Neurocomputing 651 (2025) 131014

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Survey paper

A comprehensive review of deep learning techniques for interaction-aware trajectory prediction in urban autonomous driving

Iago Pachêco Gomes*, Denis Fernando Wolf

University of São Paulo, Institute of Mathematics and Computer Sciences, Avenida Trabalhador São-carlense, 400 - Centro, São Carlos, 13566-590, São Paulo, Brazil

HIGHLIGHTS

- This paper reviews intention and interaction-aware trajectory prediction methods.
- It examines 14 key machine learning sub-areas applied to trajectory prediction.
- It analyzes the top three techniques on public benchmarks for trajectory prediction.
- It presents a comprehensive review of public datasets for trajectory prediction.
- It discusses key future directions and considerations in trajectory prediction.

ARTICLE INFO

Communicated by S. Bonechi

Keywords:
Trajectory prediction
Maneuver intention
Interaction
Autonomous vehicles

ABSTRACT

Autonomous vehicles can improve urban transport by using multiple components that accurately represent their surroundings and improve decision-making processes. One essential component is trajectory prediction, which estimates the future states of traffic participants and anticipates hazardous scenarios. There are different approaches for trajectory prediction, in which Intention-aware and Interaction-aware approaches represent the state-of-the-art since they involve better representation of the surroundings. This paper reviews the literature on Interaction-Aware Trajectory Prediction for autonomous vehicles. It explores how incorporating maneuver intentions and interactions can improve prediction accuracy, and it examines the techniques and datasets employed in this field.

1. Introduction

Autonomous vehicles (AV) are intelligent robotic vehicles designed to operate in traffic autonomously, managing various driving scenarios while adhering to traffic regulations [17]. They integrate diverse sensors and software components to build a detailed scene representation, enabling them to comprehend the environment, make decisions, and execute actions comparable to or exceeding those of human drivers [177]. Consequently, they are expected to transform urban traffic, enhancing mobility, safety, and accessibility, while also contributing to the reduction of pollutant emissions [65].

As the system's autonomy increases, they are required to handle a broader range of tasks, meeting the requirements of their current level of autonomy [225]. Because of this, the system needs more components to safely avoid collisions, since obstacle detection alone is not enough. Hence, it is also necessary to track and predict traffic participants' trajectories and behaviors, enriching the representation of the environment. With this information, AVs can redesign their trajectories and actions to, for example, prevent accidents if collision paths are detected, or plan better trajectories and maneuvers, such as lane changes, emergency stops, obstacle avoidance, and overtaking.

Trajectory prediction refers to the process of forecasting the future states of objects, based on their past behavior and surrounding environment. In urban scenarios, an object may be any of the traffic participants that share the roads with the autonomous vehicles, such as pedestrians, vehicles, and cyclists [16]. These predicted states, characterized by their time dependency and continuous domain (*e.g.*, position), are important for autonomous vehicle navigation. Fig. 1 depicts a typical urban scene with vehicles and pedestrians at an intersection. In this scenario, the trajectory prediction module estimates the positions of surrounding vehicles and pedestrians based on observed paths and road geometry. Additionally, behavior intention prediction forecasts discrete states of

* Corresponding author.
Email addresses: iagogomes@usp.br (I.P. Gomes), denis@icmc.usp.br (D.F. Wolf).

https://doi.org/10.1016/j.neucom.2025.131014
Received 11 December 2024; Received in revised form 13 April 2025; Accepted 13 July 2025
Available online 16 July 2025

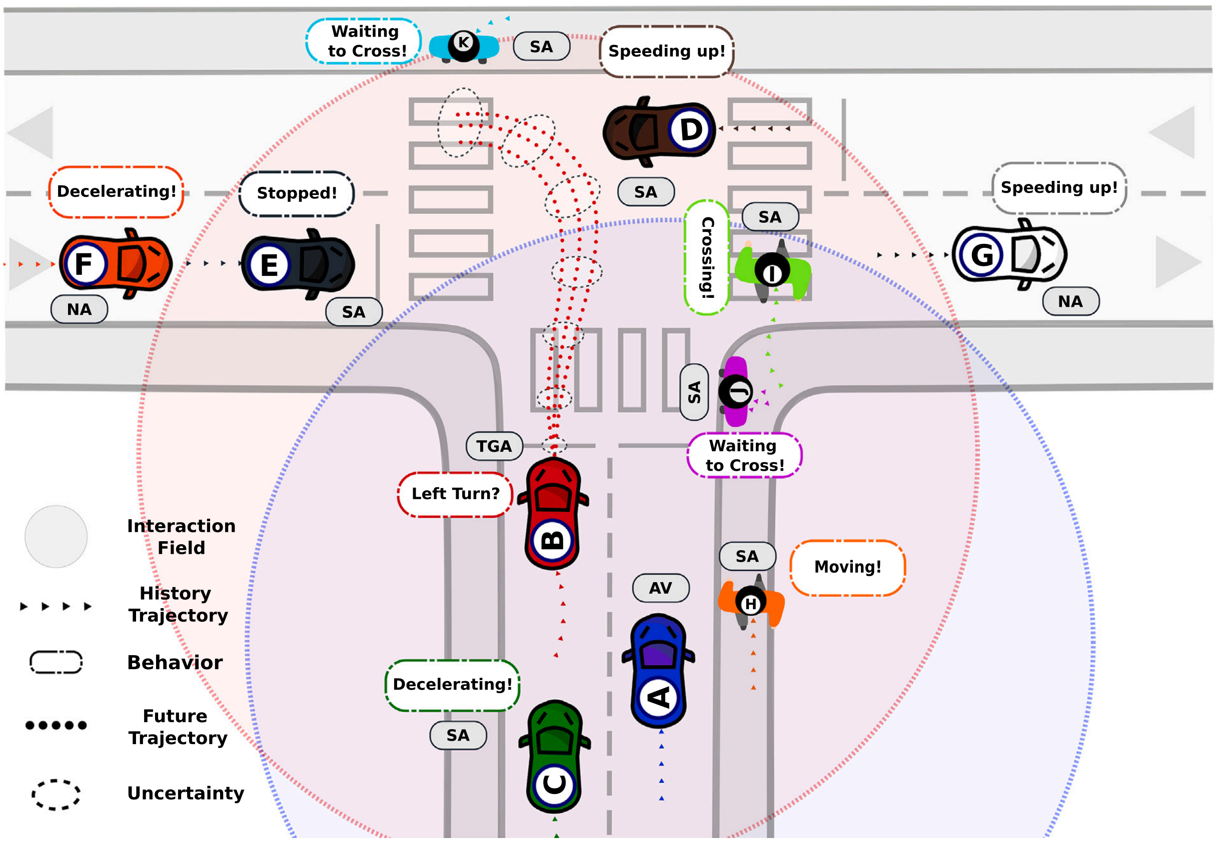


Fig. 1. General traffic scenario with interaction between agents. This scenario represents a generic situation of interaction between agents, such as vehicles and pedestrians, at an intersection. Each agent is labeled differently: **AV** (Autonomous Vehicle), **TGA** (target agent), **SA** (surrounding agents), and **NA** (non-effective agent). Within its interaction field, each agent interacts with other **SA** agents, which exert influence on its actions. **NA** agents, on the other hand, do not affect their actions. In a multi-agent prediction scenario, all surrounding agents (**SA**) is also a target agent (**TGA**).

traffic participants, such as maneuvers (*e.g.*, lane changes, left/right turns) or actions (*e.g.*, waiting to cross and crossing) [1].

Intention-aware and interaction-aware trajectory prediction represent two interconnected yet distinct approaches to forecasting the movement of traffic participants in urban scenarios. Intention-aware trajectory prediction focuses on understanding the high-level goals or discrete actions of agents. In contrast, interaction-aware trajectory prediction models the dynamic dependencies and mutual influences between agents, capturing how their trajectories adapt in response to each other [154]. While intention-aware prediction provides priors about an agent's future actions, interaction-aware prediction refines these forecasts by incorporating the interactions among agents, resulting in more context-sensitive predictions. Together, they form a complementary framework that enhances the ability of autonomous vehicles to anticipate complex behaviors in shared environments.

To address these challenges, the scientific community is actively developing advanced trajectory prediction frameworks. These efforts aim to improve prediction accuracy while considering real-time and memory constraints inherent to autonomous system architectures. In this sense, this paper conducts a literature review on Interaction-Aware Trajectory Prediction for autonomous vehicles, analyzing studies since 2008, following the DARPA Urban Challenge. The analysis focuses on how maneuver intentions and interactions enhance prediction performance, examining employed techniques, and datasets.

The research questions guiding this review are:

- **RQ1**: *How the intention prediction of traffic participants improves their trajectory prediction?* This question aims to understand how the behavior's intention improves trajectory prediction.
- **RQ2**: *How the interaction models of traffic participants improve their trajectory prediction?* This question aims to understand how the interaction models influence and improve trajectory prediction.
- **RQ3**: *What are the existing solutions for trajectory prediction that consider the intention of traffic participants and/or interaction models?* This question aims to explore the various algorithms, techniques, methodologies, and technological approaches used for trajectory prediction.

The remainder of this paper is organized as follows: Section 2 reviews related works on trajectory prediction, providing context for the study. Section 3 addresses research questions **RQ1** and **RQ2**, examining how behavioral intention and interaction modeling influence trajectory prediction. Section 4 investigates road geometry features, a key component of trajectory prediction models. Section 5 explores research question **RQ3**, analyzing the impact of deep learning techniques on prediction performance. Section 6 discusses benchmarks and challenges in trajectory prediction, identifying contributions and gaps in current approaches. Section 7 highlights the most relevant datasets used in primary studies, serving as important resources for addressing identified challenges. Section 8 synthesizes key findings, outlining directions for future research. Finally, Section 9 presents the conclusions of the review.

2. Related works

Trajectory prediction is a research field in autonomous vehicles that has been rapidly growing, driven by advancements in deep learning models, computer vision, and other machine learning techniques. In addition, the availability of various public datasets and open challenges has further contributed to the development of new technologies. As a result, there are some reviews published in the literature that summarize the progress made in this field.

Ridel et al. [219] and Brouwer et al. [21] conducted a review on pedestrian behavior and trajectory prediction, emphasizing the impact of individual pedestrian characteristics on predictions, such as age, and objects that the pedestrian is carrying, among others. Rudenko et al. [223] surveyed human motion prediction for diverse applications, introducing a taxonomy based on the *modeling approach* and *contextual cues*. The first class, *modeling approach*, focuses on how the motion is represented. Meanwhile, *contextual cue* class emphasizes the awareness of the target agent of its surroundings.

Alternatively, Bighashdel and Dubbelman [16] focused on Vulnerable Road Users (VRUs), categorizing prediction techniques into interaction-based, path-planning, and intention-based methods. *Interaction-based* approaches consider target agents as interactive entities in an environment with other agents and objects. In turn, *path-planning* methods inherit characteristics of path-planning techniques in robotics, while *intention-based* techniques take into account the motion or behavioral intention of the target agent, *e.g.*, the intention to cross the road.

Lefévre et al. [154] surveyed vehicle trajectory prediction techniques, categorizing them into physical-based, maneuver-based, and interaction-aware methods. *Maneuver-based*, also known as *intention-aware*, uses the maneuvering intention of the target vehicle as a priori information for the trajectory prediction framework. However, these techniques operate under the assumption that vehicles act independently of each other, meaning their actions do not influence the decisions of other agents. While this assumption enables long-term predictions, it is not always reliable.

In turn, *Interaction-aware* approaches consider traffic participants as interactive agents, such that their actions are interdependent. This assumption can better represent the dynamics of traffic scenarios, which makes long-term predictions more reliable than *intention-aware* methods [4,102,134,143,154]. The review presented by Lefévre et al. [154] does not provide further analysis of interaction-aware trajectory prediction because there were not many primary studies that followed this approach during the review coverage period. However, this approach has become more prevalent, especially with deep-learning techniques.

Recent efforts have updated trajectory prediction techniques, primarily leveraging deep learning [12,66,112,114,156,205,254,330]. However, these reviews are limited to a small scope of techniques compared to what is already available in the literature. Hu and Zheng [106] explored behavior and trajectory prediction in the context of Internet-of-Vehicles (IOV), highlighting the benefits of Vehicle-To-Anything (V2X) communication. The communication between traffic participants simplifies the prediction task by enabling the exchange of information regarding their states or intentions. This information is either free of noise or more reliable than the data collected by sensors belonging to the ego-vehicle. Mozaffari et al. [196] proposed a taxonomy classifying trajectory prediction studies based on input representation, output type, and prediction method, focusing on deep learning techniques. However, there are other techniques in the state-of-the-art of the research field that were not detailed by the authors in their survey.

Huang et al. [114] analyzed classical machine learning and novel deep learning techniques for trajectory prediction, focusing on the influence of interaction and road-related features, as well as different graph neural network architectures and reinforcement learning. Bharilya and Kumar [12] reviewed classical machine learning models, probabilistic approaches, and deep learning architectures, providing detailed discussions on their components and covering evaluation metrics and their characteristics. Huang et al. [112] concentrated on vision-based models that predict multiple trajectories per agent, effectively capturing uncertainty in predictions. Similarly, Fang et al. [66] examined uncertainty and recent deep learning techniques but emphasized behavior estimation, including discrete behaviors such as lane changes and pedestrian road crossings. However, these reviews often lack a comprehensive analysis of some technologies applied to trajectory prediction, including recent breakthroughs in deep learning.

Therefore, in this review, we focus on interaction-aware trajectory prediction approaches while also addressing intention-aware methods and hybrid approaches that leverage both to enhance prediction accuracy. We provide an extended classification of prediction methods and contextualize them within these frameworks. The review highlights several advanced deep learning approaches, such as semi/self-supervised models, memory-augmented networks, diffusion networks, meta-learning, large language models, transformers, multi-task learning, continual learning, and uncertainty estimation models, many of which are not extensively covered in existing reviews. We also analyze the datasets commonly used to develop and evaluate these methods, exploring how maneuver intention and interaction are incorporated. Finally, we discuss key challenges and propose future directions for advancing the field.

3. Behavior intention and interaction features

This section explores **RQ1** and **RQ2**, focusing on behavioral intention and interaction modeling's importance for trajectory prediction. Intention-aware trajectory prediction considers the target agent's behavioral intentions to forecast future states.

A maneuver, as described by Katrakazas et al. [134], represents a recurring pattern of vehicle motion, such as lane changes, left/right turns, or straight-line driving. Maneuver intention, on the other hand, reflects a driver's inclination towards executing a specific maneuver. There are many variables that influence this decision and the performance of maneuvers. According to AbuAli and Abou-Zeid [1], some stand out, being: individual resources, which are physical, social, psychological, and mental features of the driver; knowledge or skills, which represent the driver's experience and knowledge about the driving task; environmental factors, which are surrounding features such as weather, traffic conditions, and vehicle status; and workload and risk awareness, which address features of defensive driving such as risk assessment, situation awareness or traffic attention, and the effort expended for the task.

As outlined by Xing et al. [290] and Doshi and Trivedi [61], drivers' decisions are naturally hierarchical, and road-level decisions can be divided into strategical, tactical, and operational sub-levels. Strategical decisions are high-level motion plans such as routes and comfort parameters. Alternatively, tactical decisions are short-term decisions that lead to general goal satisfaction, which can also be regarded as lateral and longitudinal maneuvers. Finally, the operational level, also known as the control level, is responsible for estimating control commands.

In heterogeneous driving scenarios, the notion of maneuver intention extends to behavioral intention, increasing the diversity of possible actions. For instance, pedestrians exhibit complex motion patterns, adding challenges to anticipate their actions and predict their trajectories. For instance, these agents can change their actions abruptly, such as changing direction, stopping, or accelerating.

Intention prediction primarily focuses on tactical decision-making, particularly in lateral (*e.g.*, lane changes), and longitudinal (*e.g.*, acceleration, deceleration), maneuvers. For other agents like pedestrians, the goal is to anticipate their actions that might affect the vehicle trajectory, such as crossing the road or waiting to cross the road. However, in intention-aware trajectory prediction, each behavior is treated as an independent task, disregarding potential interactions with other agents. This assumption overlooks the dynamic nature of traffic scenarios, where interactions among agents influence decision-making [290].

In reality, drivers and pedestrians operate within an interactive environment, where actions may depend on the surrounding context. This highlights the complexity of intention-aware trajectory prediction and underscores the need for models that account for interactive behaviors among traffic participants.

In this sense, interaction-aware trajectory prediction methods extract reliable information from surroundings, enabling more accurate

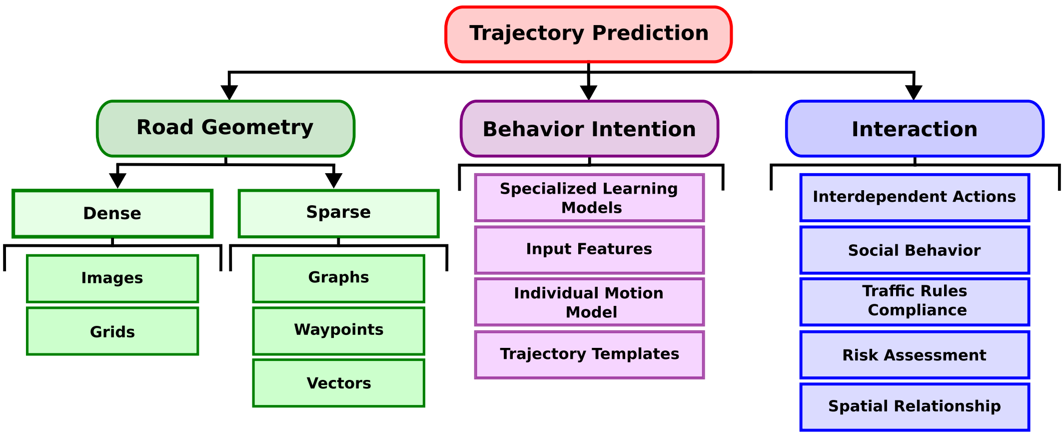


Fig. 2. A general classification for road geometry, behavior intention, and interaction features.

long-term trajectory predictions compared to other approaches. However, interaction among traffic participants is challenging for autonomous systems, given the closely interconnected system of behaviors. There are many reasons that increase the difficulty of interaction modeling, some of them are: complexity in predicting interaction-aware maneuver intention; wide range of traffic scenarios; the number of traffic rules; heterogeneity of traffic participants (*e.g.*, cars, buses, trucks, motorcyclists, cyclists, and pedestrians); traffic participants that disobey traffic rules; stochastic behavior of traffic participants; real-time constraints of the application; complexity in defining spatio-temporal relationship between traffic participants; and the different contributions that each traffic participant has in relation to the actions of the target vehicle [29,153,154].

