Communications in Transportation Research 5 (2025) 100170

Contents lists available at ScienceDirect

Communications in Transportation Research

journal homepage: www.journals.elsevier.com/communications-in-transportation-research

Full Length Article

LC-LLM: Explainable lane-change intention and trajectory predictions with Large Language Models

Mingxing Peng [a], Xusen Guo [a], Xianda Chen [a], Kehua Chen [b], Meixin Zhu [c,*], Long Chen [d], Fei-Yue Wang [d]

[a] Intelligent Transportation (INTR), Systems Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, 511400, China
[b] Division of Emerging Interdisciplinary Areas (EMIA), Academy of Interdisciplinary Studies, The Hong Kong University of Science and Technology, Hong Kong, 999077, China
[c] Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things, Guangzhou, 511400, China
[d] State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China

ARTICLE INFO

Keywords:
Lane change (LC)
Large language models (LLMs)
Intention prediction
Trajectory prediction
Fine-tuning
Interpretability
Autonomous driving

ABSTRACT

To ensure safe driving in dynamic environments, autonomous vehicles should possess the capability to accurately predict lane change intentions of surrounding vehicles in advance and forecast their future trajectories. Existing motion prediction approaches have ample room for improvement, particularly in terms of long-term prediction accuracy and interpretability. In this study, we address these challenges by proposing a Lane Change-Large Language Model (LC-LLM), an explainable lane change prediction model that leverages the strong reasoning capabilities and self explanation abilities of Large Language Models (LLMs). Essentially, we reformulate the lane change prediction task as a language modeling problem, processing heterogeneous driving scenario information as natural language prompts for LLMs and employing supervised fine-tuning to tailor LLMs specifically for lane change prediction task. Additionally, we finetune the Chain-of-Thought (CoT) reasoning to improve prediction transparency and reliability, and include explanatory requirements in the prompts during the inference stage. Therefore, our LC-LLM not only predicts lane change intentions and trajectories but also provides CoT reasoning and explanations for its predictions, enhancing its interpretability. Extensive experiments based on the large-scale highD dataset demonstrate the superior performance and interpretability of our LC-LLM in lane change prediction task. To the best of our knowledge, this is the first attempt to utilize LLMs for predicting lane change behavior. Our study shows that LLMs can effectively encode comprehensive interaction information for understanding driving behavior.

1. Introduction

Lane change prediction of surrounding vehicles is a critical task for autonomous driving systems, as it enables autonomous vehicles to anticipate the lane change intentions and future trajectories of surrounding vehicles in advance, thereby facilitating informed decision-making. Generating accurate and explainable lane change predictions is essential for the development of safer and more reliable autonomous driving systems.

In recent years, considerable advancements have been made in lane change predictions. Various approaches have been explored to enhance performance in detecting lane change maneuvers (Gao et al., 2023; He et al., 2019; Mozaffari et al., 2022; Xin et al., 2018). However, at the early stage of a lane change maneuver when lateral displacement is minimal, it is essential to rely on the interaction information between the target vehicle and its surrounding vehicles for accurate predictions. Therefore, the ability to effectively model interaction information among vehicles becomes paramount. Previous methods exhibit limited capability in understanding interactive information, which results in weak long-term prediction performance (He et al., 2019; Xin et al., 2018). Additionally, lane change trajectory prediction has witnessed notable progress with the adoption of deep learning techniques, such as Graph Neural Networks (GNN) and Transformers, leading to competitive results (Gao et al., 2023; Seff et al., 2023; Shi et al., 2022). Despite these successes, these deep

* Corresponding author. Intelligent Transportation (INTR), Systems Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, 511400, China.
E-mail address: meixin@seu.edu.cn (M. Zhu).

https://doi.org/10.1016/j.commtr.2025.100170
Received 9 September 2024; Received in revised form 4 December 2024; Accepted 16 December 2024

learning-based approaches frequently suffer from a lack of interpretability, as they generate predictions about future behaviors without offering substantial explanations for their results. In summary, enhancing the capability to model interactions among vehicles and augmenting the interpretability of prediction results are key challenges in lane change prediction for autonomous driving.

Recently, advancements in Large Language Models (LLMs) have provided new opportunities for addressing challenges associated with lane change prediction. Noteworthy progress in LLMs (Achiam et al., 2023; Ouyang et al., 2022; Touvron et al., 2023) has showcased their robust information comprehension skills and powerful common sense reasoning abilities. However, current research on lane change prediction has not yet incorporated LLMs. Given the LLMs' profound understanding of complex driving scenarios and their extensive knowledge base, including common driving knowledge, it is rational to utilize LLMs in modeling interactions between vehicles for the enhancement of future driving behavior predictions. In addition, LLMs have demonstrated the capacity to leverage their vast knowledge base to generate explanations for their predictions (Huang et al., 2023; Rajani et al., 2019). Inspired by these studies, there is potential to leverage LLMs for the explanation of predicted lane change intentions and future trajectories, thereby augmenting the interpretability of predictions in autonomous driving.

In this study, we propose a novel approach that leverages LLMs' powerful reasoning and self-explanation capabilities to address the aforementioned lane change prediction challenges, enhancing both prediction performance and interpretability in the lane change prediction task.

In essence, we reformulate the task of predicting lane change intentions and trajectories as a language modeling problem. To this end, we integrate heterogeneous inputs, including target vehicle state, surrounding vehicle state, and map information, converting them into structured prompts in natural language to input into the LLM. Then, we employ a supervised fine-tuning technique to tailor the LLM specifically for the lane change prediction task. Additionally, in the fine-tuning stage, we perform Chain-of-Thought (CoT) reasoning to enhance transparency and reliability in lane change predictions. Finally, we integrate explanatory requirements into the prompts in the inference stage, thus enabling our fine-tuned model to generate explanations for the lane change prediction results. Benefiting from the powerful capabilities of LLM, our fine-tuned model demonstrates better performance and interpretability in lane change prediction.

The main contributions of this study include:

- We propose LC-LLM, the first Large Language Model for lane change prediction. It leverages the powerful capabilities of LLMs to understand complex interactive scenarios, enhancing the performance of lane change prediction.
- Our LC-LLM achieves explainable predictions. It not only predicts lane change intentions and trajectories but also generates explanations for the prediction results. Additionally, we evaluate the accuracy of CoT reasoning to quantitatively evaluate the interpretability of our LC-LLM.
- Extensive experiments on the highD dataset demonstrate that our LC-LLM outperforms all baseline models, achieving a 16.7% improvement in lane change intention prediction, a 57.2% improvement in lateral trajectory prediction, and a 48.1% improvement in longitudinal trajectory prediction, respectively.

The remainder of the study is organized as follows. Section 2 reviews previous research related to our work. Section 3 presents the LLM-based lane change intentions and trajectory prediction approach. Section 4 provides experimental results demonstrating the performance of the proposed approach. Finally, the conclusions are drawn in Section 5.

2. Related works

The literature reviews (Chen et al., 2022; Huang et al., 2022a; Mozaffari et al., 2020) discuss various methods for predicting the future states of surrounding vehicles in autonomous driving systems. Depending on the output type, these prediction studies can be classified into two categories: intention prediction and trajectory prediction (Mozaffari et al., 2020). Concurrently, LLMs have shown significant advancements in recent years. Comprehensive overviews of LLM applications in autonomous driving are discussed in Yang et al. (2023), highlighting LLMs' potential to enhance both the transparency and accuracy of autonomous driving prediction systems.

2.1. Intention prediction

Intention prediction is the task of forecasting the high-level behaviors that the surrounding vehicles intend to perform in the upcoming time steps (Mozaffari et al., 2020). Research in this field focuses on accurately predicting the intentions of surrounding vehicles to change lanes before the maneuver is executed, thereby enhancing the autonomous driving system's ability to make better decisions about its own trajectory. Early studies primarily used machine learning techniques, such as support vector machines (SVM), to detect lane change intentions based on vehicle dynamics features such as acceleration, steering angle, and distance (Lyu et al., 2020; Mandalia and Salvucci, 2005). Additionally, integrated learning models like XGBoost have been employed to combine multiple external features, such as various traffic conditions, to improve prediction performance (Xue et al., 2022; Zhang et al., 2022). In parallel, probabilistic models like Dynamic Bayesian Networks (DBN) (He et al., 2019) were introduced as intention prediction models to better handle uncertainty in the driving environment. As the complexity of driving environments increases, more sophisticated deep learning approaches, including convolutional neural networks (CNNs) (Hong et al., 2019; Mozaffari et al., 2022), Long Short-Term Memory (LSTM) Networks (Xin et al., 2018; Zyner et al., 2017, 2018), and Transformers (Gao et al., 2023), became prevalent for intention prediction. These methods focus on capturing the complex dynamics of driving behaviors. For example, Hong et al. (2019) and Mozaffari et al. (2022) utilized CNNs as feature extractors to fuse complex contextual information from the driving environment, improving the system's ability to predict lane change intentions from rich visual and sensory inputs. On the other hand, LSTM was employed to handle sequential data. Xin et al. (2018) implemented a dual-block LSTM architecture, where the first LSTM block processes sequential trajectory data to recognize driver intentions as an intermediate indicator. Similarly, Izquierdo et al. (2019) utilized a hybrid CNN-LSTM model to capture both local and global contextual features, as well as temporal information, to forecast vehicle lane change intentions. More recently, Gao et al. (2023) proposed a Transformer-based intention prediction model to extract social interactions and vehicle states, demonstrating the potential of attention mechanisms to model complex interactions in multi-agent environments. Despite these advancements, existing methods still face challenges in accurately predicting lane change intentions well in advance, particularly in highly dynamic and uncertain driving scenarios, due to the inability of these methods to combine more traffic domain knowledge for intention prediction reasoning.

