

中图法分类号: TP18; TP391.4 文献标识码: A 文章编号: 1006-8961(2025)12-3914-13

论文引用格式: Hao W, Zuo Z B, Lu H S, Liang W, Jin H Y and Shi Z H. 2025. Enhanced attention-based joint semantic instance segmentation network for point clouds. Journal of Image and Graphics, 30(12):3914-3926(郝雯, 左占彬, 卢翰森, 梁玮, 金海燕, 石争浩. 2025. 基于增强注意力的点云语义实例联合分割. 中国图象图形学报, 30(12):3914-3926)[DOI: 10. 11834/jig. 240445]

## 基于增强注意力的点云语义实例联合分割

郝雯<sup>1,2\*</sup>, 左占彬<sup>1</sup>, 卢翰森<sup>1</sup>, 梁玮<sup>1,2</sup>, 金海燕<sup>1,2</sup>, 石争浩<sup>1,2</sup>

1. 西安理工大学计算机科学与工程学院, 西安 710048; 2. 陕西省网络计算与安全技术重点实验室, 西安 710048

**摘要:** **目的** 针对已有的特征融合策略难以充分挖掘语义—实例特征之间相关性问题, 提出一种基于增强注意力的语义实例联合分割网络。**方法** 首先, 通过构建基于注意力机制的语义特征提取模块, 有效捕获点间的局部上下文信息。然后, 利用编码—解码模块获得点云的初始语义特征和初始实例特征, 其中编码—解码模块包括基于注意力池化的 PointNet++ 集合抽象层、PointConv 的编码层、解码层以及 PointNet++ 的特征传播层。接着, 设计一个增强注意力模块, 利用双重注意力机制, 自适应地学习中心特征与邻近特征的相似性, 动态确定注意力权重, 并将得到的双重注意力权重进行求和, 与初始语义特征相乘, 获得增强后的语义特征。最后, 将增强注意力模块内嵌到联合分割模块的语义分支中, 将语义特征与实例特征有效融合, 提高语义—实例联合分割的精度。**结果** 相较于对比算法中最优值, 本文方法在 S3DIS (stanford large-scale 3D indoor semantics dataset) 数据集中语义分割平均交并比与实例分割平均加权覆盖率指标分别提升 3.5% 和 1.2%。在 ScanNet 数据集中, 语义分割平均交并比与实例分割中平均加权覆盖率指标分别提升 3.2% 和 2.8%。**结论** 本文提出的网络模型能够有效融合提取的语义特征与实例特征, 其语义分割与实例分割的准确性明显优于现有的联合分割方法。

**关键词:** 深度学习; 点云; 语义分割; 实例分割; 增强注意力机制

### Enhanced attention-based joint semantic instance segmentation network for point clouds

Hao Wen<sup>1,2\*</sup>, Zuo Zhanbin<sup>1</sup>, Lu Hansen<sup>1</sup>, Liang Wei<sup>1,2</sup>, Jin Haiyan<sup>1,2</sup>, Shi Zhenghao<sup>1,2</sup>

1. Department of Computer Science, Xi'an University of Technology, Xi'an 710048, China;

2. Shaanxi Key Laboratory for Network Computing and Security Technology, Xi'an 710048, China

**Abstract: Objective** With the rapid advancement of 3D sensing technologies such as LiDAR (light detection and ranging) and depth cameras, large-scale 3D point clouds have emerged as a crucial data source for a wide range of applications, including autonomous driving, robotic navigation, augmented reality, and urban scene reconstruction. Compared to 2D images, point clouds offer precise spatial geometry and provide a comprehensive representation of the environment without perspective distortion. Additionally, they are robust to variations in lighting and texture. Point cloud segmentation plays a crucial role in scene analysis and interpretation. The segmentation can be categorized into three types: semantic segmenta-

收稿日期: 2024-08-12; 修回日期: 2025-05-15; 预印本日期: 2025-05-22

\* 通信作者: 郝雯 haowensxf@163.com

**基金项目:** 国家自然科学基金项目(61602373); 陕西省教育厅青年创新团队项目(24JY116); 陕西省自然科学基金基础研究计划(2024JC-YBMS-477); 中国国家留学基金资助

**Supported by:** National Natural Science Foundation of China(61602373); Shaanxi Provincial Department of Education Youth Innovation Team Project(24JY116); Natural Science Basic Research Program of Shaanxi Province(2024JC-YBMS-477); China Scholarship Council Program

tion, instance segmentation, and joint semantic-instance segmentation. Semantic segmentation partitions a 3D scene into informative regions and assigns each region to a specific class. Instance segmentation identifies and separates individual objects at the point level, including those that belong to the same semantic category. In recent years, researchers have increasingly focused on combining the two tasks to achieve more consistent and informative scene-level interpretations. Joint semantic-instance segmentation leverages the intrinsic correlation between semantic and instance-level segmentation, enabling the two tasks to complement and reinforce each other. In 3D point cloud contexts, this joint approach substantially improves the capability of the system to comprehend complex environments and offers strong technical support for the development of intelligent systems. Consequently, this approach has become an area of growing interest and active research. However, most existing methods for joint semantic-instance segmentation rely on simplistic feature fusion strategies, which limit their effectiveness in fully capturing the potential relationship between semantic and instance features. Aiming to address this limitation, an enhanced attention-based joint semantic-instance segmentation network is proposed. This network is designed to effectively model and utilize the correlation between semantic and instance information. **Method** The enhanced attention-based joint semantic-instance segmentation neural network (EAJS-Net) incorporates a semantic feature extraction module based on an attention mechanism. This module focuses on the local neighborhood of each point and dynamically adjusts attention weights to emphasize key information, thereby enhancing the extraction of semantic features across points. Additionally, an attention-enhanced semantic/instance feature fusion module is introduced, which adaptively learns the similarity between central and adjacent features. This design reinforces key characteristics and effectively captures the correlation between instance and semantic segmentation, ultimately improving overall segmentation accuracy. EAJS-Net integrates PointNet++ and PointConv as its backbone network and comprises three main components: a point feature enhancement module, an encoder-decoder module, and an enhanced attention-based joint segmentation module. The input to EAJS-Net includes  $N \times 9$  dimensional point cloud data, where  $N$  represents the number of points, and the nine dimensions include coordinate values ( $XYZ$ ), color information ( $RGB$ ), and normalized coordinates. A semantic feature extraction module based on an attention mechanism is employed to effectively capture local contextual information between points. The enhanced features extracted by this module are then fed into the encoding layer, which includes four encoding modules: one attention pooling-based set abstraction layer adapted from PointNet++ and three feature encoding layers derived from PointConv. The corresponding decoding layer comprises four decoding modules: three deep feature decoding layers derived from PointConv and one feature propagation layer from PointNet++. By utilizing the attention pooling-based set abstraction layer from PointNet++, the network effectively captures spatial geometric relationships among features. Through the combination of the encoding and decoding layers, the initial semantic and instance features of the point cloud are extracted, laying the foundation for accurate joint segmentation. An enhanced attention module is designed to adaptively learn the similarity between central and neighboring features through dual attention mechanisms, which dynamically compute attention weights. These dual attention weights are summed and applied to the initial semantic features, resulting in enhanced semantic representations. This module is embedded within the semantic branch of the joint segmentation module, enabling more effective integration of semantic and instance features to improve joint segmentation accuracy. The encoded features are then upsampled through two parallel decoder branches to generate an instance feature matrix and a semantic feature matrix, which serve as inputs to the joint segmentation module. Within this module, the semantic and instance branches are integrated using the enhanced attention mechanism. The final output comprises instance embeddings and semantic predictions, supporting precise and consistent segmentation results. **Result** The proposed network is evaluated on the Stanford large-scale 3D indoor spaces (S3DIS) dataset and ScanNet V2 to assess its performance on point cloud segmentation tasks. Six fold cross-validation is performed on the S3DIS dataset, and the results of EAJS-Net are compared with those of the state-of-the-art (SOTA) methods. For semantic segmentation on the S3DIS dataset, EAJS-Net achieves a mean intersection over union (mIoU) of 65.9%, overall accuracy (oAcc) of 89.1%, and mean accuracy (mAcc) of 76.0%. Compared to JSNet++, these results represent improvements of 3.5% (mIoU), 0.4% (oAcc), and 3.2% (mAcc). For instance segmentation, EAJS-Net reaches a weighted coverage rate of 61.1%, outperforming JSNet++ by 4.1% (mean weighted coverage, mWCov), 4.6% (mean coverage, mCov), and 1.2% (mean recall, mRec). On the ScanNet dataset, EAJS-Net improves the mIoU for semantic segmentation by 3.2% and increases the weighted coverage

