

DOI:10.13382/j.jemi.B2508498

二维分割驱动的电力设施三维分割方法*

张丹丹¹ 宋睿² 任超³ 刘高华³ 李帅^{4,5} 李威⁵

(1. 天津大学福州国际联合学院 天津 300072; 2. 中国电力科学研究院有限公司 北京 100192;

3. 天津大学电气与信息工程学院 天津 300072; 4. 浙江大学机械工程学院 杭州 310030;

5. 国网瑞嘉(天津)智能机器人有限公司 天津 300480)

摘要:面向电力机器人带电作业与设备巡检等任务需求,三维分割技术将复杂电力场景分解为多个独立的语义或实例级别的三维空间区域,为电力机器人提供了精准的空间认知,是保障电力系统自动化运维、提升带电作业安全性与效率的关键场景理解技术。现有方法因高成本的三维数据采集和模型实时处理能力不足,难以满足实际电力场景的需求。因此,提出了首个无需三维数据输入的低成本、快速三维分割框架,仅需以多视角的单目二维光学图像和目标提示词为输入就可实现电力设施的三维分割,有效解决了现有方法的局限性。首先利用 Grounded-SAM2 的强泛化特性以及所设计的颜色标注机制,获取输入图像中电力设施的 RGB 标识化分割掩码;随后,经 Spann3r 重建模型融合几何及颜色特征生成电力场景点云;最终,依据颜色筛选准则获得电力设施候选点云集,并基于 DBSCAN 密度约束聚类方法对候选点云进行处理,实现电力设施的实例区分与噪声抑制,输出精准的三维分割结果。实验结果表明,该方法在真实电力场景中验证可实现 92.6% 以上平均交并比(mIoU)与 93.4% 以上平均准确率(mAcc)的三维分割精度,为电力机器人智能化运维提供了一种无需三维传感器支持的低成本且快速的解决方案。

关键词: 三维分割; 二维分割; 三维重建; 电力设施

中图分类号: TP391.41; TM76

文献标识码: A

国家标准学科分类代码: 520.2050

2D segmentation-driven 3D power facility reconstruction method

Zhang Dandan¹ Song Rui² Ren Chao³ Liu Gaohua³ Li Shuai^{4,5} Li Wei⁵

(1. School of International Joint Institute, Tianjin University, Tianjin 300072, China; 2. China Electric Power Research Institute,

Beijing 100192, China; 3. School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China;

4. School of Mechanical Engineering, Zhejiang University, Hangzhou 310030, China; 5. State Grid Ruijia (Tianjin) Intelligent

Robot Co., Ltd., Tianjin 300480, China)

Abstract: Three-dimensional segmentation decomposes complex power environments into semantic and instance-level regions, enabling precise spatial awareness for power robots in automated and safer live-line operations. However, existing methods face high 3D data acquisition costs and limited real-time processing, hindering practical deployment. To address these challenges, we introduce the first low-cost, rapid 3D segmentation framework using only multi-view monocular images and text prompts, without requiring 3D sensors. Specifically, we leverage Grounded-SAM2's robust generalization and a custom color-annotation scheme to produce RGB segmentation masks for each view. These masks are then fed into the Spann3r reconstruction model, fusing geometry and color to reconstruct a dense scene point cloud. Candidate point sets corresponding to power equipment are extracted via color filtering, and a density-constrained DBSCAN clustering step separates individual instances while suppressing noise. Experimental validation in real-world power scenarios shows that, using only monocular RGB imagery, our method achieves over 92.6% mean intersection over union (mIoU) and over 93.4% mean accuracy (mAcc) for 3D segmentation. It offers the low-cost, and fast solution for the intelligent operation and maintenance of power robots without 3D sensor support.

Keywords: 3D segmentation; 2D segmentation; 3D reconstruction; power facilities

收稿日期: 2025-06-24 Received Date: 2025-06-24

* 基金项目: 国家电网公司总部科技项目(5700-202440335A-2-1-ZX)资助

0 引言

作为人工智能感知和理解三维空间的关键技术手段之一,三维分割一直是学术研究的热点,其目标是通过构建能预测三维场景中物体细粒度标签的算法模型,获得特定目标的类别及在三维空间中占有的区域^[1],从而为下游任务进一步从三维数据中精准提取目标物体的语义信息、几何结构及空间位姿提供支撑。

在电力作业场景中,配网带电作业机器人、电力巡检无人机、变电站巡检机器人等智能体对电力设施区域的准确识别存在迫切需求^[2]。尽管成熟的二维分割技术能够在二维图像中实现电力设施区域的高精度分割,但其因深度信息和空间拓扑表征能力的缺失,无法有效识别电力设施作业目标的三维区域,难以满足电力机器人在复杂电力场景中空间感知的语义理解需求^[3-5]。三维分割技术作为三维空间中的场景理解技术,通过融合点云与多视角图像信息,不仅能实现电力设施的三维像素级语义分割,为场景感知提供更全面的空间上下文信息,还可为视觉语言动作模型(VLA)提供结构化环境表征^[6],支撑其推理与决策能力。

然而,在实际应用中,三维分割技术作为场景理解的底层支撑,其实用性仍面临显著挑战。现有电力机器人中的三维分割模型性能严重受限于高成本的三维数据(点云数据、RGB-D)获取与标注流程,以及复杂的模型训练过程。从数据采集层面看,基于激光雷达测距或深度相机的三维数据采集方案在远距探测、遮挡及反光等工况下^[7],获取高质量点云/RGB-D数据的成本显著增加,大幅削弱了现有三维分割方法^[8-14]的实际可行性;从模型训练层面看,当前面向点云的分割方法^[7,15]依赖真实标签数据的监督训练,而点云数据的标注难度和成本远高于二维图像的像素级标注,且由于点云的稀疏性、不规则性和高维特性,模型训练过程变得更加复杂,对计算资源的需求也显著提高^[16],难以满足电力边缘侧设备对低时延的要求。因此,亟需构建一种不依赖昂贵的三维数据获取和注释过程的三维分割方法,以突破高成本及高时延的技术瓶颈。

将二维分割映射到三维空间为解决上述局限性提供了一条新的研究路径。现有的代表性工作如 SA3D 方法^[12],通过神经辐射场模型获得场景的三维表示,将二维分割掩码逆渲染至三维空间,实现了在三维场景中的物体分割。但该方法仍需依赖预训练三维场景表示,本质上依然需要三维数据的输入,且在动态电力场景下存在渲染算力大、模型更新延迟的问题。为此,基于电力机器人主要依靠二维图像进行环境感知的特点,本文提出了一种从二维分割中直接重建三维分割结果的新方法,

该方法作为首个无需以三维数据为输入的三维分割方法,有望成为三维分割技术领域的一种新范式。本文所提出方法主要依赖于两项关键技术:一方面,当前二维分割技术已具备高度泛化能力,以 SAM(segment anything model)^[17]为代表的高性能模型(此外包括 Grounded-SAM2^[18]等)可精准生成目标物的像素级分割掩码;另一方面,DUST3R^[19]和 Spann3r^[20]等通过双视图或多帧图像直接回归三维点云的新兴三维重建方法能够很好解决被遮挡物体的重建问题,因此其可以进一步实现对二维分割掩码的三维重建表征。