2.2. Trajectory prediction

Trajectory prediction refers to forecasting a series of future locations of surrounding vehicles over a time window (Mozaffari et al., 2020). Research on trajectory prediction in autonomous driving systems is a critical area of study, aiming to enhance the safety of self-driving vehicles by forecasting the future trajectory points of surrounding vehicles. Early approaches employed LSTM networks, with one LSTM serving as an encoder to extract features from historical trajectories and another LSTM acting as a decoder to predict future trajectories (Altché and de La Fortelle, 2017; Deo and Trivedi, 2018; Xin et al., 2018). These LSTM-based approaches successfully modeled the temporal dynamics of vehicle movement but were still limited in modeling complex interactions and spatial relationships among vehicles and the environment. Subsequently, several studies focused on trajectory prediction by utilizing bird's-eye view images of traffic scenarios as input and applying CNNs to process the rasterized scene data (Chai et al., 2020; Cui et al., 2019), while they suffered from poor performance with sparse data. More recent works represented scenes with vectorized data, such as points and polylines, and processed them with GNNs (Diehl et al., 2019; Gao et al., 2020; Jeon et al., 2020; Liang et al., 2020; Mo et al., 2022, 2024; Zhao et al., 2021) or Transformers (Gao et al., 2023; Huang et al., 2022b; Shi et al., 2022). For example, VectorNet (Gao et al., 2020) and TNT (Zhao et al., 2021) used GNNs to extract spatial relationships and interaction features from vectorized high-definition (HD) maps and historical vehicle trajectories, thereby avoiding lossy rendering and computationally intensive CNN encoding. However, these approaches still overlook the edge attributes between vehicle nodes, which may limit their ability to fully capture the complex interactions between vehicles. To address this limitation, Mo et al. (2022) proposed a heterogeneous edge-enhanced graph attention network that incorporates edge attributes, enabling the model to better capture interaction features. Meanwhile, Shi et al. (2022) introduced the Motion Transformer framework, which treats trajectory prediction as a joint optimization problem involving both global intention localization and local movement refinement. More recently, Seff et al. (2023) framed multi-agent motion prediction as a language modeling task by representing continuous trajectories as sequences of discrete motion tokens. This approach is a novel attempt to leverage the power of language models in understanding the dynamics of multi-agent motion. While these deep learning-based approaches achieve competitive results, their predictions often lack interpretability, which hampers the development of safer and more transparent autonomous driving systems.

2.3. LLMs for autonomous driving

Recent advancements in LLMs have demonstrated remarkable capabilities in extensive knowledge storage, logical reasoning, and question-answering. These competencies suggest a natural extension for applying LLMs to enhance the field of autonomous driving (Yang et al., 2023). For example, Wu et al. (2023) integrated LLMs with 3D detection and tracking tasks by using language prompts as semantic cues, thereby improving the understanding of dynamic environments. In a related study, Mao et al. (2023) proposed a prompting-reasoning-finetuning strategy, enabling LLMs to generate driving trajectories and showcasing the detailed numerical reasoning capabilities of GPT-3.5 in motion planning. Meanwhile, Chen et al. (2023) designed an object-level multimodal LLM architecture that combines vectorized numeric modalities with pre-trained LLMs, thereby facilitating more accurate scene understanding and decision-making processes. Additionally, Xu et al. (2024) presented an interpretable autonomous driving system employing LLMs capable of processing multi-frame video inputs and textual queries, facilitating the interpretation of vehicle actions, offering pertinent reasoning, and effectively addressing a diverse range of questions posed by users. While LLMs have been widely explored in the domains of autonomous driving perception and planning, their application in intention prediction and trajectory prediction remains underexplored. These gaps present a promising avenue for future research, where the integration of LLMs could potentially improve prediction performance and system interpretability in autonomous driving.

3. Methodology

In this section, we introduce our LC-LLM, a model based on LLM designed for predicting lane change intentions and trajectory in autonomous driving systems. The whole pipeline of our LC-LLM is clearly illustrated in Fig. 1. As depicted in the figure, the green trajectory in the "Lane change event" is the future trajectory of the target vehicle, which the model aims to predict. For the vehicle label information, "C" for car and "T" for truck. We reconceptualize the task of predicting intentions and trajectory as a language modeling problem. To this end, we articulate observations using natural language as prompts for input into the LLM and leverage supervised fine-tuning techniques to tailor the LLM to this specific task. During the inference stage, we incorporate explanatory requirements into the prompt. As a result, our fine-tuned model, LC-LLM, not only forecasts the lane change intentions and future trajectories but also provides CoT reasoning and explanations for the predictions, thereby enhancing their interpretability.

3.1. Problem formulation

In this study, we propose a predictive model based on LLMs to simultaneously predict the lane change intention and trajectory of the target vehicle. This model serves as a prediction module within an autonomous driving system, with the goal of forecasting the future trajectories of surrounding vehicles around the ego vehicle. A representative scenario is illustrated in Fig. 2, where the ego vehicle (autonomous vehicle) is shown in yellow, and the target vehicle, one of the surrounding vehicles within the ego vehicle's vicinity, is represented in green. The future behavior of the target vehicle presents potential risks to the ego vehicle. Therefore, accurately predicting the target vehicle's lane change intention and trajectory in advance is crucial for ensuring the safe driving of the ego vehicle. The proposed predictive model aims to forecast the trajectory of the target vehicle over the next t future timesteps and to determine its lane change intentions within the same temporal horizon, as indicated by the orange line in Fig. 2.

We define the state of the target vehicle as $\boldsymbol{tv}$, the states of the surrounding vehicles as $\boldsymbol{sv}$, and the map information as $\boldsymbol{m}$, which serve as the inputs to our model. The outputs are the CoT reasoning $\mathbb{C}$, the lane change intention $\mathbb{I}$, and the future trajectory $\mathbb{T}$. The process of the model $\mathbb{F}$ can be formulated as

$$\mathbb{C}, \mathbb{I}, \mathbb{T} = \mathbb{F}(\boldsymbol{tv}, \boldsymbol{sv}, \boldsymbol{m}) \tag{1}$$

where $\mathbb{I} \in \{0, 1, 2\}$, 0 denotes "keep lane", 1 denotes "left lane change", and 2 denotes "right lane change". The future trajectory of prediction horizon t can be denoted by $\mathbb{T} = \{(x_1, y_1), (x_2, y_2), \ldots, (x_t, y_t)\}$. And the corresponding CoT reasoning $\mathbb{C}$ is composed of notable features and potential behaviors. The target vehicle state $\boldsymbol{tv}$ consists of a historical trajectory of this vehicle and its current velocity, type, etc. The surrounding vehicle states $\boldsymbol{m}$ contain each vehicle's current velocity, type, and their spatial relationship relative to the target vehicle. The map information $\boldsymbol{m}$ includes details such as lane identifiers and lane markings.

We reframe the task of predicting intentions and trajectory as a language modeling problem. By utilizing natural language to describe both the input data and output results, we can represent them as a sequence of tokens. The input data which describes the current frame driving scenario can be denoted as a sequence of tokens T_s:

$$T_s = K(\boldsymbol{tv}, \boldsymbol{sv}, \boldsymbol{m}) \tag{2}$$

The prediction result which describes the CoT reasoning, lane change intentions and trajectory can also be represented as a sequence of tokens $\{T_1, T_2, \ldots, T_n\}$:

$$\{T_1, T_2, \ldots, T_n\} = K(\mathbb{C}, \mathbb{I}, \mathbb{T}) \tag{3}$$

where K is a language tokenizer used to transform input data and output results into tokens. T_i represents the i-th token in the sequence. By adopting this language-based representation, we can redefine the prediction problem as a language modeling problem, with the loss function resembling that of language modeling (Qiu et al., 2020):

$$\mathbb{L} = -\sum_{i=1}^{n} \log P\left(T_i^* \middle| T_{<i}, T_s\right) \tag{4}$$

