Journal of Artificial Intelligence Practice (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/jaip.2023.060504
ISSN 2371-8412 Vol. 6 Num. 5

Vehicle Driving Intent Recognition Based on Enhanced Bidirectional Long Short-Term Memory Network

Dong He*, Maojie Zhao, Zinan Wang

School of Mechatronics and Vehicle Engineering, Chongqing Jiaotong University, Chongqing, 400074, China

**Corresponding author: HD15283947963@163.com*

Keywords: Autonomous Driving; Multi-Vehicle Interaction; Driving Intention Recognition; Bidirectional Long Short-Term Memory

Abstract: In the context of high-speed mixed traffic and intricate multi-vehicle interaction, existing driving intention recognition models for research vehicles inadequately address crucial factors, such as driving style and vehicle-vehicle interaction information. This paper introduces a novel driving intention recognition model based on an enhanced bidirectional long- and short-term memory network (Bi LSTM). The proposed model leverages the driving trajectory sequence of the target vehicle, driving style, and interaction features of surrounding vehicles as inputs for effective training and learning. It facilitates the classification and recognition of the driving intention feature dataset, specifically considering diverse driving styles. Additionally, the whale optimization algorithm is employed to optimize pivotal hyperparameters, encompassing the number of hidden layer nodes and learning rate, effectively mitigating the adverse impacts of manual parameter adjustment. The model's efficacy is validated using the NGSIM dataset, exhibiting an impressive recognition accuracy of 97.5% in precisely identifying vehicle driving intentions.

1. Introduction

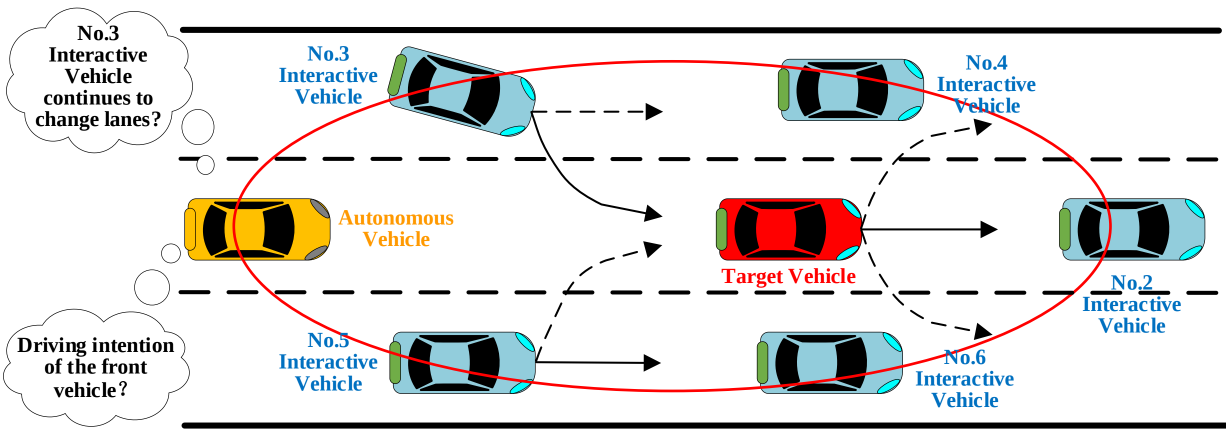


Figure 1: Analysis of Vehicle Interaction Behavior in High-speed Mixed Traffic Environment

Vehicle interaction in a high-speed mixed traffic environment is a complex behavior, and autonomous driving can share each other's driving intentions in real time and understand the current driving environment through vehicle networking technology. However, artificial driving cars cannot exchange information with the vehicle network, and the subjective driving intentions of artificial drivers (going straight, changing lanes left, changing lanes right, etc.) are unknown, and automatic driving cars can only infer their intentions through their external behavioral characteristics as shown in Figure 1. Driving intent recognition of surrounding vehicles can both predict the driving trajectory of surrounding vehicles and provide decision-making reference information for self-driving cars. Accurate driving intent recognition helps to improve the driving safety of vehicles, and how to accurately recognize the driving intent of surrounding vehicles is also one of the current hot spots in the research of automatic driving systems.

