

Article

A General Framework for Reconstructing Full-Sample Continuous Vehicle Trajectories Using Roadside Sensing Data

Guimin Su ^{1,2,3}, Zimu Zeng ¹, Andi Song ¹, Cong Zhao ^{1,*} , Feng Shen ², Liangxiao Yuan ² and Xinghua Li ¹

¹ Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China

² Shanghai SEARI Intelligent System Co., Ltd., Shanghai 200063, China

³ Shanghai SH Intelligent Automotive Technology Co., Ltd., Shanghai 201805, China

* Correspondence: zhc@tongji.edu.cn

Abstract: Vehicle trajectory data play an important role in autonomous driving and intelligent traffic control. With the widespread deployment of roadside sensors, such as cameras and millimeter-wave radar, it is possible to obtain full-sample vehicle trajectories for a whole area. This paper proposes a general framework for reconstructing continuous vehicle trajectories using roadside visual sensing data. The framework includes three modules: single-region vehicle trajectory extraction, multi-camera cross-region vehicle trajectory splicing, and missing trajectory completion. Firstly, the vehicle trajectory is extracted from each video by YOLOv5 and DeepSORT multi-target tracking algorithms. The vehicle trajectories in different videos are then spliced by the vehicle re-identification algorithm fused with lane features. Finally, the bidirectional long-short-time memory model (LSTM) based on graph attention is applied to complete the missing trajectory to obtain the continuous vehicle trajectory. Measured data from Donghai Bridge in Shanghai are applied to verify the feasibility and effectiveness of the framework. The results indicate that the vehicle re-identification algorithm with the lane features outperforms the vehicle re-identification algorithm that only considers the visual feature by 1.5% in mAP (mean average precision). Additionally, the bidirectional LSTM based on graph attention performs better than the model that does not consider the interaction between vehicles. The experiment demonstrates that our framework can effectively reconstruct the continuous vehicle trajectories on the expressway.

Keywords: vehicle trajectory reconstruction; re-identification; missing trajectory completion; bidirectional long-short-time memory model; graph attention mechanism



Citation: Su, G.; Zeng, Z.; Song, A.; Zhao, C.; Shen, F.; Yuan, L.; Li, X. A

General Framework for Reconstructing Full-Sample Continuous Vehicle Trajectories Using Roadside Sensing Data. *Appl. Sci.* **2023**, *13*, 3141. <https://doi.org/10.3390/app13053141>

Academic Editor: Yosoon Choi

Received: 6 February 2023

Revised: 23 February 2023

Accepted: 26 February 2023

Published: 28 February 2023



Copyright: © 2023 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

In recent years, autonomous driving has become an important topic in research, and many studies exist on positioning [1], perception [2], decision-making [3,4], motion planning [5,6], and control [7,8]. Among these, decision-making in autonomous driving is closely related to the operating conditions of other vehicles on the road, and road facilities can capture these operating conditions (including position, speed, pavement condition, etc.). Therefore, the coordination and cooperation of road facilities, autonomous driving, and traffic control have become important research topics [9,10].

An important topic in vehicle–infrastructure cooperation is the extraction of vehicle trajectories. Continuous vehicle trajectory information can play an important role in autonomous driving and intelligent traffic control [11]. It can provide the trajectory position of other vehicles for autonomous vehicles, which could help autonomous vehicles judge the current and future trajectories of surrounding vehicles and determine the next action. Meanwhile, continuous vehicle trajectory data can assist intelligent traffic control, risk assessment, and early warnings and help traffic personnel analyze road operation conditions [12].

Currently, extracting continuous vehicle trajectory information is challenging [13]. Video sensors are used in most regions to obtain motion information, such as the location and speed of road traffic participants. However, the video sensor data used today are strongly affected by occlusion and perspective [14]; some have low positioning accuracy, and the detection results are strongly affected by the environment. Additionally, common sensor data problems exist, such as different IDs marking in different sensors, which prevents direct identification of the same vehicle by different sensors. Therefore, it is difficult to directly obtain continuous vehicle trajectories from the video data.

Many studies have aimed to extract continuous vehicle trajectories from dynamic and static sensors [15,16]. This task has several directions in research. Some studies concern extracting vehicle trajectories from single roadside visual sensing data, whereas others concern finding vehicles with the same ID in different sensor data. Some studies also concern predicting and completing vehicle trajectories. Although these studies solve the corresponding problems effectively, few have proposed an entire continuous vehicle trajectory reconstruction framework, and some did not consider the particularity of the traffic scene. When special problems occur (such as multiple similar vehicles appearing on the road simultaneously), the model cannot provide high accuracy. Therefore, it is essential to study the actual traffic scene and propose the corresponding vehicle trajectory reconstruction framework.

In this research, we choose the expressway as the actual traffic scene and use the visual sensor data to construct the continuous vehicle trajectory. Constructing the vehicle trajectory on the expressway involves many problems. Firstly, the distance between the sensors on the expressway is long, resulting in gaps in the detected vehicle trajectory. Secondly, similar vehicles sometimes appear simultaneously on the expressway, making it difficult to distinguish them. To solve the above problems, we propose the corresponding reconstruction framework. The framework comprises three parts: single-region video trajectory extraction, multi-camera cross-region trajectory splicing, and vehicle trajectory completion. The framework first obtains a vehicle trajectory in a single region through video data. It then uses vehicle re-identification technology to match vehicle trajectories between different regions and splices the trajectories belonging to the same vehicle. After trajectory splicing, the framework uses a trajectory completion model based on graph attention to complete the missing parts of the vehicle trajectories.

The main contributions of this study are summarized as follows:

- We consider the spatio-temporal characteristics of the vehicle in the original vision-based vehicle re-identification model. We fuse the model with lane features to help it distinguish similar vehicles that appear simultaneously.
- We applied the bidirectional LSTM network based on graph attention. This model considers the interaction between vehicles and uses bidirectional trajectory completion. Our model can solve the problem of trajectory completion better than other baseline models.
- We propose a framework for continuous vehicle trajectory reconstruction based on video data, which uses the original video data to obtain the vehicle trajectory of a single region, employs the vehicle re-identification technology to splice the continuous vehicle trajectory, and uses the bidirectional long-short-time memory model (LSTM) based on graph attention to complete the incomplete trajectory.

The remainder of this paper is organized as follows. Section 2 reviews the previous research on continuous vehicle trajectory reconstruction. Section 3 details the continuous vehicle trajectory reconstruction framework. Section 4 describes and discusses the application effect of the framework in the actual case of the Donghai Bridge. Finally, Section 5 provides the conclusions and future prospects of the framework.

2. Related Work

This section provides an overview of the complete assessment process for the continuous vehicle trajectory reconstruction framework, including single-region video trajectory extraction, multi-camera cross-region trajectory splicing, and vehicle trajectory completion.

2.1. Single-Region Video Trajectory Extraction

The sources of vehicle trajectory data in a single region are diverse [17–19], and visual sensor data is the most widely used. Currently, various vehicle trajectory tracking methods are available for visual sensors [20–25]. Most of the methods are multi-target vehicle trajectory tracking using a matching algorithm for the detected objects [26,27], such as the DeepSORT algorithm [28]. Another multi-target tracking method is to use the neural network to analyze the motion information of the object to solve the association problem [29–31], such as the TrackFormer network based on the tracking-by-attention paradigm [32], which applies an attention mechanism to data association and tracking. It avoids dependence on information such as appearance features and realizes implicit trajectory association. There are also some multi-target tracking methods based on detection and tracking. For example, Viktor Kocur proposed a multi-target tracking algorithm based on CenterTrack object detection and neural network tracking [33]. However, fluctuations and gaps exist in the trajectories due to the accuracy of the video sensors and the error of the algorithm [14]. It is necessary to improve the continuity of the trajectories on the basis of obtaining accurate trajectories or using higher-precision sensor data.

