

整合迁移学习与全卷积网络的小样本机载激光雷达点云分类

雷相达, 王宏涛*, 赵宗泽

河南理工大学测绘与国土信息工程学院, 河南 焦作 454000

摘要 针对基于深度学习的机载激光雷达(LiDAR)点云分类方法训练时间长、对样本数据需求量较大等问题,提出了一种整合迁移学习与全卷积神经网络(FCN)的小样本机载 LiDAR 点云分类方法。首先,对机载 LiDAR 点云的光谱信息进行补充,并提取点云数据的归一化高程、强度值和植被指数特征构建三通道点云特征图。然后,通过设置不同的邻域大小和投影方向,生成多尺度和多投影特征图,并采用密集连接网络 DenseNet201 预训练模型提取多尺度、多投影深度特征。最后,通过池化操作提取深度特征中的全局特征,用 FCN 进行初步分类后基于图割优化策略实现机载 LiDAR 点云的高精度分类。在标准测试数据集上的实验结果表明,在训练样本数量约占数据集 1.4% 的情况下,本方法的总体分类精度可达到 89.91%。

关键词 遥感; 机载激光雷达点云; 光谱数据; 迁移学习; 小样本; 图割优化

中图分类号 P237

文献标志码 A

doi: 10.3788/CJL202148.1610001

1 引言

机载激光雷达(LiDAR)是一种集成激光扫描测距系统、全球卫星定位系统和高精度惯性导航测量系统的主动遥感设备,可获取各种地物目标的高精度、高密度三维点云^[1]。近年来,随着无人机平台以及传感器的快速发展,机载 LiDAR 点云数据已被广泛应用于测绘、林业以及电力等行业。但从离散、不规则分布且包含噪声的机载 LiDAR 点云数据中自动、准确分类和解译出各种典型地物目标的难度较大^[2],成为近年来相关领域的研究热点之一。

目前,点云分类方法主要分为非监督分类和监督分类两种。非监督分类主要是基于规则的点云分类方法,需要设定较多的阈值,且分类结果易受噪声分布、点云密度和场景复杂度等因素的影响^[3],普适性较差。而监督分类首先利用训练样本进行特征训练,进而构建相应的模型进行分类,具有分类精度高、普适性强等特点,是目前使用最多的一种分类方

法。通常情况下,监督分类需要从点云中选取一定数量的样本区域,然后利用机器学习算法进行特征分析,进而根据不同的特征进行地物分类。如 Zhang 等^[4]和释小松等^[5]从点云数据中提取各种几何特征,并用支持向量机对点云进行分类;何曼芸等^[6]从点云数据中获取几何特征和光谱特征,并用随机森林等分类器进行点云分类;Hackel 等^[7]和 Ramiya 等^[8]对点云的邻域选择方法进行改进,充分考虑了邻域信息,并基于传统分类器对点云进行分类;Ni 等^[9]对点云进行初步分割形成点簇,再基于点簇进行特征提取及点云分类。为了保证分类精度,监督分类往往需要获取足够多的特征信息,在选取训练样本区域后由分类器构建复杂的分类模型。因此,该分类方法有较高的分类精度及鲁棒性,但其分类精度受限于训练样本的数量以及分类器的性能。近年来,随着深度学习的发展,基于深度学习的点云分类方法也相继产生。该方法主要通过级联神经网络逐层抽象得到特

收稿日期: 2020-11-28; 修回日期: 2020-12-29; 录用日期: 2021-02-07

基金项目: 河南省自然科学基金面上项目(182300410115)、河南省科技攻关(172102310350)、河南理工大学博士基金(B2017-13)、河南理工大学基本科研业务费专项(NSFRF170908)

通信作者: *211804010013@home.hpu.edu.cn

征,在提高运算效率的同时得到比传统机器学习方法更高的分类精度^[10]。但深度学习方法通常需要较多的参数量和训练数据,训练时间较长,从而影响最终的处理效率。综上所述,监督分类方法存在的主要问题:1)特征提取过程需要大量的计算,且往往需要人工干预,影响了分类过程的自动化程度;2)分类过程往往需要大量的样本数据,以保证最终的分类精度;3)分类器往往需要设置较多的参数,使分类过程变得繁琐复杂。

