



Article

Interactive Vehicle Trajectory Prediction for Highways Based on a Graph Attention Mechanism

Zhenyu Song and Yubin Qian *

School of Mechanical and Automotive Engineering, Shanghai University of Engineering Science,
Shanghai 201620, China; m310121477@sues.edu.cn

* Correspondence: qianyub@sues.edu.cn

Abstract: Precise trajectory prediction is pivotal for autonomous vehicles operating in real-world traffic conditions, and can help them make the right decisions to ensure safety on the road. However, state-of-the-art approaches consider limited information about the historical movements of vehicles. On highways, drivers make their next judgments according to the behavior of the ambient vehicles. Thus, vehicles need to consider temporal and spatial interactions to reduce the risk of future collisions. In the current work, a trajectory prediction method is put forward in accordance with a graph attention mechanism. We add the absolute and relative motion information of vehicles to the input of the model to describe the vehicles' past motion states more accurately. LSTM models are employed to process the historical motion information of vehicles, as well as the temporal correlations in interactions. The graph attention mechanism is applied to capture the spatial correlations between vehicles. Utilizing a decoder rooted in an LSTM framework, the future trajectory distribution is generated. Evaluation on the NGSIM US-101 and I-80 datasets substantiates the superiority of our approach over existing state-of-the-art algorithms. Moreover, the predictions of our model are analyzed.

Keywords: autonomous vehicles; trajectory prediction; vehicle interactions; graph attention mechanism



Citation: Song, Z.; Qian, Y. Interactive Vehicle Trajectory Prediction for Highways Based on a Graph Attention Mechanism. *World Electr. Veh. J.* **2024**, *15*, 96. <https://doi.org/10.3390/wevj15030096>

Academic Editor: Joeri Van Mierlo

Received: 29 January 2024

Revised: 23 February 2024

Accepted: 29 February 2024

Published: 5 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Autonomous technology is regarded as an important solution to decrease the presence of traffic accidents and increase traffic safety. Cognition and decision-making constitute the foundational elements of autonomous vehicular navigation [1]. On the road, autonomous vehicles detect and sense their surroundings by analyzing the information collected by sensors to maintain their stability. However, there is a challenging task between these two modules—vehicle trajectory forecasting. The goal of vehicle trajectory forecasting is to ensure that self-driving cars are able to forecast the trajectory layout in the future when facing complex traffic scenarios, and to help autonomous vehicles understand what is going to happen in the future and make effective decisions to enhance driving safety. Such prediction becomes difficult due to the uncertainty of human driving behavior.

Vehicle trajectory prediction is usually inferred on the basis of the motion characteristics of the vehicles in the past [2]. However, people often focus only on the historical positions of the vehicles and ignore other useful information. Employing extensive historical data enriches the environmental context surrounding the subject vehicle, thereby enhancing the precision of trajectory forecasting to facilitate a more judicious and secure navigational pathway. In fact, there are a large number of reasons that may influence driving action, of which the primary category are the reasons associated with vehicle kinematics, such as vehicle speed, acceleration, etc. [3]. However, the forthcoming trajectory of a vehicle is influenced not merely by its own kinematic state but significantly by the dynamics of surrounding vehicles. As the environment becomes more complex, experienced drivers can choose an appropriate road to drive on by judging the intentions of the drivers of the ambient vehicles. Therefore, the movement states of the surrounding vehicles also need

to be considered. The secondary set of variables encompasses the kinematic attributes of proximate vehicles relative to the subject vehicle, such as their spatial positioning and velocity vectors. The third category is the factors caused by the drivers themselves, such as psychological factors [4]. We use the former two factors in this paper to increase the input information and make the vehicle's historical information as rich as possible.

Furthermore, intervehicle interactions serve as an additional critical determinant for trajectory prediction. Within a shared operational environment, individual vehicles do not operate in isolation; rather, their actions reciprocally impact each other's trajectory forecasts. Initially, prioritization should be accorded to the temporal window most impactful to the subject vehicle's future trajectory, thereby enabling efficacious navigational decisions to preempt potential collisions. Subsequently, the focus should be directed towards quantifying the influence exerted by surrounding vehicles on the subject vehicle. This analytical approach aids in isolating high-impact variables, thereby allowing drivers to concentrate primarily on the most influential external agents, minimizing navigational distractions. For instance, in a lane-changing scenario, the vehicles within the destination lane are accorded elevated attention relative to those in alternative lanes. Consequently, a nuanced understanding of the temporal dynamics and spatial relevance of intervehicular interactions is central to our research endeavors.

Aiming at solving these problems, we built an LSTM model according to a graph attention mechanism (GA-LSTM), which can focus on key time series and vehicle interactions in temporal and spatial terms, respectively. In the current traffic scenario, the graph attention (GAT) mechanism offers an effective framework for capturing spatial interactions across vehicles within a single temporal snapshot [5]. LSTM models are adopted to encode historical information about different vehicles simultaneously and generate future vehicle trajectory estimation within a predictive scope after aggregating the features of all vehicles. The important results of the present article are shown below:

1. We enhance the input information of the model by introducing the absolute and relative motion information of vehicles, enhancing the vehicle interaction relationship, and providing more comprehensive information on the historical motion of vehicles.
2. We propose graph-attention-integrated LSTM for trajectory prediction (GA-LSTM) to achieve the representation of temporal and spatial dependencies between the subject as well as the ambient vehicles on the highway.

2. Related Research

Recently, plenty of scientists have accomplished the approach to vehicular trajectory forecasting. We will summarize these existing methods, focusing on the latter two.

