用于铁路场景语义分割的改进动态图卷积神经网络

王卫东, 刘延, 邱实, 刘贤华, 魏晓, 王劲*

(中南大学土木工程学院 长沙 410075)
(22020163@csu.edu.cn)

摘 要: 针对目前在铁路场景语义分割中存在的数据获取成本高、分割精度低、泛化能力差等问题, 提出了一种基于改进动态图卷积神经网络的铁路场景语义分割方法. 首先利用高分辨率的无人机采集铁路场景的多视角图像, 并通过结构运动恢复与基于面片的多视角立体视觉算法生成铁路场景的三维点云; 然后在动态图卷积神经网络中引入空间注意力模块, 增强网络结构的分割精度与泛化性; 最后通过改进后的图卷积神经网络对预处理后的铁路场景点云完成高精度的语义分割. 实验阶段采用的铁路场景包括桥梁段、路基段与联络线, 共计 11 个区域. 以平均交并比为评价指标, 与动态图卷积神经网络、PointNet++进行对比, 研究结果表明, 基于图像点云训练的改进动态图卷积神经网络对于铁路场景语义分割具有更高的精度, 与动态图卷积神经网络、PointNet++相比, 分割精度分别提高 3.3 个百分点与 6.0 个百分点, 且具有更好的泛化能力.

关键词: 铁道工程; 点云语义分割; 无人机点云; 卷积神经网络
中图法分类号: U212.24 **DOI:** 10.3724/SP.J.1089.2023-00034

Improved Dynamic Graph Convolutional Neural Network for Semantic Segmentation of Railway Scenes

Wang Weidong, Liu Yan, Qiu Shi, Liu Xianhua, Wei Xiao, and Wang Jin*

(*School of Civil Engineering, Central South University, Changsha* 410075)

Abstract: A railway scene semantic segmentation method based on improved dynamic graph convolutional neural network is proposed to address the issues of high data acquisition cost, low segmentation accuracy, and poor generalization ability in current railway scene semantic segmentation. Firstly, high-resolution drones are used to capture multi view images of railway scenes, and 3D point clouds of railway scenes are generated through structural motion recovery and patch based multi view stereo vision algorithms. Then, a spatial attention module is introduced into the dynamic graph convolutional neural network to enhance the segmentation accuracy and generalization of the network structure. Finally, an improved graph convolutional neural network is used to achieve high-precision semantic segmentation of the pre-processed railway scene point cloud. The railway scenes used in the experimental stage include bridge sections, roadbed sections, and connecting lines, totaling 11 areas. Compared with dynamic graph convolutional neural networks and PointNet++, the research results show that the improved dynamic graph convolutional neural network based on image point cloud training has higher accuracy in semantic segmentation of railway scenes, with an average intersection over union ratio as the evaluation indicator. Compared with dynamic graph convolutional neural

收稿日期: 2023-04-20; 修回日期: 2023-10-18. **王卫东**(1971—), 男, 教授, 博士生导师, 主要研究方向为铁路线路规划、土木工程 BIM、地质灾害防灾减灾、轨道交通智能建造; **刘延**(1999—), 男, 硕士研究生, 主要研究方向为铁路线路智能测量; **邱实**(1985—), 男, 教授, 博士生导师, 主要研究方向为计算机信息技术与智能交通基础设施的交叉融合; **刘贤华**(1998—), 男, 硕士研究生, 主要研究方向为神经网络的改进与应用; **魏晓**(1994—), 男, 博士研究生, 主要研究方向为铁路隧道的智能监测; **王劲**(1988—), 男, 博士, 论文通信作者, 主要研究方向为铁路基础设施参数化建模.

neural networks and PointNet++, the segmentation accuracy is improved by 3.3 percentage points and 6.0 percentage points, respectively, and it has better generalization ability.

Key words: railway engineering; semantic segmentation of point clouds; UAV point cloud; convolutional neural network

随着我国铁路路网规模的逐步扩大, 铁路系统运维管理的复杂度与现实要求也显著提升, 从大量复杂数据中准确地提取关键信息, 对铁路设施进行有效的管理和维护是铁路行业面临的一个挑战. 三维点云以数字化的方式对复杂场景真实三维信息进行记录与呈现, 具有成本低、效率高且适用性强等优势, 大量研究已经证明三维点云在工程领域具有更为广阔的应用前景[1-5]. 在铁路场景中, 快速地获取点云数据并实现主要构件的精确分割是参数检测、资产管理的基础. 例如, 李维刚等[3]利用点云的坐标信息实现不同区域的轨面与枕木的自动检测, 实现铁路轨道的自动化巡查; Lin 等[4]通过车载扫描仪采集高速铁路接触网系统的点云数据, 并基于深度学习理论与方法对不同场景的接触网部件进行识别分割, 实现对接触网系统的检测管理; Lee 等[5]采用基于图分层的图卷积神经网络(graph-based hierarchical convolutional neural network, HGCNN)模型实现对铁路桥梁三维点云的语义分割, 将扫描的得到的点云数据分类为桥梁组件, 然后将其导出并重构为建筑信息模型(building information model, BIM), 以进行资产管理和高效维护工作.

传统的点云获取设备如车载雷达和测站式扫描仪容易受到地形与天窗时间的限制, 存在采集范围小、密度不均匀、数据处理难度大等问题, 这极大制约了点云在铁路领域的应用. 相较于传感器采集点云的方式, 无人机图像点云具有采集效率高、适用范围广、处理难度低等优点, 在工程领域中获得了广泛的应用[6-9]. 例如, Pan 等[6]针对无人机图像采用结构运动恢复和多视图立体视觉(structure from motion, multi-view stereo, SFM-MVS)算法生成古桥的密集点云, 并提出了一种基于超体素结构和全局图优化的点云分割方法, 将整个点云分割成一致的分割段, 利用泊松曲面重建算法建立结构表面模型. Sahebdivani 等[7]通过 SFM-MVS 算法生成铁路轨道的点云数据, 并结合随机样本一致性(random sample consensus, RANSAC)算法与拟合傅里叶曲线, 创建轨道的连续三维模型, 并对轨道中心线的检测进行了研究. Gabara 等[8]提出了一种基于图像点云的铁路轨道几何参数测量方法, 其通过多视图立体与半全局匹配技术重建三维点云, 并比较从图像点云中获得的结果与直接大地测量点云的结果, 其结果满足铁路规范.

