The Fortieth AAAI Conference on Artificial Intelligence (AAAI-26)

KnowLCP: Knowledge Augmented Lane Change Prediction for Autonomous Driving

Yuhuan Lu[1,2], Pengpeng Xu[3], Wei Wang[1,4*], Zhen Zhang[1], Han Liu[5], Xiping Hu[1,4*]
[1]Artificial Intelligence Research Institute, Shenzhen MSU-BIT University
[2]Yunnan Key Laboratory of Intelligent Logistics Equipment and Systems, Kunming University
[3]School of Traffic and Transportation Engineering, Central South University
[4]School of Medical Technology, Beijing Institute of Technology
[5]School of Software, Dalian University of Technology
ryanlu32@gmail.com, pengpengxu@yeah.net, ehomewang@ieee.org,
zhangzhen19@lzu.edu.cn, liu.han.dut@gmail.com, huxp@bit.edu.cn

Abstract

Lane change prediction, encompassing both intention recognition and trajectory forecasting, is essential for the safe operation of autonomous vehicles in mixed-traffic environments. Existing models predominantly follow a data-driven paradigm, learning directly from historical vehicle states through an end-to-end approach. Inspired by the emerging paradigm of enhancing model generalizability through domain knowledge, we propose KnowLCP to explicitly model and integrate driving knowledge into the lane change prediction task. Specifically, we incorporate three types of knowledge: traffic risk awareness to improve intention prediction, vehicle kinematics to ensure the physical feasibility of predicted trajectories, and intention intensity to refine trajectory forecasting. Furthermore, we introduce a novel knowledge injection strategy that enhances mutual information during integration and proves superior to the traditional parallel input mechanism, which simply feeds knowledge features alongside historical states. Extensive experiments on two real-world trajectory datasets demonstrate that KnowLCP achieves average improvements of 8.3-10.3% in intention prediction and 10.1-10.3% in trajectory prediction over the best-performing baselines.

Introduction

During the shift from human-driven to fully autonomous road systems, a major challenge arises in mixed-autonomy traffic: autonomous vehicles (AVs) must safely interact with unpredictable human-driven vehicles (HDVs) (Wu et al. 2017). Human drivers' sudden lane changes significantly affect road safety, contributing to roughly 450,000 crashes annually in the U.S. (Chen et al. 2021). Therefore, reliable HDV lane change prediction is essential for enhancing AV safety and performance.

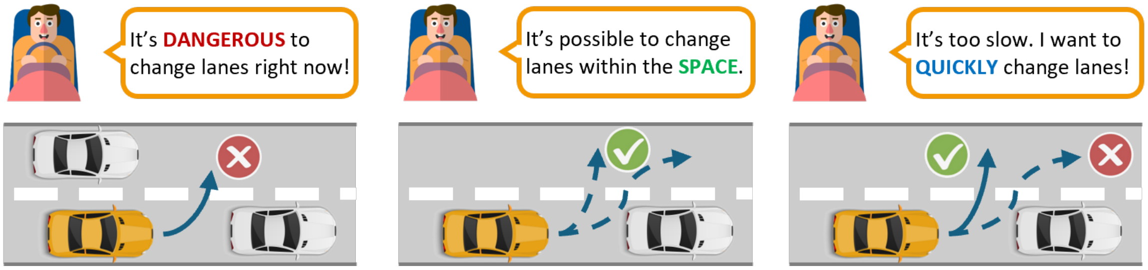
Lane change prediction is generally divided into two tasks: lane change intention prediction and lane change trajectory prediction (Xing et al. 2019). For intention prediction, earlier studies primarily focus on capturing the temporal interdependence in individual vehicle dynamics and identifying temporal variations to recognize lane change intentions (Shi and Zhang 2021; Xing et al. 2020). However, these approaches often neglect the mutual influence between vehicles, which is critical for perceiving subtle lane change behaviors, thereby limiting predictive performance. The rise of deep learning has facilitated methods incorporating the relative states between vehicles (Gao et al. 2023), which serve as strong cues to enhance intention prediction. For trajectory prediction, traditional approaches rely on physical kinematics (Ammoun and Nashashibi 2009) or maneuver constraints (Schreier, Willert, and Adamy 2016), confining their adaptability to individual-level perception. More recent research investigates modeling vehicle interactions using Graph Neural Networks (GNNs) (Jia et al. 2023; Lu et al. 2025a) to capture geometric relationships between vehicles, and self-attention mechanisms (Liao et al. 2024; Wang et al. 2025) to enable holistic feature fusion, thereby enriching contextual understanding for more accurate trajectory forecasting.

From the above discussion, it is evident that current lane change prediction methods primarily adopt a *data-driven* approach, learning directly from the historical states of vehicles via an end-to-end pipeline to predict future lane change intentions or trajectories. However, recent studies have demonstrated that incorporating domain knowledge can improve both the generalizability and reasoning capabilities of autonomous driving systems (Liu and Feng 2024; He et al. 2023; Lu et al. 2025d). This reflects that a top-down approach can help bridge information gaps that are not captured by raw data. Motivated by this, we propose to explicitly model and integrate driving knowledge into the lane change prediction task to enhance both intention and trajectory prediction performance. Nonetheless, two key challenges arise: (i) what types of knowledge should be incorporated, and (ii) how can such knowledge be effectively integrated into the existing lane change prediction pipeline?

Against this background, we incorporate three types of driving knowledge inspired by psychological factors influencing human drivers during lane changes (Yang et al. 2025; Lu et al. 2025c), as well as physical characteristics that constrain vehicle movement (Wang et al. 2025; Cui et al. 2020).

---

[*]Corresponding author.



(a) Risk awareness. (b) Kinematic constraints. (c) Intention guidance.

Figure 1: Examples of the utility of the three driving knowledge considered in this study.