Methods based on physics: Physics-driven frameworks characterize vehicles as dynamic entities compliant with mechanical laws, primarily utilizing kinematic and dynamic equations to forecast the future trajectory of the moving entity. These models usually consider vehicle speed, acceleration, and external environmental conditions such as the road friction coefficient. Veeraraghavan et al. [6] combined the unscented transformation sampled by a switched Kalman filter to provide an accurate trajectory inference at traffic junctures. Yu [7] amalgamated a 4-DoF vehicle model with a trace-free Kalman filter to enhance predictive fidelity. However, such methods can only accomplish trajectory prediction for a short period of time, and it is hard to acquire perfect accuracy. The models cannot use the interactions between the vehicles to predict the change in motion. The cooperation of autonomous vehicles is also very important. Semsar-Kazerooni et al. [8] used an artificial potential function to design a controller for cooperative adaptive cruise control by defining appropriate control laws where the system state is always driven to the minimum of the designed potential function. Liang [9] proposed a multi-agent system based distributed control architecture together with a hierarchical controller for the CAVs cooperation control system. Longitudinal, lateral, and yaw integration control of CAVs was realized by combining an artificial potential field with the distributed model predictive control algorithm. An optimal solution strategy was introduced to solve the CAVs cooperation

problem, and multiple constraints were designed to ensure the safety of vehicle spacing and vehicle stability.

Methods based on maneuvering: In these models based on maneuverability, the subject vehicle makes a series of actions according to the information from other vehicles on the road. These models usually consist of two parts: the first part is a maneuver recognition module, and the second part is a trajectory prediction module that makes better predictions of future trajectories of the vehicles based on the recognized maneuvers. The parts perform a similar task to classifiers, using the vehicles' motion states and locations as the input features, while the output is the vehicles' positions under different maneuvers in the future. Scholars have used classifiers (e.g., hidden Markov models, Bayesian networks, heuristic-based classifiers, and random forest classifiers [10–14] for maneuver recognition). Xing [15] used an unsupervised clustering algorithm to identify three driving styles and generate vehicle-specific driving styles through a Gaussian mixture model. The shortcomings of this method are that, as the traffic scenes become more complex, it is difficult for the models to distinguish different behaviors of vehicles. Moreover, manually marking the trajectories is very time-consuming, which tends to affect the accuracy of the model classifications.

Recurrent neural network-based methods: As trajectory forecasting is considered as a time series regression or classification issue, methods in accordance with recurrent neural networks are increasingly applied to such tasks. The long short-term memory (LSTM) architecture, a specialized form of recurrent neural network, effectively captures long-term dependencies between features and decides selectively whether the information is retained or not by gating units. In recent years, different architectures of LSTM networks have been used for vehicle trajectory prediction [16,17]. Altche [18] and Zyner [19] both used a single LSTM for modeling. Xin [20] used a double LSTM. Two core modules were delineated: the first ascertained driver intent through behavioral feature extraction, while the second extrapolated future trajectories. In vehicle interaction simulation, Deo et al. [21] employed an encoder–decoder schema for trajectory estimation. An additional convolutional social pooling layer was added to the social tensor to describe the interactions between vehicles. Finally, the decoder generated a multi-modal trajectory distribution based on the six driving behaviors. Alahi [22] integrated a social pooling layer to aggregate the LSTM's hidden states, thereby extracting inter-pedestrian correlations. Liu et al. [23] devised a vehicular risk map, capturing interactive dynamics to determine the subject vehicle's trajectory risk index. However, these methods lack specificity in portraying interactions between the subject vehicle and adjacent vehicles, and fail to quantify the influence exerted by adjacent vehicles on the subject vehicle.

Graph neural networks (GNNs): GNNs represent frameworks for learning directly from graph-structured information. GNNs have made significant breakthroughs in many different areas [24]. Li et al. [25] used static and dynamic graphs, respectively, to forecast the trajectories of different traffic participants to reduce the probability of autonomous vehicle accidents. There are some methods [26,27] that apply GNNs to spatiotemporal data. The graph attention (GAT) mechanism [5] is one of these methods, which represents the influence of neighboring nodes by assigning them different importance. Huang [28] applied GAT to research on pedestrian trajectory prediction and obtained excellent results. For our problem, we use GAT to model the spatial information of the vehicles. Additionally, the graphs are designed to characterize complex interactions.

3. Problem Description

To anticipate the probabilistic spatial positioning of the subject vehicle in future instances, both absolute and relative vehicular motion data from historical timestamps are essential.

3.1. Coordinate System

We use a static coordinate system, as indicated in Figure 1. The x -axis denotes the travel direction of the subject vehicle along the highway, while the y -axis is oriented perpendicular to the x -axis. This allows our model to be more independent of the curvature of the road.

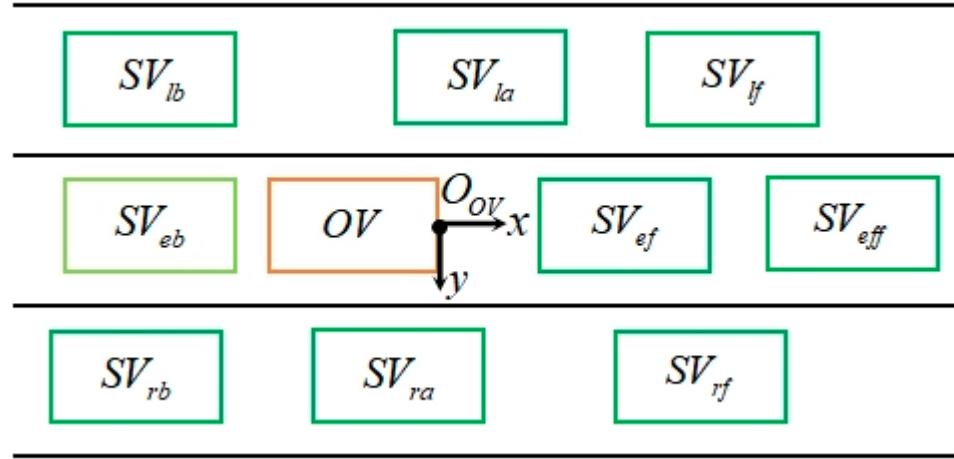


Figure 1. The local traffic scene graph with the subject vehicle at the present moment as the origin.

