

doi: 10.19562/j.chinasae.qcgc.2024.11.001

面向自动驾驶的大模型对齐技术:综述*

唐小林¹, 甘露¹, 李国法¹, 李志强², 褚文博^{3,4,5}

(1. 重庆大学机械与运载工程学院, 重庆 400044; 2. 清华大学车辆与运载学院, 北京 100084;
3. 国汽(北京)智能网联汽车研究院有限公司, 北京 100176; 4. 重庆理工大学机械检测技术与装备教育部工程研究中心, 重庆 400054;
5. 西部科学城智能网联汽车创新中心(重庆)有限公司, 重庆 401329)

[摘要] 随着Transformer注意力机制的出现,以GPT为代表的通用基础大模型实现了智能的“涌现”,给自动驾驶迈向更高级别发展带来了曙光。受限于传统从头预训练方式需要大规模、高质量、多样性自动驾驶数据和高昂训练成本的困扰,“大模型+对齐技术”范式衍生。对齐技术作为通用基础大模型与自动驾驶之间的纽带,通过微调或提示工程等定制化方式,可高效、专业地解决自动驾驶领域内的工程性问题。对齐技术已是大模型在垂直领域发展的研究热点,但缺乏系统研究成果。基于此,本文首先对自动驾驶发展与大模型技术进行概述,从而衍生出对齐技术。然后,分别从微调和提示工程两个角度进行综述,系统化梳理并剖析各分类技术的结构或性能特点,同时给出实际的应用案例。最后,基于现有研究提出了对齐技术的研究挑战与发展趋势,为促进自动驾驶迈向更高级别发展提供参考。

关键词: 自动驾驶;大模型;综述;对齐技术;微调;提示工程

Large Model Alignment Technology for Autonomous Driving: A Review

Tang Xiaolin¹, Gan Lu¹, Li Guofa¹, Li Keqiang² & Chu Wenbo^{3,4,5}

1. College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing 400044;
2. School of Vehicle and Mobility, Tsinghua University, Beijing 100084;
3. China Intelligent and Connected Vehicles (Beijing) Research Institute Co., Ltd., Beijing 100176;
4. Engineering Research Center of Mechanical Testing Tech. and Equip. Ministry of Education,
Chongqing University of Technology, Chongqing 400054;
5. Western China Science City Innovation Center of Intelligent and Connected Vehicles (Chongqing) Co., Ltd., Chongqing 401329

[Abstract] With the emergence of the Transformer attention mechanism, general-purpose large models represented by GPT have achieved the "emergence" of intelligence, bringing a dawn to the advancement towards higher levels of autonomous driving. Limited by the traditional from-scratch pre-training approach, which requires large-scale, high-quality, diverse autonomous driving data and incurs high training cost, the "large model + alignment technology" paradigm has been derived. As a bridge between general-purpose large models and autonomous driving, alignment technology, through customization methods such as fine-tuning or prompt engineering, achieves efficient and professional solutions to engineering problems within the field of autonomous driving. Alignment technology has become a hot research topic in the development of large models in vertical fields, but it lacks systematic research results. Based on this, this article firstly provides an overview of the development of autonomous driving and large model technology, thereby deriving alignment technology. Then, it reviews from the perspectives of fine-tuning and prompt engineering, systematically reviewing and analyzing the structure or performance characteristics of each classification technology, while providing actual application cases. Finally, based on existing research, the research

* 国家重点研发计划(2022YFB2503205)、国家自然科学基金(52372377, 52272421, 52222215, 52072051)、重庆市自然科学基金(CSTB2023NSCJOX0003, CSTB2023NSCQ-MSX0985)和智能绿色车辆与交通全国重点实验室开放基金课题(KFZ2409)资助。

原稿收到日期为2024年01月23日,修改稿收到日期为2024年03月26日。

通信作者:褚文博,博士,研究员, E-mail: chuwenbo@wicv.cn。

challenges and development trends of alignment technology are proposed, offering references for promoting the advancement towards higher level of autonomous driving development.

Keywords: autonomous driving; large model; review; alignment technology; fine-tuning; prompt engineering

前言

随着《“十四五”交通领域科技创新规划》、《自动驾驶汽车运输安全服务指南(试行)》(征求意见稿)、《上海市加快智能网联汽车创新发展实施方案》、《关于开展智能网联汽车准入和上路通行试点工作的通知》等一系列政策的提出,自动驾驶已然成为国家科技战略的重要一环。自动驾驶的发展极大地推动了经济发展和社会进步^[1],其核心目标是通过高度自动化的驾驶来提高道路安全性、提升交通效率,并最终实现减少交通拥堵和环境污染的愿景。

以 LLaMA^[2]、GPT^[3]、Gemini^[4]、文心一言^[5]、盘古^[6]等为代表的通用基础大模型取得了巨大的成功,在模拟类人智能方面显示出巨大的潜力^[7],能够解决社会生活中普遍的上游任务,在自然语言处理、计算机视觉、多模态交互等方面取得了巨大的成功,例如文本生成、翻译、总结、问答、对话生成、代码生成等普遍上游。然而对于一些需要定制输出、垂直领域等专业化、特定需求的下游任务却显得不尽人意,尤其体现在需要具备高度可靠性和鲁棒性的自动驾驶领域^[8],如驾驶员行为分析、3D场景识别、实时动态建图、多模态交互等。面对通用基础类大模型的“全而不精”,传统从头开始的预训练方式尽管可以解决这一问题,但受限于大规模、高质量、多样性自动驾驶数据集和高昂的训练成本,亟需一种新范式的出现。

对齐技术以通用基础大模型为着力点,通过微调或提示工程等定制化方式,能够精确、高效、专业地解决自动驾驶领域内的工程性问题,例如 DriveLLM^[9]、DriveGPT4^[10]、Think Twice^[11]、GPT-4V^[12]、DriveGPT^[13]等。因此,对齐技术作为基础大模型与自动驾驶间的纽带,可以有效弥合自动驾驶汽车与人类驾驶人行为和思维的差异,是未来发展自动驾驶大模型的一个重要方向。

尽管对齐技术在专业化大模型发展中是一个研究热点,但目前尚缺乏系统性的研究成果。本文首先对自动驾驶发展与大模型技术进行概述,从而衍生出对齐技术。接着,从微调和提示工程两个角度

出发,系统地回顾和分析了不同分类技术的结构和性能特征,并给出实际的应用案例。最后,基于现有研究,提出了对齐技术未来的发展方向,旨在为自动驾驶的进一步发展提供有价值的参考。

1 自动驾驶发展与大模型技术概述

1.1 自动驾驶发展

自动驾驶系统的发展是人工智能和机器学习领域的一项杰出成就,它标志着人类进入了新的技术时代。自动驾驶技术的演化可以追溯到20世纪80年代,当时的技术尚处于原型机和概念验证阶段。随着时间的推移,自动驾驶技术经历了从初步实验到高级自动化的转变,这一过程紧密地与大型神经网络模型技术的发展相互促进。

在21世纪初,随着深度学习的兴起,自动驾驶系统的发展迎来了飞跃。这一时期,卷积神经网络(CNN)和递归神经网络(RNN)等深度神经网络模型^[14]成为自动驾驶技术不可或缺的一部分。这些网络模型在图像识别、物体检测、语义分割以及自然语言处理等方面展现出了卓越的性能,极大地提升了自动驾驶系统的感知和决策能力。

自动驾驶汽车依赖于大量的传感器数据,包括来自摄像头、雷达、激光雷达(LiDAR)等的输入。深度神经网络模型通过对这些数据进行深入分析和学习,能够实现对外部环境的精确理解。这包括识别道路标志、行人、其他车辆以及各种障碍物,同时还能够对交通情况进行预测并据此作出驾驶决策。随着技术的进步,自动驾驶系统开始采用更加复杂和先进的Transformer注意力模型,进一步提高了自动驾驶系统的性能,使其能够更好地处理不确定性和复杂性,从而在更多样化的环境和场景下安全高效地运行。

1.2 大模型技术

大模型一般指的是通过海量数据训练、具有过亿参数量的大型神经网络模型。它以Transformer^[15]注意力机制模型为架构,通过自监督方式自学了大量“世界知识”,实现智能“涌现”,具有良好可扩展性和适应性,同时表现出强大的语言能力(如Bert擅长

的“文字填空”和 ChatGPT 精通的“文字接龙”),能够解决社会生活中普遍的上游任务。

在当今人工智能的研究领域中,大模型根据其处理的数据模态不同,可以分为大型语言模型、大型视觉模型以及大型多模态模型等。在本文中,这些模型通常被视作基础模型,并在自动驾驶等特定垂直领域中通过“大模型+对齐技术”的范式得到应用。大型语言模型,作为以 Transformer 为基础的人工智能算法之一,通过在广泛的文本数据集上进行训练,展现了其在理解自然语言及执行文本生成和理解等复杂任务方面的强大能力。这种模型在翻译、文本、问答等多项自然语言处理任务上均展现出卓越的性能。在自动驾驶领域,由于决策过程与人类驾驶员的推理过程极其相似,大型语言模型被用于通过文本描述和解释环境场景,以做出安全、合理且舒适的驾驶决策。

另一方面,大型视觉模型在目标检测、跟踪等计算机视觉任务中取得了显著的成就。例如,基于 Vision Transformer 架构的 Grounding DINO^[16]通过自监督学习,实现了从局部图像块预测全局图像特征的能力。Grounded-SAM^[17]模型则通过对数十亿个分割掩码的训练,展现了在没有样本输入的情况下进行分割的“涌现”能力,能够在给定合理提示的情况下进行精准的目标分割。扩散模型^[18],一种通过在图像上迭代添加噪声并进行反向扩散的过程,从随机噪声中恢复出高度真实图像的大型视觉模型,已广泛应用于图像和视频生成领域。

