

DOI:10.13196/j.cims.2023.0729

车身焊装场景下高密度点云数据的半监督语义分割方法

韩松杰, 刘银华⁺, 李彦征, 陈浩

(上海理工大学 机械工程学院, 上海 200093)

摘要:数字化工艺仿真模型的准确性是车身焊装工艺开发的核心,焊装场景下基于点云的工艺装备等分割识别是实现物理作业环境与仿真环境虚实一致性的关键,针对焊装场景点云分割中点云密度不均衡、局部特征差异大、依赖标记样本等问题,提出一种少样本条件下基于生成对抗网络的半监督点云语义分割方法,通过融合使用标记和无标记点云数据来提升分割精度。通过改进 RandLA-Net,采用最远点采样并调整编码解码结构以增强复杂特征学习能力;引入对抗结构与自训练机制,充分利用无标记样本信息;通过引入平滑性约束,选择高可靠性伪标签,降低引入错误标签的概率。最后,在自采的焊装工位点云等数据集上开展对比实验验证了所提方法的优越性。

关键词:焊装工位; 点云语义分割; 半监督; 生成对抗网络

中图分类号:U466;TP391.41;TG44 **文献标识码:**A

Semi-supervised semantic segmentation method of 3D high-density point cloud in welding scene of auto bodies

HAN Songjie, LIU Yinhu⁺, LI Yanzheng, CHEN Hao

(School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The accuracy of digital process simulation model is the core of body welding process development, and the segmentation and identification of process equipment based on point cloud in the welding scene is the key to realize the consistency between physical operation environment and simulation environment. For the problems of uneven point cloud density, large local feature differences and dependence on labeled samples in point cloud segmentation of welding assembly site, a semi-supervised point cloud semantic segmentation method based on generative adversarial network under the condition of few samples was proposed, and the segmentation accuracy by fusing the labeled and unlabeled point cloud data was improved. The RandLA-Net was improved by adopting the farthest point sampling and adjusting the coding and decoding structure to enhance the complex feature learning ability. The adversarial structure and self-training mechanism were introduced to utilize the unlabeled sample information fully. Furthermore, the segmentation performance was improved by introducing smoothness constraints and selecting highly reliable pseudo-labels to reduce the introduction of mislabeled information. The experimental validation was performed on the self-generated welding station point cloud. Results showed that on the welding station dataset, the proposed method achieved a comparable segmentation performance to that of the fully supervised method, and the segmentation accuracy was improved by 8.40% compared with the semi-supervised point cloud segmentation method.

Keywords: welding scene; point cloud semantic segmentation; semi-supervised; generative adversarial network

收稿日期:2023-12-04;修订日期:2024-08-15。Received 04 Dec. 2023; accepted 15 Aug. 2024.

基金项目:国家自然科学基金面上资助项目(51875362);上海市自然科学基金资助项目(21ZR1444500);上海市浦江人才计划资助项目(22PJD048)。**Foundation items:** Project supported by the National Natural Science Foundation, China (No. 51875362), the Natural Science Foundation of Shanghai Municipality, China(No. 21ZR1444500), and the Shanghai Pujiang Program, China(No. 22PJD048).

0 引言

焊装工艺是白车身复杂结构形成的关键工艺,直接影响车辆的结构完整性、安全性和外观质量。车身往往由数百台焊接机器人和工装夹具经过数十个焊装工位形成,焊装过程具有工艺装备繁多、焊装场景复杂、焊接机器人空间狭窄等特点。车身焊装工艺规划常采用数字化工艺仿真方式进行工艺设计与机器人离线编程,因此焊装场景中机器人、工装等工艺装备在虚实场景的几何一致性直接影响焊装工艺设计的有效性。

近年来,随着三维扫描技术的广泛应用,利用焊装车间的扫描点云数据对焊接工位、工艺装备等进行自动分割和配准成为提升数字化工艺模型准确性的主要手段^[1]。该方法的主要思想是通过点云分割自动提取焊装场景中的机器人、工装等工艺设备,进一步采用配准技术进行虚实场景下工艺设备的位姿偏差计算与校准,提高虚实一致性。可见,车身复杂焊装场景中的点云分割效果是提高焊装机器人数字化工艺开发效率与准确性的关键。

焊装场景点云主要有以下特点:①焊装工位点云通过拼接多站激光扫描仪的扫描数据获得,为提高拼接精度,部分区域存在重复扫描的情况,导致不同区域点云密度不同;②不同焊装机器人位姿差异较大,导致相同语义物体的局部特征具有多样性;③焊装工艺装备种类繁多,不同装备之间的特征差异大,难以建立统一标准的分割规则。

由于上述焊装场景点云分割的特殊性,传统分割方法如基于属性数据的方法、基于模型的方法、基于区域的生成方法和基于边缘的方法等^[2-3],难以胜任焊装场景点云分割对泛化性和鲁棒性的需求。近年来,基于深度神经网络(Deep Neural Networks, DNN)的点云分割方法逐渐取代了传统方法,其主要分为投影方法^[4-6]、体素化方法^[7-8]和基于点的方法^[9-13]。基于投影的方法无法捕捉复杂焊装场景的局部结构,而基于体素的方法存在计算成本高、训练时间长的问

题,不适用于大规模高密度点云条件下的焊装场景,基于点的方法克服了以上缺点,直接在三维空间中处理点云数据,其代表工作通过设计有效的特征提取和聚合机制,能够更加精准地表征语义信息特征,从而显著提升分割性能。然而针对焊装场景仍存在两方面问题:①局部特征提取模块感受野不足;②依赖大量高质量人工标注样本。

半监督方法^[14-15]作为有效解决少样本问题的方法之一,通过融合使用标注样本与未标注样本信息,降低了对人工标注样本的依赖。生成对抗网络(Generative Adversarial Network, GAN)^[16]作为其代表方法之一,已成功应用于图像分割、图像生成等领域^[15,17]。ODENA 等^[17]将 GAN 扩展到半监督环境,在显著提高生成样本质量的同时减少了生成器的训练时间。随后的研究^[18-20]为了增强 GAN 的稳定性,引入基于全卷积结构的鉴别器,从未标记的样本中学习,并通过加入多种损失函数产生多种对抗效果,从而优化生成器的性能。虽然上述方法在图像分割领域将 GAN 与半监督训练方法结合显著提升了模型性能,但是针对复杂场景下的点云分割问题仍无成熟的解决方法。