Fig. 2 summarizes some applications of behavioral intention and interaction for trajectory prediction that stood out in the primary studies listed in the review. The main use of behavioral intention features is as conditional variables for either probabilistic models or multimodal deep learning approaches. Prior to the widespread adoption of deep learning techniques in prediction tasks, researchers relied on behavioral intentions to develop motion models tailored to specific classes.

- **Trajectory templates**: Group trajectories based on motion pattern similarity using algorithms like clustering or probabilistic models.
- **Individualized motion models**: Build motion models for each behavior class using curve regressions or motion equations in filters or sampling techniques.
- **Input features**: Intention prediction serves as an input feature for trajectory predictors, used as a conditional variable or part of the feature vector in probabilistic, autoencoder, or deep learning frameworks.
- **Specialized learning models**: Estimates the vehicle's future trajectory using learning models specialized in each motion pattern, such as multi-modal encoder-decoder architectures.

In traffic interaction, the focus often lies on two main aspects: how to represent the interaction as input features, such as grids, images, or graphs; and how to develop machine learning models to extract features from these structures, such as graph neural networks, pooling operations, or attention mechanisms. Primarily, these models aim to capture spatial and temporal relationships. However, some researchers have also explored representing risk assessment (*e.g.*, collision detection), traffic rules compliance, and decision-making features derived from the interactions among traffic participants.

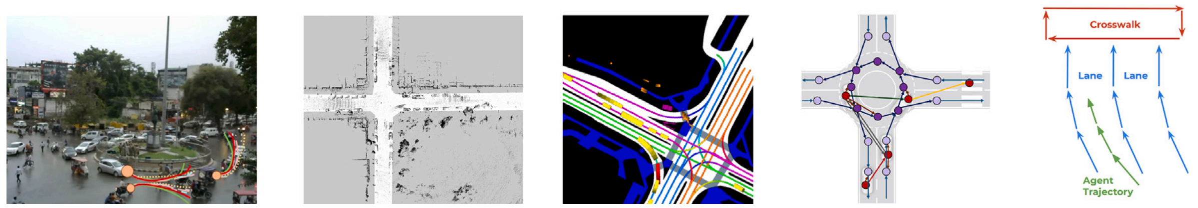
- **Spatial relationship**: This represents how the position of each traffic participant influences the target vehicle's future position using dense or sparse data structures like Occupancy Grids or graph structures.
- **Risk assessment**: Utilizes metrics like time-to-collision to assess the interaction between traffic participants, serving as input features or interaction heuristics.
- **Traffic rules compliance**: Considers compliance with traffic rules as input features or heuristic assessment criteria, limited to rules like speed limits or stop signs.
- **Social behaviors**: Represents widely followed navigation rules such as spacing between vehicles or average driving speed, serving as features for prediction or trajectory evaluation.
- **Interdependent actions**: Models the influence of each participant's actions on the traffic scenario using probabilistic models or game theory to estimate conditional probability distributions or interactions among agents.

Intention-aware and interaction-aware trajectory approaches complement each other by combining high-level behavioral cues with dynamic agent dependencies. Intention-aware models predict an agent's future actions, such as a vehicle's intention to change lanes or a pedestrian's plan to cross the street. Interaction-aware models refine these predictions by considering the mutual influences and responses of surrounding agents, resulting in context-sensitive outputs.

Fig. 1 illustrates a scenario involving vehicles **B** (red), **E** (black), **D** (brown), and pedestrian **K** (light blue) in a complex interaction. An intention-aware approach predicts vehicle **B**'s future position based on its identified intention to *turn left*, without considering other agents. An interaction-aware approach, in contrast, estimates vehicle **B**'s position based on its relationships with surrounding agents, though it may struggle with increased uncertainty at intersections. A hybrid approach combines both methods, considering vehicle **B**'s intention and how vehicles **D**, **E**, and pedestrian **K** influence its behavior, leading to more adaptive predictions such as speed adjustments or changes in the sharpness of its turn.

Several studies address these interactions in trajectory prediction. Choi et al. [46] introduced a goal-oriented multi-modal approach using spatio-temporal graphs to model agent interactions, predict maneuver intentions, and generate heatmaps of possible trajectories using a Conditional Variational Autoencoder (CVAE). Similarly, Hu et al. [108] proposed a framework focusing on vehicle interactions. Their method used Dynamic Time Warping (DTW) to compare historical and reference trajectories for intention prediction.

Feng et al. [72] employed an LSTM network to extract features from historical trajectories and Time-To-Collision (TTC) estimations for behavior intention prediction. They used a CVAE to generate future trajectories based on these features. Li et al. [164] introduced a framework that integrates maneuver intention using a Bidirectional LSTM for intention forecasting and interaction features from a Graph Neural Network



(a) Images (Chandra, Bhattacharya, Bera, & Manocha, 2019b) (b) LiDAR BEV (Choi et al., 2021) (c) Semantic Grids (Pathiraja et al., 2022) (d) Graphs (Z. Sun et al., 2024) (e) Vectors (J. Gao et al., 2020)

Fig. 3. Types of road geometry representation.

(GNN). They developed a "Behavior Attention Mechanism" to compute attention scores based on vehicle intentions. Chandra et al. [31] presented an interaction-aware trajectory prediction model focusing on maneuvers such as overspeeding, underspeeding, or neutral behavior. They used undirected weighted graphs to represent spatial relationships and employed a Spectrum Graph-LSTM network for intention prediction through graph spectrum analysis. These works demonstrate how combining intention-aware and interaction-aware approaches enhances trajectory prediction, providing more accurate and context-sensitive outcomes.

On the other hand, considering behavioral intention and interaction also has limitations. These approaches increase model complexity, requiring more computational resources and extended training times. They rely on large, annotated datasets to effectively capture diverse behaviors and interactions. Additionally, modeling the dynamic interplay between intention and interaction can introduce uncertainty, especially in scenarios with unpredictable agent behaviors or limited data. Balancing these factors while maintaining real-time processing remains a significant challenge for deploying these models in autonomous systems.

4. Road geometry representation

Road geometry is an important feature for behavior intention and trajectory prediction for vehicles and pedestrians on urban roads. In the first case, this representation enables behavior prediction models to anticipate the intentions of traffic participants based on their interactions with the surrounding infrastructure, such as lane changes, left/right turns, or other route choices. In turn, for trajectory prediction, the surrounding layout provides plausible future positions for the target agent, by portraying the spatial constraints, lane configurations, traffic regulations, and geometric features of the road network. Moreover, the layout and topology of the surrounding environment are also essential to enhance the representation of the interaction between traffic participants, since they can define interaction areas between vehicles (*e.g.*, merges, overtaking, and negotiations at intersections or roundabouts) or pedestrians (*e.g.*, pedestrian crossings), and their spatial limitations.

Fig. 3 presents structures used to represent road geometry in behavior and trajectory prediction. These structures are categorized as dense or sparse representations. Dense structures like images and grids (e.g., Figs. 3a–c) offer detailed road geometry features, but heavily depend on the model to extract relevant information. Conversely, sparse representations like waypoints, graphs (Fig. 3d), and vectors (Fig. 3e) explicitly represent road geometry features, but may lack certain details such as colors, textures, bumps, or elevation profiles.

5. Deep learning techniques

Several techniques have been used for interaction-aware trajectory prediction, due to the aspects related to intention prediction, interaction modeling, and time series analysis. These techniques extend to analytical, probabilistic, and deep learning models. In addition, there are also architectural designs and auxiliary techniques, which aim to improve the performance of other approaches. Hereby, to answer **RQ3**, we highlight the main approaches and topics applied in interaction-aware trajectory prediction. Fig. 4 provides an overview of these techniques, grouped by machine learning topics and methods. Tables 1–3 present a detailed classification of techniques and primary studies. Notably, recent approaches often integrate multiple techniques in a hybrid methodology.

5.1. Sequence-based models

Sequence-based models for trajectory prediction are neural network architectures designed to handle sequential data, such as historical trajectories of moving objects, and make predictions about their future paths. These models leverage the temporal dependencies present in the sequential data to learn patterns [226].

However, these techniques may encounter difficulties when dealing with large temporal sequences or when incorporating features from different agents within an interactive scenario. As a result, current models often employ sequential models initially to extract sequential features from temporal-related observations (*e.g.*, historical trajectory). Subsequently, these features are refined, and interaction dynamics are captured using more suitable approaches.

Limitations: A key limitation of traditional sequential models is their ability to handle long temporal sequences without losing context.

5.2. Spatial-based models

Spatial-based models leverage neural networks to capture spatial relationships and patterns in data. They employ convolutional layers to extract spatial features from input data, like images or grids. Another approach involves representing spatial relationships using graphs, with nodes indicating agent positions and edges as relationships, such as the relative distances. While spatial-based models may not capture temporal dependencies as effectively as sequence-based models, they remain valuable for trajectory prediction tasks. Additionally, these models can implicitly or explicitly represent interactions between traffic participants and other contextual information (*e.g.*, traffic rules and road geometry).

Limitations: The spatial-based techniques are sensitive to the quality of spatial representation, which limits their ability to extract meaningful features from sparse or noisy data. Moreover, they also lose performance in crowded environments, where the number of agents demands more processing time, such as in graph-based approaches.

5.3. Generative models

Generative models learn the underlying distribution of trajectory data and generate new plausible trajectories. However, each type of generative model has its advantages and disadvantages, and the choice of model depends on factors such as the complexity of the data distribution, and the specific requirements of the application.

For example, Conditional Variational Autoencoder (CVAE) uses an encoder-decoder structure to map input to a latent space and reconstruct it, conditioned on a variable [9]. Generative Adversarial Networks (GANs) consist of a generator and discriminator, with the generator

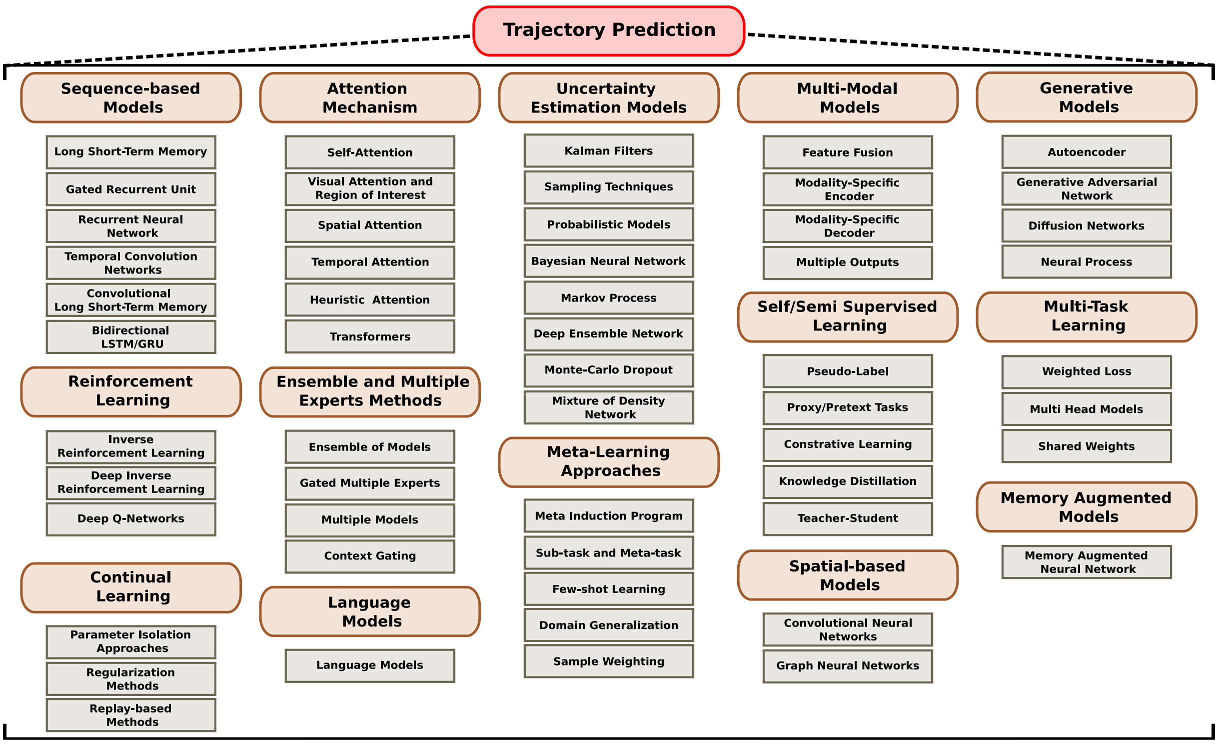


Fig. 4. An overview on techniques applied for trajectory prediction.

Table 1
Classification of primary studies – Part 1.

Topic	Techniques	Primary studies
Sequence-based models	Long Short-Term Memory (LSTM)	[10,11,14,29–31,33,35,39,41,42,55,56,64,68,88,98,100,101,105,116,119,123,130,131,135,136,149,164,169,193,203,211,224,259,271,276,278,288,289,303,315,321]
		[22,24,37,38,40,45,53,59,60,62,70,72,73,108,115,139,142,147,150,159,166,171,172,175,183,186–189,192,210,222,227,234,237,239,241,243,250,255,274,295,297,304,308,314,318]
	Gated Recurrent Units (GRUs)	[2,3,34,50,57,69,72,77,81–83,97,111,119,121,153,158,160,161,170,175,184,193,208,210,218,220,221,227,234,235,247,265,268,272,285,293,298–300,310,324]
	Temporal Convolutional Networks (TCN)	[162,181,242,250,304]
	Convolutional LSTM (Conv-LSTM)	[136,241,291]
	Bidirectional LSTM or Bidirectional GRU	[29,30,198]
	Other Recurrent Neural Networks	[52,74,102,121,124,164,167,190,191,227,243,318]
Spatial-based models	Convolutional Neural Networks (CNN)	[10,22,25,27,37,39,40,46,51,55,59,63,73,81,82,132,150,153,158,162,165,166,171,184,189,224,227,242,265,304,307,309,312,319,322]
		[15,20,24–26,35,36,49,50,58,62,64,67,70,76,76,77,87,92,93,111,122,123,128,128,130,139,140,147,148,161,168,171,174,174,175,178,180,181,181,187,192,193,197,203,209–214,217,220,221,233–235,239,241,250,266,268,278,280,286,293,301,303,308–310,313,318,321,321]
	Graph Neural Networks (GNN)	[2,6,15,31,34,40,50,53,57,82,83,92,94,105,162,164,175,193,247,250,252,255,269,274,279,293,297,300,307,312,319,322]
Generative models	Generative Adversarial Networks (GAN)	[89,100,149,159,160,169,197,222,238,276]
	Conditional Variational Autoencoder (CVAE)	[37–39,45,46,49,64,72,107–109,119–121,153,158,170,184,209,227,241,278,305]
	Other Autoencoder	[43,53,69,128,132,207,308]
	Diffusion Networks	[34,126,167,253,279]
	Neural Process	[215,327]

creating patterns following data distribution and the discriminator distinguishing real from synthetic data [91]. Diffusion models refine noise into data distribution iteratively [48].

CVAE's advantage lies in its conditional variable, allowing precise control over distribution reconstruction, although high-dimensional conditioning can pose training challenges [9]. GANs offer flexibility in generating realistic trajectories, but may face instability and mode collapse during training [144]. Diffusion models efficiently handle uncertainty but require extensive data and computational resources for training and parameter tuning [48].

Table 2
Classification of primary studies – Part 2.

