

# 融合标精地图先验的地图不确定性感知及拓扑关系推理

张想, 彭力

(江南大学物联网工程学院, 江苏 无锡 214215)

**摘要:** 自动驾驶场景理解是自动驾驶技术的关键环节之一, 地图感知和拓扑关系推理是场景理解的重要组成部分。地图感知任务主要包括道路元素感知和交通元素感知, 拓扑关系推理在地图感知的基础上构建感知结果间的拓扑关系。然而, 在传感器受到遮挡或感知范围超过传感器范围时, 传统方法的地图感知性能会受到影响。同时, 由于拓扑关系推理依赖于地图感知结果, 地图感知误差会进一步影响拓扑关系推理的准确性。为此, 提出了一种融合标精地图先验的地图感知不确定性建模方法, 并基于地图感知不确定性实现鲁棒的拓扑关系推理。首先, 通过引入标精地图先验信息, 有效提升了遮挡场景下的地图感知性能。随后, 使用 Laplace 分布建模地图感知结果, 实现了对地图感知不确定性的建模。最后, 基于地图感知结果及其不确定性, 提出了一种基于概率的拓扑关系推理方法, 有效提升了拓扑关系构建精度。在公开数据集 OpenLaneV2 上进行大量实验, 结果表明, 所提方法在地图感知和拓扑关系推理任务上性能均优于对比方法。

**关键词:** 自动驾驶; 地图感知; 不确定性建模; 拓扑关系推理; 特征融合

中图分类号: TP391

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0070312

## Map Uncertainty Perception and Topological Relationship Reasoning Integrating Standard Map Priors

ZHANG Xiang, PENG Li

(School of Internet of Things, Jiangnan University, Wuxi 214215, Jiangsu, China)

**【Abstract】** Autonomous driving scene understanding is one of the critical components of autonomous driving technology. Map perception and topological relationship inference are essential parts of scene understanding. Map perception tasks mainly include road element perception and traffic element perception. Topological relationship reasoning builds topological relationships between perception results based on map perception. However, traditional methods face challenges in map perception performance when sensors are occluded, or the perception range exceeds the sensor's limit. Additionally, since topological relationship inference in driving scenes relies on map perception results, errors in map perception can further impact the accuracy of topological relationship inference. To address this, a map perception uncertainty modeling method incorporating standard map priors is proposed, along with robust topological relationship inference based on map perception uncertainty. First, by introducing high-precision map prior information, the method effectively enhances map perception performance in occluded scenarios. Next, the map perception results are modeled using the Laplace distribution to achieve uncertainty modeling of map perception. Finally, a probabilistic topological relationship inference method is proposed based on the map perception results and their uncertainties, which effectively improves the accuracy of topological relationship construction. Extensive experiments conducted on public datasets demonstrate that the proposed method outperforms comparative methods in both map perception and topological relationship inference tasks.

**【Key words】** autonomous driving; map perception; uncertainty modeling; topological relationship reasoning; feature fusion

## 0 引言

自动驾驶技术是未来智慧交通系统的重要组成部分, 近年来受到了广泛关注<sup>[1-2]</sup>。场景理解是自动驾驶的关键技术之一, 包括地图感知<sup>[3-4]</sup>和拓扑关系推理。地图感知是自动驾驶系统获取周围道路环境

信息的重要手段, 通过地图感知, 自动驾驶系统可以获取道路、车道线、交通标志等信息<sup>[5-6]</sup>。地图感知任务包括对车道线位置、车道线类型等道路相关元素的感知以及对红绿灯状态、道路指示牌类别等交通相关元素的感知。拓扑关系推理是自动驾驶系统对周围环境进行理解的重要组成部分, 通过拓扑关

基金项目: 国家自然科学基金面上项目(61873112)。

作者简介: 张想, 男, 硕士研究生, 主研方向为模式识别; 彭力(通信作者), 教授、博士。

收稿日期: 2024-09-02

修回日期: 2024-10-22

E-mail: penglimail2002@163.com

系推理,自动驾驶系统可以推断车道与车道、红绿灯、交通标志等之间的关系<sup>[7]</sup>。通过地图感知和拓扑关系推理,自动驾驶系统可以更好地理解周围环境,如在复杂路口中,不同红绿灯控制不同车道的通行状态,且每条车道受到对应交通标识的限制。

目前,地图感知和拓扑关系推理技术已经取得了一定的进展。STSU (Structured Bird's-Eye-View Traffic Scene Understanding)<sup>[8]</sup>在单目摄像头传感器输入的基础上抽取地图特征,利用类似 DETR<sup>[9]</sup>的架构检测车道线,通过贝塞尔曲线拟合车道线检测结果。VectorMapNet<sup>[10]</sup>在多摄像头传感器输入的基础上,利用序列点列表征道路元素,使用类似 DETR 的解码器预测关键点,并采用自回归范式生成后续点列。MapTR<sup>[11]</sup>采用基于统一排列的建模方式,将每个地图元素建模为一个点集,并通过使用一组等价的排列来处理几何上的不确定性。MapTR 在建模过程中消除了序列点列表示道路元素产生的歧义,确保更准确的几何结构表示,并且在计算效率上也有显著提升。上述工作只关注对地图的感知,并未对红绿灯、交通标志等交通元素进行感知,且未对道路元素和交通元素之间的拓扑关系进行推理。在上述工作的基础上,许多研究工作进一步探索了交通元素感知以及拓扑关系推理。TopoNet<sup>[12]</sup>提出一种场景知识图模型,通过 Transformer 编码器<sup>[13]</sup>分别编码地图元素和交通元素,进而通过图神经网络推理地图元素和交通元素之间的拓扑关系。TopoMLP<sup>[14]</sup>在 TopoNet 的基础上,不再使用场景知识图建模推理拓扑关系,转而使用多层感知器(MLP)网络推理拓扑关系,实现了更高的计算效率。LaneSegNet<sup>[15]</sup>提出一种车道线注意力机制和查询向量初始化方法,并将道路元素表示替换为车道实例表示,更好地融合了道路元素的几何特征和拓扑信息,实现了端到端的地图感知和拓扑关系推理。

然而,现有工作在地图感知和拓扑关系推理中仍然存在一些问题<sup>[16-17]</sup>。首先,现有纯视觉地图感知方案依赖摄像头传感器输入。当地图元素被其他车辆或物体遮挡时,模型无法有效感知当前地图元素,容易造成感知结果的部分缺失,为下游规划决策模块带来挑战。同时,WU 等<sup>[14]</sup>证明了拓扑关系推理任务的性能依赖于地图感知任务的性能,故为模型提供额外先验信息以提升模型对被遮挡地图元素的感知性能对于拓扑关系推理任务至关重要。其次,现有工作输出的地图感知结果为确定性结果,未对地图感知的不确定性进行建模<sup>[18-19]</sup>。在训练过程