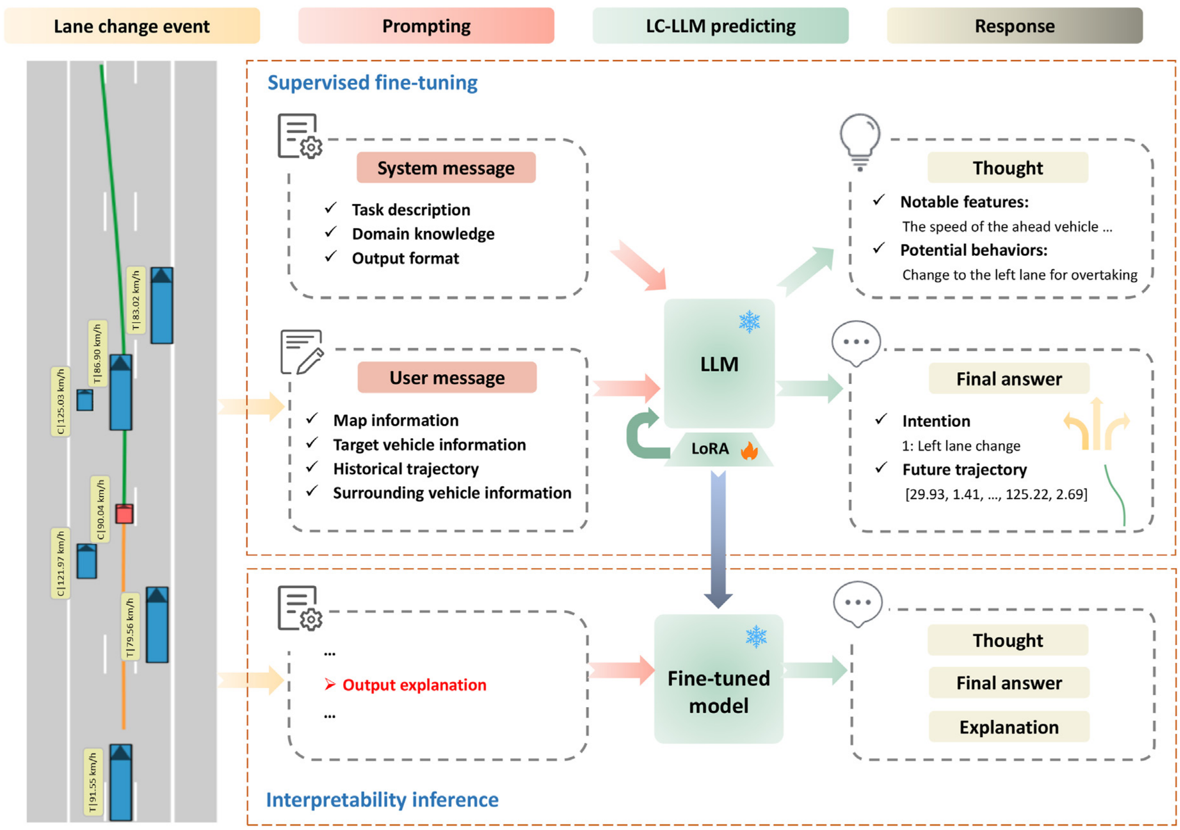

Fig. 1. Pipeline of our LC-LLM framework. Observations are input as natural language prompts into the LLM, fine-tuned to predict lane change intentions, trajectories, and CoT reasoning. During inference, prompts include explanatory requirements, enabling the model to provide interpretable predictions with reasoning and explanations.

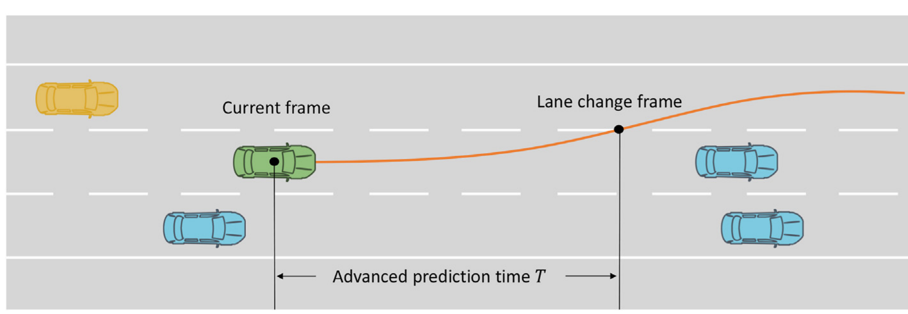

Fig. 2. Scenario description. The ego vehicle (autonomous vehicle) is depicted in yellow, while the target vehicle is shown in green and surrounding vehicles in blue. The proposed predictive model is designed to predict the future lane change trajectory of the target vehicle. The orange line illustrates the trajectory of the target vehicle in the future t timesteps. The advanced prediction time T represents the temporal interval between the current frame and the lane change frame.

where T_i^* denotes the ground truth of the next token given its historical tokens, and n represents the token number of ground truth in one sample. By maximizing the conditional probability P of the token T_i^*, our model can effectively predict lane change intentions and future trajectory.

3.2. Prompting

Generally, LLMs receive inputs in the form of natural language rather than unprocessed vectorized data. Consequently, the formulation of effective prompts that describe the current observations in natural language is critical. Several studies have sought to harness the deep reasoning capabilities of the LLMs through clever prompt design (Cui et al., 2024; Mao et al., 2023). Building on these prior efforts, we have crafted prompts that are clearer, more intelligent, and better structured. Fig. 3 provides an example of our input prompts.

As illustrated in Fig. 3, the input prompts consist of a system message displayed in the upper text block and a user message presented in the lower text block. The system message remains consistent across diverse driving scenarios. It delineates the designated role of the LLM, provides details of the coordinate system, and outlines the information and format for the LLM's output. In our study, the assigned role for the LLM is that of a predictive model integrated within an autonomous driving system. The coordinate system in each driving scenario is the vehicle coordinate system, which is centered on the current position of the target vehicle. The expected output includes predictions of lane change intentions and trajectory points over a future time horizon of 4 s, as well as thought reasoning, which consists of notable features and potential behavior.

The user message provides a description of the observations specific to the current frame, thus varying with each driving scenario. It includes information about the map, the state of the target vehicle, and the spatial

Role: You are an expert driving prediction model of an autonomous driving system, that can predict the future driving intention and future 4-s driving trajectory for a given target vehicle, avoiding collision with other vehicles and obstacles on the road.
Context:
- Coordinates: y-axis is perpendicular, and x-axis is parallel to the direction the target vehicle is facing. The target vehicle's current position is (0,0). Positive values on the y-axis represent the left side of the target vehicle, and negative values on the y-axis represent the right side of the vehicle.
Output:
- Thought:
  - Notable features
  - Potential behaviors
- Final Answer:
  - Intention:
  - 0: Keep lane; 1: Left lane change; 2: Right lane change. The final answer should be one of the three modes.
  - Trajectory (MOST IMPORTANT): 4 points, one every 1 s
  - $[(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)]$

Base

Map information

Surrounding vehicle infor

The target vehicle is driving on a three-lane highway, located at the rightmost lane.
The information of target vehicle is as follow:
  - Velocity (km/h): $vx = 90.04$, $vy = -0.47$
  - Acceleration (m/s²): $ax = 0.36$, $ay = 0.07$
  - Type: Car, with width of 1.92 m and length of 4.24 m
  - Historical position of the last 2 s (one point every 0.4 s): [(−48.46, 0.02), (−39.62, 0.03), (−29.80, 0.06), (−19.93, 0.07), (−9.98, 0.05), (0.0, 0.0)]

The information of its surrounding vehicles (with a range of 200 m) are listed as follow:
  - Ahead: a Truck traveling at 86.90 km/h of x-axis, with a distance of 21 m.
  - Behind: a Truck traveling at 91.55 km/h of x-axis, with a distance of 68 m.
  - Left front: a Car traveling at 125.03 km/h of x-axis, with a distance of 26 m.
  - Left rear: a Car traveling at 121.07 km/h of x-axis, with a distance of 13 m.
  - Right front: a Truck traveling at 83.02 km/h of x-axis, with a distance of 41 m.
  - Right rear: a Truck traveling at 79.56 km/h of x-axis, with a distance of 32 m.

Fig. 3. Example of our input prompts in the fine-tuning stage. The input prompts comprise a system message displayed in the upper text block and a user message presented in the lower text block. The user message mainly contains map information, the target vehicle's state, and the spatial relationships between the target vehicle and its surrounding vehicles.

relationships between the target vehicle and its surrounding vehicles. Map information in the prompts primarily denotes the number of lanes in the scenario and indicates whether the target vehicle is located in the leftmost, middle, or rightmost lane. Prompts related to the target vehicle's state are generated by detailing the target vehicle's historical trajectory over the past 2 s, its current velocity, and its vehicle type. Given that most vehicles do not have a large lateral displacement 4 s before changing lanes, the model mainly relies on the interaction information between the target vehicle and surrounding vehicles to predict lane change intentions in advance. Therefore, comprehending the information pertaining to surrounding vehicles is crucial for accurately predicting the lane change intentions of the target vehicle. We denote the information of the nearest vehicles in eight directions surrounding the target vehicle as the surrounding vehicle prompts. These directions include ahead, left front, right front, left side, right side, rear, left rear, and right rear. The surrounding vehicle information in each direction encompasses details such as vehicle type, current speed, and relative distance from the target vehicle. This comprehensive surrounding vehicle information serves as a critical prompt for the predictive model, enabling a thorough analysis of the contextual interactions influencing the target vehicle's lane change intentions.

3.3. Reasoning

Recently, CoT reasoning has demonstrated remarkable capabilities in carrying out more intricate reasoning tasks (Kojima et al., 2022; Mao et al., 2023; Wei et al., 2022; Zhou et al., 2022). Drawing inspiration from GPT-Driver (Mao et al., 2023), which utilizes a CoT reasoning strategy in motion planning tasks to enhance transparency throughout the planning procedure, we employ CoT reasoning in lane change prediction tasks to improve the reliability of predictions and explanations. Unlike GPT-Driver, which only labels thoughts based on whether the ego vehicle and surrounding vehicles will collide in the future, we also consider domain knowledge in driving, traffic rules, and traditional lane change model rules (Gipps, 1986; Hidas, 2002) for labeling CoT reasoning. In this way, our LC-LLM model can learn this knowledge and these rules to provide more reliable and accurate predictions and explanations. The details of CoT reasoning are shown in Fig. 4.

Our CoT reasoning consists of notable features and potential behaviors. The annotation of CoT reasonings are as follows.

3.3.1. Labeling of notable features

Firstly, we labeled notable features such as significant lateral movement when lateral velocity exceeds 1.5 km/h, and high longitudinal acceleration when it surpasses 0.5 m/s^2. Secondly, the relative speeds of surrounding vehicles compared to the target vehicle in the directions ahead, left front, and right front are considered notable features. Vehicles in these positions with a speed higher than the target vehicle are labeled as free. Conversely, vehicles with a speed lower than the target vehicle are labeled as blocked, indicating potential congestion or the need for the target vehicle to overtake. Lastly, the presence of a truck ahead within 100 m is also notable due to its impact on traffic flow and overtaking maneuvers (Gipps, 1986). Additionally, in right lane change scenarios, the characteristic of the target vehicle being a truck is also labeled as notable due to the traffic rule that trucks typically travel in the right lane.