2.2. Multi-Camera Cross-Region Trajectory Splicing

Currently, the main method of vehicle trajectory splicing is finding the trajectories of the same vehicle in different videos and then splicing the trajectories. For vehicle re-identification, the basic method is based on sensors, such as using wireless magnetic sensors to determine vehicle identity [34]. With the development of emerging sensors and technologies, such as identification tags and GPS (global positioning system), vehicles can be re-identified directly. For example, Prinsloo proposed a vehicle re-identification algorithm based on radio-frequency identification tags, which is suitable for various toll stations [35].

Considering the low penetration of emerging sensors and the large-scale use of visual sensors on the road, many studies have proposed methods to realize vehicle re-identification from videos. The early method is based on the method of artificial design features [36,37]. Researchers directly extract specific vehicle features from vehicle images to calculate the similarity to re-identify the vehicle ID (specific vehicle features: vehicle texture, vehicle color distribution, etc.), such as the BOW-SIFT method based on texture features and the BOW-CN method based on color features proposed by Liu et al. [38]. After CNN (convolution neural network) was suggested, scholars proposed using deep learning to obtain multidimensional information about vehicles to measure the similarity between vehicle images [39–43]. Simultaneously, considering the spatial and temporal information of vehicles, scholars have proposed methods such as screening time and geographical correlations to increase the accuracy of vehicle re-identification [44,45]. Although the current models have high accuracy, they struggle to solve some problems in practical applications, such as the appearance of multiple similar vehicles simultaneously. Therefore, in the framework of vehicle trajectory reconstruction, the vehicle re-identification model needs improvements to overcome these practical problems.

2.3. Vehicle Trajectory Completion

The completion of missing trajectories can be generalized as a trajectory prediction problem. In the early studies, trajectory prediction did not consider the interaction between individuals and only used historical data as the basis to predict the future trajectory [46,47]. In recent studies, advanced deep learning methods used networks such as RNNs and graph attention and considered the interactions between individuals in the model, combined

with individual historical trajectory data, to predict future trajectories [48,49]. Zhao et al. proposed a vehicle trajectory prediction model based on graph attention [50]. The graph attention network simulates the interaction between vehicles and improves the authenticity and accuracy of the predicted trajectory. Experiments show that the model based on graph attention can effectively simulate the interaction between vehicles, and the model has a better effect on the task of vehicle trajectory prediction. In addition, geofencing has been shown to be a reliable method for trajectory prediction of vehicles (ground or aerial). Using the new kinematic model, combined with the geo-fence algorithm, the trajectory of vehicles in the area can be predicted [51–53]. There are also studies based on traffic flow and individual complete trajectories to complete the overall traffic flow trajectories. Chen proposed a macro–micro integration framework for completing and reconstructing vehicle trajectories on the highway in a multi-source data environment [54]. The framework completes the vehicle trajectories using fixed-position sensor data to collect the position and speed of vehicles at a certain location and a small amount of probe vehicle (PV) data that generates continuous trajectories. Considering the feasibility of practical applications, we hope to improve the accuracy of the vehicle trajectory prediction model based on graph attention in our framework.

3. Methodology

This section introduces our framework for continuous vehicle trajectory reconstruction based on visual sensors, as shown in Figure 1. The continuous vehicle trajectory reconstruction framework mainly comprises three modules: single-region video trajectory extraction, multi-camera cross-region trajectory splicing, and vehicle trajectory completion.

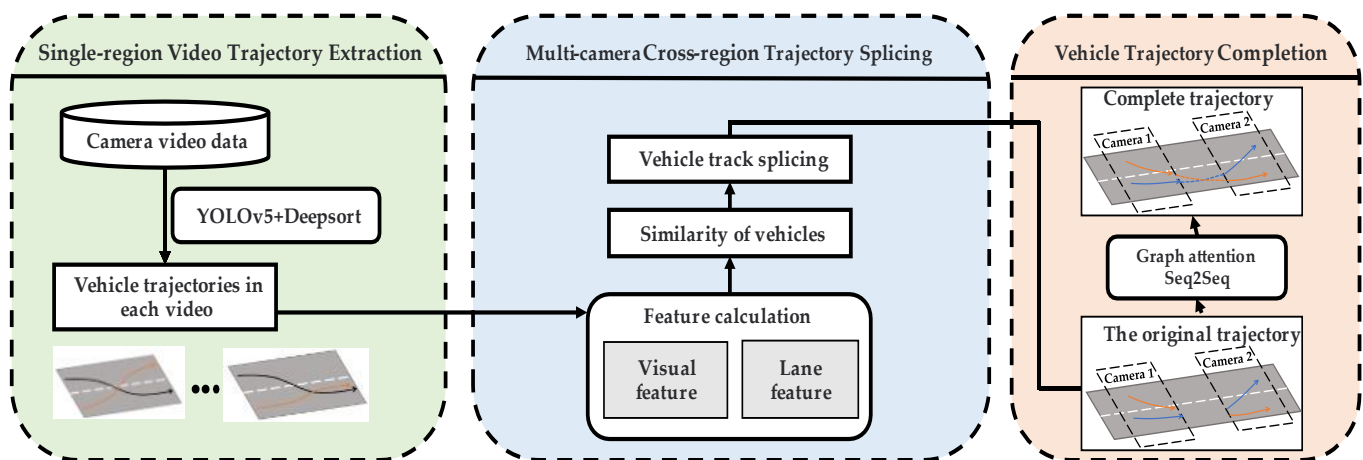


Figure 1. The framework for continuous vehicle trajectory reconstruction based on video data.

For video data, the framework uses a single-region video track extraction module to obtain vehicle trajectory data for each visual sensor. It then uses the multi-camera cross-region trajectory splicing module to match the vehicle trajectories of different sensors to achieve trajectory splicing within the entire region. We import the missing vehicle trajectory data into the bidirectional LSTM model based on graph attention for trajectory completion and finally reconstruct the continuous vehicle trajectory. The following section introduces each module in detail.

3.1. Single-Region Video Trajectory Extraction

Many challenges are involved in extracting vehicle trajectories from visual sensors. The videos of visual sensors used in this experiment have problems such as insufficient video pixels, video jitter, and mutual occlusion between vehicles. To obtain more accurate single-region video trajectory data, we adjusted the videos for deshaking and color contrast to stabilize their vehicle trajectory data. Aiming for the occlusion between vehicles, we

used the DeepSORT multi-target tracking model in this module [28]. When the vehicle is temporarily occluded in the video, it can still be recognized and tracked by the model. Figure 2 shows the model structure.