中,若被感知地图元素受到大面积遮挡,模型则无法从输入中提取有效信息,但训练时仍然会以真实标签监督模型。这会使得模型无法有效学习到输入与输出之间的映射关系,导致模型无法收敛或学习到错误的映射关系。而引入不确定性可以在受到遮挡或其他噪声干扰时,将模型输出结果的置信度降低,从而避免在这些情况下对模型进行过度监督。最后,现有的拓扑关系推理方法无法有效地利用地图感知的不确定性信息,导致拓扑关系推理的性能受到限制。

为解决上述问题,本文提出一种融合标精地图先验的地图感知不确定性建模及拓扑关系推理方法。为了提升地图感知的性能,本文提出在感知阶段引入标精地图先验。标精地图具有较低的获取和更新成本,同时包含了关键的道路结构信息和语义信息。将标精地图作为先验信息,可以为模型提供详细的道路结构和语义数据,从而弥补在纯视觉地图感知方案中因部分地图元素被遮挡导致的输入信息缺失,提升模型的感知性能,为下游规划决策模块提供优质感知结果。为了提升地图感知的鲁棒性,本文提出在感知阶段通过 Laplace 分布<sup>[20]</sup>建模道路元素,从而通过 Laplace 分布的方差引入地图感知的不确定性。同时,本文提出在拓扑关系推理阶段融合地图感知的不确定性信息,提升拓扑关系推理的性能。不确定性建模能够提高模型对异常值的鲁棒性,对于受干扰较大的车道线,例如在阴影、强光或者道路标志重叠的情况下,这种建模方式可以更好地处理异常值,从而减少错误的检测结果。为了在拓扑关系推理阶段融合地图感知的不确定性信息,本文提出了一种基于概率的拓扑关系推理方法,以此在拓扑关系推理阶段引入地图感知的不确定性信息,提升拓扑关系推理的性能。本文的主要贡献如下:

1) 本文在地图感知阶段引入标精地图先验,给出了一种融合标精地图先验的地图感知方法,提升了地图感知的性能。

2) 在此基础上,本文提出了一种基于 Laplace 分布的地图感知不确定性建模方法,实现了对地图感知结果的不确定性建模。

3) 本文提出了一种基于概率的拓扑关系推理方法,以此在拓扑关系推理阶段引入地图感知的不确定性信息,提升拓扑关系推理的性能。

## 1 地图感知和拓扑关系推理任务描述

地图感知和拓扑关系推理任务是自动驾驶系

统中的重要组成部分。在自动驾驶任务中, 自车需要感知周围的道路环境, 包括道路元素和交通元素, 用于规划自车的行驶轨迹。除此之外, 自车还需要推理道路元素和交通元素之间的拓扑关系, 用于理解周围道路环境, 保证自车规划的行驶轨迹符合交通规则。如在复杂路口中, 不同红绿灯控制不同车道的通行状态, 且每条车道受对应交通标识的限制。

因此, 地图感知和拓扑关系推理任务主要可分为两个阶段: 地图感知阶段和拓扑关系推理阶段。地图感知阶段需要模型根据摄像头传感器输入数据, 检测自车周围的道路元素和交通元素。道路元素包括车道中心线、车道分隔线、路沿、人行横道线等。其中, 车道分隔线需要检测其类型, 如实线、虚线。道路元素检测结果由道路元素在三维自车坐标系下的坐标表示。交通元素包括红绿灯、交通标志、地面标志等。其中, 红绿灯需要检测其类型, 如红灯、绿灯、黄灯; 交通标志需要检测其类型, 如直行、禁止左转、禁止右转、禁止调头; 地面标志需要检测其类型, 如左转箭头、右转箭头、直行箭头、调头箭头。交通元素的检测结果由其在二维像素坐标系下

的边界框表示。拓扑关系推理阶段需要模型根据地图感知的结果推理道路元素之间的关系以及道路元素和交通元素之间的关系。道路元素之间的关系包括车道线与车道线之间的挂接关系、车道分叉关系、车道合并关系。道路元素和交通元素之间的关系包括车道线与红绿灯之间的关系、车道线与交通标志之间的关系。拓扑关系推理的结果使用图结构表示, 图中的节点表示道路元素和交通元素, 图中的边表示道路元素和道路元素之间的关系以及道路元素和交通元素之间的关系。

## 2 主要方法

本文提出一种融合标精地图先验的地图感知不确定性建模及拓扑关系推理方法。该方法的主要流程如图 1 所示(彩色效果见《计算机工程》官网 HTML 版, 下同)。在本章中, 首先在第 2.1 节中概述了本文所提方法的整体流程, 然后分别在第 2.2、2.3、2.4 节中详细介绍了标精地图先验融合的方法、感知不确定性建模的方法和基于概率的拓扑关系推理的方法, 最后在第 2.5 节中介绍了训练损失函数的设计。

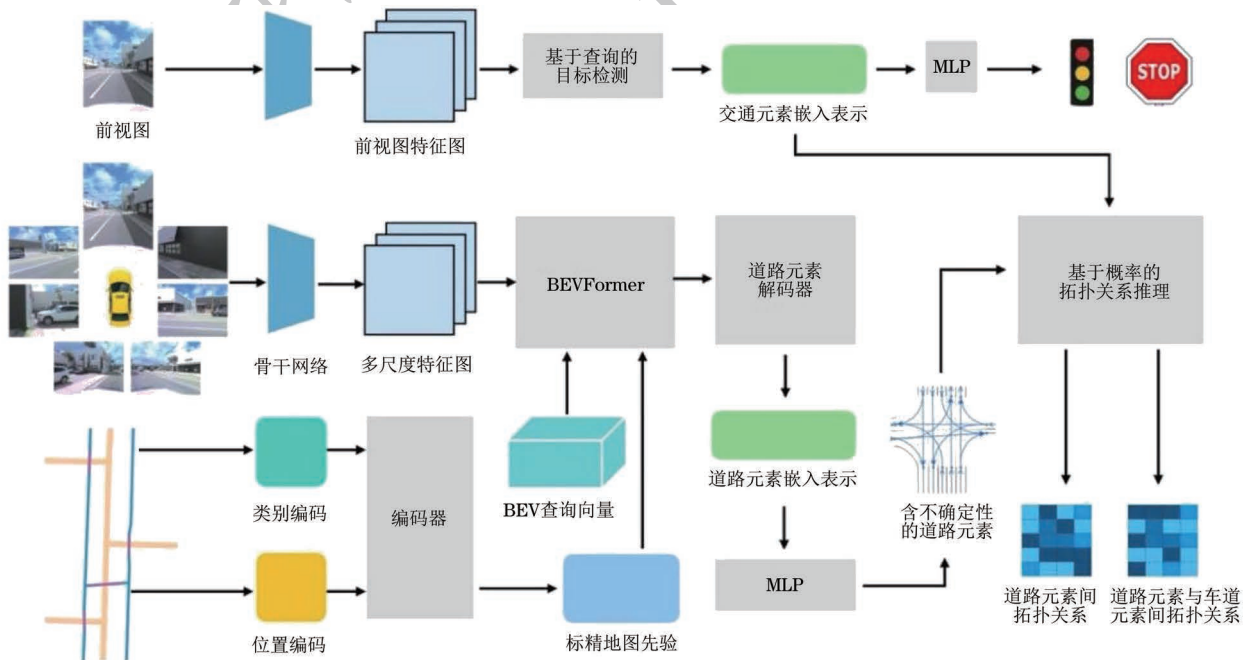


图 1 融合标精地图先验的地图不确定性感知及拓扑关系推理框图

Fig. 1 Block framework for map uncertainty perception and topological relationship reasoning that integrates standard map priors