3.3.2. Labeling of potential behaviors

Potential behaviors are classified into eight categories: "Change to the left lane for overtaking" occurs when encountering a slower vehicle ahead and occupying the rightmost or middle lane; "Change left to the fast lane" corresponds to significant acceleration from the target vehicle; "Irregular left lane change" applies to scenarios in the highD dataset that involve a left lane change but do not fit the first two categories; "Change to the right lane for overtaking" is when the target vehicle occupies the leftmost or middle lane and the lane ahead is blocked; "Change right to the slow lane" is associated with significant deceleration from the target

Thought:
- Notable feature: the speed of the ahead vehicle is 86.90 km/h, slower than that of the target's, ahead is block.
- Notable feature: the type of ahead vehicle is "Truck ".
- Notable feature: left front is free.
- Notable feature: the speed of the right front vehicle is 83.02 km/h, slower than that of the target's.
- Potential behavior: change to the left lane for overtaking.

Final answer:
- Intention: "1: Left lane change"
- Trajectory: "[(25.21, −0.06), (50.75, 0.24), (76.69, 1.49), (103.26, 2.69)]"

Thought
Final answer

Fig. 4. Example of the output of our LC-LLM. The content in the yellow text block is the CoT reasoning.

vehicle or when the target vehicle is a truck; “Irregular right lane change” covers right lane change scenarios in the highD dataset that do not fit the first two rightward categories; “Following and keep lane” occurs when following a slower vehicle ahead; “Normal keep lane” is when the lane ahead is unobstructed and it is a lane keeping scenario in the highD dataset. These categories assist in predicting driving patterns and providing reliable explanations.

3.4. Fine-tuning

In our work, we utilize an open-source foundational language model, Llama-2-13b-chat (Touvron et al., 2023), as our pre-trained LLM. To achieve parameter-efficient fine-tuning, we adopt the Low-Rank Adaptation (LoRA) (Hu et al., 2021) strategy, which freezes the pre-trained model weights and injects trainable rank decomposition matrices into each layer of the Transformer architecture. Additionally, we customize the LLM for our specific prediction task using supervised fine-tuning techniques. The raw dataset utilized in our work originates from the highD dataset (Krajewski et al., 2018), which captures human naturalistic vehicle trajectories on German highways. Following the fine-tuning instructions provided for Llama-2, each data sample is formatted to include an input prompt and a corresponding answer, separated by a special token. This formatting is illustrated in Fig. 5. The answer for each sample is derived from the ground truth obtained from the highD dataset, encompassing programmatically labeled future trajectories, driving intentions, and CoT reasoning. In summary, we process the highD raw dataset using natural language, format each sample into the Llama format, and then feed it into LLM. Finally, we fine-tune the LLM by aligning the LLM's output $\{\mathbb{C}, \mathbb{I}, \mathbb{T}\}$ with the corresponding ground truth labels $\{\mathbb{C}^*, \mathbb{I}^*, \mathbb{T}^*\}$. This fine-tuning process is executed through the language modeling loss $\mathbb{L}$ as defined in Eq. (4). During fine-tuning, we mask the loss for tokens within the input prompts, focusing backpropagation only on the tokens comprising the answers, in alignment with the fine-tuning approach of Llama-2. By following this approach, our fine-tuned model, LC-LLM, can predict human driving behaviors and corresponding CoT reasoning. This capability is crucial for ensuring the safety of ego vehicles within an autonomous driving system.

3.5. Interpretability of prediction

A prevalent limitation in contemporary autonomous driving prediction models lies in their constrained interpretability. This is attributed to the fact that these models generate predictions about the future behaviors or trajectories of the target vehicle through black-box neural networks, offering little explanation for their prediction results. In our research, we address this limitation in two distinct ways. First, we fine-tune the CoT reasoning capabilities of our LC-LLM. By doing so, our model learns the domain knowledge and rules of driving, which enhances the transparency and reliability of lane change predictions, leading to more accurate and explainable predictions. Second, we incorporate explanatory requirements into the input prompts during the inference stage, where the weights of our fine-tuned model are fixed. Benefiting from the self-explanation (Huang et al., 2023; Rajani et al., 2019) capabilities of LLMs, our fine-tuned model, LC-LLM, not only predicts lane change intentions and future trajectories but also provides corresponding CoT reasoning and explanations for its prediction results, thereby enhancing the interpretability of these results. This approach facilitates a more transparent and understandable prediction process, which is crucial for the practical application of autonomous driving systems. Some visual examples of the interpretability of our LC-LLM model are shown in Fig. 6.

4. Experimental results and analysis

This section outlines the evaluations of our LC-LLM model. First, we detail the dataset processing and the experimental setup. Subsequently, we elucidate the evaluation metrics. Next, we compare the proposed model with the baseline models and analyze the results quantitatively. Moreover, we perform an ablation study on the key components of our proposed method. Finally, we conduct an interpretability study and a zero-shot study to evaluate the interpretability and effectiveness of our approach.

4.1. Dataset processing

To evaluate the performance of the model proposed in this study, the highD dataset was used for both training and testing. The highD dataset is a large-scale natural vehicle trajectory dataset collected on German highways. It encompasses 16.5 h of data from six distinct locations, involving 110,000 vehicles, covering a total distance of 45,000 km, and capturing 5600 documented complete lane changes. The highD dataset comprises data extracted from 60 recordings. For our study, we selected data from the first 50 recordings for training the model, while the remaining 10 recordings were reserved for testing. Furthermore, we adopt a target-centric strategy that normalizes all position inputs to the coordinate system centered on the current frame position of the target vehicle.

```
<s>[INST] <<SYS>>
system massage
<</SYS>>

user massage [/INST] answer massage </s>
```

Fig. 5. Data sample of Llama format. Special tokens, such as [INST], serve as delimiters distinguishing system messages, user messages, and answer messages. The details of these messages are elucidated in the Prompting section.

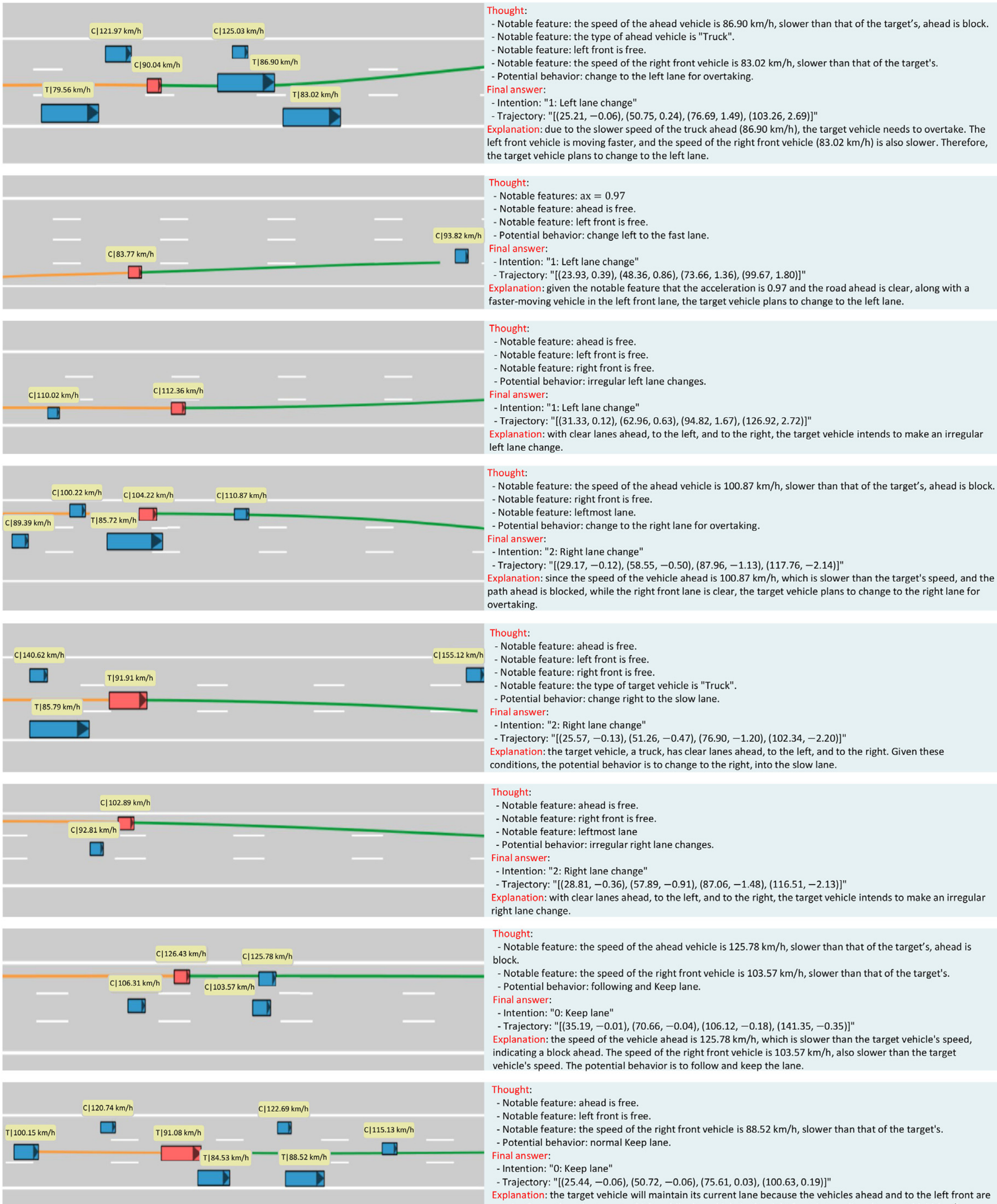

Fig. 6. Some visual examples demonstrating the interpretability of our LC-LLM. In each example, the image on the left depicts the driving scenario, and the text on the right presents the corresponding output of our model, which includes the prediction results and their reasoning and explanations.

