



自动化学报

ACTA AUTOMATICA SINICA

人工智能大模型及单点技术专题序言

赖剑煌

Guest editorial of the special session on artificial intelligence foundation models and specific techniques

Lai Jian-Huang

在线阅读 View online: <https://doi.org/10.16383/j.aas.c202608>

您可能感兴趣的其他文章

人工智能大模型及单点技术专题序言

赖剑煌^{1, 2, 3, 4}

算力、海量多源数据与预训练框架的持续突破,推动了人工智能大模型技术快速发展并进入规模化落地阶段,它凭借通用表征、多模态生成与复杂推理能力正广泛赋能各领域。但是,现有体系仍存在诸多核心瓶颈:静态训练模式易引发灾难性遗忘,难以实现开放环境持续增量学习;全参数微调算力开销巨大,轻量化适配成本难以平衡;图结构与文本语义融合不足、视频生成三维空间一致性差、多模态模型推理强化学习训练低效等问题突出;工业场景下基础模型与任务专用模型协同机制尚不完善。行业亟需融合控制、优化、机器学习交叉理论,攻克持续学习、参数高效微调、可控生成、强化学习优化、图学习与大模型协同建模等单点关键技术,搭建虚实融合验证平台,推动大模型从静态通用生成向可终身学习、低成本适配、多模态可控、行业深度适配的可靠智能体系跨越。

为此,《自动化学报》组织“人工智能大模型及单点技术”专题,促进基础模型理论、底层优化算法、多模态生成、行业落地应用的交叉融合,为科研人员与产业从业者提供展示前沿理论、核心单点技术与工程实践成果的学术平台。本专题经同行严格评议共录用 6 篇论文,涵盖多类代表性成果:大模型持续学习、大模型参数高效微调、文本属性图学习、工业缺陷检测等体系化综述,多模态级联强化学习、3D 可控视频扩散等原创研究,系统呈现了理论演进、算法创新与工业落地的多维进展。

西安交通大学徐宗本院士和孟德宇教授等阐述了面向大模型时代的持续学习方法论演变,重点针对传统的非预训练模型持续学习方法和大模型时代

下逐步演进的预训练模型持续学习方法研究进展进行详细综述,主要从四个层面对比非预训练模型和预训练模型方法的异同点,即数据层面、模型层面、损失/优化层面以及理论层面。着重分析从应用非预训练模型的方法发展到应用预训练模型的方法的技术变化,并分析出现此类差异的内在本质,总结并展望未来持续学习发展的趋势。

北京大学林宙辰教授和唐岸达阐述了大模型参数高效微调方法的进展。针对大模型全参数微调算力、存储开销过高的难题,参数高效微调 (PEFT) 成为核心解决方案。该综述将现有 PEFT 划分为添加、局部、重参数化、融合四大范式,剖析各类方法机理与适用场景;在统一实验条件下定量对比各方案性能与参数效率,梳理其从单一算法创新走向算力、存储、性能综合权衡的发展规律,同时介绍 PEFT 在语言、视觉、多模态领域的落地应用,最后总结现存问题并指明轻量化微调技术后续研究方向。

中国人民大学魏哲巍教授和郑艳萍等阐述了大模型驱动文本属性图学习研究进展。传统图神经网络仅依靠浅层文本嵌入表征图节点,无法充分挖掘文本属性图内丰富语义、上下文关联,图结构与文本信息融合存在明显短板。依托大模型强大语义理解、知识存储与逻辑推理能力,学界形成大模型驱动文本属性图学习全新研究范式。该综述围绕图生成、图预测、图推理三大核心任务,系统梳理主流技术路线、通用数据集与评测基准,明确各任务现存痛点:图生成存在结构与文本表达低效问题,图预测缺少成熟生成方案、评估体系不完善,图推理难以完成复杂多跳关系建模。该文为后续研究提供系统化参考与整体性视角。