3.2. Construction of Local Traffic Scenes

There are two methods for constructing local traffic scenes. The first one is to construct based on the distance of the subject vehicle, although this approach is commonly used, it is too subjective. Therefore, we adopt the second method, which is constructed according to the spatial proximity relationship of the subject vehicle. Figure 1 shows a local traffic scene that we constructed, consisting of ten vehicles, where the front vehicle of the front vehicle of the subject vehicle is also taken into account. This method is not limited by distance, and we extract the surrounding vehicles only if they appear at the corresponding locations near the subject vehicle. If no vehicles appear at these specific locations, they are not considered.

3.3. Inputs and Outputs

The model's input is divided into two discrete segments.

The first part is the historical motion information of the subject vehicle C_{ov} , which includes the positions, velocities, and accelerations:

$$C_{ov} = [c_{ov}^{t_{obs}-t_h}, \dots, c_{ov}^{t_{obs}-\Delta t}, c_{ov}^{t_{obs}}] \quad (1)$$

where,

$$c_{ov}^{t_{obs}} = [x_{ov}^{t_{obs}}, y_{ov}^{t_{obs}}, v_{ov}^{t_{obs}}, a_{ov}^{t_{obs}}] \quad (2)$$

where, $x_{ov}^{t_{obs}}$ and $y_{ov}^{t_{obs}}$ denote the position of the subject vehicle at time t_{obs} , respectively, $v_{ov}^{t_{obs}}$ denotes the speed of the subject vehicle at time t_{obs} , $a_{ov}^{t_{obs}}$ denotes the acceleration of the subject vehicle at time t_{obs} .

The second part is the previous motion data of the neighboring vehicles C_{svi} , incorporating positional, velocity, and acceleration data of both surrounding and adjacent vehicles relative to the subject vehicle:

$$C_{svi} = [c_{svi}^{t_{obs}-t_h}, \dots, c_{svi}^{t_{obs}-\Delta t}, c_{svi}^{t_{obs}}] i \in \{lb, la, lf, eb, ef, eff, rb, ra, rf\} \quad (3)$$

where,

$$c_{svi}^{t_{obs}} = [x_{svi}^{t_{obs}}, y_{svi}^{t_{obs}}, v_{svi}^{t_{obs}}, a_{svi}^{t_{obs}}, \Delta x_{svi,ov}^{t_{obs}}, \Delta y_{svi,ov}^{t_{obs}}, \Delta v_{svi,ov}^{t_{obs}}, \Delta a_{svi,ov}^{t_{obs}}] \quad (4)$$

where, $x_{svi}^{t_{obs}}$ and $y_{svi}^{t_{obs}}$ denote the positions of the surrounding vehicles at time t_{obs} , respectively, $v_{svi}^{t_{obs}}$ denotes the velocities of the surrounding vehicles at time t_{obs} , $a_{svi}^{t_{obs}}$ denotes the accelerations of the surrounding vehicles at time t_{obs} , $\Delta x_{svi,ov}^{t_{obs}}$ and $\Delta y_{svi,ov}^{t_{obs}}$ denote the positions of the surrounding vehicles relative to the subject vehicle at time t_{obs} , respectively, $\Delta v_{svi,ov}^{t_{obs}}$ denotes the velocities of the surrounding vehicles relative to the subject vehicle at time t_{obs} , $\Delta a_{svi,ov}^{t_{obs}}$ denotes the accelerations of the surrounding vehicles relative to the subject vehicle at time t_{obs} .

The model yields a probabilistic distribution delineating the forthcoming spatial coordinates of the subject vehicle:

$$Y = [y^{t_{obs}+\Delta t}, \dots, y^{t_{obs}+t_f}] \quad (5)$$

where,

$$y^{t_{obs}} = [x_{ov}^{t_{obs}}, y_{ov}^{t_{obs}}] \quad (6)$$

Given the inherent unpredictability of trajectory prediction, it is posited that the subject vehicle's positions conform to a Gaussian distribution throughout the predictive time horizon.

$$\theta = [\theta^{t_{obs}+\Delta t}, \theta^{t_{obs}+2\Delta t}, \dots, \theta^{t_{obs}+t_f}] \quad (7)$$

representing the Gaussian distribution parameters of the positions of the subject vehicle across the predictive timeframe, including its mean vector and covariance matrix.

4. Model

The architecture of our proposed model is illustrated in Figure 2, encompassing an LSTM encoder, a graph attention mechanism, and an LSTM decoder. A homologous encoder–decoder structure is applied to Deo [21]. In lieu of convolutional pooling strata, we incorporate a graph attention mechanism to augment model efficacy.

