

整合集成预测约束与错误预测熵最大化的 MLS点云分类方法

雷相达¹, 管海燕¹, 董震²

1. 南京信息工程大学 遥感与测绘工程学院, 南京 210044;
2. 武汉大学 测绘遥感信息工程国家重点实验室, 武汉 430079

摘要: 许多深度学习点云分类方法通过增加点云特征聚合模块, 增强点云特征的表达能力, 但该类方法往往会带来训练参数增加以及模型过拟合的问题。针对该问题, 本文提出了一个整合集成预测约束与错误预测熵最大化的深度学习用于移动激光扫描MLS (Mobile Laser Scanning) 点云分类。方法通过集成预测约束分支以及错误预测熵最大化分支可以在不增加训练参数的情况下, 增强基线网络的点云特征表达, 提高模型泛化能力。其中集成预测约束分支首先通过记录点云在训练过程中的预测值, 生成集成预测值, 然后采用一致性约束增强模型的点云特征表达。错误预测熵最大化分支鼓励模型对错误预测点进行熵值最大化, 增加该点的不确定性, 提高模型的泛化能力。所提方法在多个公开MLS点云数据集上进行验证, 结果表明所提方法可以在不增加训练参数的情况下, 提高基线方法的分类性能。与对比方法相比, 所提方法在Toronto3D、WHU-MLS、Paris数据集上获得了最优的平均交并比(83.68%、65.85%、44.19%), 表明了方法的有效性。

关键词: 遥感, MLS点云分类, 深度学习, 集成预测约束, 错误预测熵最大化

中图分类号: P237/P2

引用格式: 雷相达, 管海燕, 董震. 2025. 整合集成预测约束与错误预测熵最大化的MLS点云分类方法. 遥感学报, 29(1): 329-340
Lei X D, Guan H Y and Dong Z. 2025. Integrating ensemble prediction constraints and error prediction entropy maximization for MLS point cloud classification. National Remote Sensing Bulletin, 29(1): 329-340 [DOI: 10.11834/jrs.20233174]

1 引言

数字孪生城市是新一代信息技术在城市的综合集成应用, 是实现数字化治理和发展数字经济的重要载体。作为构建数字孪生城市的基础数据, 城市场景的语义信息是实现城市要素数字化、连接物理世界与数字世界, 建立城市物理空间与数字空间映射的关键, 可为数字孪生城市的可视化、仿真模拟以及智能分析与决策等提供数据支撑(中国信息通信研究院, 2020)。目前传统的测绘手段(如全站仪或GNSS等)外业劳动强度大、效率低, 无法满足城市场景语义信息获取的需求。

移动激光扫描MLS (Mobile Laser Scanning) 系统能够快速获取大规模场景高精度高密度的三

维激光点云数据(谭贲等, 2012), 具有采集速度快、定位精度高以及点云密度大等优势(杨必胜等, 2021)。其获取的高精度、高密度点云可以直接表征场景中精细的空间几何结构, 特别是为城市场景中道路交通标志牌、交通信号灯、路灯等小尺寸要素的精确几何形状描述提供了可能性。这些优势使得MLS系统迅速发展并逐渐应用于各行各业(Campbell等, 2010; 杨必胜等, 2017), 特别是城市场景的语义信息获取方面(靳兴浩等, 2022; 刘如飞等, 2022; Luo等, 2020)。但是因为城区MLS点云通常具有场景复杂、数据量大、空间分布不均匀等特点, 所以对其进行精准分类获取语义信息还存在重大挑战(Han等, 2021)。

传统的点云分类方法依靠人工提取特征, 设

收稿日期: 2023-05-25; 预印本: 2023-08-18

基金项目: 国家自然科学基金(编号: 41971414); 江苏省研究生科研创新项目(编号: KYCX23_1361)

第一作者简介: 雷相达, 研究方向为遥感数据智能处理。E-mail: leixd@nuist.edu.cn

通信作者简介: 管海燕, 研究方向为遥感数据智能处理。E-mail: guanhy.nj@nuist.edu.cn

计相应的分类规则或者采用传统机器学习分类器进行点云分类。如 Xu 等 (2015) 从点云数据中获取 13 个人工特征, 然后根据先验知识对场景进行地物分类。Hackel 等 (2016) 通过 K 邻近算法确定最优邻近距离, 然后基于特征向量、邻域矩阵以及高程值计算点云特征, 最后采用随机森林算法进行点云分类。Weinmann 等 (2015) 则提出了一个基于香农熵的最优邻域选择方法, 并在最优邻域的基础上计算点云特征用于点云分类。由于该类方法受限于方法设计者的先验知识以及分类器的分类性能, 无法快速且准确地进行大规模点云分类 (景庄伟 等, 2021)。

近年来, 深度学习技术以其强大的特征学习能力, 被广泛应用于三维点云数据处理, 并取得了大量的进展 (Guo 等, 2021)。如 Hu 和 Yuan (2016) 将点云转化为高程特征图, 然后将特征图输入到二维图像卷积神经网络中进行地面点以及非地面点分类。Qin 等 (2020) 首先对点云进行像素以及体素的转化, 再采用深度学习网络 VPNet 获取点云的深度特征以及预测类别。虽然上述方法可以获得较好的分类精度, 但是点云在形式转换的过程中, 会造成三维特征损失以及计算冗余的问题 (龚靖渝 等, 2023)。为了避免上述问题, 直接对点云进行分类的 PointNet (Qi 等, 2017a)、PointNet++ (Qi 等, 2017b) 深度学习方法被提出。随后基于点云的深度学习分类方法得到了快速发展, 如 DGCNN 方法 (Wang 等, 2019) 采用图卷积提取点云的局部特征, KPConv 方法 (Thomas 等, 2019) 应用核点卷积获取点云特征, PCAN (Zhang 和 Xiao, 2019) 通过注意力机制聚合点云特征。尽管上述方法可以提升点云分类精度, 但是它们多需要分块预处理, 不仅会增加数据处理时间, 还会影响点云几何特征的完整表达 (Du 等, 2021)。

为解决上述问题, 学者提出了直接对大规模点云进行操作的深度学习分类方法。如 RandLA-Net (Hu 等, 2022) 采用随机采样策略以及局部特征聚合模块, 可以快速且准确地进行大规模点云分类; FG-Net (Liu 等, 2023) 采用可变形卷积建模几何结构, 以及点态注意聚合模块挖掘点云之间的相关特征, 进而完成大规模点云分类; SCF-Net (Fan 等, 2021) 在上述方法的基础上, 通过局部极值表达点云的局部上下文信息, 其后采用双距离注意力模块以及全局上下文模块获取点云局部

以及全局特征; BAF-LAC 网络 (Shuai 等, 2021) 则在 RandLA-Net 网络基础上采用向后注意融合机制以及局部聚合分类器, 弥补原始网络投影过程中信息损失以及分类结果存在噪声的缺陷。虽然上述大规模点云分类方法通过增加不同模块提高了点云特征表达的能力, 但是复杂特征聚合模块会带来计算成本增加以及模型过拟合的问题。

本文提出一个整合集成预测约束与错误预测熵最大化的 MLS 点云分类方法, 可在不增加基线模型训练参数的情况下, 增强点云特征表达, 提高模型泛化能力。方法首先借鉴点云弱监督集成预测策略 (Wang 和 Yao, 2022; Lei 等, 2022), 在 RandLA-Net 基础上构建全监督集成预测约束分支; 该分支通过记录点云训练过程中的预测值生成集成预测值, 采用一致性约束当前预测以及集成预测值的差异, 增强点云的特征表达。其次, 针对模型容易出现过度置信预测的问题, 本文提出错误预测熵最大化策略, 对训练过程中的错误预测点进行熵值最大化, 增加复杂场景中易错点的不确定性, 从而防止模型过拟合, 提高点云分类精度。