Topic	Techniques	Primary studies
Reinforcement learning	Inverse Reinforcement Learning (IRL)	[90,117,231]
		[107,244]
	Deep Inverse Reinforcement Learning (Deep IRL)	[73]
	Deep Q-Networks (DQN)	[132]
Attention mechanism	Squeeze-and-Excitation Attention Mechanism (SE)	[24]
	Self-Attention and Multi-Head Attention Mechanism	[6,37,38,40,57,70,74,75,82,97,115,125,126,133,139,141,142,166,175,186–189,200,236,252,265,272,275,285,295,299,303,314,324,329]
	Cross-Attention Mechanism	[36,70,75,81,82,92–94,115,125,133,151,199,200,215,217,236,247,272,274,275,305,323,324,329]
	Spatial Attention	[35,125,133,139,151,158,162,188,189,210,252,265,274,295,314]
	Temporal Attention	[36,94,125,133,151,162,186,197,200,252,272,274,285,324]
	Social Attention	[6,35,36,40,81,82,142,186,188,189,200,210,252,272,285,314,324]
	Visual Attention and Region of Interest	[29,158,210,233,298]
	Heuristic Attention	[3,52,131,164,317,326]
	Transformers	[13,34,36,43,75,86,87,101,115,116,126,127,133,151,152,165,167,170,175,179,197,199,200,207,216,236,277,281,297,299,305,309,310,317,323,325,329]
Uncertainty estimation models	Kalman Filters (KF)	[8,135,155,161,287,311]
	Sampling Techniques	[103,163,202,257,284]
	Bayesian Network (BN)	[80,84,99,176,228–230]
	Gaussian Mixture Models (GMM)	[5,27,54,75,98,109,110,113,115,209,227,234,236,240,255,275,281,283,289,311]
	Bayesian Neural Networks (BNN)	[249]
	Monte-Carlo Dropout	[14,59]
	Deep Ensemble Network	[234,251]
	Gaussian Process (GP)	[95,176,256,257,284]
	Markov Process	[218,240]
	Mixture of Density Network (MDN)	[68,69,74,110]
Self-supervised and semi-supervised learning	Knowledge Distillation	[272,277,302,322]
	Proxy/Pretext Tasks	[15,43,128,151,168,207,214]
	Contrastive Learning	[150]
	Pseudo-Labels	[148,243,266,292,297]
	Teacher-Student	[150,322]
Continual learning	Regularization Methods	[44,204,238]
	Parameter Isolation Approaches	[118,260,261,268]
	Replay-based Methods	[296]

Limitations: GANs face instability and mode collapse during training, and CVAEs struggle with high-dimensional latent spaces. Diffusion models are computationally expensive, requiring extensive resources for training. Moreover, all generative approaches demand large, diverse datasets for effective learning, which can limit their applicability in data-scarce domains.

5.4. Reinforcement learning

These techniques learn optimal decision-making policies in a highly interactive learning strategy, in which the agent must continuously interact with the environment. Each action causes a transition between states, and is evaluated by a reward function. The purpose of the learning process is to maximize the reward function [248]. In trajectory prediction, this technique has been applied especially for learning human-like behaviors regarding interactions and decision-making.

The approaches usually vary from conventional Reinforcement Learning (RL) to more advanced Deep Reinforcement Learning (Deep RL) techniques. RL methods like Q-learning or Policy Gradient often require manually crafted features and may struggle to capture intricate spatial and temporal dependencies in trajectory data. On the other hand, Deep RL models can handle high-dimensional input spaces, effectively capturing complex dependencies, thus proving more suitable for real-world applications with dynamic environments. Additionally, Deep RL models implicitly encode traffic rules and explicitly model interactions between agents. Nevertheless, designing effective reward functions and ensuring stable training remain challenges in both RL and Deep RL methodologies.

Limitations: These techniques heavily rely on carefully designed reward functions, which are often scenario-specific and difficult to generalize. In addition, training is computationally intensive, requiring extensive simulations or real-world interactions. Moreover, stability issues and exploration–exploitation trade-offs remain challenges, especially in dynamic traffic scenarios.

5.5. Attention mechanism

The attention mechanism is a machine learning technique initially used in Natural Language Processing (NLP) to address the limitations of recurrent neural networks (RNNs) in processing long-term sequences. This mechanism is inspired by the selective process of human cognition, where certain pieces of information are selectively prioritized over others [201]. Therefore, it estimates weights from the input data, representing each input's contribution to generating the desired output.

There are many applications of attention for trajectory prediction. They can learn the most relevant temporal and spatial features for prediction [186,295,300] or differentiate the impact of each surrounding vehicle on the target vehicle's trajectory [29,161,188]. In

Table 3
Classification of primary studies - Part 3.

Topic	Techniques	Primary studies
Multi-task learning	Weighted Loss	[35,36,36,70,76,166,172,213,233,242, 297,304]
	Multi-Head Models	[22,35,36,132,166,213,297,304]
	Shared Weights	[122,214]
Ensemble and multiple-experts	Context Gating	[259]
	Interactive Multiple Models	[54,155]
	Bootstrap aggregating	[173]
	Gated Ensemble of Models	[140,195]
	Mixture of Experts	[77,119,263]
Meta-learning approaches	Meta Induction Program	[60]
	Sub-task and Meta-task	[237]
	Few-Shot Learning	[62,120,327]
	Domain Generalization	[71,170,208,277]
	Sample Weighting	[214]
Memory-augmented models	Memory Augmented Neural Networks	[184,185,292,296]
Language models		[7,125,232,246]

intention-aware frameworks, attention mechanisms can also identify the importance of features for each maneuver intention. Additionally, this technique is used in image processing to highlight features in specific areas of the image [273,320].

Furthermore, transformers are a promising approach for trajectory prediction, built on the attention mechanism. Initially proposed for machine translation in NLP, transformers outperformed recurrent neural networks for long-term sequences [262]. They have since been applied to various tasks such as object detection, image segmentation, pose estimation, tracking, and trajectory prediction [96].

A Transformer network consists of an encoder-decoder architecture with multi-head self-attention mechanisms and feed-forward networks [96]. The encoder learns long-term relationships within input features through positional encoding and attention score matrices. Finally, the decoder uses attention to map the encoder's output to the Transformer's output space in a sequence-to-sequence model. However, attention mechanisms introduce challenges related to computational complexity, overfitting, and hyperparameter tuning, which must be addressed to ensure models' effectiveness and robustness.

Yuan et al. [305] propose AgentFormer, an agent-aware Transformer that models socio-temporal dependencies among multiple agents. It captures temporal relationships to link past agent states to future states and social interactions to represent how agents influence one another. Alternatively, Zhang et al. [309] present Gatformer, a Graph Attention Transformer for trajectory prediction. Gatformer represents traffic scenes as sparse graphs, extracts spatial features using CNNs, and employs a Graph Attention Network (GAT) to model agent-agent and agent-infrastructure interactions. A Transformer network then predicts trajectories. In contrast, Nayakanti et al. [199] introduce Wayformer, which uses attention-based scene encoders and decoders to model diverse input features, including road geometry, lane connectivity, and agent interactions. It explores various fusion strategies for input modalities within the attention framework. Karim et al. [133] propose DESTINE, a method for predicting temporally consistent road user trajectories. DESTINE dynamically predicts agent goals and employs temporal transductive alignment for map-compliant predictions, generated in a coarse-to-fine manner with an attention module for temporal alignment. Mu et al. [197] present Multi-modality Scene Tokenization (MOST), which tokenizes the visual world into a compact set of scene elements for motion prediction. It integrates pre-trained image foundation models and LiDAR neural networks to encode scene elements effectively. Finally, Liu et al. [179] introduce mmTransformer, a Transformer framework for multimodal motion prediction. This architecture employs stacked Transformers to model multimodality at the feature level using fixed independent proposals. A region-based training strategy further encourages multimodality in the generated proposals.

Limitations: The high computational cost of attention mechanisms, particularly transformers, limits their scalability for large datasets or real-time applications. Moreover, they are prone to overfitting, requiring careful regularization. Finally, hyperparameter tuning is complex and resource-intensive.

5.6. *Uncertainty estimation models*

Uncertainty Estimation Models are techniques designed to assess the uncertainty associated with predictions, aiding in the quantification of a model's confidence or reliability. Various factors, including limited data, model misrepresentation, or inherent stochasticity, can contribute to this uncertainty. In trajectory prediction, the dynamic and interactive nature of traffic scenarios and the sensors' noise are major sources of uncertainty. Moreover, longer prediction horizons tend to accumulate more uncertainty in predicted trajectories. Techniques like Bayesian Neural Networks (BNN), non-linear filtering, such as the Extended Kalman Filter (KF), and probabilistic models, integrate uncertainty into their processes. However, some uncertainty estimation methods may introduce computational complexity during inference, potentially resulting in slower prediction times.

Li et al. [163] propose a Constrained Mixture Sequential Monte Carlo (CMSMC) method for simultaneous tracking and prediction of multi-agent behaviors. The method integrates a probabilistic graphical model for behavior recognition with a state evolution module, supporting multi-modality and addressing uncertainties such as occlusion and sensor failures. Similarly, Tang et al. [249] develop a Bayesian Recurrent Neural Network (BRNN) for vehicle trajectory prediction, ensuring physical feasibility and adaptability to specific human driving behaviors over long prediction horizons. Huang et al. [115] introduce GameFormer, a hierarchical game-theoretic model for multi-agent interactive prediction and planning. The model uses a Transformer encoder-decoder architecture to iteratively capture and refine agent interactions. Game theory principles guide the hierarchical decoding process, improving integration of agent interdependencies in trajectory predictions. The approach employs multi-modal trajectory predictions, representing potential future states as Gaussian mixture models (GMMs), with each mode corresponding to a possible trajectory and its likelihood. Finally, Shao et al. [234] present a unified prediction-planning approach that explicitly incorporates various forms of uncertainty into both prediction and planning modules to enhance safety and reliability. They categorize uncertainty into short-term, long-term, and epistemic components.

Limitations: Some uncertainty estimation methods exhibit high computational complexity during inference or depend on strong statistical assumptions, such as linearity in Kalman Filters. Additionally, these methods struggle to scale in complex environments with multiple agents.

5.7. *Self-supervised and semi-supervised learning*

Self-supervised and Semi-supervised Learning methodologies leverage unlabeled or partially labeled data to improve model performance. In the former, the model generates its supervisory signal from the input data, such as predicting missing or corrupted segments or clustering similar samples. Alternatively, Semi-Supervised Learning combines supervised learning with unlabeled data to improve performance.

These approaches are particularly relevant for intention-aware and interaction-aware trajectory prediction, given the lack of datasets with behavioral intention annotations. Consequently, semi-supervised and self-supervised models can mitigate the data availability limitations in the main trajectory prediction benchmarks. However, these strategies may introduce additional computational complexity and prolonged training durations compared to fully supervised methods, due to the integration of unsupervised loss functions.

Wang et al. [272] propose a trajectory prediction framework called Partial Observations Prediction (POP) for congested urban road scenarios. The framework uses a self-supervised pre-task to reconstruct missing observations. It then applies feature distillation, transferring knowledge from a teacher model trained on complete observations to a student model that operates with partial observations. Alternatively, Park et al. [207] address distribution shifts in trajectory prediction with a test-time training method. This approach employs a masked autoencoder (MAE) to learn robust representations of complex interactions despite distribution shifts. The MAE reconstructs masked portions of input trajectory data to improve generalization. Lamm et al. [150] investigate semi-supervised models for vehicle trajectory prediction. They compare contrastive learning with teacher-student methods and analyze the performance of models predicting a small number of trajectories versus those estimating probabilities over larger trajectory sets. Finally, Wang et al. [266] present a self-supervised approach for class-agnostic motion prediction using unlabeled LiDAR point clouds. The method generates coarse pseudo-motion labels by using an optimal transport solver to identify correspondences between current and future point clouds.

Limitations: These methods increase training complexity and computational cost by incorporating additional unsupervised loss functions or using unlabeled data.

5.8. Continual learning

Continual Learning (CL) involves the study, development, and evaluation of techniques and methodologies for learning over time [157]. This means continuously learning new concepts or tasks and improving knowledge and performance on already learned tasks [157,206]. This research field has many applications in trajectory prediction, especially in interaction-aware frameworks. These frameworks benefit from continual learning algorithms because they have multiple components that can improve over time. For instance, continual learning can enhance behavior or intention prediction by learning new maneuvers not previously known [229]. It can also improve interaction modeling by learning new interaction behaviors and their consequences. Additionally, continual learning can use feedback from the tracking algorithm to improve predictor performance by learning the individual characteristics of each target vehicle.

Limitations: Continual learning techniques face significant challenges, including the tendency for models to experience catastrophic forgetting, wherein previously learned knowledge is lost. Additionally, replay-based methods require effective memory management, while regularization methods must address instability during adaptation. These challenges also add complexity and computational overhead to the techniques.

5.9. Multi-task learning

Multi-task learning (MTL) is a machine learning approach that simultaneously learns two or more related tasks [316]. Zhang and Yang [316] explain that this approach leverages the correlation between tasks to improve overall generalization. To implement MTL, researchers use methods such as shared parameters between models and combining different loss functions. The learning process must carefully update parameters to avoid bias towards specific tasks, aiming for balanced performance across all tasks [258].

In trajectory prediction, MTL is particularly useful because autonomous driving involves many correlated tasks. For instance, intention prediction and trajectory prediction can be combined in a multi-task learning framework. Similarly, multi-target tracking and trajectory prediction can enhance each other's performance. Additionally, a multi-objective loss function can optimize various goals while learning to predict trajectories.

Limitations: The primary challenge in multi-task learning lies in designing a loss function that balances tasks and ensures stable model training, which often demands extensive parameter tuning and computational resources. Designing effective architectures for complex multi-task setups remains an ongoing research challenge.

5.10. Multi-modal models

Multi-modality has various interpretations in trajectory prediction. It can denote multi-modal machine learning, where models integrate diverse data sources like images, point clouds, historical trajectories, and maps. In this case, feature fusion techniques extract and integrate information from these modalities. For feature fusion, concatenation of feature vectors is commonly employed, along with cross-attention mechanisms between modalities. However, despite the inherently multimodal nature of predictor inputs, feature vector fusion is not widely explored in trajectory prediction, and further investigation is needed to determine the best strategies (in addition to concatenation and cross-attention).

In multimodal machine learning for trajectory prediction, the predominant approach involves developing modality-specific encoders and task-specific decoders. Modality-specific encoders are tailored to capture semantic information from various data sources, such as historical trajectory data, road geometry, interaction patterns, and temporal features. Task-specific decoders, on the other hand, are designed to address specific prediction tasks like trajectory and behavioral intention prediction. Furthermore, decoders can be specialized for different agent types, such as vehicles and pedestrians.

Alternatively, multi-modality can also refer to predicting multiple trajectories. These models provide a more comprehensive representation of uncertainty, enabling the model to account for multiple potential future outcomes. However, this approach increases the complexity of the model, and requires careful interpretation and handling of the multiple predicted trajectories. Moreover, such models often incorporate uncertainty estimation, which assesses trajectory probability to identify the most likely candidate. Consequently, evaluation metrics have been adapted or proposed to assess these probability estimations.

Limitations: Models that process inputs from different modalities (*e.g.*, images and point clouds) are more complex and require careful feature fusion to combine information effectively. Similarly, prediction approaches that generate multiple trajectories for each agent increase complexity and demand more computational resources. Maintaining real-time inference becomes increasingly challenging as the number of agents grows in dynamic scenes.

5.11. Ensemble and multiple-experts methods

Ensemble and Multiple Experts Methods involve combining predictions from multiple models or experts to improve overall performance. Common ensemble techniques include averaging predictions, bagging (bootstrap aggregating), boosting, and stacking. Each base model may be trained independently or with different subsets of the data, and the final prediction is typically a combination of individual model outputs. This approach can improve prediction accuracy and robustness by leveraging diverse models and reducing the risk of overfitting. However, it increases computational complexity and training time, especially for large ensembles or complex base models. In addition, it requires careful selection and tuning of base models and ensemble methods to achieve optimal performance.

Conversely, multiple experts methods involve training multiple specialized models or experts, each focused on different aspects or subtasks of the prediction problem. The predictions of these experts are then combined using a gating mechanism or weighting scheme to produce the final prediction. While this approach enables models to capture diverse aspects of the prediction problem by leveraging specialized expertise, it also increases model complexity and computational overhead, especially if complex gating mechanisms are used.

Limitations: Ensemble methods increase computational costs during training and inference, especially for large ensembles. In turn, multiple-expert approaches require complex gating mechanisms and careful expert selection, adding complexity to the implementation. Additionally, maintaining model interpretability in these frameworks remains a significant challenge.

5.12. Meta-learning approaches

Meta-learning is a paradigm focused on the development of models capable of efficiently generalizing knowledge and rapidly adapting to novel tasks or datasets, even with limited training data [79,182]. In trajectory prediction, meta-learning approaches seek to train models that exhibit quick adaptation to new scenarios or unseen data distributions, thereby enhancing prediction accuracy and generalization performance [137]. Some of the methods can adapt to most of the prediction architectures with minimal changes. However, they may be sensitive to the choice of hyperparameters and training procedures, requiring careful tuning to achieve optimal performance.

Park et al. [208] propose a method using Neural Stochastic Differential Equations (NSDE) to enhance the transferability of trajectory prediction models across datasets. The method employs continuous representations to accommodate arbitrary time steps and stochastic representations to account for detector and tracker errors, addressing discrepancies in data acquisition strategies. Alternatively, Ivanovic et al. [120] present a method for efficiently adapting behavior prediction models to new environments. The approach uses meta-learning, specifically Bayesian regression, to integrate an adaptive layer into existing behavior prediction models, enabling better environmental adaptability. Similarly, Shi et al. [237] introduce MetaTraj, a meta-learning framework designed to improve the transferability of trajectory prediction models to unseen scenes and objects. The framework defines sub-tasks, consisting of dynamic-length sequence prediction tasks from observations, and meta-tasks, which ensure the learned prior information contributes to accurate long-term predictions. This approach enables models to generalize effectively to new scenarios.

Limitations: Meta-learning is computationally expensive and highly sensitive to hyperparameter selection. In this sense, adapting to complex, real-world scenarios with noisy or imbalanced data remains challenging. Additionally, training requires diverse datasets that capture the variability of potential environments.

5.13. Memory-augmented models

Memory Augmented Models represent a class of neural network architectures that incorporate external memory components to enhance the model's capacity to retain and access information over time [138]. This class of techniques is particularly well-suited for tasks involving sequential or temporal data, such as language modeling, time-series prediction, and sequential decision-making, as they can store contextual and long-term information, retrieving this knowledge when necessary. However, the representation of stored knowledge and its significance for future inferences are inherent challenges in the area. Furthermore, the models are more complex and require adequate management in addition to more memory space.