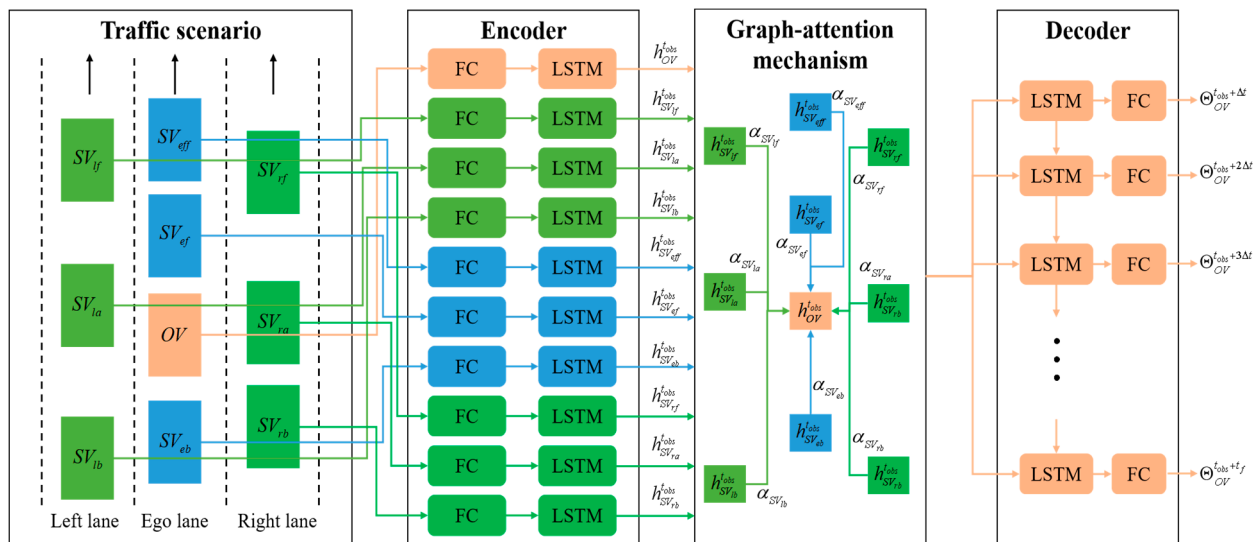


Figure 2. Proposed model: The LSTM encoder ingests past vehicular motion data, while the graph attention mechanism contextualizes spatial interactions between the subject and surrounding vehicles. The decoder then extrapolates the future trajectory of the subject vehicle.

4.1. LSTM Encoder

The encoder scrutinizes past vehicular movements within the existing traffic milieu. It comprises fully connected and LSTM layers, where vehicle-specific weights are universally applied. Across historical time intervals, each vehicle's motion metrics are fed through the encoder:

$$e_i^{t_{obs}} = FC(c_{ov}^{t_{obs}}, c_{svi}^{t_{obs}}; W_{emb}) \quad (8)$$

$$h_{svi}^{t_{obs}} = LSTM(h_{svi}^{t_{obs}-\Delta t}, e_i^{t_{obs}}; W_{encoder}) \quad (9)$$

where, $FC()$ is a fully connected function with the activation function LeakyReLU, W_{emb} and $W_{encoder}$ are the weights of its embedding layer as well as the encoder. $h_{svi}^{t_{obs}}$ and $h_{ov}^{t_{obs}}$ denotes the contemporaneous hidden vectors for the surrounding and subject vehicles, respectively.

4.2. Graph Attention Mechanism

However, the interactions between vehicles cannot be represented by using only LSTMs. To share information among vehicles on the highway, we consider the vehicles as nodes on the graph. Because the graph attention mechanism can collect information from neighboring nodes for aggregation by assigning different levels of importance to them according to their influence, we chose to use GAT as our sharing mechanism. For the graph attention mechanism, nodes and edges are the two most important constituent elements. As shown in Figure 2, in our model, each node represents the feature vector $h^{t_{obs}} = [h_{ov}^{t_{obs}}, h_{svi}^{t_{obs}}]$ encoded by each vehicle at t_{obs} , and each edge represents the weight of the ambient vehicles on the subject vehicle.

GAT computes the features of nodes by focusing on each node's neighbors and combining them with the data from the graph structure. Multiple graph attention layers are stacked into GAT.

During the observation period, $h_{ov}^{t_{obs}}$ is sent to the graph attention layer. For the node pair (OV, SV_i) , its weight in the attention mechanism can be represented as:

$$\alpha_{ov,sv_i}^{t_{obs}} = \frac{\exp(\text{LeakyReLU}(a^T [Wh_{ov}^{t_{obs}} || Wh_{sv_i}^{t_{obs}}])))}{\sum_{sv_i \in N_i} \exp(\text{LeakyReLU}(a^T [Wh_{ov}^{t_{obs}} || Wh_{sv_i}^{t_{obs}}])))} \quad (10)$$

where, $||$ denotes the splicing operation between vectors, T denotes the transposition of a matrix, $\alpha_{ov,sv_i}^{t_{obs}}$ denotes the attention weight of the node SV_i compared to the node OV at time t_{obs} , N_i denotes the set of all neighboring nodes of the node OV . $W \in R^{F' \times F}$ denotes the learnable shared-weight matrix, and $a \in R^{2F'}$ denotes the learnable weight vector. This is normalized by applying the LeakyReLU activation function.

Following acquiring the attention weights, the output of the node OV in the single graph attention layer at time t_{obs} is represented as:

$$\bar{h}_{ov}^{t_{obs}} = \sigma(\sum_{SV_i \in N_i} \alpha_{ov,sv_i}^{t_{obs}} Wh_{sv_i}^{t_{obs}}) \quad (11)$$

where, σ is a nonlinear function. Equations (10) and (11) exhibit how a single graph attention layer operates. $\bar{h}_{ov}^{t_{obs}}$ is the feature vector generated by aggregating the spatial information of all ambient vehicles for the subject vehicle at time t_{obs} .

4.3. LSTM Decoder

The decoder acquires important information about the vehicles according to the feature vector. It is employed to give a predicted probability regarding the subject vehicles' positions in the future during the following time t_f by outputting the Gaussian distribution parameters:

$$\theta^{t_{obs}} = \Lambda(LSTM(\bar{h}_{ov}^{t_{obs}-\Delta t}, W_{dec})) \quad (12)$$

where $\theta^{t_{obs}}$ denotes the parameters of output in terms of the subject vehicle position distribution at t_{obs} , $\Lambda()$ denotes a completely connected function with the activation function LeakyReLU, and W_{dec} denotes the LSTM decoder's weights.

4.4. Training and Implementation Details

The number of units in both the encoder and the decoder is 128. The size of the embedding vector is 32. In addition, we utilize the Adam optimization algorithm [29] with a learning rate of 0.001 and the ReLU activation function with $\alpha = 0.1$ to train the model. The batch size is 128. The model is implemented using Pytorch [30].