At present, the research methods of driving intention recognition at home and abroad can be divided into two categories: driving intention recognition methods based on traditional machine learning and driving intention recognition methods based on deep learning, both of which have their own advantages and disadvantages. Among the traditional machine learning methods based on traditional machine learning, support vector machine model (SVM) and markov model are more widely used [1]. Zhu Liling et al. [2] proposed an SVM-based driving intention classification model, which is able to classify three types of driving intentions: vehicle following, lane change preparation, and lane change execution, but the information may be lost when using principal component analysis to process high-dimensional driving data, and it is also difficult for SVMs to solve the problem of multi-classification. Liu et al. [3] proposed a new model combining markov model and SVM to obtain better recognition accuracy and real-time performance than a single model for lane changing, lane keeping, and overtaking on highways, and the improvement lies in the use of a two-layer algorithm to fully utilize the performance advantages of both classifiers. Among the deep learning-based methods, neural network model and long short-term memory network model (LSTM) are more widely used [4]. Huang et al. [5] applied deep neural network (DNN) to lane changing behavior recognition, which can effectively fit the complex features of lane changing behaviors, but the DNN network has a weak ability to capture the temporal features, so the model has received less attention in recent years. Phillips et al. [6] used LSTM to build an intention recognition model on the collected intersection traffic data to realize the prediction of left-turn, right-turn, and straight ahead intentions, and verified the ability of LSTM to capture the features of real collected time-sequence data. Huifei et al [7] used bidirectional long- and short-term memory network (Bi LSTM) to identify abnormal states in driving behavior, and their innovation lies in combining Bi LSTM with fully connected neural networks, giving full play to the advantages of both.

In terms of the research process and the corresponding indicators, deep learning models perform better than traditional machine learning models in general. The deep learning model represented by the long and short-term memory neural network model, with its strong fitting ability and ability to capture time series features, makes its performance in driving intention recognition overall stronger than that of the traditional machine learning model. However, some studies still have difficulties that need to be solved, such as insufficient consideration of surrounding vehicle interaction features, neglecting driving styles, and difficulties in manual parameterization.

2. Improved model

2.1. Bidirectional long and short-term memory neural network

Long short-term memory neural network is a class of variants of recurrent neural network (RNN), compared with RNN it controls the flow of information by introducing the "gate" structure and the concept of cells, thus overcoming the gradient explosion and gradient disappearance that plague RNN.

The most basic cell structure of the network is shown in Figure 2, which differs from RNN in that there are three special "gate" structures, Forget Gate, Input Gate and Output Gate, as well as a cell that is used for storing and transmitting information about the state of the previous moments. The basic unit realizes the flow of information mainly through these special structures [8].

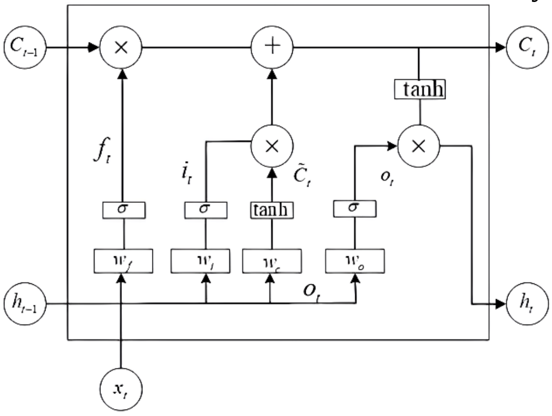


Figure 2: LSTM Basic Cell Structure

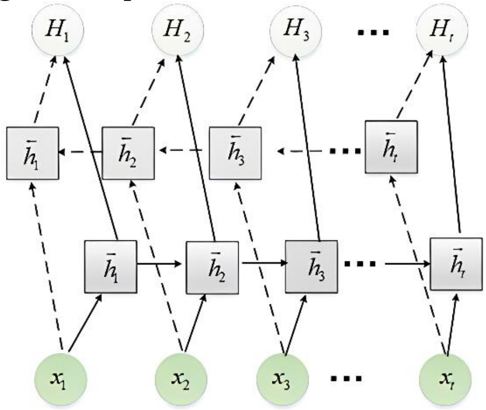


Figure 3: Bi LSTM Structure Flow Chart

Although LSTM solves the problems existing in RNN, it can only utilize the past feature information by itself and ignore the future information, so Bi-directional Long Short Term Memory Neural Network was born on this basis [9]. The model contains two independent LSTM networks, and the inputs of the model are fed into the two LSTM networks in the form of forward and reverse respectively, and then the feature vectors extracted from the two networks are spliced to obtain the final feature vector of the model; after the forward and reverse feature extraction, the final spliced vector possesses the information of the past and the future at the same time, and its structural flow is shown in Figure 3, with the $\vec{h}_1$ to $\vec{h}_t$ as forward operation process, $\overleftarrow{h}_t$ to $\overleftarrow{h}_1$ is the reverse operation process.