针对上述问题,人们开始研究基于迁移学习的点云分类方法^[3]。迁移学习可将网络在原任务中学习到的成果迁移到目标任务中,在提取有效特征的同时减少深度学习网络的复杂度,简化训练过程中的参数设置。因此,本文采用迁移学习方法对描述

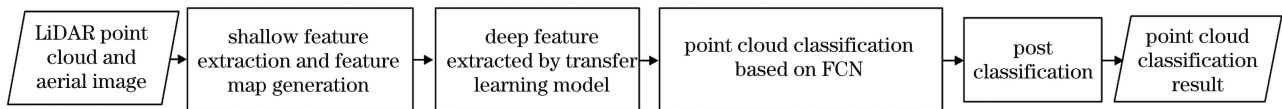


图 1 基于迁移学习的小样本点云分类流程

Fig. 1 Flow chart of the small sample point cloud classification based on transfer learning

2.1 点云特征提取与特征图的生成

2.1.1 点云特征

目前,应用到机载 LiDAR 点云分类中的点云特征主要包含几何特征、光谱特征、高程特征三类。在选取点云特征时,Niemeyer 等^[11]和 杨书娟等^[12]对各类特征的重要性进行评价,结果表明,归一化高度是机载 LiDAR 点云数据分类中最重要的特征。此外,光谱特征中的强度信息是识别不同表面材料的有效特征,融合多光谱航空影像光谱信息计算的归一化差异植被指数(NDVI)可以很好地区分地物与植被。此外,几何特征中的曲率变化率、表面粗糙度以及平整度等特征,可以区别建筑物、道路等平面物体与植被灌木等不规则物体。但计算点云几何特征时,容易受到邻域条件、点云密度以及点云噪声的影响,无法提取具有稳定表征性的特征。而归一化高度、强度值、NDVI 三个特征具有一定的稳定性。因此,本方法选择归一化高度、强度值、NDVI 三个特征作为点云的浅层特征。

1) 强度值的提取:机载 LiDAR 系统可提供激光脉冲的回波强度信息,因此,本方法采用原始数据中的强度信息作为特征值。

2) 归一化高程值的计算:采用准确且高效的布料模拟滤波算法^[13]对原始点云进行滤波,得到地面点及非地面点。为克服实验区域地形对滤波效果的

三维点特征的多尺度、多投影特征图进行深度特征提取,从而在小样本数据的情况下,最大程度地减少训练时间,同时保证点云特征的有效性。此外,综合考虑点云的局部特征和全局特征,采用浅层深度学习网络以及考虑上下文信息的后处理方法,完成了机载 LiDAR 点云的高精度分类。

2 分类方法与原理

本方法的主要步骤如图 1 所示,首先,用机载 LiDAR 点云以及航空影像提取的浅层特征生成多尺度、多投影特征图;然后,通过迁移学习进行深层特征的提取,再构建全卷积网络(FCN)进行点云分类;最后,通过图割优化对分类结果进行后处理,得到分类结果。

影响,首先,对点云进行分块处理,用布料模拟滤波算法分别对平缓地形和陡峭地形进行滤波处理,并对地面点进行合并。然后,对地面点进行内插处理生成数字高程模型(DEM),并逐点获取归一化高度。

3) NDVI 的计算:由于 NDVI 的计算需要多光谱信息,因此先将机载 LiDAR 点云与相应区域的多光谱航空影像进行融合,补充点云的多光谱信息^[6]。数据融合方法将三维点云的坐标以及多光谱航空影像的外方位元素代入共线条件方程,以计算每个三维点在对应影像上的像素位置 (x, y) ,该过程可表示为

$$x = -f \frac{a_1(X - X_s) + b_1(Y - Y_s) + c_1(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)},$$

$$y = -f \frac{a_2(X - X_s) + b_2(Y - Y_s) + c_2(Z - Z_s)}{a_3(X - X_s) + b_3(Y - Y_s) + c_3(Z - Z_s)},$$
(1)

