

中图法分类号: TP309; TP399 文献标识码: A 文章编号: 1006-8961(2024)05-1117-02

论文引用格式: 王井东, 桑农, 程明明, 张兆翔, 李玺, 高常鑫, 余昌黔, 陈秀妍. 2024. 《中国图象图形学报》图像/视频语义分割专栏简介. 中国图象图形学报, 29(05):1117-1118 [DOI:10.11834/jig.2400005]

## 《中国图象图形学报》 图像/视频语义分割专栏简介

王井东<sup>1</sup>, 桑农<sup>2</sup>, 程明明<sup>3</sup>, 张兆翔<sup>4</sup>, 李玺<sup>5</sup>, 高常鑫<sup>2</sup>, 余昌黔<sup>6</sup>, 陈秀妍<sup>7\*</sup>

1. 百度, 北京 100085; 2. 华中科技大学人工智能与自动化学院, 武汉 430074; 3. 南开大学计算机学院, 天津 300350;
4. 中国科学院自动化研究所, 北京 100190; 5. 浙江大学计算机科学与技术学院, 杭州 310013; 6. 昆仑万维, 北京 100005;
7. 《中国图象图形学报》编辑部, 北京 100190

语义分割是计算机视觉领域的基础问题, 对分析和理解图像/视频场景数据有着重要的作用和意义, 可广泛应用于安防监控、自动驾驶、工业质检、辅助医疗、遥感分析、图片视频编辑等实际应用。同时, 图像/视频语义分割在智慧城市、智慧教育、智慧医疗及智慧制造等领域起到关键作用, 符合国家“四个面向”重大需求。2017年国务院印发的《新一代人工智能发展规划》一文中也提到要重点突破“复杂动态场景感知与理解”等技术。

为了更好地报道图像/视频语义分割的发展, 及时发布我国学者在相关领域的重要学术前沿技术进展, 促进学术交流和技术创新, 《中国图象图形学报》邀请业内专家共同策划推出“图像/视频语义分割”专栏, 主要收录国内学者在相关领域具有创新性、突破性的研究成果。以期为相关领域的研究人员提供参考。

经过严格评审, “图像/视频语义分割”专栏共收录10篇论文, 包括: 4篇综述和6篇研究论文。

4篇综述中, 《深度学习实时语义分割综述》(作者: 高常鑫, 徐正泽, 吴东岳, 余昌黔, 桑农\*)对深度学习中实时语义分割算法进行了全面的论述和分析: 介绍了语义分割和实时语义分割任务的基本概念、应用场景和面临的问题; 详细介绍了实时语义分割算法中常用的技术和设计; 全面整理和归纳了现阶段实时语义分割算法, 并分析了各类算法的特点和局限性; 提供了完整的实时语义分割评价体系; 最后

分析了实时语义分割领域仍存在的挑战, 对未来可能的研究方向提出了相应见解。论文提及的算法、数据集和评估指标已汇总至 <https://github.com/xzz777/Awesome-Real-time-Semantic-Segmentation>。

《基于深度学习的弱监督语义分割方法综述》(作者: 项伟康, 周全\*, 崔景程, 莫智懿, 吴晓富, 欧卫华, 王井东, 刘文予)首先, 介绍了弱监督语义分割任务, 以及相关数据集和评估指标。接着, 根据弱标注类型和受关注程度, 从图像级标注、其他弱标注以及大模型辅助这3个方面回顾和讨论了弱监督语义分割的研究进展。最后, 分析了弱监督语义分割领域存在的问题与挑战, 并就其未来研究方向提出建议, 旨在进一步推动弱监督语义分割领域研究的发展。

《车路两端纯视觉鸟瞰图感知研究综述》(作者: 周松燃, 卢焯昊, 励雪巍, 傅本尊, 王井东, 李玺\*)对自动驾驶感知领域中的关键技术——纯视觉BEV感知这一前沿新兴领域的技术脉络与最新发展趋势进行详细综述。从相机部署端和相机视角出发, 将纯视觉BEV感知方法划分为车端单视角感知、车端环视视角感知和路端固定视角感知三大方向, 同时结合其通用处理流程, 针对数据集输入、BEV模型和任务推理输出三大模块展开了综述, 以提供纯视觉BEV感知领域的总体解决方案。最后, 对目前BEV领域实际落地面临的技术挑战和部署挑战进行了总结, 并且从多个角度对BEV感知技术未来的发展方向进行了展望。

