

基于图神经网络的自动驾驶场景理解

张志鹏

华中科技大学 湖北武汉 430070

摘要: 本文旨在探讨图神经网络 (Graph Neural Network, GNN) 在自动驾驶场景理解中的应用。随着自动驾驶技术的快速发展,对复杂交通场景的准确理解成为实现高级自动驾驶功能的关键。图神经网络作为一种处理图结构数据的强大工具,能够捕捉场景元素之间的复杂关系和相互作用,为自动驾驶场景理解提供了新的视角和方法。本文首先介绍了图神经网络的基本概念、架构及应用案例,然后详细阐述了如何将自动驾驶场景建模为图结构,并利用图神经网络进行特征提取、信息融合和关系推理,最终实现场景的理解。通过训练优化和性能提升策略的讨论,本文为基于图神经网络的自动驾驶场景理解提供了全面的框架和实用的方法。

关键词: 图神经网络; 自动驾驶; 场景理解; 特征提取; 关系推理

引言

自动驾驶技术作为智能交通系统的重要组成部分,近年来取得了显著进展。但要实现完全自动驾驶,车辆必须能够准确理解复杂的交通场景,包括识别道路、车辆、行人、交通信号等元素,并预测它们的行为和交互。传统的计算机视觉和深度学习方法在处理这类复杂场景时存在局限性,难以捕捉场景元素之间的复杂关系和动态变化。图神经网络作为一种新兴的深度学习方法,以其处理图结构数据的独特优势,在自动驾驶场景理解中展现出巨大潜力。本文将深入探讨图神经网络在自动驾驶场景理解中的应用,为相关研究和实践提供理论支持和技术参考。

一、图神经网络,基础梳理

(一) 图神经网络的基本概念与原理

图神经网络是一种专门用于处理图结构数据的深度学习模型。图结构数据由节点(或顶点)和边组成,节点代表实体或对象,边表示节点之间的关系或交互。与传统的网格结构数据(如图像)和序列结构数据(如文本)不同,图结构数据具有不规则性和复杂性,难以用传统的深度学习方法直接处理。图神经网络通过引入邻接矩阵或图卷积操作,能够有效地捕捉图结构数据中的节点特征和节点之间的关系,实现图数据的表示学习和推理。其基本原理是通过聚合邻居节点的信息来更新当前节点的信息,从而逐层传递信息,最终得到整个图的表示。

(二) 图神经网络的架构与组成要素

图神经网络的架构通常包括输入层、图卷积层(或图嵌入层)、池化层(可选)、全连接层和输出层。其中,

图卷积层是图神经网络的核心组件,负责提取图结构数据中的特征。图卷积操作可以通过邻接矩阵和节点特征矩阵的乘积来实现,也可以通过更复杂的注意力机制或消息传递机制来进行。池化层用于减少图的规模,提取图的全局特征。全连接层则用于将图特征映射到具体的任务空间,如分类、回归或生成等。图神经网络的组成要素还包括损失函数、优化算法和正则化技术等,它们共同决定了图神经网络的性能和效果。

(三) 图神经网络在领域内的应用案例

图神经网络在多个领域内展现出了广泛的应用前景,如社交网络分析、推荐系统、生物信息学、化学信息学以及交通领域等。在社交网络中,图神经网络可以用于用户行为预测、社区发现和影响力最大化等任务。在推荐系统中,图神经网络可以利用用户-物品交互图来捕捉用户的兴趣和偏好。在生物信息学中,图神经网络可以用于蛋白质结构预测、基因表达分析和药物发现等。在化学信息学中,图神经网络可以用于分子性质预测、反应路径规划和药物合成设计等。在交通领域,图神经网络可以应用于交通流量预测、路况监测和智能交通信号控制等。这些应用案例充分展示了图神经网络在处理图结构数据方面的强大能力和广泛应用潜力。

二、场景建模,图化表示

(一) 自动驾驶场景的元素识别与分类

自动驾驶场景理解的首要任务是识别并分类场景中的元素。这些元素包括但不限于道路、车辆、行人、交通信号、标志标线等。元素识别可以通过计算机视觉技术实现,如使用卷积神经网络(CNN)对图像进行特征

提取和分类。然而，单纯的元素识别并不足以理解场景，还需要考虑元素之间的关系和相互作用。因此，在识别元素的基础上，还需要对元素进行分类，如将车辆分为轿车、卡车、公交车等，将行人分为行人、骑行者等，以便后续的图化表示和关系推理。