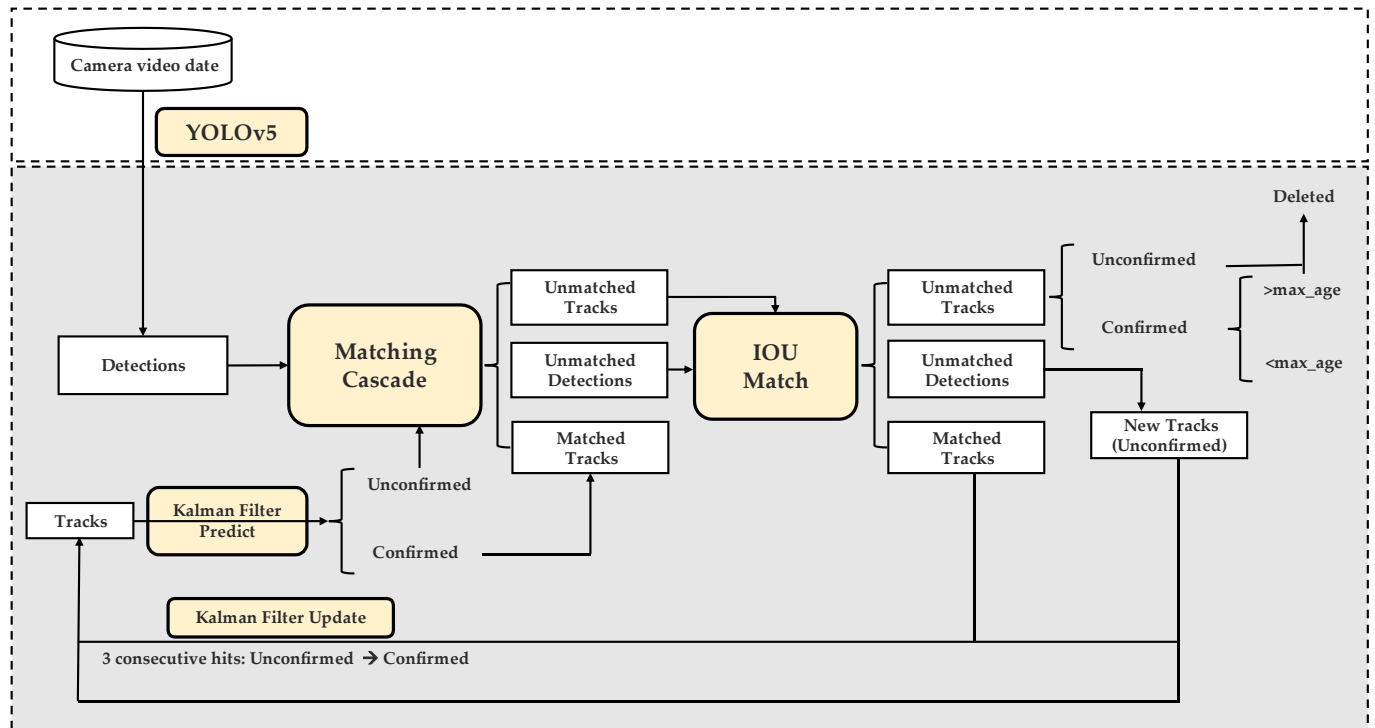


Figure 2. The structure of the single-region video trajectory extraction module.

Firstly, the model uses the YOLOv5 algorithm to extract the vehicle position information from the current frame in the video; it then uses Kalman filtering to predict the position of each vehicle in the next frame. The model then reuses YOLOv5 to extract the vehicle position information from the next frame. The model then uses the Hungarian algorithm to perform positional and visual data association of detected and predicted vehicle positions, determining the matching relationship between vehicles in the preceding and following video frames. After the matching is completed, the model updates the tracked vehicle position box. The model continuously loops to track the vehicle trajectory in the video.

The vehicle trajectory obtained from the video is in the pixel coordinate system, but the required trajectory coordinates should be in the world coordinate system. It is necessary to convert the coordinate system of the vehicle trajectory. The conversion of the pixel coordinate system to the world coordinate system usually conforms to the following Equation (1):

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{Z} H \begin{pmatrix} U \\ V \\ 1 \end{pmatrix} = \frac{1}{Z} \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{pmatrix} U \\ V \\ 1 \end{pmatrix} \quad (1)$$

$$H = A(R_1 \ R_2 \ T)$$

In Equation (1), H is the homography matrix, which is obtained by multiplying the internal and external parameters of the camera. Here, A represents the intrinsic parameters of the camera, R_1 represents the first column of the rotation matrix, R_2 represents the second column of the rotation matrix, and T represents the translation vector. Z represents the scale factor, u and v represent pixel coordinates, and U and V represent the corresponding world coordinates.

According to the design standard, we can obtain a priori information on the relative positions of the lane lines in space (in China, the longitudinal length of lane lines along the lane direction is 6 m, the longitudinal distance between lane lines along the lane direction

is 9 m, and the lane width is 3–4 m.) [55]. We can select four points located on the lane line from the video image. The distance between these four points in the world coordinate system should satisfy the prior information on the lane lines. We choose one of the points as the origin and use the prior information to obtain the coordinates of the other three points in the world coordinate system. H and Z can be calculated by inserting the pixel coordinate system coordinates and world coordinate system coordinates of these points into the formula. The transformation matrix can then be used to convert the pixel coordinates to world coordinates for the video. Figure 3 shows the matrix calibration process.

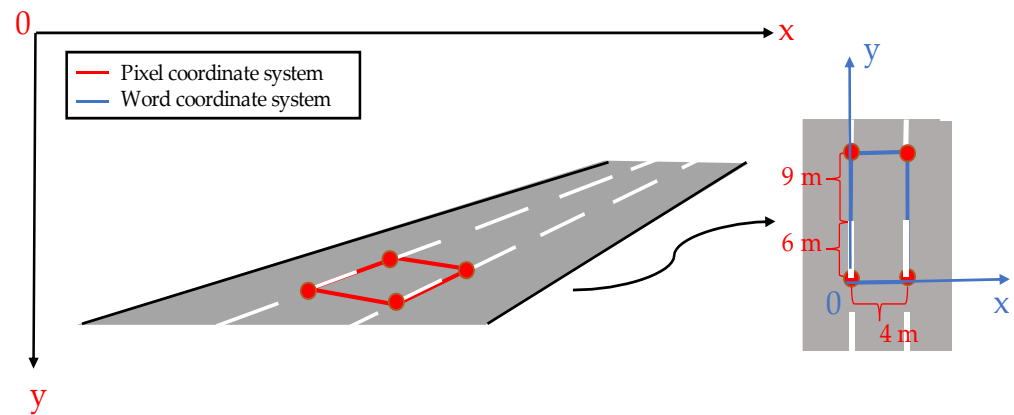


Figure 3. Calibration of the coordinate system transformation matrix.

3.2. Multi-Camera Cross-Region Trajectory Splicing

Multi-camera cross-region trajectory splicing is an important step in continuous vehicle trajectory reconstruction. Since most of the visual sensors deployed on the road network today cannot directly capture the ID information of the vehicle, it is difficult to determine the matching relationships between the vehicle trajectories extracted from different visual sensors. An additional module is required to match the trajectories.

Considering the above issues, a multi-camera cross-region trajectory splicing module is constructed in the framework. In this module, considering that similar trucks may appear on the road simultaneously, reducing the accuracy of trajectory splicing, we built a new model. The model can extract the vehicle's visual and lane features from the videos and calculate the feature similarity of vehicles in different videos to find the same vehicle and splice the trajectory.

As shown in Figure 4, the visual feature extraction part of the network uses the Reid network proposed by Zheng et al. [56]. The model uses the latest network pre-trained on ImageNet as the backbone module, including ResNeXt101, ResNeXt101_32x8l_wsl, and ResNet50_IBN_a. The lane feature extraction part of the network uses a fully connected layer to convert the number vector of the lane where the vehicle is located into a feature vector with the same length as the vehicle's visual feature. We then use the fusion feature of the visual feature and lane feature to calculate the similarity between vehicles.

We fuse the vehicle's visual features and lane features according to the following Equation (2):

$$F_{fusion} = F_{vision} + k * F_{lane} \quad (2)$$

In Equation (2), F_{fusion} is the fusion feature of the vehicle, F_{vision} is the visual feature of the vehicle, and F_{lane} is the lane feature of the vehicle; k is the scale factor, which is obtained by model training.