针对非结构化复杂焊装场景下点云分割任务中存在的点云密度不均衡、局部特征差异大等问题,本文综合考虑局部特征提取模块感受野、焊装场景密度等因素,提出一种适应于焊装场景的点云语义分割模型,针对焊装场景点云语义分割模型依赖标记样本的问题,考虑置信区域选择,将自训练机制与 GAN 结合,提出一种基于 GAN 的半监督点云语义分割方法来分割车身制造工艺装备。

1 模型的构建

本文的主要目的是在给定焊装场景点云时自动识别焊装机器人、工装夹具、围栏等工艺装备,为后续工艺装备的虚实一致性校准奠定基础。为此,本文构建了如图 1 所示的 GAN 模型。

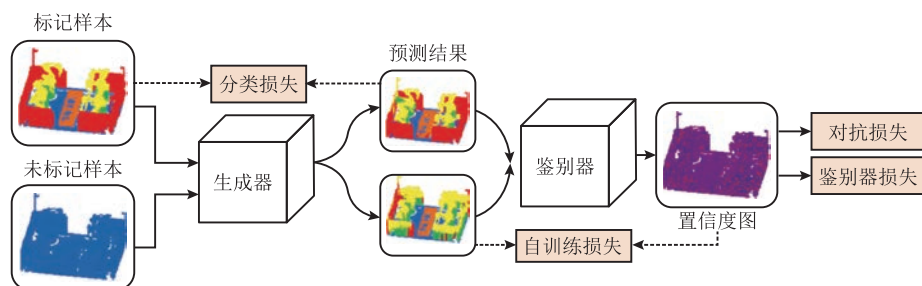


图1 生成对抗网络模型

1.1 生成器模型

为提高机器人、工装等多种类工艺装备的局部特征提取能力, 本文提出基于改进 RandLA-Net^[13] 的生成器模型, 并融合多分类交叉熵损失函数提升模型对不平衡类别的学习能力。其中 RandLA-Net 主要从两个方面进行改进: ①下采样过程采用最远点采样^[11] (Farthest Point Sampling, FPS) 算法来保留重要特征; ②针对焊装机器人相同区域点云局部

特征差异大的特点, 提出深度特征提取编码层, 提高网络对不同区域的表征能力。生成器模块网络框架图如图 2 所示, 其中 N 为点云数目。总体来说, 生成器模块采用编码—解码结构, 通过共享多层感知器 (Multi-Layer Perceptron, MLP) 层实现编码—解码模块的连接, 最后采用 3 个全连接层 (Fully Connected layer, FC) 和一个随机失活层 (Dropout, DP) 实现对每个点的语义信息预测。

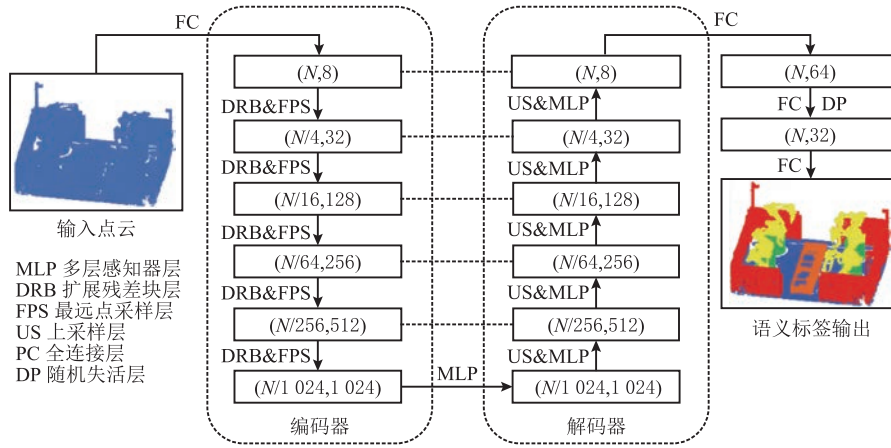


图2 生成器模型框架

车身焊装场景点云数据密度不均衡, 在编码下采样过程中容易丢失点云数量较少的工艺装备特征信息。为此, 在编码层用扩展残差块 (Dilated Residual Block, DRB) 和 FPS 模块, 下采样过程在保留输入点云的整体形状和特征的同时扩大了感受野, 从而更好地捕获局部和全局邻域信息之间的相关性。生成器使用 $L_{Encoder}$ 个编码层对三维点云进行采样, 编码层表达式为

$$Encoder(F) = DRB_{\{L_{Encoder}\}}(FPS_{\{L_{Encoder}\}}(\cdots(DRB_1(FPS_1(F)))))) \quad (1)$$

式中: F 为编码层的输入; $L_{Encoder}$ 为编码层的层数。

在编码阶段, 采用 K 最近邻 (K -Nearest Neighbor, KNN) 算法精确地查找每个点的最临近点, 上采样则采用最临近点插值算法^[11], 保持了点云的空间连续性和几何结构, 提高了模型对不同焊装场景中不同物体和表面的区分能力, 从而提高了分割精度。解码过程中, 点云数量逐层恢复为 N 。设解码层共有 $L_{Decoder}$ 层, 表达式如下:

$$Q = MLP(Encoder(F)); \quad (2)$$

$$Decoder(Q) = MLP_{\{L_{Decoder}\}}(US_{\{L_{Decoder}\}}(\cdots MLP_1(US_1(Q)))) \quad (3)$$

式中 US 为最临近点插值模块。

网络的最后 3 层为全连接层, 旨在将学到的特征映射到样本标签空间。全连接层的后面增加了一个 Dropout 网络层, 通过在训练过程中随机关闭网络中的神经元, 避免网络过深带来的过拟合效应, 并提升模型在新的焊装场景点云数据上的泛化能力。最后网络输出大小为 $N \times n_{class}$ 的所有点的预测语义类别。