Yang et al. [296] address catastrophic forgetting in trajectory prediction models when adapting to new scenarios. They propose Continual Learning-based Trajectory Prediction with Memory Augmented Networks (CLTP-MAN). CLTP-MAN incorporates an external memory module to store prior knowledge and a multi-hop attention mechanism to retrieve relevant information for trajectory prediction. Alternatively, Marchetti et al. [185] introduce a Social Memory Module (SMEMO) for trajectory forecasting, focusing on human interactions. SMEMO employs a Memory Augmented Neural Network to store past information from multiple agents in a shared external memory, which it uses for predictions. The model is designed to learn and explain cause–effect relationships between the motions of different agents.

Limitations: These models demand significant computational resources and memory management techniques. The interpretability of stored knowledge and its relevance for prediction remain challenging.

5.14. Language models

Language models can predict trajectories for autonomous vehicles by analyzing large datasets of vehicle movements, traffic conditions, and environmental cues. These models effectively process complex, unstructured data, enabling them to predict potential vehicle paths in various scenarios [267]. Their advantages include handling diverse data types and adapting to different contexts without extensive retraining. Large Vision Models (LVMs), a variant designed to process images, offer additional features [264]. Besides trajectory prediction, LVMs can use text queries to extract other relevant traffic scene information, such as the most likely agent posing a threat to the ego-vehicle. However, these techniques face challenges such as interpretability issues and require significant computational resources.

Sun et al. [246] propose State Transformer (STR), which unifies motion prediction and planning as a sequence modeling task. STR arranges observations, states, and actions into a single sequence, inspired by the success of large language models. It uses a conditional causal transformer backbone, similar to GPT-2, to learn representations from input sequences. These sequences include embeddings of encoded maps, past trajectories, and traffic light states, which condition the generation of future state sequences. Alternatively, Seff et al. [232] introduce MotionLM, a model that frames multi-agent motion prediction as a language modeling task. MotionLM represents continuous trajectories as sequences of discrete motion tokens and employs a language modeling objective to generate joint distributions over future agent interactions. Bae et al. [7] present LMTraj, a language-based multimodal trajectory predictor. LMTraj reformulates trajectory prediction as a question-answering task, converting trajectory coordinates into natural language prompts and using image captioning to describe scene images as text. Finally, Jia et al. [125] propose a motion forecasting method based on GPT-style next token prediction. This method represents both input and output in a unified space, facilitating autoregressive predictions. It incorporates three factorized attention modules and various position encoding styles to capture complex spatial–temporal and semantic relationships in driving scenes.

Limitations: Language models require vast computational resources and large datasets for training, which can limit their applicability in resource-constrained environments. Interpretability is another challenge, as their predictions often lack transparency. Moreover, fine-tuning these models for domain-specific tasks demands significant effort and expertise.

6. Benchmarks and prediction challenges

Benchmark datasets such as Argoverse 1[1] and 2,[2] NuScenes,[3] and Waymo[4] are important for advancing trajectory prediction research field. These benchmarks provide standardized environments for testing and comparing the performance of various prediction algorithms. In this section, we review the Top-3[5] methods, highlighting their methodology and employed techniques. Table 4 presents the benchmarks.

HPNet [252] improves trajectory prediction accuracy and stability by leveraging historical frames. Its core innovation, the Historical Prediction Attention module, encodes the dynamic relationship between successive predictions using a Multi-Head Attention Mechanism.

1 https://eval.ai/web/challenges/challenge-page/454/leaderboard/1279.
2 https://eval.ai/web/challenges/challenge-page/1719/leaderboard/4761.
3 https://eval.ai/web/challenges/challenge-page/591/leaderboard/1659.
4 https://waymo.com/open/challenges/2023/motion-prediction/.
5 Top-3 according to each benchmark and with available papers.

Table 4
Benchmark on trajectory prediction.

Benchmark	Technique	minFDE (K)	minADE (K)	minFDE (1)	minADE (1)
Argoverse 1	HPNet [252]	1.09	0.75	3.47	1.60
	SEPT [151]	1.06	0.73	3.18	1.44
	QCNet [324]	1.07	0.73	3.34	1.52
Argoverse 2	QCNet [324]	1.02	0.50	2.29	0.94
	SEPT [151]	1.14	0.55	3.22	1.27
	HeteroGCN + + [77]	1.46	0.69	3.05	1.23
Waymo (2023)	MGTR (Ens) [75]	1.20	0.58	–	–
	MTR + + [236]	1.12	0.56	–	–
	AMP [125]	1.19	0.60	–	–
Nuscene	SemanticFormer [247]	–	0.87	3.88	–
	UniTraj [71]	–	0.84	5.40	–
	Goal-LBP [299]	–	0.93	9.20	–

HPNet also employs Graph Neural Networks (GNN) to model relationships between agents and road geometry features. SEPT [151] uses self-supervised learning for spatio-temporal understanding in traffic scenes. It pretrains a scene encoder with three tasks: Masked Trajectory Modeling (MTM) to capture kinematic dependencies, Masked Road Modeling (MRM) for road network topology, and Tail Prediction (TP) which is a simplified version of motion prediction aimed at understanding agent–road interactions. The pretrained encoder is then fine-tuned on motion forecasting tasks. Alternatively, QCNet [324] enhances trajectory prediction by introducing a query-centric paradigm for scene encoding. This approach reduces redundant computations and enables faster inference by learning representations independent of the global space–time coordinate system. Despite its advantages, QCNet may face challenges with training stability and motion pattern compliance.

HeteroGCN + + [77] uses dynamic heterogeneous graphs to improve motion forecasting in dynamic autonomous driving scenarios. Its main contribution is modeling evolving spatio-temporal dependencies by encoding diverse scenario components and their interactions over time. However, the approach may face challenges with computational complexity and real-time performance. MGTR [75] improves motion prediction in autonomous driving by effectively utilizing multi-granular context features from LiDAR data and other multimodal inputs. The framework employs a Transformer encoder-decoder architecture, however, its complexity and computational demands may limit real-time applications. In turn, MTR + + [236] employs transformer encoder layers with local self-attention for multi-agent motion prediction, preserving locality structure and improving memory efficiency. It integrates learnable intention queries to predict multimodal future trajectories, enhancing accuracy. However, its computational complexity and extensive training data requirements may be limiting factors. AMP [125] introduces an autoregressive motion prediction paradigm using GPT-style next-token prediction training. Its main contribution is unifying observed and future states into ego-centric token representation. However, there is a performance gap between autoregressive and independent generation, suggesting the need for incorporating classical state estimation methods.

SemanticFormer [247] predicts multimodal trajectories by reasoning over a semantic traffic scene graph using multiple attention mechanisms and graph neural networks. It employs a hierarchical heterogeneous graph encoder to capture spatio-temporal and relational information, a probability predictor for trajectory decoding, and a refinement module. The complexity of constructing and maintaining the knowledge graph and associated computational overhead are drawbacks. Alternatively, UniTraj [71] unifies various datasets, models, and evaluation criteria to address vehicle trajectory prediction challenges. Its main contribution is facilitating cross-dataset experiments, revealing significant performance drops when models transfer to new datasets. Finally, Goal-LBP [299] is a goal-based, local behavior-guided model for predicting trajectories in autonomous driving. It employs Transformer encoders and an attention mechanism to represent interactions and uses an encoder-decoder GRU model for trajectory prediction. Challenges include accurately modeling diverse future behaviors and uncertain agent intentions.

These works demonstrate that the current focus on trajectory prediction research is the integration of advanced machine learning techniques, such as self-supervised learning, attention mechanisms, and graph neural networks, to capture complex spatio-temporal dependencies and agent interactions. However, the field faces several challenges. Many SOTA models, such as those using transformers and graph neural networks, have high computational demands, which can hinder real-time applications. Models often experience significant performance drops when applied to new datasets, indicating a need for approaches that can generalize across different environments and scenarios. In addition, ensuring stable training processes, particularly for models employing novel architectures like QCNet's query-centric paradigm, remains a challenge. Extensive training data is necessary for models like MTR + +, highlighting the need for efficient data utilization and augmentation techniques. Finally, constructing and maintaining complex models, such as SemanticFormer's semantic traffic scene graph, introduces additional computational overhead and design complexity.

7. Datasets

The evaluation of software components is an important step in developing autonomous and intelligent vehicles. Before real-world testing, it is essential to verify each algorithm's performance using appropriate methodologies. For trajectory prediction, simulation platforms and datasets with real traffic information are relevant tools for development and subsequent evaluation [143]. Moreover, in scientific research, publicly available data are important to allow for comparison between different studies.

This section presents public datasets for trajectory prediction collected from primary studies mentioned earlier and newly published datasets with extensive sensor and traffic scenario data. Table 5 summarizes these datasets. Most datasets provide data from multiple sensors (e.g., cameras and LiDAR) and cover various driving conditions (e.g., highways, intersections, roundabouts, and urban streets). Additionally, some recent datasets include interactive scenarios with diverse agents (e.g., pedestrians, cyclists, and vehicles). Information on the maneuver intentions of traffic participants is also available, though only a few datasets include such annotations.

8. Remarks discussion

Different frameworks address trajectory prediction by considering the intentions and interactions of traffic participants. These frameworks use various technologies and methods, from analytical and probabilistic models to machine learning. They also differ in their inputs and outputs, using historical trajectories, road geometry, traffic laws, risk assessment metrics, and surrounding representations such as frontal cameras, bird's eye view images, or graphs. The outputs can be future trajectories (multi-modal or unimodal), grid maps, or control inputs for

Table 5
Overview of trajectory prediction datasets.

Dataset	Traffic scenarios	Sensors	Number of sequences	Sequence length	Traffic participants	Maneuvers	References	Observations
NGSIM	Highway	Trajectory, BEV images	–	45 min (I-80) 45 min (US-101)	Vehicles	Lane-change, lane-keeping, merging, yield	[129]	Very noisy [47]
HighD	Highway	Trajectory, BEV images	–	16.5 h (~13.6 s per vehicle)	Vehicles	Free-driving, vehicle following, critical maneuver, lane-change	[145]	Good weather conditions
Nuscene	Urban	Cameras, LiDAR, GPS (trajectory), IMU, RADAR, HD-Map	1000	20 s each sequence	23 object classes	Stopped, parked, moving	[23]	Various weather and lighting conditions
Lyft dataset	Urban	Cameras, LiDAR, GPS (trajectory), RADAR, HD-Map, BEV images,	170,000	25 s each sequence	Vehicles, pedestrians, riders	–	[104]	It has semantic and aerial map
ApolloScape (Trajectory)	Urban	Cameras, LiDAR, GPS (trajectory)	–	53 min (train) 50 min (test)	Vehicles, pedestrians, riders	–	[270]	Various lighting conditions and traffic densities
Argoverse 1 (Motion Forecasting)	Urban	GPS (trajectory), HD-Map	323,557	5 s each sequence	Vehicles	–	[32]	Challenging segments
Argoverse 2 (Motion Forecasting)	Urban	GPS (trajectory), HD-Map	250,000	11 s each sequence	Vehicles, pedestrians, riders	–	[282]	Kinematically and socially unusual behavior
Waymo Open (Motion)	Urban	Cameras, LiDAR, GPS (trajectory), HD-Map	103,354	20 s each sequence	Vehicles, pedestrians, riders	–	[245]	Various weather and lighting conditions
Interaction dataset	Intersections roundabouts highway	Trajectory, BEV images, HD-Map	40,054	991 min	Vehicles, pedestrians	Lane-change, lane-keeping, merging, yield, left/right turn	[306]	Highly interactive scenarios
BLVD	Urban highway	Cameras, LiDAR, GPS (trajectory), BEV images	6004	–	Vehicles, pedestrians, riders	Lane-change, lane-keeping, stopping, accelerating, decelerating, left/right turn	[294]	Various lighting conditions and traffic densities
inD	Intersections	trajectory, BEV images, HD-Map	–	589 min	Vehicles, pedestrians, riders	Left/right turn, merging, yield	[19]	Good weather conditions
TRAF	Urban	Trajectory, cameras	50	–	Vehicles, pedestrians, riders, rickshaw	–	[28]	Various lighting conditions and traffic densities
roundD	Roundabouts	Trajectory, BEV images	–	6 h	Vehicles, pedestrians, riders	–	[146]	Good weather conditions
KITTI (Tracking)	Urban	Images, LiDAR, GPS (trajectory)	21 (train) 29 (test)	–	Vehicles, pedestrians	–	[78]	Manually annotated
LOKI	Urban	Images, LiDAR, GPS (trajectory), HD-Map	644	~12.6 s Each sequence	Vehicles, pedestrians, riders	Stopped, parked, left/right turn, lane-change, lane-keeping	[85]	Intention annotation
exiD	Highway	Trajectory, BEV images, HD-Map	92	>16 h	Vehicles	Lane-change, merging	[194]	Good weather conditions
uniD	Urban	Trajectory, BEV images, HD-Map	12	–	Vehicles, pedestrians, riders	–	[18]	Good weather conditions
VisDrone	Urban	Trajectory, BEV images	400	265,228 total frames	Vehicles, pedestrians, riders	–	[328]	Various weather and lighting conditions

kinematically feasible trajectory estimation. Moreover, frameworks that consider interactions between traffic participants also use a variety of approaches, as interaction is a major challenge in autonomous navigation. This section highlights key remarks from studies on interaction-aware trajectory prediction.

- **Traffic rules**: Traffic rules significantly influence drivers' decision-making, impacting both the interaction with other traffic participants and the surrounding environment. These rules are crucial for intention prediction and interaction modeling. One remark worth mentioning is that for more realistic interaction models, a driver may disobey some traffic rules.
- **Graphs**: Graphs are versatile data structures that model both the road geometry and interactions between traffic participants. Spatio-temporal and heterogeneous graphs are particularly effective for representing temporal dependencies and interactions among different participants, such as pedestrians, cyclists, cars, and trucks.
- **Attention mechanism**: Attention mechanisms improve the performance of deep learning architectures and produce significant results in trajectory prediction. They highlight important features and differentiate the influence of each surrounding vehicle on the target vehicle. Additionally, they can extract relevant and intricate relationships between traffic participants or the multi-modal inputs of the framework.
- **Continual learning**: Continuous learning offers many possibilities for trajectory prediction. This approach can use feedback from tracking systems to improve predictor performance, or incrementally learn new interaction patterns and maneuver intentions.
- **Driving styles**: Driving styles are an important factor in trajectory prediction and tactical decision-making. Research has demonstrated

significant differences in the trajectories of drivers with different styles, such as aggressive versus calm drivers. However, predicting trajectories is challenging because it requires estimating a driver's style from only a few seconds of observation, and there is a lack of labeled datasets.

- **Traffic heterogeneity**: Traffic heterogeneity is an important challenge for trajectory predictors, as different traffic participant categories have unique decision-making processes and motion patterns. Technologies such as heterogeneous graphs, multi-task learning, and attention mechanisms can explicitly address this problem.
- **Road geometry**: Road geometry is relevant information for predicting trajectories and maneuver intentions. However, the many variations in road geometry, especially at intersections, make prediction tasks and vehicle interaction modeling challenging. Directed graphs can represent some road geometry variations by including lane directions. Bird's eye view images from High-Definition Maps (HD-Maps) also capture detailed road features. However, there is no consensus on the best way to represent roads, leaving room for future research.
- **Goal-oriented or anchor-based prediction**: One interesting approach to predicting maneuver intention and trajectory is to consider that the entire trajectory of the traffic participant can be divided into sub-goals, and take them into account during inference. Sampling techniques can estimate some goal hypotheses, which can help extract features from road geometry, motion dynamics, and spatial interaction.
- **Partially observable states and stochastic behaviors**: There are some gaps to explore in approaches that assume either partially observable states or stochastic behavior for trajectory prediction. In this case, sampling and probabilistic models can be integrated into deep learning approaches to account for the uncertainty in the observed input.
- **Highly interactive and unusual driving scenes**: Many techniques for trajectory prediction are already achieving good metrics scores. However, their benchmark focus on data from datasets with usual and simple driving scenes. Therefore, it is important to address the issue of predicting trajectory in highly interactive scenarios and, even, with scenes with unusual driving behaviors. This analysis allows a deeper analysis on the performance of state-of-art predictors.
- **Multi-task learning**: Autonomous systems encompass a range of correlated tasks, including object detection, tracking, behavior prediction, trajectory prediction, and risk assessment. Employing multitasking methodologies, such as weighted loss functions, weight sharing mechanisms, or multi-head models, offers a means to consolidate these tasks within more efficient and resilient architectures, by enabling tasks to mutually contribute with each other's learning processes.
- **Self/semi-supervised learning**: Self-supervised and semi-supervised learning methods are valuable tools for improving trajectory prediction tasks by using datasets that may lack complete labels. This is especially important given the lack of data with multi-modal information, such as behavior intentions and driving scenarios. Therefore, developing effective strategies to address these limitations can enhance predictive accuracy by providing training methodologies and incorporating self-supervision tasks.
- **Meta-learning and domain generalization**: Meta-learning and domain generalization techniques enable trajectory prediction models to adapt quickly to new environments and scenarios, while avoiding overfitting. Therefore, these techniques can improve the robustness and reliability of prediction architectures.

9. Conclusion

To improve safety in autonomous navigation, it is important not only to detect obstacles but also to predict their future trajectories. This prediction enhances environmental representation and helps foresee hazardous situations like collisions. Various methods exist for trajectory prediction. This paper reviews approaches that consider the intentions and interaction models of traffic participants.