多模态 Transformer 的兴起促进了大型多模态模型的发展,这类模型能够处理包括文本、图像、音频、点云在内的跨模态输入,并用于执行更复杂的任务,如从图像生成文本描述、利用多模态输入进行综合分析和推理等。典型的多模态模型包括 CLIP^[19]、Video-LLaMA^[20]、Flamingo^[21]等,它们在多模态理解和生成任务中展现了前所未有的能力,为人工智能

领域带来了新的突破和应用前景。

相较于小模型来说,大模型不仅体现在参数量大、模型规模大,主要还具有以下特点:(1)智能“涌现”——遵循“缩放法则”,随着参数量、数据集大小和计算量的增加,大模型的性能会呈现出接近指数型的增长,其精度提升的速度远超线性增长预期。(2)领域应用“自适应”——大模型通过在多样化的大数据集上进行预训练,形成了庞大的语义信息库,使其能够轻松在不同领域间迁移应用。例如,可以摆脱自动驾驶“长尾”场景数据难以挖掘、小模型碎片化场景无法复用等局限。(3)多任务学习——大模型能够学习并执行多种任务,如感知、决策、规划等,通过共享特征表示来提升数据效率、泛化能力和整体性能。(4)标签数据“低依赖”——大模型采用自监督学习方式预训练,通过特征提取、表示学习等设计,揭示数据的内在结构和模式,极大减少了对标签数据的依赖。

2 对齐技术的衍生

技术变革推动自动驾驶行业迈向新的发展阶段,带来创新的思路和模式。在此背景下,大模型凭借其智能“涌现”、领域“自适应”和多任务学习的能力,为自动驾驶领域带来了创新的技术路径。目前,面向自动驾驶的大模型技术主要探索两种方向:一是为特定场景和任务预训练的专用大模型,这些模型在其专长领域内表现优异,但面对复杂多变的实际道路环境和交通条件时,挑战重重;二是预训练一个能涵盖所有自动驾驶场景和任务的通用大模型,理论上可行,但因须采集和标注极广泛的交通场景数据(大规模、高质量、多样性),实际操作中成本高昂,难度大。

鉴于这两种方法的局限性,自动驾驶领域开始寻求新的解决方案,即“大模型+对齐技术”的策略,以实现更高效、专业的应用。如图 1 所示,这一策略

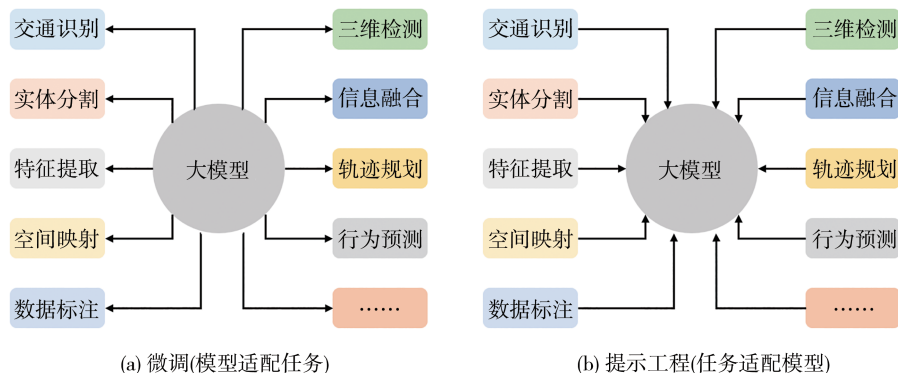


图 1 对齐技术

通过微调或提示工程等定制化方法,精确解决自动驾驶的具体问题。微调通过对少量专用数据集进行训练,从而适配特定任务,使模型与自动驾驶的下游任务(如目标检测、特征提取、轨迹规划、行为预测等)对齐。相反地,提示工程则是任务适配模型,通过设计合适的提示,激发大模型的潜能,实现特定化的行为输出,无须修改模型参数,极大降低成本。

3 微调

3.1 参数高效微调

在参数高效微调技术(PEFT)提出之前,模型微

调主要依赖于全量微调技术,即在新数据集上对预训练模型的全部参数进行重新训练。这种方法对于较小的模型而言,能有效迁移学习能力^[22],且计算成本较低。然而,随着模型规模的增大,全量微调开始遭遇到计算成本高和存储需求大的问题。为了解决这些挑战,谷歌在2019年针对BERT^[23]模型提出了PEFT^[24],标志着高效微调技术研究的开始,并随后在自动驾驶等领域得到应用。PEFT通过仅对部分参数进行微调并固定其他部分,显著降低了计算和存储成本。这种方法只须存储被微调的参数,从而在保持模型性能的同时,实现了更高的效率。PEFT的主要方法如图2所示。

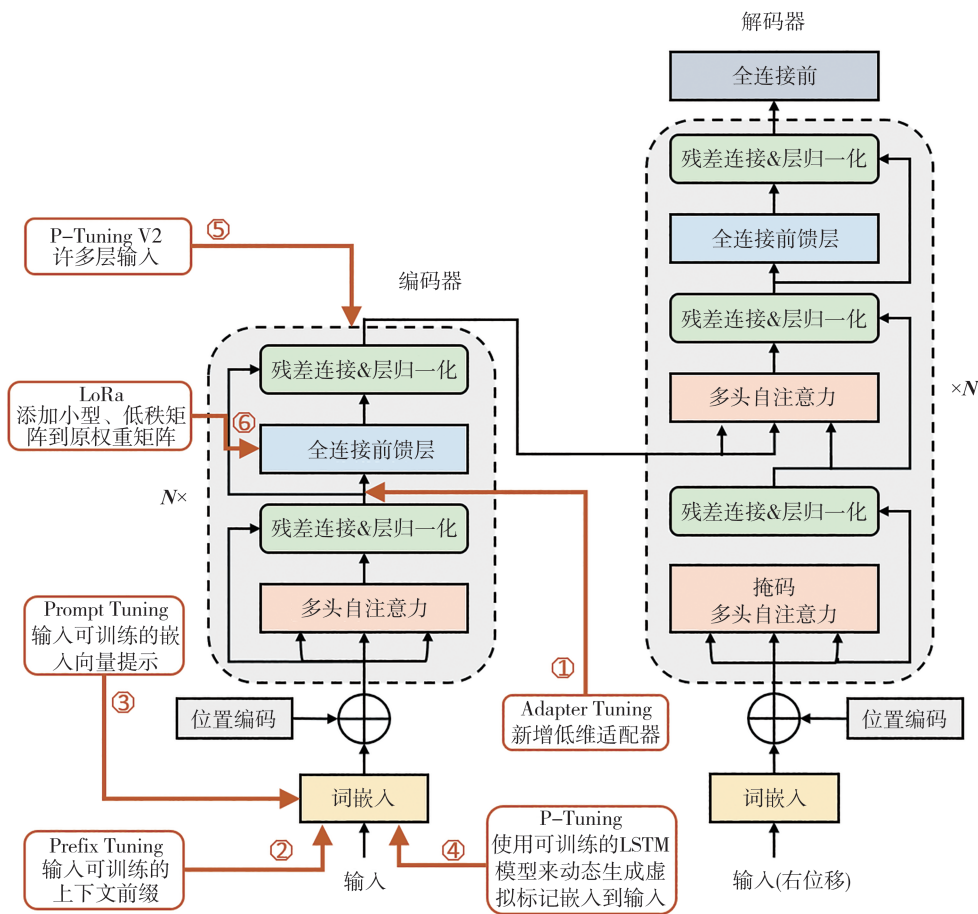


图2 高效微调技术

3.1.1 Adapter Tuning

Adapter Tuning 是一种在保持原始模型参数不变的情况下,通过增加可训练的适配器来提升模型微调效率和减少计算量的方法。这些适配器在模型的结构上增加了额外的层,但与Transformer的全连接前馈网络模块相比,适配器具有更小的隐藏维度。这种设计不仅减少了额外的计算负担,而且因为适

配器的维度较小,所以增加的参数量也很有限,从而实现了高效的微调。

适配器的工作机制包括将输入特征首先通过一个降维矩阵 W 映射到一个低维空间,然后通过一个非线性激活函数(如ReLU)进行处理,最后通过一个升维矩阵 Q 将特征映射回原来的高维空间。LV-Adapter^[25]利用了CLIP模型在视觉与语言嵌入对齐

方面的能力,通过特定任务的语言提示,将语言先验融合到视觉特征中,实现了在统一自动驾驶场景下的快速多任务学习适应。DriveAdapter^[26]通过结合BEVFusion^[27]和Mask2Former^[28]的分割技术,并在感知和规划任务的模型中插入Adapter,提升了分割性能。同时,虽然将新模态集成到预训练的大模型中可以显著增强车辆对驾驶场景的理解,但这通常需要大量的数据和计算资源。为了克服这一挑战,Driving with LLMs^[29]项目采用了向量编码器、向量模型以及冻结的大型语言模型(Adapter),通过优化向量编码器和模型权重,并减少向量结果中的误差,实现了将目标级向量模态融合到预训练的大语言模型中,有效提升了车辆对驾驶场景理解的能力。