1.1.1 局部特征聚合模块

局部特征聚合模块包括局部空间编码 (Local Spatial Encoding, LSE) 和注意力池化 (Attentive Pooling, AP)。LSE 本质上是对每个点及其周围临近点的位置进行编码, 目的是为了捕捉焊装场景点云中每个点及其周围空间的详细几何关系, 步骤为: ①寻找每个点的最临近点; ②对每个点与最临近点的位置信息编码; ③增强点云特征。给定点 P 坐标为 (x, y, z) , 采用 KNN 算法寻找其最临近点, 然后对每个点与临近点之间的位置信息进行编码, 设点 P 的最临近点的集合为 $\{P^1, P^2, \dots, P^k, P^{k+1}\}$, 用以下公式进行位置信息编码:

$$r^k = MLP\{P \oplus P^k \oplus (P - P^k) \oplus |P - P^k|\} \quad (4)$$

式中: P 和 P^k 分别为两点之间的位置信息 (x, y, z) 的坐标; $|P - P^k|$ 为计算两点之间的欧氏距离;

⊕为串联运算符; r^k 为相关点位置信息的编码值。

1.1.2 注意力池化模块

现有研究^[10,12]通常采用最大池化方法,这种方式忽略了相邻点之间的特性,无法有效集成关键特征,从而丢失关键信息。因此,通过引入注意力机制,为每个点及其邻近点的特征分配不同的权重,可以有效提升焊装车间中工艺装备关键特征的提取能力,尤其在面对少样本问题时,该方法能够有效增强模型对稀有类别的识别能力。给定局部特征集合 $\hat{F}_i = \{\hat{f}_i^1 \cdots \hat{f}_i^k \cdots \hat{f}_i^K\}$,AP的计算方式如下:

$$S_i^k = \text{softmax}(\hat{f}_i^k) = \frac{\exp(\hat{f}_i^k)}{\sum_{k=1}^K \exp(\hat{f}_i^k)}; \quad (5)$$

$$\tilde{f}_i = \sum_{i=1}^K (\hat{f}_i^k \odot S_i^k). \quad (6)$$

式中: S_i^k 为点云中第*i*个点的第*k*个邻近点特征的注意力权重; \tilde{f}_i 为聚合了第*i*个点及其*K*个邻近点特征的特征向量; \odot 表示点乘运算。

为了在降采样过程中保留点云的完整性和几何细节,引入DRB模块,如图3所示。该模块由多个LSE模块和AP模块组成,并通过跳跃连接进行特征融合。经过融合操作后,点云的有效感受野显著扩大。在密度焊装场景点云条件下,相比直接增大KNN中的*K*值,DRB模块是一种更加高效且计算成本较低的方法,可以增大感受野并促进邻域点之间的特征传播。

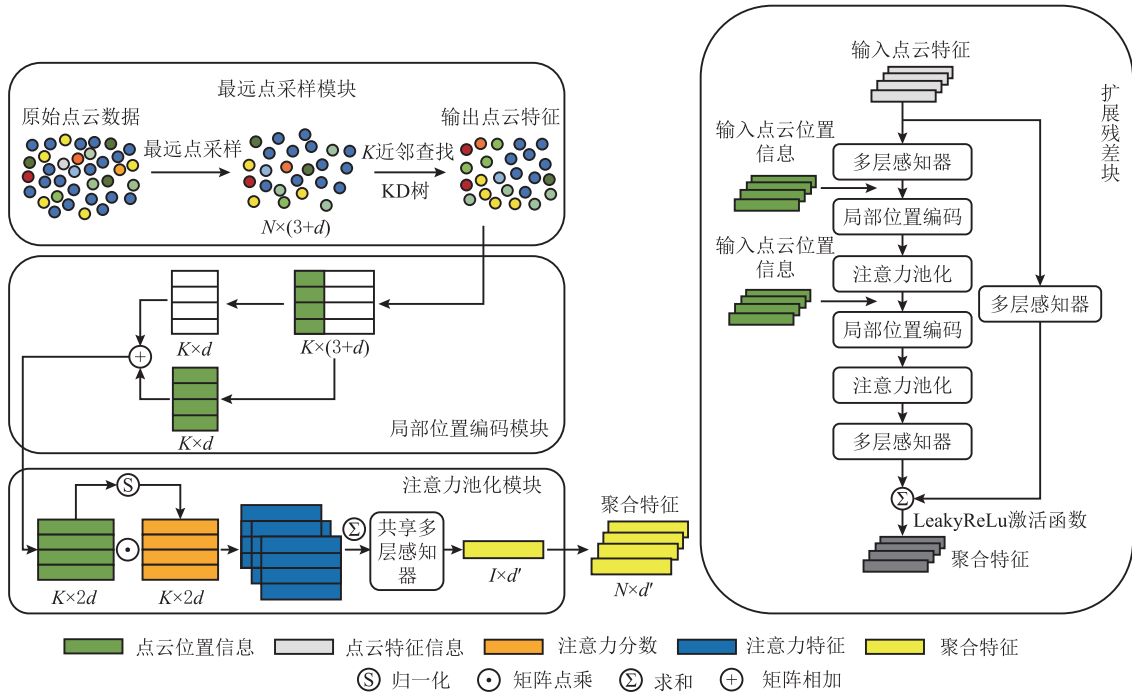


图3 局部特征聚合模块

1.2 鉴别器模型

鉴别器用来判断生成器的车间点云预测结果与真实标注结果的置信度,从而指导生成器更新参数,提高预测的质量和准确性。受深度卷积生成对抗网络^[21](Deep Convolution GAN, DCGAN)的启发,本文采用图4中的全卷积结构,旨在增加鉴别器的学习难度。该全卷积结构由一组一维卷积层组成,每个卷积层采用1x1卷积核,从而在保持特征图尺寸不变的情况下,学习特征之间的交互关系。为了增强鉴别器在对抗训练中区分真实样本和生成样本的能力,本文借鉴类似生成器的通道设计提高了鉴别

器的模型复杂度。另外,为了防止过拟合并增强鉴别器的鲁棒性,在每个卷积层后(除了最后一层)加入Dropout层。在激活函数方面则采用了LeakyReLU函数。鉴别器的输入包括带有真实标签的焊装车间点云和生成器输出的带有预测概率的点云。具体而言,鉴别器的输入为真实标签与生成器预测标签图的组合,构成了一个条件对抗网络。