5. Empirical Assessment

5.1. Dataset

For the current inquiry, the publicly available NGSIM US-101 [31] and I-80 [32] vehicular trajectory datasets serve as the experimental foundation. Each dataset is composed of trajectories from real highway scenes observed by the camera at 10 Hz in 2005. Each dataset contains three 15 min periods representing three traffic states: light congestion, moderate congestion, and full congestion during peak hours. Each dataset is partitioned into training and testing subsets, comprising approximately 75% and 25%, respectively. Trajectory sequences are segmented into 8 s intervals, utilizing the initial 3 s vehicle motion history to extrapolate the subsequent 5 s trajectory of the subject vehicle. To enhance computational efficiency, segments are down-sampled to 5 fps.

5.2. Evaluation Metrics

To validate trajectory prediction accuracy, the root mean squared error (RMSE) between the actual and predicted future trajectories across a 5 s horizon is employed, as corroborated by prior studies [17,22]. RMSE is computed utilizing the Gaussian distribution's predicted means and quantifies the divergence between real and estimated positions, defined as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{i,t_{obs}}^{pred} - x_{i,t_{obs}}^{true})^2 + (y_{i,t_{obs}}^{pred} - y_{i,t_{obs}}^{true})^2} \quad (13)$$

where, N represents the amount of trajectories within the testing set, $x_{i,t_{obs}}^{pred}$ and $y_{i,t_{obs}}^{pred}$ denote the predicted position for the trajectory i at time t_{obs} , respectively, $x_{i,t_{obs}}^{true}$ and $y_{i,t_{obs}}^{true}$ indicate the true position for the trajectory i at time t_{obs} , respectively.

5.3. Compared Models

In the following sections, trajectory prediction models are compared:

Constant Velocity (CV): The model uses a vehicle's constant speed for trajectory prediction.

Convolutional Social Pooling (CS-LSTM) [21]: The model utilizes convolutional pooling layers and generates single-mode trajectory predictions.

Non-Local Social Pooling (NLS-LSTM) [16]: This model integrates a social pooling layer to encapsulate vehicle interactions within the prevailing traffic landscape, irrespective of spatial proximity.

Multi-Head Attention Social Pooling (MHA-LSTM) [17]: This model uses a four-head attention mechanism for trajectory prediction and does not input additional vehicle information.

Dual Learning Model (DLM) [33]: The model uses a risk map to consider collision time and uses ConVLSTM to represent the spatiotemporal interactions of vehicles.

Driving Risk Map-Integrated Deep Learning (DRM-DL) [23]: The model generates trajectories based on CVAE, constructs a risk map to achieve the interactions between vehicles, and represents the probability distribution of trajectories in accordance with the trajectory risk value.

Graph Attention-LSTM (GA-LSTM): This is the model put forward in the present work.

5.4. Results

Table 1 enumerates the RMSE metrics for the evaluated models. As evidenced by the tabulated outcomes, our proposed architecture demonstrates superior performance, affirming its efficacy.

Table 1. Comparative root mean squared prediction error (RMSE) across a 5 s forecasting horizon.

Evaluation	Prediction	CV	CS-LSTM	NLS-LSTM	MHA-LSTM	DLM	DRM-DL	GA-LSTM
Metric	Horizon(s)							
RMSE (M)	1	0.73	0.61	0.56	0.56	0.41	0.42	0.27
	2	1.78	1.27	1.22	1.22	0.95	0.88	0.60
	3	3.13	2.09	2.02	2.01	1.72	1.43	0.99
	4	4.78	3.10	3.03	3.00	2.64	2.15	1.47
	5	6.68	4.37	4.30	4.25	3.87	3.07	2.15

We observe that the constant velocity (CV) model yields elevated RMSE values and its performance deteriorates with temporal progression. This decrement is attributed to the CV model's sole reliance on the vehicles' physical states, while neglecting the kinematics of surrounding vehicles. This also highlights the importance of vehicle interaction information for trajectory prediction.

It is easy to notice that MHA-LSTM performs better than CS-LSTM and NLS-LSTM, suggesting that vehicle interaction information can be captured better using attention mechanisms compared to convolutional layers.

In addition, we observe that DLM produces lower RMSE values than MHA-LSTM. The risk map in DLM portrays the uncertainty of vehicle motion better by describing the hazard level of the current traffic scenario.

Finally, our proposed model reduces the prediction error by about 30% compared to DRM-DL over the same prediction horizon. Because the input information to our model is richer and the model knows how to describe the past motion of vehicles better, the model improves the accuracy of trajectory prediction significantly. It has been shown that it is more effective to consider the relative significance of the ambient vehicles with the help of the absolute and relative motion information of vehicles than to introduce risk maps.

5.5. Qualitative Analysis

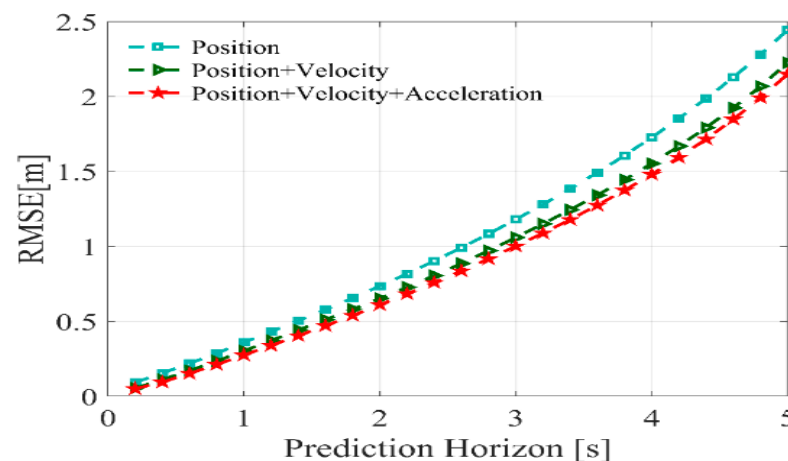
In this section, a qualitative analysis of the predictions made by our model is performed.