然而目前针对铁路场景的分割大多基于启发式算法[10-11], 包括 RANSAC、区域生长等算法, 均严重依赖几何规则和约束条件下的先验知识, 其提取特征能力和分割精度均不足, 并且只适用于元素类别较少和结构特征明显的分割任务, 泛化能力差. 随着深度学习技术在铁路基础设施上的应用日渐成熟[12-13], 基于深度学习的点云分割方法比启发式算法能更有效地提取点云的深层特征, 对复杂点云场景具有更强的适应性.

目前关于点云语义分割的方法大致可分为 3 类[14]: 基于投影的方法、基于体素化的方法、基于点的方法. 基于投影的方法对生成图像的视角敏感, 无法充分地利用点云的完整几何形状; 基于体素化的方法需要大量的内存, 因此不适用于大规模点云; 基于点的方法可直接使用点云的坐标, 解决了将卷积神经网络(convolutional neural network, CNN)直接应用到原始点云的难题. 以 PointNet[15]和 PointNet++[16]为代表的基于深度学习的方法, 在 ShapeNet part 数据集上的分割均交并比(mean intersection over union, mIoU)分别达到了 83.7%和 85.1%. Grandio 等[14]基于 PointNet++框架对 4 个铁路场景进行语义分割, 其中对车载雷达获取的点云场景的分割平均准确率达到了 90.0%以上, 但在无人机图像生成的点云场景中表现较差, 这使得无人机图像点云的高效采集与深度学习的高精度分割无法同时兼顾.

不同于传统的点网络 PointNet, 动态图卷积神经网络(dynamic graph CNN, DGCNN)[17]并不是直接生成嵌入点的特征. 其生成描述中心点与邻域点之间关系的边缘特征, 并在网络的每一层动态地构建图结构, 将每一点作为中心点来表征其与各个邻近点的边缘特征; 再将这些特征聚合得到该点的新特征; 最后将多个边缘卷积模块(edge convolution, EdgeConv)首尾连接, 通过不断地堆叠获得多层次的表征. 然而边缘特征是通过构造某

点的局部邻域图来提取局部特征，在建立图结构时，将整个点云划分为一个独立的分组，在一定程度上忽略了点云之间的相关性，对全局特征提取有所欠缺. 空间注意力模块(spatial attention module，SAM)允许点云中每个点与其他点进行交互，而不必局限于局部邻域，这使得DGCNN可以从全局信息中学习更加丰富和全面的特征表达.

基于此，本文提出了一种基于改进动态图卷积神经网络的点云分割方法，在获取低成本、高效率采集图像点云的基础上，同时实现铁路场景的高精度分割. 该方法首先根据铁路线路这一广域线性结构的自身特征，综合考虑了基础点云准确度、完整性与可识别性等三大因素，制定了针对性的图像数据采集方案. 随后，结合结构运动恢复(structure from motion，SFM)算法与基于面片的多视角立体视觉算法(patch-based multi-view stereo，PMVS)[18]生成点云数据，通过上述方法获取的点云数据在采集成本、点云覆盖率、完整性等方面明显优于传统的脉冲激光点云. 随后，对基础点云执行滤波去噪、均匀平滑、特征划定等预处理，进一步提升基础点云的质量. 最后，采用改进后的空间注意力动态图卷积神经网络(SAM-DGCNN)进行语义分割，与基于启发式规则的分割算法相比，SAM-DGCNN 具有更高的精度以及更强的泛化能力；与 DGCNN 以及 PointNet++相比，具有更高的分割精度.

1　研究方法

1.1　基本框架

本文方法包括基于多视角图像的点云数据生成和基于 SAM-DGCNN 的铁路场景语义分割. 首先基于无人机技术采集铁路场景的多视角图像，依托于采集的图像数据，通过 SFM-PMVS 算法生成点云数据；然后基于点云数据充分训练 SAM-DGCNN，以实现对铁路场景的语义分割，并与不同的网络模型进行对比. 本文方法的基本框架如图 1 所示.

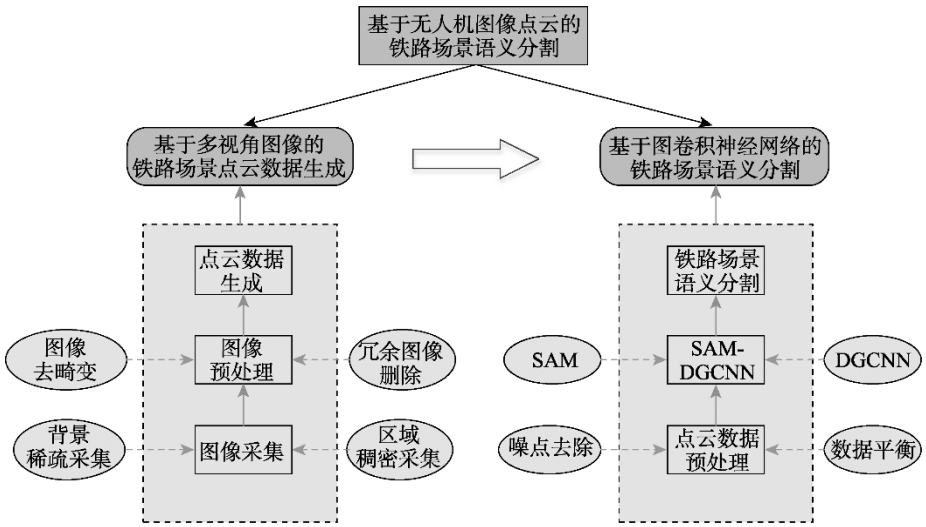


图 1　铁路场景语义分割总体框架

1.2　基于无人机获取多视角铁路场景点云数据

如图 2 所示，基于无人机的多视角铁路场景点云数据获取包括无人机图像采集、图像预处理、点云数据生成 3 个步骤.