式中, (X, Y, Z) 为地面点的三维坐标, f 为相机的焦距, (X_s, Y_s, Z_s) 为外方位元素的三个线元素, $(a_1, b_1, c_1, a_2, b_2, c_2, a_3, b_3, c_3)$ 为由外方位元素中三个角元素计算的旋转矩阵参数。

获取像点坐标后,采用最邻近插值方法得到其对应的多光谱信息,包括近红外、红色、绿色波段的信息 W_{NIR}, W_R, W_G 。每个三维点对应的 NDVI 值可表示为

$$W_{NDVI} = \frac{W_{NIR} - W_R}{W_{NIR} + W_R} \quad (2)$$

2.1.2 点特征图的生成

采用深度学习方法进行点云分类时,通常将三维点投影至二维平面,生成二维点云特征图,进而完成点云分类。这类方法中点特征图的质量决定着点特征的表达,从而影响最终的分类精度。若直接对三维点进行二维投影,不考虑点的空间关系,特征图则无法完整表达点的三维空间特征,造成特征损失,进而影响分类结果。文献[3]提出了一种考虑三维点空间关系的点特征图生成方法,可以很好地表达点云的三维特征,但较多的邻近点搜索导致生成特征图的成本较高。因此,本方法在此基础上对投影点云以网格尺寸为参数进行体素下采样,在保留三

维点特征的基础上减少计算量。图 2 显示了以 p 点为中心点的局部机载 LiDAR 点云数据集生成 p 点特征图的过程。首先,将点云投影至二维平面,并以点 p 为中心构建一个规则网格(大小为 33×33),每个网格点的平面坐标由 p 点的坐标和网格大小确定,高程值是利用每个网格点搜索 3 个与 p 点距离最近的点。然后,按照反距离加权插值得到特征点的二维坐标,如图 2(a)所示。对于每个三维网格点,从 p 点的立方体邻域内搜索与其距离最近的 3 个点,并采用反距离权重计算得到特征图中每个网格点的 NDVI、强度和归一化高程值,如图 2(b)所示。最后,对三种特征重新赋值(0~255),生成 1 张具有三个通道的特征图,如图 2(c)所示,以用于后续的深层特征提取。

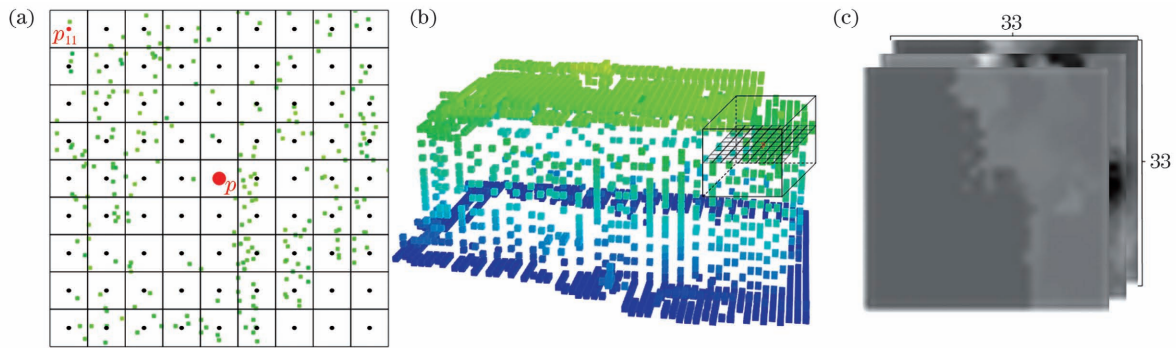


图 2 点特征图的生成过程。(a)特征图的二维坐标;(b)立方体邻域内的点特征;(c)点特征图

Fig. 2 Generation process of the point feature map. (a) Two-dimensional coordinates of the feature map; (b) point features in the cube neighborhood; (c) point feature map

2.1.3 多尺度及多投影特征的数据增强

1) 多尺度特征图的生成

为了减小地物尺寸对分类结果的影响,使提取的特征具有多尺度特性,提高特征对噪声的鲁棒性^[14]。本方法将特征图的网格高度 H 设置为 33,在此基础上,根据数据集的点云密度及实验区域的地物尺寸,分别采用 0.1, 0.3, 0.5 m 作为点特征图