5.5.1. Effects of Different Input Features

Table 2 and Figure 3 show the RMSE values of our model with various input features when considering the relative motion information of vehicles. We find that the model produces the highest RMSE values when the input information for our model is only the positions of the vehicles and the positions of the ambient vehicles relative to the subject vehicle. Because the input information of the model is relatively singular, it cannot describe the motion states of the vehicles well and affects the subsequent trajectory prediction. The model performs better when the velocity information of the vehicles is added. The performance of the model is further enhanced when the acceleration information of vehicles continues to be added. This illustrates that the relationship between vehicles is fully associated with positions, velocities, and accelerations. The motion states of the vehicles can be described by velocity and acceleration, and the inclusion of the velocity and acceleration information of the vehicles enhances the accuracy of trajectory prediction.

Table 2. Comparative evaluation of model performance utilizing varied input features.

Model	Prediction Horizon(s)				
	1	2	3	4	5
GA-LSTM [Position]	0.36	0.74	1.18	1.73	2.44
GA-LSTM [Position–Velocity]	0.30	0.65	1.06	1.55	2.23
GA-LSTM [Position–Velocity–Acceleration]	0.28	0.62	1.00	1.48	2.15

**Figure 3.** Comparison of the results of our model with various input features.

5.5.2. Effects of Different Input Lengths of Historical Trajectories

Table 3 as well as Figure 4 represent the RMSE values of the model with various input lengths of historical trajectories. Our model performs better as the input length of the historical trajectories increases. When the input length of the historical trajectories is 1 s, although the prediction effect does not differ much from that of other input lengths over a 1 s prediction horizon, the RMSE value rises faster as the prediction horizon increases. Because it contains too little trajectory information and the model is not sufficiently trained to fit well, our model, which selects a longer input length of historical trajectories appropriately, can enhance the accuracy of the trajectory forecast.

Table 3. Assessment of model performance with diverse historical trajectory durations.

Model	Prediction Horizon(s)				
	1	2	3	4	5
GA-LSTM [1s]	0.34	0.73	1.15	1.64	2.31
GA-LSTM [2s]	0.30	0.66	1.06	1.54	2.18
GA-LSTM [3s]	0.28	0.62	1.00	1.48	2.15
GA-LSTM [4s]	0.26	0.58	0.96	1.43	2.10

5.6. Visualization Outcomes

In order to highlight the performance of the vehicle trajectory forecast more intuitively, a vehicle lane change trajectory was randomly chosen, and the change in trajectory prediction results were observed over the entire process of the vehicle lane change. We took six pictures within the specified time, as described in Figure 5. Figure 5a illustrates the initial recognition of lane-changing attributes by the predicted trajectory. In Figure 5b–f, the predicted trajectories gradually show the typical features of lane changes and converge to the true trajectory as time goes on. We can see that these two trajectories are very similar over the prediction horizon, which proves the efficiency of our model.

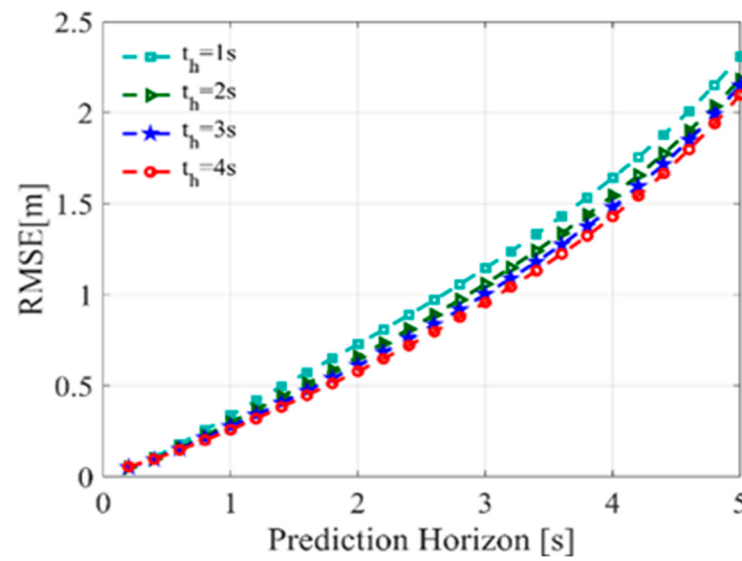


Figure 4. Comparative analysis of model outcomes based on variable historical trajectory durations.