2.2. Driving Intent Recognition Model

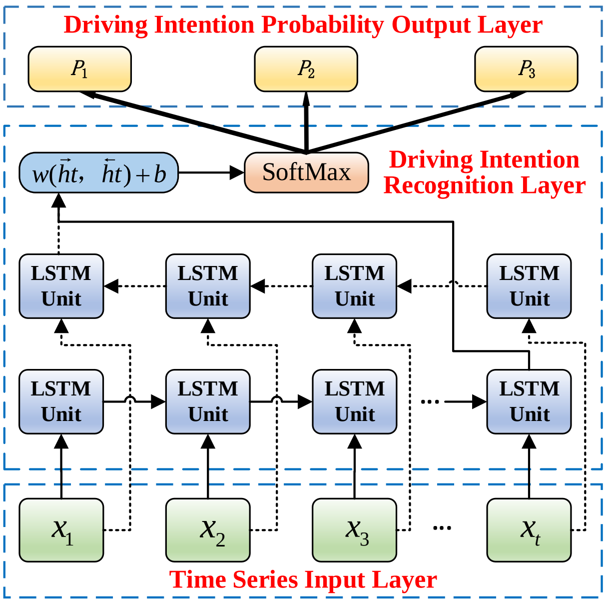


Figure 4: Bi LSTM Driving Intention Recognition Model

In this paper, the architecture of driving intention recognition model is built based on bi-directional long- and short-term memory network as shown in Figure 4.

The solid line in the above figure represents the process of forward LSTM information propagation and the dashed line represents the process of reverse LSTM information propagation. The time sequence $X_t = (x_1, x_2, x_3 ... x_t)$ is input to the forward and reverse LSTM networks respectively, and the input x_t at each moment *t* in the sequence obtains the forward output $\vec{h}_t$ and the reverse output $\overleftarrow{h}_t$ in this model, which is then spliced to obtain the vector $x_t' = [\vec{h}_t, \overleftarrow{h}_t]$. After obtaining x_t' it is input to the fully connected layer to obtain the output $y_t = wx_t' + b$, where *w*,*b* are the weights and bias of the fully connected layer. After obtaining y_t it is input to the SoftMax layer to obtain the probability of each of the three input categories left lane change, right lane change, and straight ahead through the activation function at moment *t*. The category with the highest probability is then used as the driving intention at the current moment using the classification layer.

2.3. Bi LSTM hyperparameter optimization process and results

The hyperparameters of Bi LSTM network have a large impact on the performance of the model, and the correct selection of the two important hyperparameter indicators, namely the number of nodes in the hidden layer and the learning rate, can improve both the training speed and the model accuracy. Therefore, in order to obtain the optimal solution of these two hyperparameters, this paper uses the Whale Optimization Algorithm (WOA) to find the optimization of the loss function of Bi LSTM, and when the iterative curve of the algorithm converges and the loss function is the lowest the corresponding hyperparameter is the optimal solution. The algorithm flow of WOA-Bi LSTM is shown in Figure 5.