基于上述二维分割泛化能力与新兴三维重建技术的可行性,本文方法的实施步骤为首先,利用具有强泛化能力的二维分割模型 Grounded-SAM2 对多视角的二维电力场景图像进行分割处理,根据目标提示词生成准确的目标电力设施的像素级分割掩码。在生成掩码的过程中,引入本文设计的颜色标注机制,使二维分割掩码具备显著的颜色特征。之后,借助 Spann3r 三维重建过程,将带有颜色信息的二维分割掩码图像重建映射至三维空间,重建出电力场景点云。最后,利用颜色信息作为筛选条件生成电力设施候选点云集,通过基于密度的空间聚类应用噪声(DBSCAN)方法分析被筛选出的候选点位,提取满足邻域关系的目标点并分配同一标签,实现同类目标的区分和噪声点过滤,进而获得目标电力设施在三维空间中的区域范围和几何结构。

本文借助成熟的二维分割技术和直接回归三维点云的快速重建技术,不仅能够消除现有三维分割方法对高质量三维数据的依赖,还有效解决了此前 SA3D 中神经场渲染重建所需高计算资源的问题,提高了三维分割的效率和准确性。实验表明,与现有的先进方法 SA3D 相比,本文所提出的方法能直接从现场所获取的多视角二维光学图像中推导出所需的三维分割结果,并展现出更精细准确的三维分割效果。此外,本文方法还具备重建与分割速度快、便于优化更新的优势,可以更直接的满足实际应用需求,节省了传感器费用成本、算力资源成本和三维模型构建时间成本。值得一提的是,本方法依靠 Grounded-SAM2 的强大学习能力,仅需少量二维电力设施图像数据即可完成模型的微调,使其更容易推广到实际场景中。

本文提出了首个无需三维数据输入的三维分割新方法,仅通过输入多视角的二维光学图像和目标提示词就可推导出电力设施目标的三维分割结果,有效解决了高成本三维数据依赖问题,为电力机器人提供即时目标分割与空间感知能力。本文首次将强泛化能力的二维分割与增量式的快速三维重建技术相融合,实现了低成本且快速的三维分割过程,显著降低计算开销与延迟。设计了 RGB 标签并构建了基于颜色编码的筛选策略,充分利

用映射中的显著颜色特征获得三维分割结果。构建了实际电力设施场景测试集,实验验证所提方法在真实场景中的准确性与鲁棒性,为三维分割技术落地提供新方法。

1 相关工作

1.1 三维分割工作

三维分割任务的目标是划定空间中特定物体所占据区域。以深度学习强大的拟合能力为基础,现有的三维分割方法如图1所示,主要以点云等三维数据作为网络的输入,例如 Pointnet 系列^[9-10]、SAM3D^[11]、SA3D^[12]等。除此之外,基于 RGB-D 图像^[21]和基于点云变换^[22-26]的

三维分割方法也有相应研究,但这些方法对优质完整的三维数据都具有较强的依赖。另外,在对这些模型进行训练时,需要为其提供大量带有真实标签的三维数据,涉及非常高的标注成本。同时,因为三维分割模型通常比二维分割模型更为复杂,所以在实际应用中,边缘设备的计算资源有限性会进一步制约这些算法的性能,难以满足实时处理的需求。因此,如何在保证分割精度的前提下,降低对数据质量和计算资源的依赖,一直是三维分割领域的重要研究方向。本文提出了直接通过高精度的二维分割结果重建出目标三维模型的三维分割方法,以避免对完整三维点云输入和三维样本全监督训练的依赖。

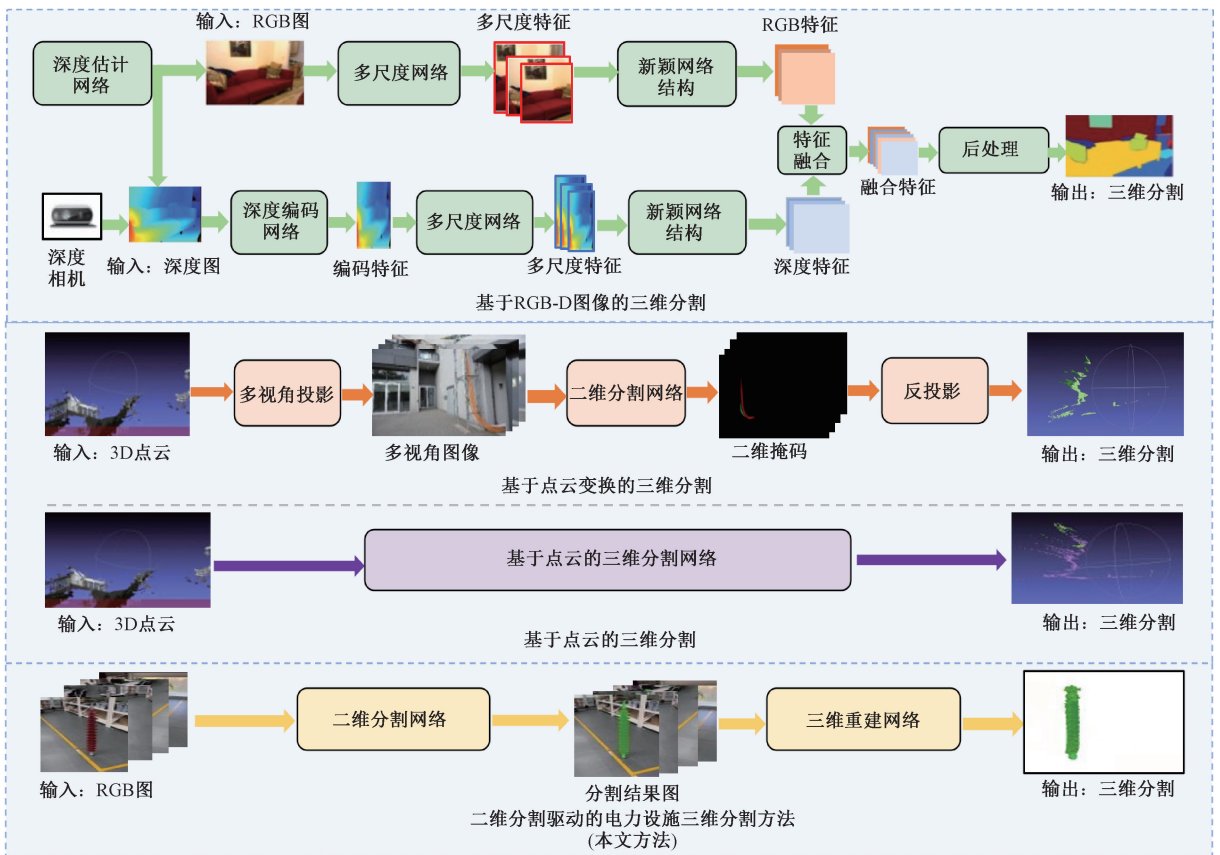


图1 三维分割方法

Fig. 1 3D segmentation methods

1.2 二维分割工作

交互式分割模型 SAM^[17]及其扩展方案凭借在开放场景中卓越的泛化表现和显著的分割灵活性,已广泛应用于二维场景理解。Grounded-SAM2^[18]整合多模态提示与开放集检测技术,支持复杂场景下的任意分割任务。同时,轻量化改进如 EfficientSAM^[27]与鲁棒性增强方案如 SAMURAI^[28]平衡了分割性能与效率。此外, DINO-X^[29]提出开放世界检测框架,探索无监督语义理解范式,

