 4 GRU layers, hidden layer size of 128, and dropout of 0.2. Training parameters consist of 60 iterations, batch size of 1024, mean squared error loss function, Adam optimizer with a learning rate of 0.001, and weight decay of 0.0001.

Based on the original datasets, only the target vehicle information is retained and utilized as input for comparing the performance impact using two distinct data inputs, with terminal displacement error serving as the metric. Trajectory prediction plots and endpoint displacement error box plots are generated separately for comparative analysis, as shown in Table. 3 and Figure. 5.

Table 3. Comparison of different feature selection methods

Different methods	Displacement Errors/m				
	1s	2s	3s	4s	5s
Single feature	0.67	1.31	1.96	2.65	3.41
Multiple features	0.58	1.09	1.56	2.05	2.67

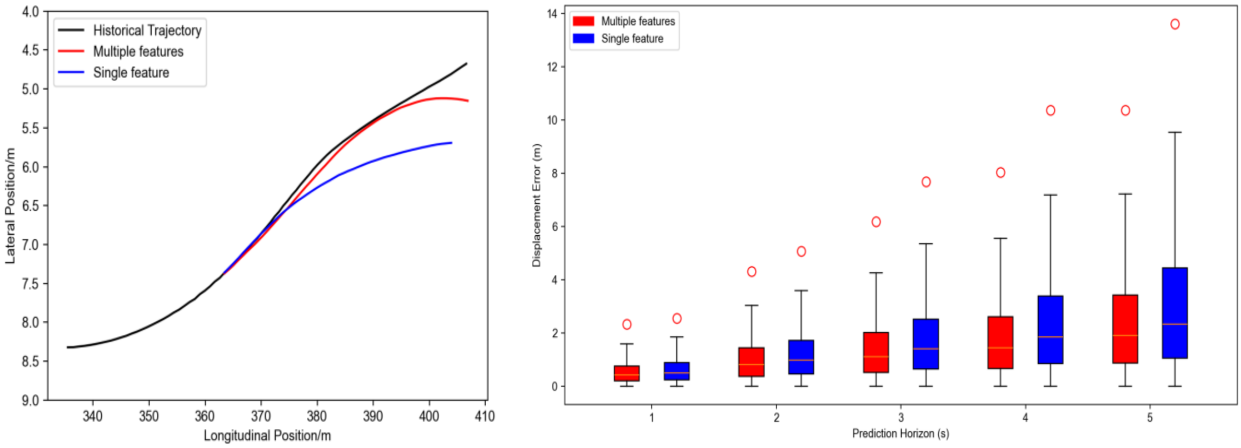

Figure 5. Comparison of different feature selection methods

The figures illustrate that the trajectory predicted with multi-feature input aligns more closely with the actual trajectory, resulting in a shorter box plot, reduced abnormal average values, and an overall 20.5% decrease in displacement error. Compared to models relying solely on the target vehicle's historical data, the MTF-GRU's integration of multiple features—including surrounding vehicle interactions—significantly improves trajectory prediction accuracy. This result demonstrates that incorporating interaction dynamics between vehicles enhances the model's ability to learn complex spatiotemporal patterns.

4.2. Performance comparison of different models

In order to further evaluate the performance of the proposed model, S-LSTM, S-GAN, M-LSTM were selected respectively to compare with the MTF-GRU model proposed in this study, and the evaluation indicators were obtained as shown in Table 4 and Figure. 6.

Table 4. Comparison result of different models

Different models	Displacement Errors/m				
	1s	2s	3s	4s	5s
S-LSTM	0.63	1.51	2.16	3.14	4.43
S-GAN	0.54	1.31	2.21	3.23	4.41
M-LSTM	0.56	1.23	2.14	3.25	4.67
ST-LSTM	0.52	1.16	1.89	2.71	3.61
MTF-GRU	0.51	0.97	1.36	1.75	2.31

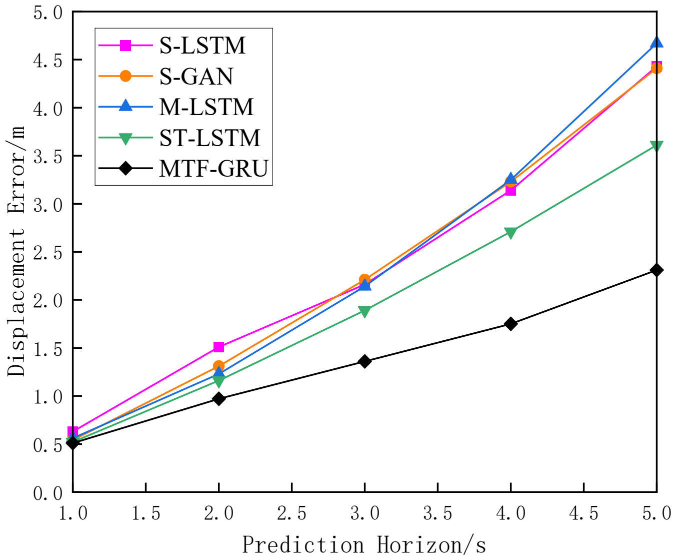

Figure 6. Result of different models

As presented in Table 4, the MTF-GRU model demonstrates minimal discrepancy from ground-truth values during short-term trajectory prediction compared to state-of-the-art approaches. Notably, it exhibits significantly higher accuracy in long-term predictions where other models show exponential error growth over time.