rate for instance segmentation by 2.8% compared to JSNet. Visual comparisons between EAJS-Net and other SOTA methods are also presented, demonstrating that EAJS-Net consistently achieves superior segmentation results, even in complex indoor scenes. In addition, ablation experiments are conducted to validate the effectiveness of individual modules within the network. The enhanced attention-based joint segmentation module in EAJS-Net dynamically adjusts attention weights to effectively capture various features, successfully integrating semantic and instance features into the semantic feature space. This integration notably enhances the performance of the semantic segmentation task. **Conclusion** Aiming to address the limitations of existing feature fusion strategies that fail to fully capture inter-instance semantic correlations, this paper proposes a novel semantic-instance joint segmentation network, EAJS-Net, based on an enhanced attention mechanism. A new semantic feature extraction module is designed to capture contextual relationships among points. Additionally, an enhanced attention module is introduced to effectively aggregate instance features into the semantic feature space. This improved feature fusion strategy boosts the performance of joint semantic-instance segmentation. Experimental results demonstrate that EAJS-Net effectively integrates semantic and instance features, substantially improving the accuracy of both segmentation tasks compared to SOTA methods.

**Key words:** deep learning; point cloud; semantic segmentation; instance segmentation; enhanced attention-based mechanism

## 0 引言

点云分割是将点云数据划分成具有相似特征的不同部分,它是处理点云数据的关键步骤,已广泛应用于机器人导航、自动驾驶和文化遗产保护等多个领域,成为当前研究的热点(He等,2025)。点云分割可以分为语义分割、实例分割和语义实例联合分割3种类型。点云语义分割的目标是通过点对每个点的细致分类,为场景中的每一个点赋予特定的类别标签,从而实现对三维场景中各类别语义信息的准确划分和识别(Qi等,2017a;Wang等,2019b;Zhang等,2023;Xu等,2024;朱仲杰等,2024)。点云实例分割不仅需要点对点云数据中的每个点赋予类别标签,还需要区分同一语义类别中的不同实例(Vu等,2022;Ngo等,2023;Yao等,2025;Zhang等,2024)。由于不同类别的点属于不同实例,同一实例的点属于同一类别,语义实例联合分割利用了两者的相关性,促进实例分割和语义分割的协同进行(Wang等,2019a;Chen等,2022;刘苏毅等,2023)。点云场景语义实例联合分割的研究不仅能提升系统对三维世界的感知能力,还为智能化系统的实现提供了坚实的技术支持。因此,语义实例联合分割吸引了众多研究学者的关注。

ASIS (associatively segmenting instances and semantics)(Wang等,2019a)首次将语义分割和实例分割两个任务结合,利用多层感知机将语义特征映

射到实例特征空间。同时,利用K近邻(K-nearest neighbor,KNN)查找算法查找属于同一实例的点,融合其语义特征,促进每个点的语义预测,但仅利用简单的融合策略并不能充分挖掘两个任务之间潜在的相关性。随后,SASO(Tan等,2021a)利用多尺度语义关联模块提取点云不同尺度的语义特征,通过显著点的选择和优化聚类,将具有相似语义特征的点聚合为实例,实现对点云的语义与实例分割。BAN (bi-directional attention network)(Wu等,2020)通过构建双向注意力模块,使用一个任务的特征测量的相似性矩阵来帮助另一个任务聚合非局部信息,避免潜在的特征排斥和任务冲突。ISSF(instance separation and semantic fusion)(Zhong和Zeng,2021)通过特定实例的质心位置来补充位置不变性语义特征,以分离不同的实例。同时,设计了一个基于注意力的语义融合模块,在实例嵌入空间中对注意力进行编码,并将其用于融合语义特征空间中的语义信息。HCFS3D (hierarchical coupled feature selection network for 3D semantic and instance segmentation)(Tan等,2021b)提出一种基于层次耦合特征选择的网络框架。为了避免过拟合,提出一种自适应的平滑损失函数,为不同样本自适应地分配不同的权重。JSNet(joint segmentation network)(Zhao和Tao,2020)将PointNet++(Qi等,2017b)和PointConv(Wu等,2019)相结合作为骨干网络,设计了一个联合分割模块,利用特征拼接和相加操作提升语义分割与实例分割的性能。随后,JSNet++(Zhao和Tao,2023)基于

几何结构和密度信息设计了动态滤波器卷积(dynamic filters for convolution, DFConv),以提取点的结构化特征。并设计点对相关模块对JSNet原有的联合分割模块进行改进,进一步提高分割精度。Oneformer3d(Kolodiaznyi等,2024)通过设计一组可学习的核,生成一个实例或语义类别的掩码。这些核由基于Transformer的解码器进行训练,解码器以统一的实例和语义查询作为输入,从而实现点云场景实例、语义和全景分割。然而,目前已有的语义实例联合分割方法往往将点云固有的几何属性直接输入到网络中,并未考虑点间的语义特征。同时,在特征聚合阶段,平等地对待邻域特征,将邻域特征简单地融合在一起,难以充分地挖掘语义分割任务和实例分割任务之间的潜在关系。

针对已有的特征融合策略难以充分挖掘语义—实例之间相关性问题,本文提出一种基于增强注意力机制的语义实例联合分割网络(enhanced attention-based joint semantic-instance segmentation neural network, EAJS-Net)。本文的主要工作为:1)设计了一种基于注意力机制的语义特征提取模块,关注每个点的邻域区域,动态调整注意力权重,关注关键信息,更好地提取点间的语义特征;2)提出基于增强注意力机制的语义/实例特征融合模块,本文提出的融合模块能够自适应地学习中心特征与邻域特征的相似性,强化重要特征,充分挖掘实例分割和语义分割之间的相关性,以提高分割精度;3)将EAJS-Net在S3DIS和ScanNet V2两个公共数据集上进行验证,实验结果表明,EAJS-Net在总体准确率(overall

accuracy, oAcc)、平均准确率(mean accuracy, mAcc)、平均覆盖率(mean coverage, mCov)、平均加权覆盖率(mean weighted coverage, mWCov)等多项指标优于其他最先进的方法(state-of-the-art, SOTA),具有较好的分割性能。

## 1 基于增强注意力机制的联合分割

### 1.1 网络整体结构

图1为EAJS-Net的整体网络结构。EAJS-Net将PointNet++(Qi等,2017b)与PointConv(Wu等,2019)相结合作为骨干网络。整个网络由3部分组成:基于注意力机制的语义特征提取模块、编码—解码模块以及基于增强注意力的联合分割模块。EAJS-Net将 $N \times 9$ 维点云数据作为输入,其中, $N$ 是点的个数,9维分别代表坐标值XYZ,颜色信息RGB以及归一化的坐标信息。首先,将点云数据输入到基于注意力机制的语义特征提取模块,可以得到 $N \times 16$ 维的语义特征。然后,将该语义特征输入到EAJS-Net的编码模块,整个编码模块包括4层:1个基于注意力池化的集合抽象层和PointConv的3个特征编码层。每层输出的特征大小分别是 $N_a \times 64$ 、 $N_b \times 128$ 、 $N_c \times 256$ 和 $N_d \times 512$ ,其中 $N_a$ 、 $N_b$ 、 $N_c$ 和 $N_d$ 分别是每个模块下采样的点数。相应地,整个解码模块包括4层:3个PointConv的深度特征解码层和1个PointNet++的特征传播层,每个模块输出的特征大小分别是 $N_d \times 512$ 、 $N_c \times 256$ 、 $N_b \times 128$ 和 $N_a \times 128$ 。通过两个