通过自主学习对象语义表征,显著提升模型零样本泛化能力,推动视觉系统向无监督式通用场景理解演进。基于 YOLO 架构的改进方法(如 YOLOv11^[30])通过轻量化设计,在保留实时性的同时提升了密集目标的分割精度。二维分割技术的迭代升级将加速开发高鲁棒性、资源高效的模型架构,为本文三维分割方法提供更优化的底层支撑。

本文基于对二维分割精度与灵活性的需求,将

Grounded-SAM2 作为二维分割工具。其出色的多元提示分割能力,为所提出的三维分割方法提供了可靠的二维分割标注,并使本方法能够更灵活地完成复杂场景下的目标分割任务。

1.3 三维重建工作

三维重建的目标是从二维图像重构高密度三维结构。传统流程涵盖特征检测匹配、运动恢复结构、集束调整和 multiview stereo 4 个阶段,重建结果易受误差积累影响。近年来,基于可微分渲染的 NeRF 系列方法^[31-36]通过深度神经网络学习场景连续体积表征,能在稀疏视图下合成新视角并隐式建模辐射场,但其训练效率与计算复杂度仍为关键瓶颈。3DGS (3D Gaussian Splatting)^[37]创新性结合离散分布与连续优化,显著提升渲染速度与质量,但依赖于相机参数精确初始化。为突破这一限制, DUST3R^[19]提出端到端非参数化重建框架,直接从图像对映射三维空间点图。Spann3r^[20]进一步扩展该框架,引入

空间记忆模块实现实时增量重建,无需逐场景优化。

本文依托 Spann3r 的增量式三维重建过程,创新性地提出了二维分割模型与三维重建的联合框架,利用二维分割生成场景掩码并经过重建过程获得三维分割结果,为多领域应用提供了全新的解决方案。

2 方法设计

本方法是一种基于二维分割结果的创新电力设施三维分割方法,方法框架如图 2 所示。其流程从图像数据及提示词输入开始,首先,通过目标提示词对多视角电力场景二维图像进行分割,并对分割后的区域赋予颜色标注;其次,通过三维重建将带有颜色标注的二维分割结果图升维映射转换到三维空间;最终,经过颜色筛选机制以及密度聚类处理,输出可用于后续电力设施管理和分析的三维分割结果。

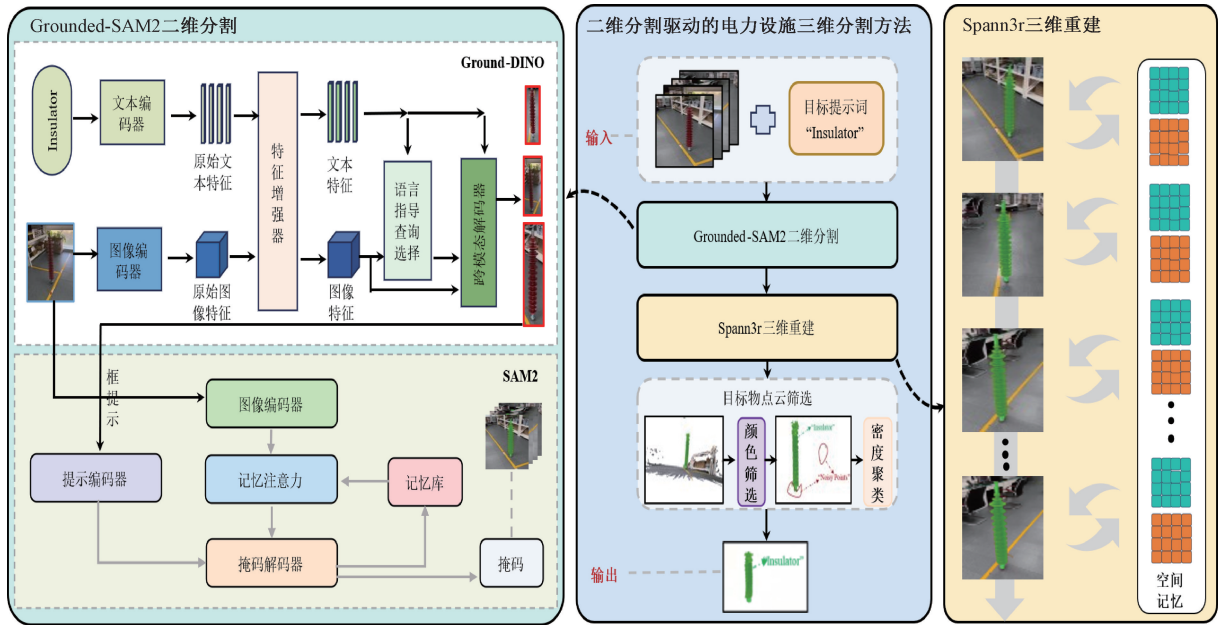


图 2 二维分割驱动的电力设施三维分割方法框架图

Fig. 2 Framework of the 3D segmentation method for power facilities driven by 2D segmentation

2.1 二维图像分割与标注

本文引入 RGB 标签作为分割掩码的标注信息。以下,本文将详细说明基于 Grounded-SAM2 的逐帧预测二维图像分割掩码及 RGB 标签分配方法。

由 Grounding-DINO 和 SAM2 两大模型组成的 Grounded-SAM2 框架,在开放集上具有出色的任务性能,能够实现基于任意文本输入检测和分割图片或视频中相对应的任何区域。受这些先进视觉任务模型的激励,工作的目标是基于 Grounded-SAM2 模型逐帧生成带有显著标注信息的精准二维分割掩码。为此,本文引入具有预

设 (R, G, B) 数值和透明度参数的颜色掩码作为 RGB 标签即为颜色标签,通过固定颜色空间实现目标物的唯一性标注,可直接通过重新定义二维图像分割模型输出的掩码的颜色和透明度来实现。获得带有 RGB 标签标注的二维分割掩码的实现过程为首先,给定多视角的输入视频帧图像和目标提示文本,利用 Grounding-DINO 在第 1 帧图像中生成目标框信息;随后,目标框作为提示信息被输入到即时视觉分割方法 SAM2 中,生成细粒度的掩码注释,完成第 1 帧图像目标区域的精准分割,并基于此分割结果获得后续所有帧图像的分割结果,确保全视频

的分割一致性;最后,将预设的 RGB 标签叠加至每一帧的掩码区域,赋予分割结果可区分的颜色特征,并输出带有 RGB 标注的多视角分割结果图,如图 3 所示。

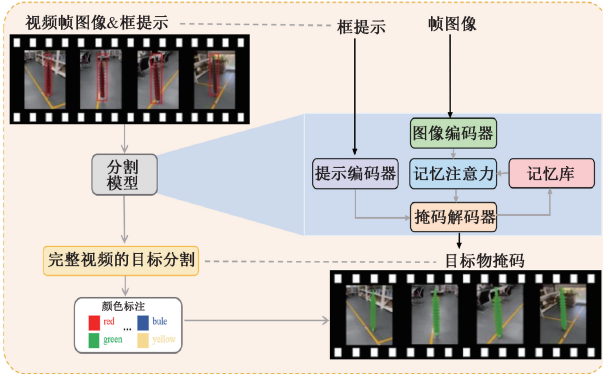


图3 二维分割掩码获取及 RGB 标注过程

Fig. 3 2D segmentation mask acquisition and RGB annotation process

在上述过程中,二维分割结果主要依赖 Grounded-SAM2 中的语言引导机制来识别目标物体,而非颜色信息。Grounded-SAM2 本身具备较强的泛化能力,能够应对目标物体与背景颜色相似的情况。此外,文中引入的颜色标注机制仅附加在已分割的掩码上,用于标注目标丰富目标的特征信息,不会影响模型的识别过程。因此,颜色标注本身不会导致目标误检或漏检。在将带有 RGB 标签的分割结果用于三维重建时,少量背景点会被映射为目标的标注颜色,形成噪声点。但这些噪声点通常在后续 DBSCAN 聚类步骤中可以被有效剔除,不会影响最终的分割效果。