Step1. 设定合适的无人机航高、相机角度以及重叠率等航摄参数，随后以预设的航线规划采集铁路场景的倾斜影像，同时在测量范围内布设像控点，以便建立统一的高程基准.

Step2. 为提高计算效率，针对采集图像中重叠率过高的冗余图像进行删减，同时对出现畸变的图像进行优化.

Step3. 分别采用 SFM 算法和 PMVS 算法进行稀疏重建和稠密重建.

1.3　基于 SAM-DGCNN 的铁路场景语义分割

本文在采用 DGCNN 提取边缘特征的基础上，引入 SAM 来捕捉点间的语义特征[19-22]，其网络架构如图 3 右下部分所示. 对于给定的特征向量 $\boldsymbol{A}\in\mathbb{R}^{C\times H\times W}$，将 $\boldsymbol{A}$ 分别输入 2 个独立 1×1 卷积层降维得到 $\boldsymbol{B}^{C\times H\times W}$ 和 $\boldsymbol{C}^{C\times H\times W}$，将 $\boldsymbol{C}$ 转置并与 $\boldsymbol{B}$ 矩阵相乘，归一化后得到空间注意力权重矩阵

$$\boldsymbol{U}_{ij}=\exp\left(\boldsymbol{B}_i\cdot\boldsymbol{C}_j^{\mathrm{T}}\right)\Big/\sum_{i=1}^{N}\exp\left(\boldsymbol{B}_i\cdot\boldsymbol{C}_j^{\mathrm{T}}\right).$$

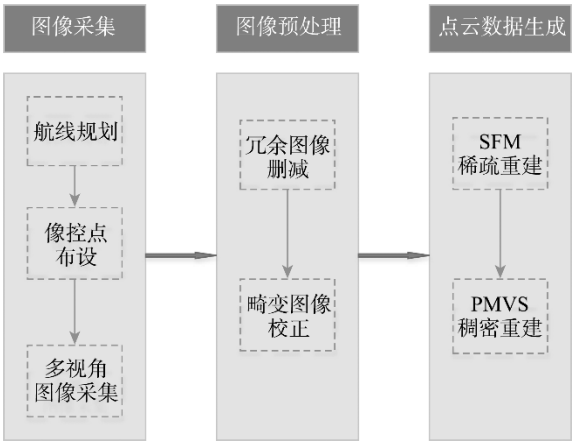


图 2 多视角铁路场景点云数据获取

$\boldsymbol{U}_{ij} \in \mathbb{R}^{N\times N}$，它衡量的是第 i 个点与第 j 个点之间的相似性. 其中，$\boldsymbol{B}_i$ 和 $\boldsymbol{C}_j$ 分别是点 i 和 j 在 $\boldsymbol{B}$ 和 $\boldsymbol{C}$ 中的特征表示. 对特征向量 $\boldsymbol{A}$ 应用卷积操作得到新的特征向量 $\boldsymbol{D}^{C\times H\times W}$，并将其与空间注意力权重矩阵 $\boldsymbol{U}_{ij}$ 进行矩阵相乘操作；最后与尺度参数 α 相乘，并且与输入特征 $\boldsymbol{A}$ 进行逐元素求和操作，得到增强的输出特征 $\boldsymbol{F}=\alpha\sum_{i=1}^{N}(\boldsymbol{U}_{ij}\boldsymbol{D}_j)+\boldsymbol{A}_i$. 其中，尺度参数 α 的初始值为 0.

SAM 通过自适应地调整权重来区分不同特征的重要程度，从而实现对不同特征的加权处理. 这使得改进后网络更加关注铁路等狭长结构物的边缘和细节等重要部分，避免了传统方法中容易出现的错分和漏分等问题.

改进后的网络架构图如图 3 所示，包括 4 个边缘卷积层，4 个 SAM 和 3 个全连接层；每一层由边缘卷积模块提取的局部几何特征和利用 SAM 提取点间的语义特征组成. 融合上一层提取的局部几何特征与点间语义特征，并将其作为下一层边缘卷积模块的输入，(·)中的数字代表特征的输出维度. 采用多层感知器(multi-layer perceptron, MLP)提取各层的特征并进行拼接融合，将融合特征输入到一个共享的全连接层，通过最大池化操作得到一个 1 024 维全局特征描述符；最后使用 3 个全连接层(512, 256, s)得到目标所属类别的 s 维向量.

基于 SAM-DGCNN 分割铁路场景，主要包括点云预处理、网格采样、神经网络分割、输出预测结果等步骤，基本流程如图 4 所示.