### 2.1 方法概述

本文提出一种融合标精地图先验的地图感知不确定性建模及拓扑关系推理方法, 该方法能够端到端地感知道路元素和交通元素, 并推理道路元素和交通元素之间的拓扑关系。具体来说, 本文的方法主要包括以下几个步骤: 首先将自车环视图输入骨

干网络中, 再将骨干网络的输出输入特征金字塔网络(FPN)中, 得到不同尺度的特征图。然后, 将 FPN 的输出输入 BEVFormer<sup>[21]</sup> 中, 将环视图特征转换为统一的鸟瞰图(BEV)特征。接着, 使用 BEV 特征进行道路元素感知。本文使用基于查询的道路元素检测方法, 通过初始化查询向量, 使用交叉注意

力机制在 BEV 特征图中检测道路元素并更新查询向量。更新后的查询向量可由多个 MLP 检测头预测道路元素的位置和类别。在道路元素感知阶段,本文引入标精地图先验,提升了道路元素感知的性能,并且使用 Laplace 分布建模道路元素,实现了对道路元素的不确定性建模。交通元素检测和道路元素检测并行,但是由于该任务仅关注交通元素在前视图二维像素坐标系下的边界框,因此不需要使用 BEV 特征。在上述 FPN 抽取的特征图中,可以得到前视图对应的特征图。使用基于查询向量的目标检测方法,通过初始化查询向量,使用交叉注意力机制在前视图特征图中检测交通元素并更新查询向量,更新后的查询向量可由多个 MLP 检测头预测交通元素的边界框坐标和类别。得到道路元素和交通元素的检测结果后,本文利用道路元素和交通元素检测的更新后的查询向量进行拓扑关系推理。拓扑关系推理分为道路元素之间的拓扑关系推理和道路元素与交通元素之间的拓扑关系推理。本文利用两个 MLP 分别预测道路元素之间的拓扑关系和道路元素与交通元素之间的拓扑关系。为了在拓扑关系推理阶段利用引入的地图感知不确定性信息,本文提出了一种基于概率的拓扑关系推理方法。最终,本文使用多任务损失函数训练整个模型,包括道路元素检测损失、交通元素检测损失、道路元素之间拓扑关系推理损失和道路元素与交通元素之间拓扑关系推理损失。

## 2.2 标精地图先验融合

利用自行车周视图进行道路元素感知已经取得了一定的进展,但是在遮挡、强光照射、雷雨天气等极端场景下,模型无法获得有效输入,导致道路元素感知性能下降。为了提升道路元素感知的性能,本文提出了一种融合标精地图先验的方法。标精地图中包含了大量的道路结构信息和道路语义信息,可以作为道路元素感知的先验知识,且标精地图获取成本低、更新频率高,适合用于道路感知任务。

标精地图由一组折线组成,其中每条折线代表一个道路元素,如车道、人行横道、路沿等。由于每条折线的长度不固定,模型无法直接处理折线,因此将折线等分为一组固定长度  $L$  的点列,故所有标精地图元素可表示为  $\mathbf{P} = \{p_1, p_2, \dots, p_n\}$ , 其中  $p_i = \{(x_i, y_i)\}_{i=1}^L$ 。本文采用正余弦位置编码的方式将标精地图  $\mathbf{P}$  中的每个点列映射到一个固定维度的向量,表示为  $\mathbf{V} = \{v_1, v_2, \dots, v_n\}$ , 其中  $v_i \in \mathbb{R}^d$ 。为了区分不同类别的折线,为每条折线添加一个类别独热(one-hot)向量,表示为  $\mathbf{C} = \{c_1, c_2, \dots, c_n\}$ , 其

中  $c_i \in \mathbb{R}^{|\mathbf{C}|}$ ,  $|\mathbf{C}|$  为类别数量。将位置编码向量和类别 one-hot 向量拼接,得到标精地图特征  $\mathbf{M}$ , 表示为  $\mathbf{M} = \{m_1, m_2, \dots, m_n\}$ , 其中  $m_i \in \mathbb{R}^{d+|\mathbf{C}|}$ 。为了进一步获得标精地图的全局信息,采用标准 6 层 Transformer 编码器对标精地图特征进行提取,得到最终的标精地图特征  $\mathbf{M}_{\text{map}} \in \mathbb{R}^{n \times d}$ 。

原始的 BEVFormer<sup>[21]</sup> 采用基于查询向量的 BEV 特征构建方法。首先初始化三维查询向量  $\mathbf{Q}_{\text{bev}}$  表示三维 BEV 特征,  $\mathbf{Q}_{\text{bev}} \in \mathbb{R}^{p \times q \times h}$ , 其中  $p, q, h$  分别表示 BEV 感知范围的长、宽和高。然后,将查询向量  $\mathbf{Q}_{\text{bev}}$  送入时间自注意力层融合上一时刻和当前时刻的信息,得到更新后的查询向量送入空间交叉注意力层,融合环视图提取出的特征,得到最终的 BEV 特征。本文在 BEVFormer 的基础上,在空间交叉注意力层后增加了一个地图交叉注意力层,用于融合标精地图特征。具体来说,将标精地图特征  $\mathbf{M}_{\text{map}}$  送入地图交叉注意力层,将上一层输出的 BEV 特征作为查询向量,标精地图特征作为键向量和值向量,使 BEV 特征和标精地图进行交互,得到融合标精地图先验的 BEV 特征  $\mathbf{Q}_{\text{bev}}^M$ 。带有地图交叉注意力层的 BEVFormer 的结构如图 2 所示。