为提升标注的视觉显著性及三维空间中的复用性,优选的 RGB 标签的 (R, G, B) 数值应选取 $(255, 0, 0)$ 、 $(0, 255, 0)$ 、 $(0, 0, 255)$ 等现实中少见且具有高区分度的颜色,可优化后续的目标点云筛选效果。在标注策略上,本文统一同类目标的 RGB 标签颜色,区别不同类型目标的 RGB 标签颜色,这样后续可以直接依据 RGB 标签完成对于目标种类的分。受 RGB 标定的二维像素在三维重建后将把 RGB 信息代入到三维空间,可支撑后续的目标点云的筛选。需要注意的是,在视频处理中,为避免 Grounded-SAM2 因目标追踪丢失或重识别失败而导致同一目标被多个 RGB 标签标记的问题,必须确保每种目标类别对应唯一的 RGB 标签颜色。这种一致性设置不仅可以避免标注混乱,还能进一步提升视频分割的精确度和连贯性。

2.2 基于二维分割的三维重建

Spann3r 基于显示点云学习的重建方式,可以直接从多视角的输入图像重建出三维点云场景,并且能很好解决被遮挡物体的重建问题。在本文所提出的方法中,对

二维分割结果的重建可视作被遮挡区域的重建推导,进而实现二维分割结果向三维空间的映射。以下,本文将详细说明基于 Spann3r 重建二维分割掩码的三维点云的具体过程。

Spann3r 网络由 ViT 特征编码器、Transformer 特征解码器以及空间记忆编码器组成,其中空间记忆模块可以使模型直接在公共坐标系中回归点图。在 Spann3r 模型进行重建的前向传播过程中当前帧 I_t 和前一帧生成的查询特征 f_{t-1}^Q 作为输入,ViT 特征编码器对给定图像序列 $\{I_i\}_{i=1}^N$ 中的一帧 I_t 进行编码输出当前帧的视觉特征 f_t^I :

$$f_t^I = \text{Encoder}^I(I_t) \quad (1)$$

使用查询特征 f_{t-1}^Q 结合空间记忆模块中的记忆键 f^K 和值特征 f^V 检索输出融合记忆之后的特征 f_{t-1}^C :

$$f_{t-1}^C = \text{MR}(f_{t-1}^Q, f^K, f^V) \quad (2)$$

$$\text{MR}(f_{t-1}^Q, f^K, f^V) = \text{Softmax}\left(\frac{f_{t-1}^Q (f^K)^T}{\sqrt{C}}\right) \cdot f^V + f_{t-1}^Q \quad (3)$$

当前帧的视觉特征 f_t^I 和融合记忆后的特征 f_{t-1}^C 被输入两个相互交织的 Transformer 特征解码器,经过交叉注意力推理输出目标解码器解码的特征 f_t^H 和参考解码器解码的特征 f_{t-1}^H :

$$f_t^H, f_{t-1}^H = \text{Decoder}(f_t^I, f_{t-1}^C) \quad (4)$$

特征 f_t^H 被用来生成下一步的查询特征 f_t^Q , 特征 f_{t-1}^H 被用来生成对应的点图 X_{t-1} 和置信图 C_{t-1} 。同时特征 f_{t-1}^H 和预测出的点图 X_{t-1} 被输入到记忆编码器中用来编码生成空间记忆模块里新的记忆键 f_{t-1}^K 和值特征 f_{t-1}^V :

$$f_t^Q = \text{head}_{\text{query}}^{\text{target}}(f_t^H, f_{t-1}^H) \quad (5)$$

$$X_{t-1}, C_{t-1} = \text{head}_{\text{ref}}^{\text{out}}(f_{t-1}^H) \quad (6)$$

$$f_{t-1}^K = \text{head}_{\text{ref}}^{\text{key}}(f_{t-1}^H, f_{t-1}^I) \quad (7)$$

$$f_{t-1}^V = \text{Encoder}^{\text{Memory}}(X_{t-1}) + f_{t-1}^K \quad (8)$$

每一帧图像经过与记忆模块中的特征的交互能够生成全局坐标系下的点图,在整个重建过程中实现增量式重建。Spann3r 可以解决掩码区域推理重建问题,除了受空间记忆模块存储的历史特征信息的引导,更重要的是依靠训练损失函数 L 中置信度损失 L_{conf} 的巧妙设计。网络训练损失由置信度损失和尺度损失组成:

$$L = L_{\text{conf}} + L_{\text{scale}} \quad (9)$$

$$L_{\text{conf}} = \sum_t \sum_{i \in V} C_t^i L_{\text{reg}}(i) - \lambda \log C_t^i \quad (10)$$

$L_{\text{reg}}(i)$ 计算图像中每个有效像素的预测值和真实值之间的欧氏距离作为回归损失, V 为所有有效像素的集合。在 L_{conf} 中将每个像素的置信度定义为 $C_t^i = 1 + \exp(\hat{C}_t^i)$ (\hat{C}_t^i 为网络原始输出的像素置信度)可以确保严格的正置信度。因此,在正则化项的作用下会迫使网络在较困难的区域(置信度低的区域)进行更多的外推以

获得更合理的预测。这也是本方法中能够实现用二维分割获得的掩码直接重建出三维分割结果的根本原因,利用掩码周围区域的像素推导出掩码区域的重建结果。参数 λ 用来控制模型需要分配给每个像素损失的总置信度分数。

尺度损失 $L_{scale} = \text{MAX}(0, \bar{X}, \bar{X}_{gt})$, 其中 \bar{X} 为预测点图中所有点到原点的平均距离, \bar{X}_{gt} 为真实点图中所有点到原点的平均距离。尺度损失能够约束预测尺度小于真实尺度,以防止模型学习到无意义的大尺度点云。

本文方法对整个视频序列应用 Grounded-SAM2 二维图像分割与标注过程后,生成带有标注的二维分割结果图像。将这些图像送入 Spann3r 三维重建模型,为在重建质量与计算效率之间取得平衡,重建过程采用 Spann3r 的演示命令,按照每隔 10 帧选取 1 帧关键帧的策略,对这些关键帧执行显式点云重建,以基于标注的几何纹理信息和标注周围图像的上下文信息还原出被标注区域的三维形态。在上述二维分割结果映射到三维空间的过程中,Spann3r 将分割区域视为遮挡区域,主要依赖重建损失中的置信度项进行推理重建,因此存在一定的不确定性从而引入精度损失。为缓解该问题,本文在分割掩码的颜色标注中设置了透明度系数,以保留更多原始纹理信息,帮助模型更准确地重建分割区域。