1.3 模型训练

在GAN中,以标记点云与无标记点云作为输入,预测每个类别的概率。鉴别器*D*得到经过One-Hot的点云的真实标签和生成器的输出预测

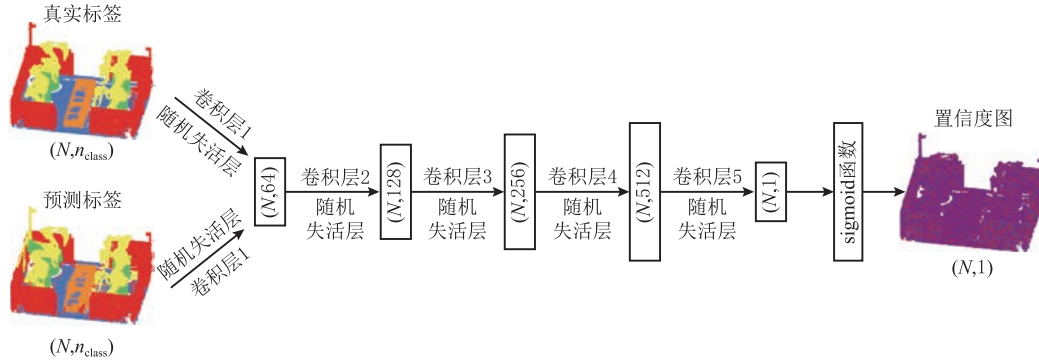


图4 鉴别器模型框架

结果, 任务是将预测的分割结果与真实标签分开。在训练的初始阶段, 采用带有标记的点云对生成器和鉴别器进行训练, 随着时间的推移, 生成器与鉴别器趋于稳定, 再逐渐引入无标记点云进一步训练生成器。训练的具体流程如算法 1 所示。

算法 1 基于 GAN 的半监督方法训练流程。

输入: 标记点云集 L ; 未标记点云集 U ; 标签真值 G ; 迭代数 M ; 生成器 S ; 鉴别器 D ; 生成器优化器 S_{optim} ; 鉴别器优化器 D_{optim} ; 自训练开始迭代次数 T 。

输出: 每个点的语义类别 P 。

```

1: For epoch = 1: M do
2: For l ∈ L do
3: 生成器优化器  $S_{\text{optim}}$  与鉴别器优化器  $D_{\text{optim}}$  梯度置零
4: 标记样本  $l$  输入生成器  $S$  得到预测结果  $l_{\text{pred}}$ 
5: 预测结果  $l_{\text{pred}}$  输入鉴别器得到置信度图  $\text{Map}^{\text{label}}$ 
6: 根据式(9)和式(10)分别计算标记样本的分类损失  $l_{ce}$  与对抗损失  $l_{adv}$ 
7: 取一批次无标记样本  $u \subset U$ , 输入生成器  $S$  得到预测结果  $U_{\text{pred}}$ 
8: 预测结果  $U_{\text{pred}}$  输入鉴别器  $D$  得到置信度图  $\text{Map}^{\text{unlabel}}$ 
9: 根据式(10)计算对抗损失  $U_{adv}$ 
10: if epoch > T then
11: 根据  $\gamma$  对  $\text{Map}^{\text{unlabel}}$  进行筛选, 得到可信预测结果  $\text{Pred}_T$ 
12: 根据式(11)计算自训练损失  $L_{st}$ 
13: else
14:  $L_{st} = 0$ 
15: end if
16: 根据式(7)计算生成器总损失  $L_S$ 
17: 生成器损失  $L_S$  反向传播
18: 根据式(13)得到真实标签独热编码  $G_{\text{hot}}$ 
19: 鉴别器  $D$  对独热编码结果  $G_{\text{hot}}$  进行置信度判断, 得到  $\text{Map}_{\text{gt}}$ 
20: 根据式(10)计算对抗损失  $G_{adv}$ 
21: 根据式(14)计算鉴别器总损失  $L_D$ 
22: 鉴别器损失  $L_D$  反向传播
23: 生成器优化器  $S_{\text{optim}}$  优化, 鉴别器优化器  $D_{\text{optim}}$  优化
24: end for

```

25: end for

1.3.1 生成器训练

给定一个由标记点云 \mathbf{X}^n 及其真实标签 \mathbf{Y}^n 和未标记点云 \mathbf{X}^m 组成的序列集合, 本文用 $S(\cdot)$ 表示生成器的输出, $D(\cdot)$ 表示鉴别器的输出。

作为分割网络, 生成器通过最小化损失函数 L_S 进行训练, 该损失函数由多分类加权交叉熵损失 L_{ce} 、对抗损失 L_{adv} 、自训练损失 L_{st} 组成, 计算公式为

$$L_S = L_{ce} + \lambda_{adv} L_{adv} + \lambda_{st} L_{st}. \quad (7)$$

训练初期, 生成器在预测未标记点云时误差较大, λ_{st} 随着迭代次数逐步增加,

$$\lambda_{st}(t) = \begin{cases} 0, & t < T_1; \\ \frac{t - T_1}{T_2 - T_1} \alpha, & T_1 \leq t \leq T_2; \\ \alpha, & T_2 \leq t. \end{cases} \quad (8)$$

式中: λ_{st} 表示自训练损失的权重系数; t 为当前迭代数; α 为不同数据集下的权重参数; T_1 为自训练的迭代次数; T_2 为再次增加自训练损失权重时的迭代次数。

首先, 针对有标记点云下焊装场景中工装等类别样本数量少的问题, 采用 L_{ce} 反应生成器输出 $S(\mathbf{X}^n)$ 与真值 \mathbf{Y}^n 之间的差异, 以不断优化生成器, 使生成器的输出结果逼近真实标签,

$$L_{ce} = - \sum \omega \cdot \mathbf{Y}^n \log(S(\mathbf{X}^n)). \quad (9)$$

式中 ω 为不同类别的点云权重系数。

其次, 在使用完全卷积鉴别器时, 计算对抗损失 L_{adv} 以完成对抗过程, 使生成器输出的预测分布更接近实际标注, 对抗损失公式为

$$L_{adv} = - \sum \log(D(S(\mathbf{X}))). \quad (10)$$

式中 \mathbf{X} 表示标记点云 \mathbf{X}^n 和未标记点云 \mathbf{X}^m 。

将训练结果趋于稳定的生成器和鉴别器中加入

未标记点云,形成一个自训练框架,其主要思想是:首先,通过已训练的鉴别器 $D(S(\mathbf{X}^m))$ 生成一个置信度判断点云图,该图可用于推断与真实标注分布足够接近的区域;然后,对置信度图进行二值化处理,突出显示值得信任的区域,将经过生成器预测后该区域中各点的概率矩阵记为 $\mathbf{U}=[u_1 u_2 u_3 \cdots u_n]$,选取概率最大的类别作为该点伪标签 Y^* ,其中 n 为预测类别的数量;最后,将带有伪标签的焊装工位点云加入生成器模型计算自训练损失,再次优化生成器模型。自训练损失计算如下:

$$L_{st} = \begin{cases} -\sum w \cdot Y^* \log(S(\mathbf{X}^m)), & D(S(\mathbf{X}^m)) \geq \gamma; \\ 0, & \text{其他。} \end{cases} \quad (11)$$