The first is **risk awareness**. As illustrated in Figure 1a, human drivers typically make lane change decisions based on their perceived safety of the driving environment to avoid collisions. To model this behavior, we incorporate classical traffic safety metrics, such as time-to-collision (TTC) and its extensions (Minderhoud and Bovy 2001), to quantify risk awareness, thereby enhancing intention prediction. The second is **kinematic constraints**. As shown in Figure 1b, HDVs must adhere to physical laws of motion when executing lane changes. We therefore integrate vehicle kinematic equations (Cui et al. 2020) to constrain the prediction space, ensuring that the generated trajectories are physically plausible. Finally, we incorporate **intention guidance**. As shown in Figure 1c, the intensity of lane change intention, which reflects the urgency of a driver's desire to change lanes, can significantly affect the choice of future trajectories. To capture this, we leverage neural ordinary differential equations (ODEs) (Chen et al. 2018) to model the continuous variability of intention context over time. This allows us to gauge the evolving tendency to change lanes and incorporate it into trajectory prediction. Building upon the modeling of driving knowledge, we further propose an innovative knowledge integration philosophy. Unlike existing approaches in autonomous driving, which typically combine historical states with driving knowledge as parallel inputs to a feature extractor (Wang et al. 2025; Lu et al. 2025c), we draw inspiration from the demonstrated superiority of fine-tuning over in-context learning in large language models (LLMs) (Liu et al. 2022). Accordingly, we design a dual-stream architecture that separates the processing pipelines for historical states and driving knowledge. Instead of feeding them in parallel, we integrate the knowledge into intermediate layers of the feature extractor. This design enables more effective fusion of contextual cues and driving knowledge, which, as shown in this study, improves lane change prediction performance. We also provide a theoretical justification for this design from the perspective of mutual information theory.

To conclude, the contributions of this work are summarized as follows: (1) We propose **KnowLCP**, a knowledge-augmented model for joint lane change intention and trajectory prediction. KnowLCP adopts a dual-Transformer architecture, with separate branches for intention and trajectory prediction, each taking as input the historical states of vehicles. This design is inspired by the most advanced lane change prediction model presented in (Gao et al. 2023), which also serves as the strongest baseline in our experimental evaluations. (2) Building on this dual-Transformer framework, we incorporate three types of driving knowledge to enhance prediction performance via top-down information enrichment: risk awareness to improve intention prediction, and kinematic constraints and intention guidance to refine trajectory prediction. (3) We introduce a novel knowledge integration strategy that incorporates driving knowledge into the Transformer layers, inspired by the fine-tuning paradigm in LLMs, thereby expanding the model's perception and utilization of domain knowledge. Specifically, we encode the driving knowledge into fixed-length key-value pairs, which can be seamlessly injected into each attention layer of the Transformer to influence the attention computation at multiple levels. (4) Extensive experiments on two widely used datasets demonstrate the superiority of KnowLCP over state-of-the-art baselines and validate the effectiveness of our proposed design choices.

Related Work

Lane Change Intention Prediction. Lane change intention prediction aims to infer the future intentions of human drivers over a defined time horizon based on the observable historical driving states of multiple vehicles. Recurrent Neural Networks (RNN) and their variants, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have been widely utilized to capture temporal variations in driving features, demonstrating strong performance in intention prediction (Xing et al. 2020; Yuan et al. 2023; Wang et al. 2024; Liao et al. 2022; Li et al. 2024a). More recently, Transformer, a revolutionary architecture in natural language processing, has been adopted in autonomous driving applications (Vaswani et al. 2017). Inspired by its success, Transformer-based models have also been explored for lane change intention prediction, outperforming traditional RNN-based approaches (Lu et al. 2025c; Gao et al. 2024; Lu et al. 2025b).

Lane Change Trajectory Prediction. Traditionally, lane change trajectory prediction has been treated as a subtask of general trajectory prediction. As a result, many generic trajectory prediction methods have been directly applied to lane change scenarios. Early approaches based on physics models (Ammoun and Nashashibi 2009; Xie et al. 2017) and maneuver-based strategies (Schreier, Willert, and Adamy 2016; Deo, Rangesh, and Trivedi 2018) primarily focus on modeling the individual dynamics of vehicles. However, these methods often overlook the influence of surrounding vehicles, resulting in overly idealized and less reliable predictions. With the advancement of deep learning, there has been a surge in methods that explicitly model interactions among vehicles. For instance, (Jia et al. 2023) and (Lu et al. 2025a) employ GNNs to capture geometric relationships among vehicles, while (Liao et al. 2024; Wang et al. 2025; Liu et al. 2024; Zhang et al. 2024) utilize Transformer architectures to model high-order inter-vehicle interactions. Although these methods have achieved strong performance in lane keeping scenarios, their effectiveness in lane change scenarios remains limited. This is largely due to their reliance on abundant training data from lane keeping behaviors, which may not generalize well to the more sophisticated and less frequent lane change cases. To address this gap, recent efforts have focused on integrating maneuver-specific features into Transformer-based models, thereby enabling more accurate and context-aware predictions tailored to lane change trajectories (Li et al. 2024b). These developments highlight the need for a dedicated trajectory prediction framework specifically designed for lane change scenarios.

Joint Lane Change Prediction. Currently, joint intention-trajectory prediction models primarily focus on capturing interactions between intention and trajectory pipelines (Gao et al. 2023; Do, Han, and Choi 2023). Our experiments confirm that joint prediction offers advantages for both tasks, primarily because the mutual awareness between the two pipelines allows complementary information to be effectively leveraged. Therefore, in this study, we adopt the joint prediction framework proposed in (Gao et al. 2023) as the foundational structure upon which we build our knowledge-augmented paradigm.