3.1.2 Prefix Tuning

Prefix Tuning 是一种在不改变原有模型结构和参数的情况下,通过为模型的每一层增加任务特定的可训练前缀,以降低训练成本并实现高效精调的方法。具体而言,这种方法在模型的每一层的注意力机制中,为键(K)和值(V)添加预设的向量(分别是 P_k 和 P_v),从而形成新的注意力机制。这样,模型在处理输入时,会首先考虑这些前缀向量,从而调整其输出行为以更好地适应特定任务。

Prefix Tuning 的主要优势在于其无须调整模型的全部权重,而是通过简单地在输入中添加前缀向量来实现模型行为的调整,这大大降低了模型调优的复杂度和成本。在实际应用中,PATT^[30]是一种利用前缀调优技术的并行注意力模块,它能有效地实现自然语言文本和视频数据的对齐。另一个例子是 PM-DETR^[31],该模型通过引入提示记忆对齐机制,利用从提示域内存中提取的特定领域知识来减少不同领域之间的差异,从而提高模型在特定任务上的表现。这些应用展示了 Prefix Tuning 不仅在理论上具有创新性,而且在实际应用中也显示出了强大的灵活性和效率。

3.1.3 Prompt Tuning

Prompt Tuning 的核心是在预训练模型的输入中加入可学习的嵌入向量作为提示,通过优化这些向量使模型输出更适合特定任务的响应。这与 Prefix Tuning 类似,后者也在输入中添加可学习向量,但用途不同。Prefix Tuning 的可学习前缀主要提供直接的上下文信息,影响模型的整体行为;而 Prompt Tuning 的可学习向量(Prompt Tokens)旨在模仿自然语言提示,引导模型针对特定任务生成目标输出,这些向量更侧重于用较少的数量模拟自然语言提示。

VE-Prompt^[32]是一个有效的多任务学习框架,通过引入特定任务的视觉示例提示来指导模型学习高质量的任务特定表示,已在自动驾驶的目标检测、语义分割、可驾驶区域分割和车道检测中实现统一感知。GT-Prompt^[33]通过学习通用提示并为特定任务生成特定提示,帮助模型捕获任务的不变性和特定信息,在自动驾驶的3D多任务感知中得到成功应用。DA-Pro^[34]使用可学习的领域自适应提示生成每个领域的动态检测头,有效地将有标注的源域训练检测器推广到无标注的目标域,广泛应用于域自适应目标检测。P2P^[35]通过将三维点云转换为彩色二维图像,实现了用更少的可学习参数将二维预训练知识迁移到三维领域。

3.1.4 P-Tuning

与 Prompt Tuning 相似,P-Tuning 也利用“提示”或“指导”来改善模型在特定任务上的表现,但它们在嵌入提示的方式上有所不同。Prompt Tuning 采用静态和可训练的虚拟标记嵌入,这些嵌入在初始化后大多保持不变,仅在训练过程中更新一组固定的嵌入参数,操作简便但在处理复杂任务或需要细粒度控制时可能存在局限。P-Tuning 则通过一个可训练的 LSTM 模型(Prompt Encoder)动态生成虚拟标记嵌入,根据输入数据的差异产生不同的嵌入,从而提供更高的灵活性和适应性,适用于需要精细控制和深入理解复杂上下文的任务。Munir 等^[36]将 P-Tuning 应用于自动驾驶轨迹预测,通过提供可解释的驾驶情境推理,成功实现了对复杂驾驶场景的精细控制和理解。

3.1.5 P-Tuning V2

P-Tuning V2 是 P-Tuning 的升级版,主要改进在于引入多层连续提示而非仅在输入层插入,提升了模型微调时的参数量和适应性。这种设计允许在模型的多个层级中独立插入连续提示,增加了可训练参数,提高了对复杂任务的处理能力和在小型模型上的表现。MvNet^[37]采用了 P-Tuning V2 方法,通过将3D点云编码为多视角图像特征,并设计了一种基于 P-Tuning V2 的多视角提示融合模块,有效整合不同视角信息,减少了对大规模标注3D数据的依赖,提高了小样本分类的性能。

3.1.6 LoRA

LoRA 通过在模型的关键层添加小型低秩矩阵,调整模型行为而非改变其结构。这主要体现在多头自注意力和前馈神经网络部分,通过引入低秩矩阵 A 和 B 的乘积 AB ,实现原始权重矩阵的低秩近似调

整,适用于需要实时性的任务。基于LoRA,UDA^[38]模型通过无监督域自适应和生成对抗网络,有效处理源域和目标域之间的复杂域移位,提高目标分割性能。LoRA作为动态网络替代方案^[39],被应用于DyRoNet^[40]等动态路由网络,通过基于LoRA的微调实现特定环境下的快速响应和高精度检测。同时,LoRA也适用于危险驾驶行为分析,例如MM-LMF^[41]采用低秩多模态数据融合方法,有效识别危险驾驶行为。为提升模型效率,LoRA还可与模型压缩技术结合,如剪枝、量化等。

3.2 监督式微调

3.2.1 有监督微调

如图3所示,监督式微调^[42]首先在源数据集上预训练神经网络模型,然后创建目标模型,复制除输

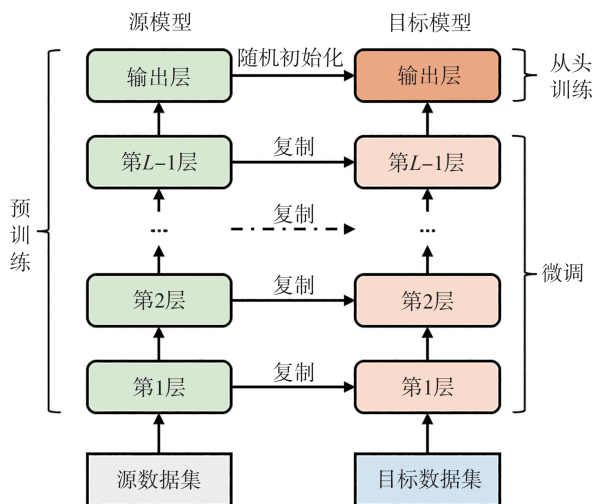


图3 监督式微调示意图

出层外的所有设计和参数。这些参数携带源数据集的知识,适用于目标数据集。由于源模型输出层与其数据集标签紧密相关,目标模型将替换该层以适应新的类别,并在目标数据集上进行训练,微调除输出层外的所有参数。例如,DriveLM^[43]通过在大量图视觉问答数据集上微调,显著提升了对驾驶场景的理解,并实现端到端自动驾驶。DriveMLM^[44]利用大量标注的行为规划数据集,通过自然语言指令和相应决策的人际互动,增强模型理解和反应人类指令的能力。尽管监督式微调有效弥补了模型与自动驾驶任务间的差异,但获取和标注大量驾驶场景数据既耗时又费力,这促使自动化标注技术的发展。

3.2.2 无监督微调

无监督微调利用无标签数据集,通过学习数据的内在结构或生成特征来提升模型的表示能力。例如,Rote-DA^[45]通过分析重复路线的驾驶数据和生成伪标签,实现3D物体检测。TriCC^[46]通过三角形约束对比度框架,利用多模态信息和时间序列动态学习3D无监督表示。Cohere3D^[47]采用对比学习算法,从长期输入序列中识别相干实例,实现2D到3D对象状态的表示。自监督微调如AD-PT^[48],利用未标记的点云数据生成统一表示,解耦预训练过程与微调任务,提升性能可扩展性和跨数据集应用。

然而,无监督微调面临着由于数据密度低导致的特征表示损坏和目标域表示不良的挑战^[49]。为解决这一问题,采用了3种策略,如图4所示。

(1)稀疏源数据重放^[50-52]:通过融合源域数据与

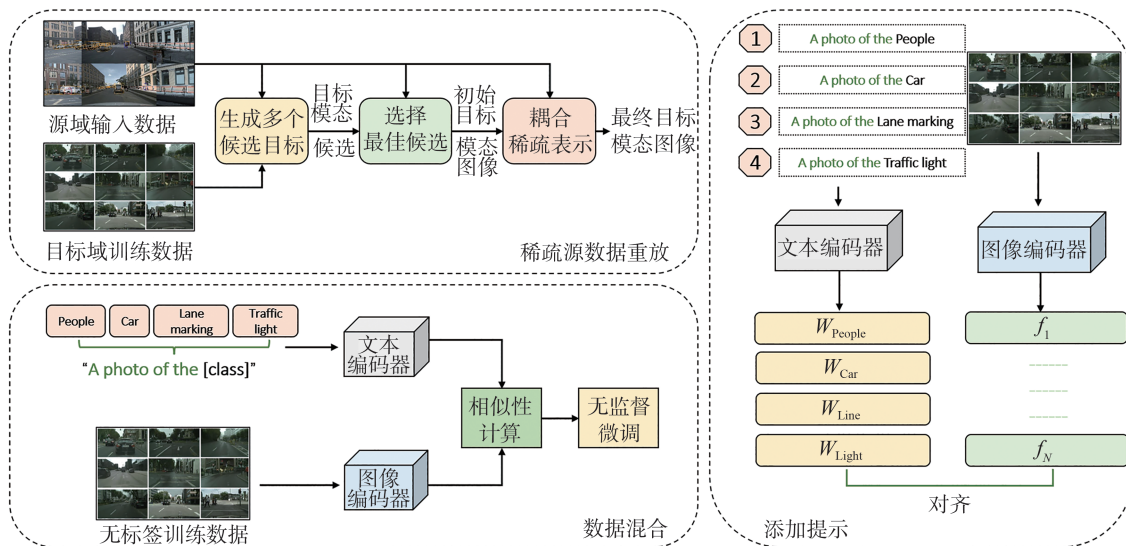


图4 无监督微调的3种可靠策略示意图

目标域数据避免特征表示损坏。

(2)数据混合^[53-56]:通过增加正样本和对齐项提高不常见特征的匹配权重,改善无监督学习效果。

(3)添加提示^[57-59]:结合对齐提示和目标数据的离散分布实现无监督微调。

3.3 强化学习微调

3.3.1 基于人类反馈的强化学习微调

基于人类反馈的强化学习微调(RLHF)是一种通过人类偏好来训练模型,以学习复杂行为的方法。

例如,WebGPT^[60]通过RLHF对GPT-3进行微调,增强了网络导航和信息检索能力。这种方法依据人类反馈训练奖励模型,使大模型输出更贴近用户期望。如图5所示,OpenAI通过在GPT系列发展中,如InstructGPT^[61]、ChatGPT^[62]到GPT-4^[63],使用不同规模的专业注释者提供的反馈作为奖励信号,更新模型参数。常见的奖励机制包括PPO^[64-65]、DQN^[66-67]、DDPG^[68]和DPO^[69-70]。

