

Received May 5, 2022, accepted May 25, 2022, date of publication May 30, 2022, date of current version June 6, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3179007

Driver Lane Change Intention Recognition Based on Attention Enhanced Residual-MBi-LSTM Network

ZHANQIAN WU¹, KAICHONG LIANG¹, DENGCHENG LIU²,
AND ZHIGUO ZHAO¹, (Member, IEEE)

¹School of Automobile, Tongji University, Shanghai 201804, China

²Nanchang Automotive Institute of Intelligence and New Energy, Nanchang, Jiangxi 330052, China

Corresponding author: Zhiguo Zhao (zhiguozhao@tongji.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 52172390, and in part by the Perspective Study Funding of Nanchang Automotive Institute of Intelligence and New Energy under Grant TPD-TC202110-07.

ABSTRACT Accurate identification of lane-changing intention can effectively assist intelligent driving vehicles in terms of decision-making and trajectory planning, which plays a significant role in enhancing driving safety by reducing traffic accidents caused by lane-changing. Based on the trajectory characteristics and vehicle interaction information, an attention-enhanced bidirectional multi-layer residual long-short term memory neural network (Attention Enhanced Residual-MBi-LSTM) model is proposed for lane change intention recognition in this paper. Firstly, an EWMA filter is employed to smooth the noisy data collected from the vehicle. Then a four-layer bidirectional residual LSTM memory (Residual-MBi-LSTM) structure is used to extract lane-changing features from the historical driving trajectories of ego-vehicle and vehicle interaction information. Besides, the attention mechanism is added to adjust the weight of data in different time frames. After that, the current lane-changing probability is calculated and output by the Softmax function. Finally, the vehicle lane-changing intention recognition model is firstly trained and then verified in the HighD dataset. According to the HIL experiment, the proposed model has the ability to identify driver intention on average of 2.07 seconds in advance.

INDEX TERMS Lane change, advanced driver assistance system (ADAS), driving intention recognition, long short-term memory (LSTM), attention mechanism.

I. INTRODUCTION

According to statistics, human drivers' operational errors are one of the principal origins of most traffic accidents. Among all the traffic accidents, 18% occur in the lane changing process [1]–[3]. With the rapid development of artificial intelligence, advanced driving assisted systems (ADAS) have been widely developed, hence significantly reducing the probability of traffic accidents. Driving intention recognition has been reckoned as one of the core technologies of next-generation ADAS products. By recognizing driving intention, ADAS can give early warning to vehicle collisions and take collision avoidance operations to avoid accidents [4].

There is a clear difference between lane change behavior perception and intention recognition: lane-changing behavior

The associate editor coordinating the review of this manuscript and approving it for publication was Jie Gao¹.

perception refers to perceiving the process by sensors when the vehicle has started to change lanes [5], while lane-changing intention recognition refers to the process of identifying or predicting possible upcoming behaviors such as lane-changing, cutting, overtaking or turning before starting. But both of them are essentially the process of classifying multidimensional time series (or vectors) constructed from vehicle motion information, driver gaze (or head-turning) information, or traffic context information [6].

Many scholars have carried out research on the recognition of driver lane change intention. Focused on the relationship between ego-vehicle and directly surrounding vehicles, Schlechtriemen *et al.* [7] used a generative model based on an improved naive Bayesian approach, which was trained and validated on a test drive in the real world, to classify the lane-changing intention. Based on Support Vector Machine (SVM), Kumar *et al.* [8] proposed a recognition model, which

can predict lane change intention on average 1.3s in advance for online recognition. However, these above methods only classify according to the current state parameters and cannot use the context information of the continuous dynamic driving process to mine deep driving intentions.

The Hidden Markov method (HMM) [9] overcomes the shortcomings of the above simple classifiers and is able to model time series of arbitrary length, but its accuracy is relatively lower than current mainstream methods. Furthermore, the above machine learning methods need a priori model, which may cause bias due to inaccurate prior assumptions [10]. In conclusion, for the lane-changing intention recognition problem, statistical-based machine learning methods have certain defects and limitations.

The rapid development of high-performance computing and big data technology in recent years has made data-driven lane change intention recognition possible. However, earlier studies paid less attention to the correlation between time series information [11].

With the success of LSTM in speech recognition and machine translation, realizing the advantages of LSTM in solving time series problems, some scholars have introduced LSTM into the field of lane changing intention recognition. Jian *et al.* [12] proposed a driving intention detection LSTM based approach with better accuracy in driver intention recognition. Xie *et al.* [13] proposed an LSTM based lane change implementation (LCI) model, however, the proposed LCI model cannot recognize the lane changing intention until the vehicle is about to cross the lane marks. In order to improve the recognition accuracy in an early stage, classical single-layer-LSTM was extended to multi-layer-LSTM by Tang *et al.* [14]. However, these methods do not take the weight assignment of data in different time frames into account. Due to the weighting scheme of LSTM [15], lane changing intention can only be detected after significant lateral movement. In order to improve the timeliness, the attention mechanism was introduced to reallocate the weight. Guo *et al.* [17] proposed an AT-LSTM Based intention recognition model. In addition to vehicle motion parameters, Guo also introduces human features such as pupil gaze point and head pose by using the eye tracker, which is not suitable for a real driving environment. Similar to AT-LSTM, Hao *et al.* [18] also proposed Attention-based GRU for intention recognition. However, the use of single-layer-GRU inhibits the improvement in recognition accuracy.

Based on the trajectory characteristics and vehicle interaction information, this paper proposes an attention-enhanced residual MBi-LSTM model for driver lane change intention recognition.

Our contributions in this study are twofold:

(1) A novel attention enhanced Residual-MBi-LSTM Network is proposed for driver lane change intention recognition. In this paper, A 4-layer sequential bidirectional LSTM neural network with short cut structure is proposed to extract lane change intention from the history trajectory of ego-vehicle and the interaction between the ego-vehicle and

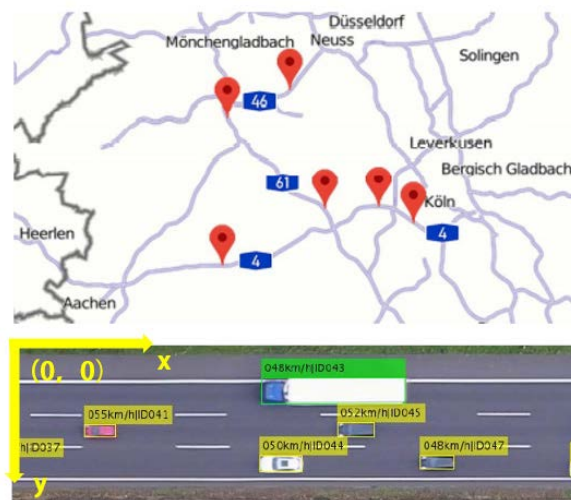


FIGURE 1. Schematic diagram of the research area of the HighD dataset [16].

the preceding vehicle, while a soft attention mechanism is introduced to adjust the weight of different data frames.

(2) The introduction of residual block enables our model to become deeper, solving the optimization bottleneck problem.

The structure of this paper is organized as follows: in part II, the data preprocessing method are introduced; in part III, the proposed model structure, parameter determination method, and model training process is introduced; in the last part, the performance of the model is tested through ablation experiment together with other experiments, a hardware-in-loop simulation is also conducted to test the model performance in the real working condition; finally, a conclusion is given in part VII.

II. DATASET CONSTRUCTION

A. HIGHD INTRODUCTION

The HighD dataset [16] is used to extract the trajectory information of the ego-vehicle and surrounding vehicles in this paper. The HighD dataset is an unmanned aerial vehicle (UAV) sampled large-scale vehicle natural trajectory database.