Methodology

Problem Formulation

In this study, we aim to build a lane change prediction model for an autonomous vehicle (referred to as the *ego vehicle*) that simultaneously forecasts the short-term lane change intentions and trajectories of its nearby vehicles (i.e., *target vehicles*). Given the current time t, we define the historical states of vehicles as $\mathcal{X}^{t-T_h+1:t} = \left\{\mathbf{X}_1^{t-T_h+1:t}, \ldots, \mathbf{X}_i^{t-T_h+1:t}, \ldots\right\}$. Here, T_h represents the traceback time window. Specifically, the historical states can be expressed as:

$$\mathbf{X}_i^{t-T_h+1:t} = \left\{\mathbf{p}_i^{t-T_h+1:t}, \mathbf{v}_i^{t-T_h+1:t}, \mathbf{a}_i^{t-T_h+1:t}, \mathbf{l}_i\right\}, \quad (1)$$

where the terms $\mathbf{p}_i^{t-T_h+1:t}$, $\mathbf{v}_i^{t-T_h+1:t}$, and $\mathbf{a}_i^{t-T_h+1:t}$ represent the time series of 2D coordinates, velocities, and accelerations of vehicle i, respectively. $\mathbf{l}_i$ denotes the vehicle size. By capturing intricate relationships among these historical states, the proposed lane change prediction model generates the predicted lane change intentions for the next time step $t+1$ and target vehicle trajectories over a future time horizon T_f:

$$\mathbf{Y}_{int}^{t+1} = \left\{\mathbf{y}_{int,i}^{t+1} | i \in [1, n]\right\}, \quad (2)$$

$$\mathbf{Y}_{traj}^{t+1:t+T_f} = \left\{\mathbf{y}_{traj,i}^{t+1:t+T_f} | i \in [1, n]\right\}, \quad (3)$$

where n denotes the total number of target vehicles. $\mathbf{y}_{int,i}^{t+1}$ represents the probability distribution over three lane change intention modes: lane keeping, left lane change, and right lane change, ensuring that the sum of its components equals 1. Meanwhile, $\mathbf{y}_{traj,i}^{t+1:t+T_f}$ denotes the time series of 2D coordinates, representing the most likely trajectory of the target vehicle.

Overall Framework

The overall framework of the proposed KnowLCP is illustrated in Figure 2. Specifically, we design two branches: a vanilla lane change prediction branch (indicated by blue arrows) and a knowledge augmentation branch (indicated by purple arrows). The main branch comprises two parallel pipelines for intention and trajectory prediction, each implemented using a decoder-only Transformer architecture, following the design of (Gao et al. 2023). For a given target vehicle i, the input consists of the historical states $\left\{\mathbf{X}_i^{t-T_h+1:t}, \mathcal{X}_{\mathcal{N}_i}^{t-T_h+1:t}\right\}$, where $\mathcal{N}_i$ denotes the set of surrounding vehicles of the target vehicle i. The dual Transformers process these inputs to produce intention and trajectory context representations, which are subsequently passed through Multilayer Perceptrons (MLPs) to generate the final predictions. In the knowledge augmentation branch, the three types of domain knowledge are first modeled and encoded into learnable embeddings using MLPs or LSTMs, resulting in key-value pairs. These key-value pairs are then directly injected into the hidden attention layers of the Transformers. Detailed descriptions of the knowledge encoding and injection mechanisms are provided in the following sections.

Knowledge Encoding

We consider the three types of domain knowledge: risk awareness for intention prediction, and kinematic constraints and intention guidance for trajectory prediction.

Risk Awareness Perceived safety aids in capturing the underlying psychological processes of human drivers, serving as a crucial indicator of future driving intentions. In the traffic safety domain, perceived safety is quantified using three risk metrics (Minderhoud and Bovy 2001): Time-to-Collision (TTC), Time Exposed Time-to-Collision (TET), and Time Integrated Time-to-Collision (TIT). While these metrics effectively capture risk relationships between vehicles, they have notable limitations: they do not account for the varying number of surrounding vehicles for a target vehicle, and they are designed primarily for single-lane (1D) scenarios. To overcome these limitations, we propose enhanced formulations of these three metrics that extend their applicability to multi-vehicle and 2D settings. For clarity, we present the modified calculation methods below and provide

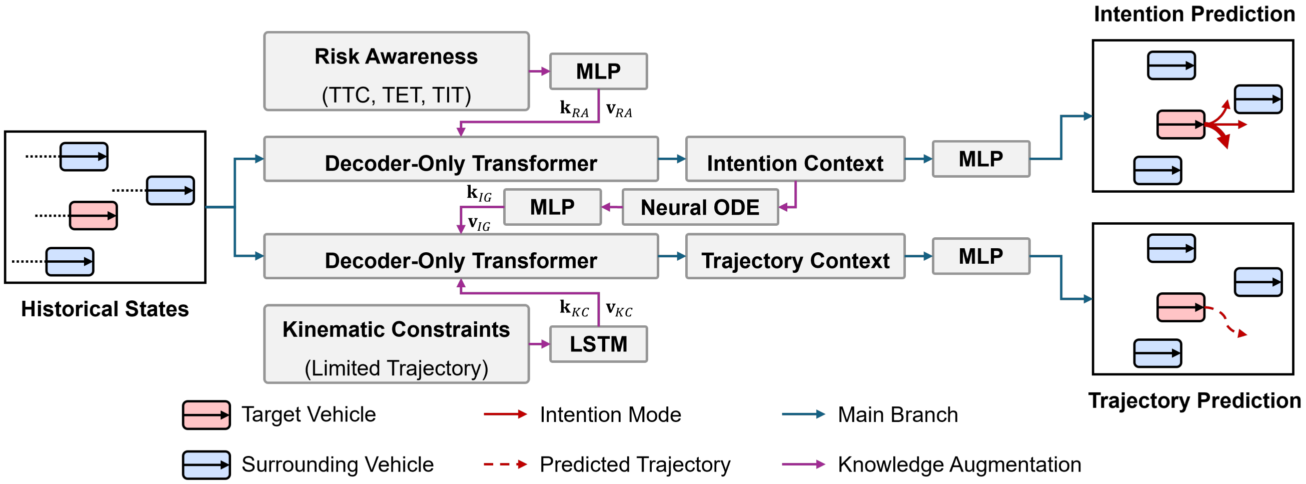


Figure 2: Overall framework of KnowLCP. MLP: Multilayer Perceptron; LSTM: Long Short-Term Memory network.

the theoretical justification for generalizing from 1D to 2D in Appendix A.