Our analysis of key studies highlights existing efforts and outlines advancements in the field. There is considerable research potential in developing methods that account for both intention and interaction, handle diverse traffic situations, and consider interactions among heterogeneous traffic participants while adhering to traffic laws, such as signalized intersections, speed limits, and lane change permissions.

CRediT authorship contribution statement

Iago Pacheco Gomes: Writing – review & editing, Methodology, Conceptualization, Writing – original draft, Investigation, Validation, Formal analysis, Software, Data curation. **Denis Fernando Wolf:** Project administration, Writing – review & editing, Funding acquisition, Supervision, Formal analysis, Resources, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Iago Pacheco Gomes reports that financial support was provided by State of Sao Paulo Research Foundation. Iago Pacheco Gomes also reports that financial support was provided by Coordination of Higher Education Personnel Improvement. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) for the financial support under grant 88887.500344/2020-0, and the São Paulo Research Foundation (FAPESP) for the financial support under grants 2019/27301-7 and 2022/04473-0.

Data availability

No data were used for the research described in the article.

References

[1] N. AbuAli, H. Abou-Zeid, Driver behavior modeling: developments and future directions, Int. J. Vehic. Technol. (2016) 6952791.
[2] J. An, W. Liu, Q. Liu, L. Guo, P. Ren, T. Li, Dginet: dynamic graph and interaction-aware convolutional network for vehicle trajectory prediction, Neural Netw. 151 (2022) 336–348.
[3] K. Apoorva, R. Dhanya, A.K. Anjana, S. Natarajan, Trajectory forecasting of entities using advanced deep learning techniques, in: Advanced Computational and Communication Paradigms, Springer, 2018.
[4] A. Artun Edo, Decision-Making Strategies for Automated Driving in Urban Environments, Springer Nature, 2020.
[5] D. Augustin, M. Hofmann, U. Konigorski, Prediction of highway lane changes based on prototype trajectories, Forsch. im Ing. 83 (2019) 149–161.
[6] G. Aydemir, A.K. Akan, F. Güney, Adapt: efficient multi-agent trajectory prediction with adaptation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
[7] I. Bae, J. Lee, H.-G. Jeon, Can language beat numerical regression? Language-based multimodal trajectory prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024.
[8] M. Bahari, I. Nejjar, A. Alahi, Injecting knowledge in data-driven vehicle trajectory predictors, Transp. Res. Part C Emerg. Technol. 128 (2021) 103010.
[9] P. Baldi, Autoencoders, unsupervised learning, and deep architectures, in: Proceedings of ICML Workshop on Unsupervised and Transfer Learning, 2012.
[10] Y. Ban, X. Li, G. Rosman, I. Gilitschenski, O. Meireles, S. Karaman, D. Rus, A deep concept graph network for interaction-aware trajectory prediction, in: 2022 International Conference on Robotics and Automation (ICRA), 2022.
[11] A. Benterki, M. Boukhnifer, V. Judalet, C. Maaoui, Artificial intelligence for vehicle behavior anticipation: hybrid approach based on maneuver classification and trajectory prediction, IEEE Access 8 (2020) 56992–57002.
[12] V. Bharilya, N. Kumar, Machine learning for autonomous vehicle's trajectory prediction: a comprehensive survey, challenges, and future research directions, Vehic. Commun. 46 (2024) 100733.
[13] M. Bhat, J. Francis, J. Oh, Trajformer: trajectory prediction with local self-attentive contexts for autonomous driving. arXiv preprint arXiv:2011.14910, 2020.

[14] A. Bhattacharyya, M. Fritz, B. Schiele, Long-term on-board prediction of people in traffic scenes under uncertainty, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
[15] P. Bhattacharyya, C. Huang, K. Czarnecki, Ssl-lanes: self-supervised learning for motion forecasting in autonomous driving, in: Conference on Robot Learning, 2023.
[16] A. Bighashdel, G. Dubbelman, A survey on path prediction techniques for vulnerable road users: from traditional to deep-learning approaches, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019.
[17] K. Bimbraw, Autonomous cars: past, present and future a review of the developments in the last century, the present scenario and the expected future of autonomous vehicle technology, in: 2015 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO), 2015.
[18] J. Bock, R. Krajewsk, T. Moers, L. Vater (n.d.). The Unid Dataset, https://www.unid-dataset.com/
[19] J. Bock, R. Krajewski, T. Moers, S. Runde, L. Vater, L. Eckstein, The ind dataset: a drone dataset of naturalistic road user trajectories at German intersections, in: 2020 IEEE Intelligent Vehicles Symposium (IV), 2019.
[20] F.A. Boulton, E.C. Grigore, E.M. Wolff, Motion prediction using trajectory sets and self-driving domain knowledge. arXiv preprint arXiv:2006.04767, 2020.
[21] N. Brouwer, H. Kloeden, C. Stiller, Comparison and evaluation of pedestrian motion models for vehicle safety systems, in: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), 2016.
[22] T. Buhet, E. Wirbel, A. Bursuc, X. Perrotton, Plop: probabilistic polynomial objects trajectory prediction for autonomous driving, in: Conference on Robot Learning, 2021.
[23] H. Caesar, V. Bankiti, A.H. Lang, S. Vora, V.E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, O. Beijbom, Nuscenes: a multimodal dataset for autonomous driving. arXiv preprint arXiv:1903.11027, 2019.
[24] Y. Cai, Z. Wang, H. Wang, L. Chen, Y. Li, M.A. Sotelo, Z. Li, Environment-attention network for vehicle trajectory prediction, in: IEEE Transactions on Vehicular Technology, 2021.
[25] D. Cao, J. Li, H. Ma, M. Tomizuka, Spectral temporal graph neural network for trajectory prediction, in: IEEE International Conference on Robotics and Automation (ICRA), 2021, pp. 1839–1845.
[26] S. Carrasco, D.F. Llorca, M. Sotelo, Scout: Socially-consistent and Understandable Graph Attention Network for Trajectory Prediction of Vehicles and Vrus, 2021, pp. 1501–1508.
[27] Y. Chai, B. Sapp, M. Bansal, D. Anguelov, Multipath: multiple probabilistic anchor trajectory hypotheses for behavior prediction. In L.P. Kaelbling; D. Kragic, K. Sugiura, (Eds.), in: Proceedings of the Conference on Robot Learning, PMLR, Oct Nov 30–01 2020.
[28] R. Chandra, U. Bhattacharya, A. Bera, D. Manocha, Traphic: trajectory prediction in dense and heterogeneous traffic using weighted interactions, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
[29] R. Chandra, U. Bhattacharya, A. Bera, D. Manocha, Traphic: trajectory prediction in dense and heterogeneous traffic using weighted interactions, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
[30] R. Chandra, U. Bhattacharya, C. Roncal, A. Bera, D. Manocha, Robusttp: end-to-end trajectory prediction for heterogeneous road-agents in dense traffic with noisy sensor inputs, in: ACM Computer Science in Cars Symposium, 2019.
[31] R. Chandra, T. Guan, S. Panuganti, T. Mittal, U. Bhattacharya, A. Bera, D. Manocha, Forecasting trajectory and behavior of road-agents using spectral clustering in graph-lstms, in: IEEE Robotics and Automation Letters, 2020.
[32] M.-F. Chang, J.W. Lambert, P. Sangkloy, J. Singh, S. Bak, A. Hartnett, D. Wang, P. Carr, S. Lucey, D. Ramanan, J. Hays, Argoverse: 3d tracking and forecasting with rich maps, in: Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
[33] G. Chen, L. Hu, Q. Zhang, Z. Ren, X. Gao, J. Cheng, St-lstm: spatio-temporal graph based long short-term memory network for vehicle trajectory prediction, in: 2020 IEEE International Conference on Image Processing (ICIP), 2020.
[34] K. Chen, X. Chen, Z. Yu, M. Zhu, H. Yang, Equidiff: a conditional equivariant diffusion model for trajectory prediction, in: 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC), 2023.
[35] S. Chen, Z. Ma, X. Zhu, C. Wang, L. Zheng, L. Huang, J. Bai, Mmtp: multi-modal trajectory prediction with interaction attention and adaptive task weighting, in: 2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC), 2022.
[36] X. Chen, H. Zhang, F. Zhao, Y. Cai, H. Wang, Q. Ye, Vehicle trajectory prediction based on intention-aware non-autoregressive transformer with multi-attention learning for internet of vehicles, in: IEEE Transactions on Instrumentation and Measurement, 2022.
[37] H. Cheng, W. Liao, X. Tang, M.Y. Yang, M. Sester, B. Rosenhahn, Exploring dynamic context for multi-path trajectory prediction, in: 2021 IEEE International Conference on Robotics and Automation (ICRA), 2021.
[38] H. Cheng, W. Liao, M.Y. Yang, B. Rosenhahn, M. Sester, Amenet: attentive maps encoder network for trajectory prediction, ISPRS J. Photogramm. Remote Sens. 172 (2021) 253–266.
[39] H. Cheng, W. Liao, M.Y. Yang, M. Sester, B. Rosenhahn, MCENET: multi-context encoder network for homogeneous agent trajectory prediction in mixed traffic, in: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), 2020.
[40] H. Cheng, M. Liu, L. Chen, H. Broszio, M. Sester, M.Y. Yang, Gatraj: a graph-and attention-based multi-agent trajectory prediction model, ISPRS J. Photogramm. Remote Sens. 205 (2023) 163–175.
[41] H. Cheng, M. Sester, Mixed traffic trajectory prediction using lstm-based models in shared space, in: The Annual International Conference on Geographic Information Science, 2018.
[42] H. Cheng, M. Sester, Modeling mixed traffic in shared space using lstm with probability density mapping, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018.
[43] J. Cheng, X. Mei, M. Liu, Forecast-mae: self-supervised pre-training for motion forecasting with masked autoencoders, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
[44] Y. Cheng, W. Zhao, C. Liu, M. Tomizuka, Human motion prediction using semi-adaptable neural networks, in: 2019 American Control Conference (ACC), 2019.
[45] K. Cho, T. Ha, G. Lee, S. Oh, Deep predictive autonomous driving using multi-agent joint trajectory prediction and traffic rules, in: IROS, 2019.
[46] C. Choi, S. Malla, A. Patil, J.H. Choi, Drogon: a trajectory prediction model based on intention-conditioned behavior reasoning, in: PMLR (Proceedings of Machine Learning Research), 2021, pp. 49–63.
[47] B. Coifman, L. Li, A critical evaluation of the next generation simulation (ngsim) vehicle trajectory dataset, Transp. Res. Part B Methodol. 105 (2017) 362–377.
[48] F.-A. Croitoru, V. Hondru, R.T. Ionescu, M. Shah, Diffusion models in vision: a survey, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023.
[49] A. Cui, S. Casas, A. Sadat, R. Liao, R. Urtasun, Lookout: diverse multi-future prediction and planning for self-driving, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021.
[50] A. Cui, S. Casas, K. Wong, S. Suo, R. Urtasun, Gorela: go relative for viewpoint-invariant motion forecasting, in: 2023 IEEE International Conference on Robotics and Automation (ICRA), 2023.
[51] H. Cui, V. Radosavljevic, F.-C. Chou, T.-H. Lin, T. Nguyen, T.-K. Huang, J. Schneider, N. Djuric, Multimodal trajectory predictions for autonomous driving using deep convolutional networks, in: International Conference on Robotics and Automation, 2019.
[52] S. Dai, L. Li, Z. Li, Modeling vehicle interactions via modified lstm models for trajectory prediction, IEEE Access 7 (2019) 38287–38296.
[53] V.M. Dax, J. Li, E. Sachdeva, N. Agarwal, M.J. Kochenderfer, Disentangled neural relational inference for interpretable motion prediction, in: IEEE Robotics and Automation Letters, 2023.
[54] N. Deo, A. Rangesh, M.M. Trivedi, How would surround vehicles move? a unified framework for maneuver classification and motion prediction, in: IEEE Transactions on Intelligent Vehicles, 2018.
[55] N. Deo, M.M. Trivedi, Convolutional social pooling for vehicle trajectory prediction, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018.
[56] N. Deo, M.M. Trivedi, Multi-modal trajectory prediction of surrounding vehicles with maneuver based lstms, in: 2018 IEEE Intelligent Vehicles Symposium (IV), 2018.
[57] N. Deo, E. Wolff, O. Beijbom, Multimodal trajectory prediction conditioned on lane-graph traversals, in: Conference on Robot Learning, 2022, pp. 203–212.
[58] F. Diehl, T. Brunner, M.T. Le, A. Knoll, Graph neural networks for modelling traffic participant interaction, in: 2019 IEEE Intelligent Vehicles Symposium (IV), 2019.
[59] N. Djuric, V. Radosavljevic, H. Cui, T. Nguyen, F.-C. Chou, T.-H. Lin, N. Singh, J. Schneider, Uncertainty-aware short-term motion prediction of traffic actors for autonomous driving, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2020.
[60] C. Dong, Y. Chen, J.M. Dolan, Interactive trajectory prediction for autonomous driving via recurrent meta induction neural network, in: International Conference on Robotics and Automation (ICRA), 2019, https://doi.org/10.1109/ICRA.2019.8794392.
[61] A. Doshi, M.M. Trivedi, Tactical driver behavior prediction and intent inference: a review, in: 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2011.
[62] R.G. Dsouza, Deep Learning Based Motion Forecasting for Autonomous Driving (Master's thesis), The Ohio State University, 2021.
[63] L. Du, Z. Wang, L. Wang, Z. Zhao, F. Su, B. Zhuang, N.V. Boulgouris, Adaptive visual interaction based multi-target future state prediction for autonomous driving vehicles, in: IEEE Transactions on Vehicular Technology, 2019.
[64] A. Dulian, J.C. Murray, Multi-modal anticipation of stochastic trajectories in a dynamic environment with conditional variational autoencoders. arXiv preprint arXiv:2103.03912, 2021.
[65] D.J. Fagnant, K. Kockelman, Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations, Transp. Res. Part A Policy Pract. 77 (2015) 167–181.
[66] J. Fang, F. Wang, J. Xue, T.-S. Chua, Behavioral intention prediction in driving scenes: a survey, in: IEEE Transactions on Intelligent Transportation Systems, 2024.
[67] L. Fang, Q. Jiang, J. Shi, B. Zhou, Tpnet: trajectory proposal network for motion prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.
[68] C. Fei, X. He, X. Ji, Multi-modal vehicle trajectory prediction based on mutual information, IET Intell. Transp. Syst. 2019 14, 148–153.
[69] C. Fei, X. He, S. Kawahara, N. Shirou, X. Ji, Conditional wasserstein auto-encoder for interactive vehicle trajectory prediction, in: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), 2020.
[70] C. Feng, H. Zhou, H. Lin, Z. Zhang, Z. Xu, C. Zhang, B. Zhou, S. Shen, Macformer: map-agent coupled transformer for real-time and robust trajectory prediction, in: IEEE Robotics and Automation Letters, 2023.
[71] L. Feng, M. Bahari, K.M.B. Amor, É. Zablocki, M. Cord, A. Alahi, Unitraj: A Unified Framework for Scalable Vehicle Trajectory Prediction, 2024, pp. 106–123.