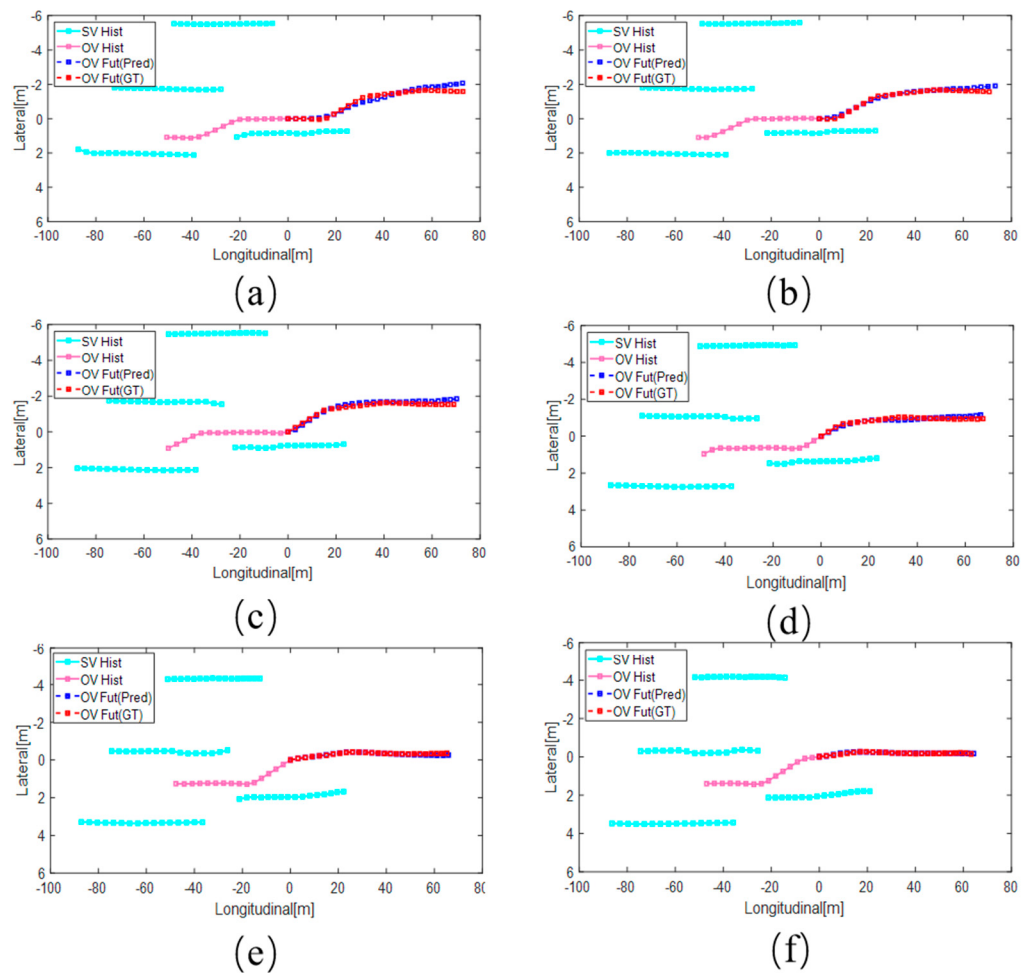


Figure 5. Visual comparison between predicted and actual trajectories for a representative case.

6. Conclusions and Future Work

The current investigation introduces an interactive methodology for forecasting the future trajectory of the subject vehicle. Initially, the model assimilated both absolute and relative kinetic parameters to provide a multidimensional description of the vehicle's

historical motion. Subsequently, long short-term memory (LSTM) networks were employed to encapsulate the historical motion data and discern temporal inter-dependencies in vehicle interactions. Concurrently, a graph attention mechanism was implemented to delineate the spatial interplay between the subject vehicle and its surrounding counterparts. The decoding component ultimately generated a Gaussian distribution, representing the future trajectory of the subject vehicle, based on the graph attention mechanism's output.

In comparison with existing trajectory prediction models, we find that our model is superior to other models with respect to RMSE values on two public natural vehicle trajectory datasets. Qualitative analysis shows that our model performs better with the addition of the absolute and relative vehicle motion information, demonstrating the validity of the input information. The input length of vehicle historical trajectories also affects the effectiveness of the model. The graphical outputs substantiate that our model proficiently identifies lane-changing behavior, thereby corroborating the prediction's fidelity.

One shortcoming of our method is that it is only applicable in highway scenarios. Future work will focus on expanding the method to other traffic scenarios, including intersections and roundabouts. In addition, we consider extending our proposed approach to complex traffic scenarios with various agents (e.g., bicycles, pedestrians, or trucks).

In addition, the graph attention network (GAT) for processing interaction features is more suitable for node-invariant scenarios; it needs to be defaulted that the surrounding entities of the subject vehicle will not change during the historical observation time and the future to-be-predicted time, but in real traffic environments, there is no guarantee that a certain interacting vehicle will keep traveling near the subject vehicle during the prediction time, and the surrounding vehicles can undergo lane-changing behaviors to move away from the subject vehicle and quit the domain range of the subject vehicle. In the future, the graph attention network for extracting interaction features can be improved and structurally optimized to make it suitable for prediction scenarios with variable nodes.

Author Contributions: Conceptualization, Z.S. and Y.Q.; methodology, Z.S.; software, Z.S.; validation, Z.S. and Y.Q.; formal analysis, Z.S.; investigation, Y.Q.; resources, Y.Q.; data curation, Z.S.; writing—original draft preparation, Z.S.; writing—review and editing, Y.Q.; visualization, Z.S.; supervision, Z.S.; project administration, Y.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

LSTM: long short-term memory, GAT: graph attention, CAVs: cooperation of autonomous vehicles, OV: subject vehicle, SV_i : surrounding vehicle, GNNs: graph neural networks, GA-LSTM: graph attention-LSTM, RMSE: root mean squared error, CV: constant velocity, CS-LSTM: convolutional social pooling, NLS-LSTM: non-local social pooling, MHA-LSTM: multi-head attention social pooling, DLM: dual learning model, DRM-DL: driving risk map-integrated deep learning.

References

1. Hang, P.; Lv, C.; Huang, C.; Cai, J.; Hu, Z.; Xing, Y. An integrated framework of decision making and motion planning for autonomous vehicles considering social behaviors. *IEEE Trans. Veh. Technol.* **2020**, *69*, 14458–14469. [\[CrossRef\]](#)
2. Sheng, Z.; Xu, Y.; Xue, S.; Li, D. Graph-based spatial-temporal convolutional network for vehicle trajectory prediction in autonomous driving. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 17654–17665. [\[CrossRef\]](#)
3. Hu, Y.; Zhan, W.; Tomizuka, M. Probabilistic prediction of vehicle semantic intention and motion. In Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 2018; pp. 307–313.