2 方 法

本文方法的点云分类流程如图 1 所示。首先输入点云通过骨干网络获得点云预测概率分布, 然后输入基础监督分支、集成预测约束分支以及错误预测熵最大化分支分别计算加权交叉熵、集成预测以及错误预测熵最大化损失共同约束模型训练。其中:

(1) 基础监督分支根据真实标签、预测概率分布以及类别权重计算加权交叉熵损失, 为模型训练提供基础全监督信号。并且该分支通过对不同类别点进行加权, 缓解训练过程中点云类别不平衡问题。

(2) 集成预测约束分支首先整合训练过程中点云在不同阶段的预测值生成集成预测值, 然后计算其与当前预测值之间的差异获得集成预测损失。该分支通过集成预测值以自监督的形式丰富点云的特征, 提高模型的点云表征能力。

(3) 错误预测熵最大化分支基于真实标签和预测伪标签获取错误预测点, 进行错误预测熵值最大化计算熵最大化损失。该分支通过增加错误预测点的熵值, 提高复杂场景中模型预测的不确定性, 防止模型过拟合。

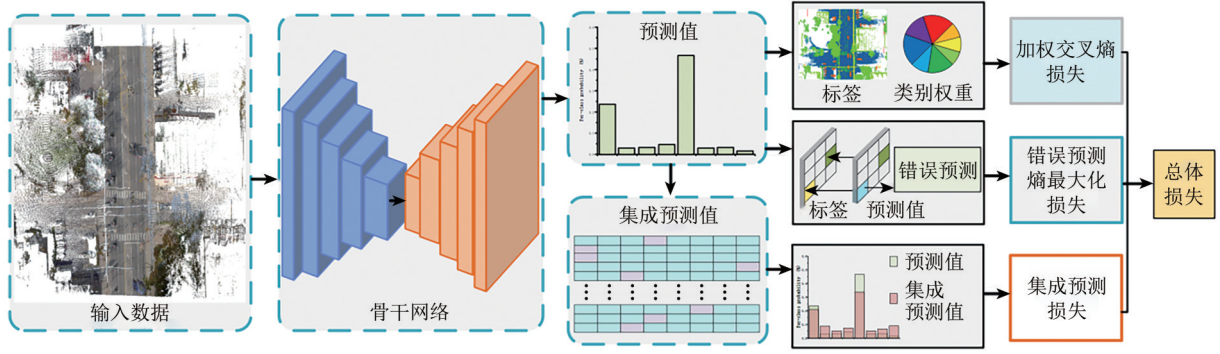


图1 点云分类流程图

Fig. 1 Flow chat of point cloud classification

2.1 骨干网络与基础监督分支

本文采用 RandLA-Net 网络作为骨干网络进行点云特征提取。如图 2 所示, RandLA-Net 采用经典的编码—解码结构。其中编码器由随机采样和局部特征聚合模块构成, 解码器由上采样和多层感知机构成, 编码以及解码器对应层之间由进行残差连接, 最后采用全连接分类器进行类别特征提取。由于 RandLA-Net 网络采用快速且复杂度低的随机采样方法, 所以该网络可以对大规模点云进行高效的点云分类。此外, 为了缓解点云类别不平衡问题, RandLA-Net 网络采用点云类别数量的比重计算权重。然而, 相较于 RandLA-Net 采用的权重计算方法, Han 等 (2021) 发现采用比重的

平方根可以获得更加平滑的权重, 进而提高点云的分类精度。因此本文采用平方根类别权重计算加权交叉熵损失 L_{se} , 具体的计算公式如下所示:

$$W_c = \frac{\sqrt{\sum_{i=1}^K N_i}}{\sqrt{N_c}} \quad (1)$$

$$L_{se} = -\frac{1}{|P_l|} \sum_i |P_l| W_i \sum_c y_{ic} \log p_{ic} \quad (2)$$

式中, W_c , N_c 为训练数据中类别为 c 的样本点权重和个数, N_i 为类别为 i 的样本点个数, K 为类别总体个数, P_l 为标记样本点集合, $|l|$ 为点集中点的个数, W_i 为第 i 个点的类别权重, y_{ic} 和 p_{ic} 分别为第 i 个点的标签以及属于第 c 类的预测概率。

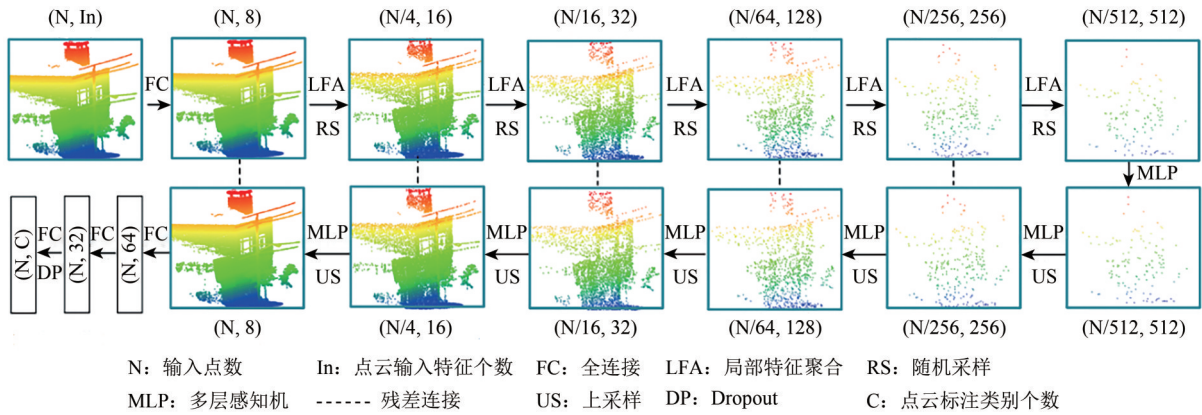


图2 RandLA-Net网络(点云数据由高程值进行显示)

Fig. 2 RandLA-Net network (Point clouds are displayed by elevation values)

2.2 集成预测约束分支

为了增强点云特征表达, 本文采用集成预测约束策略优化点云训练。如图 3 所示, 由于 RandLA-Net 网络输入具有一定的随机性, 同一个点在不同时序输入时, 其相对位置以及上下文信

息会发生相对变化, 即可以看作是一次数据增强。所以记录不同时序模型预测结果的集成预测值更加鲁棒, 特征表达更加丰富。本文通过一致性约束的方式, 鼓励模型预测向集成预测靠近, 增强点云特征表达。值得注意的是, 集成预测操作可

以在一次前向传递过程中完成,不会额外增加模型参数量,造成大量的训练计算损耗。具体生成集成预测的方法如图3所示:首先输入点云经过骨干网络预测输出概率分布后,将其传递到集成预测向量进行判断,即当前点的集成预测向量为