[72] X. Feng, Z. Cen, J. Hu, Y. Zhang, Vehicle trajectory prediction using intention-based conditional variational autoencoder, in: IEEE Intelligent Transportation Systems Conference (ITSC), 2019.
[73] T. Fernando, S. Denman, S. Sridharan, C. Fookes, Deep inverse reinforcement learning for behavior prediction in autonomous driving: accurate forecasts of vehicle motion, IEEE Signal Processing Magazine, 2020.
[74] M. Fu, T. Zhang, W. Song, Y. Yang, M. Wang, Trajectory prediction-based local spatio-temporal navigation map for autonomous driving in dynamic highway environments, IEEE Trans. Intell. Transp. Syst. 23 (2021) 6418–6429.
[75] Y. Gan, H. Xiao, Y. Zhao, E. Zhang, Z. Huang, X. Ye, L. Ge, Multi-Granular Transformer for Motion Prediction with Lidar, 2024, pp. 15092–15098.
[76] J. Gao, C. Sun, H. Zhao, Y. Shen, D. Anguelov, C. Li, C. Schmid, Vectornet: encoding hd maps and agent dynamics from vectorized representation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.
[77] X. Gao, X. Jia, Y. Li, H. Xiong, Dynamic scenario representation learning for motion forecasting with heterogeneous graph convolutional recurrent networks, in: IEEE Robotics and Automation Letters, 2023.
[78] A. Geiger, P. Lenz, R. Urtasun, Are we ready for autonomous driving? the kitti vision benchmark suite, in: Conference on Computer Vision and Pattern Recognition (CVPR), 2012.
[79] H. Gharoun, F. Momenifar, F. Chen, A.H. Gandomi, Meta-learning approaches for few-shot learning: a survey of recent advances, ACM Comput. Surv. 56 (12) (2024) 1–41.
[80] J.S. Gill, P. Pisu, M.J. Schmid, A probabilistic framework for trajectory prediction in traffic utilizing driver characterization, in: 2019 IEEE 2nd Connected and Automated Vehicles Symposium (CAVS), 2019.
[81] T. Gilles, S. Sabatini, D. Tsishkou, B. Stanciulescu, F. Moutarde, Home: Heatmap output for future motion estimation, in: International Intelligent Transportation Systems Conference (ITSC), 2021.
[82] T. Gilles, S. Sabatini, D. Tsishkou, B. Stanciulescu, F. Moutarde, Thomas: Trajectory Heatmap Output with Learned Multi-Agent Sampling, 2021.
[83] T. Gilles, S. Sabatini, D. Tsishkou, B. Stanciulescu, F. Moutarde, Gohome: graph-oriented heatmap output for future motion estimation, in: 2022 International Conference on Robotics and Automation (ICRA), 2022.
[84] K. Gillmeier, F. Diederichs, D. Spath, Prediction of ego vehicle trajectories based on driver intention and environmental context, in: 2019 IEEE Intelligent Vehicles Symposium (IV), 2019.
[85] H. Girase, H. Gang, S. Malla, J. Li, A. Kanehara, K. Mangalam, C. Choi, Loki: long term and key intentions for trajectory prediction, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021.
[86] R. Girgis, F. Golemo, F. Codevilla, M. Weiss, J.A. D'Souza, S.E. Kahou, F. Heide, C. Pal, Latent Variable Sequential Set Transformers for Joint Multi-Agent Motion Prediction, 2021.
[87] C. Gómez-Huélamo, M.V. Conde, R. Barea, L.M. Bergasa, Improving multi-agent motion prediction with heuristic goals and motion refinement, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
[88] C. G'Omez-Hu'elamo, M.V. Conde, M. Ortiz, Exploring map-based features for efficient attention-based vehicle motion prediction. arXiv preprint arXiv:2205.13071, 2022.
[89] C. G'Omez-Hu'elamo, M.V. Conde, M. Ortiz, S. Montiel, R. Barea, L.M. Bergasa, Exploring attention Gan for Vehicle Motion Prediction, 2022, pp. 4011–4016.
[90] D.S. Gonz'alez, J.S. Dibangoye, C. Laugier, High-speed highway scene prediction based on driver models learned from demonstrations, in: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), 2016.
[91] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative Adversarial Networks, 2014.
[92] J. Gu, C. Hu, T. Zhang, X. Chen, Y. Wang, Y. Wang, H. Zhao, Vip3d: end-to-end visual trajectory prediction via 3d agent queries, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
[93] J. Gu, C. Sun, H. Zhao, Densetnt: end-to-end trajectory prediction from dense goal sets, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021.
[94] C. Guo, S. Fan, C. Chen, W. Zhao, J. Wang, Y. Zhang, Y. Chen, Query-informed multi-agent motion prediction, Sensors 24 (1) (2023) 9.
[95] Y. Guo, V.V. Kalidindi, M. Arief, W. Wang, J. Zhu, H. Peng, D. Zhao, Modeling multi-vehicle interaction scenarios using gaussian random field, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019.
[96] K. Han, Y. Wang, H. Chen, X. Chen, J. Guo, Z. Liu, Y. Tang, A. Xiao, C. Xu, Y. Xu, et al. A survey on visual transformer. arXiv preprint arXiv:2012.12556, 2020.
[97] Z. Hao, X. Huang, K. Wang, M. Cui, Y. Tian, Attention-based gru for driver intention recognition and vehicle trajectory prediction, in: 2020 4th CAA International Conference on Vehicular Control and Intelligence (CVCI), 2020.
[98] M. Hasan, A. Solernou, E. Paschalidis, H. Wang, G. Markkula, R. Romano, Maneuver-aware pooling for vehicle trajectory prediction. arXiv preprint arXiv:2104.14079, 2021.
[99] G. He, X. Li, Y. Lv, B. Gao, H. Chen, Probabilistic intention prediction and trajectory generation based on dynamic bayesian networks, in: 2019 Chinese Automation Congress (CAC), 2019.
[100] C. Hegde, S. Dash, P. Agarwal, Vehicle trajectory prediction using gan, in: 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2020.
[101] L. Hou, S.E. Li, B. Yang, Z. Wang, K. Nakano, Structural transformer improves speed-accuracy trade-off in interactive trajectory prediction of multiple surrounding vehicles, in: IEEE Transactions on Intelligent Transportation Systems, 2022.
[102] L. Hou, L. Xin, S.E. Li, B. Cheng, W. Wang, Interactive trajectory prediction of surrounding road users for autonomous driving using structural-lstm network, in: IEEE Transactions on Intelligent Transportation Systems, 2019.
[103] A. Houenou, P. Bonnifait, V. Cherfaoui, W. Yao, Vehicle trajectory prediction based on motion model and maneuver recognition, in: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2013.
[104] J. Houston, G. Zuidhof, L. Bergamini, Y. Ye, L. Chen, A. Jain, S. Omari, V. Iglovikov, P. Ondruska, One thousand and one hours: self-driving motion prediction dataset. arXiv preprint arXiv:2006.14480, 2020.
[105] H. Hu, Q. Wang, L. Du, Z. Lu, Z. Gao, Vehicle trajectory prediction considering aleatoric uncertainty, in: Knowledge-Based Systems, 2022.
[106] X. Hu, M. Zheng, Research progress and prospects of vehicle driving behavior prediction, World Electr. Veh. J. 12 (2) (2021) 88.
[107] Y. Hu, L. Sun, M. Tomizuka, Generic prediction architecture considering both rational and irrational driving behaviors, in: IEEE Intelligent Transportation Systems Conference (ITSC), 2019.
[108] Y. Hu, W. Zhan, L. Sun, M. Tomizuka, Multi-modal probabilistic prediction of interactive behavior via an interpretable model, in: 2019 IEEE Intelligent Vehicles Symposium (IV), 2019.
[109] Y. Hu, W. Zhan, M. Tomizuka, A framework for probabilistic generic traffic scene prediction, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018.
[110] Y. Hu, W. Zhan, M. Tomizuka, Probabilistic prediction of vehicle semantic intention and motion, in: 2018 IEEE Intelligent Vehicles Symposium (IV), 2018.
[111] Y. Hu, W. Zhan, M. Tomizuka, Scenario-transferable semantic graph reasoning for interaction-aware probabilistic prediction, IEEE Trans. Intell. Transp. Syst. 23 (12) (2022) 23212–23230.
[112] R. Huang, H. Xue, M. Pagnucco, F.D. Salim, Y. Song, Vision-based multi-future trajectory prediction: a survey, in: IEEE Transactions on Neural Networks and Learning Systems, 2025.
[113] X. Huang, S.G. McGill, B.C. Williams, L. Fletcher, G. Rosman, Uncertainty-aware driver trajectory prediction at urban intersections, in: International Conference on Robotics and Automation (ICRA), 2019, https://doi.org/10.1109/ICRA.2019.8794282.
[114] Y. Huang, J. Du, Z. Yang, Z. Zhou, L. Zhang, H. Chen, A survey on trajectory-prediction methods for autonomous driving, IEEE Trans. Intell. Veh. 7 (3) (2022) 652–674.
[115] Z. Huang, H. Liu, C. Lv, Gameformer: game-theoretic modeling and learning of transformer-based interactive prediction and planning for autonomous driving, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
[116] Z. Huang, X. Mo, C. Lv, Multi-modal motion prediction with transformer-based neural network for autonomous driving, in: International Conference on Robotics and Automation (ICRA), 2022.
[117] Z. Huang, J. Wu, C. Lv, Driving behavior modeling using naturalistic human driving data with inverse reinforcement learning, in: IEEE Transactions on Intelligent Transportation Systems, 2021.
[118] M. Huynh, G. Alaghband, Aol: adaptive online learning for human trajectory prediction in dynamic video scenes. arXiv preprint arXiv:2002.06666, 2020.
[119] B. Ivanovic, A. Elhafsi, G. Rosman, A. Gaidon, M. Pavone, Mats: an interpretable trajectory forecasting representation for planning and control. In J. Kober; F. Ramos, C. Tomlin, (Eds.), in: Proceedings of the 2020 Conference on Robot Learning, vol. 155, PMLR, 2021 Nov 16–18, pp. 2243–2256.
[120] B. Ivanovic, J. Harrison, M. Pavone, Expanding the deployment envelope of behavior prediction via adaptive meta-learning, in: IEEE International Conference on Robotics and Automation (ICRA), 2023, https://doi.org/10.1109/ICRA48891.2023.10161155.
[121] B. Ivanovic, K.-H. Lee, P. Tokmakov, B. Wulfe, R. Mcllister, A. Gaidon, M. Pavone, Heterogeneous-agent trajectory forecasting incorporating class uncertainty, (2022) 12196–12203.
[122] S. Jawed, A. Rashed, L. Schmidt-Thieme, Multi-step forecasting via multi-task learning, in: 2019 IEEE International Conference on Big Data (Big Data), 2019.
[123] H. Jeon, J. Choi, D. Kum, Scale-net: scalable vehicle trajectory prediction network under random number of interacting vehicles via edge-enhanced graph convolutional neural network, in: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.
[124] Y. Jeong, K. Yi, Bidirectional long shot-term memory-based interactive motion prediction of cut-in vehicles in urban environments, IEEE Access 8 (2020) 106183–106197.
[125] X. Jia, S. Shi, Z. Chen, L. Jiang, W. Liao, T. He, J. Yan, Amp: autoregressive motion prediction revisited with next token prediction for autonomous driving. arXiv preprint arXiv:2403.13331, 2024.
[126] C. Jiang, A. Cornman, C. Park, B. Sapp, Y. Zhou, D. Anguelov, et al. Motiondiffuser: controllable multi-agent motion prediction using diffusion, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
[127] T. Jiang, Y. Liu, Q. Dong, T. Xu, Intention-aware interactive transformer for real-time vehicle trajectory prediction in dense traffic, Transp. Res. Rec. 2677 (3) (2022) 946–960.
[128] R. Jiao, X. Liu, T. Sato, Q.A. Chen, Q. Zhu, Semi-supervised semantics-guided adversarial training for robust trajectory prediction, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
[129] J.C. John Halkias, Next Generation Simulation (ngsim), 2020. http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm
[130] C. Ju, Z. Wang, C. Long, X. Zhang, D.E. Chang, Interaction-aware kalman neural networks for trajectory prediction, in: 2020 IEEE Intelligent Vehicles Symposium (IV), 2020.

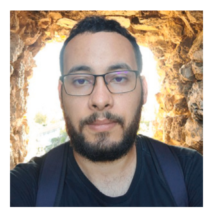
[131] N. Kamra, H. Zhu, D.K. Trivedi, M. Zhang, Y. Liu, Multi-agent trajectory prediction with fuzzy query attention, Adv. Neural Inf. Process. Syst. 33 (2020) 22530–22541.
[132] E. Kargar, V. Kyrki, Increasing the efficiency of policy learning for autonomous vehicles by multi-task representation learning, in: IEEE Transactions on Intelligent Vehicles, 2022.
[133] R. Karim, S.M.A. Shabestary, A. Rasouli, Destine: Dynamic Goal Queries With Temporal Transductive Alignment for Trajectory Prediction, 2024, pp. 2230–2237.
[134] C. Katrakazas, M. Quddus, W.-H. Chen, L. Deka, Real-time motion planning methods for autonomous on-road driving: state-of-the-art and future research directions, Transp. Res. Part C Emerg. Technol. 60 (2015) 416–442
[135] A. Kawasaki, A. Seki, Multimodal trajectory predictions for urban environments using geometric relationships between a vehicle and lanes, in: 2020 IEEE International Conference on Robotics and Automation (ICRA), 2020, https://doi.org/10.1109/ICRA40945.2020.9196738.
[136] M. Khakzar, A. Rakotonirainy, A. Bond, S.G. Dehkordi, A dual learning model for vehicle trajectory prediction, IEEE Access 8 (2020) 21897–21908.
[137] A.G. Khoee, Y. Yu, R. Feldt, Domain generalization through meta-learning: a survey, Artif. Intell. Rev. 57 (10) (2024) 285.
[138] S. Khosla, Z. Zhu, Y. He, Survey on memory-augmented neural networks: cognitive insights to ai applications. arXiv preprint arXiv:2312.06141, 2023.
[139] B. Kim, S.H. Park, S. Lee, E. Khoshimjonov, D. Kum, J. Kim, J.S. Kim, J.W. Choi, Lapred: lane-aware prediction of multi-modal future trajectories of dynamic agents, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
[140] G. Kim, D. Kim, Y. Ahn, K. Huh, Hybrid approach for vehicle trajectory prediction using weighted integration of multiple models, IEEE Access 9 (2021) 78715–78723.
[141] H. Kim, D. Kim, G. Kim, J. Cho, K. Huh, Multi-head attention based probabilistic vehicle trajectory prediction, in: 2020 IEEE Intelligent Vehicles Symposium (IV), 2020, https://doi.org/10.1109/IV47402.2020.9304741.
[142] S. Kim, D. Kum, J. Won Choi, Recup net: recursive prediction network for surrounding vehicle trajectory prediction with future trajectory feedback, in: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), 2020.
[143] P. Koopman, M. Wagner, Autonomous vehicle safety: an interdisciplinary challenge, in: IEEE Intelligent Transportation Systems Magazine, 2017.
[144] Y. Kossale, M. Airaj, A. Darouichi, Mode collapse in generative adversarial networks: an overview, in: 2022 8th International Conference on Optimization and Applications (ICOA), 2022.
[145] R. Krajewski, J. Bock, L. Kloeker, L. Eckstein, The highd dataset: a drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems, in: 21st International Conference on Intelligent Transportation Systems, 2018.
[146] R. Krajewski, T. Moers, J. Bock, L. Vater, L. Eckstein, The round dataset: a drone dataset of road user trajectories at roundabouts in Germany, in: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), 2020.
[147] M. Krüger, A.S. Novo, T. Nattermann, T. Bertram, Interaction-aware trajectory prediction based on a 3d spatio-temporal tensor representation using convolutional–recurrent neural networks, in: 2020 IEEE Intelligent Vehicles Symposium (IV), 2020.
[148] S. Kumar, Y. Gu, J. Hoang, G.C. Haynes, M. Marchetti-Bowick, Interaction-based trajectory prediction over a hybrid traffic graph, in: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2021.
[149] Y.-L. Kuo, X. Huang, A. Barbu, S.G. McGill, B. Katz, J.J. Leonard, G. Rosman, Trajectory prediction with linguistic representations, in: 2022 International Conference on Robotics and Automation (ICRA), 2022.
[150] N. Lamm, S. Jaiprakash, M. Srikanth, I. Drori, Vehicle trajectory prediction by transfer learning of semi-supervised models. arXiv preprint arXiv:2007.06781, 2020.
[151] Z. Lan, Y. Jiang, Y. Mu, C. Chen, S.E. Li, Sept: towards efficient scene representation learning for motion prediction, in: The 12th International Conference on Learning Representations, 2023.
[152] B. Lange, J. Li, M.J. Kochenderfer, Scene Informer: anchor-Based Occlusion Inference and Trajectory Prediction in Partially Observable Environments, 2024, pp. 14138–14145.
[153] N. Lee, W. Choi, P. Vernaza, C.B. Choy, P.H. Torr, M. Chandraker, Desire: distant future prediction in dynamic scenes with interacting agents, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[154] S. Lefévre, D. Vasquez, C. Laugier, A survey on motion prediction and risk assessment for intelligent vehicles, ROBOMECH J. 1 (2014) 1–14.
[155] V. Lefkopoulos, M. Menner, A. Domahidi, M.N. Zeilinger, Interaction-aware motion prediction for autonomous driving: a multiple model kalman filtering scheme, IEEE Robot. Autom. Lett. 6 (2021) 80–87.
[156] F. Leon, M. Gavrilescu, A review of tracking, prediction and decision making methods for autonomous driving, arXiv preprint arXiv:1909.07707, 2019.
[157] T. Lesort, V. Lomonaco, A. Stoian, D. Maltoni, D. Filliat, N. Díaz-Rodríguez, Continual learning for robotics: definition, framework, learning strategies, opportunities and challenges, Inf. Fusion. 58 (2020) 52–68.
[158] J. Li, H. Ma, M. Tomizuka, Conditional generative neural system for probabilistic trajectory prediction, in: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019a.
[159] J. Li, H. Ma, M. Tomizuka, Interaction-aware multi-agent tracking and probabilistic behavior prediction via adversarial learning, in: International Conference on Robotics and Automation, 2019b.
[160] J. Li, H. Ma, W. Zhan, M. Tomizuka, Coordination and trajectory prediction for vehicle interactions via bayesian generative modeling, in: 2019 IEEE Intelligent Vehicles Symposium (IV), 2019.
[161] J. Li, H. Ma, Z. Zhang, J. Li, M. Tomizuka, Spatio-temporal graph dual-attention network for multi-agent prediction and tracking, IEEE Trans. Intell. Transp. Syst. 23 (8) (2021) 10556–10569.
[162] J. Li, H. Ma, Z. Zhang, M. Tomizuka, Social-wagdat: interaction-aware trajectory prediction via wasserstein graph double-attention network. arXiv preprint arXiv:2002.06241, 2020.
[163] J. Li, W. Zhan, Y. Hu, M. Tomizuka, Generic tracking and probabilistic prediction framework and its application in autonomous driving, in: IEEE Transactions on Intelligent Transportation Systems, 2019.
[164] L. Li, W. Zhao, C. Wang, Q. Chen, F. Chen, Bram-ed: vehicle trajectory prediction considering the change of driving behavior, in: IEEE/ASME Transactions on Mechatronics, 2022.
[165] L.L. Li, B. Yang, M. Liang, W. Zeng, M. Ren, S. Segal, R. Urtasun, End-to-end contextual perception and prediction with interaction transformer, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.
[166] L. Li, X. Wang, D. Yang, Y. Ju, Z. Zhang, J. Lian, Real-time heterogeneous road-agents trajectory prediction using hierarchical convolutional networks and multi-task learning, in: IEEE Transactions on Intelligent Vehicles, 2023.
[167] R. Li, C. Li, D. Ren, G. Chen, Y. Yuan, G. Wang, Bcdiff: bidirectional consistent diffusion for instantaneous trajectory prediction, Adv. Neural Inf. Process. Syst. 36 (2024) 14400–14413.
[168] R. Li, H. Shi, Z. Fu, Z. Wang, G. Lin, Weakly supervised class-agnostic motion prediction for autonomous driving, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
[169] X. Li, G. Rosman, I. Gilitschenski, C.-I. Vasile, J.A. DeCastro, S. Karaman, D. Rus, Vehicle trajectory prediction using generative adversarial network with temporal logic syntax tree features, in: IEEE Robotics and Automation Letters, 2021.
[170] X. Li, F. Huang, Z. Fan, F. Mou, Y. Hou, C. Qian, L. Wen, Metatra: meta-learning for generalized trajectory prediction in unseen domain. arXiv preprint arXiv:2402.08221, 2024.
[171] X. Li, X. Ying, M.C. Chuah, Grip: graph-based interaction-aware trajectory prediction, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019.
[172] Z. Li, J. Gong, C. Lu, Y. Yi, Interactive behavior prediction for heterogeneous traffic participants in the urban road: a graph-neural-network-based multitask learning framework, in: IEEE/ASME Transactions on Mechatronics, 2021.
[173] Z. Li, Y. Lin, C. Gong, X. Wang, Q. Liu, J. Gong, C. Lu, An ensemble learning framework for vehicle trajectory prediction in interactive scenarios, in: IEEE Intelligent Vehicles Symposium, 2022, https://doi.org/10.1109/IV51971.2022.9827070.
[174] M. Liang, B. Yang, R. Hu, Y. Chen, R. Liao, S. Feng, R. Urtasun, Learning lane graph representations for motion forecasting, in: Computer Vision-ECCV 2020: 16th European Conference, August 23–28, 2020, Proceedings, Part II 16, Glasgow, UK, 2020, pp. 541–556.
[175] H. Liao, S. Liu, Y. Li, Z. Li, C. Wang, Y. Li, S.E. Li, C. Xu, Human Observation-Inspired Trajectory Prediction for Autonomous Driving in Mixed-Autonomy Traffic Environments, 2024, pp. 14212–14219.
[176] J. Liu, Y. Luo, H. Xiong, T. Wang, H. Huang, Z. Zhong, An integrated approach to probabilistic vehicle trajectory prediction via driver characteristic and intention estimation, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019.
[177] S. Liu, L. Li, J. Tang, S. Wu, J.-L. Gaudiot, Creating Autonomous Vehicle Systems, Morgan Claypool Publishers, 2017.
[178] Y. Liu, B. Li, X. Wang, C. Sammut, L. Yao, Attention-aware social graph transformer networks for stochastic trajectory prediction, in: IEEE Transactions on Knowledge and Data Engineering, 2024.
[179] Y. Liu, J. Zhang, L. Fang, Q. Jiang, B. Zhou, Multimodal motion prediction with stacked transformers, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
[180] Y. Liu, X. Qi, E.A. Sisbot, K. Oguchi, Multi-agent trajectory prediction with graph attention isomorphism neural network, in: 2022 IEEE Intelligent Vehicles Symposium (IV), 2022.
[181] Y. Lu, W. Wang, X. Hu, P. Xu, S. Zhou, M. Cai, Vehicle trajectory prediction in connected environments via heterogeneous context-aware graph convolutional networks, in: IEEE Transactions on Intelligent Transportation Systems, 2022.
[182] Y. Ma, S. Zhao, W. Wang, Y. Li, I. King, Multimodality in meta-learning: a comprehensive survey, Knowl.-Based Syst. 250 (2022) 108976.
[183] Y. Ma, X. Zhu, S. Zhang, R. Yang, W. Wang, D. Manocha, Trafficpredict: trajectory prediction for heterogeneous traffic-agents, in: Proceedings of the AAAI Conference on Artificial Intelligence, 2019.
[184] F. Marchetti, F. Becattini, L. Seidenari, A. Del Bimbo, Multiple trajectory prediction of moving agents with memory augmented networks, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.
[185] F. Marchetti, F. Becattini, L. Seidenari, A. Del Bimbo, Smemo: social memory for trajectory forecasting, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.
[186] J. Mercat, T. Gilles, N. El Zoghby, G. Sandou, D. Beauvois, G.P. Gil, Multi-head attention for multi-modal joint vehicle motion forecasting, in: 2020 IEEE International Conference on Robotics and Automation (ICRA), 2020.
[187] K. Messaoud, N. Deo, M.M. Trivedi, F. Nashashibi, Trajectory prediction for autonomous driving based on multi-head attention with joint agent-map representation, in: 2021 IEEE Intelligent Vehicles Symposium (IV), 2021.