We use the fused features to calculate the cosine similarity between different vehicles and find the vehicle-matching result with the best similarity matching.

$$dist(A, B) = 1 - \cos(A, B) = \frac{\|A\|_2 * \|B\|_2 - AB}{\|A\|_2 * \|B\|_2} \quad (3)$$

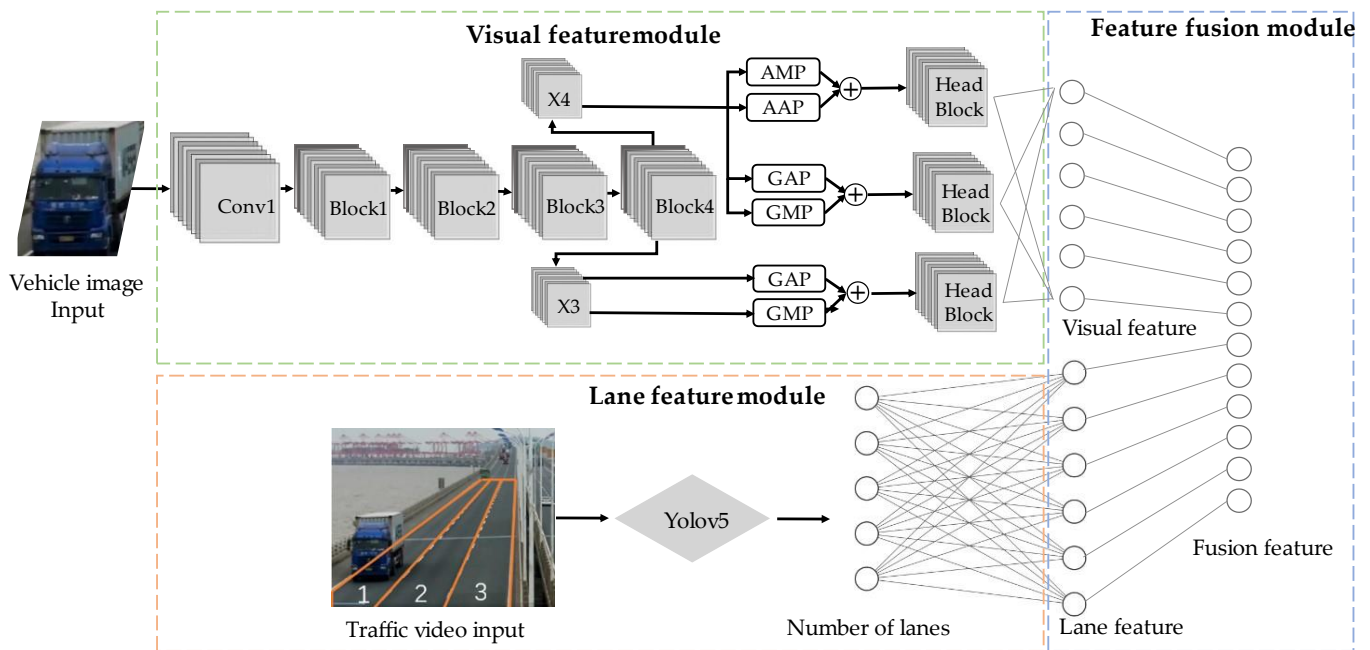


Figure 4. The structure of the vehicle re-identification model.

In Equation (3), A and B are the fusion features of two vehicles, and $dist(A, B)$ is the feature similarity of the two vehicles.

After obtaining the vehicle-matching results, we splice the trajectories belonging to the same vehicle.

3.3. Vehicle Trajectory Completion

Due to the insufficient coverage region of the video sensor, gaps exist in the obtained continuous vehicle trajectories, so the missing trajectories must be completed. In this module, we refer to the STGAT model proposed by Huang et al. [57]. Considering the mutual interaction of vehicles, we add the graph attention mechanism to the Seq2Seq architecture. This model combines the temporal features and the spatial interaction features between vehicles and improves the accuracy of vehicle trajectory completion. Figure 5 shows the bidirectional LSTM vehicle trajectory completion model based on graph attention.

We propose a double-stream mechanism. The two trajectories before and after the missing trajectory are input into the encoder module to obtain the temporal features. Then the combination of temporal features is input into the decoder module to obtain the final completed trajectory. In our model, two completed vehicle trajectories are generated from the two ends of the missing trajectory towards the middle, respectively. When the Euclidean distance between the two trajectories is less than $Distance_{low}$ or the maximum number of completion frames is reached (set to $Frame_{max}$), the completion is stopped.

The encoding module consists of an LSTM module and a GAT module. Considering the vehicle's temporal characteristics, we use the LSTM module in the encoder to extract the timing characteristics of the trajectory sequence, focusing on the motion mode of each vehicle. The preliminary encoding results are shown as Equation (4):

$$m_i^t = LSTM(m_i^{t-1}, e_i^t; W_m) \quad (4)$$

In Equation (4), m_i^t is the eigenvector in the network corresponding to vehicle i at time t , e_i^t is the embedded coding result of the position coordinate of vehicle i at time t , and W_m represents the weight of the model.

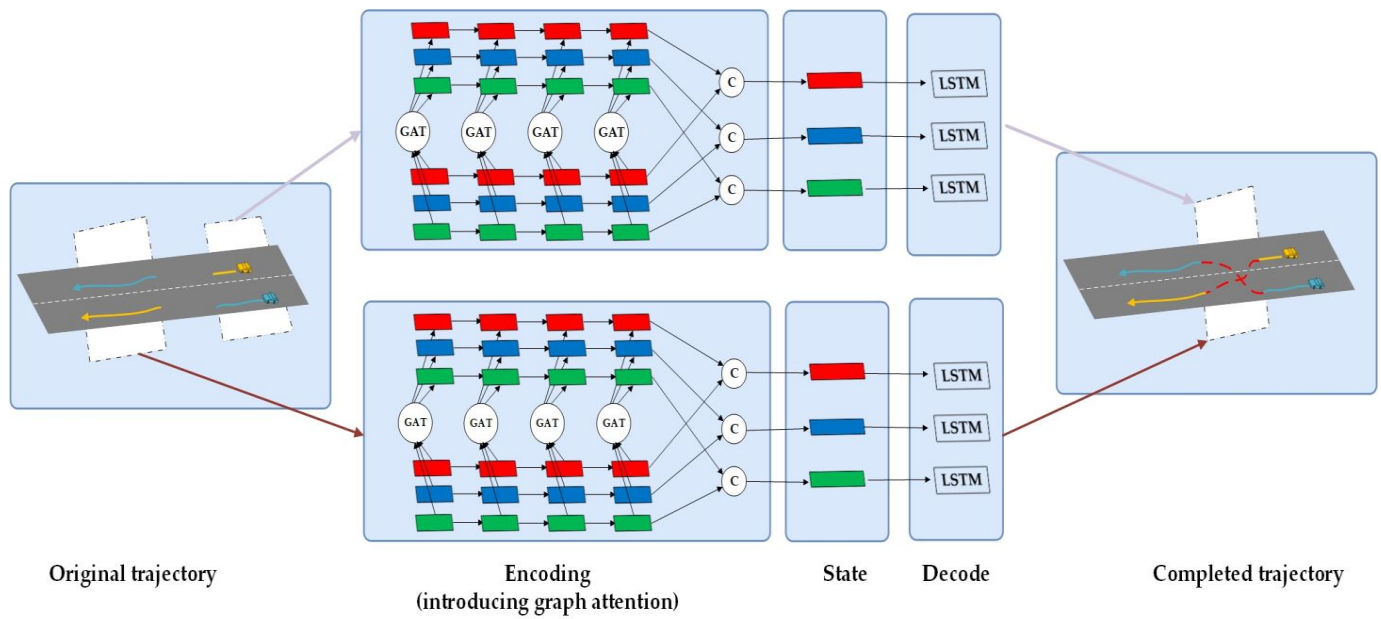


Figure 5. The neural network structure of the bidirectional LSTM vehicle trajectory completion model based on graph attention.

For vehicle interaction characteristics, this model uses the graph attention model in the encoder. The model takes all vehicles as nodes and assigns importance to aggregate neighbor vehicle information. The preliminary temporal encoding results are input into the graph attention model, and the output is

$$\hat{m}_i^t = \sigma(\sum_{j \in N_i} GAT(m_i^t, m_j^t) W_j m_j^t) \quad (5)$$

$$GAT(m_i^t, m_j^t) = \frac{\exp(\text{LeakyReLU}(a^T [W_i m_i || W_j m_j]))}{\sum_{k \in N_i} \exp(\text{LeakyReLU}(a^T [W_i m_i || W_k m_k]))} \quad (6)$$

In Equations (5) and (6), $GAT(m_i^t, m_j^t)$ is the graph attention coefficient at time t of node pair (i, j) , W is the dimension mapping matrix, N_i is the set of neighbor nodes of node i , and σ is a nonlinear activation function.