的像素网格尺寸,即立方体的邻域边长分别为 3.3, 9.9, 16.5 m, 得到不同尺度的点特征图如图 3 所示。多尺度特征图可以增加点特征的表达,实现不同大小地物目标的准确分类。

2) 多投影特征图的生成

对于三维点云的分类,若只考虑一个投影方向,则无法准确刻画地物目标的三维特征,从而影响点

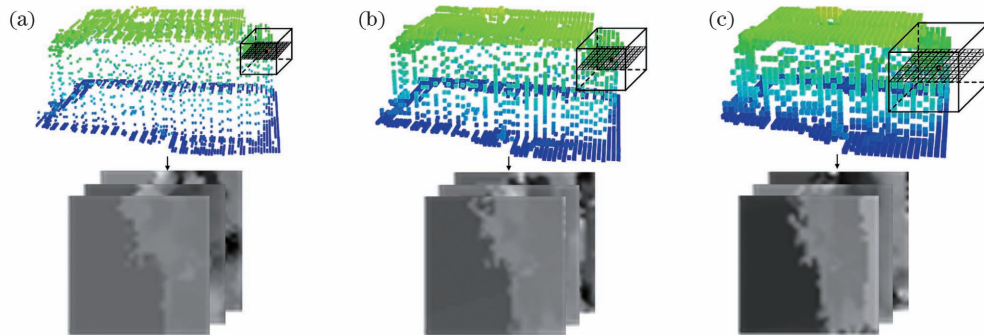


图 3 多尺度特征图的生成。(a)网格尺寸为 0.1 m;(b)网格尺寸为 0.3 m;(c)网格尺寸为 0.5 m

Fig. 3 Generation of the multi-scale feature maps. (a) Grid size of 0.1 m; (b) grid size of 0.3 m; (c) grid size of 0.5 m

云的分类精度。通常将原始数据先进行小范围旋转,再进行特征图的生成,从而得到不同视角的特征图^[15]。若只进行 Z 方向的投影以及小范围旋转,得到的多视角特征图无法提取点云其他方向的特征,且旋转角度的设置需结合一定的先验知识。因此,本方法首先将像素网格尺寸设置为 0.1 m,在上述

点特征图生成方法的基础上,采用转换投影方向的方式,分别将点云向 YZ 及 XZ 平面进行投影,生成 X、Y 方向的投影特征图,如图 4 所示。最终每个点生成 5 张三通道的浅层特征图($5 \times 33 \times 33 \times 3$),包括基于 YZ 平面三个尺寸的多尺度特征图以及 XZ 和 YZ 平面的多投影特征图。

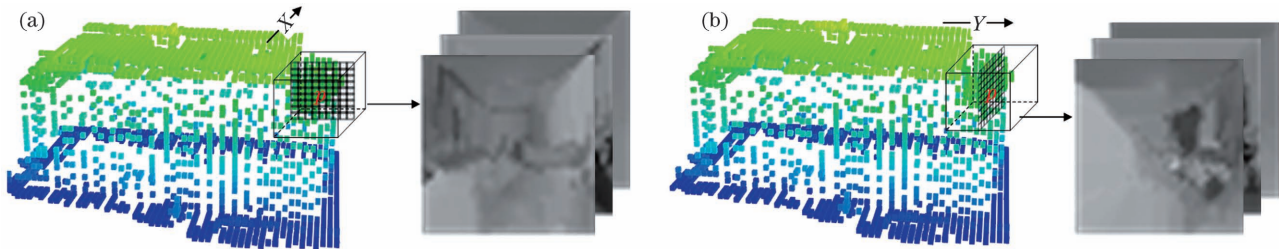


图 4 多投影示意图。(a) X 方向;(b) Y 方向

Fig. 4 Schematic diagram of the multi-projection. (a) X direction; (b) Y direction