The dataset includes trajectory data of 110,000 vehicles at 6 locations, as is shown in Figure 1, around Cologne, Germany, for 16.5 hours, with a total mileage of 45,000 km, including 5,600 complete lane change records.

The dataset includes data information such as time t , vehicle ID, vehicle ordinates, lateral and longitudinal speed etc., which can fully meet the requirements of this paper. It also provides complete surrounding vehicle information and vehicle interaction information, for example, surrounding vehicle IDs, TTC, THW, etc.

B. DATA PREPROCESSING

As mentioned above, the dataset is collected with a UAV located directly above the road segment. Due to

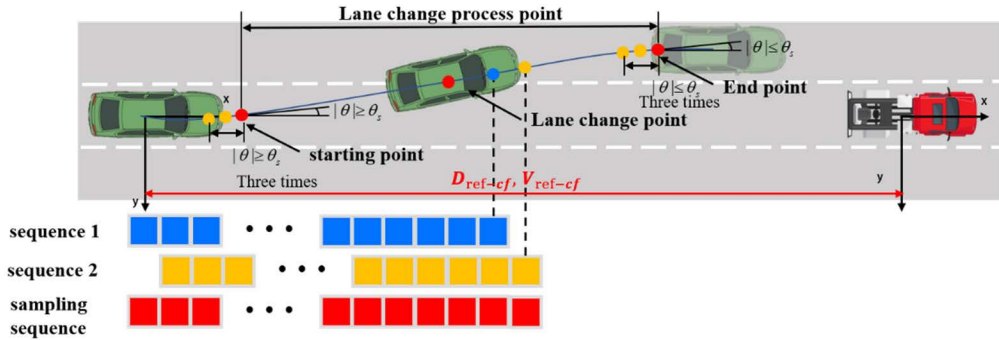


FIGURE 2. Schematic diagram of sliding time window sampling method.

the rotation and movement of the UAV, the data set has some measurement noise. For the purpose of improving the data quality, a symmetrical exponentially weighted moving average (EWMA) filter is adopted to smooth the noisy raw data.

The method is shown as follows [14]:

$$\bar{x}_\alpha(t_i) = \frac{\sum_{k=i-D}^{I+D} x_\alpha(t_k) e^{-\frac{|i-k|}{\Delta}}}{\sum_{k=i-D}^{I+D} e^{-\frac{|i-k|}{\Delta}}} \quad (1)$$

where $\bar{x}_\alpha(t_i)$ denotes the filtered value at time t_i , D denotes sliding time window length and Δ denotes the average of sliding time window.

The weighted moving average (WMA) refers to the method of assigning larger weight to recent data and less weight to past data. In this paper, the recent observations have a greater impact on the predicted value (PV), hence being able to reflect the trend better. The EWMA method means that the weighting coefficient of each value decreases exponentially with time.

After smoothing, a typical vehicle (id = 1-13) trajectory is presented in Figure 3(a), while the smoothing effect of vehicle speed (id = 1-13) and angular velocity (id = 1-13) is presented in Figure 3(b), (c). As can be noticed from Figure 3 (a) and (c), both the Savitzky-Golay method and the EWMA method can better reflect the natural trajectory. Therefore, considering the filtering effect and hardware resource consumption, the EWMA method is adopted to smooth the vehicle trajectory, velocity, and angular velocity.

For the intention recognition module, the extracted trajectory segments need to be divided into three categories: left lane change, right lane change, and non-lane change, and then corresponding labels are attached respectively [19].

As shown in Figure 2, the intersection of the vehicle trajectory and the lane line is defined as the lane change point.

Once the lane change point has been determined, the start-point and end-point of the lane changing process can be determined by calculating the heading angle θ of the vehicle.

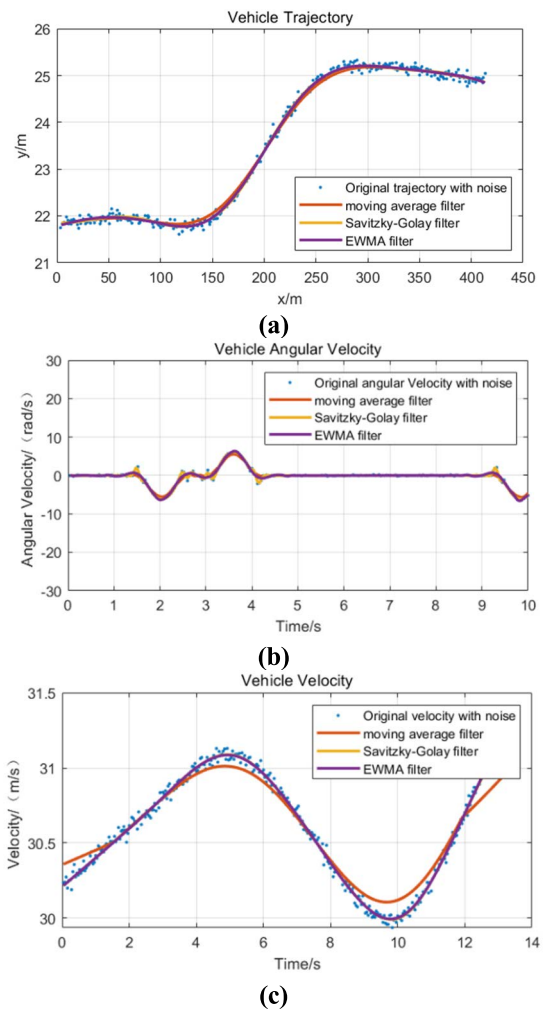


FIGURE 3. Filtering effect of (a) trajectory (b) angular velocity (c) velocity.

Among that, the heading angle θ can be defined as follows:

$$\theta(t) = \arctan\left(\frac{x^{(t)} - x^{(t-3)}}{y^{(t)} - y^{(t-3)}}\right) \quad (2)$$

where (x, y) are the coordinates of the vehicle.

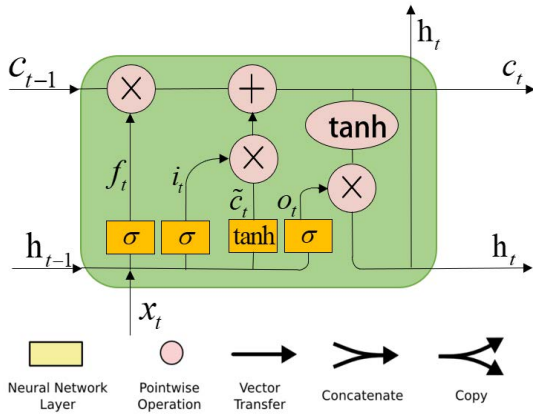


FIGURE 4. LSTM cell structure.

Traverse all the heading angle θ from the lane changing point along both sides of the timeline. If $|\theta| \leq \theta_e$ (where θ_e is the threshold of the end heading angle for lane change) three times continuous, the point that reaches the threshold for the first time is defined as the end-point. If $|\theta| \geq \theta_s$ (where θ_s is the threshold of the initial heading angle for lane change) three times continuous, the point that reaches the threshold for the first time is defined as the start-point.

To make full use of the data, the sliding time window segmentation method is employed to process the vehicle trajectory segments with equal length [20], the time window length is T_p , the sampling frequency is 25Hz, and the number of sampling points is $n = 25 \cdot T_p$.

Since the total number of lane change trajectories is far less than that of non-lane change trajectories, random down-sampling is used to reduce the number of non-lane change trajectories. Finally, the number of left lane change, non-lane change, and right lane change data segments are 32828, 26419 and 10614, respectively. The three types of data and corresponding labels are randomly sorted as an entire data set, which is then divided into training set, validation set and test set according to the proportion of 70%, 15% and 15%.