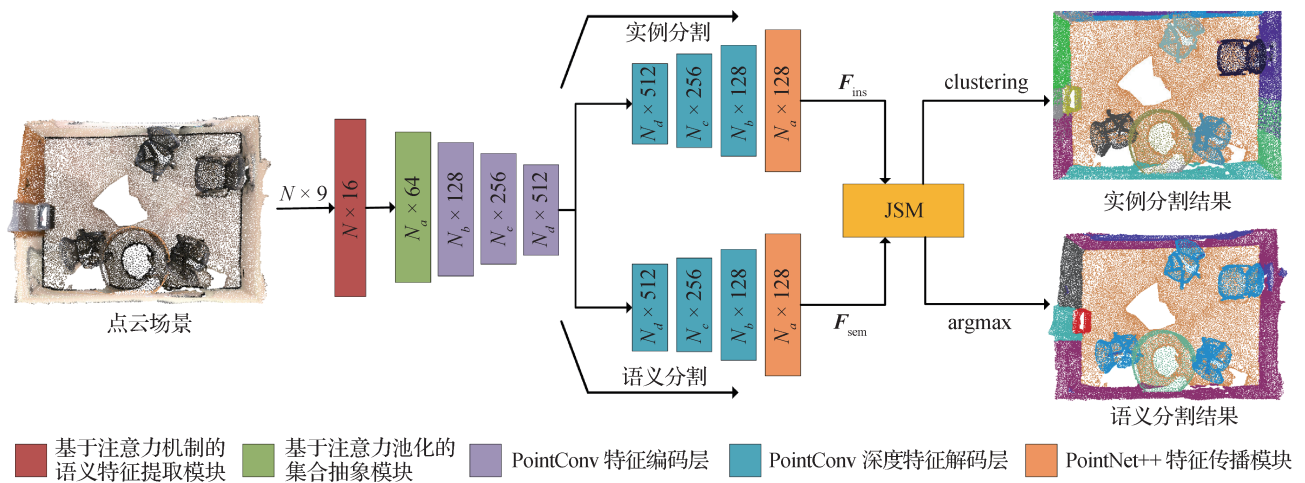


图1 EAJS-Net网络结构

Fig. 1 Architecture of the EAJS-Net

实例特征  $F_{ins}$  和语义特征  $F_{sem}$ 。然后将实例特征  $F_{ins}$  和语义特征  $F_{sem}$  输入到增强注意力机制的联合分割模块(joint segmentation module, JSM), 最终得到实例嵌入  $E_{ins}$  和语义预测  $P_{sem}$ 。

## 1.2 基于注意力机制的语义特征提取

为了丰富点云数据的表征, 提升分割性能, 本文提出一种基于注意力机制的语义特征提取模块(attention mechanism-based semantic feature extraction module, ASFE), 通过分析点与邻域点间的空间关系, 有效捕获点云之间的语义相关性, 获取点间的语义特征。ASFE 模块的结构如图 2 所示。

首先, 利用 KNN 查找点  $p_i$  的  $k$  近邻, 构造点与邻近点的边信息  $e_{ij} = (p_i - p_j)$ 。将边信息送入多层感知机(multilayer perceptron, MLP), 得到特征  $F_{c1}$ 。同时, 将点云数据作为输入, 经过两层 MLP, 得到特征  $F_{c2}$ 。然后, 将  $F_{c1}$  送入 MLP 后与  $F_{c2}$  相加, 利用 softmax 函数对求和结果进行归一化, 得到权重矩阵。最后, 将特征  $F_{c1}$  与权重矩阵相乘, 得到最终的语义特征  $F_c$ 。

与原始输入点云相比, ASFE 模块在分析点云局部结构信息的同时, 根据特征相关性分配邻域点的权重, 自适应地聚合局部邻域特征, 有效提取点云的上下文信息, 这种丰富的特征表示将作为 EAJS-Net 编码模块的输入。

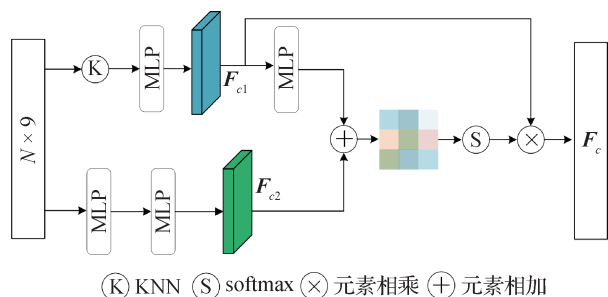


图 2 基于注意力机制的语义特征提取模块

Fig. 2 Attention mechanism-based semantic feature extraction module

## 1.3 编码—解码模块

EAJS-Net 采用与 JSNet 类似的网络架构, 使用 PointNet++ 的集合抽象层作为编码器的第 1 层, 后 3 层则采用 PointConv 的 3 个特征编码层。相应地, 解码层由 PointConv 的 3 个深度特征解码层和 PointNet++ 的 1 个特征传播层组成。目前已有网络的编

码器第 1 层, 往往利用最大池化或平均池化操作来整合提取的特征。最大池化操作仅保留局部最大值, 易受到局部极大值的影响。平均池化则计算局部平均值, 易受噪声影响。因此, 本文提出的 EAJS-Net 在编码器第 1 层引入注意力池化操作(attention pooling, AP)(Hu 等, 2020), 根据不同的输入特征自适应地调整权重分配, 通过加权平均的方式, 对噪声和异常值进行抑制, 在保留更多信息的同时, 捕获关键特征。

给定特征  $F_i$  与其邻域特征集合  $F_i^K = \{f_i^1, f_i^2, \dots, f_i^k, \dots, f_i^K\}$ , 邻域特征  $f_i^k$  的特征维度为  $N_x \times K \times d$ , 其中,  $N_x$  表示采样密度, 即输入点云的采样点数,  $K$  表示每个采样点邻域半径内的邻居点数,  $d$  表示编码层特征空间维度。注意力池化模块首先定义一个基于多层感知机的函数  $g(\cdot)$  和一个权重矩阵  $W_a$ , 计算每个邻域点特征的注意力权重得分  $a_i^k$ 。具体为

$$a_i^k = g(f_i^k, W_a) \quad (1)$$

为确保所有点特征的注意力权重得分之和为 1, 对每个邻居点特征注意力权重得分  $a_i^k$  进行 softmax 归一化操作, 具体为

$$S_i^k = \frac{\exp(a_i^k)}{\sum_{k=1}^K a_i^k} \quad (2)$$

将归一化后的注意力权重得分  $S_i^k$  与邻域点特征  $f_i^k$  相乘, 并加权求和, 得到特征向量  $f_i^w$ , 具体为

$$f_i^w = \sum_{k=1}^K (S_i^k \times f_i^k) \quad (3)$$

接着, 经过 PointConv 的 3 个特征编码层、3 个解码层以及 PointNet++ 特征传播模块, 得到实例特征矩阵  $F_{ins}$  和语义特征矩阵  $F_{sem}$ 。

## 1.4 基于增强注意力的联合分割模块

如果仅仅对语义特征和实例特征进行简单的拼接或相加, 并不能充分挖掘两种特征之间的相关性, 为更好地提取实例分割和语义分割之间的关联性, 提高分割精度, 本文提出一种基于增强注意力机制的特征融合策略, 旨在促进两种任务的联合分割。

如图 3 所示, 增强注意力模块(enhanced attention module, EAM)将语义特征  $F_{sem}$  作为输入。首先, 分别将语义特征  $F_{sem}$  送入一层 MLP, 得到 3 个语义特征  $F_1$ 、 $F_2$  和  $F_3$ 。然后, 将特征  $F_1$  进行转置并与特征  $F_2$  相乘, 得到语义特征  $F_{s1}$ 。同样地, 将特征  $F_3$  进行