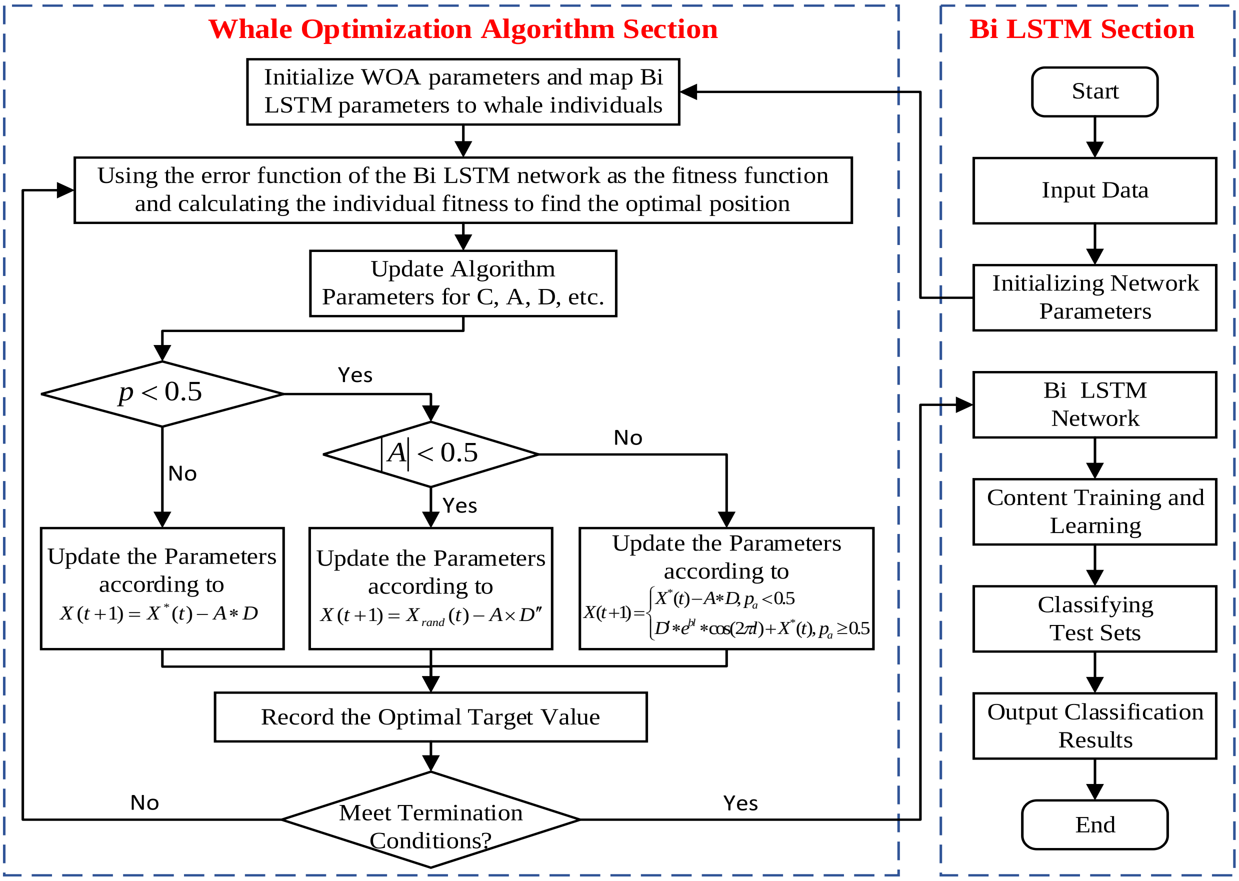


Figure 5: WOA-Bi LSTM Algorithm Flow

The whale optimization algorithm updates the optimization parameters by simulating the hunting behaviour of the whale community, which includes three parts: encircling the prey, attacking the prey and randomly searching for the prey [10]. The position of each whale represents a feasible solution, and for N parameters to be optimized, the position of the whale can be set as $X=(x_1,x_2,x_3...x_N)$. Optimization of the loss function of the Bi LSTM network was performed using the WOA model. The number of iterations of the optimization algorithm is set to 50, the optimizer uses Adam, the loss function is cross-entropy, the number of whale populations in the WOA-Bi LSTM model is set to 30, and the number of iterations is set to 20, and the input of the Bi LSTM network is the driving intention dataset constructed by the characteristic equation of the motion state of the target vehicle. According to the output results of the iterative convergence of the WOA model, the whale position vector corresponding to the fitness value at the convergence of the fitness curve is the optimal solution, the corresponding optimal number of hidden layer nodes is 82, and the optimal learning rate is 0.0016.