2.2 基于迁移学习的点云深度特征提取

传统机器学习以及深度学习方法在提取深层次特征的过程中,需要消耗大量的时间和样本数据。迁移学习可认为是小样本分类的一种有效策略,包括基于实例的迁移、基于特征的迁移以及基于模型共享参数的迁移等。其中,基于模型共享参数的迁移学习方法可将在大规模数据集中预训练好的网络模型嵌入其他任务模型中作为特征提取器,从而有效提取小样本数据的深度特征信息,以进行分类^[16]。因此,本方法通过模型共享参数的迁移学习方法对提取的浅层特征图进行深度特征提取,得到泛化能力和表达能力更强的特征,从而提高分类精度。常用于迁移学习的模型有

LeNet、AlexNet、VGG (Visual geometry group)、Inception、残差网络 (ResNet) 及密集连接网络 (DenseNet)。其中, DenseNet 由密集学习模块 (Dense block) 重复堆积而形成,通过建立密集连接 (Dense connection) 实现密集学习^[17],使特征和梯度的传递更有效,网络更容易训练。常用的预训练数据集 ImageNet 可为迁移学习提取具有泛化能力的高级特征提供数据保障^[18],因此,本方法采用 DenseNet201 在 ImageNet 数据集上的预训练模型进行多尺度及多投影特征图的深度特征提取,如图 5 所示。其中,输入为多尺度及多投影特征图,通过 DenseNet201 的预训练模型生成 5 个 1920 维的深层特征。

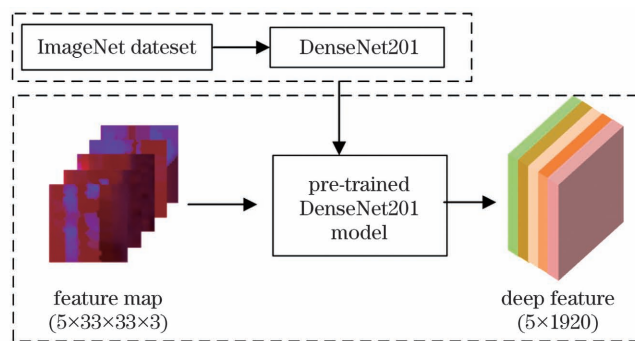


图 5 基于迁移学习的深度特征提取

Fig. 5 Deep feature extraction based on transfer learning

2.3 基于卷积神经网络的点云分类

点云特征图的深度特征由 DenseNet201 基于 ImageNet 的预训练模型迁移得到,但目前关于 ImageNet 数据集与点特征图的相关性还是未知的,因此,直接用点特征图提取的点特征进行点云分类的精度较低。文献[19]的研究结果表明,全连接网

络可以很好地分类深度特征。但相比全连接网络,卷积神经网络 (CNN) 可以在数据降维过程中进行有效的数据挖掘,同时减少计算成本、增加分类精度。因此,本方法参照基于深度学习模型的 PointNet^[20],设计了一个基于 FCN 的浅层深度学习模型,对深度特征进行分类,分类过程如图 6 所

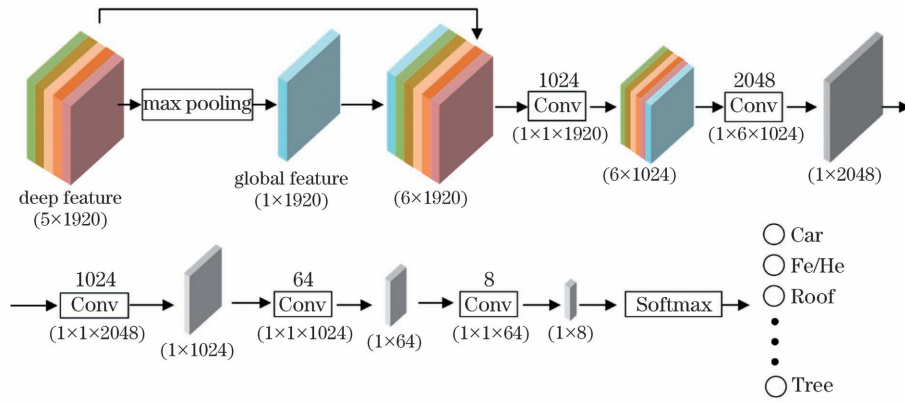


图 6 基于 FCN 的点云分类

Fig. 6 Point cloud classification based on FCN

示。首先,对迁移学习得到的深层特征(5×1920)进行最大池化(Max pooling)操作,以获取全局特征(1×1920);其次,用卷积层(Conv)对迁移学习得到的深层特征以及池化操作得到的全局特征进行降维(6×1024);然后,用 4 个卷积层对 6×1024 维的特征进行卷积运算,提取更深层次的有效特征(1×8);最后,采用输出层常用的 Softmax 函数得到每个激光脚点的类别标签。其中,卷积层上方的数字表示卷积核的数量,下方数字表示卷积核的尺寸,特征下方的数字表示卷积运算后特征的尺寸。