（二）场景元素的图结构表示方法

将自动驾驶场景建模为图结构是图神经网络应用的基础。场景元素可以作为图的节点，元素之间的关系或交互可以作为图的边。具体来说，可以将道路作为图的骨架，将车辆、行人等元素作为节点连接在道路上，根据元素之间的相对位置和运动状态确定边的存在和权重。此外，还可以引入属性图的概念，为节点和边添加属性信息，如节点的类型、速度、方向等，边的距离、角度等。由此，自动驾驶场景就被表示为一个包含丰富信息的属性图，为后续的图神经网络处理提供了基础。

（三）场景图模型的构建与优化策略

构建场景图模型需要考虑多个因素，如节点的选择、边的定义、属性的设置等。为了准确反映场景的真实情况，需要选择合适的节点和边来表示场景元素及其关系。同时，为了减少图的复杂性和计算量，需要对图进行简化或裁剪，如去除冗余节点和边、合并相似节点等。此外，还可以利用先验知识或数据驱动的方法来优化场景图模型，如通过聚类算法对节点进行分组、通过图嵌入技术提取图的低维表示等。这些优化策略可以提高场景图模型的准确性和效率，为后续的图神经网络处理提供更好的输入。

三、特征提取，信息融合

（一）图神经网络中的节点特征提取

节点特征是图神经网络处理的重要输入。在自动驾驶场景中，节点特征可以包括元素的类型、位置、速度、方向等基本信息，还可以包括元素的历史轨迹、行为模式等高级信息。为了提取节点特征，可以使用卷积神经网络、循环神经网络或变换器等深度学习模型对原始数据进行处理。同时，还可以利用图神经网络中的邻接矩阵或注意力机制来聚合邻居节点的信息，从而丰富节点特征的表示。通过节点特征提取，可以为后续的关系推理和场景理解提供有力的支持。

（二）图神经网络中的边特征表示方法

边特征在图神经网络中同样重要，它反映了节点之间的关系或交互。在自动驾驶场景中，边特征可以包括节点之间的距离、角度、相对速度等物理量，还可以包括节点之间的历史交互记录、行为相似性等高级特征。

为了表示边特征，可以引入边嵌入的概念，将边映射到低维向量空间中。同时，还可以利用图神经网络中的消息传递机制或注意力机制来传递边特征信息，从而增强图神经网络的表示能力。通过边特征表示，可以更准确地捕捉场景元素之间的复杂关系和动态变化。

（三）多源信息融合的图神经网络策略

自动驾驶场景理解需要融合多源信息，如图像、雷达、激光雷达（LiDAR）等传感器数据，以及地图、交通规则等先验知识。为了融合这些信息，可以构建多模态的图神经网络模型，将不同来源的信息作为图的节点或边特征输入到模型中。同时，还可以利用注意力机制或门控机制来动态调整不同信息源的重要性，从而实现信息的有效融合。通过多源信息融合，可以提高自动驾驶场景理解的准确性和鲁棒性，为后续的决策和控制提供更好的支持。

四、关系推理，场景理解

（一）图神经网络中的关系推理机制

关系推理是图神经网络的核心功能之一。在自动驾驶场景中，关系推理可以用于推断元素之间的相互作用和行为模式。例如，可以通过图神经网络预测车辆之间的跟驰行为、超车行为等，或者预测行人的过街意图、行走路径等。为了实现关系推理，图神经网络需要学习节点和边之间的复杂关系，并通过逐层传递信息来更新节点的表示。同时，还可以利用图神经网络中的注意力机制或图卷积操作来捕捉节点之间的长距离依赖关系，从而提高关系推理的准确性。

（二）场景元素间的相互作用分析

场景元素间的相互作用是自动驾驶场景理解的关键。通过分析元素之间的相互作用，可以揭示场景的动态变化和潜在风险。例如，可以分析车辆之间的相对速度和距离来判断是否存在追尾风险，或者分析行人与车辆之间的相对位置和运动轨迹来判断行人是否可能穿越马路。为了进行相互作用分析，可以利用图神经网络中的关系推理机制来推断元素之间的行为模式和交互关系。同时，还可以结合先验知识和实时数据来动态调整分析策略，从而提高相互作用分析的准确性和实时性。

（三）基于图神经网络的情景理解模型

基于图神经网络的情景理解模型可以将上述的特征提取、信息融合和关系推理等步骤集成在一起，形成一个完整的框架。从实际来看，可以先通过卷积神经网络或变换器等模型提取图像、雷达等传感器数据中的特征，然后将这些特征作为图的节点或边输入到图神经网络中。