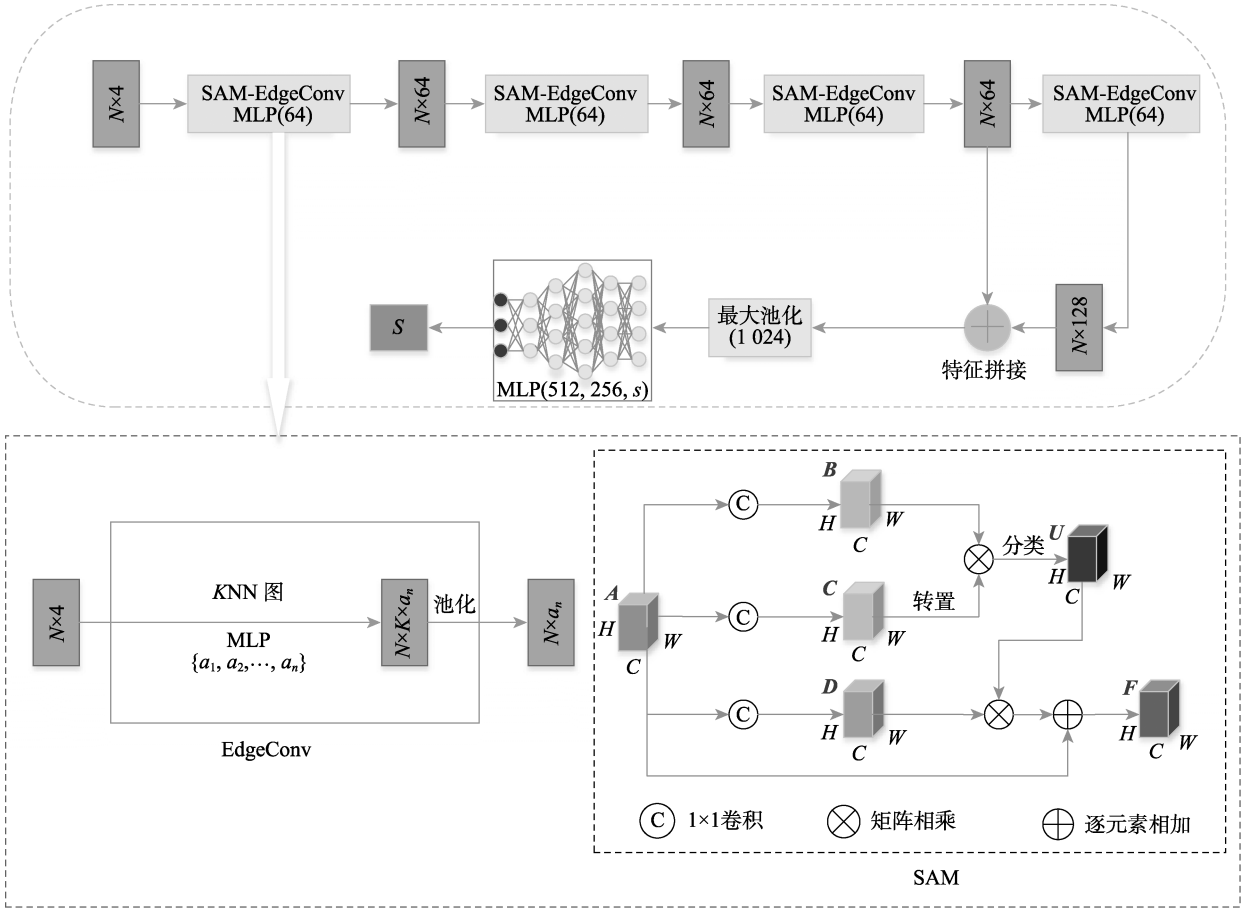


图 3 SAM-DGCNN 架构

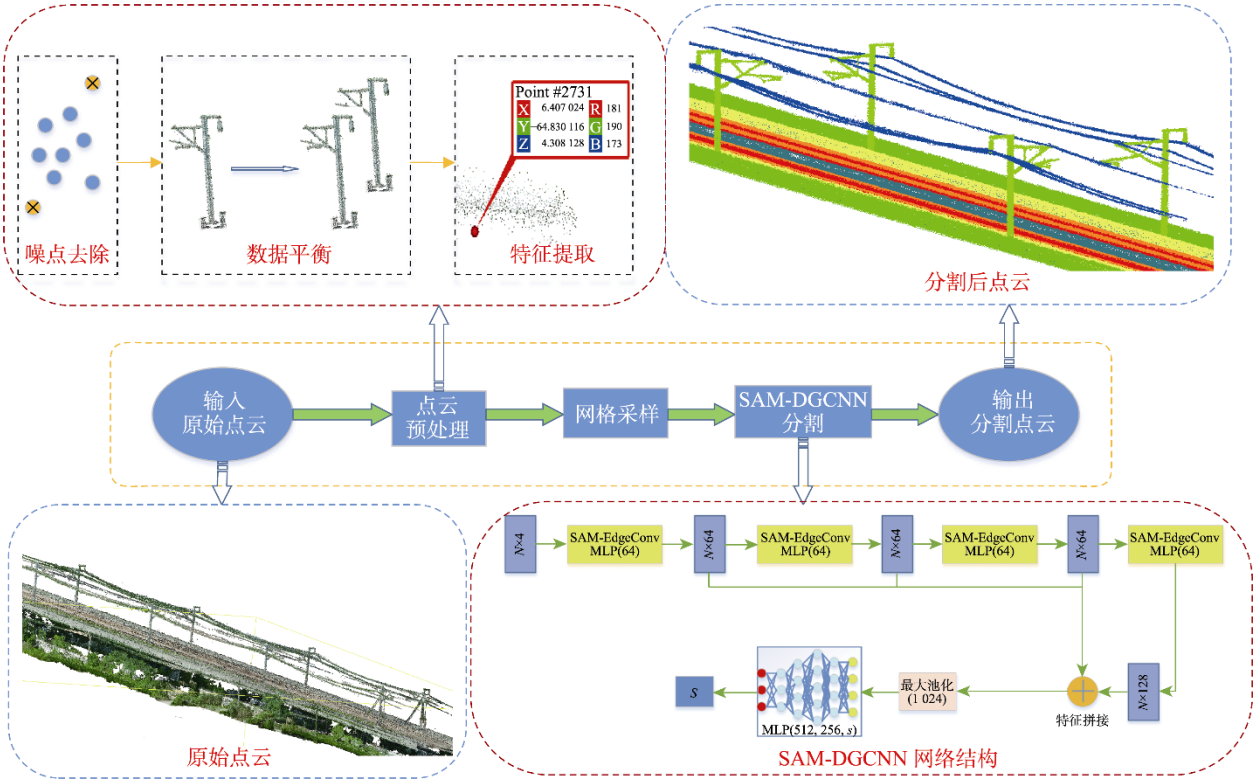


图 4　铁路场景三维点云分割流程

2　案例实现

2.1　点云数据获取

本文采用大疆无人机(Phantom RTK 版)进行铁路场景图像采集，在无人机底部安装的 3 轴云台俯仰角度可达 120°，可用于较大范围的图像采集. 在实际拍摄前，首先通过测试设置合理的无人机最佳飞行高度、相机角度、航向重叠率与旁向重叠率等参数，以保证飞行安全以及铁路场景数据的完整性，无人机航摄参数如表 1 所示.