III. VEHICLE LANE CHANGE INTENTION RECOGNITION MODEL

A. RESIDUAL MULTI-BILSTM NETWORK STRUCTURE

1) LSTM NETWORK STRUCTURE

LSTM is a special kind of recurrent neural network (RNN). LSTM overcomes the problems of gradient vanishing and exploding gradients in long-time sequence training [21], hence having better performance in long-time sequence training.

The structure of a single LSTM unit is shown in Figure 4.

The calculation process of an LSTM unit is as follows:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [x_t, h_{t-1}] + b_f) \\ i_t &= \sigma(W_i \cdot [x_t, h_{t-1}] + b_i) \\ o_t &= \sigma(W_o \cdot [x_t, h_{t-1}] + b_o) \end{aligned}$$

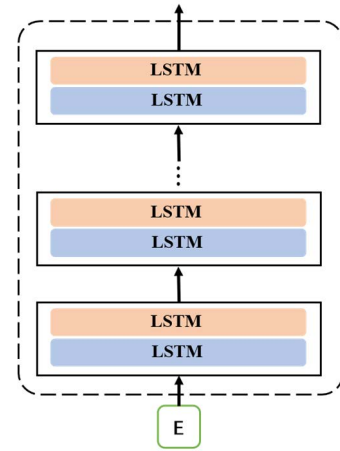


FIGURE 5. Multi-layer LSTM structure.

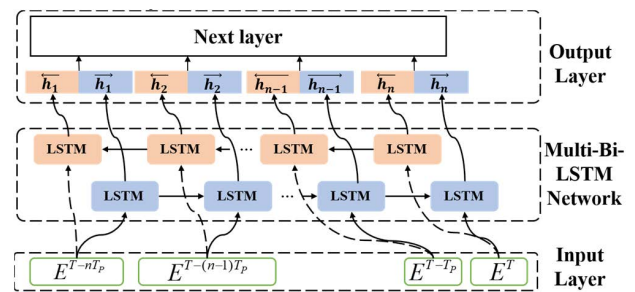


FIGURE 6. Multi-Bi-LSTM structure.

$$\begin{aligned} \tilde{c}_t &= \tanh(W_c \cdot [x_t, h_{t-1}] + b_c) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \quad (3)$$

where \odot denotes the dot product, $\sigma(\cdot)$ denotes the sigmoid function.

Among them, f_t is the forget gate, which activates the current input state x_t and the incoming hidden state h_{t-1} of the previous layer unit through the sigmoid function; i_t is the input gate, which activates the current input state x_t and the previous layer unit through the sigmoid function and passes in the hidden state h_{t-1} , but its weight W_c is different from the forget gate weight W_f . The input gate i_t selects the current layer information \tilde{c}_t as a new memory to add to the current memory state c_t , o_t denotes the output gate, outputting the hidden state h_t of the current layer.

2) MULTI-BILSTM NETWORK STRUCTURE

As is shown in Figure 5, the Multi-LSTM refers to a network structure formed by stacking layers of LSTM. Compared with the traditional single layer LSTM, Multi-LSTM structure can significantly increase the nonlinear components and hence improve generalization ability and robustness [14].

The BiLSTM network consists of two sub-layers: the forward LSTM (FW LSTM) layer and the backward LSTM (BW LSTM) layer. Unlike the traditional LSTM

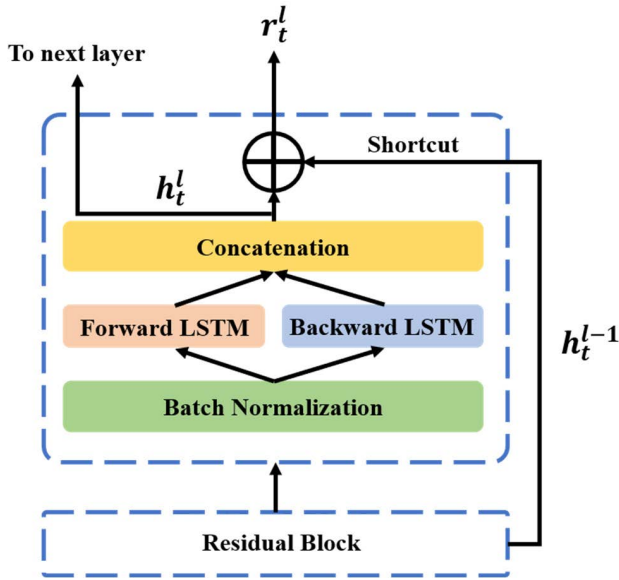


FIGURE 7. Schematic diagram of Residual block.

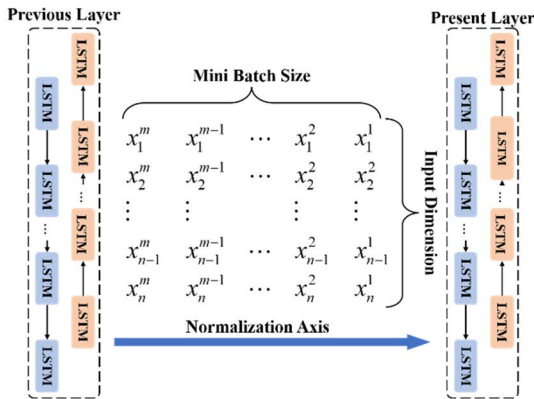


FIGURE 8. Schematic diagram of BN algorithm.

neural network backward propagation method, the BiLSTM network has bidirectional characteristics so that it can effectively use the past features through BW LSTM and future features through FW LSTM. BiLSTM can help solving long-term dependency problems and improve prediction accuracy by establishing two-way connections in neural networks [22]–[24]. The BiLSTM structure is shown in Figure 6. The final output of the forward and reverse LSTM units is concatenated as follows:

$$(h_t^{fw}, c_t^{fw}) = LSTM(e_t, h_{t-1}, c_{t-1}) \quad (4)$$

$$(h_{t-t_h}^{bw}, c_{t-t_h}^{bw}) = LSTM(e_{t-t_h}, h_{t-t_h+1}, c_{t-t_h+1}) \quad (5)$$

$$h_t = concat(h_t^{fw}, h_t^{bw}) \quad (6)$$

3) RESIDUAL BLOCK

Due to the deepening and increasing complexity of the network, the learning efficiency shall be reduced, hence making the accuracy not being effectively improved [25].

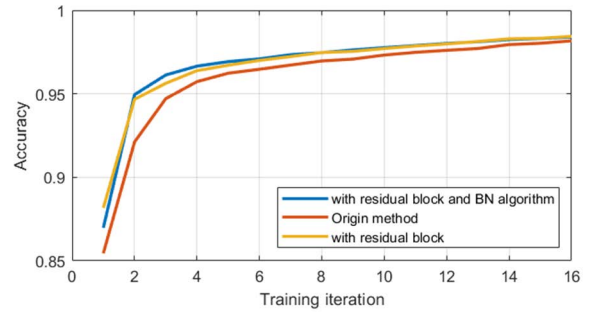
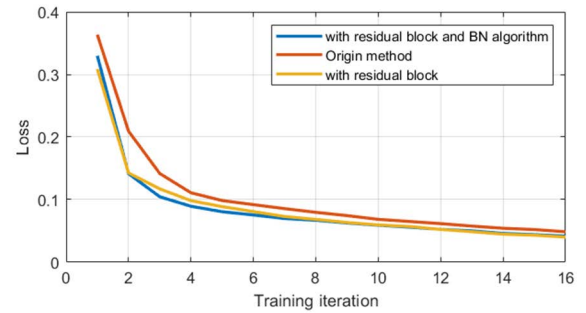


FIGURE 9. Training process with BN algorithm and residual block (a) loss (b) accuracy.