1) Time-to-Collision (TTC): TTC represents the time remaining until two vehicles collide if they continue on their current trajectories. It serves as an indicator of impending risk and functions as an early warning system. The TTC for the target vehicle i with respect to its surrounding vehicle j at time t is calculated as follows:

$$\delta_{TTC_{i,j}^t} = -\frac{d_{i,j}^t}{\hat{d}_{i,j}^t}. \tag{4}$$

$d_{i,j}^t$ denotes the distance between vehicles i and j and $\hat{d}_{i,j}^t$ is the change rate of $d_{i,j}^t$:

$$d_{i,j}^t = \sqrt{\left(p_i^t - p_j^t\right)^\top \left(p_i^t - p_j^t\right)}, \tag{5}$$

$$\hat{d}_{i,j}^t = \frac{1}{d_{i,j}^t}\left(p_i^t - p_j^t\right)^\top \left(v_i^t - v_j^t\right), \tag{6}$$

where (p_i^t, v_i^t) and $\left(p_j^t, v_j^t\right)$ represent the 2D coordinates and velocities of vehicles i and j, respectively. To accommodate the varying number of surrounding vehicles, the aggregated TTC for vehicle i is computed using a learnable operation:

$$\delta_{TTC_i^t} = \sum_{j \in \mathcal{N}_i} \alpha_{i,j}^t \delta_{TTC_{i,j}^t}, \tag{7}$$

where $\alpha_{i,j}^t$ is the attention weight of vehicle j to vehicle i learned by Graph Attention Networks (GATs) (Veličković et al. 2017).

2) Time Exposed Time-to-Collision (TET): TET quantifies the time during which a vehicle is exposed to safety-critical TTC values over the specified duration T_h:

$$\delta_{TET_i^t} = \sum_{t_k = t - T_h + 1}^{t} \zeta_i\left(t_k\right) \cdot \tau_{\text{sc}}, \tag{8}$$

where τ_{sc} is set to 0.1s, representing the minimum time step during which the measured TTC values remain unchanged. $\zeta_i\left(t_k\right)$ is a switching variable that indicates whether the TTC at time t_k falls within the safety-critical threshold:

$$\zeta_i(t_k) = \begin{cases} 1, & \forall 0 \leq \delta_{TTC_i^{t_k}} \leq \delta_{TTC^*}, \\ 0, & \text{else.} \end{cases} \tag{9}$$

where δ_{TTC^*} denotes the threshold safety-critical TTC value, which is set to 2.5s in this study.

3) Time Integrated Time-to-Collision (TIT): TET does not differentiate between varying risk levels, as it treats all TTC values within the threshold as equally critical. To address this limitation, TIT integrates the TTC profiles to more effectively represent the degree of safety:

$$\begin{gathered} \delta_{TIT_i^t} = \sum_{t_k = t - T_h + 1}^{t} \left[\delta_{TTC^*} - \delta_{TTC_i^{t_k}}\right] \cdot \tau_{\text{sc}}, \\ \forall 0 \leq \delta_{TTC_i^{t_k}} \leq \delta_{TTC^*}. \end{gathered} \tag{10}$$

A lower TTC and higher values of TET and TIT indicate greater exposure to potential collision risks, leading to a reduction in perceived safety. To encode this perceived safety, we employ two separate MLPs to process the combined safety features of the target vehicle, denoted as $\left\{ \delta_{TTC_i^{t_k}} \Big|_{t_k = t - T_h + 1}^{t}, \delta_{TET_i^t}, \delta_{TIT_i^t} \right\}$. These MLPs produce the corresponding key and value embeddings, $\tilde{\mathbf{k}}_{RA}$ and $\tilde{\mathbf{v}}_{RA}$, which are used to inject risk-awareness information into the Transformer layers for intention prediction.

Kinematic Constraints To augment the physical feasibility of predicted trajectories, we employ a two-axle vehicle kinematics model (Wang et al. 2025; Cui et al. 2020), predicated on the vehicle's center of mass. This model is instrumental in refining motion dynamics over the look-ahead horizon, serving as a pivotal reference for the trajectory prediction stage. The state of the target vehicle i at the next time step $t + 1$ is determined using the state transition equations (Wang et al. 2025). Through an iterative process, the kinematics-based future vehicle trajectory is generated and denoted as $\dot{\mathbf{p}}_i^{t+1:t+T_f}$. To encode the corresponding physical constraints, we employ two separate LSTM networks that transform the trajectory into key and value embeddings (hidden states in LSTMs), denoted as $\tilde{\mathbf{k}}_{KC}$ and $\tilde{\mathbf{v}}_{KC}$, respectively.

Intention Guidance Intention intensity plays a critical role in shaping the driving maneuvers of human drivers, thereby influencing the future trajectory of the target vehicle. To model this latent tendency, we utilize a neural ODE to capture the continuous evolution of intention. This is achieved by applying the neural ODE to the intention context learned by the Transformer in the main intention prediction branch:

$$\hat{\mathbf{h}}_{IG,i}^{t-T_h+1:t} = \varphi_{LN}\left(\varphi_{ODE}\left(\bar{\mathbf{X}}_{IP,i}^{t-T_h+1:t}\right)\right), \qquad (11)$$

where $\bar{\mathbf{X}}_{IP,i}^{t-T_h+1:t}$ denotes the temporal sequence of the extracted intention context over the traceback time window and $\varphi_{LN}(\cdot)$ represents Layer Normalization (Ba, Kiros, and Hinton 2016), which is applied to enhance training stability. The neural ODE is implemented using an LSTM, and $\hat{\mathbf{h}}_{IG,i}^{t-T_h+1:t}$ represents the densified hidden states, with twice the number of original time intervals (i.e., of size $2T_h$), to more meticulously capture the variability of intention tendency. In summary, incorporating intention intensity contributes to more accurate trajectory prediction by aligning the output with the driver's behavioral patterns, thereby improving the plausibility and rationality of the predicted trajectories. Afterward, two separate MLPs are applied to the intention-guided representation $\hat{\mathbf{h}}_{IG,i}^{t-T_h+1:t}$ to generate the corresponding key and value embeddings, denoted as $\tilde{\mathbf{k}}_{IG}$ and $\tilde{\mathbf{v}}_{IG}$, respectively.