表 1　无人机航摄参数

类型	参数	值
相机参数	照片最大分辨率/像素	5 472 : 3 648 (3 : 2)
	照片格式	JPEG
	焦距/mm	35
飞行参数	飞行高度/m	50
	飞行速度/(m·s^{-1})	3
	正向重叠率/%	75
	侧向重叠率/%	70
	拍摄角度/(°)	45

为了降低计算机的计算成本，首先针对采集后的图像进行预处理，人工删除重复以及模糊的图像. 在此基础上，结合 SFM 算法与 PMVS 算法获得铁路场景的点云数据，具体流程如图 5 所示. 首先利用 SIFT 算法进行铁路场景特征点的提取；其次，为提高匹配效率，采用 *K*D-Tree 对特征点进行匹配；然后利用光束法区域平差空中三角测量，获得稀疏点云模型；最后采用 PMVS 算法构建面片，通过面片的迭代扩散和过滤，将稀疏点云扩散为稠密点云.

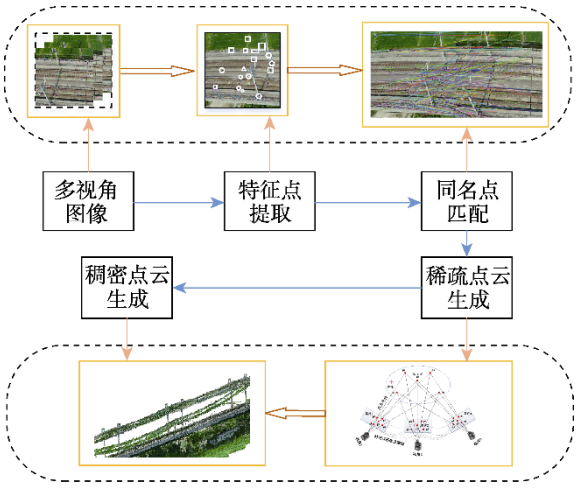
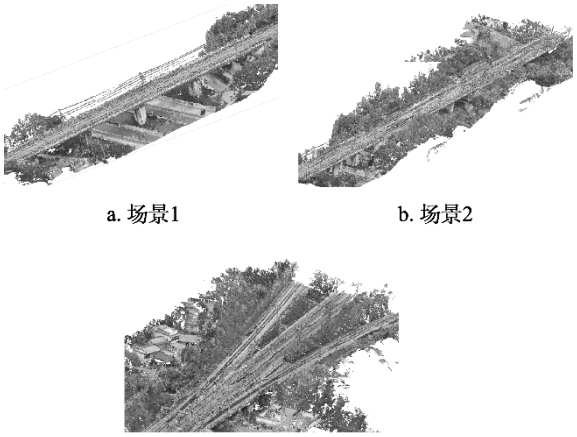


图 5　铁路场景三维点云构建流程

本文采集的 3 个实验场景如图 6 所示. 场景 1 为一段直线铁路桥梁，结构类型较为单一；场景 2 结构较复杂，包括 3 个桥梁段和 2 个路基段；场景



a. 场景1 b. 场景2

c. 场景3

图 6 3 个场景的图像点云数据

3 在正线的基础上, 增加了 2 段联络线, 联络线为有砟轨道.

为方便网络模型的训练和测试, 将 3 个场景的区域进行了划分. 对于场景 1, 根据线路长度将其划分为 3 个区域; 对于场景 2, 根据不同的轨道基础支撑结构将其划分为 5 个区域; 对于场景 3, 根据 3 条不同线路将其划分为 3 个区域. 在模型训练和测试中, 将单个场景中的其中一个区域作为测试集, 其余区域作为训练集.

考虑选择单一场景训练模型容易导致泛化能力不足的情况, 本文在 3 个场景的基础上, 增设了一个多场景融合的数据集, 即分别从 3 个场景中各取 50%的点云数据用于训练, 其余用于测试.

2.2 数据预处理

图像点云生成后, 还需要对点云数据预处理, 预处理的步骤包括噪点去除、数据平衡与特征划定 3 个步骤.

Step1. 噪点去除. 生成的图像点云中存在大面积的无用背景点以及大量的离散点和孤立点, 因此需要进行噪点去除. 本文首先采用直通滤波去除远离铁路场景的大面积背景噪点, 之后再采用统计滤波来去除靠近基础设施的邻近噪声点以及远离基础设施的离散点.

Step2. 数据平衡. 在本文中, 用于训练神经网络的数据集共区分了 6 个类别, 分别是背景、钢轨、接触柱、接触网、轨道板和护栏. 未进行数据平衡前, 数据集中不同类别的点数是不平衡的, 背景点云占了很高的比例, 接触柱、接触网和钢轨等其他类别只占少量, 导致神经网络只会从大比例类别中学习特征, 而忽略其他类别, 最终影响网络的分割精度. 针对此不足, 本文采用数据增强的方式来平衡数据. 首先针对在训练集中占比大的类别, 通过裁剪和均匀降采样的方式来降低其在数据集中所占的比例; 其次针对在训练集中占比较小的类别, 通过复制、旋转和小范围裁剪等几何变换增大其所占比例, 最终使各个类别比例接近, 达到最佳的训练结果. 以融合数据集为例, 平衡前后的点云数据如表 2 所示.

表 2 不同类别的点云数量

数据平衡处理	全部点云	背景	钢轨	接触柱	接触网	轨道板	护栏
数据平衡前	45 768 579	34 490 377	2 218 148	1 814 548	1 365 889	5 275 988	603 629
数据平衡后	36 456 083	18 235 435	4 248 653	4 865 417	3 579 641	3 842 167	1 684 752

Step3. 特征划定. 由于点云可能具备多个属性来表征每个点, 这增加了网络训练的计算成本, 因此需要选择点云的部分属性作为特征输送到网络中. 本文中神经网络考虑了三维坐标与强度值, 并将其作为输入特征向量

$$\boldsymbol{X}_{n\times 4}=\begin{bmatrix} x_1 & y_1 & z_1 & I_1 \\ \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & z_n & I_n \end{bmatrix}.$$

其中, n 表示输入网络的点的数量; (x, y, z)表示每个输入点的坐标; I 表示每个输入点的强度值.