In our work, the lane change frame T_{lc} is defined as the frame in which the lane id changes. The advanced prediction time T denotes the temporal interval between the lane change frame and the current frame, and it can be expressed as $T = T_{lc} - T_{current}$, as illustrated in Fig. 2. For the purpose of training and evaluating our proposed LC-LLM, we extract both lane change (LC) and lane keeping (LK) scenarios from the highD dataset. The LK scenarios comprise data samples wherein the lane id remains constant throughout, while the LC scenarios consist of data samples where the advanced prediction time T falls within the range of [0 s, 4 s]. Previous works have primarily focused on predicting lane change behavior within a 4-s advance time (Xing et al., 2019), and if the advanced prediction time is too long, it becomes difficult for downstream decision systems to determine how to react to the potential lane change (Song and Li, 2022). Therefore, we selected 4 s as a reasonable advanced prediction time. Furthermore, in order to evaluate the performance of our LC-LLM in different T, we divide LC scenarios into four parts according to the four intervals of T ($T \in [0, 1]$, $T \in (1, 2]$, $T \in (2, 3]$, $T \in (3, 4]$). For the training dataset, we randomly select 48,000 lane keeping (LK) samples, 12,000 left lane change (LLC) samples from each T interval, and 12,000 right lane change (RLC) samples from each T

interval, resulting in a total of 144,000 samples. Similarly, for the testing dataset, we randomly select 8000 LK samples, 2000 LLC samples from each T interval, and 2000 RLC samples from each T interval, accumulating a total of 24,000 samples.

4.2. Experimental setup

We utilize the training dataset extracted from the highD dataset to train our proposed LC-LLM, along with reimplemented baseline models: an LSTM-based model (Xin et al., 2018), which employs a dual LSTM network for intention and trajectory prediction; a Transformer-based model (Gao et al., 2023), which utilizes a dual Transformer module for intention and trajectory prediction; and a GNN-based model (Mo et al., 2022), which incorporates GNNs to extract vehicles interaction features, followed by a Transformer module for trajectory prediction. All models are implemented using the PyTorch framework and leverage information from the target vehicle, surrounding vehicles and their interaction. Additionally, we employ the DeepSpeed (Rasley et al., 2020) library for distributed training of our LC-LLM. The training process is conducted on eight A800 GPUs and takes 11 h for the 13B model. The hyperparameters used for model training are detailed in Table 1. To ensure the stability of the experimental results, each experiment is repeated five times, and the average value is reported as the final result.

4.3. Evaluation metrics

Similar to previous works, we utilize precision, recall, F1 score, and macro average metrics to evaluate the performance of the intention prediction task. Additionally, we use the Root Mean Square Error (RMSE) metric to evaluate the performance of the trajectory prediction task.

4.3.1. Evaluation metrics of intention prediction

Precision: The ratio of correctly predicted positive instances to the total predicted positives.

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}} \tag{5}$$

Recall: The ratio of correctly predicted positive instances to the total actual positives.

$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}} \tag{6}$$

F1 Score: The harmonic mean of precision and recall.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

Macro Avg: The average precision, recall, and F1 score calculated across all classes. It is calculated as the arithmetic mean of precision, recall, and F1 score.

4.3.2. Evaluation metrics of trajectory prediction

RMSE: A measure of the average deviation between predicted and actual values.

Table 1
Hyperparameters for training our LC-LLM.

Hyper-parameter	Value
Learning rate	5×10^{-4}
Batch size	8
Training epochs	2
LoRA rank, r	64
LoRA alpha, α	16
Gradient accumulation steps	8
Warmup steps	600
Load in 8-bit	✓

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{8}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of instances. In trajectory prediction evaluation, we assess the RMSE in both the lateral and longitudinal directions, denoted as RMSE (lat) and RMSE (lon), respectively.

4.4. Overall performance

We evaluate the performance of our proposed LC-LLM on two tasks: intention prediction and trajectory prediction. The evaluation results are presented as follows.

4.4.1. Intention prediction results and analysis

Table 2 shows the overall performance of the proposed LC-LLM as well as baseline models in intention prediction with four different advanced prediction times T. We can note that our proposed LC-LLM significantly outperforms baseline models for all T, especially for the longer T. For example, when $T \in (2, 3]$, the average F1 score of our LC-LLM is 9.1% higher than that of the GNN model, 11.9% higher than that of the Transformer model and 14.9% higher than that of the LSTM model. Furthermore, when $T \in (3, 4]$, our LC-LLM surpasses the GNN model by 16.7%, the Transformer model by 20.3%, and the LSTM model by 25.0% in terms of average F1 score. These results highlight the superior performance of our LC-LLM, even during the early stages of a lane change maneuver when lateral displacement is minimal.

4.4.2. Trajectory prediction results and analysis

We assessed the performance of our LC-LLM in the trajectory prediction task and compared them with baseline models. The results are depicted in Fig. 7. The left figure illustrates the trends of lateral RMSE for trajectory points across different models and prediction horizons, while the right figure shows the trends of longitudinal RMSE. The results indicate a notable improvement in both lateral and longitudinal RMSE for our LC-LLM compared to the baselines. Notably, with larger prediction horizons, our model demonstrates increasing advantages over the baselines in both lateral RMSE and longitudinal RMSE, indicating better robustness with respect to long-term prediction horizons.

The results of the lane change intention prediction task and the trajectory prediction task demonstrate that our LC-LLM outperforms all baseline models in long-term (4 s) prediction. Specifically, it achieves a 16.7% improvement in lane change intention prediction, a 57.2% improvement in lateral trajectory prediction, and a 48.1% improvement in longitudinal trajectory prediction.

4.5. Ablation study

4.5.1. Component ablation studies

We explored the impact of incorporating different components on the trajectory prediction task, and the results are illustrated in Fig. 8. A series of ablation experiments were conducted, progressively introducing different components: Target Vehicle Information, Map Information, Surrounding Vehicle Information, and CoT Reasoning. The comprehensive model, integrating all contextual prompts and the CoT reasoning component, exhibited superior performance compared to the base model with only Target Vehicle Information, achieving a substantial 10.64% reduction in lateral RMSE, and a remarkable 29.57% decrease in longitudinal RMSE. Specifically, the addition of Surrounding Vehicle Information proved crucial for future trajectory forecasting. Furthermore, the incorporation of CoT reasoning yielded notable enhancements in trajectory prediction.

4.5.2. Multi-task ablation study

We further examined the efficacy of multi-task learning in

Table 2
Comparison of the intention prediction performance of the proposed LC-LLM and baseline models on the highD dataset.

Model	Intention	$T \in$ [0,1]			$T \in$ (1, 2]			$T \in$ (2,3]			$T \in$ (3,4]			Avg. ($T \in$ [0,4])		
		P (%)	*R* (%)	F1 (%)	*P* (%)	*R* (%)	F1 (%)	*P* (%)	*R* (%)	F1 (%)	*P* (%)	*R* (%)	F1 (%)	*P* (%)	*R* (%)	F1 (%)
LSTM	LK	90.9	97.4	94.0	85.7	97.1	91.0	70.6	96.4	81.5	57.7	96.8	72.3	73.9	96.9	83.8
	LLC	98.2	90.3	94.1	97.9	85.8	91.4	96.9	69.3	80.8	94.0	50.5	65.8	97.1	74.0	84.0
	RLC	99.0	99.9	99.5	98.5	97.6	98.1	98.1	90.2	94.0	97.0	76.1	85.3	98.2	91.0	94.5
	Macro avg.	96.0	95.9	95.9	94.0	93.5	93.5	88.5	85.3	85.4	82.9	74.5	74.4	89.7	87.3	87.4
Transformer	LK	97.6	96.5	97.1	90.8	96.2	93.4	74.6	95.9	83.9	61.1	95.8	74.6	78.4	96.1	86.1
	LLC	97.9	98.1	98.0	97.4	91.3	94.3	96.5	76.1	85.1	93.9	58.2	71.9	96.7	80.9	88.1
	RLC	98.6	99.6	99.1	98.1	98.4	98.3	97.8	90.6	94.1	95.3	77.4	85.4	97.6	91.5	94.4
	Macro avg.	98.0	98.1	98.0	95.4	95.3	95.3	89.7	87.6	87.7	83.4	77.1	77.3	90.9	89.5	89.6
GNN	LK	96.3	96.1	96.2	94.3	96.9	95.6	79.7	95.5	86.9	63.8	96.8	76.9	81.3	96.3	88.2
	LLC	97.5	96.3	96.9	98.3	94.3	96.3	95.1	80.2	87.0	93.5	60.6	73.5	96.4	82.8	89.1
	RLC	98.5	100.0	99.3	98.0	99.3	98.6	97.7	93.8	95.7	97.5	81.5	88.8	98.0	93.6	95.7
	Macro avg.	97.4	97.4	97.4	96.9	96.8	96.8	90.9	89.8	89.9	85.0	79.6	79.7	91.9	90.9	91.0
LC-LLM	LK	99.6	96.4	97.9	99.9	97.1	98.5	98.7	96.7	97.7	85.9	96.6	90.9	95.6	96.7	96.2
	LLC	97.9	99.4	98.6	98.6	99.6	99.1	97.7	98.7	98.2	97.1	89.3	93.0	97.8	96.7	97.3
	RLC	98.1	99.8	99.0	98.2	100.0	99.1	98.0	99.0	98.5	97.2	92.9	95.0	97.9	97.9	97.9
	Macro avg.	98.5	98.5	**98.5**	98.9	98.9	**98.9**	98.1	98.1	**98.1**	93.4	92.9	**93.0**	97.1	97.1	**97.1**