2.4 分类结果的后处理

上述分类过程中并未考虑上下文信息,容易使点云分类结果中含有噪声,从而影响分类精度。因此,本方法将机载 LiDAR 点云中的相邻点连接,从而构建一个以各个 LiDAR 点为顶点的无向图模型,并将能量函数 $E(L)$ 定义为

$$E(L) = \sum_{p \in P_a} D_p(l_p) + \lambda \sum_{(p,q) \in N} V_{p,q}(l_p, l_q) \cdot B_{(p,q)}, \quad (3)$$

式中,等式右侧第一项为数据项, L 为点云的标签 Label, $D_p(l_p)$ 为将点 p 标记为相应类别 l_p 时的代价,由 p 点及其邻域内点所属地物类别的概率计算得到;第二项为平滑项, $V_{p,q}(l_p, l_q)$ 为将相邻点 (p, q) 标记为相同类别或不同类别时的代价,当标记为相同类别时,其值为 0;反之,其值为 1。 $B_{(p,q)}$ 为 e 指数的倒数,其指数值对应于相邻点之间的三维空间距离。参数 λ 为数据项和平滑项对总能量函数的贡献, P_a 为所有点的集合, N 为邻域点的集合。

采用图割优化策略最小化能量函数^[21],从而在上下文信息约束下优化分类结果。相比其他后处理方法,图割优化的后处理方法能在点云密度不均匀

和含有噪声点的情况下实现复杂环境下的点云分类后处理^[22]。

3 实验与分析

3.1 实验数据

实验采用国际摄影测量与遥感协会(ISPRS)网站提供的数据集,包含机载 LiDAR 点云及其对应的多光谱航空影像^[11]。多光谱航空影像的地面分辨率为 8 cm,同时提供了影像的内外方位元素;机载 LiDAR 点云数据由徕卡 ALS50 系统采集,每平方米中约有 4 个点,每个激光脚点包括三维坐标、反射强度及类别标签等信息。数据集包含两块区域的数据,分别记为训练数据集和测试数据集,如图 7 所示。其中,训练数据集有 753876 个点、测试数据集有 411722 个点。该数据集将地物分为 9 个类别,包括电力线(Pow)、低矮植被(Low_veg)、不透水表面(Imp_sur)、车辆(Car)、栅栏(Fe/He)、屋顶(Roof)、外墙(Facade)、灌木(Shrub)、树木(Tree)。由于电力线的点数过少(仅有 1146 个点),无法进行有效的分析,因此,实验只对其他 8 类地物进行分析。

实验从包含类别标签的原始数据集中选取训练样本,经过反复实验、平衡样本数量以及分类精度的情况下,从每个类别中随机选取 2000 个点(共 16000 个点)作为训练样本,并用测试数据集进行实验分析。基于 FCN 的训练阶段,本方法采用可根据训练数据迭代更新神经网络权重的 Adam 优化算法,初始学习率设置为 0.0001,批次大小设置为 128,训练 100 个 epoch。用总体精度(OA)、F1 分数(F_1)以及平均 F1 分数(Avg F_1)作为精度评价指标。 F_1 是综合考虑查准率 P 以及查全率 R 的评价指标,可表示为

