DOI: 10.7544/issn1000-1239.ps20240901
61(9): 2127−2127, 2024

真实场景中的物体识别与环境理解：开放世界的视觉问题

查红彬
（北京大学　北京　100871）

机器视觉研究的主要目的是提出能够在真实场景中开展有效工作的物体识别与环境理解方法与系统.近年来,随着深度学习与大模型技术的广泛使用,机器视觉已经在结构化静态物体与环境的识别等领域取得了巨大的进步,为信息技术产业的发展提供了坚实的技术基础.但是,真实场景中经常会出现复杂的环境变化与传感器运动,这些具有高度不确定性的因素又给基于深度学习的视觉系统提出了挑战.开放世界的机器视觉是对这一类问题的模式化抽象,是促使视觉系统从封闭环境到开放世界环境转变的有效途径.中国科学院计算技术研究所的陈熙霖研究员、王瑞平研究员团队对开放世界中物体识别与检测问题的现状与挑战进行了细致的分析,并指出了其未来的发展方向,对机器视觉研究的进一步开展具有重要的意义.

文章的核心技术贡献包括:

1）广义开放世界物体识别与检测任务的定义.文章定义了一个综合考虑5个开放性维度（开集学习、零样本学习、小样本学习、长尾学习、增量学习）的任务设定,以模拟真实世界的复杂性和多样性.这种定义超越了当前研究中普遍设定的单一维度评估,为开放世界物体识别与检测提供了新的综合性评测视角.

2）核心难度指标的构建.文章基于5个开放性维度提出了相应的核心难度指标,能够量化任务的难度,并根据不同指标取值生成具有不同难度级别的数据集,有效地评估和比较不同方法的性能.

3）仿真环境数据集的构建.文章利用AI2-THOR仿真平台,构建了一个多样化的大规模数据集,该数据集通过控制场景中的物体位置、姿态、材料和相机的拍摄视角等,实现了域随机化,增强了数据集的多样性和真实性.

4）现实世界挑战的模拟.文章使用仿真环境生成指定类型数据的方式,克服了传统数据收集方法的局限,模拟了现实世界中物体的尺度变化、位姿多样性、多物体共生关系、遮挡以及背景干扰等复杂视觉场景,为开放世界物体检测的技术发展和性能评估提供了有效的实验验证条件.

5）实验结果的深入分析.文章不仅在仿真数据集上验证了所提出指标的有效性,而且将实验结果扩展到真实的公开物体检测数据集上,验证了相关指标对开放世界物体检测性能的影响.

综上所述,文章通过较为综合系统性的研究,为开放世界物体识别与检测问题提供了新的思路和方法,采用仿真环境构建数据集并提出多维度评估指标,有效模拟和解决真实世界的挑战.文章的工作为该领域研究人员提供了新的评测方法和基准数据集,对于促进后续理论方法的研究具有积极的推动作用.

评述专家



查红彬,教授,博士生导师.主要研究方向为智能科学技术、计算机视觉.

亮点论文

聂晖,王瑞平,陈熙霖.开放世界物体识别与检测系统：现状、挑战与展望[J].计算机研究与发展,2024,61(9):2128−2141. DOI: 10.7544/issn1000-1239.202440054