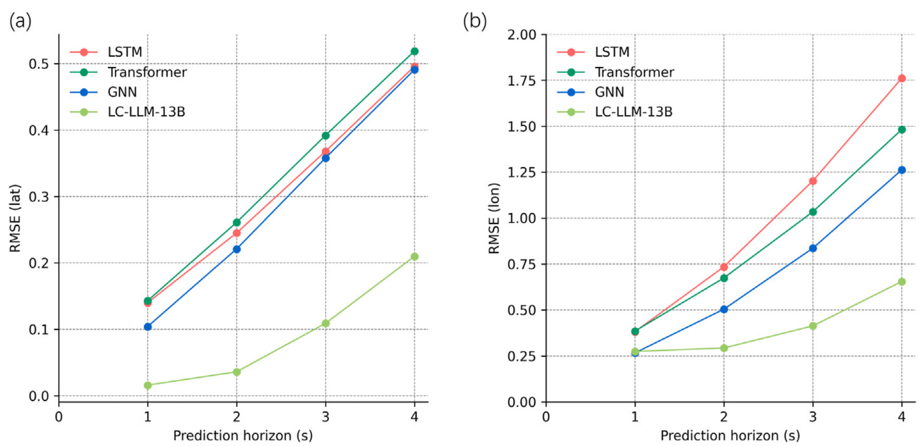

Fig. 7. Comparison of the trajectory prediction performance of the proposed LC-LLM and baseline models on the highD dataset. Results for (a) RMSE (lat) and (b) RMSE (lon).

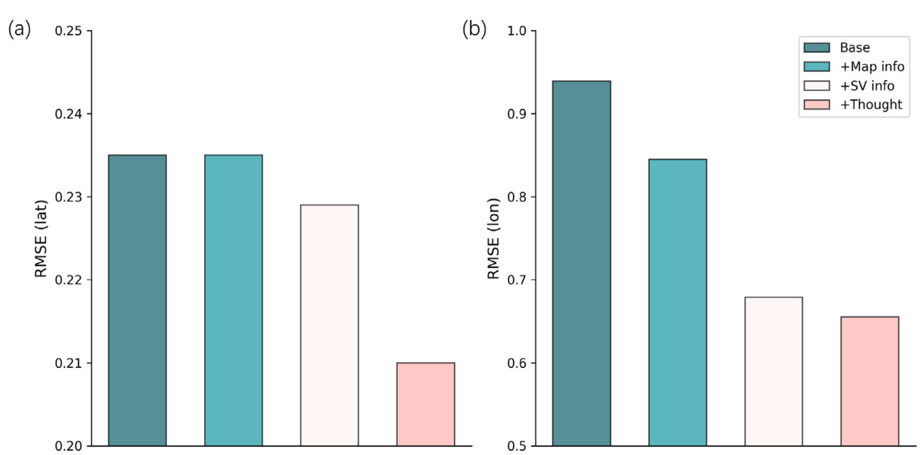

Fig. 8. Component ablation studies for 4-s prediction performance. 'Base' refers to the model results obtained using only the target vehicle information for prompting. '+Map info', '+SV info', and '+Thought' indicate the model results when map information, surrounding vehicle information, and CoT reasoning are sequentially incorporated, respectively. Results for (a) RMSE (lat) and (b) RMSE (lon).

concurrently predicting lane change intentions and future trajectories, as depicted in Table 3. The findings demonstrate superior performance in the multi-task setting compared to the isolated trajectory prediction task, emphasizing the benefits of integrating high-level intention prediction to achieve more accurate trajectory predictions. Furthermore, in comparison to the standalone intention prediction task, adding trajectory prediction not only elevates the overall task complexity but also enables the model to capture more nuanced details, consequently, enhancing the accuracy of intention prediction.

Table 3
Multi-task ablation study.

Task	Intention result	Trajectory result	
	F1 (avg)	RMSE (lat)	RMSE (lon)
Only intention	0.948	—	—
Only trajectory	—	0.257	0.798
Intention + trajectory	**0.971**	**0.210**	**0.655**

Table 4
Zero-shot study.

Method	Intention prediction		Trajectory prediction		
	F1 (avg)	No. of failed cases	RMSE (lat)	RMSE (lon)	No. of failed cases
Zero-shot (Llama-13B-chat)	0.210	4	1.405	63.366	4
Fine-tune (LC-LLM)	**0.971**	**0**	**0.210**	**0.655**	**0**

4.6. Interpretability study

4.6.1. Visualization of results

In order to demonstrate the interpretability of our LC-LLM, we visualize several sample scenarios alongside the corresponding outputs from our model, as depicted in Fig. 6. The visualization illustrates that our LC-LLM is capable of forecasting lane change intentions and future trajectories of the target vehicle while also providing corresponding CoT reasoning and explanations for the prediction results. For example, in the first case in the figure, with a slow-moving truck positioned ahead, the target vehicle needs to overtaking to gain a speed advantage. And given the high velocity of the left front vehicle and the low velocity of the right front vehicle, the target vehicle has an opportunity to execute a left lane change. The explanation provided by our LC-LLM is consistent with this driving scenario. These explanations provided by our model for these scenarios serve as evidence that our LC-LLM possesses a deep understanding of the driving scenario and makes reasonable and reliable predictions for the target vehicle. In contrast to previous methodologies that solely generate prediction results, our LC-LLM model exhibits better interpretability.

4.6.2. Quantitative evaluation

To quantitatively evaluate the interpretability of our LC-LLM, we evaluate the accuracy of CoT reasoning. We randomly selected 100 samples from the test dataset and manually compared the accuracy of the CoT reasoning output by our LC-LLM with the ground truth in the dataset, scoring them accordingly. Scoring rule is as follows: 10 points are deducted for each error, omission, or addition of a notable feature, and 50 points are deducted for a potential behavior error. Our quantitative evaluation result shows that our model scores 97.2. This result demonstrates that our LC-LLM excels in generating accurate and interpretable predictions, confirming the effectiveness of our methods in enhancing the transparency and reliability of autonomous driving systems.

4.7. Zero-shot study

Table 4 presents a comparative analysis between zero-shot and fine-tuned experiments across two tasks. The fine-tuned model demonstrated substantial improvements, yielding markedly higher F1 scores and significantly lower RMSE compared to the zero-shot experiment. These results underscore the efficacy of model fine-tuning in incorporating domain-specific knowledge into LLMs. This affirms the suitability of fine-tuning as the preferred approach for enabling LLMs to excel in specific domains in future applications.

4.8. Robustness evaluations

As mentioned in Feng et al. (2023), safety-critical events are important for testing the model's robustness. In this experiment, we evaluated the robustness of our LC-LLM by creating four safety-critical scenarios not

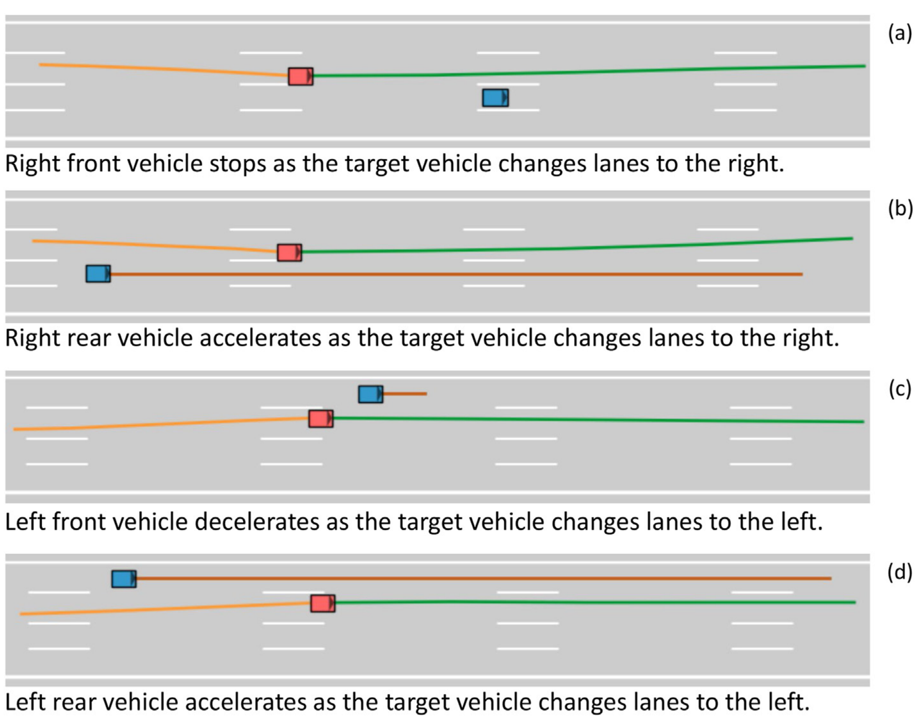

Fig. 9. Some visual examples evaluating the robustness of our LC-LLM. These test samples are from outside the distribution of the highD dataset. (a) Right front vehicle stops as the target vehicle changes lanes to the right; (b) right rear vehicle accelerates as the target vehicle changes lanes to the right; (c) left front vehicle decelerates as the target vehicle changes lanes to the left; (d) left rear vehicle accelerates as the target vehicle changes lanes to the left.

included in the highD dataset. The first scenario involves a left lane change with the left front performing emergency braking, with speeds ranging from 0 to 50 km/h in 10 km/h increments and relative distances from 10 to 100 m in 10-m increments, resulting in 60 samples. The second scenario tests a right lane change with the right front vehicle emergency braking, using the same speed and distance ranges, also resulting in 60 samples. The third scenario involves a left lane change with the left rear vehicle accelerating, with speeds of 100–150 km/h in 10 km/h increments and the same distance range, resulting in 60 samples. The fourth scenario examines a right lane change with the right rear vehicle accelerating, using the same speed and distance ranges as the third scenario, producing 60 samples.