第一次赋值时,将当前预测概率分布直接赋予集成预测。若当前点的集成预测已经存储了预测值,则进行指数滑动平均计算获得集成预测值,如式(3)所示。

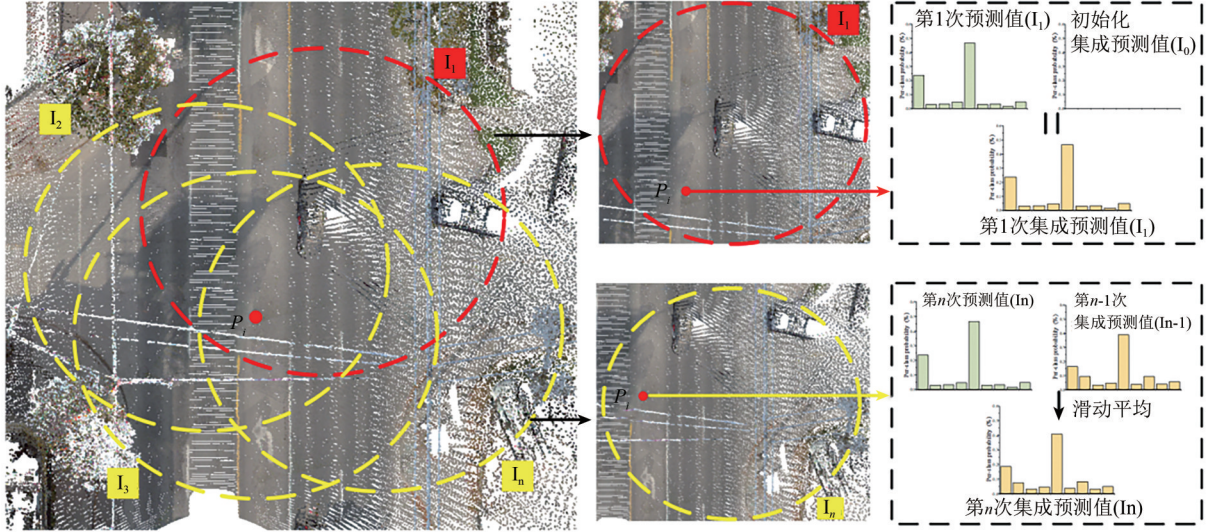


图3 集成预测流程图

Fig. 3 Flow chat of ensemble prediction

$$\bar{p}_N = \begin{cases} p_N, & N = 1 \\ \alpha \bar{p}_{N-1} + (1 - \alpha) p_N, & N > 1 \end{cases} \quad (3)$$

式中, \bar{p} 为集成预测值(概率分布), p 为当前预测值(概率分布), N 为集成预测值赋值的次数, α 为滑动指数, 本文设置为 0.9。在获得集成预测值后, 本文采用 KL 散度 KLD (Kullback-Leibler Divergence) 衡量集成预测值与当前预测值之间的差异 V , 然后计算集成预测损失, 并通过最小化该损失约束模型训练。其中两个预测值之间的差异以及集成预测损失 L_{ep} 的计算公式如下所示:

$$V(P_i) = D_{KL}(p_i \| \bar{p}_i) = \sum_c p_{ic} \log \left(\frac{p_{ic}}{\bar{p}_{ic}} \right) \quad (4)$$

$$L_{ep} = \frac{1}{M} \sum_i V(P_i) \quad (5)$$

式中, $V(P_i)$ 为 P_i 点的 KL 散度, p_{ic} 、 \bar{p}_{ic} 为 P_i 点在当前预测以及集成预测中被预测为 c 类别的概率, M 为输入点参与损失计算的个数。

2.3 错误预测熵最大化分支

在点云分类任务中, 复杂模型以及交叉熵损失对标签的 One-hot 预处理方法, 容易导致模型过度置信, 进而影响分类精度。为了缓解该问题, 传统深度学习网络通常采用软化 One-hot 的标签平

滑策略进行网络训练 (Müller 等, 2019)。该策略中隐含了对所有预测点进行熵值最大化处理, 即通过缓解模型对于真实标签的过度依赖, 提高模型泛化能力。Larrazabal 等 (2021) 通过在医学图像领域中的实验表明, 相较于对全部预测像素进行熵最大化处理, 仅对错误预测像素进行熵最大化可以更加有效地防止模型过拟合, 提高模型精度。受该方法的启发, 本文提出将错误预测熵最大化用于点云分类网络中优化模型训练, 提高模型分类的泛化能力。该方法首先将点云的真实标签以及当前预测的伪标签进行对比, 获取错误预测点集。然后鼓励错误预测点的熵值最大化, 提高该点的混乱度, 以减少模型对于易错点的过度置信预测, 提高模型泛化能力。本文采用 KL 散度衡量错误预测点预测概率分布与最大熵分布之间的差异 D_{KL} , 然后基于差异值计算熵最大化损失 L_{er} , 并通过最小化该损失进行熵值最大化, 其中熵最大化损失的计算公式如下:

$$L_{er} = \frac{1}{|P_e|} \sum_i D_{KL}(p_i \| p_K) = \frac{1}{|P_e|} \sum_i \sum_c p_{ic} \log \left(\frac{p_{ic}}{p_{Kc}} \right) \quad (6)$$

式中, P_e 为错误预测的点集, $|P_e|$ 为点集中点的个数, K 为点云中类别个数, p_i 和 p_{ic} 为第 i 个点的预

测概率分布以及预测为 c 类别的概率, p_k 为最大熵分布, 即所有类别的预测概率 p_{kc} 均为 $1/K$ 。

本文将上述3个分支获得的损失进行整合, 获得总体损失 L_{all} 用于模型训练。具体的计算公式, 如式(7)所示:

$$L_{all} = L_{se} + \lambda_1 L_{er} + \lambda_2 L_{ep} \quad (7)$$

式中, λ_1 、 λ_2 分别为熵最大化损失、集成预测损失的平衡权重。本文将 λ_1 、 λ_2 设置为1。

3 实验分析

3.1 数据集

如图4所示, Toronto3D包含了许多车道以外

的目标以及拥有丰富的属性信息, 因此本文选择该数据集作为实验的主要数据 (Tan 等, 2020)。该数据采源于加拿大城区, 包含约1 km长度的街道数据, 共有约7830万个点, 其中每个点包含了XYZ坐标信息、RGB颜色信息、强度信息、GPS时间、扫描角度以及标签信息。如图4所示, 数据集被划分为4个区域, 并且每个点被标注到9个类别, 分别是未分类 (Unclassified)、道路 (Road)、道路标记 (Road marking, Rd m.)、自然植被 (Natural)、建筑物 (Building, Build.)、公共线 (Utility line, Ut l.)、汽车 (Car) 以及栅栏 (Fence)。

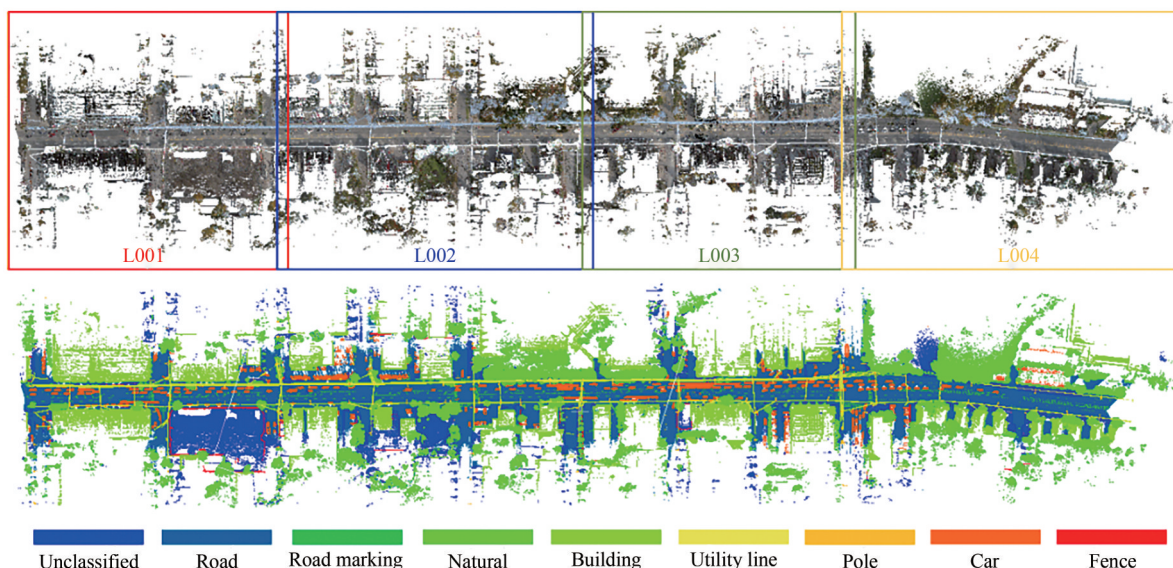


图4 Toronto3D数据集(第一行数据由RGB信息进行显示, 第二行数据由标签信息进行显示)