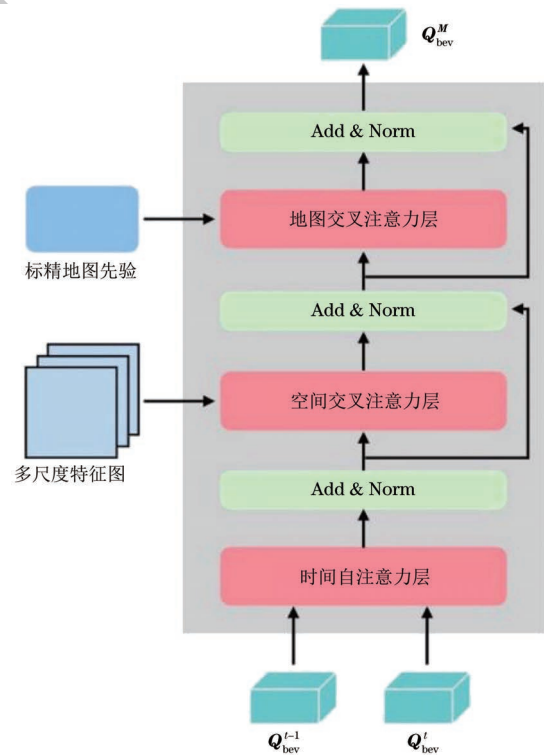


图 2 带地图交叉注意力层的 BEVFormer 结构图

Fig.2 BEVFormer structure diagram with map cross attention layer

## 2.3 感知不确定性建模

传统方法如 TopoNet<sup>[12]</sup>、TopoMLP<sup>[14]</sup> 等在地图感知阶段通过一个简单的 MLP 作为回归头,直

接预测道路元素的坐标。这种确定性的建模方式无法有效处理数据集中存在的极端场景数据,如在遮挡情况下仍然要求模型输出准确的检测结果。这类异常场景数据会使模型输出结果不稳定,从而影响模型的鲁棒性和泛化性。同时,对于下游任务来说,确定性建模无法反映模型预测结果可能存在的误差,导致下游任务的性能受到限制。为了解决这一问题,本文提出了一种基于 Laplace 分布的地图感知不确定性建模方法。相较于高斯分布,Laplace 分布是一种具有尖峰分布,能够更好地适应数据集中存在的极端场景数据,且本文用于计算道路元素预测结果与真值之间误差的损失函数为 L1 损失,与 Laplace 分布的尖峰分布相适应。具体来说,本文将每个道路元素视为固定长度  $L$  的点列,记作  $R = \{r_1, r_2, \dots, r_n\}$ ,其中每个道路元素  $r_j = \{(x_i, y_i)\}_{i=1}^L, x_i, y_i$  分别为每个点的横纵坐标。本文使用 Laplace 分布建模每一个道路元素,则道路元素  $r_j$  的概率密度函数为:

$$p(r_j | \mu, b) = \prod_{l=1}^L \frac{1}{2b^{(l)}} \exp\left(-\frac{|r_j^{(l)} - \mu^{(l)}|}{b^{(l)}}\right) \quad (1)$$

式中: $\mu^{(l)}$  表示道路元素  $r_j$  的第  $l$  个点的位置参数; $b^{(l)}$  表示道路元素  $r_j$  的第  $l$  个点的尺度参数。

本文模型使用基于查询向量的道路元素检测方法,通过初始化查询向量,使用交叉注意力机制在 BEV 特征图中检测道路元素并更新查询向量。更新后的查询向量通过两个 MLP 分别预测道路元素的位置参数和尺度参数。为了训练模型输出道路元素的位置参数和尺度参数,需要最大化道路元素的对数似然函数,即最小化负对数似然损失函数。具体来说,对于每个道路元素  $r_j$ ,定义其负对数似然损失函数为:

$$\mathcal{L}_{\text{NLL}}(r_j, \mu, b) = -\log_a p(r_j | \mu, b) = \sum_{l=1}^L \log_a(2b^{(l)}) + \frac{|r_j^{(l)} - \mu^{(l)}|}{b^{(l)}} \quad (2)$$

式中: $r_j^{(l)}$  表示道路元素  $r_j$  的第  $l$  个点的真实位置参数; $\mu^{(l)}$  表示道路元素  $r_j$  的第  $l$  个点的预测位置参数; $b^{(l)}$  表示道路元素  $r_j$  的第  $l$  个点的预测尺度参数。最终,将所有道路元素的负对数似然损失函数求和,得到整体的负对数似然损失函数。

#### 2.4 基于概率的拓扑关系推理

拓扑关系推理主要分为道路元素之间的拓扑关系推理和道路元素与交通元素之间的拓扑关系推理。在道路元素之间的拓扑关系推理中,传统方法考虑确定性的道路元素检测结果,直接使用 MLP 处理各个道路元素的确定坐标,以此推理道路元素

之间的拓扑关系。然而,这种确定性的建模方式无法反映在道路元素检测中可能存在的误差,导致拓扑关系推理的性能受到限制。为了解决这一问题,本文提出了一种基于概率的拓扑关系推理方法,在拓扑关系推理任务中引入道路元素不确定性。当道路元素受到其他车辆或树木遮挡时,其坐标位置感知结果可能不准确,模型会通过不确定性范围对这些坐标位置进行合理的推测与约束,避免因个别异常点或偏移点导致整体拓扑关系推理的误差过大。具体来说,本文记道路元素  $p$  和道路元素  $q$  之间的拓扑关系为  $A_{pq}, A_{pq} \in \{0, 1\}$ ,其中  $A_{pq} = 1$  表示道路元素  $p$  和道路元素  $q$  之间存在拓扑关系, $A_{pq} = 0$  表示道路元素  $p$  和道路元素  $q$  之间不存在拓扑关系。道路元素之间的拓扑关系的概率分布为:

$$P(A_{pq} | \{\mu_i^p, b_i^p\}_{i=1}^L, \{\mu_j^q, b_j^q\}_{j=1}^L) = C(A_{pq} | \{\mu_i^p, b_i^p\}_{i=1}^L, \{\mu_j^q, b_j^q\}_{j=1}^L) \quad (3)$$

式中: $\{\mu_i^p, b_i^p\}_{i=1}^L$  表示由  $L$  个点列组成的道路元素  $p$  的位置参数和尺度参数; $\{\mu_j^q, b_j^q\}_{j=1}^L$  表示由  $L$  个点列组成的道路元素  $q$  的位置参数和尺度参数; $C(\cdot)$  表示类别分布。则所有道路元素之间的拓扑关系的概率分布为:

$$P(A | \{\mu_i, b_i\}_{i=1}^L, \{\mu_j, b_j\}_{j=1}^L) = \prod_{p=1}^n \prod_{q=1}^n P(A_{pq} | \{\mu_i^p, b_i^p\}_{i=1}^L, \{\mu_j^q, b_j^q\}_{j=1}^L) \quad (4)$$

本文使用 MLP 预测道路元素之间拓扑关系的概率分布,将道路元素的位置参数和尺度参数编码后送入 MLP,得到道路元素之间的拓扑关系预测  $\hat{A}$ :

$$\hat{A} = \text{MLP}(\mathcal{E}(\mathcal{F}(\{\mu_i, b_i\}_{i=1}^L, \{\mu_j, b_j\}_{j=1}^L))) \quad (5)$$