The decoder module is proposed to aggregate temporal and spatial features and then generate the final completed trajectory. At the time point t of node i , these characteristics are combined as the input of the LSTM module: the initial coding result $\bar{m}_i = \delta_1(m_i^t)$ and the output $\bar{g}_i = \delta_2(g_i^t)$ after the graph attention and LSTM coding. Based on this spatiotemporal information, the missing trajectory can be completed.

4. Experiments

4.1. Scene Description and Data Preprocessing

1. Donghai Bridge dataset [58]: The Donghai Bridge (Shanghai, China) is 32.5 km long and equipped with more than 200 cameras to support the commercial operation of self-driving trucks in Yangshan Port, as shown in Figure 6. This dataset contains video data during peak and off-peak hours throughout the day. In this experiment, we obtained data from two visual sensors on the Donghai Bridge to realize vehicle trajectory reconstruction. The distance between the visual sensors is 1020 m. The videos were collected during the peak hours of the day and lasted for half an hour, and the traffic flow in the videos was restricted. We have adjusted the debouncing and color contrast to make the videos more suitable for multi-target tracking. We captured 101 sets of vehicle images from two videos to create a dataset for the vehicle re-identification

experiment. We also obtained complete short-distance trajectory data from the videos and created a dataset to evaluate the vehicle trajectory completion model.

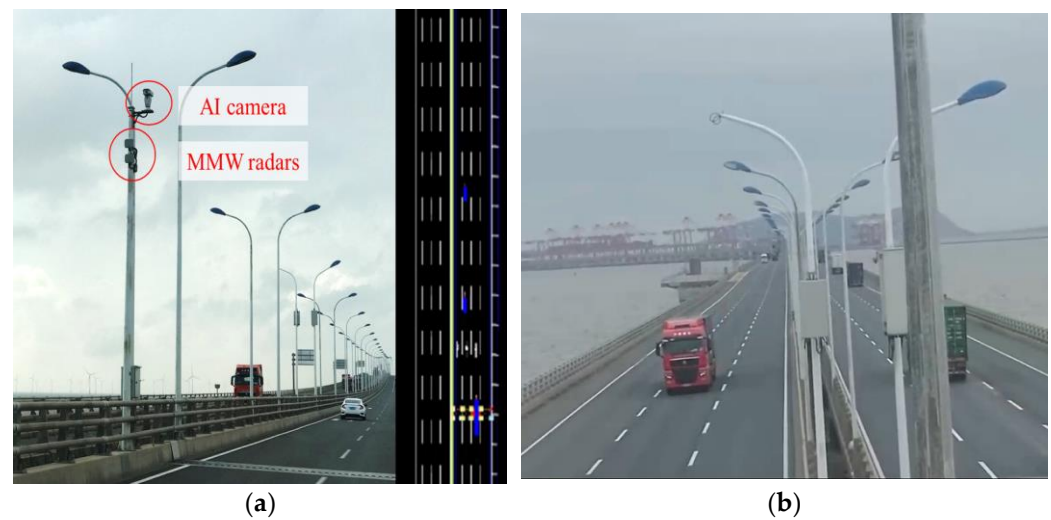


Figure 6. A real scenario of roadside perception on Donghai Bridge, Shanghai, China: (a) roadside MMW radar and AI cameras; (b) Donghai Bridge video data.

2. CityFlow dataset [59]: To train the vehicle re-identification model, we used the CityFlow dataset. The dataset is captured by 40 cameras in the traffic surveillance environment. It contains 56,277 bounding boxes in total. The training set comprises 36,935 bounding boxes from the 333 vehicles, and the test set comprises 18,290 bounding boxes from the other 333 identities. The remaining 1052 images are the queries. The dataset contains a variety of vehicles and can be used as training data for the vehicle re-identification model.
3. The Argoverse bidirectional trajectory completion dataset: To train and evaluate the trajectory completion model, we modified the Argoverse motion forecasting dataset [60] and named it the Argoverse bidirectional trajectory completion dataset. The Argoverse motion forecasting dataset can be used to support vehicle trajectory forecasting tasks, providing rich trajectories and lane centerlines collected by a fleet of AVs in Pittsburgh and Miami. The dataset contains 333 K 5 s long sequences. We split it into 200 K for training, 66 K for validation, and 67 for testing. The vehicle trajectories of the Argoverse bidirectional trajectory completion dataset are sampled at 10 Hz. We selected the (0, 2] s and the (4, 5] s as the ground truth and the (2, 4] s for prediction.

4.2. Description of Evaluation Metrics and Model Parameters

In this experiment, we used the following model evaluation metrics:

- mAP: mean average precision. The average precision is calculated as the weighted mean of precisions at each threshold; the weight is the increase in recall from the prior threshold. The mAP is calculated by finding the average precision (AP) for each class and then averaging over several classes.

$$mAP = \frac{1}{N} \left(\sum_{i=1}^N AP_i \right) \quad (7)$$

- CMC: cumulative matching characteristics. This is a classic evaluation index in the re-identification problem. The abscissa of the curve is the rank, and the ordinate is the recognition accuracy; $rank_n$ indicates the top n results in descending order of the similarity of the recognition results containing the target. The recognition accuracy is the ratio of the number of $rank_n$ to the total number of query samples.

- ADE: average displacement error. This is defined as the Euclidean distance (L_2 distance) between the ground truth and prediction of the vehicle locations over all predicted time steps:

$$L_2 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (8)$$

In Equation (8), (x_1, y_1) and (x_2, y_2) are the locations of the ground truth and prediction of the vehicle locations, respectively.

Table 1 shows some of our model parameter settings.

Table 1. Some model parameters.

Module	Parameter	Value
Single-region vehicle trajectory extraction	Confidence (YOLOv5)	0.8
Multi-camera cross-region trajectory splicing	k (vehicle re-identification)	0.05
Vehicle trajectory completion	$Distance_{low}$	0.05
	$Frame_{max}$	200

4.3. Single-Region Vehicle Trajectory Extraction

The video data is imported into the single-region vehicle trajectory extraction module in the framework, and the DeepSORT algorithm is used to obtain the trajectories of all vehicles in the video in the pixel coordinate system. Figure 7 shows the visualization of some trajectories.