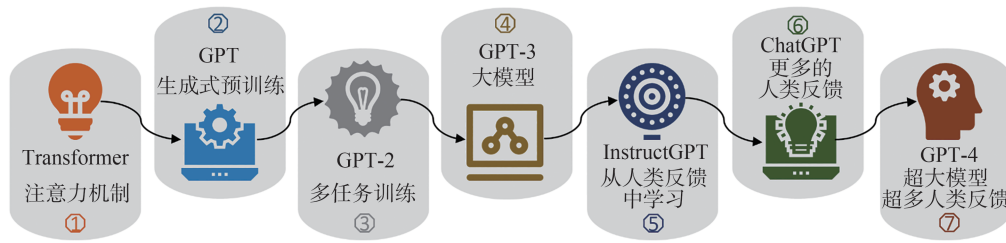


图5 GPT系列的演变

基于GPT思想,MTD-GPT^[71]模型通过专家数据收集、GPT训练和评估3个阶段,利用PPO算法和注意力机制训练专家模型,并将“状态-动作-奖励”元组转换为Token格式,以自然语言处理任务的形式输入GPT模型,实现复杂场景下的准确多任务决策。NuScenes-MQA^[72]采用RLHF构建视觉问答数据集,提高语言生成和场景识别的准确性。

然而,大模型可能因训练数据偏差、无法访问实时信息或理解生成限制而出现“幻觉”问题,尤其在自动驾驶等应用中可能导致安全风险。Sun等^[73]提出的事实增强RLHF对齐算法,通过引入额外的事

实信息增强奖励模型,缓解奖励黑客现象。RLHF-V^[74]通过细化人类反馈一致性,增强模型可信度。Volcano^[75]通过视觉信息反馈自然语言,调整初始响应,减少多模态幻觉。

3.3.2 基于AI反馈强化学习微调

为了解决RLHF中高成本和低效率的问题,基于AI反馈的强化学习微调(RLAIF)^[76]应运而生。如图6所示,RLAIF与RLHF原理相似,但反馈来源是AI,利用AI大模型代替人类进行评级,实现了无须人工标注的情况下,既超越了传统监督式微调的效果,又达到了与RLHF相当的性能。

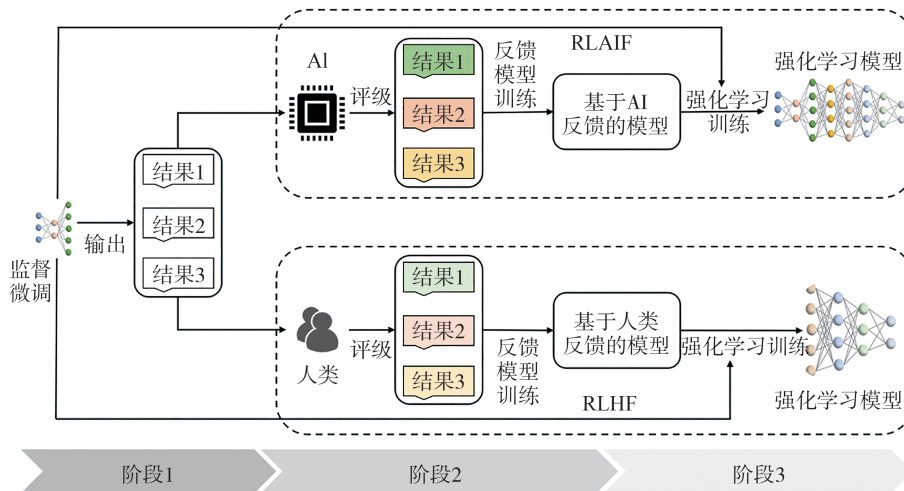


图6 RLHF与RLAIF原理示意图

RLAIF 通常在 AI 评级的数据集上训练奖励模型,借助比奖励模型更强大的 AI 标注器,以模型蒸馏的形式高效实现^[77]。此外,RLAIF 还可以直接使用 AI 反馈作为强化学习的奖励信号^[78],虽然这种方法对计算要求较高。作为 RLHF 的有效替代^[79],RLAIF 的目标是通过 AI 智能指导,最大化 AI 生成的偏好与人类偏好的一致性,减少对人工标注的依赖,加速大模型的训练和部署。

4 提示工程

提示工程利用精心设计的提示指令激发大模型的潜能,实现无须更新模型参数即可提升性能,具有低成本的优势。其核心在于通过调整大模型生成 Token 的概率,针对性地优化在特定领域的表达能力。常见的 Token 选择策略包括贪婪采样^[80]、光束

搜索^[81]、概率采样^[82]和基于温度采样^[83]。如图 7 所示,美国西北大学^[84]结合提示工程和 GPT 的强大知识库、推理及人机交互能力,有效减轻自动驾驶在复杂场景下的安全担忧。

4.1 指令提示

4.1.1 思维链

思维链(chain of thought, COT)^[85]通过模仿人类大脑的思维方式,引导大模型进行多步推理,从而提升其在少样本或零样本学习场景下的性能。这种方法不仅增强了大模型的推理能力,还促进了其“涌现”能力的展现。例如,DiLU^[86]项目通过 COT 从现实世界数据集中提取经验,将知识驱动能力引入自动驾驶系统。SurrealDriver^[87]利用基于大语言模型的生成式框架模拟人类驾驶行为,并通过 COT 提示开发的“教练代理”来评估和辅助“驾驶员代理”,使其积累经验并形成人性化的驾驶风格。

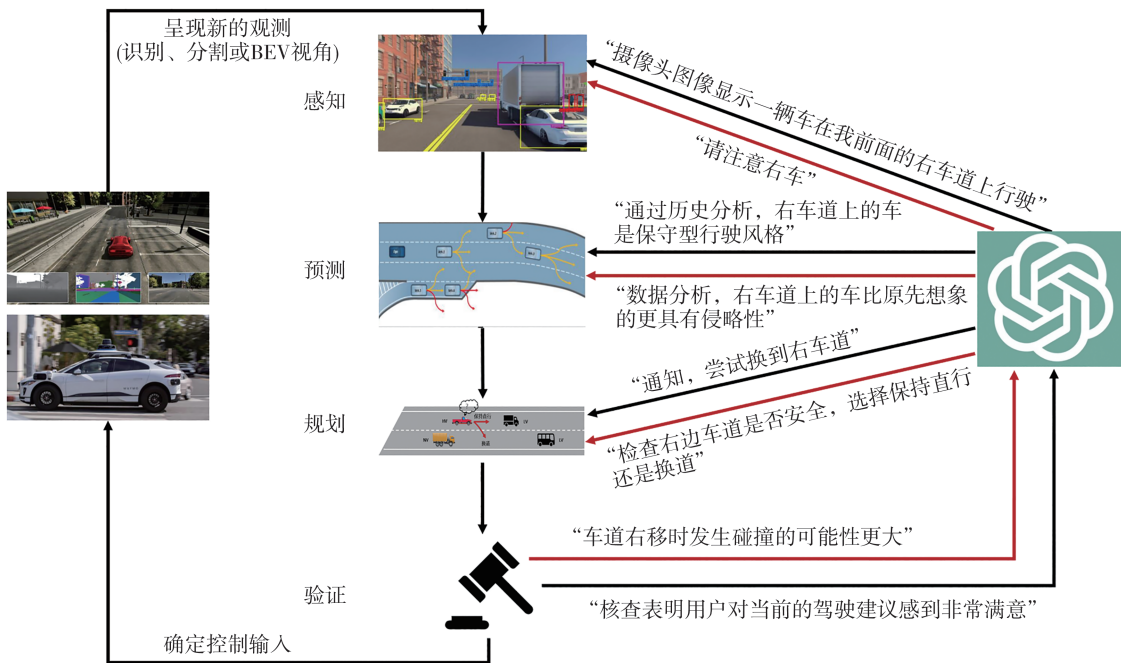


图7 面向自动驾驶的提示工程技术应用示例

然而,由于大模型生成的随机性,并不能保证每一次生成都是正确的,如何提高其鲁棒性和准确性成了另一个研究方向。为此,Hao 等^[88]提出了自一致性思维链,通过“自一致性”的解码策略,让模型从多种角度产生答案,并通过加权投票选择最可靠的答案,以提高输出的可信度。此外,还有循序渐进式^[89]、自我提问式^[90]、计划与解决式^[91]、元提示^[92-93]、知识增强提示^[94]和迭代型提示^[95]等策略,进一步提升大模型的性能。