式中: $\mathcal{F}(\cdot)$  将所有道路元素两两组合,形成道路元素对并拼接; $\mathcal{E}(\cdot)$  将道路元素对编码。则损失函数可定义为所有道路元素之间的拓扑关系二元交叉熵损失函数之和:

$$\mathcal{L}_{\text{RR}} = -\sum_{p=1}^n \sum_{q=1}^n A_{pq} \log_a \hat{A}_{pq} + (1 - A_{pq}) \log_a (1 - \hat{A}_{pq}) \quad (6)$$

与道路元素之间的拓扑关系推理不同,道路元素与交通元素之间的拓扑关系推理需要进一步考虑两者所在特征空间不一致的问题:道路元素的表示空间为三维世界空间,而交通元素的表示空间为二维像素空间。为了解决这一问题,本文在概率建模的基础上,引入特征对齐模块,通过融入三维世界空间和二维像素空间的映射关系促使模型自发学习道路元素与交通元素之间的空间对

齐关系。除此之外,道路元素与交通元素之间的拓扑关系推理与道路元素之间的拓扑关系推理类似。记道路元素  $p$  和交通元素  $q$  之间的拓扑关系为  $B_{pq}, B_{pq} \in \{0, 1\}$ 。道路元素与交通元素之间拓扑关系的概率分布为:

$$P(B | \{\mu_i, b_i\}_{i=1}^L, (x, y, h, w)) = \prod_{p=1}^n \prod_{q=1}^m P(B_{pq} | \{\mu_i^p, b_i^p\}_{i=1}^L, (x_q, y_q, h_q, w_q)) \quad (7)$$

式中:  $\{\mu_i^p, b_i^p\}_{i=1}^L$  表示由  $L$  个点列组成的道路元素  $p$  的位置参数和尺度参数;  $(x_q, y_q, h_q, w_q)$  表示交通元素  $q$  的边界框坐标。同样使用 MLP 预测道路元素与交通元素之间拓扑关系的概率分布,将道路元素的位置参数和尺度参数编码后送入 MLP,并融合空间对齐映射关系,得到道路元素与交通元素之间的拓扑关系预测  $\hat{B}$ :

$$Q = \text{MLP}(\dots) \quad (8)$$

$$\hat{B} = \text{MLP}(\mathcal{E}(\mathcal{F}(\{\mu_i, b_i\}_{i=1}^L, Q))) \quad (9)$$

式中:

$\mathcal{F}(\cdot)$  将所有道路元素与交通元素组合,形成道路元素与交通元素对并拼接;  $\mathcal{E}(\cdot)$  将道路元素与交通元素对编码。则损失函数可定义为所有道路元素与交通元素之间拓扑关系的二元交叉熵损失函数之和:

$$\mathcal{L}_{\text{RT}} = - \sum_{p=1}^n \sum_{q=1}^m B_{pq} \log_a \hat{B}_{pq} + (1 - B_{pq}) \log_a (1 - \hat{B}_{pq}) \quad (10)$$

## 2.5 训练损失函数

本文的训练损失函数由道路元素检测损失、交通元素检测损失、道路元素之间拓扑关系推理损失和道路元素与交通元素之间拓扑关系推理损失组成。道路元素检测损失  $\mathcal{L}_R$  包括道路元素位置参数和尺度参数的负对数似然损失、道路元素分类损失。交通元素检测损失  $\mathcal{L}_T$  包括交通元素边界框坐标的负对数似然损失、交通元素分类损失。道路元素之间拓扑关系推理损失  $\mathcal{L}_{\text{RR}}$  包括道路元素之间的拓扑关系二元交叉熵损失。道路元素与交通元素之间拓扑关系推理损失  $\mathcal{L}_{\text{RT}}$  包括道路元素与交通元素之间的拓扑关系二元交叉熵损失。整体的训练损失函数为:

$$\mathcal{L} = \mathcal{L}_R + \mathcal{L}_T + \mathcal{L}_{\text{RR}} + \mathcal{L}_{\text{RT}} \quad (11)$$

## 3 实验

### 3.1 数据集和评价指标

OpenLaneV2 是首个致力于解决自动驾驶场景

下拓扑关系推理任务的大规模数据集。OpenLaneV2 数据集包含了大量的车道线、道路元素、车道线-车道线关系、车道线-道路元素关系标注数据,以及对应的多视角图像和标精地图数据。本文实验使用 OpenLaneV2 数据集,其中包含了 22 477 个训练样本、4 806 个验证样本以及 4 816 个测试样本,每个样本中都包含了 7 张环视图像以及对应地点周围的标精地图。

模型评价指标使用 OpenLaneV2 UniScore (OLUS) 指标,该指标综合考虑了道路元素检测、交通元素检测、道路元素间拓扑关系推理、道路元素与交通元素间拓扑关系推理 4 个任务的性能。在道路元素中,由于车道相关元素使用折线表示、区域相关元素使用多边形表示,因此将道路元素检测任务分为车道线检测和区域检测两个子任务,故评价指标具体表示为:

$$I_{\text{OLUS}} = \frac{1}{5} (I_{\text{DET}_l} + I_{\text{DET}_{\text{te}}} + I_{\text{DET}_a} + f(I_{\text{TOP}_{\text{ll}}}) + f(I_{\text{TOP}_{\text{lt}}})) \quad (12)$$

式中:  $I_{\text{DET}_l}$ 、 $I_{\text{DET}_{\text{te}}}$ 、 $I_{\text{DET}_a}$ 、 $I_{\text{TOP}_{\text{ll}}}$ 、 $I_{\text{TOP}_{\text{lt}}}$  分别表示车道线检测、交通元素检测、区域检测、道路元素间拓扑关系推理和道路元素与交通元素间拓扑关系推理任务性能;  $f$  表示平方根函数。具体来说,道路元素检测指标使用弗雷切特距离计算检测结果与真值之间的相似性,并以此和阈值计算均值平均精度 (mAP); 交通元素检测使用交并比计算检测结果与真值之间的相似性,并以此和阈值计算 mAP; 拓扑关系推理则是采用图领域中的邻边预测准确率作为评价指标。

### 3.2 基线方法

本文选取的基线方法分别是 STSU<sup>[8]</sup>、VectorMapNet<sup>[10]</sup>、MapTR<sup>[11]</sup>、TopoNet<sup>[12]</sup>、TopoMLP<sup>[14]</sup>、LaneSegNet<sup>[15]</sup>。其中, STSU、VectorMapNet、MapTR 仅检测道路元素以及道路元素之间的拓扑关系, TopoNet、TopoMLP、LaneSegNet 则在检测道路元素的基础上对交通元素进行检测,并推理道路元素与交通元素之间的拓扑关系。此外, TopoNet 和 TopoMLP 原文中使用的数据集是 OpenLaneV2 的初代版本,仅要求模型检测车道中心线,而不考虑道路分隔线及其虚实分类,而本文使用的数据集是 OpenLaneV2 的 Map Element Bucket 版本,该版本综合考虑了车道线、道路分隔线、交通标志、交通灯等多种道路元素,因此对模型的检测和推理能力提出了更高的要求。