转置并与特征  $F_2$  相乘, 得到语义特征  $F_{s2}$ 。接着, 分别对  $F_{s1}$  和  $F_{s2}$  进行 softmax 操作并相加, 得到权重矩阵  $S_F$ 。最后, 将权重矩阵与特征  $F_{sem}$  相乘并送入 MLP, 得到增强后的语义特征  $F_{sem}^1$ 。

$$F_1 = F_2 = F_3 = f_{Conv1D}(F_{sem}) \quad (4)$$

$$F_{s1} = F_1^T \otimes F_2 \quad (5)$$

$$F_{s2} = F_3^T \otimes F_2 \quad (6)$$

$$S_F = f_{softmax}(F_{s1}) \oplus f_{softmax}(F_{s2}) \quad (7)$$

$$F_{sem}^1 = f_{Conv1D}(S_F \otimes F_{sem}) \quad (8)$$

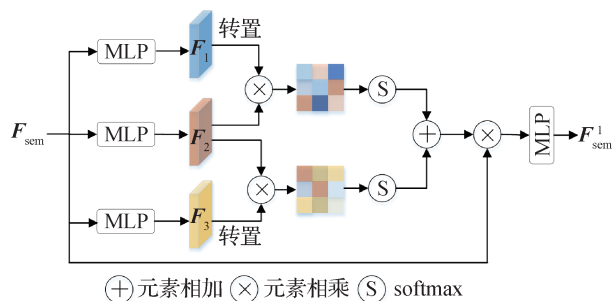


图3 增强注意力模块

Fig. 3 Enhanced attention module

基于增强注意力机制的联合分割模块(JSM)能够有效地整合语义特征和实例特征, 提高语义分割与实例分割性能。JSM 模块如图4所示。

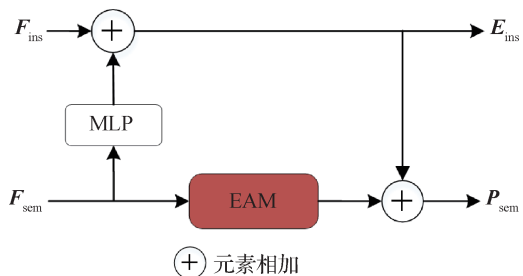


图4 联合分割模块

Fig. 4 Joint segmentation module

在实例分割任务中, 将实例特征矩阵  $F_{ins}$  作为联合分割模块的实例分支输入。经过一层 MLP, 将语义特征映射到实例特征空间后与实例特征  $F_{ins}$  融合, 最终得到实例嵌入  $E_{ins}$ 。

$$E_{ins} = F_{ins} \oplus f_{Conv1D}(F_{sem}) \quad (9)$$

在语义分割任务中, 将语义特征  $F_{sem}$  作为联合分割模块中语义分支的输入, 经增强注意力模块后, 得到语义矩阵  $F_{sem}^1$ 。然后将实例嵌入  $E_{ins}$  与  $F_{sem}^1$  相加, 得到最终的语义预测  $P_{sem}$ 。

$$P_{sem} = F_{sem}^1 \oplus E_{ins} \quad (10)$$

## 2 实验结果与分析

### 2.1 实验设置与评价指标

本文将提出的 EAJIS-Net 在 S3DIS (Stanford large-scale 3D indoor spaces) (Armeni 等, 2016) 和 ScanNet V2 (Dai 等, 2017) 两个公共数据集上进行评估。S3DIS 数据集包含 6 个室内区域, 共 272 个房间。数据集定义了 13 种语义类别, 如椅子、桌子、地板和墙壁等。在训练过程中, 将点云场景分为  $1 \text{ m} \times 1 \text{ m}$  的数据块, 每个数据块随机采样 4 096 个点。ScanNet V2 数据集由 1 201 个训练场景和 312 个测试场景组成, 由于已有机器硬件的限制, 如果每个数据块采样 8 192 个点, 则无法运行实验。因此, 本文采用与 S3DIS 相同的策略, 将场景分为  $1 \text{ m} \times 1 \text{ m}$  的数据块, 每个数据块随机采样 4 096 个点。

本文所有实验均在配备 Intel (R) i7-8700 CPU, 64 G 内存和单块 NVIDIA RTX 2080Ti GPU 的电脑上运行。在训练过程中, 动量 (momentum) 设置为 0.9, batch size 设置为 12, 迭代训练次数 (epoch) 设置为 100, 初始学习率 (learning rate) 设置为 0.001, 初始学习率每迭代 30 000 次降低为原来的一半, 采用 Adam 优化器优化模型。

对于语义分割任务, 评价指标包括总体准确率 oAcc、平均准确率 mAcc 和平均交并比 (mean intersection-over-union, mIoU)。

对于实例分割任务, 本文采用平均精确度 (mean precision, mPrec)、平均召回率 (mean recall, mRec)、平均覆盖率 mCov 和平均加权覆盖率 mWCov 作为评价标准。Cov 是根据标准值实例大小加权匹配平均实例 IoU 预测, 进而得到 WCov。给定标准区域  $G$  和预测区域  $O$ , Cov 和 WCov 可以表示为

$$Cov(G, O) = \sum_{m=1}^{|G|} \frac{1}{|G|} \max_n IoU(r_m^G, r_n^O) \quad (11)$$

$$WCov(G, O) = \sum_{m=1}^{|G|} \omega_m \max_n IoU(r_m^G, r_n^O) \quad (12)$$

$$\omega_m = \frac{|r_m^G|}{\sum_k |r_k^G|} \quad (13)$$

式中,  $|r_m^G|$  表示标准区域  $m$  中所包含的点数。

## 2.2 实验结果与分析

### 2.2.1 S3DIS数据集实验结果与分析

表1展示了EAJS-Net在S3DIS数据集上6折交叉验证的实例分割结果。为了展现本文提出的EAJS-Net的分割性能,选取目前较先进的方法进行比较,包括ASIS(Wang等,2019a)、3DCFS(3D point clouds segmentation framework via coupled feature selection)(Du等,2020)、BAN(Wu等,2020)、ISSF(Zhong和Zeng,2021)、SASO(semantic association and salient point clustering optimization)(Tan等,2021a)、JSNet(Zhao和Tao,2020)、JSNet++(Zhao和Tao,2023)、AM-ASIS(attention mechanism based associatively segmenting instances and semantics)(项学泳等,2023)与JPCP(joint prototype and coefficient prediction for 3D instance segmentation)(Royen等,2024)。EAJS-Net在mWCov、mCov、mPrec以及mRec指标中分别达到61.1%、57.4%、65.5%和54.6%。与JSNet++相比,EAJS-Net在mWCov、mCov与mRec指标中分别提升4.1%、4.6%以及1.2%。与JPCP相比,EAJS-Net在mPrec与mRec指标上分别低5%与1.9%,表明所提模型在实例边界细化和小目标识别方面仍存在一定不足。在mWCov与mCov两个指标上,EAJS-Net分别提升1.2%与4.9%,表明所提出的方法在整体覆盖效果和大尺寸目标实例的分割方面具有较好的表现。这主要得益于EAJS-Net引

表1 6折交叉验证实例分割结果

Table 1 Instance segmentation results for 6-fold cross-validation

方法	mWCov	mCov	mPrec	mRec
3DCFS	57.1	53.1	63.7	49.1
ASIS	55.1	51.2	63.6	47.5
BAN	56.2	52.1	63.4	51.0
ISSF	58.1	54.2	65.3	50.8
SASO	58.3	54.5	64.2	50.8
JSNet	58.0	54.1	66.9	53.9
JSNet++	57.0	52.8	70.2	53.4
AM-ASIS	57.1	53.1	65.2	52.8
JPCP	59.9	52.5	<b>70.5</b>	<b>56.5</b>
EAJS-Net(本文)	<b>61.1</b>	<b>57.4</b>	65.5	54.6