2.3 目标物点云筛选

本文方法在重建出带有掩码信息的三维点云场景后,以二维分割结果中的 RGB 标签的 $(R, G, B,)$ 数值为依据,筛选与目标电力设施相关的点云。为了提升算法的适应性和泛化能力,本文依据透明颜色混合规则计算出 $(R_{masked}, G_{masked}, B_{masked})$ 数值筛选范围应对该问题。定义原图像 RGB 数值为 $(R_{img}, G_{img}, B_{img})$, 二维分割结果中的 RGB 标签为 $(R, G, B,)$, 透明度为 α , 则带掩码的目标物的 RGB 数值 $(R_{out}, G_{out}, B_{out})$ 为:

$$R_{out} = \alpha \cdot R + (1 - \alpha) \cdot R_{img} \quad (11)$$

$$G_{out} = \alpha \cdot G + (1 - \alpha) \cdot G_{img} \quad (12)$$

$$B_{out} = \alpha \cdot B + (1 - \alpha) \cdot B_{img} \quad (13)$$

则对于任一带掩码的目标, RGB 筛选范围数值 $(R_{masked}, G_{masked}, B_{masked})$ 为:

$$\alpha \cdot R \leq R_{masked} \leq 255/2 + \alpha \cdot R \quad (14)$$

$$\alpha \cdot G \leq G_{masked} \leq 255/2 + \alpha \cdot G \quad (15)$$

$$\alpha \cdot B \leq B_{masked} \leq 255/2 + \alpha \cdot B \quad (16)$$

本文取 $\alpha = 0.5$, 在二维分割过程中选取了现实中少见且特征鲜明的颜色(如正红色、正绿色、正蓝色等)作为掩码的 RGB 标签,这些颜色在分割和筛选中易于区分,同时有助于提升点云筛选的精确度和稳定性。

通过颜色筛选实现目标物三维空间的初步分割后,还需对同类目标的不同实例进行区分,并滤除 Spann3r 引入的重建噪声。本文采用 DBSCAN^[38] 聚类算法,其相

比需预设簇数的 K-means^[39] 更适配目标数量未知的场景。在对单个目标进行三维分割时, DBSCAN 通过调整邻域半径和最小样本数,剔除低密度噪声,提升分割精度。在处理同类目标中的不同实例时,针对点云数据的分布特点,本文通过调整邻域半径避免相邻实例被误合并;同时,通过增大最小样本数的阈值,有效过滤掉零星噪声点而不损失实例内部连通性。两种参数协同优化后, DBSCAN 能准确地将高密度点云划分为独立簇,显著降低实例误合并与误分割的风险。

3 电力设施三维分割实验

3.1 数据集与评估指标

为了定量验证方法的有效性和实用性,本文使用了公开数据集 SPIn-NeRF 和多个真实电力场景数据。SPIn-NeRF 数据集包含了若干个正向视角的场景,对于每个场景, SPIn-NeRF 对一些数据进行手动标注,以便评估交互式三维分割性能。对于真实电力场景数据,本文采集了典型电力场景的视频数据,实验数据涵盖了绝缘子、电力场景消防箱、电力场景消防栓、标识牌等共计 4 类关键目标对象,覆盖了电力设施、室外电力作业环境、室内电力走廊等多种真实环境。每类目标均采集了多个视频片段,并构建了三维点云数据集,用于评估分割性能。每个场景数据集不仅提供了二维视频帧序列,还包含了手动注释的重建点云标签,以支持定量分析。本文采用分割结果的平均类别交并比(mIoU)和平均类别精度(mAcc)作为评价指标评估本文方法在三维分割任务中的性能。

3.2 实验结果

在本文方法中,二维分割结果的质量直接影响三维空间中分割的精度。为此,本文首先对 Grounded-SAM2 在真实电力场景中的分割性能进行了可视化验证。如图 4 所示, Grounded-SAM2 在绝缘子、电力场景消防箱、电力场景消防栓、标识牌场景下均能生成边界清晰、语义准确且与实际物体轮廓高度吻合的二维分割掩码,能够为后续三维重建阶段提供可靠输入。

基于上述 Grounded-SAM2 二维分割质量的定性验证,对本文方法在真实电力场景中的整体性能进行了量化评估。表 1 中展示的本方法在真实电力场景数据上的评估结果均是基于类别级计算,例如,绝缘子 mIoU 和 mAcc 是基于所有绝缘子样本的总体性能计算,未细分到具体型号。实验结果表明,本方法在真实电力场景数据上的评估结果均表现出了有竞争力的性能,充分证明了其在真实电力场景中的鲁棒性和泛化能力。

为进一步验证本方法的先进性,将本方法与现有的



图 4 电力设施二维分割结果可视化

Fig. 4 Visualization of 2D segmentation results for power facilities

先进方法——基于神经辐射场的 SA3D 进行对比,在 SPIn-NeRF 数据集中选择测试具有较全面视角覆盖的 4 个场景 (Flower、Fortress、Orchids、Room),以保证稳定地重建高质量点云并进行准确分割,突出在相同条件下的性能对比。实验结果如表 2 所示,处理 4 个相同场景中的三维分割任务时,本方法 mAcc 与 SA3D 方法相近而 mIoU 提升了 11.8%,尤其在 Orchids 场景中,mIoU 提升了 27.8%,表明本方法在分割精度上相较于 SA3D 更具有优势。SPIn-NeRF 数据集的其他场景本文也进行了实验,但由于场景可用视角图数量较少,导致点云稀疏度较高、重建质量波动显著。因此,为避免因结果波动过大而影响对比结论,表 2 中未列出这些波动较大的场景。

表 1 本方法在真实电力场景数据上的评估结果

Table 1 Performance evaluation results of the proposed method on real-world power scenes datasets (%)

场景	mIoU	mAcc	场景	mIoU	mAcc
Fair_Box	96.3	96.4	Hydrant	95.5	96.2
Insulator	97.3	97.8	Sign	92.6	93.4

表 2 本方法与 SA3D 在 SPIn-NeRF 数据的评估结果

Table 2 Performance comparison between the proposed method and SA3D on the SPIn-NeRF datasets (%)

场景	SA3D		本文	
	mIoU	mAcc	mIoU	mAcc
Flower	94.6	99.7	99.4	99.5
Fortress	97.1	99.6	99.7	99.8
Orchids	70.4	93.3	98.2	98.8
Room	87.5	98.3	99.6	99.7
平均值	87.4	97.7	99.2	99.5

此外,为分析方法在真实场景中的实际应用效率,本文进一步在真实电力场景数据上进行了推理效率评估,结果如表 3 所示。虽在表 1 中的 4 类目标本文均完成了

推理实验,但由于 SA3D 方法在“电力场景消防箱”和“标识牌”两个场景中重建失败,无法生成可用的三维模型及分割结果,故表 3 仅保留了 SA3D 成功推理的“绝缘子”和“电力场景消防栓”两个场景以保证可比性。实验结果表明,本文方法在绝缘子与消防栓两个场景中的推理速度相较 SA3D 分别提升了 9.67 倍和 28.48 倍,凸显了其在真实电力场景中的效率优势。

表 3 本方法与 SA3D 在真实电力场景数据上的运行时间

Table 3 Runtime performance of the proposed method and SA3D across real-world power scenes (s)

场景	SA3D	本文
Insulator	586.74	60.65
Hydrant	545.22	19.14

综合评估实验结果,本文所提出的方法在处理不同场景下的三维分割任务时,不仅具有较高的分割精度和较短的运行时间,还表现出优异的泛化能力,能够有效适应多样化的环境。充分验证了本方法在场景理解和三维建模任务中的应用潜力,并为实际场景中的自动化处理提供了有效的解决方法。

3.3 三维分割实验可视化结果

所提方法在电力场景中对电力设施的三维分割效果如图 5 所示。图 5 中各场景的绝缘子所用标注颜色虽不尽相同,但仅用于丰富可视化效果,颜色本身不代表不同型号。结果显示,本文方法能够精准分割三维点云,既完整保留了目标的几何结构,又能有效抑制复杂背景干扰,从而充分体现了其在真实场景下的适用性。