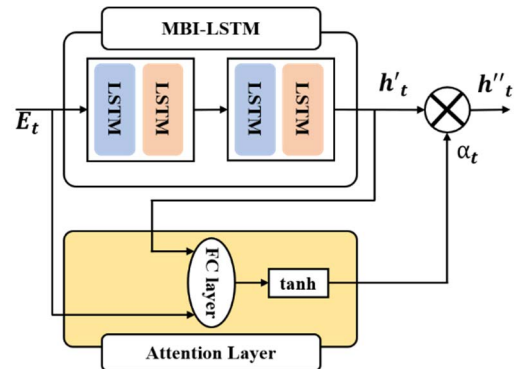


FIGURE 10. Schematic diagram of attention mechanism.

Inspired by Resnet [26], to solve the optimization bottleneck problem, this paper adds a Residual Block to the MBi-LSTM structure. As is shown in the Figure 7, The final output of the l -th residual block can be written as follows:

$$r_t^l = h_t^{l-1} + h_t^l \quad (7)$$

where h_t^l represents the hidden state of the l -th BiLSTM layer.

B. BATCH NORMALIZATION ALGORITHM

Batch Normalization (BN) [27] is an algorithm that speeds up the training process by fixing the distribution of layer inputs and reducing internal covariate shifts. The BN layer has trainable weights that normalize the layer output by setting the mean to 0 and the variance to 1. The standardization process is shown in Figure 8. The calculation process can be

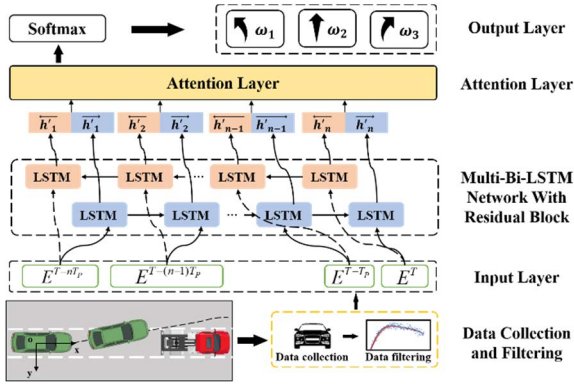


FIGURE 11. Schematic diagram of driver lane change Intention recognition based on attention enhanced residual-MBi-LSTM Network.

TABLE 1. Main parameters of vehicle lane change process.

Symbol	Definition
x	vehicle lateral coordinate
y	vehicle longitudinal coordinate
v_x	lateral velocity
v_y	longitudinal velocity
a_x	lateral acceleration
a_y	longitudinal acceleration
θ	vehicle heading angle
ω	heading angular velocity

constructed as follow:

$$\hat{x}_i^k = \frac{x_i^k - E(x_i)}{\sqrt{\delta(x_i)}} \quad (8)$$

where \hat{x}_i^k denotes the i -th row (dimension of the input vector) the k -th column (order in the mini-batch) of normalized input value; x_i^k denotes the i -th row and the k -th column of the original value, $E(x_i)$ denotes the batch data mean, $\delta(x_i)$ denotes the batch data variance.

After normalization, x_i is restricted to a normal distribution, which reduces the performance of the network. Therefore, a scale transformation factor γ and an offset transformation factor β are introduced as follows:

$$y_i^k = \gamma_i^k \cdot \hat{x}_i^k + \beta_i^k \quad (9)$$

The training process after introducing the residual block and BN algorithm is shown in Figure 9. Figure 9 (a) is the loss of the training process, while Figure 9 (b) is the accuracy of the training process. As shown in Figure 9, after introducing the residual block and adopting the BN algorithm, the loss decreases more rapidly during the first several training iterations, hence making the model converges faster.

C. ATTENTION MECHANISM

For a long time series, the importance of different data in different time frames are usually unequal. In order to allocate

the weight of data in different time frames, a soft attention mechanism [28] is introduced.

The structure of the soft attention mechanism is presented in Figure 10.

As shown in the figure, h' is the output of the MBi-LSMT network with residual block, where $h' = \{h'_1, h'_2, \dots, h'_{n-1}, h'_n\}$, s_t denotes the importance of the t -th time frame. The calculation process can be constructed as follows:

$$s_t = U_s \tanh(W_{xs}E_t + W_{os}h'_t + b_s) + b_{us} \quad (10)$$

where U_s , W_{xs} , W_{os} are the learned weight matrices; b_s , b_{us} are the bias vectors; E_t denotes the spatiotemporal feature input; while h'_t denotes the spatiotemporal information extracted by the MBi-LSMT network with residual block. The importance value of k -th frame is calculated as:

$$\alpha_t = \frac{e^{s_t}}{\sum_{i=1}^k e^{s_i}} \quad (11)$$

The importance value α_t is the attention weight, and the larger α_t is, the more important the frame data is.

The final output can be calculated as follow:

$$h''_t = \alpha_t \odot h'_t \quad (12)$$

The weights of the final output will be assigned by data frame importance, not equal weight output.

D. DRIVER LANE CHANGE INTENTION RECOGNITION BASED ON ATTENTION ENHANCED RESIDUAL-MBI-LSTM NETWORK

This paper proposes an attention-enhanced bidirectional multi-layer residual long short-term memory neural network lane-changing intention recognition model, whose structure is presented in the Figure 11. The model consists of data preprocessing layer, input layer, residual-MBi-LSTM layer, soft-attention layer, Softmax layer, and output layer.

Table 1 shows the main parameters [29], [30]. while considering the motion parameters of ego-vehicle, surrounding vehicles, especially the vehicle in front of the ego-vehicle, have a significant impact on the driving intention and driving trajectory of the ego-vehicle [31], [32]. Therefore, the relative distance D_{ref_cf} and relative speed V_{ref_cf} of the ego vehicle and the preceding vehicle are considered.

$$D_{ref_cf} = D_{ref} - D_{cf} \quad (13)$$

$$V_{ref_cf} = V_{ref} - V_{cf} \quad (14)$$

where D_{ref} , D_{cf} are the absolute coordinates of the ego-vehicle and the preceding vehicle, respectively; V_{ref} , V_{cf} are the speeds of the ego-vehicle and the preceding car, respectively.

The model input parameters are constructed as follows:

$$E = [e^{t-th}, \dots, e^{t-1}, e^t] \quad (15)$$

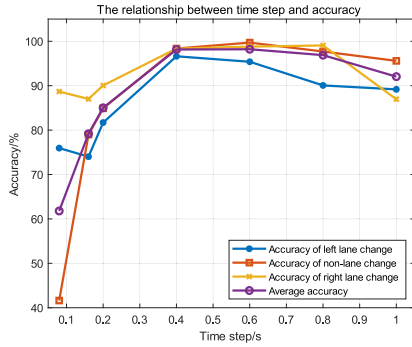


FIGURE 12. The relationship between model recognition accuracy and time step.

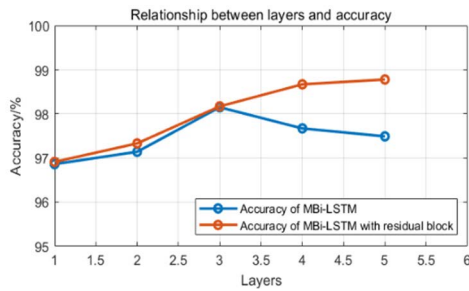


FIGURE 13. The relationship between model recognition accuracy and the number of Bi-LSTM layers.

where e^t is the characteristic parameter of vehicle lane change, and t^h is the length of the selected time series.