注:加粗字体表示各列最优结果。

入的注意力机制,使得网络在特征提取过程中能够更关注关键区域和整体形状的连续性。

表2列出了EAJS-Net在S3DIS数据集上6折交叉验证的语义分割结果。EAJS-Net在mIoU、oAcc和mAcc指标中分别达到65.9%、89.1%和76%,高于JSNet++与AM-ASIS方法。与JSNet++相比,mIoU、oAcc和mAcc指标分别提升3.5%、0.4%和3.2%。相比AM-ASIS,EAJS-Net在mIoU和oAcc指标上分别提升4.2%和1.0%。实验结果验证了EAJS-Net在语义分割任务上的有效性。

表2 6折交叉验证语义分割结果

Table 2 Semantic segmentation results for 6-fold cross-validation

方法	mIoU	oAcc	mAcc
3DCFS	60.3	86.3	72.4
ASIS	59.3	86.2	70.1
BAN	60.8	87.0	71.7
ISSF	60.9	86.7	71.6
SASO	61.1	87.0	72.8
JSNet	61.7	88.7	71.7
JSNet++	62.4	88.7	72.8
AM-ASIS	61.7	88.1	-
EAJS-Net(本文)	<b>65.9</b>	<b>89.1</b>	<b>76.0</b>

注:加粗字体表示各列最优结果。“-”表示未给出实验结果。

图5展示了EAJS-Net与ASIS、BAN、JSNet在S3DIS数据集上实例分割可视化结果。第1列为点云场景,第2列为分割真值,随后各列分别表示ASIS、BAN、JSNet与EAJS-Net在不同场景下的实例分割结果。在场景1中,ASIS、BAN以及JSNet未能将邻近的椅子单独分割出来。另外,BAN和JSNet对墙面上的黑板存在误分的情况,JSNet未能将右侧墙边的柜子正确分割,而EAJS-Net完成了单个椅子、柜子与黑板的准确分割。在场景2中,ASIS和BAN仍然无法精准分割单个椅子,而JSNet与EAJS-Net则实现了对单个椅子的准确分割。在场景3中,ASIS并未将每个墙面单独提取出来,BAN对左上侧物体进行了过分割,JSNet则将左侧墙面上的黑板分成了两部分。在场景4中,ASIS和BAN将相邻的两把椅子误分为一簇,BAN和JSNet对于墙体和门框

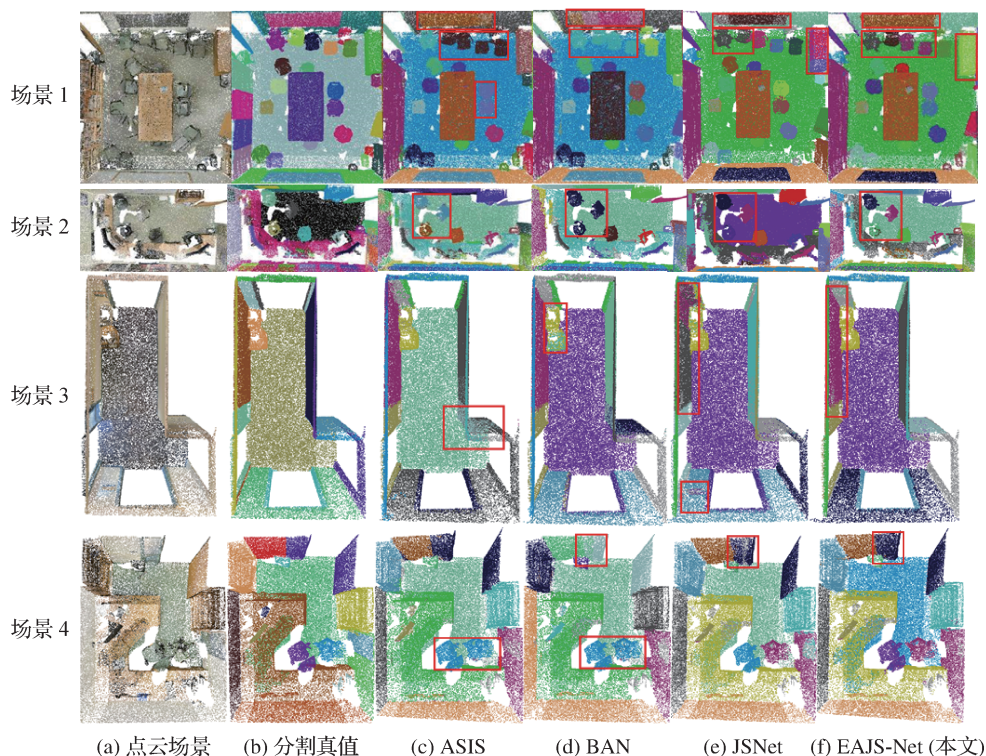


图5 不同方法在S3DIS数据集场景实例分割可视化对比

Fig. 5 Visual comparison of scene instance segmentation for different methods on the S3DIS dataset

((a) scene point clouds; (b) segmentation ground truth; (c) ASIS; (d) BAN; (e) JSNet; (f) EAJS-Net(ours))

的分割精度较低,两者掺杂有局部杂乱点。与ASIS、BAN和JSNet相比,EAJS-Net对场景3与场景4的分割结果更接近于真值。

图6展示了EAJS-Net与ASIS、BAN、JSNet在S3DIS数据集上语义分割可视化结果。在场景1中,与ASIS、BAN和JSNet相比,EAJS-Net对下侧墙面的分割更接近于真值。场景2中,ASIS和BAN未能对右侧墙体中的黑板准确分割,JSNet将黑板中间的部分点错误划分为其他类别。同时,它们将垃圾桶和椅子分为一类。在场景3中,由于右侧墙面上的两块黑板距离很近,导致ASIS、BAN和JSNet都未能对相邻的黑板实现精准分割。在场景4中,ASIS、BAN与JSNet均未能对场景左上侧区域(红框紫色区域)进行准确分割。从图6可以看出,本文提出的EAJS-Net对S3DIS数据集的语义分割结果更接近于分割真值,这主要是由于EAJS-Net中基于增强注意力机制的联合分割模块,能够更好地捕捉关键特征,动态调整各个特征的权重,将语义特征与实例特征有效地聚合到语义特征空间,促进语义分割任务。

### 2.2.2 ScanNet数据集实验结果与分析

ScanNet V2数据集存在更严重的遮挡与数据缺

失问题。由于已有机器硬件的限制,无法运行高分辨率的点云数据,因此将ScanNet V2数据集场景划分为 $1\text{ m} \times 1\text{ m}$ 的数据块,每个数据块随机采样4 096个点。从表3可以看出,在实例分割任务中,本文提出的EAJS-Net在mCov、mWCov、mRec和mPrec指标上分别达到34.2%、35.4%、34.4%和36.7%,明显优于ASIS、BAN和JSNet。与JSNet相比,EAJS-Net在mCov、mWCov、mRec和mPrec指标上分别提升2.8%、2.8%、2.4%和0.8%。在语义分割任务中,EAJS-Net在mAcc、oAcc和mIoU指标上分别达到55.8%、74.3%和43.5%,相比于ASIS,提升7.1%、1.3%和5.3%。在语义分割和实例分割任务中性能的明显改进,表明EAJS-Net能更好地捕捉语义和实例特征之间的关系,有效融合语义特征与实例特征,提升分割性能。

图7展示了ScanNet V2数据集的实例分割结果。从场景1可以看出,ASIS与BAN将右下角的马桶分为两部分,JSNet则将上侧的墙面分割为多个簇。EAJS-Net则可以将马桶与上侧的墙面准确分割出来,这主要归功于语义特征提取模块,该模块能够丰富点云的特征信息,增强了对细节特征的提取能