4.1.2 思维树

思维树(tree of thought, TOT)^[96]是对思维链(chain of thought, COT)的进一步扩展,它通过与大模型的多次交互,采用连贯的子步骤逐步解决复杂问题。TOT 的核心思想在于引导大模型将问题分解为多层次的树状结构,并动态地选择最佳路径进行解决,如采用广度优先搜索^[96-97]或深度优先搜索等策略^[98]。这种方法使得大模型能够通过评估多种推理路径和自我评估来决定下一步的最佳行动方案,

必要时还可以进行前瞻或回溯,以做出全面考虑的决策。

例如,Zheng等^[99]通过将复杂的交通环境和参与者的历史轨迹信息转化为图像提示,利用TOT从大型语言模型中提取丰富的交通上下文信息,从而提升运动预测的准确性。同样,Yang等^[100]构建了一个包含自然驾驶实验和驾驶后访谈的人类驾驶员行为自然语言数据集,通过TOT使得驾驶员代理的行为与人类驾驶风格保持一致。

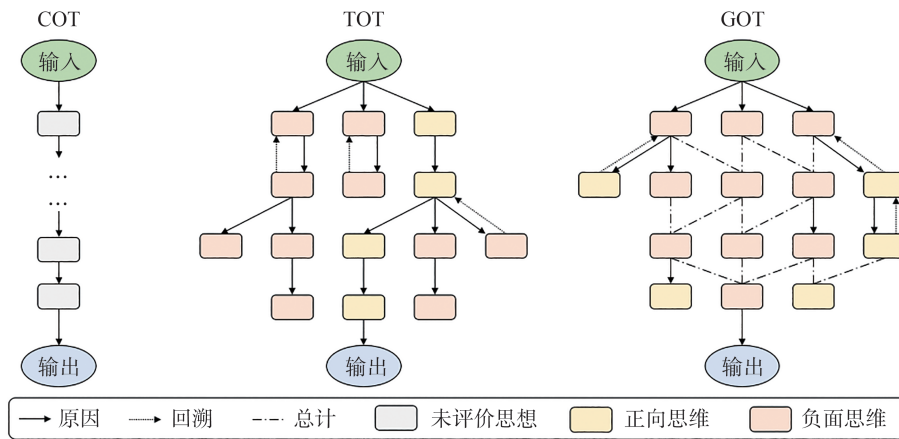


图8 思维链(COT)、思维树(TOT)和思维图(GOT)对比示意图

例如,DriveLM^[43]项目利用了一个包含语言信息的自动驾驶数据集,并结合大型语言模型的强大推理能力,采用GOT的思维图结构来构建驾驶场景的问答集。它通过提出“假设性”问题来推理未来可能发生的事件,从而实现自动驾驶的决策制定和确保规划的可解释性。

然而,由于TOT和GOT的实施过程需要与大模型进行多次交互,这不仅导致了高昂的计算成本,而且整体的交互逻辑也相对复杂。为了降低TOT或GOT查询交互过程的复杂度,提出了思维算法(algorithm of thought, AOT)^[104]。AOT策略通过直接在大模型内部完成迭代过程,实现了过程的轻量化,有效减少了计算资源的消耗。

4.2 少样本与零样本提示

4.2.1 少样本提示

少样本提示(few-shot prompting)^[105]通过向大模型提供少量但高质量的示例,包括针对目标任务的输入及其期望输出,旨在帮助大模型更准确地把握用户的意图,进而生成与用户预期一致的结果。例如,Liang等^[106]提出的多任务统一感知框架VE-Prompt,就是通过向模型提供特定任务的视觉示例

4.1.3 思维图

鉴于思维树(tree of thought, TOT)的严格树状结构在一定程度上限制了模型的推理能力,研究者们借鉴了知识图谱^[101]的概念,设计出了思维图(graph of thought, GOT)^[102]结构。GOT的创新之处在于其能够灵活转换不同的思考路径,进而优化推理过程。如图8所示,与TOT和COT相比,GOT能够有效地整合多个思考路径为一个统一的框架^[103],在提高单个思路的精度度的同时,也展现了更好的泛化能力。

作为提示,引导模型学习到高质量的任务特定表示。然而,这种方法面临着示例选择的质量与数量的挑战,通常需要专家对特定任务进行细致的尝试和调整。

此外,传统的添加提示示例方法常受限于文本输入窗口的长度限制。为了克服这一限制,提出了StreamingLLM^[107]框架,该框架使得模型处理无限长度文本成为可能,从而突破了文本窗口长度的约束,为模型提供了更大的灵活性和应用范围。

4.2.2 零样本提示

零样本提示(zero-shot prompting)^[108]涉及直接将任务输入模型,而不提供任何示例来明确所需的输出形式,这就对提示的质量提出了更高的要求。例如,Mao等^[109]提出的GPT-Driver便是一个典型例子,其中规划者的输入和输出通过精确的语言提示进行表达,完全不依赖于任何场景示例。这种方法在处理新颖和未见过的驾驶场景时,展现了卓越的泛化能力。相比于少样本提示,零样本提示虽然在性能上可能略有不足,但它在处理长文本的输入或输出方面具有明显优势,不会受到上下文长度的限制。

此外,为了提高提示质量的选择,自动提示工程

(automatic prompt engineering, APE)^[110]的提出实现了显著的改进。通过自动化的方式优化提示,APE在提高模型性能和适应性方面发挥了重要作用,但同样带来了高额的算力。

4.3 检索增强生成

检索增强生成(retrieval-augmented generation, RAG)^[111]是自然语言处理领域的一项主要技术,它通常涉及两个阶段:检索上下文相关信息和利用这些信息指导生成过程。RAG为大型语言模型与外部世界的有效互动提供了一种前景广阔的解决方案。依托外部知识,大型语言模型能够生成更加准确、可靠的答案。作为当前最受欢迎的应用框架之一,RAG通过结合检索模型和生成模型,利用来自私有或公开数据源的信息辅助文本生成,从而丰富了大模型的原始知识库。如图9所示,该技术通过从数据库、知识图谱等更多数据源添加背景信息,并通过训练来强化,不仅能有效克服文本窗口长度的限制,还能减少大模型产生的错误信息和应用的复杂度^[112],生成更准确且更符合实际场景的输出。

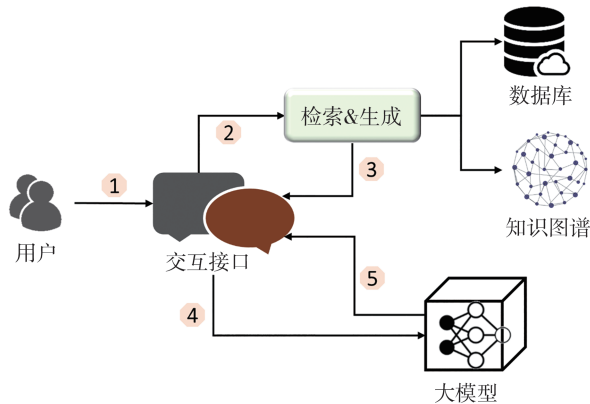


图9 检索增强生成技术的实现流程

RAG-Driver^[113]是一种新型的检索增强多模态大型语言模型,通过上下文学习实现了高性能、可解释性以及可扩展性的自动驾驶。BEV-CLIP^[114]则是一种利用描述性文本检索相应场景的多模态鸟瞰图检索方法,有效地解决了自动驾驶场景中的全局特征表示不足和文本检索能力不佳的问题。该方法利用大型语言模型的语义特征提取能力进行零样本检索,并结合了知识图谱的半结构化信息,以增强语言嵌入的语义丰富性和多样性。RealGen^[115]通过RAG技术检索交通场景行为,并以无梯度方式组合这些行为来合成新场景,广泛应用于可控流量场景生成领域。

然而,由于RAG在实时任务中可能引入高延迟,混合检索增强生成(HybridRAG)^[116]框架应运而生。这一框架旨在通过合并额外的上下文信息来增强传统语言模型的性能,同时尽量减少计算延迟,以适应更广泛的应用场景。

4.4 推理与行动

尽管大模型在理解语言和进行交互式决策方面展现了卓越的能力,但它们在推理(如COT提示技术)和执行(如轨迹生成技术)方面的能力通常被单独研究。推理与行动(ReAct)^[117]项目致力于解决大型模型在语言理解和交互式决策任务中如何将推理和行动能力结合使用的挑战。与先前的模型(例如RAG)不同,这些模型虽然能够进行任务分解、推理以及使用外部工具,但它们未能将这些能力有机整合成一个闭环系统。ReAct框架通过使大型模型以交错的方式生成推理轨迹和针对特定任务的行动,实现了对行动计划的指导、跟踪和更新,同时能够处理异常情况。此外,ReAct通过多次迭代执行,将模型功能与外部资源结合,利用外部工具收集和交互信息,最终产生更加可靠和真实的反应。

作为一个精心设计的Prompt,ReAct能够理解和分析自动驾驶环境,并与之交互,展现了在处理推理和长尾场景时的卓越能力。这为类似人类的自动驾驶发展提供了宝贵的洞见。例如,Fu等^[118]提出的自动驾驶系统,展示了利用大型语言模型以类人的方式理解驾驶环境的可能性,并探索了其在应对复杂场景时的推理、解释和记忆能力。如图10所示,Cui等^[119]提出的以人为本的自动驾驶大模型,促进了车辆与乘客的互动,并适应了乘客的偏好,实现了一个以用户为中心、透明和自适应的自动驾驶生态系统。