Fig. 4 Toronto3D dataset (The first row of data is displayed by RGB information, the second row of data is displayed by label information)

3.2 实验参数

实验选择Adam优化器, 将初始学习率、邻近点个数和格网下采样尺寸分别设定为0.01、16和0.05 m, 采用0.95作为学习率的衰减指数训练模型100个轮次。并根据显存大小, 设置输入点数为65536, 批次大小为4。模型下采样参数以及输出特征数分别设置为 $[4, 4, 4, 4, 2]$ 、 $[16, 32, 128, 256, 512]$ 。参照数据集的原始设定, 本文选择L002区域作为测试集, 其余区域作为训练数据, 并采用坐标信息以及颜色信息作为输入特征进行网络训练。为了更加准确的评价模型测试精度, 本文采用总体精度OA (Overall Accuracy)、交并比IoU (Intersection over Union) 以及平均交并比mIoU (mean Intersection over Union) 作为评价指标

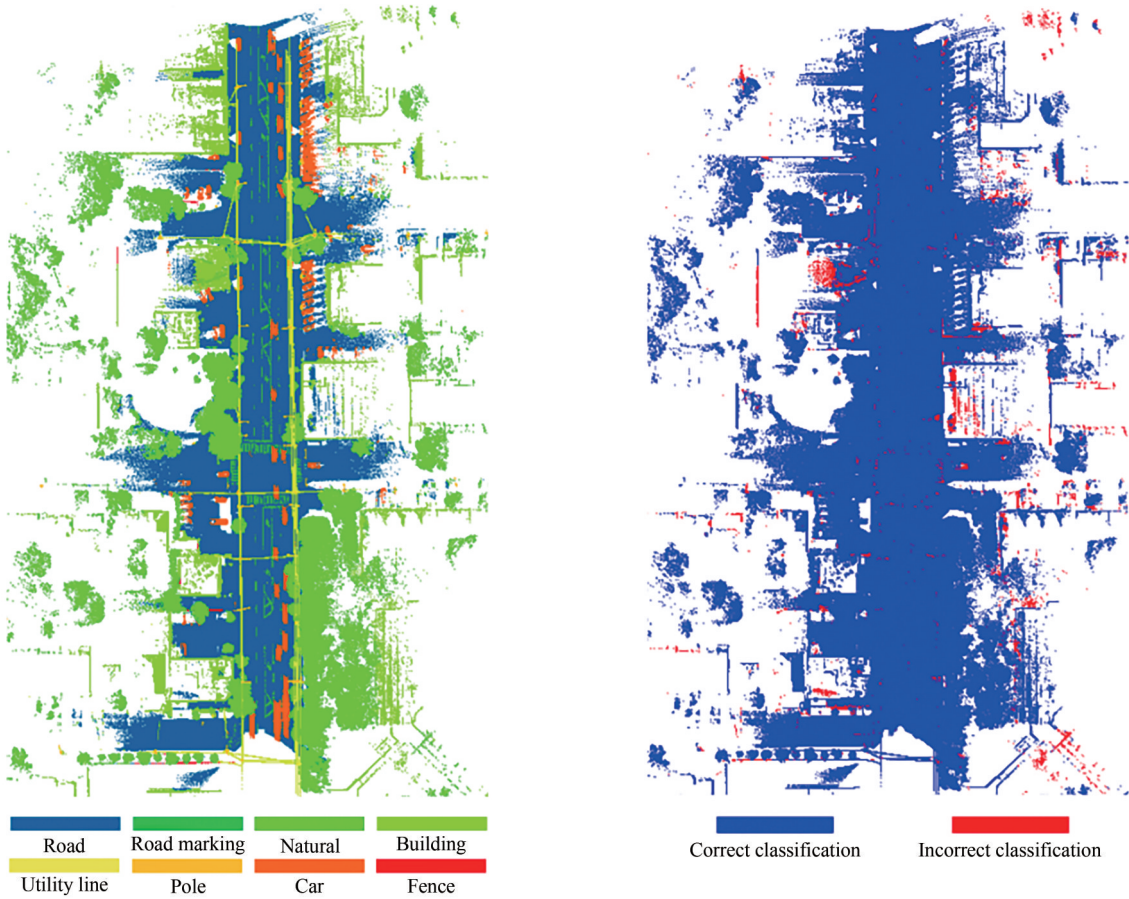
分析实验结果。本文所有实验均在NVIDIA RTX 2080 Ti图形处理器 (11 GB显存) 上进行。

3.3 实验结果分析

具体的实验结果如图5以及表1所示, 本文方法可以准确分类多数地物点, 其中OA以及mIoU分数可以达到97.71%和83.68%。但是从图5(b)错误预测示意图中可以看出, 所提方法在道路、公共线以及远离道路中心的地物分类上出现了错分现象。如图6首行中对道路标记的分类, 错误分类主要出现在道路标记的边缘区域。我们分析这是由于道路标记边缘点与道路的几何特征相似, 且颜色信息与真实标注存在一定误差导致的。从图6第二行对公共线的分类结果细节图中可以看

出，公共线与杆状物以及公共线与树冠重叠区域的特征相似，容易导致模型误分类。此外，如图6第三行细节图中所示，当地物距离道路中心较远

时，点云无法完整地表达地物特征，所以导致模型无法准确学习到该区域的地物特征，导致对该区域不完整建筑物、树木以及栅栏分类结果较差。



(a) 分类结果示意图 (b) 错误分类结果示意图
(a) Schematic diagram of classification results (b) Schematic diagram of incorrect classification results

图5 Toronto3D数据集分类结果
Fig. 5 Toronto3D dataset classification results

表1 Toronto3D定量对比结果
Table 1 Toronto3D quantitative comparison results

方法	OA	mIoU	IoU								/%
			Road	Road Marking	Natural	Building	Utility line	Pole	Car	Fence	
MappingC	94.72	82.89	97.15	67.87	97.55	93.75	86.88	82.12	93.72	44.11	
Rim et al.	83.60	71.03	92.84	27.43	89.90	95.27	85.59	74.50	44.41	58.30	
RandLA-Net	97.15	81.88	96.69	64.10	96.85	94.14	88.03	77.48	93.21	44.53	
基线方法*	96.36	78.69	95.63	57.84	96.56	93.21	86.61	75.69	90.84	33.14	
RFCR*	97.27	81.17	96.82	65.28	97.06	93.60	86.46	81.65	94.03	34.44	
本文方法*	97.71	83.68	97.40	70.00	97.42	94.20	86.24	82.60	94.75	46.85	

注：“*”表明该方法采用了本文的超参数进行模型训练以及测试。数字加粗表明方法在该指标中获得最优分数。

为验证所提方法优越性，本节将所提方法与其他主流深度学习方法在Toronto3D数据集上的分类结果进行对比分析。本文首先选择了已发表，且采用坐标和颜色信息作为输入特征的MappingConvSeg (MappingC) (Yan 等，2022)、Rim 等 (2021) 所