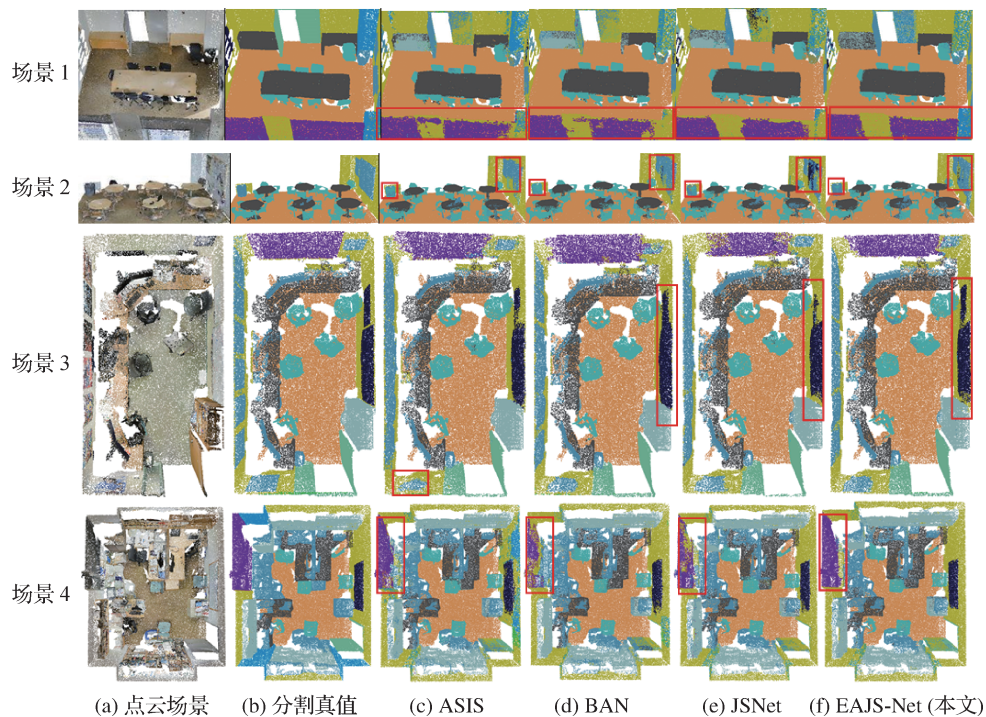


图6 不同方法在S3DIS数据集语义分割可视化对比

Fig. 6 Visual comparison of scene semantic segmentation for different methods on the S3DIS dataset  
((a) scene point clouds; (b) segmentation ground truth; (c) ASIS; (d) BAN; (e) JSNet; (f) EAJS-Net(ours))

表3 ScanNet数据集不同方法的语义—实例分割结果对比

Table 3 Comparison of semantic-instance segmentation results on the ScanNet dataset for different methods

方法	mWCov	mCov	mPrec	mRec	mIoU	oAcc	mAcc
BAN	28.8	27.6	30.4	26.3	38.2	72.8	49.4
ASIS	29.0	27.8	33.2	26.1	38.2	73.0	48.7
JSNet	32.6	31.4	35.9	32.0	40.3	73.2	52.3
<b>EAJS-Net (本文)</b>	<b>35.4</b>	<b>34.2</b>	<b>36.7</b>	<b>34.4</b>	<b>43.5</b>	<b>74.3</b>	<b>55.8</b>

注:加粗字体表示各列最优结果。

力。在场景2中,EAJS-Net对椅子的分割效果与分割真值更为接近,ASIS、BAN和JSNet都存在邻近椅子欠分割的情况。在场景3中,ASIS和BAN将并排的相邻椅子误分为一把椅子,未将单个椅子单独分割出来。同时,ASIS、BAN与JSNet都将地面分成了两部分。EAJS-Net对场景3的分割结果与分割真值最接近。在场景4中,ASIS和BAN未能实现对单个椅子的准确分割,4种方法均对左侧墙面过分割,主要是由于有小物体(真值图中蓝色点)非常靠近左侧墙面下面部分,导致这部分墙面提取的特征与上面部分不同,从而致使整个墙面被分为两部分。

图8展示了EAJS-Net在ScanNet V2数据集上的语义分割结果。在场景1中,EAJS-Net能够准确地将马桶分割出来,而ASIS和BAN将马桶盖与墙上的物体划分为一类,且存在杂乱点,JSNet则将马桶坐垫分为两部分。但是ASIS、BAN、JSNet和EAJS-Net均未能将左下角的柜子正确分类。这主要是因为该柜子由平面点组成,又紧靠墙面,EAJS-Net提取的语义与实例特征无法有效区分这些平面点,导致该场景语义分割结果不准确。在场景2中,EAJS-Net可以较为完整地将沙发分割为同一类别。在场景3中,相较于ASIS、BAN和JSNet,EAJS-Net对最左侧墙面的分割更接近于真值。在场景4中,ASIS、BAN和JSNet都没有准确地分割出墙角的垃圾桶,而EAJS-Net则能准确地分割出垃圾桶,与分割真值更为接近。实验结果表明,对于复杂室内场景,EAJS-Net仍然能够获得较好的分割结果。

## 2.3 消融研究

### 2.3.1 模块有效性

为了进一步验证EAJS-Net中各模块的有效性,本文在S3DIS数据集的区域5进行消融实验。表4展示了不同模块对语义分割和实例分割性能的影响。目前已有的语义实例联合分割方法在编码层往

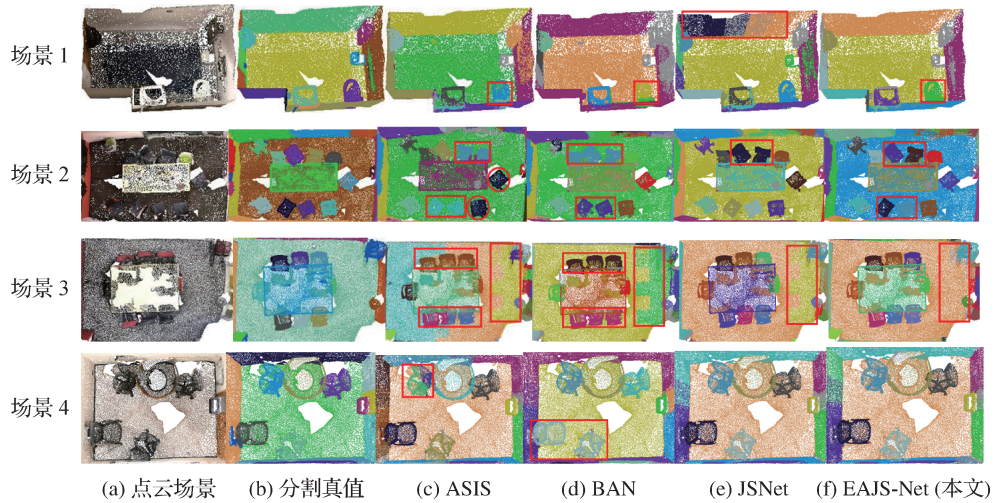


图7 不同方法在ScanNet数据集实例分割可视化对比

Fig. 7 Visual comparison of scene instance segmentation for different methods on ScanNet dataset

((a) scene point clouds; (b) segmentation ground truth; (c) ASIS; (d) BAN; (e) JSNet; (f) EAJS-Net(ours))