[188] K. Messaoud, I. Yahiaoui, A. Verroust, F. Nashashibi, Attention based vehicle trajectory prediction, in: IEEE Transactions on Intelligent Vehicles, 2020.
[189] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, F. Nashashibi, Non-local social pooling for vehicle trajectory prediction, in: 2019 IEEE Intelligent Vehicles Symposium (IV), 2019.
[190] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, F. Nashashibi, Relational recurrent neural networks for vehicle trajectory prediction, in: IEEE Intelligent Transportation Systems Conference (ITSC), 2019.
[191] K. Min, H. Kim, J. Park, D. Kim, K. Huh, Interaction aware trajectory prediction of surrounding vehicles with interaction network and deep ensemble, in: 2020 IEEE Intelligent Vehicles Symposium (IV), 2020.
[192] X. Mo, Y. Xing, C. Lv, Interaction-aware trajectory prediction of connected vehicles using cnn-lstm networks, in: The 46th Annual Conference of the IEEE Industrial Electronics Society, 2020a.
[193] X. Mo, Y. Xing, C. Lv, Recog: a deep learning framework with heterogeneous graph for interaction-aware trajectory prediction. arXiv preprint arXiv:2012.05032, 2020.
[194] T. Moers, L. Vater, R. Krajewski, J. Bock, A. Zlocki, L. Eckstein, The exid dataset: a real-world trajectory dataset of highly interactive highway scenarios in Germany, in: 2022 IEEE Intelligent Vehicles Symposium (IV), 2022, https://doi.org/10.1109/IV51971.2022.9827305
[195] T. Mortlock, A.V. Malawade, K. Tsujio, M.A. Al Faruque, Castnet: a context-aware, spatio-temporal dynamic motion prediction ensemble for autonomous driving, ACM Trans. Cyber-Phys. Syst. 8 (2) (2024) 1–20, https://doi.org/10.1145/3648622
[196] S. Mozaffari, O.Y. Al-Jarrah, M. Dianati, P. Jennings, A. Mouzakitis, Deep learning-based vehicle behavior prediction for autonomous driving applications: a review, in: IEEE Transactions on Intelligent Transportation Systems, 2020.
[197] N. Mu, J. Ji, Z. Yang, N. Harada, H. Tang, K. Chen, C.R. Qi, R. Ge, K. Goel, Z. Yang, et al. Most: multi-modality scene tokenization for motion prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024.
[198] S. Mukherjee, S. Wang, A. Wallace, Interacting vehicle trajectory prediction with convolutional recurrent neural networks, in: International Conference on Robotics and Automation (ICRA), 2020.
[199] N. Nayakanti, R. Al-Rfou, A. Zhou, K. Goel, K.S. Refaat, B. Sapp, Wayformer: motion forecasting via simple efficient attention networks, in: 2023 IEEE International Conference on Robotics and Automation (ICRA), 2023.
[200] J. Ngiam, B. Caine, V. Vasudevan, Z. Zhang, H.-T.L. Chiang, J. Ling, R. Roelofs, A. Bewley, C. Liu, A. Venugopal, et al. Scene Transformer: A Unified Architecture for Predicting Multiple Agent Trajectories, 2022.
[201] Z. Niu, G. Zhong, H. Yu, A review on the attention mechanism of deep learning, Neurocomputing 452 (2021) 48–62.
[202] K. Okamoto, K. Berntorp, S. Di Cairano, Driver intention-based vehicle threat assessment using random forests and particle filtering, IFAC-PapersOnLine 50 (1) (2017) 13860–13865.
[203] J. Pan, H. Sun, K. Xu, Y. Jiang, X. Xiao, J. Hu, J. Miao, Lane-attention: predicting vehicles' moving trajectories by learning their attention over lanes, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.
[204] Z. Pang, D. Ramanan, M. Li, Y.-X. Wang, Streaming motion forecasting for autonomous driving, in: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2023.
[205] S. Paravarzar, B. Mohammad, Motion prediction on self-driving cars: a review. arXiv preprint arXiv:2011.03635, 2020.
[206] G.I. Parisi, R. Kemker, J.L. Part, C. Kanan, S. Wermter, Continual lifelong learning with neural networks: a review, Neural Netw. 113 (2019) 54–71.
[207] D. Park, J. Jeong, S.-H. Yoon, J. Jeong, K.-J. Yoon, T4p: Test-Time Training of Trajectory Prediction via Masked Autoencoder and Actor-Specific Token Memory, 2024, pp. 15065–15076.
[208] D. Park, J. Jeong, K.-J. Yoon, Improving transferability for cross-domain trajectory prediction via neural stochastic differential equation, in: Proceedings of the AAAI Conference on Artificial Intelligence, 2024.
[209] D. Park, H. Ryu, Y. Yang, J. Cho, J. Kim, K.-J. Yoon, Leveraging Future Relationship Reasoning for Vehicle Trajectory Prediction, 2023.
[210] S.H. Park, G. Lee, J. Seo, M. Bhat, M. Kang, J. Francis, A. Jadhav, P.P. Liang, L.-P. Morency, Diverse and admissible trajectory forecasting through multimodal context understanding, in: European Conference on Computer Vision, 2020.
[211] B. Pathiraja, S. Munasinghe, M. Ranawella, M. De Silva, R. Rodrigo, P. Jayasekara, Class-aware attention for multimodal trajectory prediction. arXiv preprint arXiv:2209.00062, 2022.
[212] T. Phan-Minh, E.C. Grigore, F.A. Boulton, O. Beijbom, E.M. Wolff, Covernet: multimodal behavior prediction using trajectory sets, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020.
[213] J. Phillips, J. Martinez, I.A. Bârsan, S. Casas, A. Sadat, R. Urtasun, Deep multi-task learning for joint localization, perception, and prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
[214] M. Pourkeshavarz, C. Chen, A. Rasouli, Learn tarot with mentor: a meta-learned self-supervised approach for trajectory prediction, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
[215] S. Qin, J. Zhu, J. Qin, W. Wang, D. Zhao, Recurrent attentive neural process for sequential data. arXiv preprint arXiv:1910.09323, 2019.
[216] A. Quintanar, D. Fernández-Llorca, I. Parra, R. Izquierdo, M. Sotelo, Predicting vehicles trajectories in urban scenarios with transformer networks and augmented information, (2021) 1051–1056.
[217] X. Ren, T. Yang, L.E. Li, A. Alahi, Q. Chen, Safety-aware motion prediction with unseen vehicles for autonomous driving, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021.
[218] N. Rhinehart, R. McAllister, K. Kitani, S. Levine, Precog: prediction conditioned on goals in visual multi-agent settings, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019.
[219] D. Ridel, E. Rehder, M. Lauer, C. Stiller, D. Wolf, A literature review on the prediction of pedestrian behavior in urban scenarios, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018.
[220] J. Roh, C. Mavrogiannis, R. Madan, D. Fox, S. Srinivasa, Multimodal trajectory prediction via topological invariance for navigation at uncontrolled intersections. In J. Kober; F. Ramos, C. Tomlin, (Eds.), in: Proceedings of the 2020 Conference on Robot Learning, PMLR, Nov 2021, pp. 16–18.
[221] L. Rowe, M. Ethier, E.-H. Dykhne, K. Czarnecki, Fjmp: factorized joint multi-agent motion prediction over learned directed acyclic interaction graphs, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
[222] D. Roy, T. Ishizaka, C.K. Mohan, A. Fukuda, Vehicle trajectory prediction at intersections using interaction based generative adversarial networks, in: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019.
[223] A. Rudenko, L. Palmieri, M. Herman, K.M. Kitani, D.M. Gavrila, K.O. Arras, Human motion trajectory prediction: a survey, Int. J. Robot. Res. 39 (8) (2020) 895–935.
[224] A. Sadeghian, F. Legros, M. Voisin, R. Vesel, A. Alahi, S. Savarese, Car-net: clairvoyant attentive recurrent network, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018.
[225] SAE, S, J3016 Standard: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems, 2014.
[226] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, S. Valaee, Recent advances in recurrent neural networks. arXiv preprint arXiv:1801.01078, 2017.
[227] T. Salzmann, B. Ivanovic, P. Chakravarty, M. Pavone, Trajectron++: dynamically-feasible trajectory forecasting with heterogeneous data, In A. Vedaldi; H. Bischof; T. Brox, J.-M. Frahm, (Eds.), Computer Vision – ECCV, Springer International Publishing, 2020.
[228] M. Schreier, V. Willert, J. Adamy, Bayesian, maneuver-based, long-term trajectory prediction and criticality assessment for driver assistance systems, in: 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2014.
[229] J. Schulz, C. Hubmann, J. Löchner, D. Burschka, Interaction-aware probabilistic behavior prediction in urban environments, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018.
[230] J. Schulz, C. Hubmann, J. Löchner, D. Burschka, Multiple model unscented kalman filtering in dynamic bayesian networks for intention estimation and trajectory prediction, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018b.
[231] W. Schwarting, A. Pierson, J. Alonso-Mora, S. Karaman, D. Rus, Social behavior for autonomous vehicles, in: Proceedings of the National Academy of Sciences, 2019.
[232] A. Seff, B. Cera, D. Chen, M. Ng, A. Zhou, N. Nayakanti, K.S. Refaat, R. Al-Rfou, B. Sapp, Motionlm: multi-agent motion forecasting as language modeling, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
[233] M. Shah, Z. Huang, A. Laddha, M. Langford, B. Barber, S. Zhang Sida, C. Vallespi-Gonzalez, R. Urtasun, Liranet: end-to-end trajectory prediction using spatio-temporal radar fusion. In J. Kober; F. Ramos, C. Tomlin, (Eds.), in: Proceedings of the 2020 conference on robot learning, vol. 155, PMLR, Nov 2021, pp. 16–18 (pp. 31–48.
[234] W. Shao, J. Xu, Z. Cao, H. Wang, J. Li, Uncertainty-aware Prediction and Application in Planning for Autonomous Driving: Definitions, Methods, and Comparison, 2024.
[235] Z. Sheng, Y. Xu, S. Xue, D. Li, Graph-based spatial-temporal convolutional network for vehicle trajectory prediction in autonomous driving, in: IEEE Transactions on Intelligent Transportation Systems, 2022.
[236] S. Shi, L. Jiang, D. Dai, B. Schiele, Mtr++: multi-agent motion prediction with symmetric scene modeling and guided intention querying, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.
[237] X. Shi, H. Zhang, W. Yuan, R. Shibasaki, Metatraj: meta-learning for cross-scene cross-object trajectory prediction, in: IEEE Transactions on Intelligent Transportation Systems, 2023.
[238] W. Si, T. Wei, C. Liu, Agen: adaptable generative prediction networks for autonomous driving, in: 2019 IEEE Intelligent Vehicles Symposium (IV), 2019.
[239] H. Song, W. Ding, Y. Chen, S. Shen, M.Y. Wang, Q. Chen, Pip: planning-informed trajectory prediction for autonomous driving, in: European Conference on Computer Vision, 2020.
[240] W. Song, B. Su, G. Xiong, S. Li, Intention-aware decision making in urban lane change scenario for autonomous driving, in: 2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES), 2018.
[241] N. Sriram, B. Liu, F. Pittaluga, M. Chandraker, Smart: Simultaneous multi-agent recurrent trajectory prediction, in: European Conference on Computer Vision, 2020.
[242] J. Strohbeck, V. Belagiannis, J. Müller, M. Schreiber, M. Herrmann, D. Wolf, M. Buchholz, Multiple trajectory prediction with deep temporal and spatial convolutional neural networks, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020.
[243] J. Sun, Y. Li, H.-S. Fang, C. Lu, Three steps to multimodal trajectory prediction: modality clustering, classification and synthesis, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021.