Knowledge Injection

After obtaining the key-value pairs corresponding to the three types of driving knowledge, we inject this information into each attention layer of the dual Transformers. Specifically, let the input to the l-th attention layer be $\mathbf{Z}^l = \left[\mathbf{z}_1^l, \ldots, \mathbf{z}_j^l, \ldots, \mathbf{z}_{\mathcal{M}}^l\right]$, where $\mathcal{M} = (\mathcal{N}_i + 1) \times T_h$. The output of the attention layer is then computed as follows, where we consider a single attention head for simplicity:

$$\tilde{\mathbf{z}}_j^l = \frac{\sum_{k=1}^{\mathcal{M}} \exp\left(w_{j,k}^l\right)\mathbf{v}_k^l + \sum_{p\in\mathcal{P}} \exp\left(\tilde{w}_{j,p}^l\right)\tilde{\mathbf{v}}_p^l}{\sum_{k=1}^{\mathcal{M}} \exp\left(w_{j,k}^l\right) + \sum_{p\in\mathcal{P}} \exp\left(\tilde{w}_{j,p}^l\right)}, \qquad (12)$$

where $w_{j,k}^l = \left\langle \mathbf{q}_j^l, \mathbf{k}_k^l \right\rangle / \sqrt{D}$, $\tilde{w}_{j,p}^l = \left\langle \tilde{\mathbf{q}}_j^l \tilde{\mathbf{k}}_p^l \right\rangle / \sqrt{D}$, and $\tilde{\mathbf{q}}_j^l = \tilde{\mathbf{W}}_Q^l \mathbf{z}_j^l$. Here, $\tilde{\mathbf{W}}_Q^l$ is the learnable query head initialized from the vanilla query head $\mathbf{W}_Q^l$, and D denotes the embedding size. The set $\mathcal{P}$ refers to the injected key-value pairs, where $\mathcal{P} = \{RA\}$ is used for intention prediction, and $\mathcal{P} = \{KC, IG\}$ is used for trajectory prediction. This design enables the domain knowledge to attend to the hidden state features within each attention layer, thereby enhancing the utility and influence of the incorporated knowledge. To further elucidate the superiority of this injection strategy over the conventional parallel input of historical states and knowledge features, we provide a theoretical justification from the perspective of mutual information theory in Appendix B.

Model Training

We adopt a multi-task learning strategy, where the overall loss function for KnowLCP is defined as:

$$\begin{aligned}\mathcal{L} =& \mathcal{L}_{IP} + \mathcal{L}_{TP} \\ =& -\frac{1}{n}\sum_{i=1}^{n}\sum_{c=1}^{3} \bar{y}_{int,i,c}^{t+1} \log y_{int,i,c}^{t+1} \\ &+ \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{1}{T_f}\sum_{t_k=t+1}^{t+T_f}\left(\bar{y}_{traj,i}^{t_k} - y_{traj,i}^{t_k}\right)^2}.\end{aligned} \qquad (13)$$

The cross-entropy loss is employed for training the intention prediction pipeline, while the RMSE loss is used for optimizing the trajectory prediction pipeline. Here, $\bar{y}_{int,i,c}^{t+1}$ represents the ground-truth label for the c-th intention mode and $\bar{y}_{traj,i}^{t_k}$ denotes the ground-truth trajectory point at time t_k. In this study, as both intention and trajectory predictions are of equal importance, no trade-off parameters are introduced between the two loss terms.

Experiments

Experimental Setup

We evaluate KnowLCP using two widely adopted real-world trajectory datasets: NGSIM (Deo and Trivedi 2018) and HighD (Krajewski et al. 2018). Following prior studies (Gao et al. 2023), the lane change intention is assumed to be triggered 2 seconds before the execution of a lane change. Accordingly, we define a lane change sample as the 2 seconds before and 4 seconds after the start of the lane change maneuver. The traceback time window T_h and the prediction horizon T_f are both set to 3 seconds. Based on this definition, we obtain 10,126 lane change/lane keeping samples from the NGSIM dataset and 15,432 samples from the HighD dataset. Of these, 70% are used for model training, 15% for validation, and the remaining 15% for evaluation. To ensure a fair comparison, we adhere to the established evaluation metrics (Gao et al. 2023; Li et al. 2024a): Precision (P) and Recall (R) for intention prediction, and RMSE and ADE for trajectory prediction. Moreover, we implement KnowLCP using the Adam optimizer with a batch size of 128 and a learning rate ranging from 10^{-3} to 10^{-5}. The model is trained for 120 epochs on the NGSIM dataset and 200 epochs on the HighD dataset. We compare KnowLCP against a comprehensive set of state-of-the-art baselines across three categories: intention prediction, trajectory prediction, and joint prediction. The hyperparameter analysis and inference time of KnowLCP are presented in Appendices C-D.

Overall Performance

We evaluate the performance of KnowLCP from the perspectives of both intention prediction and trajectory predic-

Method		NGSIM				HighD			
		Left LC		Right LC		Left LC		Right LC	
		P↑	R↑	P↑	R↑	P↑	R↑	P↑	R↑
Intention Prediction	Bi-LSTM (Xing et al. 2020)	76.3%	77.3%	72.7%	72.2%	70.7%	70.4%	67.5%	67.1%
	TMMOE (Yuan et al. 2023)	75.2%	74.6%	72.5%	72.9%	73.4%	73.5%	70.5%	70.6%
	VWC (Wang et al. 2024)	80.1%	81.3%	76.6%	76.1%	76.7%	77.0%	74.6%	74.5%
	H-LSTM (Liao et al. 2022)	77.2%	75.7%	76.9%	75.8%	78.1%	77.5%	75.0%	76.1%
	P-GHMM (Li et al. 2024a)	80.1%	81.2%	81.3%	80.7%	77.3%	75.9%	76.5%	77.1%
	STMF (Gao et al. 2024)	78.4%	79.1%	80.6%	81.7%	80.4%	80.5%	78.6%	77.9%
	DUAL (Lu et al. 2025c)	90.2%	91.4%	89.7%	89.4%	88.5%	87.9%	87.8%	87.3%
Joint Prediction	MMAE (Do, Han, and Choi 2023)	82.5%	80.3%	78.6%	79.4%	77.5%	77.0%	75.1%	74.7%
	RFormer (Gao et al. 2023)	90.5%	90.4%	89.9%	90.1%	89.0%	88.2%	88.5%	87.9%
	KnowLCP	**98.1%**	**98.2%**	**97.7%**	**97.9%**	**97.8%**	**97.6%**	**97.4%**	**97.3%**

Table 1: Intention prediction comparison on NGSIM and HighD datasets. The underline and **bold** values denote the best results in baselines and overall, respectively.