The results, depicted in Fig. 9, highlight the robustness of our LC-LLM across these four challenging scenarios. In scenario (a) in Fig. 9, where the right front vehicle stops as the target vehicle changes lanes to the right, the LC-LLM accurately predicted the necessary adjustments to avoid a collision. Scenario (b) in Fig. 9 shows the right rear vehicle accelerating as the target vehicle changes lanes to the right; the LC-LLM successfully anticipated the increased speed of the rear vehicle and adjusted the trajectory accordingly. In scenario (c) in Fig. 9, the model handled the left front vehicle's deceleration effectively during the target vehicle's left lane change, maintaining safe distances and smooth transitions. Finally, in scenario (d) in Fig. 9, the LC-LLM demonstrated its ability to manage the left rear vehicle's acceleration during the target vehicle's left lane change, ensuring a safe maneuver. These results affirm the mode's capability to handle out-of-distribution cases, enhancing its applicability in real-world autonomous driving scenarios.

4.9. Limitations

Our LC-LLM, while demonstrating significant improvements in lane change intention and trajectory prediction accuracy and interpretability, has limitations. Firstly, our model has been tested exclusively on the highD dataset, which primarily comprises highway scenarios. To ensure the robustness and generalizability of our model, it is essential to evaluate its performance on more diverse datasets, such as NuScenes (Caesar et al., 2020) and Waymo (Ettinger et al., 2021), which encompass a wider variety of complex urban and suburban traffic scenarios. While incorporating detailed map information from these complex urban scenarios into input prompts presents a considerable challenge, the development of an appropriate prompt that accurately describes high-definition maps remains an objective for our future work. Secondly, our LC-LLM exhibits a slower inference speed compared to the baseline models. Specifically, the inference time for LC-LLM-13B is 0.83 s, which is considerably longer than the 0.34, 0.44, and 0.5 ms required by the LSTM, Transformer, and GNN models, respectively. This increased inference time is primarily attributed to the larger number of model parameters, which poses a challenge for real-time applications. However, several techniques could mitigate this issue, such as knowledge distillation (Gu et al., 2024), parameter pruning (Frantar and Alistarh, 2023; Liu et al., 2023b), and post-training quantization (Liu et al., 2023a). As demonstrated in recent studies (Wan et al., 2023), these methods have the potential to optimize the model for faster inference.

5. Conclusions

In this study, we introduce LC-LLM, an explainable lane change prediction model that not only forecasts lane change intentions and trajectories but also provides CoT reasoning and explanations for its predictions. We reconceptualize the task of lane change prediction as a language modeling problem and employ a supervised fine-tuning technique to tailor the LLM specifically for this task. In this way, we successfully leverage the strong common sense reasoning capabilities and self-explanation abilities of the LLM to address the challenge of lane change prediction. Our extensive experiments conducted on the highD dataset show that our LC-LLM improves the accuracy of predicting lane change intention and trajectory as well as significantly enhances the interpretability of the prediction results. Future work includes extending our approach to urban driving scenarios, reducing inference times through knowledge distillation to improve model speed and efficiency, and developing methods to predict lane change intentions and trajectories for multiple vehicles simultaneously.

CRediT authorship contribution statement

Mingxing Peng: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Xusen Guo:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Conceptualization. **Xianda Chen:** Writing – review & editing, Writing – original draft, Data curation. **Kehua Chen:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Meixin Zhu:** Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **Long Chen:** Project administration, Investigation. **Fei-Yue Wang:** Conceptualization.

Replication and data sharing

The dataset and model fine-tuning code are available at https://github.com/Pemixing/LCLLM.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (No. 52302379), Guangdong Provincial Natural Science Foundation-General Project (No. 2024A1515011790), Guangdong Province General Universities Youth Innovative Talents Project (No. 2023KQNCX100), Guangzhou Municipal Science and Technology Project (No. 2023A03J0011), and Nansha District Key R&D Project (No. 2023ZD006).

References

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., et al., 2023. GPT-4 technical report. https://doi.org/10.48550/arXiv.2303.08774.

Altché, F., de La Fortelle, A., 2017. AnLSTM network for highway trajectory prediction. In: 2017 IEEE 20th International Conference on Intelligent Transportation Systems, pp. 353–359.

Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., et al., 2020. nuscenes: a multimodal dataset for autonomous driving. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11621–11631.

Chai, Y., Sapp, B., Bansal, M., Anguelov, D., 2020. MultiPath: multiple probabilistic anchor trajectory hypotheses for behavior prediction. In: Proceedings of the Conference on Robot Learning, pp. 86–99.

Chen, L., Li, Y., Huang, C., Li, B., Xing, Y., Tian, D., et al., 2022. Milestones in autonomous driving and intelligent vehicles: survey of surveys. IEEE T. Intell. Veh. 8, 1046–1056.

Chen, L., Sinavski, O., Hünermann, J., Karnsund, A., Willmott, A.J., Birch, D., et al., 2023. Driving withLLMs: fusing object-level vector modality for explainable autonomous driving. https://doi.org/10.48550/arXiv.2310.01957.

Cui, C., Ma, Y., Cao, X., Ye, W., Wang, Z., 2024. Receive, reason, and react: drive as you say, with large language models in autonomous vehicles. IEEE Intell. Transp. Syst. Mag. 16, 81–94.

Cui, H., Radosavljevic, V., Chou, F.C., Lin, T.H., Nguyen, T., Huang, T.K., et al., 2019. Multimodal trajectory predictions for autonomous driving using deep convolutional networks. In: 2019 International Conference on Robotics and Automation, pp. 2090–2096.

Deo, N., Trivedi, M.M., 2018. Multi-modal trajectory prediction of surrounding vehicles with maneuver basedLSTMs. In: 2018 IEEE Intelligent Vehicles Symposium, pp. 1179–1184.

Diehl, F., Brunner, T., Le, M.T., Knoll, A., 2019. Graph neural networks for modelling traffic participant interaction. In: 2019 IEEE Intelligent Vehicles Symposium, pp. 695–701.

Ettinger, S., Cheng, S., Caine, B., Liu, C., Zhao, H., Pradhan, S., et al., 2021. Large scale interactive motion forecasting for autonomous driving: the waymo open motion dataset. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 9710–9719.

Feng, S., Sun, H., Yan, X., Zhu, H., Zou, Z., Shen, S., et al., 2023. Dense reinforcement learning for safety validation of autonomous vehicles. Nature 615, 620–627.

Frantar, E., Alistarh, D., 2023. SparseGPT: massive language models can be accurately pruned in one-shot. In: Proceedings of the 40th International Conference on Machine Learning, pp. 10323–10337.

Gao, J., Sun, C., Zhao, H., Shen, Y., Anguelov, D., Li, C., et al., 2020. Vectornet: encoding hd maps and agent dynamics from vectorized representation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11525–11533.

Gao, K., Li, X., Chen, B., Hu, L., Liu, J., Du, R., et al., 2023. Dual transformer based prediction for lane change intentions and trajectories in mixed traffic environment. IEEE Trans. Intell. Transport. Syst. 24, 6203–6216.

Gipps, P.G., 1986. A model for the structure of lane-changing decisions. Transport Res B-meth 20, 403–414.

Gu, Y., Dong, L., Wei, F., Huang, M., 2024. MiniLLM: knowledge distillation of large language models. https://doi.org/10.48550/arXiv.2306.08543.

He, G., Li, X., Lv, Y., Gao, B., Chen, H., 2019. Probabilistic intention prediction and trajectory generation based on dynamic bayesian networks. In: 2019 Chinese Automation Congress, pp. 2646–2651.

Hidas, P., 2002. Modelling lane changing and merging in microscopic traffic simulation. Transport. Res. C Emerg. Technol. 10, 351–371.

Hong, J., Sapp, B., Philbin, J., 2019. Rules of the road: predicting driving behavior with a convolutional model of semantic interactions. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8454–8462.

Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., et al., 2021. LoRA: low-rank adaptation of large language models. https://doi.org/10.48550/arXiv.2106.09685.

Huang, S., Mamidanna, S., Jangam, S., Zhou, Y., Gilpin, L.H., 2023. Can large language models explain themselves? a study of llm-generated self-explanations. https://doi.org/10.48550/arXiv.2310.11207.

Huang, Y., Du, J., Yang, Z., Zhou, Z., Zhang, L., Chen, H., 2022a. A survey on trajectory-prediction methods for autonomous driving. IEEE T. Intell. Veh. 7, 652–674.

Huang, Z., Mo, X., Lv, C., 2022b. Multi-modal motion prediction with transformer-based neural network for autonomous driving. In: 2022 International Conference on Robotics and Automation, pp. 2605–2611.

Izquierdo, R., Quintanar, A., Parra, I., Fernández-Llorca, D., Sotelo, M., 2019. Experimental validation of lane-change intention prediction methodologies based onCNN andLSTM. In: 2019 IEEE Intelligent Transportation Systems Conference, pp. 3657–3662.

Jeon, H., Choi, J., Kum, D., 2020. SCALE-Net: scalable vehicle trajectory prediction network under random number of interacting vehicles via edge-enhanced graph convolutional neural network. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2095–2102.

Kojima, T., Gu, S.S., Reid, M., Matsuo, Y., Iwasawa, Y., 2022. Large language models are zero-shot reasoners. In: Proceedings of the 36th International Conference on Neural Information Processing Systems, pp. 22199–22213.

Krajewski, R., Bock, J., Kloeker, L., 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, pp. 2118–2125.