5 研究挑战和发展趋势

将大模型技术应用于自动驾驶领域是一个充满活力且令人兴奋的发展方向,尽管如此,该领域仍然面临着需要进一步研究和重大突破的挑战。以下是一些未来研究可能会关注的方向。

(1)自动对齐技术。随着AI模型复杂性的增加,迫切需要开发能够自动评估和对齐模型与特定任务之间关系的系统。目前的过程,无论是微调还是提示工程,都依赖于人工干预,准确性和可靠性成为了主要挑战。

(2)数据引擎。传统的从头开始的预训练方法受限于缺乏大规模、高质量、多样化的自动驾驶数据

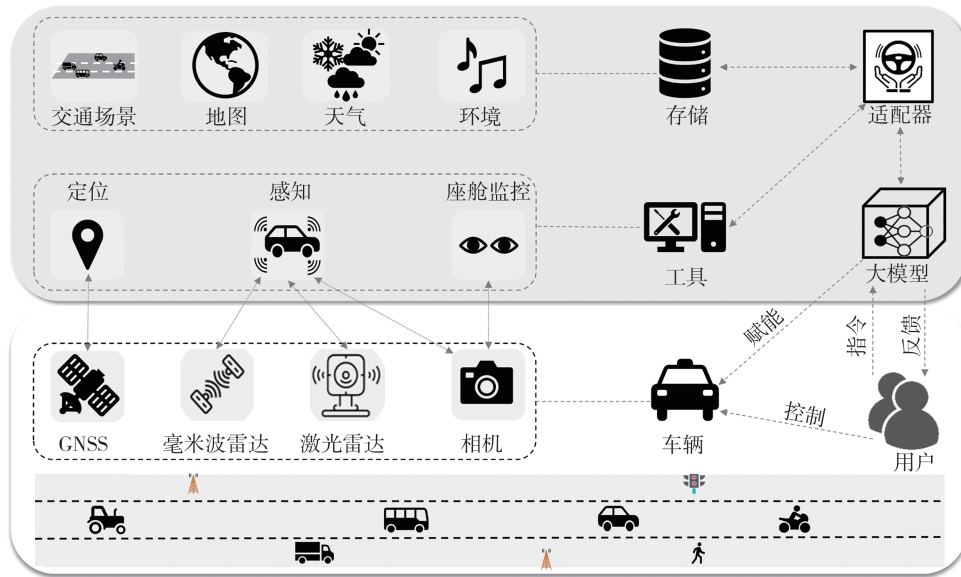


图10 以用户为中心的自动驾驶大模型集成框架

集。尽管“大模型+对齐技术”的范式在此方面具有显著优势,但这仍然只是相对的。具体而言,微调需要特定领域的数据,而提示工程需要专门的提示数据集来引导模型生成专业化的输出。数据引擎可实现数据的自动化标注,为该领域提供高效数据处理与分析。

(3)欺骗性对齐的实证研究。所谓的幻觉,指的是模型生成的内容基于不准确或虚构的信息。欺骗性对齐,也称为幻觉对齐,旨在解决数据偏差、模型无法访问最新信息,或模型在理解和生成上下文信息时的固有局限性等问题。

此外,现有的“大模型+对齐技术”范式主要停留在仿真阶段,难以在实际车辆或大规模应用中落地。同时现有大部分文献弱化了真实交通场景的多样性和复杂性,忽略了车端计算资源的局限性,尤其缺少了对安全、法律、伦理等方面的考虑等。因此,未来的研究挑战和发展趋势将集中在以下几个方面。

(1)复杂性和泛化能力。虽然RLHF、RLAIF、COT、TOT、GOT、AOT、RAG和ReAct等先进对齐技术显著提高了泛化性能,但在数据分布变化或遇到复杂多变的交通场景和环境时,需要更为强大的泛化能力。尤其是在处理一些未知的长尾场景时,也可能面临过拟合和泛化性能下降的问题。因此,需要构建并整合自动驾驶领域专用的知识图谱和向量数据库技术以提高性能。

(2)计算资源需求。尽管参数高效微调技术减少了传统从头训练或全量微调大模型的高昂训练成

本,但仍受限于车端的计算资源。因此,研究量化、剪枝、知识蒸馏等模型压缩技术,以及开发高性能的车端级芯片,是未来的重要研究方向。

(3)数据隐私和安全。随着大模型的广泛应用,对齐技术也涉及到大量的用户隐私数据。如何保护用户数据隐私,防止人车交互过程中数据的泄露和滥用,是一个重要的挑战,需要通信技术的进步以及相关法律和制度的完善。

(4)可解释性。对齐技术输出的结果往往是黑盒的,缺乏可解释性。如何解释和理解模型的决策过程是一个挑战,未来需要可解释AI技术的发展。

(5)模型偏见和公平性。对齐技术可能存在偏见和不公平性,例如在交通参与者识别中的偏见。自动驾驶可能在识别行人、自行车骑手、摩托车骑手等方面存在准确率差异,导致决策时偏向某些道路用户,而忽略其他道路用户的安全。因此,确保数据集的多样性,开发鲁棒的算法,并进行广泛测试,以确保在各种复杂多样的条件下都能公平、安全地运行。

6 结语

智能化、专业化、高效化是汽车变革的重要方向,以GPT等通用基础大模型为着力点,通过对齐技术实现了与人类驾驶人行为和思维一致的感知、预测、规划、交互等能力,使自动驾驶系统迈向更高级别发展成为了可能。本文首先对自动驾驶发展与大

模型技术进行了概述,从而衍生出对齐技术。然后,分别从微调和提示工程两个角度进行了综述,系统化梳理并剖析各分类技术的结构或性能特点。最后,基于现有研究提出了对齐技术的未来发展方向,为促进自动驾驶迈向更高级别发展提供参考。

参考文献

- [1] 李晓华. 自动驾驶的发展现状、挑战与应对[J]. 人民论坛, 2023(18): 68-72.
LI X H. Development status, challenges and responses for autonomous driving[J]. People's Tribune, 2023(18): 68-72.
- [2] TOUVRON H, LAVRIL T, IZACARD G, et al. LLaMA: open and efficient foundation language models[J]. arXiv preprint arXiv:2302.13971, 2023.
- [3] FLORIDI L, CHIRIATTI M. GPT-3: its nature, scope, limits, and consequences [J]. Minds and Machines, 2020, 30 (4) : 681-694.
- [4] TEAM G, ANIL R, BORGEAUD S, et al. Gemini: a family of highly capable multimodal models [J]. arXiv preprint arXiv: 2312.11805, 2023.
- [5] SUN Y, WANG S, FENG S, et al. Ernie 3.0: large-scale knowledge enhanced pre-training for language understanding and generation[J]. arXiv preprint arXiv:2107.02137, 2021.
- [6] REN X, ZHOU P, MENG X, et al. PanGu- Σ : towards trillion parameter language model with sparse heterogeneous computing [J]. arXiv preprint arXiv:2303.10845, 2023.
- [7] 李升波, 占国建, 蒋宇轩, 等. 类脑学习型自动驾驶决策系统的关键技术[J]. 汽车工程, 2023,45(9): 1499-1515.
LI S B, ZHAN G J, JIANG Y X, et al. Key technologies of brain inspired decision and control intelligence for autonomous driving systems[J]. Automotive Engineering, 2023,45(9): 1499-1515.
- [8] CAO D, ZOLOTAS A, WANG M, et al. Preface for feature topic on human driver behaviours for intelligent vehicles [J]. Automotive Innovation, 2024: 1-3.
- [9] CUI Y, HUANG S, ZHONG J, et al. DriveLLM: charting the path toward full autonomous driving with large language models [J]. IEEE Transactions on Intelligent Vehicles, 2023: 1-15.
- [10] XU Z, ZHANG Y, XIE E, et al. Drivegpt4: interpretable end-to-end autonomous driving via large language model [J]. arXiv preprint arXiv:2310.01412, 2023.
- [11] JIA X, WU P, CHEN L, et al. Think twice before driving: towards scalable decoders for end-to-end autonomous driving [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 21983-21994.
- [12] WEN L, YANG X, FU D, et al. On the road with GPT-4V (ision): early explorations of visual-language model on autonomous driving [J]. arXiv preprint arXiv:2311.05332, 2023.
- [13] KUMAR V, JAIN S, SONI N, et al. Drive GPT—an AI based generative driver model [C]. SAE Paper 2024-26-0025.
- [14] 张顺, 龚怡宏, 王进军. 深度卷积神经网络的发展及其在计算机视觉领域的应用[J]. 计算机学报, 2019, 42(3): 453-482.
ZHANG S, GONG Y H, WANG J J. The development of deep convolution neural network and its applications on computer vision [J]. Chinese Journal of Computers, 2019, 42(3): 453-482.
- [15] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]. Advances in Neural Information Processing Systems: Vol. 30. Curran Associates, Inc., 2017.
- [16] LIU S, ZENG Z, REN T, et al. Grounding DINO: marrying DINO with grounded pre-training for open-set object detection [J]. arXiv preprint arXiv:2303.05499, 2023.
- [17] ZHAO Z. Enhancing autonomous driving with grounded-segment anything model: limitations and mitigations [C]. 2023 IEEE 3rd International Conference on Data Science and Computer Application (ICDSCA), 2023: 1258-1265.
- [18] YANG L, ZHANG Z, SONG Y, et al. Diffusion models: a comprehensive survey of methods and applications [J]. ACM Computing Surveys, 2023, 56(4): 1-39.
- [19] YANG G, QIAO Y, SHI J, et al. Long-tailed object mining based on CLIP model for autonomous driving [C]. 2022 4th International Conference on Control and Robotics (ICCR), 2022: 348-352.
- [20] ZHANG H, LI X, BING L. Video-LLaMA: an instruction-tuned audio-visual language model for video understanding [J]. arXiv preprint arXiv:2306.02858, 2023.
- [21] ALAYRAC J B, DONAHUE J, LUC P, et al. Flamingo: a visual language model for few-shot learning [J]. Advances in Neural Information Processing Systems, 2022, 35: 23716-23736.
- [22] DAI A M, LE Q V. Semi-supervised sequence learning [J]. Advances in Neural Information Processing Systems, 2015, 28.
- [23] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [J]. arXiv preprint arXiv:1810.04805, 2018.
- [24] HOULSBY N, GIURGIU A, JASTRZEBSKI S, et al. Parameter-efficient transfer learning for NLP [C]. Proceedings of the 36th International Conference on Machine Learning. PMLR, 2019: 2790-2799.
- [25] LIANG X, WU Y, HAN J, et al. Effective adaptation in multi-task co-training for unified autonomous driving [J]. Advances in Neural Information Processing Systems, 2022, 35: 19645-19658.
- [26] JIA X, GAO Y, CHEN L, et al. DriveAdapter: breaking the coupling barrier of perception and planning in end-to-end autonomous driving [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 7953-7963.
- [27] LIANG T, XIE H, YU K, et al. BEVFusion: a simple and robust lidar-camera fusion framework [J]. Advances in Neural Information Processing Systems, 2022, 35: 10421-10434.
- [28] MARCUZZI R, NUNES L, WIESMANN L, et al. Mask-based panoptic LiDAR segmentation for autonomous driving [J]. IEEE Robotics and Automation Letters, 2023, 8(2): 1141-1148.
- [29] CHEN L, SINAVSKI O, HÜNERMANN J, et al. Driving with