### 3.3 实验设置

数据集中的图片大小为  $1\,550 \times 2\,048$  像素,为了加速训练,本文将图片缩小至  $512 \times 714$  像素。本文使用 ResNet-50<sup>[22]</sup> 作为骨干网络,配合使用 FPN<sup>[23]</sup> 抽取图像特征。训练时,本文使用 Adam 优化器,初始学习率为 0.001,学习率变化采用余弦退火策略,共训练 24 个 Epoch。在训练过程中,本文使用了数据增强技术,包括随机裁剪、随机翻转、随机旋转等。本文使用基于查询向量的检测方法,其中道路元素检测的查询向量数量为 200 个,交通元素检测的查询向量数量为 100 个,区域检测的查询向量数量为 300 个。在 6 块 NVIDIA A100 GPU 的环境下训练模型,训练过程中 Batch Size 为 1。

### 3.4 实验结果

本文在 OpenLaneV2 数据集上进行了实验,对比了多种基线方法的性能。实验结果如表 1 所示,其中,“—”表示无法提供该项指标,“\*”表示使用 YOLOv8 进行交通元素检测。从表中可以看出,本文提出的方法在道路元素检测、交通元素检测、道路元素间拓扑关系推理、道路元素与交通元素间拓扑关系推理等多个任务上均取得了最好的性能。在道路元素检测任务上,本文所提方法在车道线检测和

区域检测任务上分别取得了 0.379 和 0.267 的 mAP,相比于最好的基线方法 LaneSegNet 分别提高了 0.095 和 0.03。这主要得益于标精地图先验以及道路元素的不确定性建模。标精地图先验可以帮助模型更好地理解道路元素的位置和形状,且当道路元素受到遮挡时,模型能够根据标精地图先验推断道路元素的位置,以此有效提高检测性能,而道路元素的不确定性建模则可以帮助模型学习到更加鲁棒的特征表示,从而加快模型收敛速度,提升模型检测的准确性和鲁棒性。在拓扑关系推理任务上,本文所提方法在道路元素间拓扑关系推理和道路元素与交通元素间拓扑关系推理任务上分别取得了 0.322 和 0.239 的 mAP,相比于最好的基线方法 LaneSegNet 提高了 0.077 和 0.026。这主要得益于上游检测任务准确性的提升,本文通过引入标精地图先验和道路元素的不确定性建模,有效提高了道路元素检测的准确性,从而为下游任务提供更加准确的特征表示,提高了拓扑关系推理的准确性。同时,在对道路元素进行不确定性建模后,拓扑关系推理模块可以利用道路元素的不确定性信息,相较于确定性建模方法,拓扑关系推理模型可以更好地理解道路元素之间的关系,从而提高拓扑关系推理的准确性。

表 1 实验结果对比

Table 1 Comparison of experimental results

方法	骨干网络	$I_{DET_l}$	$I_{DET_{tc}}$	$I_{DET_a}$	$I_{TOP_{ll}}$	$I_{TOP_{lt}}$	$I_{OLUS}$
STSU	ResNet-50	0.116	—	—	0.018	—	—
VectorMapNet	ResNet-50	0.132	—	—	0.028	—	—
MapTR	ResNet-50	0.103	—	—	0.021	—	—
TopoNet	ResNet-50	0.196	0.452	0.181	0.122	0.218	0.329
TopoMLP	ResNet-50	0.253	0.464	0.203	0.163	0.197	0.354
LaneSegNet	ResNet-50	0.284	0.459	0.237	0.245	0.213	0.387
Ours	ResNet-50	0.379	0.486	0.267	0.322	0.239	0.438
Ours	SwinL	0.429	0.521	0.292	0.378	0.281	0.477
Ours*	SwinL	0.429	0.713	0.292	0.378	0.309	0.521

为了探究不同骨干网络对模型性能的影响,本文还在 SwinL<sup>[24]</sup> 上进行了实验。从表 1 中可以看出,本文所提方法在 SwinL 上取得了更强的性能。这主要得益于 SwinL 更强的特征表达能力;SwinL 在处理大规模数据集时,可以更好地学习到数据的分布特征,从而提高模型的泛化能力。此外,考虑到交通元素检测任务中大部分被检测物体均为小目标,本文还使用 YOLOv8<sup>[25]</sup> 进行交通元素检测。从表 1 中可以看出,使用 YOLOv8 进行交通元素检测可以提高交通元素检测的性能,同时提高道路元素

和交通元素之间拓扑关系推理的性能,但该方法无法实现端到端的训练。

### 3.5 消融实验

为了探寻本文所提方法的有效性,进行了多组消融实验。本文在 LaneSegNet 的基础上逐步引入标精地图先验和道路元素的不确定性建模,并在此基础上引入基于概率的拓扑关系推理模块,分别进行了实验。实验结果如表 2 所示。从表 2 中可以看出,标精地图先验有效提升了道路元素检测的性能,道路元素的不确定性建模有效提升了道路元素检测

的准确性和拓扑关系推理的准确性,基于概率的拓扑关系推理模块进一步提升了拓扑关系推理的准确性。本文提出的 3 种方法均能够使模型学习到更加

鲁棒的特征表示,提高模型的泛化能力,从而提高模型的检测性能和拓扑关系推理性能,验证了本文所提方法的有效性。

表 2 消融实验结果

Table 2 Results of ablation experiment

方法	$I_{DET_l}$	$I_{DET_{te}}$	$I_{DET_a}$	$I_{TOP_{ll}}$	$I_{TOP_{lt}}$	$I_{OLUS}$
LaneSegNet	0.284	0.459	0.237	0.245	0.213	0.387
+标精地图先验	0.362	0.476	0.259	0.296	0.227	0.398
+不确定性建模	0.377	0.478	0.267	0.308	0.231	0.422
+概率拓扑关系推理	0.379	0.486	0.267	0.322	0.239	0.438

### 3.6 可视化分析

本文首先对融合标精地图先验的有效性进行了可视化分析,如图 3 所示。在 OpenLaneV2 数据集上,分别使用 LaneSegNet 和 LaneSegNet 融合标精地图先验的方法(SD Map)感知道路元素。从图中可以发现,通过引入标精地图先验,模型可以更好地理解道路元素的位置和形状,从而提高道路元素检测的准确性。同时,当道路元素受到遮挡时,模型可以根据标精地图先验推断道路元素的位置,以此提高检测性能。在此基础上,本文进一步对道路元素的不确定性建模进行了可视化分析,如图 4 所示。本文在融合标精地图先验的基础上,使用带有不确定性建模的 LaneSegNet 感知道路元素。从图中可以发现,当道路元素清晰可见时,模型检测出的道路