Method		NGSIM				HighD			
		Lane Change		Lane Keeping		Lane Change		Lane Keeping	
		RMSE↓	ADE↓	RMSE↓	ADE↓	RMSE↓	ADE↓	RMSE↓	ADE↓
Trajectory Prediction	HDGT (Jia et al. 2023)	4.58	4.16	3.29	3.14	1.36	1.26	1.03	0.99
	BAT (Liao et al. 2024)	4.62	4.15	3.62	3.38	1.47	1.29	1.17	1.10
	LAformer (Liu et al. 2024)	4.59	4.25	3.54	3.39	1.33	1.28	1.04	0.96
	SeFlow (Zhang et al. 2024)	4.64	4.19	3.37	3.23	1.37	1.20	1.15	1.02
	HyperMTP (Lu et al. 2025a)	4.52	4.27	3.42	3.19	1.29	1.08	1.05	0.92
	WAKE (Wang et al. 2025)	4.43	4.08	**3.19**	**2.97**	1.30	1.16	**0.96**	**0.88**
	LCT-DPP (Li et al. 2024b)	4.46	4.09	3.47	3.05	1.27	1.07	1.12	1.01
Joint Prediction	MMAE (Do, Han, and Choi 2023)	4.67	4.35	3.69	3.28	1.44	1.29	1.20	1.07
	RFormer (Gao et al. 2023)	4.22	3.96	3.21	3.03	1.11	1.04	1.01	0.93
	KnowLCP	**3.87**	**3.49**	3.22	3.02	**1.02**	**0.91**	0.98	0.89

Table 2: Trajectory prediction comparison on NGSIM and HighD datasets. The evaluation metrics are measured in meters.

tion.

Intention Prediction: Table 1 presents the intention prediction results on the two datasets, with the best-performing result for each task on each dataset highlighted. Among the baselines, RFormer achieves the highest performance, demonstrating the effectiveness of jointly learning lane change intention and trajectory. Although DUAL is the only baseline that incorporates domain knowledge, its parallel input design limits its ability to fully leverage this information, highlighting the advantage of our proposed knowledge integration strategy. Overall, KnowLCP consistently outperforms all baselines, achieving an average improvement of **8.3-10.3%** over the best-performing methods across both datasets.

Trajectory Prediction: Table 2 presents the trajectory prediction results on the two datasets. Consistent with the findings for intention prediction, RFormer outperforms single-task trajectory prediction models in lane change, further validating the effectiveness of joint learning. Notably, general-purpose trajectory prediction models tend to perform better in lane keeping scenarios but fall short in lane change settings, where the lane change-specific baseline LCT-DPP demonstrates superior performance. This underscores the importance of specialized modeling for trajectory forecasting in complex lane change scenarios. WAKE, which considers vehicle kinematics, achieves the best results in lane keeping; however, its parallel input design limits its effectiveness in more dynamic lane change scenarios. In contrast, KnowLCP consistently outperforms all baselines in lane change, achieving an average improvement of **10.1-10.3%** over the best-performing alternatives across both datasets.

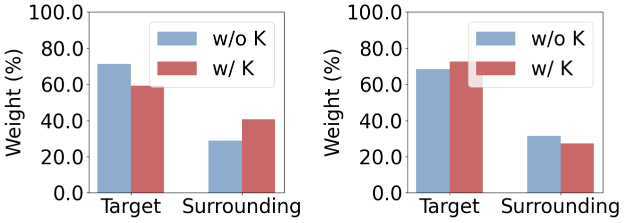
Ablation Study

To validate the effectiveness of key components in KnowLCP, we conduct an ablation study focusing on the core design choices of knowledge encoding and knowledge injection. The corresponding results are reported in Table 3.

We first design *w/o RA*, a variant that removes the risk awareness module from the intention prediction pipeline. The results show a significant decline in both intention and trajectory prediction performance, underscoring the critical role of perceived safety modeling in lane change scenarios. Capturing traffic safety levels enhances perception and refines the generation of lane change intentions and trajectories for human drivers. Similarly, *w/o KC* eliminates the kinematic constraints module from the trajectory prediction pipeline, and the results indicate that the physical constraints imposed by vehicle kinematics substantially improve the plausibility of predicted trajectories, indirectly enhancing lane change intention generation. Another variant, *w/o IG*, removes the incorporation of intention guidance, and the observed performance degradation suggests that mod-

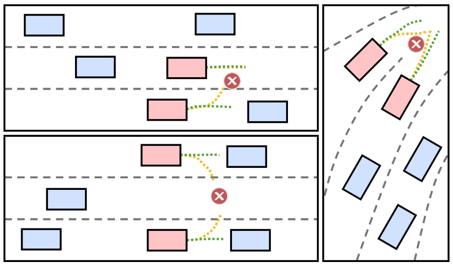
Method	NGSIM						HighD					
	Left LC		Right LC		Trajectory		Left LC		Right LC		Trajectory	
	P↑	R↑	P↑	R↑	RMSE↓	ADE↓	P↑	R↑	P↑	R↑	RMSE↓	ADE↓
KnowLCP	**98.1%**	**98.2%**	**97.7%**	**97.9%**	**3.87**	**3.49**	**97.8%**	**97.6%**	**97.4%**	**97.3%**	**1.02**	**0.91**
w/o RA	92.7%	92.9%	91.9%	91.7%	4.04	3.71	92.3%	91.8%	91.7%	91.5%	1.05	0.96
w/o KC	95.9%	95.6%	95.1%	94.8%	4.14	3.88	94.9%	94.7%	95.1%	95.2%	1.09	1.01
w/o IG	96.8%	96.7%	96.4%	96.1%	4.13	3.89	95.6%	95.1%	95.4%	95.7%	1.09	0.99
w/o KI	92.2%	92.5%	92.9%	92.3%	4.05	3.75	94.7%	93.9%	93.8%	93.2%	1.06	0.95

Table 3: Ablation results for different variants on NGSIM and HighD datasets. The trajectory task specifically refers to lane change trajectory prediction.