2.3 网络训练

在预处理步骤完成后, 对搭建的网络进行训练. 本文以 GPU 为计算核心(GPU: NVIDIA GeFore RTX 3090; CPU: AMD2990WX@3.0 GHz, RAM= 64 GB), 依托于 GitHub 开源的软件 PyTorch1.2.0, 通过交叉验证进行了超参数调整实验. 在训练数据集上获得了改进 DGCNN 的最佳超参数, 如表 3 所示.

表 3 实验超参数设置

超参数	设置
网络结构	SAM-DGCNN
初始学习率	0.001 25
指数衰减速率	0.9
参数更新算法	适应性动量估计算法
批次大小	8
训练轮次	300

3 结果对比与评估

3.1 评价指标

为了更好地评价图卷积神经网络对铁路场景的分割效果, 本文采用所有类别平均交并比的平均值(mean intersection over union, mIoU)、每个类别的平均精度(average precision, AP)作为评价指标.

mIoU 是目前最常用的点云语义分割评价指标

之一，它能够综合考虑各个类别的分类准确度，是比较网络性能的重要指标之一．其计算的关键在于 IoU，IoU 被定义为每个类别预测结果与真实结果的交集与并集之比，计算公式为

$$\mathrm{IoU}=\frac{\text{预测分割点}\cap\text{真实分割点}}{\text{预测分割点}\cup\text{真实分割点}}.$$

在铁路场景中，正确识别并分类出某些特定组件(如钢轨、接触网等)具有特殊的意义. AP 可以针对特定类别的分类效果进行评价分析，其计算的关键在于类别精确率(class precision，P)和类别召回率(class recall，R)，精确率与召回率的数学计算公式分别为精确率与召回率的计算公式分别为

$$P=\frac{\mathrm{TP}}{\mathrm{TP+FP}};$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP+FN}}.$$

其中，TP 是预测为正样本，实际为正样本的点个数；FP 是预测为正样本，实际为负样本的点个数；TN 预测为负样本，实际为负样本的点个数；FN 是预测为负样本，实际为正样本的点个数．

通过计算在不同置信度阈值下 SAM-DGCNN 在测试集上的精确率与召回率，绘制 P-R 曲线，对于每个召回率，找到其对应的最大精确率；然后将所有召回率水平下的最大精确率值相加并除以召回率的数量，得到该类别的 AP 值．

3.2　不同 K 值下的分割效果对比

在进行网格采样之后，会根据 KNN 算法对点云进行二次采样，不同的 K 值会导致不同的分割结果．K 值太小，则点包含的信息不足，导致精度不足；反之会增加计算成本、降低精度．因此，为了得到铁路场景的最佳分割效果，本文采用融合数据集，设置对照实验来寻找最佳的网格尺寸以及 K 值．图 7 给出了不同 K 值下 SAM-DGCNN 的分割效果评估．

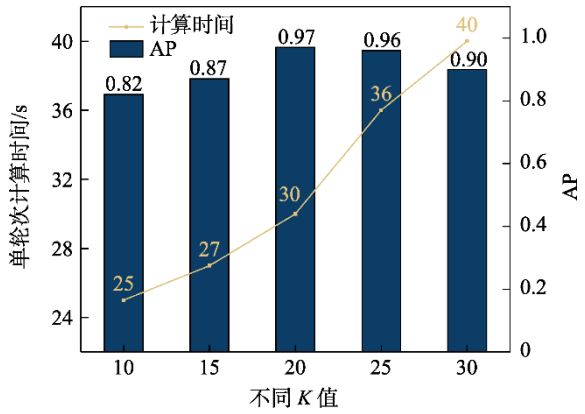


图 7　不同 K 值下的改进网络分割效果评估

从图 7 可以看出，增大 K 值通常会提高网络模型的 AP．这是因为较大的 K 值能考虑更多的局部信息，从而帮助网络模型更好地理解点云中的结构和形状．但随着 K 值的持续增大，AP 却又呈现下降的趋势；此时模型会过度关注远离中心点的其他点，使得噪声及不必要的信息被纳入网络模型中，从而降低 AP．此外，较大的 K 值也会使模型更加复杂，增加训练时间和内存开销．在本文中，在综合权衡 AP 以及计算成本的前提下，选择 K=20 作为改进网络的参数值．

3.3　SAM-DGCNN 分割效果评估

本文在充分训练改进网络的基础上，基于测试集验证了分割结果．图 8 给出了改进网络对于 3 个铁路场景不同类别的分割结果．

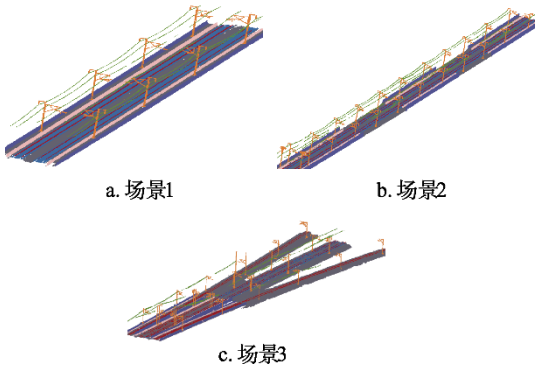


图 8　SAM-DGCNN 对 3 个场景的分割结果

在图 9 中，进一步给出了 SAM-DGCNN 对铁路场景中不同类别相应的 IoU 值．从场景 1 到场景 3，场景复杂度在增加，相对应的分割精度也在逐步降低，然而本文算法对于 3 个不同场景的点云仍保持较高的分割精度，这也证明了算法的泛化能力．从图 9 中可以看出，在 3 个场景中，护栏的分

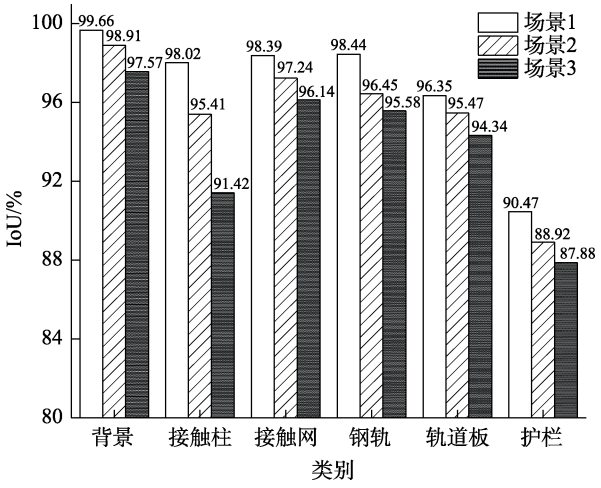


图 9　不同类别分割结果的 IoU 对比

割精度明显低于其他类别, 这是因为该类别在数据集中存在的数量太少, 从而分割中的小错误会显著地降低 IoU 值.