在图神经网络中,可以利用图卷积操作、注意力机制或消息传递机制来传递和处理信息,从而实现场景元素的关系推理和相互作用分析。最后,可以通过全连接层或输出层将图神经网络的输出映射到具体的任务空间,如场景分类、行为预测或风险评估等。通过构建基于图神经网络的场景理解模型,可以实现对自动驾驶场景的准确理解和智能决策。

五、训练优化,性能提升

(一) 图神经网络的训练数据准备

图神经网络的训练需要大量的标注数据作为支撑。在自动驾驶场景中,标注数据可以包括图像、雷达等传感器数据的标注信息,以及场景元素的行为标签、交互关系等。为了获取高质量的标注数据,可以采用人工标注、半自动标注或自动标注等方法。同时,还需要对标注数据进行预处理和增强,如数据清洗、数据增强、数据平衡等,以提高图神经网络的泛化能力和鲁棒性。通过准备充足的训练数据,可以为图神经网络的训练提供有力的支持。

(二) 图神经网络的损失函数设计

损失函数是图神经网络训练过程中的关键组成部分,它衡量了模型预测结果与真实标签之间的差异,并指导模型参数的更新。在自动驾驶场景理解中,损失函数的设计需要根据具体任务来定制。

对于场景分类任务,可以使用交叉熵损失函数。交叉熵损失函数是分类问题中常用的损失函数,它能够有效地衡量模型预测的概率分布与真实标签之间的差异。通过最小化交叉熵损失,可以使模型更加准确地预测场景的类别。

对于行为预测或风险评估等回归任务,可以使用均方误差损失函数或绝对误差损失函数。均方误差损失函数衡量了模型预测值与真实值之间的平方差异,对异常值比较敏感;而绝对误差损失函数则衡量了模型预测值与真实值之间的绝对差异,对异常值相对鲁棒。根据任务的具体需求,可以选择合适的损失函数来优化模型。

此外,在自动驾驶场景理解中,还可能存在多任务学习的情况,即同时预测场景的类别、元素的行为以及交互关系等。在这种情况下,可以设计多任务损失函数,将不同任务的损失进行加权求和,以综合考虑各个任务的表现。通过合理设计损失函数,可以引导图神经网络在训练过程中更加关注关键信息,提高模型的性能和准确性。

(三) 模型优化与性能提升的策略探讨

除了设计合适的损失函数外,还可以采用其他策略

来优化图神经网络模型,提升其性能。

1.模型架构优化:可以根据具体任务的需求,调整图神经网络的层数、节点特征维度、边特征表示方法等,以找到最优的模型架构。同时,还可以尝试引入残差连接、跳跃连接等结构,以缓解深层网络中的梯度消失或梯度爆炸问题。

2.正则化技术:为了防止过拟合,可以采用正则化技术来约束模型的参数。常见的正则化技术包括L1正则化、L2正则化以及Dropout等。通过正则化,可以使模型更加泛化,提高其在未见数据上的表现。

3.数据增强:数据增强是提高模型性能的有效手段之一。在自动驾驶场景中,可以通过对图像、雷达等传感器数据进行旋转、缩放、裁剪等操作来生成更多的训练样本。数据增强可以增加数据的多样性,使模型更加鲁棒。

4.迁移学习:迁移学习是一种将在一个任务上学到的知识迁移到另一个任务上的方法。在自动驾驶场景理解中,可以利用在大规模数据集上预训练的模型作为初始模型,然后在特定场景的数据上进行微调。通过迁移学习,可以加快模型的收敛速度,提高其性能。

5.集成学习:集成学习是通过组合多个模型来提高性能的方法。在自动驾驶场景理解中,可以训练多个图神经网络模型,并将它们的预测结果进行加权平均或投票表决。集成学习可以降低单个模型的误差,提高整体的准确性。

结束语

综上所述,通过设计合适的损失函数、优化模型架构、采用正则化技术、数据增强、迁移学习以及集成学习等策略,可以有效地提升图神经网络在自动驾驶场景理解中的性能。随着自动驾驶技术的不断发展和图神经网络研究的深入,相信基于图神经网络的自动驾驶场景理解将取得更加显著的成果。

参考文献

- [1]罗鹭.基于图神经网络的场景理解算法研究[D].电子科技大学,2020.
- [2]凌骏,赵前,张帆,等.基于异质图神经网络的知识增强文本理解研究[J].计算机与数字工程,2024,52(4):1125-1130.
- [3]孙文洁,李宗民,孙浩森.基于图神经网络的多智能体强化学习值函数分解方法[J].计算机工程,2024,50(5):62-70.DOI:10.19678/j.issn.1000-3428.0067919.