(a) Intention pipeline.

(b) Trajectory pipeline.



(c) Predicted trajectories of target vehicles by KnowLCP (green) and RFormer (yellow).

Figure 3: Further analysis of knowledge on the HighD dataset. (a)-(b) are the quantitative analyses of attention weights; and (c) is the qualitative analysis of performance in safety-critical scenarios.

eling intention variability ensures the compatibility of predicted trajectories with driver behavioral intentions, thereby improving both intention and trajectory prediction accuracy. Finally, we introduce *w/o KI*, where the proposed layer-wise knowledge injection strategy is replaced with the conventional approach of feeding knowledge features in parallel with historical states. The results demonstrate that our knowledge injection strategy substantially improves both lane change intention and trajectory prediction, underscoring its effectiveness in attending to hidden cues within each attention layer. This performance gain aligns with our mutual information-based justification. These ablation results collectively demonstrate that each component of KnowLCP contributes meaningfully to enhancing prediction performance, validating the necessity of their inclusion in the overall framework.

Further Analysis

To further investigate the effect of domain knowledge, we conduct quantitative and qualitative analyses.

Quantitative Analysis: We compute average attention weights from the final layer of the dual Transformers and compare attention allocation between target and surrounding vehicles with and without knowledge. The results are illustrated in Figures 3a-3b. For intention prediction, we observe that the inclusion of knowledge increases attention to surrounding vehicles by assigning them higher attention weights, reflecting heightened risk awareness makes human drivers more alert to their surroundings. In contrast, for trajectory prediction, kinematic constraints and intention guidance encourage the model to focus more on the target vehicle's own motion, resulting in higher attention weights on the target itself. These findings validate that the incorporated domain knowledge aligns with human-like driving behaviors and enhances the interpretability of lane changes.

Qualitative Analysis: We visualize the trajectory prediction results in safety-critical scenarios, where baseline models (e.g., RFormer) often predict trajectories that risk collisions. As shown in Figure 3c, the incorporation of domain knowledge enables KnowLCP to more accurately capture realistic driver intentions and generate feasible trajectories, underscoring the importance of knowledge-augmented modeling in sophisticated traffic scenarios to achieve socially-aware lane change prediction.

Conclusion

We propose KnowLCP, a knowledge-augmented dual-Transformer architecture for lane change prediction. KnowLCP integrates three types of driving knowledge through a novel injection strategy that enriches mutual information with historical states. Experiments on two real-world datasets show its superiority over state-of-the-art baselines. As the first model to fuse domain knowledge with joint intention and trajectory prediction, KnowLCP underscores the potential of knowledge-augmented behavior prediction in autonomous driving.

While the joint prediction paradigm enhances mutual information between tasks, misclassified intentions may adversely affect trajectory prediction, which we leave for future investigation.

Acknowledgments

This study was funded in part by the Innovation Team Project of Guangdong Province of China (No. 2024KCXTD017), in part by the Shenzhen Science and Technology Foundation (No. JCYJ20240813145816022), in part by the National Natural Science Foundation of China (Grant No.62576213), and in part by the Yunnan Science and Technology Innovation Base Construction Project (No. 202407AB110006).

References

Ammoun, S.; and Nashashibi, F. 2009. Real time trajectory prediction for collision risk estimation between vehicles. In *Proc. IEEE 5th Int. Conf. Intell. Comput. Commun. Process.*, 417–422. IEEE.

Ba, J. L.; Kiros, J. R.; and Hinton, G. E. 2016. Layer normalization. *arXiv preprint arXiv:1607.06450*.

Chen, Q.; Huang, H.; Li, Y.; Lee, J.; Long, K.; Gu, R.; and Zhai, X. 2021. Modeling accident risks in different lane-changing behavioral patterns. *Analytic methods in accident research*, 30: 100159.

Chen, R. T.; Rubanova, Y.; Bettencourt, J.; and Duvenaud, D. K. 2018. Neural ordinary differential equations. *Advances in neural information processing systems*, 31.

Cui, H.; Nguyen, T.; Chou, F.-C.; Lin, T.-H.; Schneider, J.; Bradley, D.; and Djuric, N. 2020. Deep kinematic models for kinematically feasible vehicle trajectory predictions. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 10563–10569. IEEE.

Deo, N.; Rangesh, A.; and Trivedi, M. M. 2018. How would surround vehicles move? A unified framework for maneuver classification and motion prediction. *IEEE Transactions on Intelligent Vehicles*, 3(2): 129–140.

Deo, N.; and Trivedi, M. M. 2018. Convolutional social pooling for vehicle trajectory prediction. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 1468–1476.

Do, J.; Han, K.; and Choi, S. B. 2023. Lane change–intention inference and trajectory prediction of surrounding vehicles on highways. *IEEE Transactions on Intelligent Vehicles*, 8(7): 3813–3825.

Gao, K.; Li, X.; Chen, B.; Hu, L.; Liu, J.; Du, R.; and Li, Y. 2023. Dual transformer based prediction for lane change intentions and trajectories in mixed traffic environment. *IEEE Transactions on Intelligent Transportation Systems*.

Gao, K.; Li, X.; Hu, L.; Liu, X.; Zhang, J.; Du, R.; and Li, Y. 2024. STMF-IE: A Spatial-Temporal Multi-Feature Fusion and Intention-Enlightened Decoding Model for Vehicle Trajectory Prediction. *IEEE Transactions on Vehicular Technology*.

He, X.; Wu, J.; Huang, Z.; Hu, Z.; Wang, J.; Sangiovanni-Vincentelli, A.; and Lv, C. 2023. Fear-neuro-inspired reinforcement learning for safe autonomous driving. *IEEE transactions on pattern analysis and machine intelligence*, 46(1): 267–279.

Jia, X.; Wu, P.; Chen, L.; Liu, Y.; Li, H.; and Yan, J. 2023. Hdgt: Heterogeneous driving graph transformer for multi-agent trajectory prediction via scene encoding. *IEEE transactions on pattern analysis and machine intelligence*.