3.4 与不同深度学习分割算法的对比分析

为了进一步验证改进网络分割铁路场景的优异性, 本文采用融合数据集, 并以 mIoU 为评价指标, 与现有的 2 种点云分割网络进行了对比. 图 10 给出了 SAM-DGCNN, PointNet++和 DGCNN 分割铁路场景不同类别的效果评估.

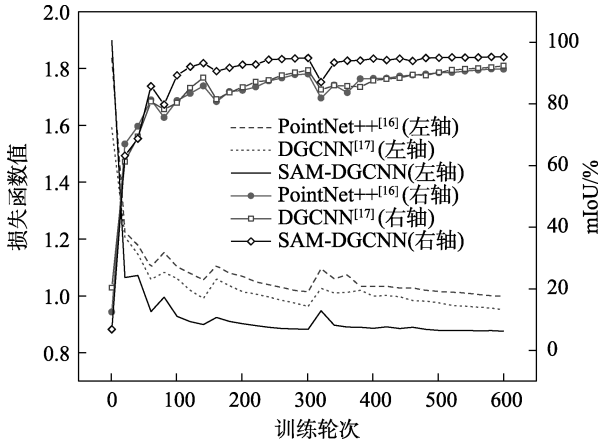


图 10 不同网络在测试集上的分割效果评估

从图 10 中可以看出, 经过近 600 个轮次的充分训练, 3 个网络的总损失均达到最低且收敛, 表明 3 个网络的训练状况均处于良好的拟合状态, 此时 SAM-DGCNN 的 mIoU 达到峰值, 为 95.88%; 比 DGCNN 高 3.3 个百分点左右, 比 PointNet++高 6.0 个百分点左右.

为了进一步验证本文算法的可行性, 将其与其他深度学习分割算法针对不同类别的 IoU 值进行了对比, 如表 4 所示. 直接的点云深度学习分割算法(PointNetSIFT[23], PointCNN[24]和 PointConv[25])与以 DGCNN 为基础框架的动态区域图卷积神经网络(dynamic region graph convolution neural network, DRGCNN)[26]相比, SAM 模块通过自适应地调整权重来区分不同特征的重要程度, 从而获得更好的特征表达, 进而提高整体分割精度. 除背景外, SAM-DGCNN 对于铁路的关键设施(钢轨、接触柱、接触网、轨道板)的 IoU 值均超过了 95%, 即使在数据集中占比最小的护栏, SAM-DGCNN 的分割精度也能达到 90%左右. 综上所述, SAM-DGCNN 对于铁路场景的语义分割能够得到较为精准的分割结果.

表 4 本文算法与其他算法分割铁路场景的 IoU 与 mIoU 对比 %

算法	mIoU	IoU					
		背景	钢轨	接触柱	接触网	轨道板	护栏
SAM-DGCNN	95.88	99.89	98.48	97.49	95.34	94.52	89.54
PointNetSIFT[23]	90.87	99.78	93.54	92.31	89.65	88.67	81.27
PointCNN[24]	89.79	98.89	89.58	92.54	88.96	87.54	81.25
PointConv[25]	90.10	98.42	90.36	92.23	87.54	89.69	82.36
DRGCNN[26]	93.55	99.78	93.68	97.25	91.06	90.36	89.18

4 结 论

本文提出了一种基于无人机图像点云的三维语义地图构建方法, 基于无人机采集的多视角图像与 SFM-PMVS 算法生成了铁路场景的三维点云数据, 并对生成的点云数据进行预处理, 并充分地训练 SAM-DGCNN, 以完成对铁路场景的语义分割, 从而构建三维语义地图, 进而有效地解决了传统语义地图构建过程中存在的获取成本高、采集难度大、分割精度低等问题.

经过充分训练的 SAM-DGCNN 获得了对不同类别的最佳 mIoU 值, 与以 PointNet++或 DGCNN 为基础框架的网络模型相比, 它在铁路场景重要设施如钢轨、接触柱、接触网、轨道板等类别上的 IoU 值具有更高的精度, 证明 SAM-DGCNN 更加适用于铁路场景的语义分割.

本文仅讨论了改进网络的泛化能力, 即对铁路场景多种设施更高精度的语义分割, 但并未对基础设施进行更为精细的划分, 如识别钢轨类别中的道岔、接触柱类别中的绝缘子等. 道岔识别对于提取铁路中心线有着重要的意义, 本文后续将继续优化神经网络, 以期可以识别道岔并进一步完成对于铁路中心线的提取工作.

参考文献(References):

[1] Zhao Ruibin, Pang Mingyong, Zhang Yanling, *et al*. Progressively extracting accurate building roofs from airborne LiDAR data[J]. Journal of Computer-Aided Design & Computer Graph-

ics, 2017, 29(04): 624-631(in Chinese)
(赵瑞斌, 庞明勇, 张燕玲, 等. 机载 LiDAR 点云中精细建筑物顶面的渐进提取[J]. 计算机辅助设计与图形学学报, 2017, 29(04): 624-631)

[2] Shen Chaohui, Huang Shisheng, Hu Shimin. Adaptive partitioning of urban facades[J]. Journal of Computer-Aided Design & Computer Graphics, 2012, 24(02): 149-151(in Chinese)
(沈超慧, 黄石生, 胡事民. 城市建筑点云的自适应分割方法[J]. 计算机辅助设计与图形学学报, 2012, 24(02): 149-151)

[3] Li Weigang, Mei Yang, Fan Xiang, *et al*. Railway track detection based on vehicle laser point cloud[J]. Chinese Journal of Lasers, 2022, 49(04): 168-179(in Chinese)
(李维刚, 梅洋, 樊响, 等. 基于车载激光点云的铁路轨道检测[J]. 中国激光, 2022, 49(04): 168-179)