式中 Y^* 为未标记点云的伪标签,

$$Y^* = \operatorname{argmax}(S(\mathbf{X}^m)). \quad (12)$$

鉴于焊装工位点云的高密度特点,本文引入基于 KD 树(K-Dimensional tree, KD-tree)的 KNN 投票机制作为平滑性约束模块,提高伪标签的质量和模型自训练的准确性,使明显预测错误的点不参与自训练损失计算,以降低错误伪标签对生成器的影响。具体过程是针对每个通过置信度评估的点,在其 KNN 领域内进行频次统计,确定出现次数最多的伪标签类别 Y^k 。若某点的伪标签 Y^* 与 Y^k 相同,则将其视为有效伪标签;反之,将其视为无效伪标签,不参与自训练损失的计算。

1.3.2 鉴别器训练

鉴别器通过最小化损失函数 L_D 进行训练, L_D 的计算公式为

$$\mathbf{Y} = \text{One_hot}(\mathbf{Y}^n); \quad (13)$$

$$L_D = -\sum (1 - R_n) \log(1 - D(S(\mathbf{X}))) + R_n \log(D(\mathbf{Y})). \quad (14)$$

式中: One_hot 表示独热编码; R_n 为鉴别器的期望值,当输入是生成器的输出时 $R_n=0$,当输入是点云的真实标签时 $R_n=1$ 。

2 案例分析

2.1 数据源介绍

通过自采的焊装工位点云数据集 WWIS (welding workshop indoor spaces) 与 S3DIS (stanford large-scale 3D indoor spaces) 公开数据集^[22]对本文方法进行验证。WWIS 数据集由三维激光扫描仪多视角扫描实际焊装工位后进行点云拼接获得,共有

8 个车身焊装工位,下采样后各工位点云数约 400 万,工装点云数约 50 万,具有高密度、样本不均衡等特点;工位平均长 8.5 m、宽 7.5 m、高 4 m,主要由焊装机器人、底座、工装、围栏、地面 5 个语义类别组成,如图 5 所示。WWIS 数据集中随机选择 6 个工位作为训练集,其余作为测试集,训练集与测试集的比例为 3:1。S3DIS 是一个大规模室内场景点云数据集,包括 6 个不同的室内场景,有 13 个语义类别,本文将第 5 个场景 Area5 作为测试集数据,其他区域作为训练集数据。

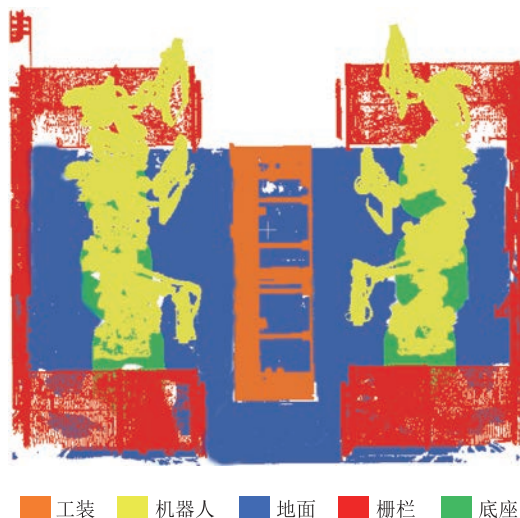


图5 焊装工位示意图

本文实验的硬件环境为 Intel(R) Core(R) i5 12400F @ 4.5 GHz, NVIDIA RTX3060 12 GB, 32 GB RAM, 软件环境为 Python 3.7, Pytorch 1.12.1。

本案例参数通过大量对比实验获得,如对抗损失权重系数 $\lambda_{adv}=0.06$ 时, WISS 分割精度最高。为确定最佳置信度阈值 γ 与自训练损失权重 α , 分别设置 $\gamma=0, 0.06, 0.006, \alpha=0.5, 0.8, 1.0$, 进行多次测试, 结果如表 1 所示。当 $\gamma=0, 0.006$ 时, 较低的置信度阈值引入错误预测分割精度较低; 当 $\gamma=0.06$ 时, 分割精度最高; 当 $\alpha=0.8$ 时, 分割效果最好。综合考虑, 选择 $\gamma=0.06, \alpha=0.8$, 具体训练参数设置如表 2 所示。

表 1 不同训练权重系数下的 WISS 分割精度

置信度 阈值 γ	自训练损失权重 α		
	0.5	0.8	1.0
0	85.45	87.09	86.06
0.006	87.96	90.76	89.84
0.060	91.46	93.74	92.93

表 2 WISS 与 S3DIS 的训练参数

参数设置	WISS	S3DIS
训练迭代次数	150	200
点云数量	4096	4096
学习率(生成器)	2.5×10^{-3}	2.5×10^{-3}
学习率(鉴别器)	1×10^{-3}	1×10^{-3}
优化器(生成器)	Adam	Adam
优化器(鉴别器)	Adam	Adam
批次大小	16	16
对抗损失权重 λ_{adv}	0.06	0.06
置信度阈值 γ	0.06	0.05
自训练损失权重 α	0.8	1.0
T_1	80	120
T_2	120	160

本文从 WWIS 训练集中随机选取 1/4 和 1/8 点云作为标注的训练数据, 其余作为未标注的训练数据, 测试集设置为训练数据的 1/4。另外, 为了评估本文方法在泛化能力的表现, 选择测试集的点云密度与训练集不同, 两种场景下训练集和测试集中

每类点云的数量如表 3 所示。通过与全监督分割方法^[9-13]、半监督学习方法^[14-15, 18, 20]对比验证本文方法, 并采用平均交并比 $MIoU$ 、全局准确率 OA 、 $F1$ 分数作为评价指标来评估本文方法的有效性。 IoU 为各个类别模型预测与真值之间的交集与其并集之比, $MIoU$ 表示 IoU 的平均值; OA 表示所有判断正确的结果占总体的比重, $F1$ 为 $Precision$ 和 $Recall$ 的调和平均数^[13]。各指标的计算公式如下:

$$IoU = \frac{TP}{FN + FP + TP}; \quad (15)$$

$$OA = \frac{TP}{TP + TN + FP + FN}; \quad (16)$$

$$Precision = \frac{TP}{TP + FP}; \quad (17)$$

$$Recall = \frac{TP}{TP + FN}; \quad (18)$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (19)$$