为进一步突出所提出方法在真实场景下的优势,SA3D 方法在相同电力场景中的三维分割效果对比如图 6 所示。使用与本方法完全一致的图像数量及视角,并参考 SA3D 文献[12]所述,通过 COLMAP 估计的相机位

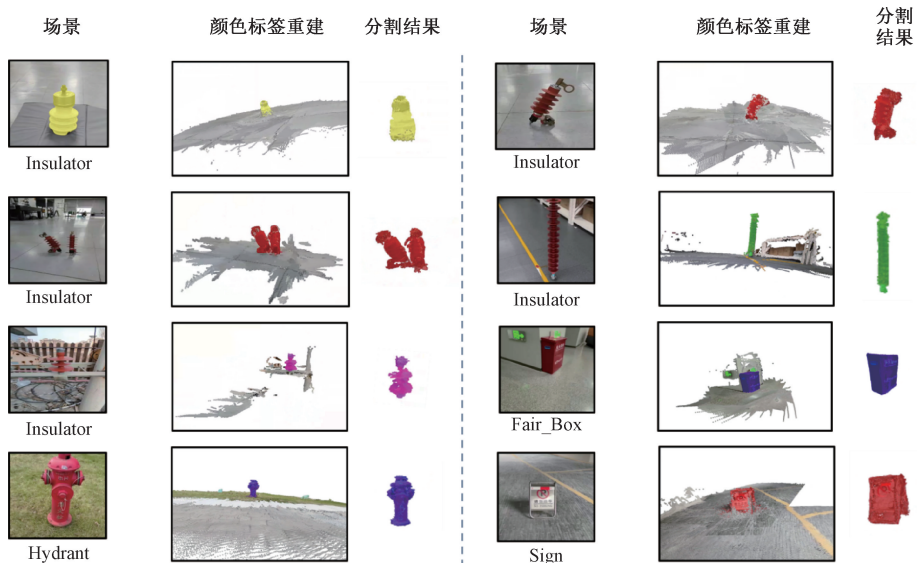


图 5 电力设施分割结果可视化

Fig. 5 Power facility segmentation results visualized

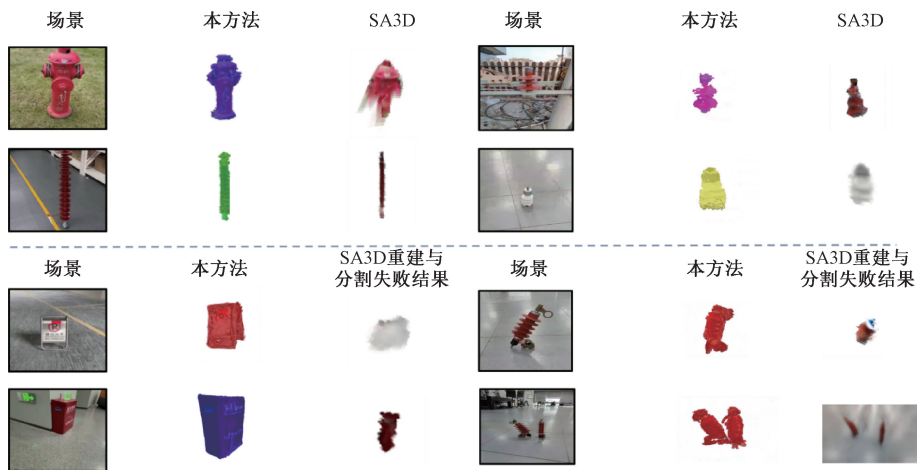


图 6 SA3D 与本方法在真实电力场景中的分割结果可视化对比

Fig. 6 Visualization comparison of segmentation results between SA3D and our method in real power scenes

姿进行 SA3D 方法的真实场景实验并对结果进行可视化。实验结果表明,在电力场景消防栓和部分绝缘子场景虽然可获得三维分割结果,但分割边缘明显模糊;而在其他场景中,SA3D 出现了重建或分割不成立的情况,无法输出有效的三维模型。相比之下,本文方法则能完整保留目标物的三维形态,而且边界清晰且无明显缺失,这也更全面地验证了本文方法在真实场景下的稳定性和鲁棒性。

在完成与 SA3D 方法的对比分析后,本文评估了方法在实例分割任务中的表现。研究主要关注方法流程与关键技术路线可行性的验证,因此实验聚焦于一般场景下的实例级三维分割。在处理实例分割任务时,对于同一类别内属性差异较大的目标,可以通过灵活的语言提

示实现属性级分割,当属性差异较小时,则依靠聚类过程实现实例分割。如图 7 所示,本文方法在实例分割任务中表现出良好的实例区分能力,能够清晰、准确地区分同类别不同实例的目标,验证了其在电力场景下的实例级三维分割能力。



图 7 实例分割结果

Fig. 7 Instance segmentation results obtained

本文通过在多个真实电力场景下的定量与定性实验验证,展现出其在精度鲁棒性、功能延展性与工程实用性等多方面的综合优势。在准确执行语义与实例分割基础任务的同时,本文方法还具备文本引导分割等扩展功能,其模块化架构能够灵活适配场景差异化需求,进一步验证了其作为三维场景分析工具的巨大潜力。

3.4 消融实验

为了验证本文方法中关键设计的有效性,进行了系统的消融实验,重点分析了 RGB 标签透明度以及不同聚类方式对分割结果的影响。表 4 为 RGB 标签透明度变化对分割性能的影响,表 5 对比了不同聚类方式下的分割效果。

表 4 RGB 标签透明度 α 消融实验评估结果

Table 4 Evaluation results of the ablation experiments on RGB label transparency α (%)

场景	$\alpha=0.0$		$\alpha=0.5$		$\alpha=1.0$	
	mIoU	mAcc	mIoU	mAcc	mIoU	mAcc
Fair_box	54.8	54.8	96.3	96.4	76.2	76.2
Hydrant	60.1	60.1	95.5	96.2	85.8	86.0
Sign	47.8	52.2	92.6	93.4	59.0	62.6
Insulator	48.9	50.0	97.3	97.8	55.3	56.3

表 5 不同聚类方法消融实验评估结果

Table 5 Evaluation results of the ablation experiments on different clustering methods (%)

场景	K-means		OPTICS		DBSCAN	
	mIoU	mAcc	mIoU	mAcc	mIoU	mAcc
Fair_box	86.0	91.7	56.7	60.0	96.3	96.4
Hydrant	94.4	96.1	46.2	53.9	95.5	96.2
Sign	88.2	93.3	49.7	53.9	92.6	93.4
Insulator	61.3	68.3	53.1	54.4	97.3	97.8

当 RGB 标签透明度设置为 0 时,二维图像中难以标记出待分割目标的区域。在重建过程中仅将原始颜色信息和纹理信息映射到三维空间,没有分割标记信息的映射,导致后续基于 RGB 标签值筛选目标点时,只能将重建过程中映射为特定颜色的噪声点误识别为目标点云,从而无法准确获取目标物的点云信息。当 RGB 标签透明度设置为 1 时,目标物区域的纹理信息被分割标记信息完全覆盖,这增加了重建过程中对目标物推理的难度,导致最终分割结果不完整,反映在评估指标上表现为性能下降。当 RGB 标签透明度设置为 0.5 时,既能保留目标区域的原始纹理信息,又能为其赋予显著的分割标记特征,从而在后续重建和筛选过程中准确保留目标物对应的点云信息,实现最优的分割效果。