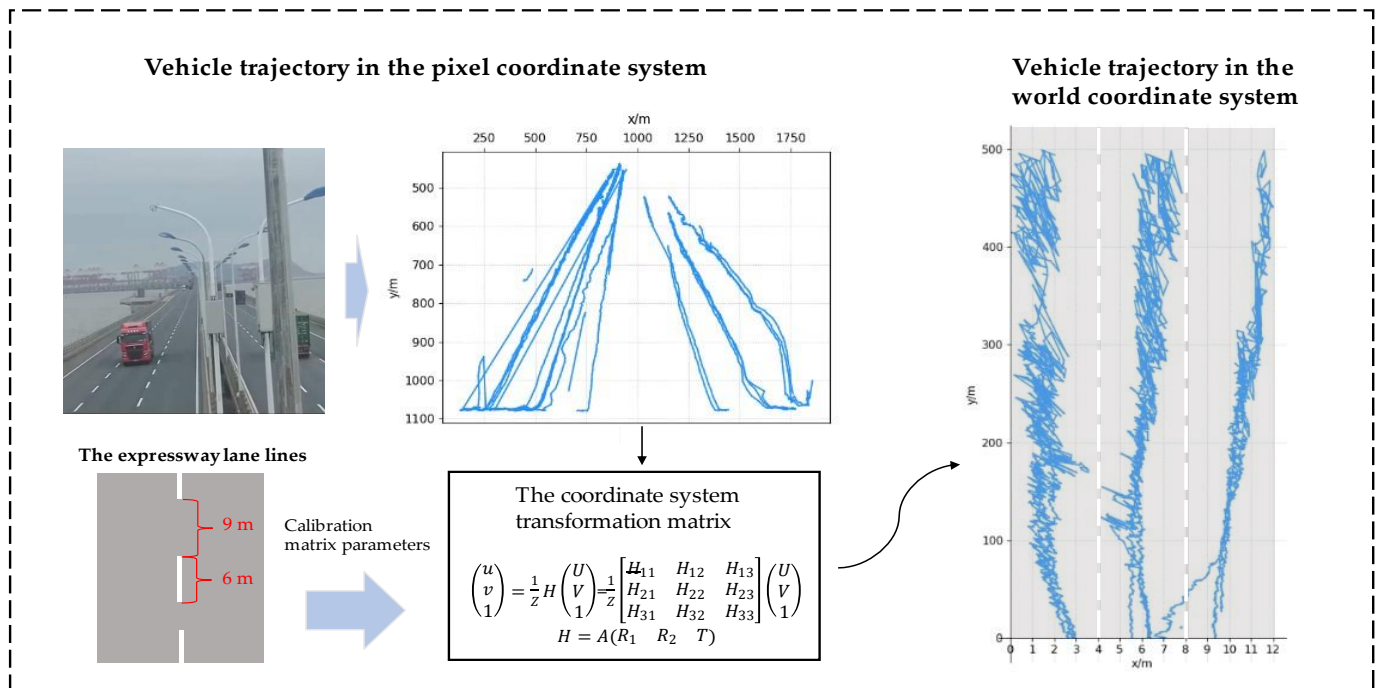


Figure 7. Single-region vehicle trajectory extraction in the world coordinate system.

4.4. Multi-Camera Cross-Region Trajectory Splicing

The multi-camera cross-region trajectory splicing module in the framework is used to assign the corresponding matching ID to the vehicle trajectories in each video on the road section.

First, the CityFlow dataset is used to train the vehicle REID network of this module to obtain the model of the vehicle visual feature vector. Part of the calibrated video trajectory data of the Donghai Bridge is then used to train the scale factor in the feature fusion formula.

After training, we used the module to match the vehicle trajectories in the two camera videos on the road section. Table 2 presents the results. The resulting mAP (mean average precision) is 73.46%, 1.5 percentage points better than the original vehicle re-identification model. This confirms that considering lane features can improve the accuracy of vehicle re-identification. Figure 8 shows the CMC (cumulative matching characteristics) diagram of the model.

Table 2. Evaluation indicators of our vehicle re-identification model fused with lane features.

Model	mAP (Mean Average Precision)
Our model	73.46%
Original vehicle re-identification model *	71.96%

* The model is taken from reference [56].

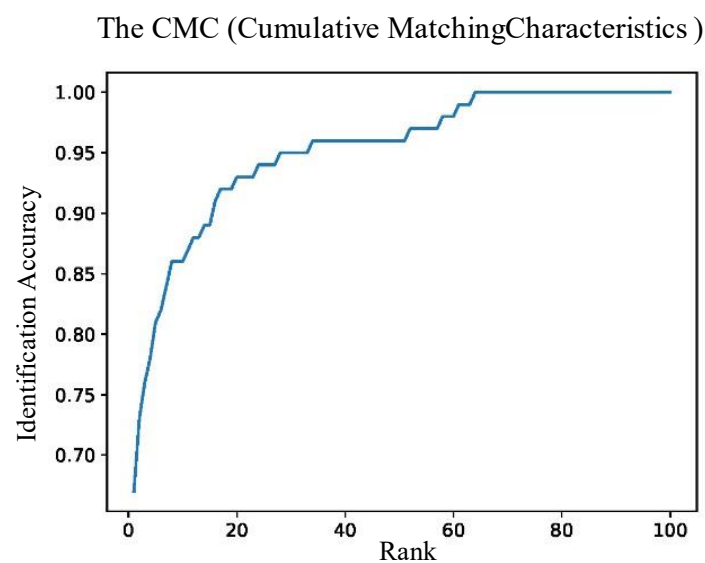


Figure 8. The CMC (cumulative matching characteristics) diagram of our vehicle re-identification model fused with lane features.

4.5. Missing Trajectory Completion

To illustrate the performance of the bidirectional LSTM vehicle trajectory reconstruction model based on graph attention in this module, we selected different baseline models and applied these to the Argoverse bidirectional trajectory completion dataset. We used ADE as the evaluating indicator. Table 3 presents the results.

Table 3. Imputation errors in the Argoverse bidirectional trajectory completion dataset.

Model	ADE (m)
Our model	0.7649
Vanilla Bi-LSTM	1.9246
CS-LSTM	1.8523
NAOMI	2.0335
STGAT	1.8139

In the comparison results, the model considering the interaction between vehicles (Vanilla Bi-LSTM) performs better than the model that does not consider the interaction (NAOMI). Meanwhile, the model using the graph to represent interactions (STGAT) outperforms other models (Vanilla Bi-LSTM, CS-LSTM). In this experiment, our bidirectional LSTM vehicle trajectory reconstruction model based on graph attention outperforms all mentioned models. This reveals that our model can effectively capture the spatio-temporal

characteristics of vehicle trajectories and the vehicle–vehicle interactions, and the bidirectional trajectory completion performs better than the unidirectional trajectory completion.

In this experiment, we applied our trained model to the vehicle trajectory completion on the Donghai Bridge. Figure 9 represents the effect of our model and some baseline models on vehicle trajectory completion on the Donghai Bridge. The result indicates that our model outperforms other models in the trajectory completion task on Donghai Bridge.

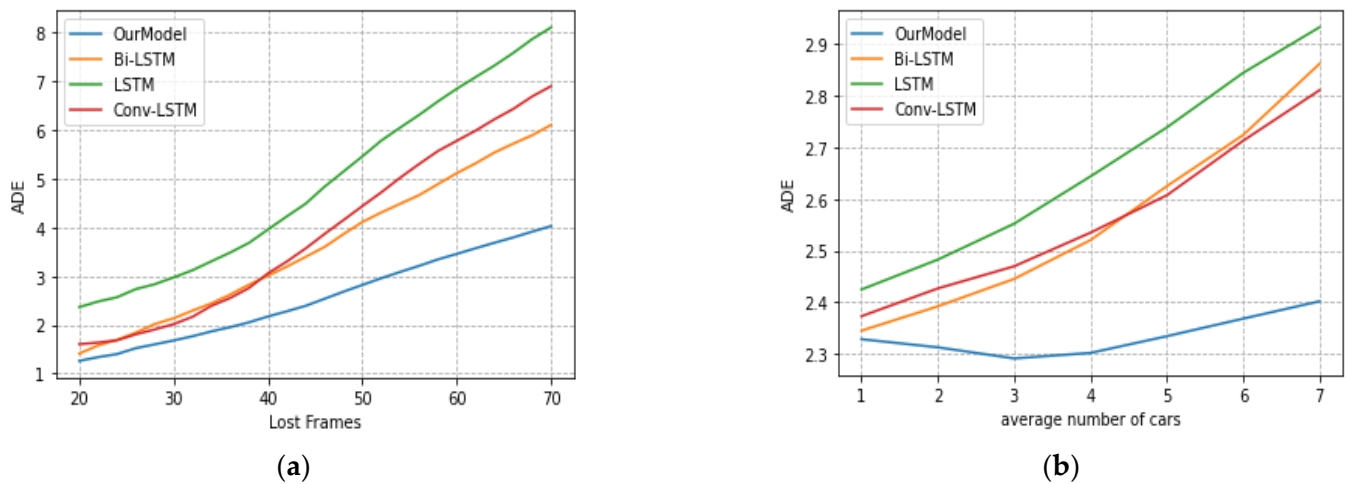


Figure 9. Evaluation of the vehicle trajectory completion effect of the models on the Donghai Bridge dataset: (a) evaluation of vehicle trajectory completion with different numbers of missing frames; (b) evaluation of vehicle trajectory completion with different numbers of missing vehicles.