Krajewski, R.; Bock, J.; Kloeker, L.; and Eckstein, L. 2018. The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In *2018 21st international conference on intelligent transportation systems (ITSC)*, 2118–2125. IEEE.

Li, Z.; Wang, Y.; Zuo, Z.; and Hu, C. 2024a. A Hierarchical Intention Inference Model for Connected and Automated Vehicles. *IEEE Transactions on Intelligent Vehicles*.

Li, Z.; Zhao, X.; Zhao, C.; Liu, Y.; and Wang, C. 2024b. A Self-Selection Personalized Lane-Changing Trajectory Prediction Approach Through a Hybrid Deep Learning Model. *IEEE Transactions on Vehicular Technology*.

Liao, H.; Li, Z.; Shen, H.; Zeng, W.; Liao, D.; Li, G.; and Xu, C. 2024. Bat: Behavior-aware human-like trajectory prediction for autonomous driving. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 10332–10340.

Liao, X.; Wang, Z.; Zhao, X.; Zhao, Z.; Han, K.; Tiwari, P.; Barth, M. J.; and Wu, G. 2022. Online prediction of lane change with a hierarchical learning-based approach. In *2022 International Conference on Robotics and Automation (ICRA)*, 948–954. IEEE.

Liu, H.; Tam, D.; Muqeeth, M.; Mohta, J.; Huang, T.; Bansal, M.; and Raffel, C. A. 2022. Few-shot parameter-efficient fine-tuning is better and cheaper than in-context learning. *Advances in Neural Information Processing Systems*, 35: 1950–1965.

Liu, H. X.; and Feng, S. 2024. Curse of rarity for autonomous vehicles. *nature communications*, 15(1): 4808.

Liu, M.; Cheng, H.; Chen, L.; Broszio, H.; Li, J.; Zhao, R.; Sester, M.; and Yang, M. Y. 2024. Laformer: Trajectory prediction for autonomous driving with lane-aware scene constraints. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2039–2049.

Lu, Y.; Wang, W.; Bai, R.; Zhou, S.; Garg, L.; Bashir, A. K.; Jiang, W.; and Hu, X. 2025a. Hyper-relational interaction modeling in multi-modal trajectory prediction for intelligent connected vehicles in smart cites. *Information Fusion*, 114: 102682.

Lu, Y.; Xu, P.; Jiang, X.; Bashir, A. K.; Gadekallu, T. R.; Wang, W.; and Hu, X. 2025b. Lane Change Prediction for Autonomous Driving With Transferred Trajectory Interaction. *IEEE Transactions on Intelligent Transportation Systems*.

Lu, Y.; Zhang, Z.; Bai, R.; Liu, H.; and Wang, W. 2025c. Dual-View Interaction-Aware Lane Change Prediction for Autonomous Driving. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 28222–28230.

Lu, Y.; Zhang, Z.; Wang, W.; Zhu, Y.; Chen, T.; Al-Otaibi, Y. D.; Bashir, A. K.; and Hu, X. 2025d. Knowledge-Driven Lane Change Prediction for Secure and Reliable Internet of

Vehicles. *IEEE Transactions on Intelligent Transportation Systems*.

Minderhoud, M. M.; and Bovy, P. H. 2001. Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis & Prevention*, 33(1): 89–97.

Schreier, M.; Willert, V.; and Adamy, J. 2016. An integrated approach to maneuver-based trajectory prediction and criticality assessment in arbitrary road environments. *IEEE Trans. Intell. Transp. Syst.*, 17(10): 2751–2766.

Shi, Q.; and Zhang, H. 2021. An improved learning-based LSTM approach for lane change intention prediction subject to imbalanced data. *Transportation research part C: emerging technologies*, 133: 103414.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; and Bengio, Y. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903*.

Wang, C.; Liao, H.; Li, Z.; and Xu, C. 2025. WAKE: Towards Robust and Physically Feasible Trajectory Prediction for Autonomous Vehicles With WAvelet and KinEmatics Synergy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Wang, X.; Zan, Y.; Zheng, J.; Wenjuan, E.; Tao, Y.; Xiao, Y.; and Li, T. 2024. Variable Weight Combination Model for Lane-changing Prediction of Human-Driven Vehicle. *IEEE Transactions on Intelligent Vehicles*.

Wu, C.; Kreidieh, A.; Vinitsky, E.; and Bayen, A. M. 2017. Emergent behaviors in mixed-autonomy traffic. In *Conference on Robot Learning*, 398–407. PMLR.

Xie, G.; Gao, H.; Qian, L.; Huang, B.; Li, K.; and Wang, J. 2017. Vehicle trajectory prediction by integrating physics- and maneuver-based approaches using interactive multiple models. *IEEE Transactions on Industrial Electronics*, 65(7): 5999–6008.

Xing, Y.; Lv, C.; Wang, H.; Cao, D.; and Velenis, E. 2020. An ensemble deep learning approach for driver lane change intention inference. *Transportation Research Part C: Emerging Technologies*, 115: 102615.

Xing, Y.; Lv, C.; Wang, H.; Wang, H.; Ai, Y.; Cao, D.; Velenis, E.; and Wang, F.-Y. 2019. Driver lane change intention inference for intelligent vehicles: Framework, survey, and challenges. *IEEE Transactions on Vehicular Technology*, 68(5): 4377–4390.

Yang, L.; Zhang, J.; Lyu, N.; and Zhao, Q. 2025. Predicting lane change maneuver and associated collision risks based on multi-task learning. *Accident Analysis & Prevention*, 209: 107830.

Yuan, R.; Abdel-Aty, M.; Xiang, Q.; Wang, Z.; and Gu, X. 2023. A Temporal Multi-Gate Mixture-of-Experts Approach for Vehicle Trajectory and Driving Intention Prediction. *IEEE Transactions on Intelligent Vehicles*.

Zhang, Q.; Yang, Y.; Li, P.; Andersson, O.; and Jensfelt, P. 2024. Seflow: A self-supervised scene flow method in autonomous driving. In *European Conference on Computer Vision*, 353–369. Springer.