[4] Lin S, Xu C, Chen L P, *et al*. LiDAR point cloud recognition of overhead catenary system with deep learning[J]. Sensors, 2020, 20(8): Article No.2212

[5] Lee J S, Park J, Ryu Y M. Semantic segmentation of bridge components based on hierarchical point cloud model[J]. Automation in Construction, 2021, 130: Article No.103847

[6] Pan Y, Dong Y, Wang D, *et al*. Three-dimensional reconstruction of structural surface model of heritage bridges using UAV-based photogrammetric point clouds[J]. Remote Sensing, 2019, 11(10): Article No.1204

[7] Sahebdivani S, Arefi H, Maboudi M. Rail track detection and projection-based 3D modeling from UAV point cloud[J]. Sensors, 2020, 20(18): Article No.5220

[8] Gabara G, Sawicki P. A new approach for inspection of selected geometric parameters of a railway track using image-based point clouds[J]. Sensors, 2018, 18(3): Article No.791

[9] Xi Zhipeng, Lou Zhuo, Li Xiaoxia, *et al*. Vision-based localization and navigation for UAV inspection in photovoltaic farms [J]. Journal of Zhejiang University(Engineering Science), 2019, 53 (05): 880-888(in Chinese)
(席志鹏, 楼卓, 李晓霞, 等. 集中式光伏电站巡检无人机视觉定位与导航[J]. 浙江大学学报: 工学版, 2019, 53(05): 880-888)

[10] Guo Baoqing, Yu Zujun, Zhang Nan, *et al*. 3D point cloud segmentation, classification and recognition algorithm of railway scene[J]. Chinese Journal of Scientific Instrument, 2017, 38(09): 2103-2111(in Chinese)
(郭保青, 余祖俊, 张楠, 等. 铁路场景三维点云分割与分类识别算法[J]. 仪器仪表学报, 2017, 38(09): 2103-2111)

[11] Sun Senzhen, Lu Xiaoping, Yang Shangbo, *et al*. A method for extraction of the feature elements point clouds based on spatial grid[J]. Engineering of Surveying and Mapping, 2015, 24(07): 53-57(in Chinese)
(孙森震, 卢小平, 杨尚波, 等. 一种基于空间网格的地物要素点云提取方法[J]. 测绘工程, 2015, 24(07): 53-57)

[12] Guo Y, Wang H, Hu Q, *et al*. Deep learning for 3D point clouds: A survey[J]. IEEE transactions on pattern analysis and machine intelligence, 2020, 43(12): 4338-4364

[13] Hu W b, Wang W D, Ai C B, *et al*. Machine vision-based surface crack analysis for transportation infrastructure. Automation in Construction, 2021, 132: Article No.103973

[14] Grandio J, Riveiro B, Soilán M, *et al*. Point cloud semantic segmentation of complex railway environments using deep learning[J]. Automation in Construction, 2022, 141: Article No.104425

[15] 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. Los Alamitos: IEEE Computer Society Press, 2017: 77-85

[16] Qi C R, Li Y, Su H, *et al*. PointNet++: deep hierarchical feature learning on point sets in a metric space[C] // Proceedings of the 31st International Conference on Neural Information Processing Systems. La Jolla: NIPS Press, 2017: 5105-5114

[17] Wang Y, Sun Y, Liu Z, *et al*. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics[J], 2019, 38(5): 1-12

[18] Furukawa Y, Ponce J. Accurate, dense, and robust multiview stereopsis[J]. IEEE transactions on pattern analysis and machine intelligence, 2010, 32(8): 1362-1376

[19] Fu J, Liu J, Tian H, *et al*. Dual attention network for scene segmentation[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Los Alamitos: IEEE Computer Society Press, 2019: 3146-3154

[20] Yu Lili, Yu Haiyang, He Zixin, *et al*. Point cloud scene segmentation based on dual attention mechanism and multi-scale features[J]. Progress in Laser and Optoelectronics, 2021, 58 (24): 471-479(in Chinese)
(于丽丽, 于海洋, 何子鑫, 等. 基于双注意力机制和多尺度特征的点云场景分割[J]. 激光与光电子学进展, 2021, 58(24): 471-479)

[21] Yang Jun, Zhang Chen. Semantic segmentation of 3D point cloud by fusing dual attention mechanism and dynamic graph convolution neural network[J/OL]. Journal of Beijing University of Aeronautics and Astronautics: 1-12[2023-4-20]. https://doi.org/10.13700/j.bh.1001-5965.2022.0775(in Chinese)
(杨军, 张琛. 融合双注意力机制和动态图卷积神经网络的三维点云语义分割[J/OL]. 北京航空航天大学学报: 1-12 [2023-4-20]. https://doi.org/10.13700/j.bh.1001-5965.2022.0775)

[22] Hao Wen, Wang Hongxiao, Wang Yang. Semantic segmentation of three-dimensional point cloud based on spatial attention and shape feature[J]. Progress in Laser and Optoelectronics, 2022, 59(08): 512-522(in Chinese)
(郝雯, 王红霄, 汪洋. 结合空间注意力与形状特征的三维点云语义分割[J]. 激光与光电子学进展, 2022, 59(08): 512-522)

[23] Jiang M, Wu Y, Zhao T, *et al*. Pointsift: a sift-like network module for 3d point cloud semantic segmentation[OL]. [2023-4-20]. https:// doi.org/10.48550/arXiv.1807.00652

[24] Li Y Y, Bu R, Sun M C, *et al*. PointCNN: convolution on X-transformed points[OL]. [2023-4-20]. https://doi.org/10.48550/arXiv.1801.07791

[25] Wu W X, Qi Z G, Li F X, *et al*. PointConv: deep convolutional networks on 3D point clouds[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Los Alamitos: IEEE Computer Society Press, 2019: 9613-9622

[26] Yue C K, Wang Y, Tang X T, *et al*. DRGCNN: dynamic region graph convolutional neural network for point clouds[J]. Expert Systems with Applications, 2022, 205: Article No.117663