5. Conclusions and Discussion

We propose a framework for full-sample vehicle trajectory reconstruction based on roadside sensing data. The framework includes multiple tasks, deep learning-based video data processing, multi-target tracking, vehicle re-identification, and missing trajectory completion. YOLOv5 and DeepSORT algorithms are applied to extract vehicle trajectories. A vehicle re-identification algorithm is then designed to splice the vehicle trajectories in different video data, where lane and visual features are fused to reduce the adverse effect on the accuracy caused by the simultaneous appearance of similar vehicles on the expressway. The LSTM model based on graph attention is then developed to complete the missing vehicle trajectory. Comprehensive experiments are conducted using roadside sensing data from the Donghai Bridge, Shanghai, China. The recognition accuracy mAP of our vehicle re-identification algorithm with lane features reaches 73.46%, 1.5 percentage points higher than the original visual-based vehicle re-identification algorithm. Meanwhile, the experimental results show that the bidirectional LSTM vehicle trajectory reconstruction model based on graph attention outperforms other models that do not consider interactions between vehicles. The experimental results show that our framework can effectively reconstruct continuous vehicle trajectories via roadside sensing data.

Future studies could follow several directions. Firstly, we could continue to improve the multi-camera cross-region tracking algorithm, use more advanced neural networks, or introduce traffic operation rules to improve the accuracy of vehicle re-identification. Secondly, the model of vehicle trajectory completion could be improved. We hope that the vehicle trajectory predicted by the model will follow the space–time law of vehicle operation and that there will be no wrongly predicted trajectory collisions caused by algorithm errors. Thirdly, since the quality of video data is strongly affected by natural factors, such as weather and light, more types of sensor data could be fused to improve the accuracy of the vehicle trajectories. Finally, when the framework is used in different traffic scenarios, by collecting new data sets and transfer learning, we can make the framework suitable for the traffic scenarios, such as urban areas and rural areas.

Author Contributions: Conceptualization, C.Z.; methodology, G.S., Z.Z. and C.Z.; software, Z.Z. and A.S.; validation, C.Z. and X.L.; formal analysis, C.Z.; investigation, Z.Z.; data curation, C.Z., F.S. and L.Y.; writing—original draft preparation, G.S. and Z.Z.; writing—review and editing, C.Z., A.S. and X.L.; visualization, Z.Z.; funding acquisition, C.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China under Grant 2022YFF0604900, and in part by the National Natural Science Foundation of China under Grant 52102383, and in part by the China Postdoctoral Science Foundation under Grant 2021M692428.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

1. Zhao, C.; Song, A.; Zhu, Y.; Jiang, S.; Liao, F.; Du, Y. Data-Driven Indoor Positioning Correction for Infrastructure-Enabled Autonomous Driving Systems: A Lifelong Framework. *IEEE Trans. Intell. Transp. Syst.* **2023**, 1–14.
2. Van Brummelen, J.; O'Brien, M.; Gruyer, D.; Najjaran, H. Autonomous vehicle perception: The technology of today and tomorrow. *Transp. Res. Part C-Emerg. Technol.* **2018**, *89*, 384–406.
3. Du, Y.; Chen, J.; Zhao, C.; Liao, F.; Zhu, M. A hierarchical framework for improving ride comfort of autonomous vehicles via deep reinforcement learning with external knowledge. *Comput.-Aided Civil Infrastruct. Eng.* **2022**, 1–20. [\[CrossRef\]](#)
4. Du, Y.; Chen, J.; Zhao, C.; Liu, C.; Liao, F.; Chan, C. Comfortable and energy-efficient speed control of autonomous vehicles on rough pavements using deep reinforcement learning. *Transp. Res. Part C-Emerg. Technol.* **2022**, *134*, 103489.
5. Zhao, C.; Zhu, Y.; Du, Y.; Liao, F.; Chan, C. A Novel Direct Trajectory Planning Approach Based on Generative Adversarial Networks and Rapidly-Exploring Random Tree. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 17910–17921.
6. Ji, Y.; Ni, L.; Zhao, C.; Lei, C.; Du, Y.; Wang, W. TriPField: A 3D Potential Field Model and Its Applications to Local Path Planning of Autonomous Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2023**, 1–14. [\[CrossRef\]](#)
7. Zhao, C.; Cao, J.; Zhang, X.; Du, Y. From search-for-parking to dispatch-for-parking in an era of connected and automated vehicles: A macroscopic approach. *J. Transp. Eng. Part A-Syst.* **2022**, *148*, 4021112.
8. Zhao, C.; Liao, F.; Li, X.; Du, Y. Macroscopic modeling and dynamic control of on-street cruising-for-parking of autonomous vehicles in a multi-region urban road network. *Transp. Res. Part C-Emerg. Technol.* **2021**, *128*, 103176.
9. Chen, J.; Zhao, C.; Jiang, S.; Zhang, X.; Li, Z.; Du, Y. Safe, efficient, and comfortable autonomous driving based on cooperative vehicle infrastructure system. *Int. J. Environ. Res. Public Health* **2023**, *20*, 893.
10. You, F.; Zhang, R.; Lie, G.; Wang, H.; Wen, H.; Xu, J. Trajectory planning and tracking control for autonomous lane change maneuver based on the cooperative vehicle infrastructure system. *Expert Syst. Appl.* **2015**, *42*, 5932–5946.
11. Li, L.; Jiang, R.; He, Z.; Chen, X.M.; Zhou, X. Trajectory data-based traffic flow studies: A revisit. *Transp. Res. Part C-Emerg. Technol.* **2020**, *114*, 225–240.
12. Zhao, C.; Ding, D.; Du, Z.; Shi, Y.; Su, G.; Yu, S. Analysis of Perception Accuracy of Roadside Millimeter-Wave Radar for Traffic Risk Assessment and Early Warning Systems. *Int. J. Environ. Res. Public Health* **2023**, *20*, 879. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Lei, C.; Zhao, C.; Ji, Y.; Shen, Y.; Du, Y. Identifying and correcting the errors of vehicle trajectories from roadside millimetre-wave radars. *IET Intell. Transp. Syst.* **2023**, *17*, 418–434. [\[CrossRef\]](#)
14. Du, Y.; Wang, F.; Zhao, C.; Zhu, Y.; Ji, Y. Quantifying the performance and optimizing the placement of roadside sensors for cooperative vehicle-infrastructure systems. *IET Intell. Transp. Syst.* **2022**, *16*, 908–925.
15. Xie, X.; van Lint, H.; Verbraeck, A. A generic data assimilation framework for vehicle trajectory reconstruction on signalized urban arterials using particle filters. *Transp. Res. Part C-Emerg. Technol.* **2018**, *92*, 364–391.
16. Wei, L.; Wang, Y.; Chen, P. A particle filter-based approach for vehicle trajectory reconstruction using sparse probe data. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 2878–2890. [\[CrossRef\]](#)
17. Chen, Y.; Wang, Y.; Qu, F.; Li, W. A graph-based track-before-detect algorithm for automotive radar target detection. *IEEE Sens. J.* **2020**, *21*, 6587–6599. [\[CrossRef\]](#)
18. Yavuz, F.; Kalfa, M. Radar target detection via deep learning. In Proceedings of the 2020 28th Signal Processing and Communications Applications Conference (SIU), Istanbul, Turkey, 5–7 October 2020.
19. Krajewski, R.; Bock, J.; Kloeker, L.; Eckstein, L. The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018.
20. Li, B.; Yan, J.; Wu, W.; Zhu, Z.; Hu, X. High performance visual tracking with siamese region proposal network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.