元素不确定性较低,而当道路元素受到遮挡或者道路元素边界模糊时,模型检测出的道路元素不确定性较高。这说明,本文所提方法可以有效地对道路元素的不确定性进行建模,从而提高道路元素检测的准确性和鲁棒性。同时,这种不确定性建模方法可以帮助模型学习到更加鲁棒的特征表示,从而加快模型收敛速度,提升模型检测的准确性和鲁棒性。如图 5 所示,本文对比了带有不确定性建模的 LaneSegNet 和不带不确定性建模的 LaneSegNet 的损失曲线。从图中可以看出,带有不确定性建模的 LaneSegNet 在训练初期损失下降速度和收敛速度更快,且在训练后期损失更小,表明带有不确定性建模的 LaneSegNet 可以更好地学习到数据的分布特征,提高模型的泛化能力。

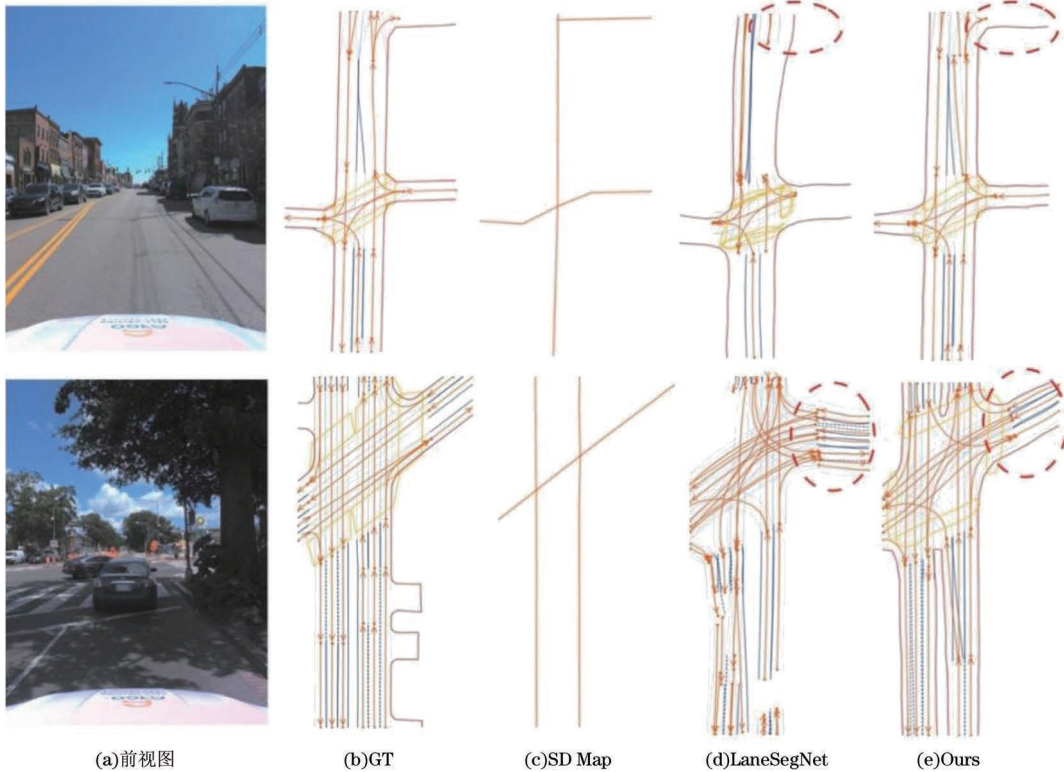


图 3 融合标精地图先验的检测结果可视化

Fig. 3 Visualization of detection results fused with standard map priors



图 4 道路元素感知不确定性建模可视化

Fig. 4 Visualization of uncertainty modeling for road element perception

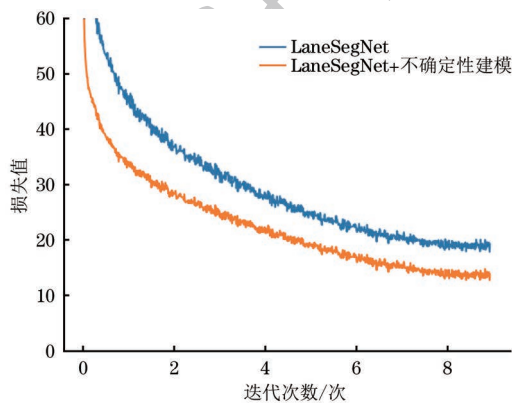


图 5 添加不确定性建模前后损失值下降对比

Fig. 5 Comparison of the decrease in loss value before and after adding uncertainty modeling

### 3.7 讨论

本文在感知模型中引入了标精地图先验,并对感知结果进行不确定性建模,所提方法可被扩展至任意架构中。本文通过交叉注意力机制在构建 BEV 特征时引入标精地图,为模型提供了道路先验信息,有效提升了模型的道路元素检测精度。该模块可直接用于任意地图感知方案,仅需在特征抽取后使用交叉注意力模块融合标精地图特征,实现即插即用。本文通过 Laplace 分布建模道路元素,利用 Laplace 分布的位置参数表示坐标值,尺度参数表示该坐标可能距真实坐标的距离的范围,即不确定性。这种建模方式可扩展至任意架构的地图感知方案中,仅需将损失函数替换为 2.3 节所提的负对

数似然损失函数。

所提方法除了能够适应不同的模型架构外,还能够适应不同数据集。OpenLaneV2 在 Argoverse 2 和 nuScenes 基础上采集了部分场景构建而成,包含了欧洲、美洲、亚洲 8 个城市的数据,数据多样性高。本文所提方法在 OpenLaneV2 数据集上相较于基线方法取得了性能提升,验证了其对于不同数据集和不同城市分布的数据具有较好的鲁棒性和泛化性。

## 4 结束语

本文提出了一种融合标精地图先验的地图感知不确定性建模及拓扑关系推理方法。本文首先提出了一种融合标精地图的先验感知方法,通过引入标精地图先验,模型可以更好地理解道路元素的位置和形状,从而提高道路元素检测的准确性。同时,当道路元素受到遮挡时,模型可以根据标精地图先验推断道路元素的位置,以此提高检测性能。在此基础上,本文提出了一种地图感知不确定性的建模方法,通过引入地图感知不确定性模型,模型能够学习到更为鲁棒的特征,从而提高道路元素检测的鲁棒性。最后,本文提出了一种基于概率的拓扑关系推理的方法,充分利用地图感知不确定性模型的输出,提升拓扑关系推理的准确性。本文在 OpenLaneV2 数据集上进行了大量实验,实验结果表明,本文提出的方法在道路元素检测和拓扑关系推理任务上均取得了较好的性能,