Where:

$$e^t = (x^t, y^t, v_x^t, v_y^t, a_x^t, a_y^t, \theta^t, \omega^t, D_{ref_cf}^t, V_{ref_cf}^t) \quad (16)$$

If there is no preceding vehicle, the corresponding virtual vehicle shall be populated, with its relative speed set to 0 and the relative distance set to $420-x$ (420 is the distance of a sector) [16].

$c = (c_1, c_2, c_3)$ is the lane-changing intention vector, c_1, c_2, c_3 represents the three categories of driving intention: left lane change, non-lane change and right lane change; $\Omega = \{\omega_1, \omega_2, \omega_3\}$ is the probability vector of the three driving intentions. $\omega_1, \omega_2, \omega_3$ is the probability of the three driving intentions. The model output is:

$$\omega_i = P(c_i/E), i = (1, 2, 3) \quad (17)$$

The data preprocessing layer is used to collect data and filter the data; the first layer of Bi-LSTM is used to extract vehicle lane change features, then the extracted lane change features are sent to the intent recognition neural network (Bi-LSTM layer2,3,4). The intention recognition neural network analyzes the temporal connection between the lane change features. After that, these features are sent to the attention layer, which assigns weights to data in different time frames. Finally, through the fully-connected (FC) layer, a Softmax function is used to calculate the probability of three driving intentions, namely the probability of left lane change and non-lane change, right lane change.



FIGURE 14. Model training process.

IV. INFLUENCE OF NETWORK STRUCTURE

A. THE RELATIONSHIP BETWEEN MODEL RECOGNITION ACCURACY AND TIME STEP

Figure 12 shows the relevance between the recognition accuracy and the time step. When the time step is no less than 0.4s, the proposed model has good accuracy for lane changing intention. The overall accuracy of driving intent recognition reaches the best when the time step equals to 0.4s. Therefore, the time step of the driving intent recognition model proposed in this paper equals to 0.4s. When the value of time step is too small, too few features will be excavated by the MBi-LSTM layer, hence resulting in the lane change intention easily affected by random interference. When the time step is less than 0.3s, the accuracy drops significantly; that is because, when the time step is too long, MBi-LSTM tends to give higher weight to the later extracted features, hence misidentifying the lane change intention as non-lane change intention.

B. THE RELATIONSHIP BETWEEN MODEL RECOGNITION ACCURACY AND THE NUMBER OF BI-LSTM LAYERS

The performances of Bi-LSTM with different layers are compared to determine the optimal number of layers. The relevance between the accuracy and the number of Bi-LSTM layers is shown in Figure 13. As can be noticed from Figure 13, when layers are greater than three, the MBi-LSTM model has an optimization bottleneck. The recognition accuracy first decreases slightly, then remains unchanged; when layers are no more than three, the prediction accuracy of the MBi-LSTM model increases with the number of layers.

Since Residual MBi-LSTM can solve the optimization bottleneck and gradient disappearance problem, after introducing the residual block structure, the recognition accuracy still keeps increasing even when layers are greater than three. Suppose the number of layers is too small, the lane-changing intention recognition model cannot fully extract the intention, resulting in the lane-changing intention being easily affected by random interference, causing a relatively lower accuracy. Considering the recognition accuracy and prediction time-consuming, a four-layer network structure is adopted in this paper.

TABLE 2. Model performance comparison table.

MODEL	BALANCED ACCURACY	F1-SCORE	RECALL	INTENTION RECOGNITION TIME
<i>Baseline</i>	65.58%	31.78%	47.84%	After significant lateral movement
<i>DCIE [33]</i>	93.07%	Not mentioned	Not mentioned	Average 2.179s before lane changing time
<i>Tang et al. [14]</i>	94.22%	88.71%	93.75%	Average 1.253s before lane changing time when the threshold is 80%
<i>AT-GRU [17]</i>	96.68%	94.05%	96.17%	Average 1.962s before lane changing time
<i>Xie et al. [13]</i>	99.32 %	Not mentioned	Not mentioned	Not recognized until the lane changing time
<i>Our method</i>	98.01%	96.18%	97.78%	Average 2.196s before lane changing time

V. MODEL TRAINING PROCESS

The model loss function is multi-class cross-entropy:

$$Loss = -\frac{1}{batch_size} \sum_{j=1}^{batch_size} \sum_{i=1}^n y_{ji} \log \hat{y}_{ji} \quad (18)$$

where y_{ji} denotes the true distribution, \hat{y}_{ji} denotes the prediction distribution, n denotes the total number of label categories.

The SGD algorithm is used in order to train the model. After 12 epochs, Loss and accuracy gradually converge. The training process curves are shown in Figure 14.

VI. EXPERIMENTAL VERIFICATION

To verify the performance of the driver lane change intention recognition method we proposed, ablation experiments to test the influence of Attention mechanism, bidirectional characteristics and residual block, together with other performance tests were conducted. Finally, hardware-in-loop simulation experiments are also carried out to test the performance of our model.

A. TEST AND ABLATION EXPERIMENT

Because of the imbalance of the dataset, the actual performance of the classifier cannot be accurately reflected by the accuracy rate. Therefore, two metrics widely used for imbalanced multi-classification problems [35], *Balanced Accuracy* and *F1-score*, are introduced. Moreover, recall rate is used in this paper as another important evaluation criterion [36].

$$\text{Balanced Accuracy} = \frac{TPR + TNR}{2} \quad (19)$$

where:

$$TPR = \frac{TP}{TP + FP} \quad (20)$$

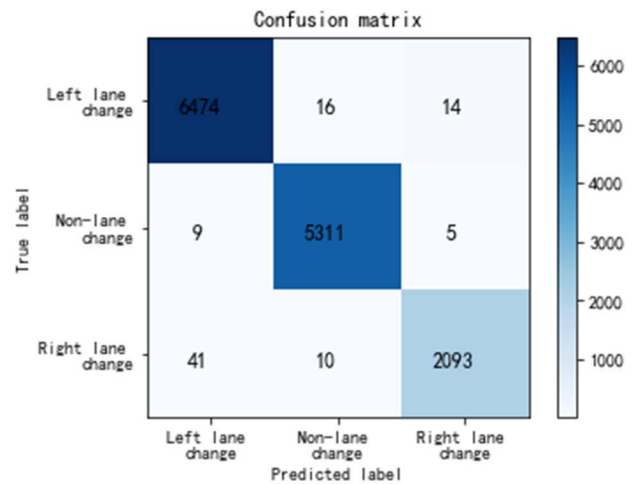
$$TNR = \frac{TN}{TN + FP} \quad (21)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (22)$$

$$\text{precision} = \frac{TP}{TP + FP} \quad (23)$$

$$\text{F1-score} = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\text{Precision} + \beta^2 \times \text{Recall}} \quad (24)$$

where, the parament $\beta = 1$.

**FIGURE 15.** Confusion matrix.

For the purpose of demonstrating the advantages of the proposed model, in this paper, we compared it with four state-of-the-art models. It is worth noting that all models mentioned above are conducted on the HighD data set. In order to ensure the fairness of the model comparison, AT-GRU based model proposed by Guo *et al.* [17], Multi-LSTM based model proposed by Tang *et al.* [14] and our model were tested on the HighD test set, the results were compared with those of DCIE [33] and basic LSTM [13] claimed in the cited literatures. The model performance comparison result is shown in Table 2.

Liu *et al.* [33] proposed the driver characteristic and intention (DCIE) based on the dynamic bayesian network (DBN). In DCIE, the driver characteristic recognition module and lane changing intention identification module are cascaded. In other words, the accuracy of the lane changing intention recognition will be affected by the accuracy of the driver characteristic recognition, hence resulting in lower recognition accuracy.