4.3. Performance comparison of different models

As shown in Figure 7, the driving intention recognition model achieves accuracies of 93.3% for left lane changes, 88.5% for lane keeping, and 93.0% for right lane changes, yielding an overall accuracy of 91.3%. Notably, left and right lane change predictions exhibit superior performance compared to lane-keeping tasks, which demonstrate the lowest recognition accuracy. Potential contributing factors include non-negligible deviations in the trajectory prediction data used for training, which may arise from sensor noise or model approximation errors.

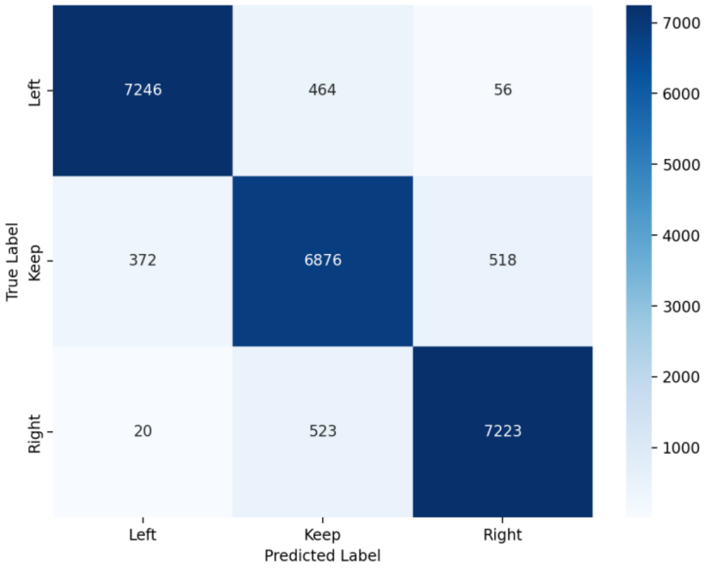

Figure 7. Confusion matrix of driving intention recognition model

5. Conclusion

This study introduces the Mixed Teaching Forcing Gated Recurrent Unit, a novel vehicle trajectory prediction model. This study findings hold significant guidance and reference value for the development of intelligent networked vehicle trajectory prediction and decision-making planning systems. This advancement provides a theoretical framework for optimizing multi-agent collaborative decision-making in intelligent transportation systems, particularly in resolving spatiotemporal conflicts under complex interaction conditions. While the model was trained on highway datasets, it lacks full consideration of the adaptability required for complex real-world road conditions including congestion and emergencies. To enhance its applicability, future research will incorporate multi-modal data (lidar and camera perception data) to enhance the model's adaptability to diverse traffic environments.

References

[1] Ji Y J, Zhang X Y, Yang Z R, et al. Trajectory Planning for Multi-Intelligent Connected Vehicles: State-of-the-Art and Prospects [J]. Chinese Journal of Mechanical Engineering, 2024, 60 (10): 129 - 146.

[2] Fang Huazhen, Liu Li, Gu Qing, et al. Automatic vehicle lane changing intention recognition research status [J]. Journal of engineering proceedings, 2024 46 - 48 (10): 18451855.

[3] Xie F, Lou J T, Zhao K, et al. Research on vehicle trajectory prediction method based on behavior recognition and curvature constraints [J]. Automotive Engineering, 2019, 41 (9): 1036 - 1042.

[4] Feng Ran, Zhang Liren, Wang Lihui. Uncertainty trajectory prediction based on second-order Markov chain [J]. Surveying, Mapping and Spatial Geographic Information, 2019, 43 (S1): 207 - 211.

[5] Jin L S, Gao M, Guo B C, et al. Driver perspective trajectory prediction based on spatiotemporal fusion LSTM network [J]. China Journal of Highway and Trans-port, 2022, 35 (4): 325 - 332.

[6] JI X W, FEI C, HE X K, et al. Intention recognition and trajectory prediction for vehicles using LSTM network [J]. China Journal of Highway and Transport, 2019, 32 (06): 34 - 42.

[7] MENG X W, TANG J, WANG Z, et al. Trajectory prediction of vehicles based on LSTM-AdaBoost model considering lane change intention [J]. Computer Engineering and Applications, 2022, 58 (13): 280 - 287.

[8] FANG Huazhen, Liu Li, Xiao Xiaofeng, et al. Vehicle Trajectory Prediction Based on Mixed Teaching Force Long Short-term Memory [J]. Journal of transportation systems engineering and information technology, 2023, 23 (4): 8087.

[9] Li Lin, Zhao Wanzhong, Wang Chunyan. Vehicle Driving Intention Analysis and Recognition Based on Bi-GLSTM Network [J]. Journal of Mechanical Engineering, 2019, 60 (10): 51 - 63.

[10] WANG Qingrong, Hao Fule, Zhu Changfeng, et al. Based on feature fusion of the vehicle trajectory prediction research [J/OL]. Computer engineering, 1 - 14 [2025 - 03 - 15].