- LLMs: fusing object-level vector modality for explainable autonomous driving[J]. arXiv preprint arXiv:2310.01957, 2023.
- [30] YU B, CHANG J, LIU L, et al. Towards a unified view on visual parameter-efficient transfer learning [J]. arXiv preprint arXiv:2210.00788, 2022.
- [31] JIA P, LIU J, YANG S, et al. PM-DETR: domain adaptive prompt memory for object detection with transformers [J]. arXiv preprint arXiv:2307.00313, 2023.
- [32] LIANG X, NIU M, HAN J, et al. Visual exemplar driven task-prompting for unified perception in autonomous driving[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 9611-9621.
- [33] LIANG X, LIANG X, XU H. Multi-task perception for autonomous driving [M]//FAN R, GUO S, BOCUS M J. Autonomous driving perception: fundamentals and applications. Singapore: Springer Nature, 2023: 281-321.
- [34] LI H, ZHANG R, YAO H, et al. Learning domain-aware detection head with prompt tuning[J]. Advances in Neural Information Processing Systems, 2023, 36: 4248-4262.
- [35] WANG Z, YU X, RAO Y, et al. P2P: tuning pre-trained image models for point cloud analysis with point-to-pixel prompting[J]. Advances in Neural Information Processing Systems, 2022, 35: 14388-14402.
- [36] MUNIR F, MIHAYLOVA T, AZAM S, et al. Exploring large language models for trajectory prediction: a technical perspective [C]. Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction. New York, NY, USA: Association for Computing Machinery, 2024: 774-778.
- [37] PENG H, LI B, ZHANG B, et al. Multi-view vision fusion network: can 2D pre-trained model boost 3D point cloud data-scarce learning? [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023: 1-1.
- [38] SAFFARI M, KHODAYAR M. Low-rank sparse generative adversarial unsupervised domain adaptation for multitarget traffic scene semantic segmentation[J]. IEEE Transactions on Industrial Informatics, 2024, 20(2): 2564-2576.
- [39] FILATOV N, KINDULOV M. Low rank adaptation for stable domain adaptation of vision transformers [J]. Optical Memory and Neural Networks, 2023, 32(2): S277-S283.
- [40] HUANG X, CHENG Z Q, HE J Y, et al. DyRoNet: dynamic routing and low-rank adapters for autonomous driving streaming perception[EB/OL]. (2024-03-08)[2024-03-20]. <https://arxiv.org/abs/2403.05050v3>.
- [41] HAO Z, LI Z, DANG X, et al. MM-LMF: a low-rank multimodal fusion dangerous driving behavior recognition method based on FMCW signals[J]. Electronics, 2022, 11(22): 3800.
- [42] HAN J, LIANG X, XU H, et al. SODA10M: towards large-scale object detection benchmark for autonomous driving [J]. 2023. DOI:10.48550/arXiv.2106.11118.
- [43] SIMA C, RENZ K, CHITTA K, et al. DriveLM: driving with graph visual question answering [J]. arXiv preprint arXiv:2312.14150, 2023.
- [44] WANG W, XIE J, HU C, et al. DriveMLM: aligning multi-modal large language models with behavioral planning states for autonomous driving[J]. arXiv preprint arXiv:2312.09245, 2023.
- [45] YOU Y, PHOO C P, LUO K, et al. Unsupervised adaptation from repeated traversals for autonomous driving [J]. Advances in Neural Information Processing Systems, 2022, 35: 27716-27729.
- [46] PANG B, XIA H, LU C. Unsupervised 3D point cloud representation learning by triangle constrained contrast for autonomous driving[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 5229-5239.
- [47] XIE Y, CHEN H, MEYER G P, et al. Cohere3D: exploiting temporal coherence for unsupervised representation learning of vision-based autonomous driving [J]. arXiv preprint arXiv:2402.15583, 2024.
- [48] YUAN J, ZHANG B, YAN X, et al. AD-PT: autonomous driving pre-training with large-scale point cloud dataset [J]. Advances in Neural Information Processing Systems, 2023, 36: 47914-47933.
- [49] LI S, CHEN D, CHEN Y, et al. Unsupervised finetuning [J]. arXiv preprint arXiv:2110.09510, 2021.
- [50] WANG J, LI W, WANG Y, et al. Representation-enhanced status replay network for multisource remote-sensing image classification [J]. IEEE Transactions on Neural Networks and Learning Systems, 2023: 1-13.
- [51] VEMULAPALLI R, VAN NGUYEN H, ZHOU S K. Unsupervised cross-modal synthesis of subject-specific scans [C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 630-638.
- [52] THIAGARAJAN J J, RAMAMURTHY K N, SPANIAS A. Multiple kernel sparse representations for supervised and unsupervised learning [J]. IEEE Transactions on Image Processing, 2014, 23(7): 2905-2915.
- [53] ZHU C, ZHANG Q, CAO L, et al. Mix2Vec: unsupervised mixed data representation [C]. 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), 2020: 118-127.
- [54] SHEN Z, LIU Z, LIU Z, et al. Un-mix: rethinking image mixtures for unsupervised visual representation learning [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(2): 2216-2224.
- [55] WISDOM S, TZINIS E, ERDOGAN H, et al. Unsupervised speech separation using mixtures of mixtures [C]. ICML 2020 Workshop on Self-supervision in Audio and Speech, 2020.
- [56] ISHIDA N, NAGATSU Y, HASHIMOTO H. Unsupervised anomaly detection based on data augmentation and mixing [C]. IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, 2020: 529-533.
- [57] TANWISUTH K, ZHANG S, ZHENG H, et al. POUF: prompt-oriented unsupervised fine-tuning for large pre-trained models