Similar to our method, Tang *et al.* [14] also employed the LSTM to model the lane changing behaviors. However, the proposed method ignores the weight allocation issues of data in different time frames. consequently, lane changing intention can only be identified after significant lateral movement. From the basic mechanism of the LSTM, Suzuki *et al.* [15] explained this limitation. The weight allocation mechanism

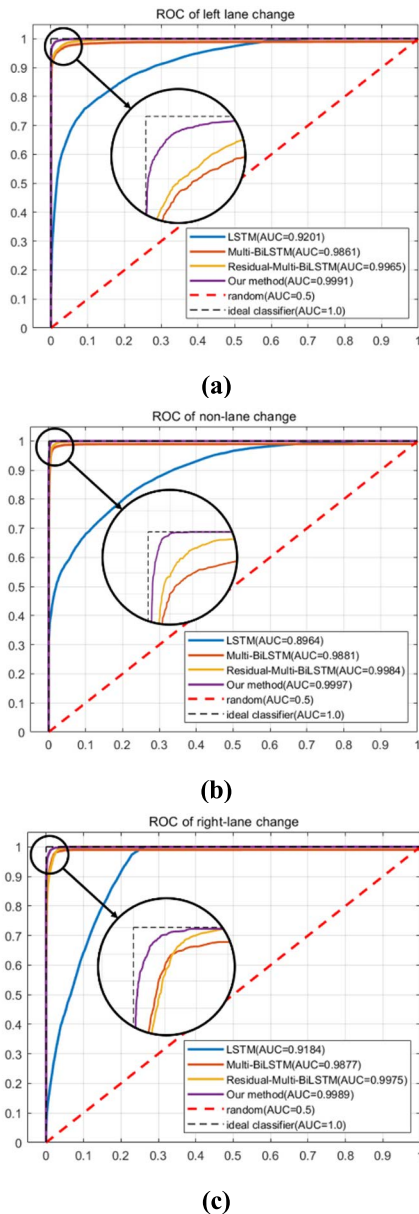


FIGURE 16. Receiver operating characteristic curve of (a) left lane change (b) non-lane change (c) right lane change.

of an LSTM network always tends to allocate higher weights to the variables which are close to the event, while variables far away from the event shall receive a tremendous penalty, hence inhibiting the model from predicting long-term lane changing intentions in advance.

The same limitation also happens in the studies of Xie *et al.* Similar to the method of Tang *et al.*, Xie *et al.* [13] adopt an LSTM network to realize intention recognition, and the accuracy of lane changing intention recognition was higher than 99%. However, their experimental setup is tricky. In their model, 5 seconds data before the lane changing time is used as the model input. That is to say, their model was unable to understand lane changing intention until the vehicle is about

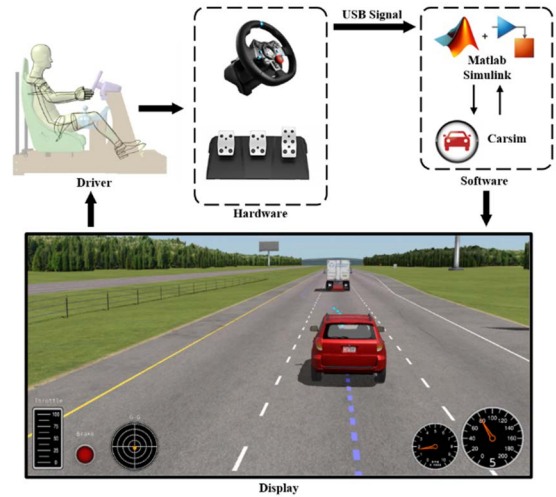


FIGURE 17. Schematic of hardware-in-loop platform.



FIGURE 18. Hardware-in-loop platform.

to cross the lane markings. Though the highest accuracy, it's obvious that their model fails to give a timely warning.

Hao *et al.* [34], Solves the above problems by introducing an attention mechanism. However, the single-layer GRU is still facing optimization bottlenecks, which limits the improvement of recognition accuracy.

In summary, the model proposed has good identification accuracy (98.01%), which is adequate with the four mentioned state-of-art models. The recognition accuracy is further improved by introducing the residual mechanism. By taking advantage of the soft attention mechanism, the proposed model can identify lane changing intentions 2.196 seconds before lane changing time on average, indicating the timeliness of the proposed model.

To visualize the data, the recognition results are plotted as a confusion matrix as shown in Figure 15. Combining Table 2 and Figure 15, it can be seen that our method has good lane-changing intention recognition accuracy.

In this paper, ablation experiments and Receiver Operating Characteristic curve (ROC) are used to test the influence of the attention mechanism, residual block and bidirectional

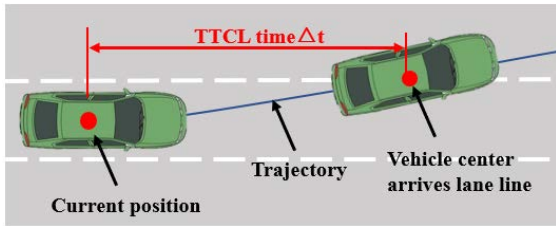


FIGURE 19. The definition of TTCL time.

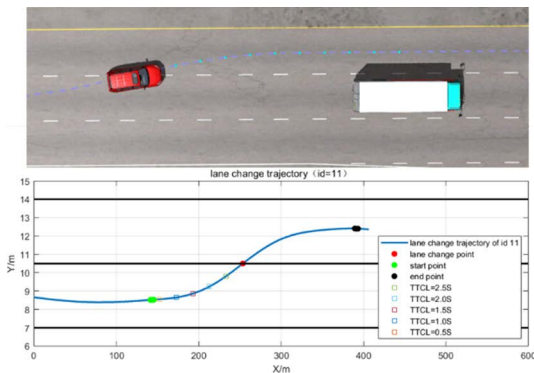


FIGURE 20. Trajectory of vehicle id 11.

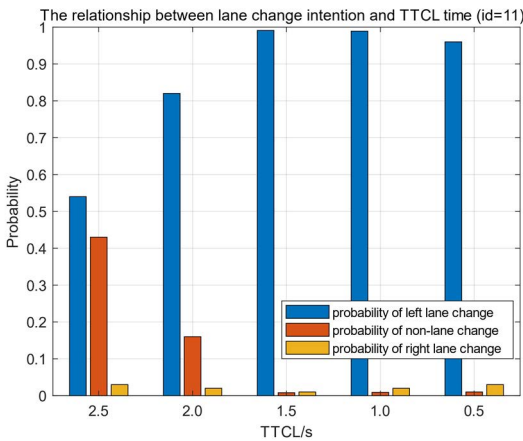


FIGURE 21. The relationship between lane change intention and TTCL time of vehicle id 11.

characteristics. The area under the ROC curve is defined as the AUC area. The closer the AUC region is to 1, the better the performance of the classifier.

As can be seen from Figure 16 (a) (b) (c), the performance of the MBi-LSTM classifier has a greater increase than that of the single-layer LSTM. In the three cases, the AUC increased by 6.6%, 9.17%, and 6.93%, respectively. That is, both the Multi structure and the bidirectional characteristics structure have greatly improved the recognition accuracy of the model. This is mainly because there are many dynamic time-varying factors that affect the vehicle lane change process, the recessive factors, such as the driver’s Driving intent or incorrect operation, etc., are difficult to analyze [37], compared with

single-layer LSTM, MBi-LSTM has more ability to infer driver intention from context information.