\* 通信作者

《深度学习实时语义分割研究进展和挑战》(作者:王卓,瞿绍军\*)对基于深度学习的实时语义分割算法进行系统的归纳总结,特别是最新的基于Transformer和剪枝的方法,全面介绍实时语义分割方法在各领域中的应用。讨论了该领域存在的挑战,并对未来值得研究的方向进行展望,为研究者解决存在的问题提供便利。

6篇研究论文中,《针对遮挡物体的轮廓细化实例分割》(李伟,黄娅\*,张馨渊,韩贵金)以BCNet(bilayer convolutional network)网络为基础,提出一种针对遮挡物体的轮廓细化实例分割算法,预测的物体轮廓更加精细,分割掩膜更加完整。

《多帧时空注意力引导的半监督视频分割》(作者:罗思涵,袁夏\*,梁永顺)为了充分利用多帧视频数据间的信息相关性,提出了一种多帧时空注意力引导的半监督视频分割方法,该方法由视频预处理(即外观特征提取网络和当前帧特征提取网络)和基于Transformer和改进的PAN模块的特征融合两部分构成。

《互补特征交互融合的RGB\_D实时显著目标检测》(作者:叶欣悦,朱磊\*,王文武,付云)为了探求RGB图像和深度图像两种模态数据中的互补信息重要性及更有效的交互方式,在分析了传统卷积网络中修正线性单元(ReLU)选通特性的基础上,设计了一种新的RGB和深度特征互补信息交互机制,并首次应用于RGB\_D显著目标检测中。

《多尺度特征融合与交叉指导的小样本语义分割》(作者:郭婧\*,王飞)提出了一种多尺度特征融合与交叉指导的小样本语义分割算法,主要包括特征提取、特征交互和双分支掩码预测3个模块。在两个主流的小样本语义分割数据集PASCAL5<sup>+</sup>和COCO-20<sup>17</sup>上进行测试,实验结果验证了所提模型的优越性。

《跨层细节感知和分组注意力引导的遥感图像语义分割》(作者:李林娟,贺赞,谢刚\*,张浩雪,柏艳红)提出了一种跨层细节感知和分组注意力引导的高分辨率遥感语义分割网络,旨在提高复杂背景下的多尺度地物语义分割的精度。

《结合潜在扩散模型和U型网络的HIFU治疗目标区域提取》(作者:翟锦涛,王润民,李昂,田峰,龚

瑾儒,钱盛友,邹孝\*)提出了一种结合潜在扩散模型和U型网络的HIFU治疗目标区域提取方法。1)利用Latent Diffusion和自动筛选机制生成高保真度的合成图像,提高了生成图像的质量;2)改进UNet网络结构和损失函数,保留更多细节和上下文信息,提高了分割准确性;3)结合生成网络和分割网络,解决训练数据稀缺问题,进一步提升了分割性能。

我们期待广大读者和科技人员通过“图像/视频语义分割”专栏,能够更深入、更全面地了解该领域的最新方法和应用,吸引更多学者从事相关研究并产生具有国际影响力的优秀成果,为本领域的发展做出新的贡献。

#### 专栏编委会:

王井东,百度计算机视觉首席科学家,研究领域为计算机视觉、多媒体以及机器学习。

桑农,华中科技大学教授,研究领域为模式识别和计算机视觉。

程明明,南开大学杰出教授,研究领域为人工智能、计算机视觉、图像视频大数据分析、计算机图形学。

张兆翔,中国科学院自动化研究所研究员,研究领域为视觉认知计算、类脑学习和面向开放环境的视觉感知与理解。

李玺,浙江大学教授,研究领域为计算机视觉、模式识别和深度学习。

高常鑫,华中科技大学教授,研究领域为计算机视觉和图像/视频理解。

余昌黔,昆仑万维高级算法专家,研究领域为计算机视觉和多模态大模型。

#### 专栏责编:

陈秀妍,编辑,主要研究方向为学术出版和媒体传播等。E-mail:chenxy@aircas.ac.cn