提方法以及RandLA-Net方法进行对比分析。然后我们在开源了代码的基线方法 (RandLA-Net)、RFCR方法 (Gong 等，2021)、SCF-Net方法以及BAF-LAC方法上进行模型训练以及测试，并设置了相同的随机种子，定量对比结果如表1和表2所示。

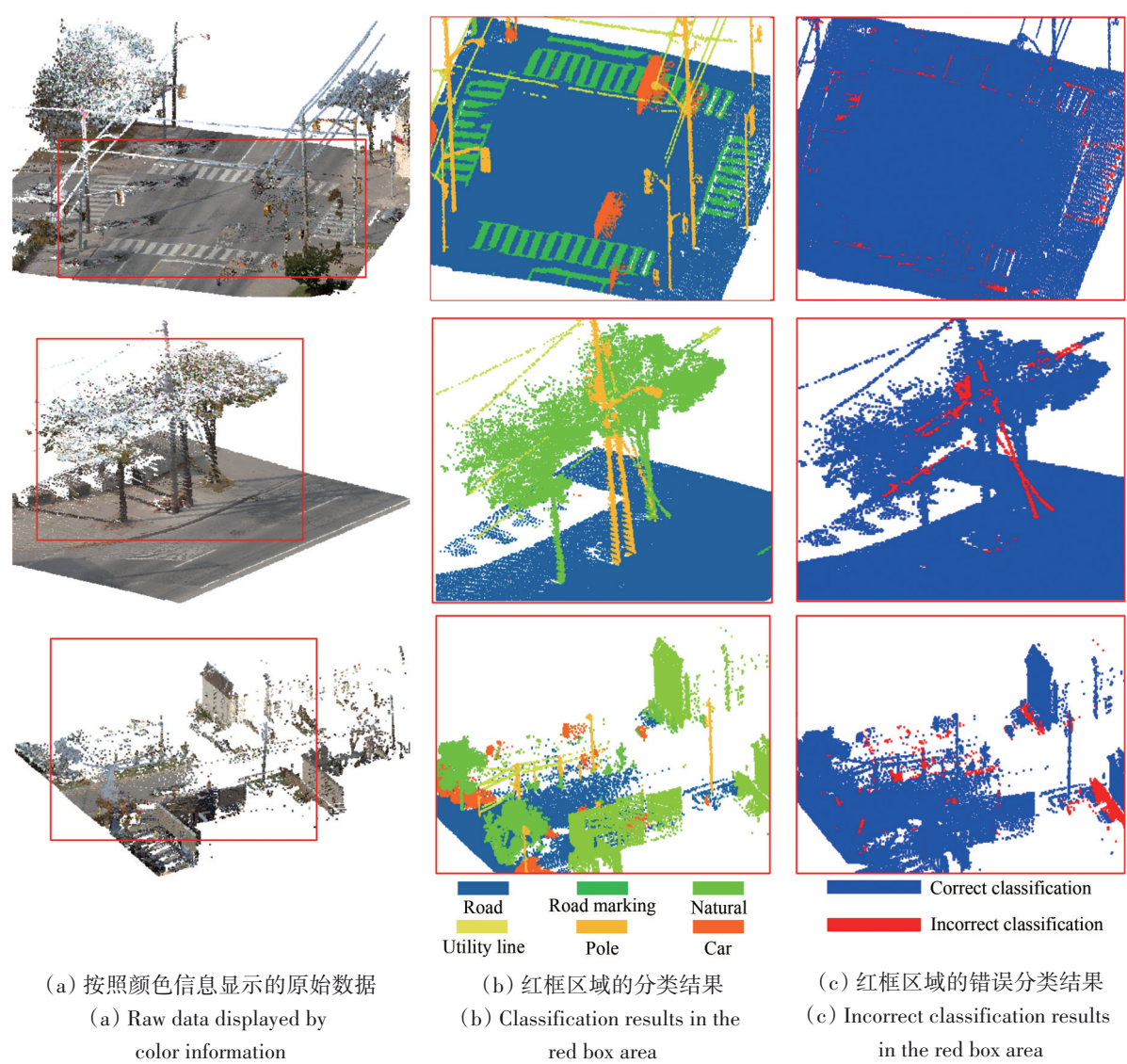


图6 分类细节图

Fig. 6 Classification detail map

表2 不同改进方法的定量对比结果

Table 2 Quantitative comparison results of different improvement methods

方法	OA	mIoU	IoU							
			Road	Road Marking	Natural	Building	Utility line	Pole	Car	Fence
SCF-Net [^]	96.39	78.46	95.8	59.14	95.91	93.36	87.37	73.63	88.23	34.24
BAF-LAC [^]	96.91	80.12	96.27	67.47	97.15	93.76	86.74	78.73	82.37	38.45
本文方法 [^]	97.57	82.03	97.28	68.46	97.10	94.02	87.30	82.29	93.65	36.11

注：“[^]”表明方法采用了批次大小为3的参数进行网络模型训练(其他超参数与本文方法保持一致)。数字加粗表明方法在该指标中获得最优分数。

由表1可知，与采用相同特征的全监督方法相比，本文所提方法在多数分类指标上优于对比方法，其中OA以及mIoU可以达到97.71%、83.68%。此外，本文采用的基线方法（RandLA-Net）相较于RandLA-Net论文精度有一定程度的降低。我们

认为这是由于硬件的限制，导致实验无法设置与论文相同的超参数，进而影响了最终的分类精度。当采用相同参数时，所提方法对比基线方法在除公共线以外的所有指标中都有提升，特别是在道路标记、杆状物、汽车以及栅栏的分类上有较大

的提升, 分别可以提高 12.16%、6.91%、3.91% 以及 13.71% 的 IoU 分数。相较于基线方法, 采用了完全监督的 RFCR 方法对网络进行了多尺度监督, 且采用特征密集 FD (Feature Densification) 方法获取更加明确的特征, 所以其精度获得了一定的提升。RFCR 方法与本文所提方法相比, 所提方法在除公共线之外的所有指标中, 均获得了精度的领先。我们认为这是由于所提方法从集成预测值中学习到了更有效的特征表达, 并且增加了易错点的不确定性, 进一步增强模型的分类能力, 提高了方法分类精度。

由于以 RandLA-Net 为基础的 SCF-Net 和 BAF-LAC 方法, 通过增加改进模块来提高模型对点云的表达能力。因此该类方法需要额外的显存进行模型训练, 进而导致实验无法在已经设置的超参数下进行。为了进行相关方法的对比, 我们将批次大小修改为 3, 并进行了模型训练以及测试, 详细结果如表 2 所示。

从表 2 中可以看出, 本文所提方法在多数评价指标中获得最优分数。特别是所提方法在汽车的分类上获得较大提升, 相较于其他两个方法至少获得了 5.42% 的提高。而 BAF-LAC 方法在植被以及栅栏上获得了精度优势、SCF-Net 在公共线的分类上获得最优结果。我们对结果进行分析后, 认为这是因为对比方法通过增加了上下文特征聚合模块加强了点云的上下文联系, 特别是 BAF-LAC 增加了局部聚合分类器对分类噪声进行了抑制, 使模型对于有立面特征的易错分类别 (栅栏) 有分类优势。而本文所提方法集合了训练过程中的特征, 并采用了熵正则化方法提升了相对独立的地物 (杆状物和汽车) 以及易错点道路以及道路标记的分类结果。结合上述所有结果来看, 本文所提方法相较于其他方法可以增强点云的特征表达, 提高模型泛化能力, 提升分类模型精度。

3.4 消融实验

(1) 所提策略的有效性。为验证所提方法中每个组成部分的有效性, 本节设置一组消融实验来评估不同模块的贡献。具体的实验组合为: 1) 将原始 RandLA-Net 网络作为基线方法; 2) 采用平方根加权损失函数 (W_{sqrt}) 替代基线网络的损失函数; 3) 增加集成预测约束 (CC); 4) 增加错误