21. Nam, H.; Han, B. Learning multi-domain convolutional neural networks for visual tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016.
22. Li, F.; Tian, C.; Zuo, W.; Zhang, L.; Yang, M. Learning spatial-temporal regularized correlation filters for visual tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
23. Keuper, M.; Tang, S.; Andres, B.; Brox, T.; Schiele, B. Motion segmentation & multiple object tracking by correlation co-clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *42*, 140–153.
24. Tang, S.; Andriluka, M.; Andres, B.; Schiele, B. Multiple people tracking by lifted multicut and person re-identification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
25. Zhu, J.; Yang, H.; Liu, N.; Kim, M.; Zhang, W.; Yang, M. Online multi-object tracking with dual matching attention networks. In Proceedings of the European conference on computer vision (ECCV), Munich, Germany, 8–14 September 2018.
26. Chen, J.; Sheng, H.; Zhang, Y.; Xiong, Z. Enhancing Detection Model for Multiple Hypothesis Tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017.
27. Dehghan, A.; Tian, Y.; Torr, P.H.; Shah, M. Target identity-aware network flow for online multiple target tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 8–10 June 2015.
28. Wojke, N.; Bewley, A.; Paulus, D. Simple online and realtime tracking with a deep association metric. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017.
29. Milan, A.; Rezatofghi, S.H.; Dick, A.; Reid, I.; Schindler, K. Online multi-target tracking using recurrent neural networks. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
30. Zhou, X.; Koltun, V.; Krähenbühl, P. Tracking objects as points. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020.
31. Chu, P.; Wang, J.; You, Q.; Ling, H.; Liu, Z. Transmot: Spatial-temporal graph transformer for multiple object tracking. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 2–7 January 2023.
32. Meinhardt, T.; Kirillov, A.; Leal-Taixe, L.; Feichtenhofer, C. TrackFormer: Multi-Object Tracking with Transformers. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–24 June 2022; Cornell University Library: Ithaca, Greece, 2022; pp. 8844–8854.
33. Kocur, V.; Ftacnik, M. Multi-Class Multi-Movement Vehicle Counting Based on CenterTrack. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Virtual, 19–25 June 2021.
34. Sanchez, R.O.; Flores, C.; Horowitz, R.; Rajagopal, R.; Varaiya, P. Vehicle re-identification using wireless magnetic sensors: Algorithm revision, modifications and performance analysis, In Proceedings of 2011 IEEE International Conference on Vehicular Electronics and Safety, Beijing, China, 10–12 July 2011.
35. Prinsloo, J.; Malekian, R. Accurate Vehicle Location System Using RFID, an Internet of Things Approach. *Sensors* **2016**, *16*, 825.
36. Ferencz, A.; Learned-Miller, E.G.; Malik, J. Building a classification cascade for visual identification from one example. In Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05), Beijing, China, 17–21 October 2005; Volume 1.
37. Zapletal, D.; Herout, A. Vehicle re-identification for automatic video traffic surveillance. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Las Vegas, NV, USA, 26 June–1 July 2016.
38. Liu, X.; Liu, W.; Ma, H.; Fu, H. Large-scale vehicle re-identification in urban surveillance videos. In Proceedings of the 2016 IEEE International Conference on Multimedia and Expo (ICME), Seattle, WA, USA, 11–15 July 2016.
39. Jiang, M.; Zhang, X.; Yu, Y.; Bai, Z.; Zheng, Z.; Wang, Z.; Wang, J.; Tan, X.; Sun, H.; Ding, E.; et al. Robust Vehicle Re-identification via Rigid Structure Prior. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Virtual, 19–25 June 2021.
40. Li, M.; Huang, X.; Zhang, Z. Self-supervised Geometric Features Discovery via Interpretable Attention for Vehicle Re-Identification and Beyond. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual, 19–25 June 2021.
41. Zhang, Y.; Liu, D.; Zha, Z. Improving triplet-wise training of convolutional neural network for vehicle re-identification. In Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), 10–14 July 2021; Hong Kong, China.
42. Guo, H.; Zhao, C.; Liu, Z.; Wang, J.; Lu, H. Learning coarse-to-fine structured feature embedding for vehicle re-identification. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018.
43. Zhu, J.; Zeng, H.; Huang, J.; Liao, S.; Lei, Z.; Cai, C.; Zheng, L. Vehicle re-identification using quadruple directional deep learning features. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 410–420. [[CrossRef](#)]
44. Liu, X.; Liu, W.; Mei, T.; Ma, H. Provid: Progressive and multimodal vehicle reidentification for large-scale urban surveillance. *IEEE Trans. Multimed.* **2017**, *20*, 645–658. [[CrossRef](#)]
45. Shen, Y.; Xiao, T.; Li, H.; Yi, S.; Wang, X. Learning deep neural networks for vehicle re-id with visual-spatio-temporal path proposals. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017.
46. Althe, F.; de La Fortelle, A. 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.
47. Dash, M.; Koo, K.K.; Gomes, J.B.; Krishnaswamy, S.P.; Rugeles, D.; Shi-Nash, A. Next place prediction by understanding mobility patterns. In Proceedings of the 2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), St. Louis, MO, USA, 23–27 March 2015; pp. 469–474.

48. Zhao, C.; Song, A.; Du, Y.; Yang, B. TrajGAT: A map-embedded graph attention network for real-time vehicle trajectory imputation of roadside perception. *Transp. Res. Part C-Emerg. Technol.* **2022**, *142*, 103787.
49. Ip, A.; Irio, L.; Oliveira, R. Vehicle trajectory prediction based on LSTM recurrent neural networks. In Proceedings of the 2021 IEEE 93rd Vehicular Technology Conference (VTC2021), Helsinki, Finland, 25–28 April 2021; pp. 1–5.
50. Zhao, Z.; Fang, H.; Jin, Z.; Qiu, Q. Gisnet: Graph-based information sharing network for vehicle trajectory prediction. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–7.
51. Kim, J.; Atkins, E. Airspace Geofencing and Flight Planning for Low-Altitude, Urban, Small Unmanned Aircraft Systems. *Appl. Sci.* **2022**, *12*, 576. [[CrossRef](#)]
52. Hermand, E.; Nguyen, T.W.; Hosseinzadeh, M.; Garone, E. Constrained control of UAVs in geofencing applications. In Proceedings of the 2018 26th Mediterranean Conference on Control and Automation (MED), Zadar, Croatia, 19–22 June 2018.
53. Zhang, X.; Zhao, C.; Liao, F.; Li, X.; Du, Y. Online parking assignment in an environment of partially connected vehicles: A multi-agent deep reinforcement learning approach. *Transp. Res. Part C-Emerg. Technol.* **2022**, *138*, 103624.
54. Chen, X.; Yin, J.; Qin, G.; Tang, K.; Wang, Y.; Sun, J. Integrated macro-micro modelling for individual vehicle trajectory reconstruction using fixed and mobile sensor data. *Transp. Res. Part C-Emerg. Technol.* **2022**, *145*, 103929.
55. Du, Y.; Qin, B.; Zhao, C.; Zhu, Y.; Cao, J.; Ji, Y. A novel spatio-temporal synchronization method of roadside asynchronous MMW radar-camera for sensor fusion. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 22278–22289. [[CrossRef](#)]
56. Zheng, Z.; Jiang, M.; Wang, Z.; Wang, J.; Bai, Z.; Zhang, X.; Yu, X.; Tan, X.; Yang, Y.; Wen, S.; et al. Going Beyond Real Data: A Robust Visual Representation for Vehicle Re-identification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 13–19 June 2020; pp. 598–599.
57. Huang, Y.; Bi, H.; Li, Z.; Mao, T.; Wang, Z. STGAT: Modeling Spatial-Temporal Interactions for Human Trajectory Prediction. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 20–26 October 2019; pp. 6272–6281.
58. Du, Y.; Shi, Y.; Zhao, C.; Du, Z.; Ji, Y. A lifelong framework for data quality monitoring of roadside sensors in cooperative vehicle-infrastructure systems. *Comput. Electr. Eng.* **2022**, *100*, 108030.
59. Tang, Z.; Naphade, M.; Liu, M.; Yang, X.; Birchfield, S.; Wang, S.; Kumar, R.; Anastasiu, D.; Hwang, J. Cityflow: A city-scale benchmark for multi-target multi-camera vehicle tracking and re-identification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8797–8806.
60. Chang, M.; Lambert, J.; Sangkloy, P.; Singh, J.; Bak, S.; Hartnett, A.; Wang, D.; Carr, P.; Lucey, S.; Ramanan, D. Argoverse: 3d tracking and forecasting with rich maps. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8748–8757.

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.