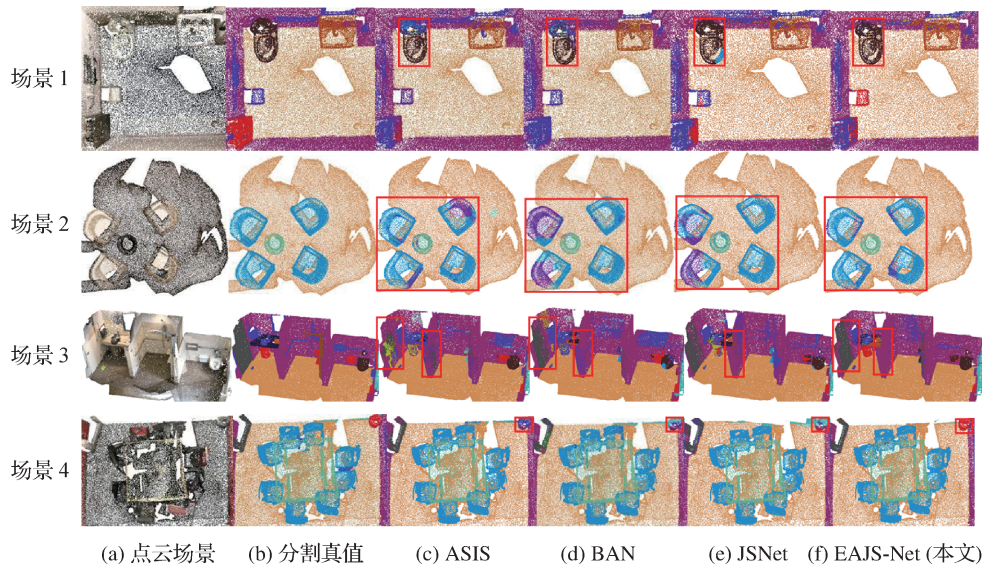


图8 不同方法在ScanNet数据集语义分割可视化对比

Fig. 8 Visual comparison of scene semantic segmentation for different methods on ScanNet dataset

((a) scene point clouds; (b) segmentation ground truth; (c) ASIS; (d) BAN; (e) JSNet; (f) EAJS-Net(ours))

往利用最大池化提取重要特征。最大池化操作仅关注区域内的最大值特征,可能会导致细节信息的丢失。因此,本文引入基于注意力池化的PointNet++和PointConv相结合的编码—解码模块(attentive pooling-based PointNet++ and PointConv, APConv)。实验结果表明,基于APConv的网络模型在mPrec和mIoU指标中分别提升1.0%和0.8%。随后引入ASFE模块,EAJS-Net在mWCov、mRec、mPrec和mAcc指标中分别提升7.0%、9.4%、6.0%和5.3%。实验验证了ASFE模块能够丰富点云的特征表示,

有效提升分割性能。考虑到简单的特征融合并不能充分挖掘语义特征与实例特征间的相关性,本文摒弃原有的基于KNN算法的简单融合策略,提出基于注意力机制的特征融合模块EAM,更好地发掘语义分割和实例分割任务间的相关性,有效融合语义特征与实例特征。可以看出,加入EAM模块后,mCov、mWCov和mIoU分别提升7.8%、7.5%和7.0%。以上消融实验结果很好地证明了EAJS-Net中基于注意力池化的PointNet++和PointConv相结合的编码—解码模块APConv、基于注意力机制的语义特征提取

表4 S3DIS数据集Area 5消融实验

Table 4 Ablation experiments of Area 5 on the S3DIS dataset

不同模块			实例分割				语义分割		
APConv	ASFE	EAM	mWCov	mCov	mPrec	mRec	mIoU	oAcc	mAcc
×	×	×	47.8	44.6	55.3	42.4	53.4	86.9	60.9
√	×	×	49.3	46.4	56.3	43.9	54.2	87.9	61.4
√	√	×	54.8	51.8	61.3	<b>51.8</b>	58.8	89.3	66.2
√	√	√	<b>55.3</b>	<b>52.4</b>	<b>62.5</b>	50.3	<b>60.4</b>	<b>90.2</b>	<b>66.6</b>

注:加粗字体表示各列最优结果。“√”表示采用此模块,“×”表示未采用。

模块 ASFE 以及基于注意力机制的特征融合模块 EAM 的有效性。

### 2.3.2 鲁棒性分析

为验证 EAJS-Net 的鲁棒性,本文将测试场景加入高斯噪声,图9展示了点云场景加入标准差为 0.001 的高斯噪声后,JSNet 与 EAJS-Net 的场景语义与实例分割结果。各列分别为:加入高斯噪声的点云场景、分割真值、JSNet 语义与实例分割结果以及

EAJS-Net 语义与实例分割结果。对于场景1(第1行),JSNet 将地面误分为桌子。对于场景2(第2行),JSNet 未将左侧、下侧以及右侧的墙面单独分割出来,EAJS-Net 虽然将左侧和下侧的墙面单独分割出来,但是右侧的墙面被过分割。JSNet 将地面分成若干部分,而 EAJS-Net 则将地面完整地分割出来。尽管在点云场景中加入高斯噪声会影响模型的分割性能,但是相较于 JSNet,EAJS-Net 的分割结果更接近于真值。

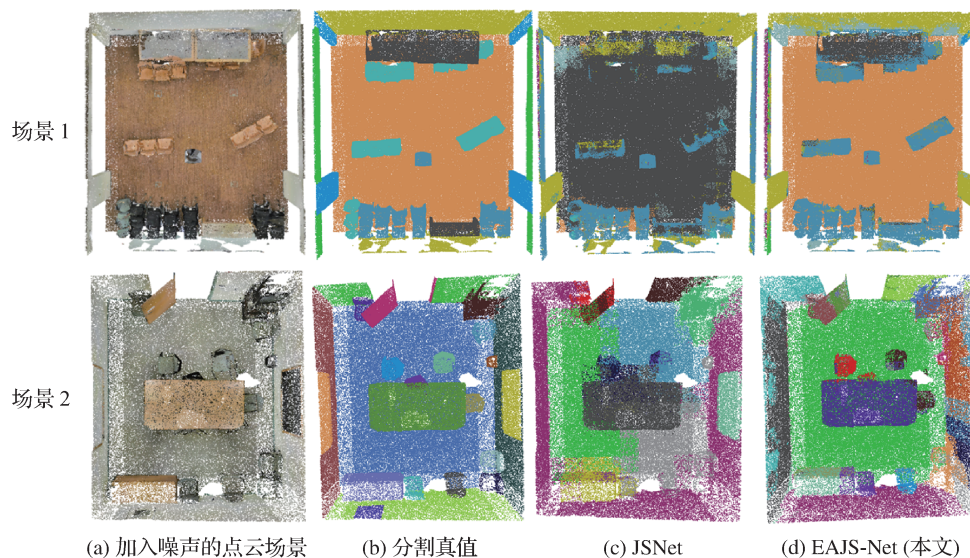


图9 加入高斯噪声后点云场景分割结果

Fig. 9 Point cloud scene segmentation results after adding Gaussian noise

((a) scene point clouds with Gaussian noise; (b) segmentation ground truth; (c) JSNet; (d) EAJS-Net(ours))

## 3 结论

目前已有的语义实例联合分割方法,大多将点云固有的几何信息输入到网络中,忽略了点间的局部几何结构与上下文信息。此外,在特征融合阶段,

采取简单的特征融合策略难以充分地挖掘语义特征和实例特征之间的潜在关系。因此,本文提出一个基于注意力机制的语义特征提取模块,对局部邻域点进行特征加权,增强关键点的特征表达能力,有效提取点间的语义特征。同时,提出一个基于增强注意力联合分割模块,将语义特征与实例特征进行有