At the same time, the residual block structure, which can solve the gradient disappearance and optimization bottleneck problems of deep neural networks, improved the AUC by 1.04%, 0.03%, and 0.98%, respectively, compared with MBi-LSTM.

B. HARDWARE-IN-LOOP SIMULATION EXPERIMENTS AND CASE ANALYSIS

The hardware-in-the-loop (HIL) platform is shown in the Figure 17 and 18. The system consists of a force feedback steering wheel, brake and throttle pedals, screen, Matlab Simulink software, Carsim software and industrial personal computer (IPC).

In the HIL system, Carsim RT dynamic Vehicle software generates virtual animation scenes projected onto a large screen. The IPC collects the states of the brake, throttle pedals and the force feedback steering wheel in real-time at a sampling rate of 25 Hz and feeds it to the intention recognition model running on the IPC.

The HIL experiment was performed on a highway with a test road of 5 km. The test road was a two-way six-lane road with a lane width of 3.75m, and there was an isolation fence in the middle of the road. The minimum speed limit and the maximum speed limit are 60km/h and 120km/h, respectively.

Before the experiment, each participant was asked to take a test drive to fully understand the HIL system. In the experiment, participants were asked to drive freely in the middle lane. When participants see an obstacle vehicle, they are free to choose between changing lanes or keeping lanes. In each experiment, the driver was required to perform at least five lane changes at any position. The test was repeated four times. This experiment also minimizes the learning effect by adjusting the position of obstacle vehicles.

As shown in Figure 19, the time to change lane (TTCL) is defined as the time interval from the current location of the lane-changing vehicle to the vehicle center’s arrival at the lane line. The earlier the vehicle cut-in intention is identified, the better the vehicle’s decision-making [38].

The lane-changing process of vehicle id11 is shown in Figure 20. In which, the ego-vehicle passes the slow-moving truck ahead from the left lane. The lane change intention during the overtaking process is shown in Figure 21, the left lane-changing intention of the ego vehicle has been detected at $\Delta t = 2.5s$, where the probability of left lane change intention $\omega_1 = 54.17\%$, slightly exceeds that of non-lane change. The left lane change intention reaches its peak $\omega_1 = 99.91\%$ at $\Delta t = 1.5s$, after that the probability of left lane change intention decreases slightly, but still remains above the threshold.

Across all simulation experiments, our model can identify driver intention on an average of 2.07 seconds before lane changing time, proving the practical value of our method.

VII. CONCLUSION

1) A driver lane change intention recognition method based on attention enhanced residual-MBi-LSTM is proposed in this paper. The introduction of a bidirectional, Multi-layer structure, attention mechanism and residual block, can help solve long-term dependence problem, optimization bottleneck and adjust information weight distribution, hence ameliorating the model's recognition accuracy.

2) This model was trained and tested on the HighD dataset. The results show that our method has a balanced accuracy rate of higher than 98.01% for the recognition of three types of driving intentions.

3) Based on the simulation experiments conducted in the hardware-in-loop platform, our method can identify driver intention on an average of 2.07 seconds in advance.