预测熵最大化方法 (ER) 训练模型。4 个实验方法依次记为基线方法、方法 A、方法 B 和本文方法, 具体结果如表 3 所示。

表 3 消融实验结果

Table 3 Ablation results

方法	W_{sqrt}	CC	ER	OA/%	mIoU/%
基线方法				96.36	78.69
方法 A	√			97.30	80.39
方法 B	√	√		97.41	81.92
本文方法	√	√	√	97.71	83.68

由表 3 可知, 将基线网络中损失函数替换为平方根加权损失函数可以将 OA 和 mIoU 分别提升 0.94% 和 1.70%。这是由于基线网络中将类别数量的占比作为权重, 该方法会导致类别数量少的标记点获得权重过大, 进而影响模型分类精度。而采用平方根加权可以有效减少过大的权重, 使得模型训练更加平滑, 提高模型分类精度。相较于方法 A, 方法 B 的 OA 以及 mIoU 分别可以提升 0.11%、1.53%。这表明增加集成预测约束可以增强点云特征表达, 提高模型分类性能。

相较于方法 B, 增加了错误预测熵最大化分支的本文方法可以进一步提高模型的测试表现。其中 OA 以及 mIoU 分别可以提高 0.30% 和 1.76%。该对比结果表明, 本文设计的熵正则化方法, 可以有效缓解模型过拟合的问题, 更加充分地挖掘点云自身特征, 进一步提高模型分类性能。

(2) 不同熵正则化策略的有效性。为进一步验证所提方法中错误熵最大化策略的有效性, 本文采用 3 种不同的熵正则化策略进行实验。具体的正则化策略有: 标签平滑策略 (LS)、特征密集策略 (FD) 以及错误预测熵最大化 (ER) 策略。其中特征密集策略 (Gong 等, 2021) 通过将熵正则化引入到隐藏特征中, 从而辅助模型进行特征表达。结果如图 7 (a) 所示, 采用错误预测熵最大化策略训练的模型获得最优的测试结果, 相较于特征密集以及标签平滑的熵正则化策略, 分别可以提高 0.92%、2.23% 的 mIoU 分数。该结果表明相较于其他已发表熵正则化策略, 本文采用的错误预测熵最大化策略可以针对性的对错误预测点进行熵值最大化, 提高模型的泛化能力, 进而提高点云分类的测试精度。

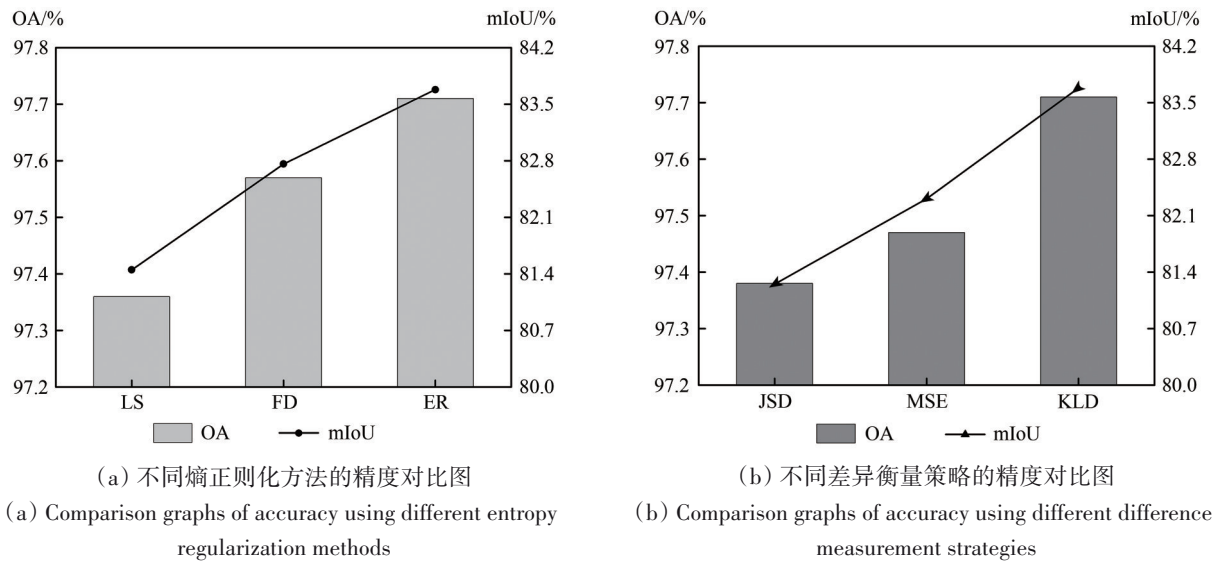


图7 采用不同训练策略的实验精度对比图

Fig. 7 Comparison charts of experimental accuracy using different training strategies

(3) 不同差异衡量策略的有效性。为验证所提方法中差异衡量策略的有效性。本文采用3种不同的差异衡量策略进行对比实验，具体的策略有：KL散度（KLD）、均方差MSE（Mean-Square Error）以及JS散度JSD（Jensen-Shannon Divergence）。具体的实验结果如图7（b）所示，本文采用的KL散度获得了最优的实验结果，其中OA以及mIoU分数分别可以提高0.24%、1.37%以上。结果表明相较于其他两个差异衡量策略，KL散度在所提方法中可以更有效地衡量预测之间的差异，增强点云分类模型的精度。

3.5 模型复杂度对比分析

为对比不同改进方法的模型复杂度，本文在采用相同超参数的情况下，选择浮点运算数FLOPs（Floating Point Operations）以及模型参数量（Parameters）作为衡量指标进行了模型复杂度对比分析，结果如表4所示。从结果中可以看出，本文所提方法与基线网络的FLOPs以及参数量相当，与对比方法相比，模型复杂度较低。结合分类精度结果，表明本文所提方法可以在不增加模型参数数量的情况下，通过集成预测约束分支以及错误预测熵最大化分支增加点云模型的分类性能。

表4 模型复杂度对比结果

Table 4 Model complexity comparison results

评价指标	RandLA-Net	SCF-Net	BAF-LAC	本文方法
FLOPs/百万	25.98	68.16	65.51	25.98
Parameters/百万	4.92	11.96	11.52	4.92

3.6 不同数据集的性能表现

为验证所提方法在不同MLS数据集上的性能表现，本文选择在公开MLS数据集WHU-MLS（杨必胜等，2021）以及Paris数据集（Deschaud等，2021）上进行模型的训练以及测试。相较于Toronto3D数据集，WHU-MLS以及Paris数据集分类场景更加复杂，分类对象更多，更加考验模型的性能。实验中训练、测试以及验证集均按照数据集的要求划分。并根据数据集提供的属性信息以及数量，将坐标以及强度信息作为WHU-MLS的输入特征，坐标以及颜色信息作为Paris数据集的输入特征。其中WHU-MLS以及Paris数据集分别将点云划分为18个和23个类别，并根据点云密度以及原论文的设置将格网下采样尺寸设置为0.08 m以及0.06 m，其余的超参数均按照本文参数设置。