Liang, M., Yang, B., Hu, R., Chen, Y., Liao, R., Feng, S., et al., 2020. Learning lane graph representations for motion forecasting. In: Proceedings of the 16th European Conference on Computer Vision, pp. 541–556.

Liu, J., Gong, R., Wei, X., Dong, Z., Cai, J., Zhuang, B., 2023a. QLLM: accurate and efficient low-bitwidth quantization for large language models. https://doi.org/10.48550/arXiv.2310.08041.

Liu, Z., Wang, J., Dao, T., Zhou, T., Yuan, B., Song, Z., et al., 2023b. Deja Vu: contextual sparsity for efficientLLMs at inference time. In: Proceedings of the 40th International Conference on Machine Learning, pp. 22137–22176.

Lyu, N., Wen, J., Duan, Z., Wu, C., 2020. Vehicle trajectory prediction and cut-in collision warning model in a connected vehicle environment. IEEE Trans. Intell. Transport. Syst. 23, 966–981.

Mandalia, H.M., Salvucci, M.D.D., 2005. Using support vector machines for lane-change detection. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 1965–1969.

Mao, J., Qian, Y., Zhao, H., Wang, Y., 2023. GPT-Driver: learning to drive withGPT. https://doi.org/10.48550/arXiv.2310.01415.

Mo, X., Huang, Z., Xing, Y., Lv, C., 2022. Multi-agent trajectory prediction with heterogeneous edge-enhanced graph attention network. IEEE Trans. Intell. Transport. Syst. 23, 9554–9567.

Mo, X., Xing, Y., Lv, C., 2024. Heterogeneous graph social pooling for interaction-aware vehicle trajectory prediction. Transp. Res. E: Logist. Transp. Rev. 191, 103748.

Mozaffari, S., Al-Jarrah, O.Y., Dianati, M., Jennings, P., Mouzakitis, A., 2020. Deep learning-based vehicle behavior prediction for autonomous driving applications: a review. IEEE Trans. Intell. Transport. Syst. 23, 33–47.

Mozaffari, S., Arnold, E., Dianati, M., Fallah, S., 2022. Early lane change prediction for automated driving systems using multi-task attention-based convolutional neural networks. IEEE T. Intell. Veh. 7, 758–770.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., et al., 2022. Training language models to follow instructions with human feedback. In: Proceedings of the 36th International Conference on Neural Information Processing Systems, pp. 27730–27744.

Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N., Huang, X., 2020. Pre-trained models for natural language processing: a survey. Sci. China Technol. Sci. 63, 1872–1897.

Rajani, N.F., McCann, B., Xiong, C., Socher, R., 2019. Explain yourself! leveraging language models for commonsense reasoning. https://doi.org/10.48550/arXiv.1906.02361.

Rasley, J., Rajbhandari, S., Ruwase, O., He, Y., 2020. Deepspeed: system optimizations enable training deep learning models with over 100 billion parameters. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 3505–3506.

Seff, A., Cera, B., Chen, D., Ng, M., Zhou, A., Nayakanti, N., et al., 2023. MotionLM: multi-agent motion forecasting as language modeling. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 8579–8590.

Shi, S., Jiang, L., Dai, D., Schiele, B., 2022. Motion transformer with global intention localization and local movement refinement. In: Proceedings of the 36th International Conference on Neural Information Processing Systems, pp. 6531–6543.

Song, R., Li, B., 2022. Surrounding vehicles' lane change maneuver prediction and detection for intelligent vehicles: a comprehensive review. IEEE Trans. Intell. Transport. Syst. 23, 6046–6062.

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., et al., 2023. Llama 2: open foundation and fine-tuned chat models. https://doi.org/10.48550/arXiv.2307.09288.

Wan, Z., Wang, X., Liu, C., Alam, S., Zheng, Y., Liu, J., et al., 2023. Efficient large language models: a survey. https://doi.org/10.48550/arXiv.2312.03863.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., et al., 2022. Chain-of-thought prompting elicits reasoning in large language models. In: Proceedings of the 36th International Conference on Neural Information Processing Systems, pp. 24824–24837.

Wu, D., Han, W., Wang, T., Liu, Y., Zhang, X., Shen, J., 2023. Language prompt for autonomous driving. https://doi.org/10.48550/arXiv.2309.04379.

Xin, L., Wang, P., Chan, C.Y., Chen, J., Li, S.E., Cheng, B., 2018. Intention-aware long horizon trajectory prediction of surrounding vehicles using dualLSTM networks. In: 2018 21st International Conference on Intelligent Transportation Systems, pp. 1441–1446.

Xing, Y., Lv, C., Wang, H., Wang, H., Ai, Y., Cao, D., et al., 2019. Driver lane change intention inference for intelligent vehicles: framework, survey, and challenges. IEEE Trans. Veh. Technol. 68, 4377–4390.

Xu, Z., Zhang, Y., Xie, E., Zhao, Z., Guo, Y., Wong, K.Y.K., et al., 2024. DriveGPT4: interpretable end-to-end autonomous driving via large language model. IEEE Rob. Autom. Lett. 9, 8186–8193.

Xue, Q., Xing, Y., Lu, J., 2022. An integrated lane change prediction model incorporating traffic context based on trajectory data. Transport. Res. C Emerg. Technol. 141, 103738.

Yang, Z., Jia, X., Li, H., Yan, J., 2023. A survey of large language models for autonomous driving. https://doi.org/10.48550/arXiv.2311.01043.

Zhang, Y., Shi, X., Zhang, S., Abraham, A., 2022. A xgboost-based lane change prediction on time series data using feature engineering for autopilot vehicles. IEEE Trans. Intell. Transport. Syst. 23, 19187–19200.

Zhao, H., Gao, J., Lan, T., Sun, C., Sapp, B., Varadarajan, B., et al., 2021. TnT: target-driven trajectory prediction. In: Conference on Robot Learning, pp. 895–904.

Zhou, D., Schärli, N., Hou, L., Wei, J., Scales, N., Wang, X., et al., 2022. Least-to-most prompting enables complex reasoning in large language models. https://doi.org/10.48550/arXiv.2205.10625.

Zyner, A., Worrall, S., Nebot, E., 2018. A recurrent neural network solution for predicting driver intention at unsignalized intersections. IEEE Rob. Autom. Lett. 3, 1759–1764.

Zyner, A., Worrall, S., Ward, J., Nebot, E., 2017. Long short term memory for driver intent prediction. In: 2017 IEEE Intelligent Vehicles Symposium, pp. 1484–1489.

Mingxing Peng received her B.S. degree in digital media technology from Huaqiao University, China in 2018, and received her M.S. degree in computer technology from Sun Yat-sen University in 2021. Currently, she is pursuing her Ph.D. degree under the Intelligent Transportation Thrust at the Systems Hub, The Hong Kong University of Science and Technology (Guangzhou). Her research interests include motion planning, trajectory prediction, and large language models for autonomous driving.

Xusen Guo received his B.S. degree in material chemistry from Sun Yat-sen University, China in 2017, and received his M.S. degree in computer science and technology from Sun Yat-sen University in 2020. Currently, he is pursuing his Ph.D. degree under the Intelligent Transportation Thrust at the Systems Hub, The Hong Kong University of Science and Technology (Guangzhou). His research interests include urban computing and the application of large language models in intelligent transportation.

Xianda Chen received his M.S. degree from The University of Hong Kong, Hong Kong, China, in 2022. He is currently pursuing his Ph.D. degree in intelligent transportation with The Hong Kong University of Science and Technology (Guangzhou). His research interests include intelligent transportation, machine learning, and data analytics.

Kehua Chen received his B.S. degree in civil engineering from Chongqing University and a dual M.S. degree in environmental sciences from the University of Chinese Academy of Sciences and the University of Copenhagen. He earned his Ph.D. degree in intelligent transportation from the Hong Kong University of Science and Technology in 2024. Currently, he is a postdoctoral scholar at the Smart Transportation Applications and Research (STAR) Lab at the University of Washington. His research interests encompass urban and sustainable computing, as well as autonomous driving.

Meixin Zhu is a tenure-track Assistant Professor in the Intelligent Transportation (INTR) under the Systems Hub at The Hong Kong University of Science and Technology (Guangzhou), and with Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things. He is also an affiliated Assistant Professor in the Department of Civil and Environmental Engineering at The Hong Kong University of Science and Technology. He obtained his Ph.D. degree in intelligent transportation at the University of Washington (UW) in 2022. His research interests include autonomous driving decision making and planning, driving behavior modeling, traffic-flow modeling and simulation, traffic signal control, and multi-agent reinforcement learning.

Long Chen (Senior Member, IEEE) received his Ph.D. degree from Wuhan University in 2013. He is currently a Professor with State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. His research interests include autonomous driving, robotics, and artificial intelligence, where he has contributed more than 100 publications. He serves as an Associate Editor for *IEEE Transaction on Intelligent Transportation Systems*, *IEEE/CAA Journal of Automatica Sinica*, *IEEE Transaction on Intelligent Vehicle*, and *IEEE Technical Committee on Cyber-Physical Systems*.

Fei-Yue Wang (Fellow, IEEE) received his Ph.D. degree in computer and systems engineering from the Rensselaer Polytechnic Institute, Troy, NY, USA, in 1990. He joined The University of Arizona, Tucson, AZ, USA, in 1990 and became a Professor and the Director of the Robotics and Automation Laboratory. In 2002, he was appointed as the Director of the Key Laboratory of Complex Systems and Intelligence Science, Chinese Academy of Sciences (CAS), and the Vice President of Institute of Automation, CAS, in 2006. He found CAS Center for Social Computing and Parallel Management in 2008 and became the State Specially Appointed Expert and the Founding Director of the State Key Laboratory for Management and Control of Complex Systems in 2011. His research interests include parallel intelligence, social computing, and knowledge automation.