3. Simulation Analysis of Driving Intent Recognition Model

3.1. Data pre-processing based on improved sliding window algorithm

Multiple traffic participants usually exist in the interaction scenario, and the self-driving car will be affected by the interactions between the surrounding vehicles, and the selection of interaction features is crucial for the accurate recognition of the intent of the artificial driving vehicle. Considering the importance of lateral distance to vehicle driving safety [11], lateral distance is introduced into the driving intention recognition task as one of the interaction features, *l* is the lateral relative distance between the target vehicle and the self-driving car, *d* is the longitudinal relative distance, and the peri-vehicle interaction feature can be expressed as equation (1).

$$\begin{cases} x_i=(l_i,d_i), i=1,2,3...6 \\ f_{social}=(x_1,x_2,x_3,x_4,x_5,x_6) \end{cases} \tag{1}$$

The interaction features describe the interaction behaviour of the target vehicle with the surrounding vehicles including the self-vehicle, identifying the driving intention also requires selecting the motion state features of the target vehicle: $f_t=(v,a,x,y,v_x,a_x)$, where *v* is the velocity of the target vehicle; *a* is the acceleration of the target vehicle; *x*, *y* denote the longitudinal and transverse coordinates of the target vehicle; v_x is the transverse velocity; a_x is the transverse acceleration.

The driving intentions generated by different styles of drivers facing the same driving scenario are also different, so accurately recognizing the driving intentions of the target vehicle also requires the introduction of its driving style features, and the driving style feature vector is shown in (2).

$$v_{style}=\begin{cases} (1,0,0), & Radicalization \\ (0,1,0), & Ordinary \\ (0,0,1), & Conservative \end{cases} \tag{2}$$

In summary, the input of the driving intention recognition model consists of three parts: interaction features, driving style type, and target vehicle motion state as shown in equation (3).

$$i_v=(f_t,f_{social},v_{style}) \tag{3}$$

After completing the selection of feature parameters, it is necessary to screen the NGSIM data and extract the corresponding feature data, and at the same time, the classification model is a supervised learning model, which needs to label the driving intention data with vehicle behaviour [12]. In this paper, a sliding window algorithm combined with lateral displacement is proposed to extract and assign labels to the corresponding data, as shown in Figure 6.

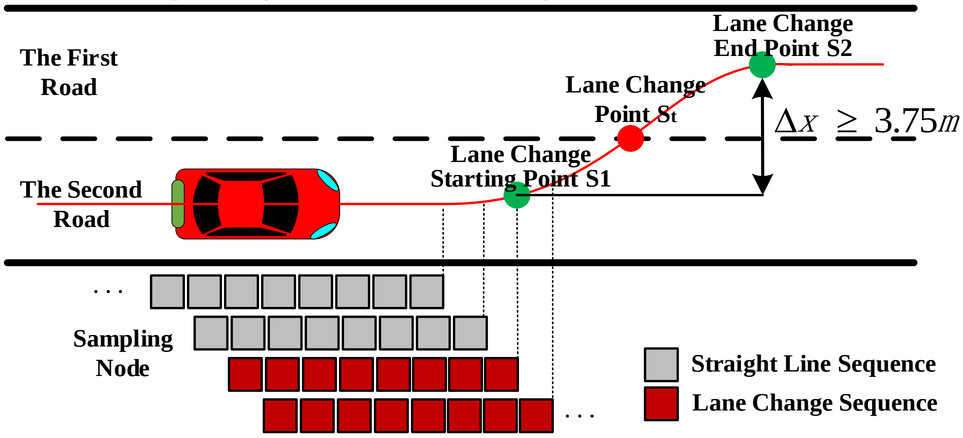


Figure 6: Improved Sliding Window Data Extraction Algorithm

As shown in the figure, the improved data extraction method is defined as follows: find the moment t corresponding to the vehicle lane-changing point S_t, then extract the sampling points corresponding to the time periods *t-4s* and *t+4s* in the sequence of vehicle trajectories and calculate the difference between the two lateral displacements. If the difference is greater than the lane width of 3.75m, the trajectory is labelled as a successful lane change. If the difference is less than 3.75m, take the moment where the lane changing point is located as the centre, take two sampling points S_1 and S_2 symmetrically at the two ends, and calculate the difference of the lateral displacement of the two; when the difference reaches 3.75m, the moment t_1 corresponding to S_1 is defined as the starting point of the lane changing, and the moment t_2 corresponding to S_2 is the ending point of the lane changing, and the sampling points falling into the interval $[t_1, t_2]$ are all the process points of the lane changing.