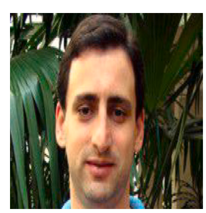
[244] L. Sun, W. Zhan, M. Tomizuka, Probabilistic prediction of interactive driving behavior via hierarchical inverse reinforcement learning, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018.
[245] P. Sun, H. Kretzschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, V. Vasudevan, W. Han, J. Ngiam, H. Zhao, A. Timofeev, S. Ettinger, M. Krivokon, A. Gao, A. Joshi, D. Anguelov, Scalability in perception for autonomous driving: waymo open dataset, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.
[246] Q. Sun, S. Zhang, D. Ma, J. Shi, D. Li, S. Luo, Y. Wang, N. Xu, G. Cao, H. Zhao, Large trajectory models are scalable motion predictors and planners. arXiv preprint arXiv:2310.19620, 2023.
[247] Z. Sun, Z. Wang, L. Halilaj, J. Luettin, Semanticformer: holistic and semantic traffic scene representation for trajectory prediction using knowledge graphs, in: IEEE Robotics and Automation Letters, 2024.
[248] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, MIT Press, 2018.
[249] C. Tang, J. Chen, M. Tomizuka, Adaptive probabilistic vehicle trajectory prediction through physically feasible bayesian recurrent neural network, in: 2019 International Conference on Robotics and Automation (ICRA), 2019, https://doi.org/10.1109/ICRA.2019.8794130.
[250] L. Tang, F. Yan, B. Zou, W. Li, C. Lv, K. Wang, Trajectory prediction for autonomous driving based on multiscale spatial-temporal graph, IET Intell. Transp. Syst. 17 (2022) 386–399.
[251] X. Tang, K. Yang, H. Wang, J. Wu, Y. Qin, W. Yu, D. Cao, Prediction-uncertainty-aware decision-making for autonomous vehicles, in: IEEE Transactions on Intelligent Vehicles, 2022, https://doi.org/10.1109/TIV.2022.3188662.
[252] X. Tang, M. Kan, S. Shan, Z. Ji, J. Bai, X. Chen, Hpnet: Dynamic Trajectory Forecasting with Historical Prediction Attention, 2024, pp. 15261–15270.
[253] Y. Tang, H. He, Y. Wang, Hierarchical vector transformer vehicle trajectories prediction with diffusion convolutional neural networks, Neurocomputing 580 (2024) 127526.
[254] I. Teeti, S. Khan, A. Shahbaz, A. Bradley, F. Cuzzolin, Vision-based intention and trajectory prediction in autonomous vehicles: a survey, in: Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, 2022, https://doi.org/10.24963/ijcai.2022/785.
[255] E. Tolstaya, R. Mahjourian, C. Downey, B. Vadarajan, B. Sapp, D. Anguelov, Identifying driver interactions via conditional behavior prediction, in: 2021 IEEE International Conference on Robotics and Automation (ICRA), 2021.
[256] Q. Tran, J. Firl, Modelling of traffic situations at urban intersections with probabilistic non-parametric regression, in: 2013 IEEE Intelligent Vehicles Symposium (IV), 2013.
[257] Q. Tran, J. Firl, Online maneuver recognition and multimodal trajectory prediction for intersection assistance using non-parametric regression, in: 2014 IEEE Intelligent Vehicles Symposium Proceedings, 2014.
[258] S. Vandenhende, S. Georgoulis, W. Van Gansbeke, M. Proesmans, D. Dai, L. Van Gool, Multi-task learning for dense prediction tasks: a survey, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
[259] B. Varadarajan, A. Hefny, A. Srivastava, K.S. Refaat, N. Nayakanti, A. Cornman, K. Chen, B. Douillard, C.P. Lam, D. Anguelov, et al. Multipath++: efficient information fusion and trajectory aggregation for behavior prediction, in: 2022 International Conference on Robotics and Automation (ICRA), 2022.
[260] D. Vasquez, T. Fraichard, C. Laugier, Growing hidden markov models: an incremental tool for learning and predicting human and vehicle motion, Int. J. Robot. Res. 28 (11–12) (2009a) 1486–1506.
[261] D. Vasquez, T. Fraichard, C. Laugier, Incremental learning of statistical motion patterns with growing hidden markov models, in: IEEE Transactions on Intelligent Transportation Systems, 2009b.
[262] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need. arXiv preprint arXiv:1706.03762, 2017.
[263] S. Veer, A. Sharma, M. Pavone, Multi-predictor fusion: combining learning-based and rule-based trajectory predictors, in: Conference on Robot Learning, 2023.
[264] J. Wang, Z. Liu, L. Zhao, Z. Wu, C. Ma, S. Yu, H. Dai, Q. Yang, Y. Liu, S. Zhang, et al. Review of large vision models and visual prompt engineering, Meta-Radiol. (3) (2023) 100047.
[265] J. Wang, P. Wang, C. Zhang, K. Su, J. Li, F-net: fusion neural network for vehicle trajectory prediction in autonomous driving, in: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021.
[266] K. Wang, Y. Wu, J. Cen, Z. Pan, X. Li, Z. Wang, Z. Cao, G. Lin, Self-supervised class-agnostic motion prediction with spatial and temporal consistency regularizations, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024.
[267] L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, Z. Chen, J. Tang, X. Chen, Y. Lin, et al. A survey on large language model based autonomous agents, Front. Comput. Sci. 18 (6) (2024) 186345.
[268] L. Wang, Y. Hu, L. Sun, W. Zhan, M. Tomizuka, C. Liu, Transferable and adaptable driving behavior prediction. arXiv preprint arXiv:2202.05140, 2022.
[269] M. Wang, X. Zhu, C. Yu, W. Li, Y. Ma, R. Jin, X. Ren, D. Ren, M. Wang, W. Yang, Ganet: goal area network for motion forecasting, in: 2023 IEEE International Conference on Robotics and Automation (ICRA), 2023.
[270] P. Wang, X. Huang, X. Cheng, D. Zhou, Q. Geng, R. Yang, The apolloscape open dataset for autonomous driving and its application, in: IEEE transactions on pattern analysis and machine intelligence, 2019.
[271] S. Wang, P. Zhao, B. Yu, W. Huang, H. Liang, Vehicle trajectory prediction by knowledge-driven lstm network in urban environments, J. Adv. Transp. 2020 (2020) 1–20.
[272] S. Wang, Y. Chen, J. Cheng, X. Mei, R. Xin, Y. Song, M. Liu, Improving autonomous driving safety with pop: a framework for accurate partially observed trajectory predictions, pp. 14450–14456 2024.
[273] W. Wang, J. Shen, Deep visual attention prediction, in: IEEE Transactions on Image Processing, 2017.
[274] X. Wang, J. Liu, H. Lin, S. Garg, M. Alrashoud, A multi-modal spatial–temporal model for accurate motion forecasting with visual fusion, Inf. Fusion 102 (2024) 102046.
[275] X. Wang, T. Su, F. Da, X. Yang, Prophnet: efficient agent-centric motion forecasting with anchor-informed proposals, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
[276] Y. Wang, S. Zhao, R. Zhang, X. Cheng, L. Yang, Multi-vehicle collaborative learning for trajectory prediction with spatio-temporal tensor fusion, in: IEEE Transactions on Intelligent Transportation Systems, 2020.
[277] Z. Wang, J. Guo, H. Zhang, R. Wan, J. Zhang, J. Pu, Bridging the gap: improving domain generalization in trajectory prediction, in: IEEE Transactions on Intelligent Vehicles, 2023.
[278] X. Weng, Y. Yuan, K. Kitani, Ptp: parallelized tracking and prediction with graph neural networks and diversity sampling, in: IEEE Robotics and Automation Letters, 2021.
[279] T. Westny, B. Olofsson, E. Frisk, Diffusion-based environment-aware trajectory prediction. arXiv preprint arXiv:2403.11643, 2024.
[280] T. Westny, J. Oskarsson, B. Olofsson, E. Frisk, Mtp-go: graph-based probabilistic multi-agent trajectory prediction with neural odes, in: IEEE Transactions on Intelligent Vehicles, 2023.
[281] J. Wiederer, J. Schmidt, U. Kressel, K. Dietmayer, V. Belagiannis, Joint out-of-distribution detection and uncertainty estimation for trajectory prediction, in: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2023.
[282] B. Wilson, W. Qi, T. Agarwal, J. Lambert, J. Singh, S. Khandelwal, B. Pan, R. Kumar, A. Hartnett, J.K. Pontes, D. Ramanan, P. Carr, J. Hays, Argoverse 2: next generation datasets for self-driving perception and forecasting, in: Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks (NeurIPS Datasets and Benchmarks 2021), 2021.
[283] F. Wirthmüller, J. Schlechtriemen, J. Hipp, M. Reichert, Teaching vehicles to anticipate: a systematic study on probabilistic behavior prediction using large data sets, in: IEEE Transactions on Intelligent Transportation Systems, 2020.
[284] C. Wissing, T. Nattermann, K.-H. Glander, T. Bertram, Interaction-aware long-term driving situation prediction, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018.
[285] S. Woo, M. Kim, D. Kim, S. Jang, S. Lee, Fimp: Future Interaction Modeling for Multi-Agent Motion Prediction, 2024, pp. 14457–14463.
[286] P. Wu, S. Chen, D.N. Metaxas, Motionnet: joint perception and motion prediction for autonomous driving based on bird's eye view maps, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020.
[287] G. Xie, H. Gao, L. Qian, B. Huang, K. Li, J. Wang, Vehicle trajectory prediction by integrating physics-and maneuver-based approaches using interactive multiple models, in: IEEE Transactions on Industrial Electronics, 2017.
[288] L. Xin, P. Wang, C.-Y. Chan, J. Chen, S.E. Li, B. Cheng, Intention-aware long horizon trajectory prediction of surrounding vehicles using dual lstm networks, in: 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2018.
[289] Y. Xing, C. Lv, D. Cao, Personalized vehicle trajectory prediction based on joint time-series modeling for connected vehicles, in: IEEE Transactions on Vehicular Technology, 2019.
[290] Y. Xing, C. Lv, H. Wang, H. Wang, Y. Ai, D. Cao, E. Velenis, F.-Y. Wang, Driver lane change intention inference for intelligent vehicles: framework, survey, and challenges, in: IEEE Transactions on Vehicular Technology, 2019.
[291] S. Xingjian, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, W.-C. Woo, Convolutional lstm network: a machine learning approach for precipitation nowcasting, Adv. Neural Inf. Process. Syst. 28 (2015) 1–9.
[292] C. Xu, W. Mao, W. Zhang, S. Chen, Remember intentions: retrospective-memory-based trajectory prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022.
[293] D. Xu, X. Shang, Y. Liu, H. Peng, H. Li, Group vehicle trajectory prediction with global spatio-temporal graph, in: IEEE Transactions on Intelligent Vehicles, 2022.
[294] J. Xue, J. Fang, T. Li, B. Zhang, P. Zhang, Z. Ye, J. Dou, Blvd: building a large-scale 5d semantics benchmark for autonomous driving, in: 2019 International Conference on Robotics and Automation (ICRA), 2019.
[295] J. Yan, Z. Peng, H. Yin, J. Wang, X. Wang, Y. Shen, W. Stechele, D. Cremers, Trajectory prediction for intelligent vehicles using spatial-attention mechanism, IET Intell. Transp. Syst. 14 (13) (2020) 1855–1863.
[296] B. Yang, F. Fan, R. Ni, J. Li, L. Kiong, X. Liu, Continual learning-based trajectory prediction with memory augmented networks, Knowl.-Based Syst. 258 (2022) 110022.
[297] B. Yang, F. Fan, R. Ni, H. Wang, A. Jafaripournimchahi, H. Hu, A multi-task learning network with a collision-aware graph transformer for traffic-agents trajectory prediction, in: IEEE Transactions on Intelligent Transportation Systems, 2024.
[298] Y. Yao, M. Xu, C. Choi, D.J. Crandall, E.M. Atkins, B. Dariush, Egocentric vision-based future vehicle localization for intelligent driving assistance systems, in: 2019 International Conference on Robotics and Automation (ICRA), 2019.
[299] Z. Yao, X. Li, B. Lang, M.C. Chuah, Goal-lbp: goal-based local behavior guided trajectory prediction for autonomous driving, in: IEEE Transactions on Intelligent Transportation Systems, 2023.

[300] L. Ye, Z. Wang, X. Chen, J. Wang, K. Wu, K. Lu, Gsan: graph self-attention network for interaction measurement in autonomous driving, in: 2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), 2020.
[301] M. Ye, T. Cao, Q. Chen, Tpcn: temporal point cloud networks for motion forecasting, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.
[302] M. Ye, J. Xu, X. Xu, T. Wang, T. Cao, Q. Chen, Bootstrap motion forecasting with self-consistent constraints, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
[303] J. Yu, M. Zhou, X. Wang, G. Pu, C. Cheng, B. Chen, A dynamic and static context-aware attention network for trajectory prediction, ISPRS Int. J. Geo-Inf. 10 (5) (2021) 336.
[304] R. Yuan, M. Abdel-Aty, Q. Xiang, Z. Wang, X. Gu, A temporal multi-gate mixture-of-experts approach for vehicle trajectory and driving intention prediction, in: IEEE Transactions on Intelligent Vehicles, 2023.
[305] Y. Yuan, X. Weng, Y. Ou, K.M. Kitani, Agentformer: agent-aware transformers for socio-temporal multi-agent forecasting, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021.
[306] W. Zhan, L. Sun, D. Wang, H. Shi, A. Clausse, M. Naumann, J. Kummerle, H. Konigshof, C. Stiller, A. de La Fortelle, et al. Interaction dataset: an international, adversarial and cooperative motion dataset in interactive driving scenarios with semantic maps. arXiv preprint arXiv:1910.03088, 2019.
[307] C. Zhang, H. Sun, C. Chen, Y. Guo, Banet: motion forecasting with boundary aware network. arXiv preprint arXiv:2206.07934, 2022.
[308] H. Zhang, Y. Wang, J. Liu, C. Li, T. Ma, C. Yin, A multi-modal states based vehicle descriptor and dilated convolutional social pooling for vehicle trajectory prediction. arXiv preprint arXiv:2003.03480, 2020.
[309] K. Zhang, X. Feng, L. Wu, Z. He, Trajectory prediction for autonomous driving using spatial-temporal graph attention transformer, in: IEEE Transactions on Intelligent Transportation Systems, 2022.
[310] K. Zhang, L. Li, Explainable multimodal trajectory prediction using attention models, Transp. Res. Part C Emerg. Technol. 143 (2022) 103829.
[311] L. Zhang, W. Xiao, Z. Zhang, D. Meng, Surrounding vehicles motion prediction for risk assessment and motion planning of autonomous vehicle in highway scenarios, IEEE Access, 8 (2020) 209356–209376.
[312] L. Zhang, P.-H. Su, J. Hoang, G.C. Haynes, M. Marchetti-Bowick, Map-adaptive goal-based trajectory prediction, in: Conference on Robot Learning, 2021.
[313] L. Zhang, P. Li, J. Chen, S. Shen, Trajectory prediction with graph-based dual-scale context fusion, in: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2022.
[314] T. Zhang, M. Fu, W. Song, Y. Yang, M. Wang, Trajectory prediction based on constraints of vehicle kinematics and social interaction, in: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020.
[315] T. Zhang, W. Song, M. Fu, Y. Yang, M. Wang, Vehicle motion prediction at intersections based on the turning intention and prior trajectories model, IEEE/CAA J. Autom. Sin. 8 (2021) 1657–1666.
[316] Y. Zhang, Q. Yang, A survey on multi-task learning, in: IEEE Transactions on Knowledge and Data Engineering, 2021.
[317] Z. Zhang, A. Liniger, C. Sakaridis, F. Yu, L.V. Gool, Real-time motion prediction via heterogeneous polyline transformer with relative pose encoding, Adv. Neural Inf. Process. Syst. 36 (2024) 57481–57499.
[318] C. Zhao, A. Song, Y. Du, B. Yang, Trajgat: a map-embedded graph attention network for real-time vehicle trajectory imputation of roadside perception, Transp. Res. Part C Emerg. Technol. 142 (2022) 103787.
[319] H. Zhao, J. Gao, T. Lan, C. Sun, B. Sapp, B. Varadarajan, Y. Shen, Y. Shen, Y. Chai, C. Schmid, et al. Tnt: target-driven trajectory prediction, Conf. Robot Learn. 155 (2021) 895–904.
[320] Q. Zhao, C. Koch, Learning saliency-based visual attention: a review, Signal Process. 93 (6) (2013) 1401–1407.
[321] Z. Zhao, H. Fang, Z. Jin, Q. Qiu, Gisnet: graph-based information sharing network for vehicle trajectory prediction, in: 2020 International Joint Conference on Neural Networks (IJCNN), 2020.
[322] Y. Zhong, Z. Ni, S. Chen, U. Neumann, Aware of the history: trajectory forecasting with the local behavior data, in: European Conference on Computer Vision, 2022, pp. 393–409.
[323] Y. Zhou, H. Shao, L. Wang, S.L. Waslander, H. Li, Y. Liu, Smartrefine: a scenario-adaptive refinement framework for efficient motion prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024.
[324] Z. Zhou, J. Wang, Y.-H. Li, Y.-K. Huang, Query-centric trajectory prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.
[325] Z. Zhou, L. Ye, J. Wang, K. Wu, K. Lu, Hivt: hierarchical vector transformer for multi-agent motion prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.
[326] D. Zhu, G. Zhai, Y. Di, F. Manhardt, H. Berkemeyer, T. Tran, N. Navab, F. Tombari, B. Busam, Ipcc-tp: utilizing incremental pearson correlation coefficient for joint multi-agent trajectory prediction, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
[327] J. Zhu, S. Qin, W. Wang, D. Zhao, Probabilistic trajectory prediction for autonomous vehicles with attentive recurrent neural process. arXiv preprint arXiv:1910.08102, 2019.
[328] P. Zhu, L. Wen, D. Du, X. Bian, H. Fan, Q. Hu, H. Ling, Detection and tracking meet drones challenge, IEEE Trans. Pattern Anal. Mach. Intell. (2021), https://doi.org/10.1109/TPAMI.2021.3119563
[329] Y. Zhu, D. Luan, S. Shen, Biff: Bi-level future fusion with polyline-based coordinate for interactive trajectory prediction, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023.
[330] Q. Zou, Y. Hou, K. Xiong, An overview of the motion prediction of traffic participants for host vehicle, in: 2019 Chinese Control Conference (CCC), 2019.

Author biography



Iago Pacheco Gomes He received a PhD degree in Computer Science in the Institute of Mathematics and Computer Science at the University of São Paulo (ICMC/USP) in 2024. He obtained his MS degree in Computer Science from the University of São Paulo (ICMC/USP) in 2019. His current research interests are Machine Learning, Autonomous Vehicles, Trajectory and Behavior Prediction.



Denis Fernando Wolf He received a PhD degree in Computer Science at the University of Southern California - USC in 2006. He is currently an Associate Professor in the Department of Computer Systems at the University of São Paulo (ICMC-USP). His current research interests are Mobile Robotics, Machine Learning, Computer Vision and Embedded Systems.