- [C]. International Conference on Machine Learning. PMLR, 2023: 33816–33832.
- [58] LIU X, JI K, FU Y, et al. P-Tuning: prompt tuning can be comparable to fine-tuning across scales and tasks[C]//MURESAN S, NAKOV P, VILAVICENCIO A. Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Dublin, Ireland: Association for Computational Linguistics, 2022: 61–68.
- [59] XU Z, WANG C, QIU M, et al. Making pre-trained language models end-to-end few-shot learners with contrastive prompt tuning[C]. Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining. New York, NY, USA: Association for Computing Machinery, 2023: 438–446.
- [60] NAKANO R, HILTON J, BALAJI S, et al. WebGPT: browser-assisted question-answering with human feedback[J]. arXiv preprint arXiv:2112.09332, 2021.
- [61] OUYANG L, WU J, JIANG X, et al. Training language models to follow instructions with human feedback[J]. Advances in Neural Information Processing Systems, 2022, 35: 27730–27744.
- [62] WU T, HE S, LIU J, et al. A brief overview of chatgpt: the history, status quo and potential future development[J]. IEEE/CAA Journal of Automatica Sinica, 2023, 10(5): 1122–1136.
- [63] CUI G, YUAN L, DING N, et al. UltraFeedback: boosting language models with high-quality feedback[J]. arXiv preprint arXiv:2310.01377, 2023.
- [64] CAO Y, IVANOVIC B, XIAO C, et al. Reinforcement learning with human feedback for realistic traffic simulation[J]. arXiv preprint arXiv:2309.00709, 2023.
- [65] LINDNER D. Algorithmic foundations for safe and efficient reinforcement learning from human feedback[D]. ETH Zurich, 2023.
- [66] LIU J, HANG P, QI X, et al. MTD-GPT: a multi-task decision-making GPT model for autonomous driving at unsignalized intersections[C]. 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2023: 5154–5161.
- [67] KWON M, XIE S M, BULLARD K, et al. Reward design with language models[J]. arXiv preprint arXiv:2303.00001, 2023.
- [68] CAI X Q, ZHANG Y J, CHIANG C K, et al. Imitation learning from vague feedback[J]. Advances in Neural Information Processing Systems, 2024, 36.
- [69] RAFAILOV R, SHARMA A, MITCHELL E, et al. Direct preference optimization: your language model is secretly a reward model[J]. Advances in Neural Information Processing Systems, 2024, 36.
- [70] YANG Y, BHATT N P, INGEBRAND T, et al. Fine-tuning language models using formal methods feedback[J]. arXiv preprint arXiv:2310.18239, 2023.
- [71] LIU J, HANG P, QI X, et al. MTD-GPT: a multi-task decision-making gpt model for autonomous driving at unsignalized intersections[C]. 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC), 2023: 5154–5161.
- [72] INOUE Y, YADA Y, TANAHASHI K, et al. NuScenes-MQA: integrated evaluation of captions and QA for autonomous driving datasets using markup annotations[C]. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024: 930–938.
- [73] SUN Z, SHEN S, CAO S, et al. Aligning large multimodal models with factually augmented RLHF[J]. arXiv preprint arXiv:2309.14525, 2023.
- [74] YU T, YAO Y, ZHANG H, et al. RLHF-V: towards trustworthy MLLMs via behavior alignment from fine-grained correctional human feedback[J]. arXiv preprint arXiv:2312.00849, 2023.
- [75] LEE S, PARK S H, JO Y, et al. Volcano: mitigating multimodal hallucination through self-feedback guided revision[J]. arXiv preprint arXiv:2311.07362, 2023.
- [76] LEE H, PHATALE S, MANSOOR H, et al. RLAI: scaling reinforcement learning from human feedback with AI feedback[J]. arXiv preprint arXiv:2309.00267, 2023.
- [77] YANG K, KLEIN D, CELIKYILMAZ A, et al. RLCD: reinforcement learning from contrast distillation for language model alignment[J]. arXiv preprint arXiv:2307.12950, 2023.
- [78] WU T, ZHU B, ZHANG R, et al. Pairwise proximal policy optimization: harnessing relative feedback for LLM alignment[J]. arXiv preprint arXiv:2310.00212, 2023.
- [79] HÖGLUND S, KHEDRI J. Comparison between RLHF and RLAI in fine-tuning a large language model[R]. Stockholm: KTH Royal Institute of Technology, 2023. <https://www.diva-portal.org/smash/get/diva2:1782683/FULLTEXT01.pdf>.
- [80] AKINWANDE V, JIANG Y, SAM D, et al. Understanding prompt engineering may not require rethinking generalization[J]. arXiv preprint arXiv:2310.03957, 2023.
- [81] PRYZANT R, ITER D, LI J, et al. Automatic prompt optimization with “gradient descent” and beam search[J]. arXiv preprint arXiv:2305.03495, 2023.
- [82] CHANG C C, REITTER D, AKSITOV R, et al. KL-Divergence guided temperature sampling[J]. arXiv preprint arXiv:2306.01286, 2023.
- [83] ZHU Y, LI J, LI G, et al. Improving code generation by dynamic temperature sampling[J]. arXiv preprint arXiv:2309.02772, 2023.
- [84] WANG Y, JIAO R, LANG C, et al. Empowering autonomous driving with large language models: a safety perspective[J]. arXiv preprint arXiv:2312.00812, 2023.
- [85] YASUNAGA M, CHEN X, LI Y, et al. Large language models as analogical reasoners[J]. arXiv preprint arXiv:2310.01714, 2023.
- [86] WEN L, FU D, LI X, et al. DiLu: a knowledge-driven approach to autonomous driving with large language models[J]. arXiv preprint arXiv:2309.16292, 2023.
- [87] JIN Y, SHEN X, PENG H, et al. SurrealDriver: designing generative driver agent simulation framework in urban contexts based on large language model[J]. arXiv preprint arXiv:2309.13193, 2023.
- [88] HAO S, GU Y, MA H, et al. Reasoning with language model is

- planning with world model[J]. arXiv preprint arXiv:2305.14992, 2023.
- [89] ZHOU D, SCHÄRLI N, HOU L, et al. Least-to-most prompting enables complex reasoning in large language models [J]. arXiv preprint arXiv:2205.10625, 2022.
- [90] PRESS O, ZHANG M, MIN S, et al. Measuring and narrowing the compositionality gap in language models [J]. arXiv preprint arXiv:2210.03350, 2022.
- [91] WANG L, XU W, LAN Y, et al. Plan-and-solve prompting: improving zero-shot chain-of-thought reasoning by large language models[J]. arXiv preprint arXiv:2305.04091, 2023.
- [92] HOU Y, DONG H, WANG X, et al. MetaPrompting: learning to learn better prompts [J]. arXiv preprint arXiv: 2209.11486, 2022.
- [93] DE WYNTER A, WANG X, GU Q, et al. On meta-prompting [J]. arXiv preprint arXiv:2312.06562, 2023.
- [94] LIU J, LIU A, LU X, et al. Generated knowledge prompting for commonsense reasoning [J]. arXiv preprint arXiv: 2110.08387, 2021.
- [95] WANG B, DENG X, SUN H. Iteratively prompt pre-trained language models for chain of thought [J]. arXiv preprint arXiv: 2203.08383, 2022.
- [96] YAO S, YU D, ZHAO J, et al. Tree of thoughts: deliberate problem solving with large language models [J]. *Advances in Neural Information Processing Systems*, 2024, 36.
- [97] WENG G, ANDRZEJAK A. Automatic bug fixing via deliberate problem solving with large language models [C]. 2023 IEEE 34th International Symposium on Software Reliability Engineering Workshops (ISSREW), 2023: 34–36.
- [98] QIN Y, LIANG S, YE Y, et al. ToolLLM: facilitating large language models to master 16000+ real-world APIs [J]. arXiv preprint arXiv:2307.16789, 2023.
- [99] ZHENG X, WU L, YAN Z, et al. Large language models powered context-aware motion prediction [J]. arXiv preprint arXiv: 2403.11057, 2024.
- [100] YANG R, ZHANG X, FERNANDEZ-LAAKSONEN A, et al. Driving style alignment for llm-powered driver agent [J]. arXiv preprint arXiv:2403.11368, 2024.
- [101] WEN Y, WANG Z, SUN J. MindMap: knowledge graph prompting sparks graph of thoughts in large language models [J]. arXiv preprint arXiv:2308.09729, 2023.
- [102] BESTA M, BLACH N, KUBICEK A, et al. Graph of thoughts: solving elaborate problems with large language models [J]. arXiv preprint arXiv:2308.09687, 2023.
- [103] LEI B, LIN PEI H, LIAO C, et al. Boosting logical reasoning in large language models through a new framework: the graph of thought [J]. arXiv preprint arXiv:2308.08614, 2023.
- [104] SEL B, AL-TAWAHA A, KHATTAR V, et al. Algorithm of thoughts: enhancing exploration of ideas in large language models [J]. arXiv preprint arXiv:2308.10379, 2023.
- [105] GU Y, HAN X, LIU Z, et al. PPT: pre-trained prompt tuning for few-shot learning [J]. arXiv preprint arXiv: 2109.04332, 2021.
- [106] LIANG X, NIU M, HAN J, et al. Visual exemplar driven task-prompting for unified perception in autonomous driving [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023: 9611–9621.
- [107] XIAO G, TIAN Y, CHEN B, et al. Efficient streaming language models with attention sinks [J]. arXiv preprint arXiv: 2309.17453, 2023.
- [108] ALLINGHAM J U, REN J, DUSENBERRY M W, et al. A simple zero-shot prompt weighting technique to improve prompt ensembling in text-image models [C]. *Proceedings of the 40th International Conference on Machine Learning*. PMLR, 2023: 547–568.
- [109] MAO J, QIAN Y, YE J, et al. GPT-Driver: learning to drive with GPT [J]. arXiv preprint arXiv:2310.01415, 2023.
- [110] ZHOU Y, MURESANU A I, HAN Z, et al. Large language models are human-level prompt engineers [J]. arXiv preprint arXiv: 2211.01910, 2022.
- [111] LEWIS P, PEREZ E, PIKTUS A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks [J]. *Advances in Neural Information Processing Systems*, 2020, 33: 9459–9474.
- [112] WU Y, ZHU J, XU S, et al. RAGTruth: a hallucination corpus for developing trustworthy retrieval-augmented language models [J]. arXiv preprint arXiv:2401.00396, 2023.
- [113] YUAN J, SUN S, OMEIZA D, et al. RAG-Driver: generalisable driving explanations with retrieval-augmented in-context learning in multi-modal large language model [J]. arXiv preprint arXiv:2402.10828, 2024.
- [114] WEI D, GAO T, JIA Z, et al. BEV-CLIP: multi-modal BEV retrieval methodology for complex scene in autonomous driving [J]. arXiv preprint arXiv:2401.01065, 2024.
- [115] DING W, CAO Y, ZHAO D, et al. RealGen: retrieval augmented generation for controllable traffic scenarios [J]. arXiv preprint arXiv:2312.13303, 2023.
- [116] XIA M, ZHANG X, COUTURIER C, et al. Hybrid retrieval-augmented generation for real-time composition assistance [J]. arXiv preprint arXiv:2308.04215, 2023.
- [117] YAO S, ZHAO J, YU D, et al. ReAct: synergizing reasoning and acting in language models [J]. arXiv preprint arXiv: 2210.03629, 2022.
- [118] FU D, LI X, WEN L, et al. Drive like a human: rethinking autonomous driving with large language models [C]. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024: 910–919.
- [119] CUI C, MA Y, CAO X, et al. Receive, reason, and react: drive as you say with large language models in autonomous vehicles [J]. arXiv preprint arXiv:2310.08034, 2023.