式中: TP, TN, FP, FN 分别表示真阳性、真阴性、假阳性、假阴性。

表 3 两种场景下的各类点云数量

场景	样本设置	工装	机器人	地面	围栏	底座	总和
I	训练点云(标记)(1/4)	525 747	4 882 443	1 208 535	1 414 870	1 248 211	9 279 806
	训练点云(未标记)(1/2)	815 875	8 235 590	3 390 168	3 174 254	2 216 529	17 832 416
	测试点云(1/4)	432 456	4 367 795	1 503 739	1 509 690	1 003 435	8 817 115
II	训练点云(标记)(1/8)	234 081	2 146 223	910 441	870 460	648 037	4 809 242
	训练点云(未标记)(5/8)	1 024 321	9 563 438	4 176 455	3 880 710	2 867 521	21 512 445
	测试点云(1/4)	432 456	4 367 795	1 503 739	1 509 690	1 003 435	8 817 115

2.2 结果分析

图 6 所示为 1/8 标记训练点云数量下不同全监督方法的分割性能。现有基于 DNN 的点云语义分割方法需要在全监督模式下训练, 仅标注过数据才能被使用; 全监督方法 PointNet++^[10], DGCNN^[12] 出现部分无法正确分割的情况; 本文方法不仅可以利用标签数据进行训练, 还可以充分利用未带标签数据, 实现了低标记样本下的高正确率分割性能。

将本文提出的生成器作为基准方法, 目前先进的全监督分割模型^[9-13]、本文的半监督方法与其他半监督方法^[14-15, 18, 20]在焊装工位点云场景上的分割性能如表 4 所示。

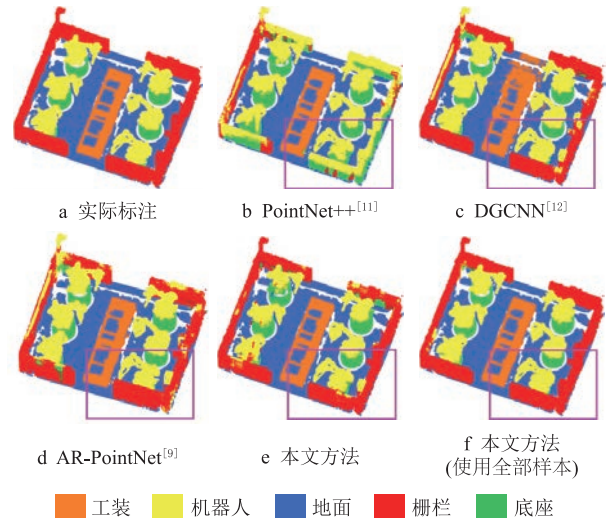


图6 不同全监督方法在1/8数据集上训练后测试的可视化结果

表 4 不同方法在 WWIS 数据集上的分割效果对比

标注比例	模型	MIoU	类别						
			工装	机器人	地面	围栏	底座	OA	F1 平均值
全	PointNet ^[10]	74.78	98.00	68.73	98.48	61.45	47.23	81.50	78.34
	PointNet++ ^[11]	81.00	74.09	94.40	70.18	82.49	83.84	91.43	85.23
	DGCNN ^[12]	92.11	96.10	95.44	90.89	91.28	86.86	96.09	96.76
	RandLA-Net ^[13]	88.10	94.49	91.92	98.19	79.88	76.03	93.73	93.04
	AR-PointNet ^[9]	93.11	97.30	90.90	98.30	83.20	95.90	95.84	96.54
	基准方法	93.57	97.66	92.67	98.52	87.86	91.14	96.34	98.00
半 1/4	Hung et al ^[20]	86.45	93.74	86.09	96.68	74.73	81.04	92.27	88.18
	S4GAN ^[18]	85.25	97.13	85.02	98.41	60.00	85.68	91.33	87.50
	Li et al ^[15]	90.06	91.52	93.49	98.78	83.39	81.29	83.13	94.46
	SSPC-Net ^[14]	92.83	96.35	93.57	96.71	86.80	90.75	96.07	97.27
	本文提出方法	93.98	96.33	94.68	97.40	91.34	90.14	96.90	98.03
半 1/8	Hung et al ^[20]	80.59	88.37	82.34	90.92	72.21	69.14	88.90	83.80
	S4GAN ^[18]	82.93	96.26	78.29	96.52	79.76	63.80	89.10	85.35
	Li et al ^[15]	86.47	94.12	87.71	96.28	79.58	74.66	92.50	92.26
	SSPC-Net ^[14]	91.66	91.57	95.16	94.07	88.15	89.35	96.00	96.20
	本文提出方法	93.74	95.80	94.44	98.11	91.34	89.04	96.86	97.59

注：“全”表示模型训练使用标注完整的训练集，“半”表示模型训练使用一定比例标注的训练集与其余未标注训练集。

由表 4 可见,本文方法的 MIoU 比其他通用半监督分割方法^[18,20]有显著提升。在 1/8 标注数据下,本文方法比当前先进的基于 GAN 的点云半监督分割方法^[15]提升 8.40%,比基于超点标签传递策略的 SSPC-Net^[14]提升 2.27%。对比本文提出的半监督方法和基准方法,1/8 标注数据已为模型训练提供足够的监督信息,取得了与全监督方法相当的性能,并超出全监督方法 0.17%。

为测试本文方法的稳定性,随机选择 6 个工位作为训练集,其余作为测试集,进行 10 次实验测试。5 种方法在焊装场景下的平均交并比和标准差如图 7 所示,可见本文方法的精度与稳定性均为最优。