REFERENCES

- [1] C. M. Martinez, M. Heucke, B. Gao, D. Cao, and F.-Y. Wang, "Driving style recognition for intelligent vehicle control and advanced driver assistance: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 666–676, Mar. 2018, doi: [10.1109/ITITS.2017.2706978](https://doi.org/10.1109/ITITS.2017.2706978).
- [2] D. Yi, J. Su, C. Liu, and W.-H. Chen, "Trajectory clustering aided personalized driver intention prediction for intelligent vehicles," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3693–3702, Jun. 2019, doi: [10.1109/TII.2018.2890141](https://doi.org/10.1109/TII.2018.2890141).
- [3] O. Scheel, L. Schwarz, N. Navab, and F. Tombari, "Situation assessment for planning lane changes: Combining recurrent models and prediction," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 2082–2088, doi: [10.1109/ICRA.2018.8462838](https://doi.org/10.1109/ICRA.2018.8462838).
- [4] L. Li and P. Li, "Analysis of Driver's steering behavior for lane change prediction," in *Proc. 11th Int. Conf. Intell. Human-Machine Syst. Cybern. (IHMSC)*, Aug. 2019, pp. 71–75, doi: [10.1109/IHMSC.2019.10112](https://doi.org/10.1109/IHMSC.2019.10112).
- [5] H. Zhang and R. Fu, "Driving behavior recognition and intention prediction of adjacent preceding vehicle in highway scene," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 20, no. 1, pp. 40–46, 2020.
- [6] Y. Xing, C. Lv, H. Wang, D. Cao, and E. Velenis, "An ensemble deep learning approach for driver lane change intention inference," *Transp. Res. C, Emerg. Technol.*, vol. 115, Jun. 2020, Art. no. 102615.
- [7] J. Schlechtriemen, A. Wedel, J. Hillenbrand, G. Breuel, and K.-D. Kuhnert, "A lane change detection approach using feature ranking with maximized predictive power," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2014, pp. 108–114, doi: [10.1109/IVS.2014.6856491](https://doi.org/10.1109/IVS.2014.6856491).
- [8] P. Kumar, M. Perrollaz, S. Lefevre, and C. Laugier, "Learning-based approach for online lane change intention prediction," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2013, pp. 797–802, doi: [10.1109/IVS.2013.6629564](https://doi.org/10.1109/IVS.2013.6629564).
- [9] S. Liu, K. Zheng, L. Zhao, and P. Fan, "A driving intention prediction method based on hidden Markov model for autonomous driving," *Comput. Commun.*, vol. 157, pp. 143–149, May 2020, doi: [10.1016/j.comcom.2020.04.021](https://doi.org/10.1016/j.comcom.2020.04.021).
- [10] S. Xiao et al., "Lane change intention recognition method based on an LSTM network," *China J. Highway Transp.*, vol. 34, no. 11, p. 236, 2021.
- [11] A. Benterki, M. Boukhnifer, V. Judalet, and C. Maaoui, "Artificial intelligence for vehicle behavior anticipation: Hybrid approach based on maneuver classification and trajectory prediction," *IEEE Access*, vol. 8, pp. 56992–57002, 2020, doi: [10.1109/ACCESS.2020.2982170](https://doi.org/10.1109/ACCESS.2020.2982170).
- [12] A. Jain, H. S. Koppula, S. Soh, B. Raghavan, A. Singh, and A. Saxena, "Brain4Cars: Car that knows before you do via sensory-fusion deep learning architecture," 2016, *arXiv:1601.00740*.
- [13] D.-F. Xie, Z.-Z. Fang, B. Jia, and Z. B. He, "A data-driven lane-changing model based on deep learning," *Transp. Res. C, Emerg. Technol.*, vol. 106, pp. 41–60, Sep. 2019.
- [14] L. Tang, H. Wang, W. Zhang, Z. Mei, and L. Li, "Driver lane change intention recognition of intelligent vehicle based on long short-term memory network," *IEEE Access*, vol. 8, pp. 136898–136905, 2020, doi: [10.1109/ACCESS.2020.3011550](https://doi.org/10.1109/ACCESS.2020.3011550).
- [15] T. Suzuki, H. Kataoka, Y. Aoki, and Y. Satoh, "Anticipating traffic accidents with adaptive loss and large-scale incident DB," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3521–3529, doi: [10.1109/CVPR.2018.00371](https://doi.org/10.1109/CVPR.2018.00371).
- [16] R. Krajewski, J. Bock, L. Kloeker, and L. Eckstein, "The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2118–2125, doi: [10.1109/ITSC.2018.8569552](https://doi.org/10.1109/ITSC.2018.8569552).
- [17] Y. Guo, H. Zhang, C. Wang, Q. Sun, and W. Li, "Driver lane change intention recognition in the connected environment," *Phys. A, Stat. Mech. Appl.*, vol. 575, Aug. 2021, Art. no. 126057.
- [18] Z. Hao, X. Huang, K. Wang, M. Cui, and Y. Tian, "Attention-based GRU for driver intention recognition and vehicle trajectory prediction," in *Proc. 4th CAA Int. Conf. Veh. Control Intell. (CVCI)*, Dec. 2020, pp. 86–91, doi: [10.1109/CVCI51460.2020.9338510](https://doi.org/10.1109/CVCI51460.2020.9338510).
- [19] X.-W. Ji, "Intention recognition and trajectory prediction for vehicles using LSTM network," *China J. Highway Transp.*, vol. 32, no. 6, p. 34, 2019.
- [20] L. Tijerina, W. Garrott, D. Stoltzfus, and E. Parmer, "Eye glance behavior of van and passenger car drivers during lane change decision phase," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1937, pp. 37–43, Jan. 2005, doi: [10.1177/0361198105193700106](https://doi.org/10.1177/0361198105193700106).
- [21] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.
- [22] G. Du, Z. Wang, B. Gao, S. Mumtaz, K. M. Abualnaja, and C. Du, "A convolution bidirectional long short-term memory neural network for driver emotion recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4570–4578, Jul. 2021, doi: [10.1109/ITITS.2020.3007357](https://doi.org/10.1109/ITITS.2020.3007357).
- [23] D. Cui, S. Yin, J. Gu, L. Liu, and S. Wei, "MSAM: A multi-layer bi-LSTM based speech to vector model with residual attention mechanism," in *Proc. IEEE Int. Conf. Electron Devices Solid-State Circuits (EDSSC)*, Jun. 2019, pp. 1–3, doi: [10.1109/EDSSC.2019.8753946](https://doi.org/10.1109/EDSSC.2019.8753946).
- [24] H. Zhang, Z. Nan, T. Yang, Y. Liu, and N. Zheng, "A driving behavior recognition model with bi-LSTM and multi-scale CNN," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Oct. 2020, pp. 284–289, doi: [10.1109/IV47402.2020.9304772](https://doi.org/10.1109/IV47402.2020.9304772).
- [25] M. Chen and W. Guo, "Multi-factor stock selecting model based on residual net and LSTM deep learning algorithm," in *Proc. IEEE Int. Conf. Adv. Electr. Eng. Comput. Appl. (AEECA)*, Aug. 2020, pp. 1049–1053, doi: [10.1109/AEECA49918.2020.9213636](https://doi.org/10.1109/AEECA49918.2020.9213636).
- [26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [27] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*.
- [28] Y. Cao, F. Yang, Q. Tang, and X. Lu, "An attention enhanced bidirectional LSTM for early forest fire smoke recognition," *IEEE Access*, vol. 7, pp. 154732–154742, 2019, doi: [10.1109/ACCESS.2019.2946712](https://doi.org/10.1109/ACCESS.2019.2946712).
- [29] J. Peng, Y. Guo, R. Fu, W. Yuan, and C. Wang, "Multi-parameter prediction of drivers' lane-changing behaviour with neural network model," *Appl. Ergonom.*, vol. 50, pp. 207–217, Sep. 2015.
- [30] R. Huang, H. Liang, P. Zhao, B. Yu, and X. Geng, "Intent-estimation-and motion-model-based collision avoidance method for autonomous vehicles in urban environments," *Appl. Sci.*, vol. 7, p. 457, May 2017.
- [31] H. Huang, Z. Zeng, D. Yao, X. Pei, and Y. Zhang, "Spatial-temporal ConvLSTM for vehicle driving intention prediction," *Tsinghua Sci. Technol.*, vol. 27, no. 3, pp. 599–609, Jun. 2022, doi: [10.26599/TST.2020.9010061](https://doi.org/10.26599/TST.2020.9010061).
- [32] Y. Xing, C. Lv, and D. Cao, "Personalized vehicle trajectory prediction based on joint time-series modeling for connected vehicles," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1341–1352, Feb. 2020, doi: [10.1109/TVT.2019.2960110](https://doi.org/10.1109/TVT.2019.2960110).
- [33] J. Liu, H. Xiong, T. Wang, H. Huang, Z. Zhong, and Y. Luo, "Probabilistic vehicle trajectory prediction via driver characteristic and intention estimation model under uncertainty," *Ind. Robot, Int. J. Robot. Res. Appl.*, vol. 48, no. 6, pp. 778–791, Nov. 2021, doi: [10.1108/IR-06-2020-0114](https://doi.org/10.1108/IR-06-2020-0114).
- [34] Z. Hao, X. Huang, K. Wang, M. Cui, and Y. Tian, "Attention-based GRU for driver intention recognition and vehicle trajectory prediction," in *Proc. 4th CAA Int. Conf. Veh. Control Intell. (CVCI)*, Dec. 2020, pp. 86–91, doi: [10.1109/CVCI51460.2020.9338510](https://doi.org/10.1109/CVCI51460.2020.9338510).
- [35] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J. Big Data*, vol. 6, p. 27, Dec. 2019.

- [36] J. V. Dueholm, M. S. Kristoffersen, R. K. Satzoda, T. B. Moeslund, and M. M. Trivedi, "Trajectories and maneuvers of surrounding vehicles with panoramic camera arrays," *IEEE Trans. Intell. Veh.*, vol. 1, no. 2, pp. 203–214, Jun. 2016, doi: [10.1109/TIV.2016.2622921](https://doi.org/10.1109/TIV.2016.2622921).
- [37] S. Dai, L. Li, and Z. Li, "Modeling vehicle interactions via modified LSTM models for trajectory prediction," *IEEE Access*, vol. 7, pp. 38287–38296, 2019, doi: [10.1109/ACCESS.2019.2907000](https://doi.org/10.1109/ACCESS.2019.2907000).
- [38] Y. Jeong and K. Yi, "Bidirectional long shot-term memory-based interactive motion prediction of cut-in vehicles in urban environments," *IEEE Access*, vol. 8, pp. 106183–106197, 2020, doi: [10.1109/ACCESS.2020.2994929](https://doi.org/10.1109/ACCESS.2020.2994929).



DENGCHENG LIU received the Ph.D. degree in vehicle engineering from the School of Automotive Studies, Tongji University, Shanghai, China, in 2020. He is currently an Industrial Strategy Researcher with the Nanchang Automotive Institute of Intelligence and New Energy. He also works as a Postdoctoral Fellow at the School of Electronic and Information Engineering, Tongji University, and Jiangling Group New Energy Vehicle Company Ltd. (JMEV). His research interests include new energy vehicle and fault diagnosis.



ZHANQIAN WU is currently pursuing the bachelor's degree in automotive engineering with Tongji University, Shanghai, China. His research interests include vehicle trajectory prediction and driver status monitor.



KAICHONG LIANG received the B.S. degree in vehicle engineering from Hunan University, Changsha, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Automotive Study, Tongji University, Shanghai. His research interests include driving style identification, vehicle trajectory prediction, and path planning.



ZHIGUO ZHAO (Member, IEEE) received the Ph.D. degree from Northwestern Polytechnical University, Xi'an, China, in 2002. He was a Postdoctoral Fellow Researcher at Tongji University, Shanghai, China, in 2004. He is currently a Professor with the School of Automotive Study, Tongji University. He has authored over 180 journal articles. He is the co-holder of 46 patent applications. His research interests include vehicle and electric vehicle dynamics and control, hybrid vehicle powertrain system control, and transmission control.

• • •