聚类过程的目标是对同一类物体的不同实例进行区分,并去除目标物周围的噪声点。实验表明,K-means 方

法需要预先设置聚类簇数,且无法有效过滤噪声点,导致分割性能下降;而 OPTICS 作为 DBSCAN 的改进方法,虽然在一定程度上解决 DBSCAN 中参数设置问题,但其会过度过滤目标物本身的点云,导致分割性能下降。此外,OPTICS 的计算复杂度显著高于 DBSCAN,难以高效处理真实场景中大规模的点云数据。相比之下,DBSCAN 在保证聚类精度的同时,能够有效平衡计算效率和噪声过滤能力,更适合实际场景中的应用需求。

实验结果表明,RGB 标签的合理引入能够实现目标电力设施三维分割精度的提升。不同的聚类策略对分割性能产生显著不同的影响。消融实验充分证明了本方法中各项设计的合理性和有效性。

4 结论

在三维分割领域,现有模型普遍存在数据收集成本高昂、运行灵活性不足以及算力需求过度的问题。本文紧扣当下电力场景中机器人主要依赖二维视觉图像感知环境的作业特性,率先提出了一种基于二维分割与三维重建的三维分割创新方法。该方法突破了传统三维分割对高质量点云数据和长时间推理的依赖,首次实现了从单目 RGB 传感器采集的多视角二维图像直接推导出目标电力设施的三维分割结果。电力机器人等智能体基于本文方法能利用即时采集的二维光学图像完成关键电力设施的识别与三维分割任务。真实电力场景下的实验结果表明,本文方法在三维分割任务中均表现出优异的性能,同时其部署简易、更新便利,实用性和可扩展性的优势,为电力机器人等智能体在实际场景中的应用提供了高效、可靠的解决方案,并为未来具身智能在空间理解方面的发展奠定了重要基础。

鉴于本文是使用二维分割与三维重建实现三维分割的首次探索工作,当前方法仍然存在一些局限性。一方面,本文方法主要依赖 Grounded-SAM2 本身对分割掩码的生成质量,未采用额外的后处理或其他专门的噪声处理或补充掩码完整性的方法,不能保证掩码存在噪声或不完整时获得完整的三维分割结果;另一方面,本文方法当前采用的三维重建方法在推理分割区域存在固有的不确定性,给最终的三维分割结果带来了不同程度的精度损失。此外,固定时间间隔的抽帧策略在动态场景或目标快速运动时难以及时捕捉所有关键信息,视角不足也会引发分割缺失。尽管如此,所提方法在初步验证中已展现出可观效果,证明了思路的可行性。针对复杂空间关系下如多实例遮挡、目标堆叠等情形,未来工作将进一步扩展实验范围,并探索更加鲁棒的参数设置与聚类策略,以提升方法在实际应用中的通用性与稳定性。为了在更具挑战性的场景中进一步提高稳定性和精度,后续

研究还将围绕增强掩码后处理、提升三维重建精度及鲁棒性、设计自适应抽帧机制以及多视角补偿等方向展开。

参考文献

- [1] DU G, WANG K, LIAN S, et al. Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: A review [J]. *Artificial Intelligence Review*, 2021, 54(3):1677-1734.
- [2] HE Y, YU H, LIU X, et al. Deep learning based 3D segmentation: A survey [J]. *ArXiv preprint arXiv: 2103.05423*, 2021.
- [3] 侯北平,李丰余,朱文,等.基于改进 U-Net 的高压电缆绝缘层图像分割研究 [J]. *电子测量与仪器学报*, 2023,37(10):232-243.
- HOU B P, LI F Y, ZHU W, et al. Research on Image Segmentation of High Voltage Cable Insulation Layer Based on Improved U-Net [J]. *Journal of Electronic Measurement and Instrumentation*, 2023, 37 (10): 232-243.
- [4] 张志君,张惊雷,贾鑫.复杂背景条件下的电气设备图像实例分割算法 [J]. *电子测量技术*, 2024, 47(1): 110-117.
- ZHANG ZH J, ZHANG J L, JIA X. Instance segmentation algorithm for electrical equipment images under complex background conditions [J]. *Journal of Electronic Measurement Technology*, 2024, 47 (1): 110-117.
- [5] 刘丹丹,梁爽,季堂煜.基于秩分解和强语义信息融合的电力巡检算法 [J]. *国外电子测量技术*, 2023, 42(2):16-22.
- LIU D D, LIANG SH, JI T Y. Power inspection algorithm based on rank decomposition and strong semantic information fusion [J]. *Foreign Electronic Measurement Technology*, 2023,42(2):16-22.
- [6] HUANG H, CHEN Y, WANG Z, et al. Chat-Scene: Bridging 3D scene and large language models with object identifiers [J]. *ArXiv preprint arXiv: 2312.08168*, 2023.
- [7] GAO B, PAN Y, LI C, et al. Are we hungry for 3D LiDAR data for semantic segmentation? A survey of datasets and methods [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 23 (7): 6063-6081.
- [8] EL GHAZOUALI S, MHIRIT Y, OUKHRID A, et al. FusionVision: A comprehensive approach of 3D object reconstruction and segmentation from RGB-D cameras using YOLO and fast segment anything [J]. *Sensors*, 2024, 24(9):2889.
- [9] QI C R, SU H, MO K, et al. Pointnet: Deep learning on point sets for 3d classification and segmentation [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017:652-660.
- [10] QI C R, YI L, SU H, et al. Pointnet ++: Deep hierarchical feature learning on point sets in a metric space [J]. *Advances in Neural Information Processing Systems*, 2017, 30:5105-5114.
- [11] ZHANG D, LIANG D, YANG H, et al. Sam3d: Zero-shot 3d object detection via segment anything model [J]. *ArXiv preprint arXiv:2306.02245*, 2023.
- [12] CEN J, FANG J, ZHOU Z, et al. Segment Anything in 3D with Radiance Fields [J]. *ArXiv preprint arXiv: 2304.12308*, 2023.
- [13] 邓磊,刘桂华,邓豪,等.三维点云分割的交联聚乙烯电缆接头参数测量 [J]. *电子测量与仪器学报*, 2022, 36(9):197-207.
- DENG L, LIU G H, DENG H, et al. Measurement of cross-linked polyethylene cable joints based on 3D point cloud segmentation [J]. *Journal of Electronic Measurement and Instrumentation*, 2022,36(9):197-207.
- [14] 王立辉,陈勇吉,韩华春,等.基于点云分割与主成分配准的换电机器人位姿估计方法 [J]. *仪器仪表学报*, 2024,45(10):123-132.
- WANG L H, CHEN Y J, HAN H CH, et al. Pose estimation method for battery swapping robots based on point cloud segmentation and principal component alignment [J]. *Chinese Journal of Scientific Instrument*, 2024,45(10):123-132.
- [15] BUI N T, HOANG D H, TRAN M T, et al. Sam3d: Segment anything model in volumetric medical images [J]. *ArXiv preprint arXiv:2309.03493*, 2023.
- [16] KOLODIAZHNYI M, VORONTOVA A, KONUSHIN A, et al. Oneformer3D: One transformer for unified point cloud segmentation [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024:20943-20953.
- [17] KIRILLOV A, MINTUN E, RAVI N, et al. Segment anything [C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023: 4015-4026.
- [18] REN T, LIU S, ZENG A, et al. Grounded SAM: Assembling open-world models for diverse visual tasks [J]. *ArXiv preprint arXiv:2401.14159*, 2024.
- [19] WANG S, LEROY V, CABON Y, et al. Dust3r: Geometric 3d vision made easy [C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024:20697-20709.
- [20] WANG H, AGAPITO L. 3d reconstruction with spatial