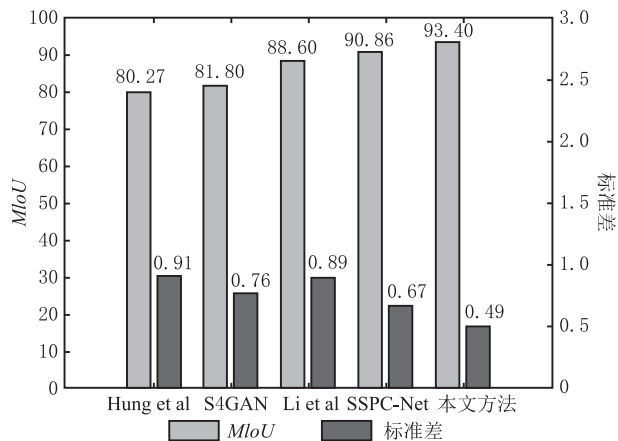


图 7 不同方法的分割精度与标准差

为进一步验证本文方法的普适性,在 S3DIS 公开数据集上进行了测试,对比结果如表 5 所示。

表 5 各类方法在 S3DIS 数据集上的 MIoU 对比

类别	方法		
	基准方法(1/8)	基准方法(全)	本文方法(1/8)
天花板	91.29	91.50	92.30
地板	96.10	97.59	97.95
墙壁	76.21	77.72	77.23
梁	0.00	0.00	0.01
柱	15.58	15.86	18.08
窗户	58.13	60.50	61.38
门	45.60	45.12	46.97
桌子	72.51	76.82	80.89
椅子	62.36	68.61	71.83
沙发	51.00	56.57	59.50
书架	41.91	45.97	49.10
黑板	56.90	60.79	55.60
其他	41.65	45.96	47.00
MIoU	54.56	57.15	58.30

由表 5 和图 8 可见,在 1/8 标注数据下,本文方法在 S3DIS 数据集上的 MIoU 超出全监督方法 1.15%。然而,由于鉴别器在某些类别上的置信度

判断性能问题, 伪标签中的错误标签对最终模型性能有较大影响, 导致某些类别的实验结果评价指标呈现升降不一的现象。

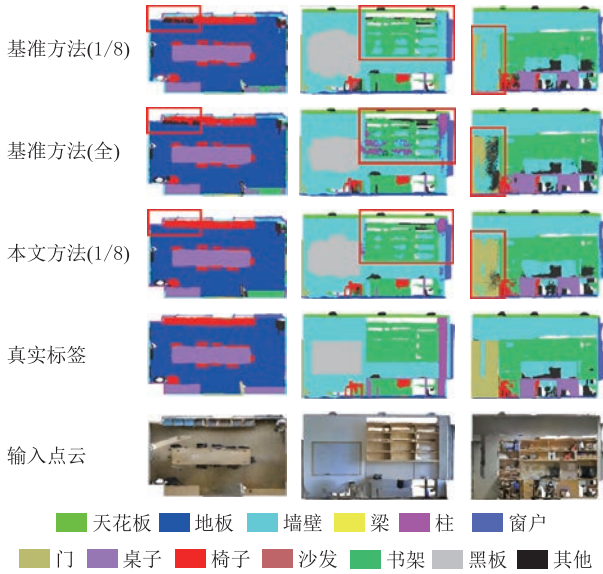


图8 S3DIS数据集分割结果可视化

通过上述分析可见, 现有半监督点云分割方法^[14-15]不适用于复杂焊装场景下的点云语义分割任务, 其未能充分利用无标记点云, 其他通用半监督语义分割方法^[18,20]未考虑点云的几何特性, 分割性能难以满足要求。本文提出的半监督点云语义分割框架充分考虑了点云的几何特性, 通过有效利用大量未标注的点云数据, 在保证分割性能的同时降低了焊装场景下的点云标注成本。

2.3 消融实验

2.3.1 生成器采样次数与网络层数分析

本节通过实验探究了不同采样方法与采样次数对生成器模型性能的影响, 实验结果如表 6 所示。在相同采样次数下, FPS 方法可以更好地保留点云下采样过程中的整体形状和几何细节特征, 分割性能相比 RS 方法提升约 5.2%。随着采样次数的增加, 生成器模型的性能得到了进一步提升。

表 6 采样方法与采样次数对生成器性能的影响

采样方法	采样次数	
	四次采样	五次采样
随机采样	88.10	88.78
最远点采样	92.69	93.57

2.3.2 生成器通道维度对 WISS 分割精度的影响

为了测试生成器不同通道维度对分割性能的影响,

设置编码器最后一层通道维度分别为 512, 1 024, 2 048 进行多次实验, 结果如图 9 所示。可见通道维度为 512 时, 精度较低; 通道数为 2 048 时, 精度达到最高, 但相比 1 024 维度, 稳定性较差, 提升效果较小。综合考虑精度、和模型稳定性, 选择通道数为 1 024 维。

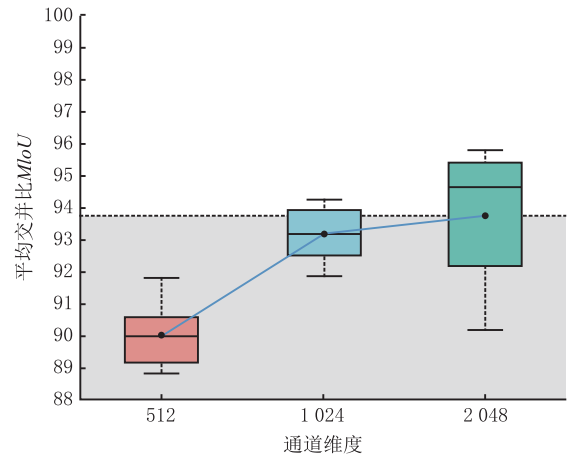


图9 不同通道维度下的分割精度

2.3.3 不同模块组合对 WISS 分割精度的影响

本节探究了不同标注比例下不同模块组合对半监督点云分割精度的影响, 如表 7 所示。在仅有 1/8 标记样本的情况下, 当添加鉴别器时, 分割性能提升 4.00%。随后, 加入自训练机制, 整个学习机制变成半监督学习模式。网络在使用标记数据的同时, 也充分利用训练数据中的未标记数据, 分割精度提高了 13.26%。最后, 加入平滑性约束, 进一步提高伪标签的可靠性, 分割精度进一步提升 2.09%, 与完全监督方法分割精度相当。

表 7 不同模块组合在 WWIS 数据集上的实验结果

标注数据比例	1/8	1/4	全
基准方法	77.95	79.95	93.57
基准方法+对抗损失	81.07	83.73	N/A
基准方法+对抗损失+自训练损失	91.82	92.57	N/A
基准方法+对抗损失+自训练损失+平滑约束	93.74	93.98	N/A