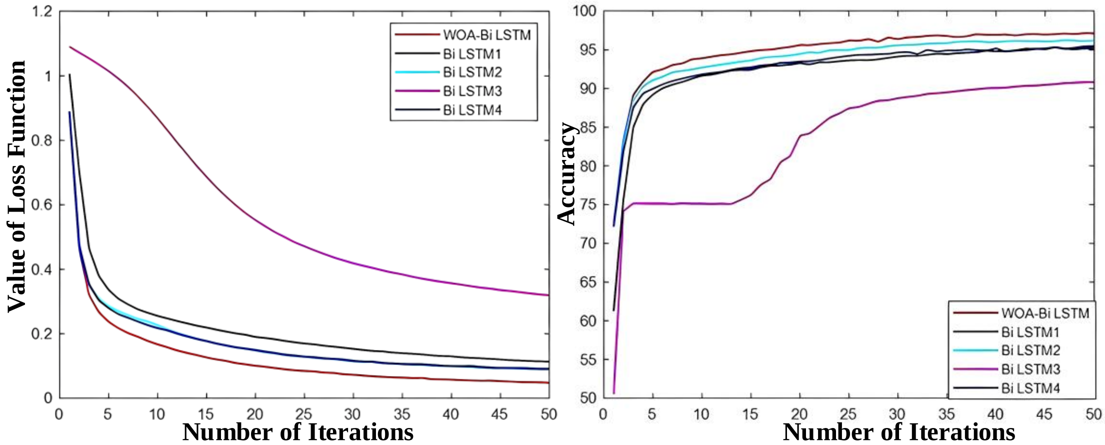
With the above method, the NGSIM data are calibrated for lane changing behaviour, and the data are processed according to the driving intention recognition model in the previous section, and the trajectory sequences are extracted as the input data of the model during the experiments, and a total of 23,800 sets of sample data are finally acquired, of which 4,977 sets of left lane-changing labels, 6,331 sets of right lane-changing labels and 12,492 sets of straight-line labels are obtained, and the data sets are divided into training sets, validation sets and test sets according to the 7:1.5 :1.5 ratio to divide the data set into training set, validation set and test set.

3.2. Simulation results analysis

Bi LSTM networks with different hyperparameters are chosen as experimental comparisons, the number of iterations of the networks are all set to 50, the loss function is the cross-entropy loss function, and numerical designators are added behind the names of the same networks for ease of differentiation, and the decreasing curves of the loss functions of each model and the increasing curves of the accuracy rates during training are shown in Figure 7.

From the above figure, WOA-Bi LSTM has the lowest loss function, the fastest convergence and the best performance among the models. The training elapsed time for each model on the training set and the accuracy of the last iteration are shown in Table 1.

In conclusion, choosing a reasonable number of hidden layer nodes and initial learning rate has a great impact on the performance of the model. The optimised Bi LSTM is used to recognise the test set data and the results are shown in Figure 8. From the figure, it can be seen that the recognition accuracy of the model on the test set is very high, reaching 97.5%, the recognition accuracy of the model on the right lane data is higher than that of the left lane, because the training data of the left lane is less than that of the right lane, and the model learns more about the understanding of the features of the right lane than that of the left lane during the training process, but both have a high recognition accuracy, and the recognition accuracy of the left lane is 95.5 per cent. The recognition accuracy of the left lane is 95.3% and the recognition accuracy of the right lane change is 97.3%. The recognition accuracy of 95.3% for the left lane change and 97.2% for the right lane change indicates that the model can recognise the left and right lane changes.



(a) Loss Function Decreasing Curve (b) Ascending Accuracy Curve

Figure 7: Curves of Each Model in Training Process

Table 1: Performance of different hyperparameter network training sets

Model	Node number	Learning rate	Accuracy rate
WOA-Bi LSTM	82	0.0016	96.6%
Bi LSTM1	41	0.0016	95.2%
Bi LSTM2	164	0.0016	96.1%
Bi LSTM3	82	0.0001	90.7%
Bi LSTM4	82	0.01	95.1%