验证了本文方法的有效性。

本文所提方法目前并未引入时间序列信息,仅对单帧传感器输入进行环境感知,下一步将继续探索如何引入时序信息提升感知结果的帧间连续性。

### 参考文献

- [1] 赵祥模国家重点研发计划(2021YFB2501200)团队. 自动驾驶测试与评价技术研究进展[J]. 交通运输工程学报, 2023, 23(6): 10-77.  
ZHAO Xiangmo's National Key R&D Program Team (2021YFB2501200). Research progress on autonomous driving testing and evaluation technologies[J]. Journal of Traffic and Transportation Engineering, 2023, 23(6): 10-77. (in Chinese)
- [2] 章军辉, 陈大鹏, 李庆. 自动驾驶技术研究现状及发展趋势[J]. 科学技术与工程, 2020, 20(9): 3394-3403.  
ZHANG J H, CHEN D P, LI Q. Research status and development trend of technologies for autonomous vehicles[J]. Science Technology and Engineering, 2020, 20(9): 3394-3403. (in Chinese)
- [3] 吴一全, 刘莉. 基于视觉的车道线检测方法研究进展[J]. 仪器仪表学报, 2019, 40(12): 92-109.  
WU Y Q, LIU L. Research and development of the vision-based lane detection methods[J]. Chinese Journal of Scientific Instrument, 2019, 40(12): 92-109. (in Chinese)
- [4] 周苏, 吴迪, 金杰. 基于卷积神经网络的车道线实例分割算法[J]. 激光与光电子学进展, 2021, 58(8): 0815007.  
ZHOU S, WU D, JIN J. Lane instance segmentation algorithm based on convolutional neural network[J]. Laser & Optoelectronics Progress, 2021, 58(8): 0815007. (in Chinese)
- [5] 张凯航, 冀杰, 蒋路, 等. 基于 SegNet 的非结构道路可行驶区域语义分割[J]. 重庆大学学报, 2020, 43(3): 79-87.  
ZHANG K H, JI J, JIANG L, et al. The semantic segmentation of driving regions on unstructured road based on segnet architecture[J]. Journal of Chongqing University, 2020, 43(3): 79-87. (in Chinese)
- [6] 刘占文, 范颂华, 齐明远, 等. 基于时序融合的自动驾驶多任务感知算法[J]. 交通运输工程学报, 2021, 21(4): 223-234.  
LIU Z W, FAN S H, QI M Y, et al. Multi-task perception algorithm of autonomous driving based on temporal fusion[J]. Journal of Traffic and Transportation Engineering, 2021, 21(4): 223-234. (in Chinese)
- [7] 刘经南, 詹骄, 郭迟, 等. 智能高精地图数据逻辑结构与关键技术[J]. 测绘学报, 2019, 48(8): 939-953.  
LIU J N, ZHAN J, GUO C, et al. Data logic structure and key technologies on intelligent high-precision map[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(8): 939-953. (in Chinese)
- [8] CAN Y B, LINIGER A, PAUDEL D P, et al. Structured bird's-eye-view traffic scene understanding from onboard images[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, Canada: IEEE Press, 2021: 15641-15650.
- [9] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with Transformers[C]//Proceedings of ECCV 2020. Berlin, Germany: Springer, 2020: 213-229.
- [10] LIU Y C, WANG Y, WANG Y L, et al. VectorMapNet: end-to-end vectorized HD map learning[C]//Proceedings of International Conference on Machine Learning. New York, USA: ACM Press, 2023: 22352-22369.
- [11] LIAO B, CHEN S H, WANG X L, et al. MapTR: structured modeling and learning for online vectorized HD map construction[C]//Proceedings of the 11th International Conference on Learning Representations. [S. l.]: ICLR, 2022: 1-16.
- [12] LI T H, CHEN L, WANG H, et al. Graph-based topology reasoning for driving scenes[EB/OL]. (2023-04-11) [2024-08-05]. <https://arxiv.org/pdf/2304.05277>.
- [13] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of NeurIPS 2017. Berlin, Germany: Springer, 2017: 6000-6010.
- [14] WU D S, CHEN J Y, JIA F, et al. TopoMLP: a simple yet strong pipeline for driving topology reasoning [C] // Proceedings of the 12th International Conference on Learning Representations. [S. l.]: ICRL, 2023: 1-12.
- [15] LI T H, JIANG P C, WANG B, et al. LaneSegNet: map learning with lane segment perception for autonomous driving [C] // Proceedings of the 12th International Conference on Learning Representations. [S. l.]: ICRL, 2023: 1-10.
- [16] 崩旋, 丁飞, 张登银. 多分支融合注意力机制的车道线检测模型[J]. 电子测量与仪器学报, 2023, 37(4): 35-43.  
KUAI X, DING F, ZHANG D Y. Lane detection model of multi-branch fusion attention mechanism [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(4): 35-43. (in Chinese)
- [17] 赵旋, 黄崇栋, 刘云涛, 等. 复杂交通环境下基于多任务学习的道路感知方法[J]. 无线电工程, 2023, 53(11): 2473-2479.  
ZHAO X, HUANG C D, LIU Y T, et al. Multi-task learning-based road perception method in complex traffic environment[J]. Radio Engineering, 2023, 53(11): 2473-2479. (in Chinese)
- [18] 伍邦谷, 张苏林, 石红, 等. 基于多分支结构的不确定性局部通道注意力机制[J]. 电子学报, 2022, 50(2): 374-382.  
WU B G, ZHANG S L, SHI H, et al. Multi-branch structure based local channel attention with uncertainty[J]. Acta Electronica Sinica, 2022, 50(2): 374-382. (in Chinese)
- [19] 高一, 王忠立, 王颖博. 基于不确定性建模的 3D 场景感知方法[J]. 机器人, 2024, 46(4): 450-464.  
GAO Y, WANG Z L, WANG Y B. A 3D scene perception method based on uncertainty modeling [J]. Robot, 2024, 46(4): 450-464. (in Chinese)
- [20] 张舒宇, 吴刘仓, 詹金龙. 基于 Laplace 分布下混合联合位置与尺度模型的参数估计[J]. 应用概率统计, 2017, 33(5): 487-496.  
ZHANG S Y, WU L C, ZHAN J L. Parameters estimation for mixture of joint location and scale models based on the Laplace distribution [J]. Chinese Journal of Applied Probability and Statistics, 2017, 33(5): 487-496. (in Chinese)
- [21] LI Z Q, WANG W H, LI H Y, et al. BEVFormer: learning bird's-eye-view representation from multi-camera images via spatiotemporal transformers[C]//Proceedings of ECCV 2022. Berlin, Germany: Springer, 2022: 1-18.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE Press, 2016: 770-778.
- [23] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE Press, 2017: 936-944.
- [24] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, Canada: IEEE Press, 2021: 9992-10002.
- [25] JOCHER G, CHAURASIA A. Ultralytics YOLO[EB/OL]. [2024-08-05]. <https://github.com/ultralytics/ultralytics>.