哈尔滨工业大学(深圳)张民教授和苏州大学杨天乐等面向视觉工业缺陷检测给出从非基础模型到基础模型的演进与协同综述。精密电子产品的视觉缺陷检测是智能制造领域关键技术。基础模型在少样本检测场景泛化优势突出,但单纯使用基础模型难以实现最优效果。现有高性能方案多融合基础模型通用表征与非基础模型面向任务的成熟学习策略。该综述系统梳理两类检测方法,多维度对比优劣,拆解非基础模型算法并将其迁移赋能基础模型,

1. 中山大学 广州 510275 2. 琶洲实验室(黄埔) 广州 510006 3. 广东省信息安全技术重点实验室 广州 510006 4. 机器智能与先进计算教育部重点实验室 广州 510006

1. Sun Yat-sen University, Guangzhou 510275 2. Pazhou Laboratory (Huangpu), Guangzhou 510006 3. Guangdong Province Key Laboratory of Information Security Technology, Guangzhou 510006 4. Key Laboratory of Machine Intelligence and Advanced Computing, Ministry of Education, Guangzhou 510006

引用格式 赖剑煌. 人工智能大模型及单点技术专题序言. 自动化学报, 2026, 52(8): 1-2

Citation Lai Jian-Huang. Guest editorial of the special session on artificial intelligence foundation models and specific techniques. *Acta Automatica Sinica*, 2026, 52(8): 1-2

DOI 10.16383/j.aas.c202608 **CSTR** 32138.14.j.aas.c202608

搭建协同适配矩阵, 分析该范式落地难点. 分析表明, 基础模型通用先验与非基础模型任务优化思路具备显著协同价值, 为工业视觉缺陷检测后续研究提供可行发展路径.

上海人工智能实验室的乔宇教授和王玮赟等提出了高效提升多模态大语言模型推理能力的级联强化学习策略. 强化学习可增强多模态大语言模型推理能力, 但在线强化学习成本高、收敛慢, 离线强化学习性能存在瓶颈. 该文融合二者优势, 提出级联强化学习策略, 分离线、在线两阶段训练: 离线阶段加速收敛、稳定训练; 在线阶段精细调优以突破性能上限. 实验表明, 该策略仅需一半训练成本即可优于纯在线方案, 在两款 InternVL 系列模型的七项多模态推理评测中, 平均准确率分别提升 6.7%、6.5%, 验证了方法高效、易扩展的优势.

清华大学朱军教授和朱泓舟提出了 3D 空间先验驱动的相机轨迹可控视频扩散模型方法. 现有相机可控视频扩散模型生成内容容易出现物体畸变、三维空间结构不一致问题, 降低视频真实感. 该文在训练与推理全流程引入 3D 空间先验优化模型: 训练阶段采用 Warp-Injection 嵌入视角形变参考帧; 推理阶段先后设计 Warp-Init 噪声校正、Warp-Guidance 能量约束策略, 逐层约束三维几何一致性. 基于 RealEstate10K 数据集开展验证, 该方法

相较现有最优模型 FVD 指标提升 18.03, 3D 结构估计失败率降至 5.2%.

上述论文大都得到国家自然科学基金、国家重点研发计划、科技创新 2030——“新一代人工智能”重大项目、鹏城实验室重大项目等资助, 其研究成果丰富了人工智能大模型及单点技术的理论和技术体系. 受时间和篇幅所限, 本专题还不能全面反映该领域的主要研究方向及最新进展, 但必将对人工智能大模型技术相关领域的研究工作起到“抛砖引玉”的推动作用. 感谢各位作者的积极投稿、认真撰写和不厌其烦地细致修改打磨, 感谢各位评审专家耐心、细致的评审工作. 同时, 感谢《自动化学报》编委会对“人工智能大模型及单点技术”专题工作的具体指导和热心帮助, 感谢《自动化学报》编辑部对征稿及出版工作的辛苦付出.



赖剑煌 中山大学计算机学院教授. 主要研究方向为计算机视觉与模式识别. E-mail: stsljh@mail.sysu.edu.cn (**LAI Jian-Huang** Professor at the School of Computer Science and Engineering, Sun Yat-sen University. His research interest covers computer vision and pattern recognition.)