此外，为了进行公平的精度对比，我们以相同的超参数进行基线方法以及RFCR方法实验，结果如表5所示。相较于基线方法，RFCR方法在Paris数据集上获得了精度的提升（OA提升1.41%、mIoU提升0.90%），但在WHU-MLS数据集上仅OA获得了提升（0.73%），而mIoU降低了10.65%。本文认为主要是WHU-MLS数据集中部分类别点云数量少且几何特征相似，导致采用完全监督的RFCR方法过度拟合，进而影响了小样本类别的分类效果。与基线方法相比，本文所提方法在Paris和WHU-MLS数据集上OA提升了2.11%和1.00%，

mIoU 提升了 3.83% 和 8.87%。相较于 RFCR 方法，所提方法 mIoU 分别提高了 2.93%、19.52%。结合 Toronto3D 数据集的结果，表明本文所提方法可以在多个 MLS 数据集上有效地提升基线方法性能。

表 5 不同数据集的实验结果
Table 5 Point cloud classification results of different datasets

方法	Paris		WHU-MLS	
	OA	mIoU	OA	mIoU
基线方法	90.81	40.36	89.77	56.98
RFCR	92.22	41.26	90.50	46.33
本文方法	92.92	44.19	90.77	65.85

4 结 论

本文提出一种整合集成预测约束以及错误预测熵最大化的城区 MLS 点云分类方法。方法首先在开源大规模点云分类网络 RandLA-Net 的基础上，引入集成预测约束策略生成相对鲁棒的集成预测值，并通过最小化当前预测值以及集成预测值之间差异，增强模型的点云特征表达能力。然后方法提出错误预测熵最大化策略，通过增加错误预测点的混乱程度，提高分类模型对于易错点的泛化能力，进而提升模型分类性能。方法在 Toronto3D 数据集上的消融实验表明，所提方法的各个分支均可以提高模型分类性能。与目前主流的深度学习点云分类方法相比，所提方法获得了最优的分类精度。相较于基线方法，其 OA 以及 mIoU 分数分别可以提高 1.35% 和 4.99%。最后通过在多个公开 MLS 点云数据集上的实验结果表明，所提方法可以在不增加模型复杂度的前提下，有效地提高基线网络的性能。尽管所提方法获得了具有竞争力的精度，但是方法还是受限于标记样本以及上下文场景有限的问题。在下一步工作中，我们将针对上述问题，进行点云弱监督分类方法以及点云数据增强相关的研究。

参考文献(References)

Campbell M, Egerstedt M, How J P and Murray R M. 2010. Autonomous driving in urban environments: approaches, lessons and challenges. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 368(1928): 4649-4672 [DOI: 10.1098/rsta.2010.0110]

China Academy of Information and Communications Technology. 2020. Digital twin cities white paper (中国信息通信研究院. 2020. 数字孪生城市白皮书)

Deschaud J F, Duque D, Richa J P, Velasco-Forero S, Marcotegui B and Goulette F. 2021. Paris-CARLA-3D: a real and synthetic outdoor point cloud dataset for challenging tasks in 3D mapping. Remote Sensing, 13(22): 4713 [DOI: 10.3390/rs13224713]

Du J, Cai G R, Wang Z Y, Huang S F, Su J H, Marcato Junior J, Smit J and Li J. 2021. ResDLPS-Net: joint residual-dense optimization for large-scale point cloud semantic segmentation. ISPRS Journal of Photogrammetry and Remote Sensing, 182: 37-51 [DOI: 10.1016/j.isprsjprs.2021.09.024]

Fan S Q, Dong Q L, Zhu F H, Lv Y S, Ye P J and Wang F Y. 2021. SCF-Net: learning spatial contextual features for large-scale point cloud segmentation//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE: 14499-14508 [DOI: 10.1109/CVPR46437.2021.01427]

Gong J Y, Lou Y J, Liu F Q, Zhang Z W, Chen H M, Zhang Z Z, Tan X, Xie Y and Ma L Z. 2023. Scene point cloud understanding and reconstruction technologies in 3D space. Journal of Image and Graphics, 28(6): 1741-1766 (龚靖渝, 楼雨京, 柳奉奇, 张志伟, 陈豪明, 张志忠, 谭鑫, 谢源, 马利庄. 2023. 三维场景点云理解与重建技术. 中国图象图形学报, 28(6): 1741-1766) [DOI: 10.11834/jig.230004]

Gong J Y, Xu J C, Tan X, Song H C, Qu Y Y, Xie Y and Ma L Z. 2021. Omni-supervised point cloud segmentation via gradual receptive field component reasoning//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE: 11668-11677 [DOI: 10.1109/CVPR46437.2021.01150]

Guo Y L, Wang H Y, Hu Q Y, Liu H, Liu L and Bennamoun M. 2021. Deep learning for 3D point clouds: a survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(12): 4338-4364 [DOI: 10.1109/TPAMI.2020.3005434]

Hackel T, Wegner J D and Schindler K. 2016. Fast semantic segmentation of 3D point clouds with strongly varying density. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, III-3: 177-184 [DOI: 10.5194/isprannals-III-3-177-2016]

Han X, Dong Z and Yang B S. 2021. A point-based deep learning network for semantic segmentation of MLS point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 175: 199-214 [DOI: 10.1016/j.isprsjprs.2021.03.001]

Hu Q Y, Yang B, Xie L H, Rosa S, Guo Y L, Wang Z H, Trigoni N and Markham A. 2022. Learning semantic segmentation of large-scale point clouds with random sampling. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(11): 8338-8354 [DOI: 10.1109/TPAMI.2021.3083288]

Hu X Y and Yuan Y. 2016. Deep-learning-based classification for DTM extraction from ALS point cloud. Remote Sensing, 8(9):730 [DOI: 10.3390/rs8090730]

Jin X H, Wang C, Hao L M and Li J J. 2022. Application of vehicle-mounted LiDAR mobile measurement system in urban parts collection. Geomatics and Spatial Information Technology, 45(2): 90-92, 96 (靳兴浩, 王超, 郝利民, 李军杰. 2022. 车载 LiDAR 移动