- memory[J]. ArXiv preprint arXiv:2408.16061, 2024.
- [21] HAZIRBAS C, MA L, DOMOKOS C, et al. Fusetnet: Incorporating depth into semantic segmentation via fusion-based CNN architecture [C]. Proceedings of the Computer Vision-ACCV 2016:13th Asian Conference on Computer Vision, 2017:213-228.
- [22] PHAM Q H, HUA B S, NGUYEN T, et al. Real-time progressive 3D semantic segmentation for indoor scenes[C]. Proceedings of the 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), 2019:1089-1098.
- [23] ANDO A, GIDARIS S, BURSUC A, et al. Rangevit: Towards vision transformers for 3d semantic segmentation in autonomous driving [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:5240-5250.
- [24] KONG L, LIU Y, CHEN R, et al. Rethinking range view representation for lidar segmentation [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023:228-240.
- [25] WU Z, SONG S, KHOSLA A, et al. 3D shapenets: A deep representation for volumetric shapes [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2015:1912-1920.
- [26] MATURANA D, SCHERER S. Voxnet: A 3d convolutional neural network for real-time object recognition [C]. Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2015:922-928.
- [27] XIONG Y, VARADARAJAN B, WU L, et al. Efficientsam: Leveraged masked image pretraining for efficient segment anything [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024:16111-16121.
- [28] YANG C Y, HUANG H W, CHAI W, et al. Samurai: Adapting segment anything model for zero-shot visual tracking with motion-aware memory [J]. ArXiv preprint arXiv:2411.11922, 2024.
- [29] REN T, CHEN Y, JIANG Q, et al. DINO-X: A unified vision model for open-world object detection and understanding [J]. ArXiv preprint arXiv: 2411.14347, 2024.
- [30] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [J]. ArXiv preprint arXiv:2410.17725, 2024.
- [31] CAI Y, WANG J, YUILLE A, et al. Structure-aware sparse-view x-ray 3d reconstruction [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024:11174-11183.
- [32] LIU X, XUE H, LUO K, et al. GenN2N: Generative NeRF2NeRF Translation [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024:5105-5114.
- [33] BARRON J T, MILDENHALL B, VERBIN D, et al. Mip-nerf 360: Unbounded anti-aliased neural radiance fields [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022:5470-5479.
- [34] JANG W, AGAPITO L. Codenerf: Disentangled neural radiance fields for object categories [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021:12949-12958.
- [35] WANG P, LIU L, LIU Y, et al. Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction [J]. ArXiv preprint arXiv: 2106.10689, 2021.
- [36] BARRON J T, MILDENHALL B, VERBIN D, et al. Zipnerf: Anti-aliased grid-based neural radiance fields [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023:19697-19705.
- [37] KERBL B, KOPANAS G, LEIMKÜHLER T, et al. 3d gaussian splatting for real-time radiance field rendering [J]. ACM Transactions on Graphics, 2023, 42(4):139:1-139:14.
- [38] KHAN K, REHMAN S U, AZIZ K, et al. DBSCAN: Past, present and future [C]. Proceedings of the Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2014), 2014:232-238.
- [39] IKOTUN A M, EZUGWU A E, ABUALIGAH L, et al. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data [J]. Information Sciences, 2023, 622:178-210.

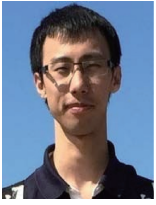
作者简介



张丹丹, 2021年于河北地质大学获得学士学位, 2024年于天津科技大学获得硕士学位, 现为天津大学博士研究生, 主要研究方向为机器人视觉感知、二维目标检测、二维分割、三维分割与三维重建。

E-mail: dandanzhang@tju.edu.cn

Zhang Dandan received her B. Sc. degree from Hebei GEO University in 2021, M. Sc. degree from Tianjin University of Science and Technology in 2024. Now she is a Ph. D. candidate in Tianjin University. Her main research interests include robot vision perception, 2D object detection, 2D image segmentation, 3D segmentation, and 3D reconstruction.



宋睿 (通信作者), 2014 年于华北电力大学获得学士学位, 2015 年于英国爱丁堡大学获得电子学硕士学位, 2019 年于英国爱丁堡大学获得博士学位。现为中国电力科学研究院有限公司高级工程师, 主要研究方向为电力边缘智能和具身人工智能。

E-mail: songrui@epri.sgcc.com.cn

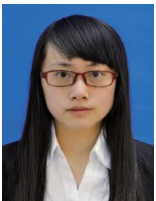
Song Rui (Corresponding author) received his B. Sc. degree from North China Electric Power University in 2014, M. Sc. degree from the University of Edinburgh in 2015, and Ph. D. degree from the University of Edinburgh in 2019. Now he is a senior engineer with China Electric Power Research Institute Company Ltd. His main research interests include power edge intelligence and embodied AI.



任超, 2009 年于山东建筑大学获得学士学位, 2011 年于哈尔滨工业大学获得硕士学位, 并于 2015 年于日本立命馆大学获得博士学位。现任天津大学副教授, 主要研究方向为机器人感知与控制。

E-mail: renchao@tju.edu.cn

Ren Chao received his B. Sc. degree from Shandong Jianzhu University in 2009, M. Sc. degree from Harbin Institute of Technology in 2011, and Ph. D. degree from Ritsumeikan University in 2015. Now he is an associate professor in Tianjin University. His main research interests include robot perception and control.



刘高华, 2010 年于青岛科技大学获得学士学位, 2013 年于天津大学获得硕士学位, 目前在天津大学攻读信息与通信工程博士学位。自 2013 年起, 她一直在天津大学电子信息工程学院工作, 主要研究方向为基于深度学习的行为分析和交互系统开发。

E-mail: suppig@126.com

Liu Gaohua received her B. Sc. degree from Qingdao Science and Technology University in 2010, M. Sc. degree from Tianjin University in 2013. Now she is a Ph. D. candidate in Tianjin University. Since 2013, she has been with the School of Electronics and Information Engineering, Tianjin University. Her main research interests include deep learning-based behavior analysis and interaction system development.



李帅, 2019 年 12 月起担任国网瑞嘉 (天津) 智能机器人有限公司研发总监, 高级工程师。现为浙江大学博士研究生, 主要研究方向为机器人智能决策与控制系统设计。

E-mail: whlls123@163.com

Li Shuai has been serving as the R&D Director of State Grid Ruijia (Tianjin) Intelligent Robot Co., Ltd. since December 2019, and he is also a senior engineer. Now he is a Ph. D. candidate at Zhejiang University. His main research interests include the design of robot intelligent decision-making and control systems.



李威, 2023 年起担任国网瑞嘉 (天津) 智能机器人有限公司研发中心主任, 电力机器人实验室主任。他主持并参与了带电作业人工智能领域多项重点项目, 主要研究方向为机器人智能决策与控制系统设计。

E-mail: liwei1@sgepri.sgcc.com.cn

Li Wei has been serving as the Director of the R&D Center and the Power Robot Laboratory at State Grid Ruijia (Tianjin) Intelligent Robot Co., Ltd. since 2023. He has led and participated in several key projects in the field of live working artificial intelligence. His main research interests include the design of robot intelligent decision-making and control systems.