4. Sarkar, A.; Czarnecki, K. A behavior driven approach for sampling rare event situations for autonomous vehicles. *arXiv* **2019**, arXiv:1903.01539.
5. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph attention networks. *arXiv* **2017**, arXiv:1710.10903.
6. Veeraraghavan, H.; Papanikolopoulos, N.; Schrater, P. Deterministic sampling-based switching kalman filtering for vehicle tracking. In Proceedings of the 2006 IEEE Intelligent Transportation Systems Conference, Toronto, ON, Canada, 17–20 September 2006; pp. 1340–1345.
7. Yu, H.; Duan, J.; STaheri Cheng, H.; Qi, Z. A model predictive control approach combined unscented Kalman filter vehicle state estimation in intelligent vehicle trajectory tracking. *Adv. Mech. Eng.* **2015**, *7*, 1687814015578361. [\[CrossRef\]](#)
8. Semsar-Kazerooni, E.; Verhaegh, J.; Ploeg, J.; Alirezaei, M. Cooperative adaptive cruise control: An artificial potential field approach. In Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV), Gothenburg, Sweden, 19–22 June 2016; pp. 361–367.
9. Liang, J.; Li, Y.; Yin, G.; Xu, L.; Lu, Y.; Feng, J.; Shen, T.; Cai, G. A MAS-Based Hierarchical Architecture for the Cooperation Control of Connected and Automated Vehicles. *IEEE Trans. Veh. Technol.* **2022**, *72*, 1559–1573. [\[CrossRef\]](#)
10. Deo, N.; Rangesh, A.; Trivedi, M.M. How would surround vehicles move? A unified framework for maneuver classification and motion prediction. *IEEE Trans. Intell. Veh.* **2018**, *3*, 129–140. [\[CrossRef\]](#)
11. Schreier, M.; Willert, V.; Adamy, J. Bayesian, maneuver-based, long-term trajectory prediction and criticality assessment for driver assistance systems. In Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), Qingdao, China, 8–11 October 2014; pp. 334–341.
12. Houenou, A.; Bonnifait, P.; Cherfaoui, V.; Yao, W. Vehicle trajectory prediction based on motion model and maneuver recognition. In Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, 3–7 November 2013; pp. 4363–4369.
13. Laugier, C.; Paromtchik, I.E.; Perrollaz, M.; Yong, M.; Yoder, J.-D.; Tay, C.; Mekhnacha, K.; Negre, A. Probabilistic analysis of dynamic scenes and collision risks assessment to improve driving safety. *IEEE Intell. Transp. Syst. Mag.* **2011**, *3*, 4–19. [\[CrossRef\]](#)
14. Schlechtriemen, J.; Wirthmueller, F.; Wedel, A.; Breuel, G.; Kuhnert, K.-D. When will it change the lane? a probabilistic regression approach for rarely occurring events. In Proceedings of the 2015 IEEE Intelligent Vehicles Symposium (IV), Seoul, Republic of Korea, 28 June–1 July 2015; pp. 1373–1379.
15. Xing, Y.; Lv, C.; Cao, D. Personalized vehicle trajectory prediction based on joint time-series modeling for connected vehicles. *IEEE Trans. Veh. Technol.* **2019**, *69*, 1341–1352. [\[CrossRef\]](#)
16. Messaoud, K.; Yahiaoui, I.; Verroust-Blondet, A.; Nashashibi, F. Non-local social pooling for vehicle trajectory prediction. In Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 9–12 June 2019; pp. 975–980.
17. Messaoud, K.; Yahiaoui, I.; Verroust-Blondet, A.; Nashashibi, F. Attention based vehicle trajectory prediction. *IEEE Trans. Intell. Veh.* **2021**, *6*, 175–185. [\[CrossRef\]](#)
18. Altche, F.; Fortelle, A.D.L. An lstm network for highway trajectory prediction. In Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017; pp. 353–359.
19. Zyner, A.; Worrall, S.; Ward, J.; Nebot, E. Long short-term memory for driver intent prediction. In Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV), Los Angeles, CA, USA, 11–14 June 2017; pp. 1484–1489.
20. Xin, L.; Wang, P.; Chan, C.Y.; Chen, J.; Li, S.E.; Cheng, B. Intention-aware long horizon trajectory prediction of ambient vehicles using dual lstm networks. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 1441–1446.
21. Deo, N.; Trivedi, M.M. Convolutional social pooling for vehicle trajectory prediction. In Proceedings of the the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 18–22 June 2018; pp. 1468–1476.
22. Alahi, A.; Goel, K.; Ramanathan, V.; Robicquet, A.; Fei-Fei, L.; Savarese, S. Social lstm: Human trajectory prediction in crowded spaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 961–971.
23. Liu, X.; Wang, Y.; Jiang, K.; Zhou, Z.; Nam, K.; Yin, C. Interactive trajectory prediction using a driving risk map-integrated deep learning method for ambient vehicles on highways. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 19076–19087. [\[CrossRef\]](#)
24. Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. *arXiv* **2018**, arXiv:1812.08434. [\[CrossRef\]](#)
25. Li, X.; Ying, X.; Chuah, M.C. GRIP++: Enhanced graph-based interaction-aware trajectory prediction for autonomous driving. *arXiv* **2020**, arXiv:1907.07792.
26. Si, C.; Chen, W.; Wang, W.; Wang, L.; Tan, T. An attention enhanced graph convolutional lstm network for skeleton-based action recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 1227–1236.
27. Yan, S.; Xiong, Y.; Lin, D. Spatial-temporal graph convolutional networks for skeleton-based action recognition. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, Hilton, NV, USA, 2–7 February 2018.
28. Huang, Y.; Bi, H.; Li, Z.; Mao, T.; Wang, Z. STGAT: Modeling spatial-temporal interactions for human trajectory prediction. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1550–1559.

29. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
30. Paszke, A.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Lerer, A. Automatic differentiation in pytorch. In Proceedings of the NIPS 2017 Autodiff Workshop: The Future of Gradient-based Machine Learning Software and Techniques, Long Beach, CA, USA, 9 December 2017.
31. Colyar, J.; Halkias, J. *US Highway 101 Dataset*; FHWA-HRT-07-030; Federal Highway Administration (FHWA): Washington, DC, USA, 2007.
32. Colyar, J.; Halkias, J. *US Highway i-80 Dataset*; FHWA-HRT-06-137; Federal Highway Administration (FHWA): Washington, DC, USA, 2006.
33. Khakzar, M.; Rakotonirainy, A.; Bond, A.; Dehkordi, S. A dual learning model for vehicle trajectory prediction. *IEEE Access* **2020**, *8*, 21897–21908. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.