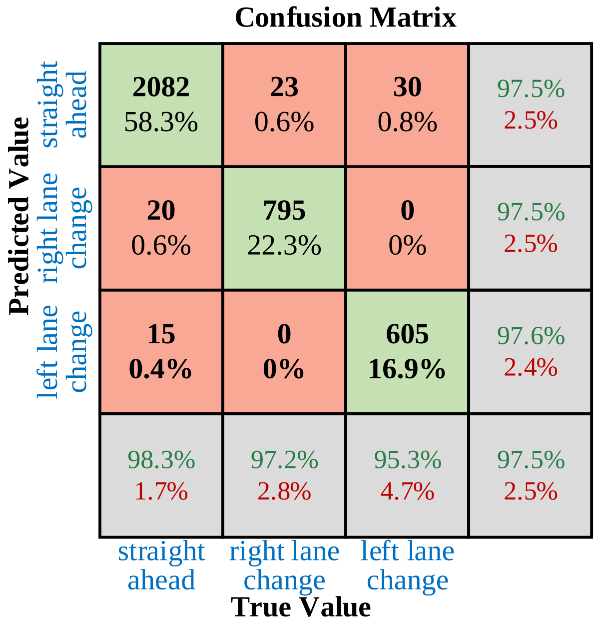


Figure 8: Test Set Confusion Matrix

4. Conclusions

This study addresses the critical task of driving intention recognition for vehicles navigating high-speed mixed traffic environments. A novel methodology is proposed, which integrates lateral displacement difference and sliding window techniques to enhance the processing of the NGSIM dataset, thereby facilitating the extraction and construction of a comprehensive driving intention feature dataset. Building upon this foundation, an improved driving intention recognition model is presented, leveraging the power of Bi LSTM. The model's performance is further optimized through meticulous hyperparameter tuning, specifically adjusting the learning rate and the number of hidden layer nodes, utilizing the whale optimization algorithm. The experimental results unequivocally demonstrate that the optimized hyperparameters acquired via the whale optimization algorithm significantly enhance the recognition accuracy of the model. On the NGSIM dataset's test set, the model achieves remarkable overall recognition accuracy of 97.5%, with precise accuracies of 95.3% for left lane changes, 97.2% for right lane changes, and 98.3% for straight-ahead driving. Notably, the recognition process is accomplished within an impressively low time span of 1.35 seconds. These findings unequivocally highlight the model's real-time capability to accurately discern driving intentions, thereby yielding valuable insights for predicting the driving trajectories of surrounding vehicles.

References

[1] Huang B, Kang W, Gu H, et al. Ultra Short Term Load Forecasting Based on Optimized Weight Cubature Kalman Filter and Support Vector Machine Combination Model[J]. Journal of Artificial Intelligence Practice, 2021, 4(1):37-45.

[2] Zhu L, Liu L, et al. Driver Behavior Recognition Based on Support Vector Machine [J]. Journal of Transportation Systems Engineering and Information Technology, 2017, 17(01):91-97.

[3] Liu Z, Wu X, et al. Driving Intention Recognition Based on HMM and SVM Cascade Algorithm[J]. Automotive Engineering, 2018, 40(07):858-864.

[4] Qin Y. BLSTM Recurrent Neural Network for Object Recognition [J]. Journal of Artificial Intelligence Practice, 2016, 1(1):25-29.

[5] Huang L, Guo H, Zhang R, et al. Capturing Drivers' Lane Changing Behaviors on Operational Level by Data Driven Methods[J]. IEEE Access, 2018, 6:57497-57506.

[6] Phillips D J, Wheeler T A, Kochender M J. Generalizable Intention Prediction of Human Drivers at Intersections [C]//2017 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2017:1665-1670.

[7] HUI F, GUO J, JIA S, et al. Detection of Abnormal Driving Behavior Based on BiLSTM. Computer Engineering and Applications, 2020, 56(24): 116-122.

[8] Kaadoud, Ikram C, et al. Knowledge Extraction from the Learning of Sequences in a Long Short Term Memory (LSTM) Architecture [J]. 2022, 235(Jan. 10):107657. 1-107657. 18.

[9] Chetan A, Anjana P, et al. Fake News Detection System Based on Modified Bi-directional Long Short Term Memory [J]. Multimedia tools and applications, 2022, 81(17):24199-24223.

[10] Mirjalili, Seyedali, Lewis, Andrew. The Whale Optimization Algorithm [J]. Advances in Engineering Software, 2016, 9551-67.

[11] Cheng G, Wang W, et al. Research Progress on Risk Assessment of Vehicle Lane-changing Behavior [J]. Journal of Harbin Institute Technology, 2023, 55(03):139-150.

[12] Deo N, Trivedi M. Multi-modal trajectory prediction of surrounding vehicles with maneuver based lstms [C]// 2018 IEEE intelligent vehicles symposium (IV). IEEE, 2018:1179-1184.