3 结束语

本文针对复杂焊装场景点云分割精度低的问题, 将生成对抗网络与半监督方法相结合, 提出一种基于生成对抗网络的半监督点云分割框架。该框架

利用基于生成对抗网络的半监督学习为无标记点云生成高置信度伪标签,并引入平滑性约束提高伪标签的可靠性,以准确训练生成器。该方法通过使用少量标注点云数据与大量未标注点云训练生成器,达到与全监督方法相近的分割性能。最后,通过自采焊工工位点云数据集与公开数据集验证了方法的有效性。

本文方法针对不同焊装场景的分割任务,置信度与超参数调整过程不尽相同,调参过程较为复杂耗时,未来考虑引入动态置信度阈值,以加速训练并提升点云分割性能。

参考文献:

- [1] SUN Zhiqiang, ZHENG Hangbin, LYU Chaofan, et al. Neural rendering-based fast scene geometry modeling and retrieval method for digital twin assets[J]. *Computer Integrated Manufacturing Systems*, 2024, 30(4):1189-1200(in Chinese). [孙志强,郑杭彬,吕超凡,等. 基于神经渲染的数字孪生产产快速场景几何建模与检索方法[J]. *计算机集成制造系统*, 2024, 30(4):1189-1200.]
- [2] LYU Chaofan, HUANG Delin, LIU Tianyuan, et al. Machining feature recognition method based on point cloud deep learning[J]. *Computer Integrated Manufacturing Systems*, 2023, 29(3):752-762(in Chinese). [吕超凡,黄德林,刘天元,等. 基于点云深度学习的加工特征识别方法[J]. *计算机集成制造系统*, 2023, 29(3):752-762.]
- [3] SUN Dianzhu, SHEN Jianghua, WANG Siteng, et al. A method for clustering and grading estimation of normal features in complex surface point clouds[J]. *Computer Integrated Manufacturing Systems*, 2021, 27(5):1440-1446(in Chinese). [孙殿柱,沈江华,汪思腾,等. 复杂型面点云的法向特征聚类分级估计方法[J]. *计算机集成制造系统*, 2021, 27(5):1440-1446.]
- [4] SU H, MAJI S, KALOGERAKIS E, et al. Multi-view convolutional neural networks for 3D shape recognition[C]//Proceedings of the IEEE International Conference on Computer Vision. Washington, D. C., USA: IEEE, 2015:945-953.
- [5] HUANG H B, KALOGERAKIS E, CHAUDHURI S, et al. Learning local shape descriptors from part correspondences with multiview convolutional networks[J]. *ACM Transactions on Graphics*, 2017, 37(1):1-14.
- [6] FENG Y F, ZHANG Z Z, ZHAO X B, et al. Gvnn: Group-view convolutional neural networks for 3D shape recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington, D. C., USA: IEEE, 2018:264-272.
- [7] HUANG J, YOU S Y. Point cloud labeling using 3d convolutional neural network [C]//Proceedings of the 23rd International Conference on Pattern Recognition. Washington, D. C., USA: IEEE, 2016:2670-2675.
- [8] DAI A, NIEßNER M. 3DMV: Joint 3D-multi-view prediction for 3D semantic scene segmentation[C]//Proceedings of the European Conference on Computer Vision. Berlin, Germany: Springer-Verlag, 2018:458-474.
- [9] LI Y Z, WANG Y N, LIU Y H. 3D point cloud segmentation based on context feature for sheet metal part boundary recognition[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72:1-10.
- [10] QI C R, SU H, MO K C, 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. Washington, D. C., USA: IEEE, 2017:652-660.
- [11] 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.
- [12] WANG Y, SUN Y B, LIU Z W, et al. Dynamic graph CNN for learning on point clouds[J]. *ACM Transactions on Graphics*, 2019, 38(5):1-12.
- [13] HU Q Y, YANG B, XIE L H, et al. Learning semantic segmentation of large-scale point clouds with random sampling[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 44(11):8338-8354.
- [14] CHENG M M, HUI L, XIE J, et al. SSPC-Net: Semi-supervised semantic 3D point cloud segmentation network[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Menlo Park, Cal., USA: AAAI Press, 2021, 35(2):1140-1147.
- [15] LI H Y, SUN Z X, WU Y J, et al. Semi-supervised point cloud segmentation using self-training with label confidence prediction[J]. *Neurocomputing*, 2021, 437:227-237.
- [16] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. *Advances in Neural Information Processing Systems*, 2014, 27:2672-2680
- [17] ODENA A. Semi-supervised learning with generative adversarial networks[EB/OL]. (2016-10-22) [2023-11-24]. <https://doi.org/10.48550/arXiv:1606.01583>.
- [18] MITTAL S, TATARCHENKO M, BROX T. Semi-supervised semantic segmentation with high-and low-level consistency[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 43(4):1369-1379.
- [19] SOULY N, SPAMPINATO C, SHAH M. Semi supervised semantic segmentation using generative adversarial network [C]//Proceedings of the IEEE International Conference on Computer Vision. Washington, D. C., USA: IEEE, 2017: 5688-5696.
- [20] HUNG W C, TSAI Y H, LIOU Y T, et al. Adversarial learning for semi-supervised semantic segmentation[EB/OL]. (2018-7-24) [2023-11-24]. <https://doi.org/10.48550/arXiv:1802.07934>.
- [21] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. (2016-01-07) [2023-11-24]. <https://doi.org/10.48550/arXiv:1511.06434>.

- [22] ARMENI I, SENER O, ZAMIR A R, et al. 3D semantic parsing of large-scale indoor spaces[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington, D. C., USA: IEEE, 2016: 1534-1543.

作者简介:

- 韩松杰(1999-), 男, 河南郑州人, 硕士研究生, 研究方向: 三维点云处理技术, E-mail: hansongjie0307@163.com;
- + 刘银华(1983-), 女, 河南许昌人, 教授, 博士, 博士生导师, 研究方向: 数字化设计与制造, 通讯作者, E-mail: liuyinhua@usst.edu.cn;
- 李彦征(1997-), 男, 吉林吉林人, 博士研究生, 研究方向: 三维环境感知与机器人智能规划, E-mail: 751361865@qq.com;
- 陈 浩(1989-), 男, 江苏扬州人, 讲师, 博士, 硕士生导师, 研究方向: 智能汽车与智能制造, E-mail: braver1989@usst.edu.cn。