效融合,促进语义—实例联合分割。实验结果表明,本文提出的EAJS-Net在S3DIS数据集和ScanNet V2数据集上具有较好的分割性能。

## 参考文献 (References)

- Armeni I, Sener O, Zamir A R, Jiang H L, Brilakis I, Fischer M and Savarese S. 2016. 3D semantic parsing of large-scale indoor spaces//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 1534-1543 [DOI: 10.1109/cvpr.2016.170]
- Chen F, Wu F, Gao G W, Ji Y M, Xu J, Jiang G P and Jing X Y. 2022. JSPNet: learning joint semantic and instance segmentation of point clouds via feature self-similarity and cross-task probability. *Pattern Recognition*, 122: #108250 [DOI: 10.1016/j.patcog.2021.108250]
- Dai A, Chang A X, Savva M, Halber M, Funkhouser T and Nießner M. 2017. ScanNet: richly-annotated 3D reconstructions of indoor scenes//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 2432-2443 [DOI: 10.1109/CVPR.2017.261]
- Du L, Tan J G, Xue X Y, Chen L L, Wen H K, Feng J F, Li J M and Zhang X L. 2020. 3DCFS: fast and robust joint 3D semantic-instance segmentation via coupled feature selection//Proceedings of 2020 IEEE International Conference on Robotics and Automation (ICRA). Paris, France: IEEE: 6868-6875 [DOI: 10.1109/icra40945.2020.9197242]
- He Y, Yu H S, Liu X Y, Yang Z G, Sun W, Anwar S and Mian A. 2025. Deep learning based 3D segmentation in computer vision: a survey. *Information Fusion*, 115: #102722 [DOI: 10.1016/j.inffus.2024.102722]
- Hu Q Y, Yang B, Xie L H, Rosa S, Guo Y L, Wang Z H, Trigoni N and Markham A. 2020. RandLA-Net: efficient semantic segmentation of large-scale point clouds//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 11105-11114 [DOI: 10.1109/CVPR42600.2020.01112]
- Kolodiazhnyi M, Vorontsova A, Konushin A and Rukhovich D. 2024. OneFormer3D: one transformer for unified point cloud segmentation//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 20943-20953 [DOI: 10.1109/CVPR52733.2024.01979]
- Liu S Y, Chi J N, Wu C D and Xu F. 2023. Recurrent slice networks-based 3D point cloud-relevant integrated segmentation of semantic and instances. *Journal of Image and Graphics*, 28(7): 2135-2150 (刘苏毅, 迟剑宁, 吴成东, 徐方. 2023. 基于递归切片网络的三维点云语义分割与实例分割. *中国图象图形学报*, 28(7): 2135-2150) [DOI: 10.11834/jig.220154]
- Ngo T D, Hua B S and Nguyen K. 2023. ISBNet: a 3D point cloud instance segmentation network with instance-aware sampling and box-aware dynamic convolution//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 13550-13559 [DOI: 10.1109/CVPR52729.2023.01302]
- Qi C R, Su H, Kaichun M and Guibas L J. 2017a. PointNet: deep learning on point sets for 3D classification and segmentation//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 77-85 [DOI: 10.1109/CVPR.2017.16]
- Qi C R, Yi L, Su H and Guibas L J. 2017b. PointNet++: deep hierarchical feature learning on point sets in a metric space//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc.: 5105-5114 [DOI: 10.5555/3295222.3295263]
- Royen R, Denis L and Munteanu A. 2024. Joint prototype and coefficient prediction for 3D instance segmentation. *Electronics Letters*, 60(5): #e13137 [DOI: 10.1049/ell2.13137]
- Tan J G, Chen L L, Wang K R, Li J M and Zhang X L. 2021a. SASO: joint 3D semantic-instance segmentation via multi-scale semantic association and salient point clustering optimization. *IET Computer Vision*, 15(5): 366-379 [DOI: 10.1049/cvi2.12033]
- Tan J G, Wang K R, Chen L L, Zhang G H, Li J M and Zhang X L. 2021b. HCFS3D: hierarchical coupled feature selection network for 3D semantic and instance segmentation. *Image and Vision Computing*, 109: #104129 [DOI: 10.1016/j.imavis.2021.104129]
- Vu T, Kim K, Luu T M, Nguyen T and Yoo C D. 2022. SoftGroup for 3D instance segmentation on point clouds//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 2698-2707 [DOI: 10.1109/CVPR52688.2022.00273]
- Wang X L, Liu S, Shen X Y, Shen C H and Jia J Y. 2019a. Associatively segmenting instances and semantics in point clouds//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 4091-4100 [DOI: 10.1109/CVPR.2019.00422]
- Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M and Solomon J M. 2019b. Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 38(5): #146 [DOI: 10.1145/3326362]
- Wu G N, Pan Z Y, Jiang P and Tu C H. 2020. Bi-directional attention for joint instance and semantic segmentation in point clouds//Proceedings of the 15th Asian Conference on Computer Vision. Kyoto, Japan: Springer: 209-226 [DOI: 10.1007/978-3-030-69525-5\_13]
- Wu W X, Qi Z G and Li F X. 2019. PointConv: deep convolutional networks on 3D point clouds//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 9613-9622 [DOI: 10.1109/cvpr.2019.00985]
- Xiang X Y, Wang L, Zong W P and Li G Y. 2023. Point cloud instance segmentation based on attention mechanism KNN and ASIS module. *Journal of Zhejiang University (Engineering Science)*, 57(5):

- 875-882 (项学泳, 王力, 宗文鹏, 李广云. 2023. ASIS模块支持下融合注意力机制KNN的点云实例分割算法. 浙江大学学报(工学版), 57(5): 875-882) [DOI: 10.3785/j.issn.1008-973X.2023.05.003]
- Xu J F, Yang S Y, Li X Z, Tang Y, Hao Y X, Hu L and Chen M. 2024. PDF: a probability-driven framework for open world 3D point cloud semantic segmentation//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 5977-5986 [DOI: 10.1109/CVPR52733.2024.00571]
- Yao L, Wang Y, Liu M Y and Chau L P. 2025. SGFormer: semantic-guided and geometric-enhanced interleaving transformer for 3D instance segmentation. IEEE Transactions on Circuits and Systems for Video Technology, 35(3): 2276-2288 [DOI: 10.1109/TCSVT.2024.3498041]
- Zhang Z H, Yang B, Wang B and Li B. 2023. GrowSP: unsupervised semantic segmentation of 3D point clouds//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 17619-17629 [DOI: 10.1109/CVPR52729.2023.01690]
- Zhang Z K, Ding J, Jiang L, Dai D X and Xia G S. 2024. FreePoint: unsupervised point cloud instance segmentation//Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 28254-28263 [DOI: 10.1109/CVPR52733.2024.02669]
- Zhao L and Tao W B. 2020. JSNet: joint instance and semantic segmentation of 3D point clouds//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA: AAAI Press: 12951-12958 [DOI: 10.1609/aaai.v34i07.6994]
- Zhao L and Tao W B. 2023. JSNet++: dynamic filters and pointwise correlation for 3D point cloud instance and semantic segmentation. IEEE Transactions on Circuits and Systems for Video Technology, 33(4): 1854-1867 [DOI: 10.1109/TCSVT.2022.3218076]
- Zhong M and Zeng G. 2021. Joint semantic-instance segmentation of 3D point clouds: instance separation and semantic fusion//Proceedings of the 25th International Conference on Pattern Recognition. Milan, Italy: IEEE: 6616-6623 [DOI: 10.1109/ICPR48806.2021.9412532]
- Zhu Z J, Zhang R, Bai Y Q, Wang Y E and Sun J M. 2024. Bilateral cross enhancement with self-attention compensation for semantic segmentation of point clouds. Journal of Image and Graphics, 29(8): 2388-2398 (朱仲杰, 张荣, 白永强, 王玉儿, 孙嘉敏. 2024. 结合双边交叉增强与自注意力补偿的点云语义分割. 中国图象图形学报, 29(8): 2388-2398) [DOI: 10.11834/jig.230430]

### 作者简介

郝雯,女,副教授,主要研究方向为点云场景分割和点云地点识别。E-mail: haowensxsf@163.com

左占彬,男,硕士研究生,主要研究方向为点云分割。E-mail: 2211221079@xaut.edu.cn

卢翰森,男,硕士研究生,主要研究方向为点云分割。E-mail: 2241221097@xaut.edu.cn

梁玮,女,讲师,主要研究方向为点云处理和图像压缩。E-mail: wliang@xaut.edu.cn

金海燕,女,教授,主要研究方向为计算机视觉、图像处理和智能信息处理。E-mail: jinhaiyan@xaut.edu.cn

石争浩,男,教授,主要研究方向为机器视觉、医学图像处理及机器学习。E-mail: ylshi@xaut.edu.cn