- 测量系统在城市部件采集的应用. 测绘与空间地理信息, 45(2): 90-92, 96 [DOI: 10.3969/j.issn.1672-5867.2022.02.024]
- Jing Z W, Guan H Y, Zang Y F, Ni H, Li D L and Yu Y T. 2021. Survey of point cloud semantic segmentation based on deep learning. *Journal of Frontiers of Computer Science and Technology*, 15(1): 1-26 (景庄伟, 管海燕, 臧玉府, 倪欢, 李迪龙, 于永涛. 2021. 基于深度学习的点云语义分割研究综述. *计算机科学与探索*, 15(1): 1-26) [DOI: 10.3778/j.issn.1673-9418.2006025]
- Larrazabal A J, Martinez C, Dolz J and Ferrante N. 2021. Maximum entropy on erroneous predictions: improving model calibration for medical image segmentation//*Proceedings of the 26th International Conference on Medical Image Computing and Computer Assisted Intervention*. Vancouver: Springer: 273-283 [DOI: 10.1007/978-3-031-43898-1_27]
- Lei X D, Guan H Y, Ma L F, Yu Y T, Dong Z, Gao K, Delavar M R and Li J. 2022. WSPointNet: a multi-branch weakly supervised learning network for semantic segmentation of large-scale mobile laser scanning point clouds. *International Journal of Applied Earth Observation and Geoinformation*, 115: 103129 [DOI: 10.1016/j.jag.2022.103129]
- Liu K C, Gao Z, Lin F and Chen B M. 2023. FG-Net: a fast and accurate framework for large-scale LiDAR point cloud understanding. *IEEE Transactions on Cybernetics*, 53(1): 553-564 [DOI: 10.1109/TCYB.2022.3159815]
- Liu R F, Ma X J, Lu X S, Wang M Y and Wang P. 2022. Automatic extraction of urban road boundaries using diverse LBP features. *National Remote Sensing Bulletin*, 26(3): 541-554 (刘如飞, 马新江, 卢秀山, 王旻烨, 王鹏. 2022. 利用多元LBP特征自动提取城市道路边界. *遥感学报*, 26(3): 541-554) [DOI: 10.11834/jrs.20209228]
- Luo H F, Khoshelham K, Fang L N and Chen C C. 2020. Unsupervised scene adaptation for semantic segmentation of urban mobile laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 169: 253-267 [DOI: 10.1016/j.isprsjrs.2020.10.002]
- Müller R, Kornblith S and Hinton G. 2019. When does label smoothing help?//*Proceedings of the 33rd International Conference on Neural Information Processing Systems*. Vancouver: Curran Associates Inc.: 4694-4703
- Qi C R, Li Y, 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: Curran Associates Inc.: 5105-5114
- Qi C R, Su H, Mo K C 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: IEEE: 77-85 [DOI: 10.1109/CVPR.2017.16]
- Qin N N, Hu X Y, Wang P Z, Shan J and Li Y J. 2020. Semantic labeling of ALS point cloud via learning voxel and pixel representations. *IEEE Geoscience and Remote Sensing Letters*, 17(5): 859-863 [DOI: 10.1109/LGRS.2019.2931119]
- Rim B, Lee A and Hong M. 2021. Semantic segmentation of large-scale outdoor point clouds by encoder-decoder shared MLPs with multiple losses. *Remote Sensing*, 13(16): 3121 [DOI: 10.3390/rs13163121]
- Shuai H, Xu X and Liu Q S. 2021. Backward attentive fusing network with local aggregation classifier for 3D point cloud semantic segmentation. *IEEE Transactions on Image Processing*, 30: 4973-4984 [DOI: 10.1109/TIP.2021.3073660]
- Tan B, Zhong R F and Li Q. 2012. Objects classification with vehicle-borne laser scanning data. *Journal of Remote Sensing (in Chinese)*, 16(1): 50-66. (谭贲, 钟若飞, 李芹. 2012. 车载激光扫描数据的地物分类方法. *遥感学报*, 16(1): 50-66) [DOI: 10.11834/jrs.20120408]
- Tan W K, Qin N N, Ma L F, Li Y, Du J, Cai G R, Yang K and Li J. 2020. Toronto-3D: a large-scale mobile LiDAR dataset for semantic segmentation of urban roadways//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Seattle: IEEE: 797-806 [DOI: 10.1109/CVPRW50498.2020.00109]
- Thomas H, Qi C R, Deschaud J E, Marcotegui B, Goulette F and Guibas L. 2019. KPConv: flexible and deformable convolution for point clouds//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul: IEEE: 6410-6419 [DOI: 10.1109/ICCV.2019.00651]
- Wang P Z, Yao W. 2022. A new weakly supervised approach for ALS point cloud semantic segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 288, 237-254 [DOI: 10.1016/j.isprsjrs.2022.04.016]
- Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M and Solomon J M. 2019. Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics (TOG)*, 38(5): 146 [DOI: 10.1145/3326362]
- Weinmann M, Jutzi B, Hinz S and Mallet C. 2015. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105: 286-304 [DOI: 10.1016/j.isprsjrs.2015.01.016]
- Xu S D, Vosselman G and Oude Elberink S. 2015. Detection and classification of changes in buildings from airborne laser scanning data. *Remote Sensing*, 7(12): 17051-17076 [DOI: 10.3390/rs71215867]
- Yan K P, Hu Q Y, Wang H Y, Huang X H, Li L and Ji S. 2022. Continuous mapping convolution for large-scale point clouds semantic segmentation. *IEEE Geoscience and Remote Sensing Letters*, 19: 6502505 [DOI: 10.1109/LGRS.2021.3107006]
- Yang B S, Han X and Dong Z. 2021. Point cloud benchmark dataset WHU-TLS and WHU-MLS for deep learning. *National Remote Sensing Bulletin*, 25(1): 231-240 (杨必胜, 韩旭, 董震. 2021. 点云深度学习基准数据集. *遥感学报*, 25(1): 231-240) [DOI: 10.11834/jrs.20210542]
- Yang B S, Liang F X and Huang R G. 2017. Progress, challenges and perspectives of 3D LiDAR point cloud processing. *Acta Geodaetica et Cartographica Sinica*, 46(10): 1509-1516 (杨必胜, 梁福逊, 黄荣刚. 2017. 三维激光扫描点云数据处理研究进展、挑战与趋势. *测绘学报*, 46(10): 1509-1516) [DOI: 10.11947/j.AGCS.2017.20170351]

Zhang W X and Xiao C X. 2019. PCAN: 3D attention map learning using contextual information for point cloud based retrieval//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and

Pattern Recognition. Long Beach: IEEE: 12428-12437 [DOI: 10.1109/CVPR.2019.01272]

Integrating ensemble prediction constraints and error prediction entropy maximization for MLS point cloud classification

LEI Xiangda¹, GUAN Haiyan¹, DONG Zhen²

1. School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China;

2. State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Abstract: Mobile Laser Scanning (MLS) systems are widely used in various fields owing to their ability of rapidly acquiring high-precision and high-density 3D point cloud data, particularly in the acquisition of urban spatial information. Given that urban MLS point clouds exhibit complex scenes, large data volumes, and uneven spatial distribution, accurate classification of large-scale urban point clouds presents significant challenges. Currently, many deep learning point cloud classification methods enhance feature representation of point clouds by adding a feature aggregation module. Nonetheless, this approach frequently results in increased training parameters and model overfitting.

We propose an MLS point cloud classification method integrating ensemble prediction constraints and error prediction entropy maximization. The proposed method can enhance the point cloud feature representation of the baseline network and improve the generalization ability of the model without increasing the training parameters. Our method consists of three main components: a basic supervision branch, an ensemble prediction constraint branch, and an error prediction entropy maximization branch. Specifically, we first employ RandLA-Net as the backbone network to obtain point cloud classification features. Then, a basic supervised branch calculates the weighted cross-entropy loss on the basis of true labels and predicts probability distributions and category weights to provide a basic fully supervised signal for model training. For the ensemble prediction constraint branch, we first generate ensemble predictions by recording the predicted values during the point cloud training process. Because the input to RandLA-Net is a random subpoint cloud, the ensemble predictions can be integrated for predictions not only at different stages but also at different relative positions. Thus, the ensemble prediction is highly robust to the current prediction. Afterward, we apply a consistency constraint to minimize the difference between the two predictions to improve the point cloud feature representation. Finally, we design an error prediction entropy maximization branch to maximize the entropy of error prediction point sets, increasing their confusion to reduce the model overfitting.

The public MLS point cloud dataset Toronto3D is chosen as the primary dataset to validate the performance of the proposed method. The qualitative result (Fig. 5) and quantitative result (Table 1) on the Toronto3D dataset show that the proposed method can correctly classify most points. To verify the validity of the method, we compare the accuracy of the proposed method with those of other popular methods. The comparison results (Table 1 and 2) indicate that the proposed method obtains the best OA (97.71%) and mIoU (83.68%). A series of ablation experiments is carried out to verify the effectiveness of each branch of the proposed method. The results (Table 3 and Fig. 7) show that each branch can effectively improve the model classification performance. The complexity analysis (Table 4) indicates that the proposed method can enhance the accuracy of the baseline method without increasing the model parameters. Furthermore, the experimental results (Table 5) on other public MLS datasets (WHU-MLS and Paris datasets) demonstrate that the proposed method can obtain competitive results on multiple datasets.

Key words: remote sensing, MLS point cloud classification, deep learning, ensemble prediction constraints, error prediction entropy maximization

Supported by National Natural Science Foundation of China (No. 41971414); Postgraduate Research and Practice Innovation Program of Jiangsu Province (No. KYCX23_1361)