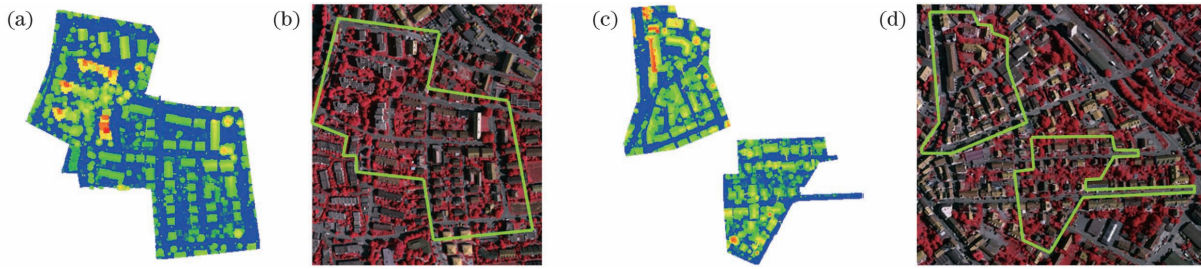


图 7 实验数据集。(a)按归一化高度显示的训练数据集;(b)训练数据集对应的航空影像;(c)按归一化高度显示的测试数据集;(d)测试数据集对应的航空影像

Fig. 7 Experimental datasets. (a) Training dataset displayed by normalized height; (b) aerial image corresponding to training dataset; (c) testing dataset displayed by normalized height; (d) aerial image corresponding to testing dataset

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R}, \quad (4)$$

式中, $P = T_p / (T_p + F_p)$, $R = T_p / (T_p + F_n)$, T_p 为预测该类别正确的样本数量, F_p 为预测该类别错误的样本数量, F_n 为该类别没有被预测出来的样本数量。

3.2 实验结果与分析

为了验证本方法的分类准确性,分别讨论了多尺度以及多投影特征、图割优化后处理以及预训练模型对分类精度的影响。

3.2.1 多尺度以及多投影特征对分类精度的影响

为验证本方法引入多尺度、多投影特征的有效性,分别采用不同特征组合进行点云分类,结果如图 8 所示。其中, S1(P1) 表示仅使用网格尺度为 0.1 m 的单尺度(单投影)特征进行实验; MS 表示采用三个尺度特征组合进行实验; MP 表示采用三个投影特征组合进行实验; MSMP 表示使用所有特征进行实验。可以发现,采用单尺度(单投影)特征分类时,各类别的 F1 分数最低;增加多尺度特征和多投影特征后,各类别的 F1 分数均有所增加,总体分类精度由 73.91% 分别增加到 80.50% 以及

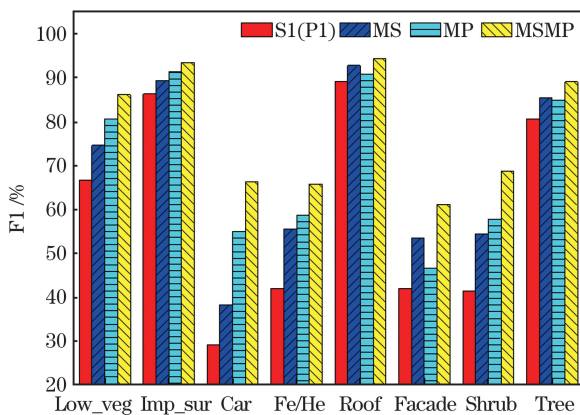


图 8 不同特征组合分类时的 F1 分数

Fig. 8 F1 scores when classifying different feature combinations

82.47%; 采用多尺度、多投影特征进行分类时,总体分类精度最高,可达到 87.76%。其中,汽车、栅栏、外墙以及灌木等地物的分类精度有明显提升。这表明引入多尺度以及多投影特征图,可对机载 LiDAR 点云进行更有效的表述,从而提高点云的分类精度。

3.2.2 不同预训练模型对分类精度的影响

为了验证本方法选择的预训练模型 DenseNet201 的有效性,采用不同的预训练模型进行实验,结果如图 9 所示。可以发现,采用 DenseNet201 预训练模型进行深度特征提取的分类结果最佳,除汽车外,各类别的 F1 分数均最大。其中,采用 DenseNet201 预训练模型的总体分类精度及平均 F1 分数分别为 87.76% 和 78.13%,比 VGG16、VGG19 和 ResNet50 预训练模型的结果分别高 2.57、1.51、1.90 个百分点及 3.41、2.10、2.22 个百分点。这表明采用 DenseNet201 预训练模型可以提取出泛化能力更强的深度特征,进一步提高分类精度。

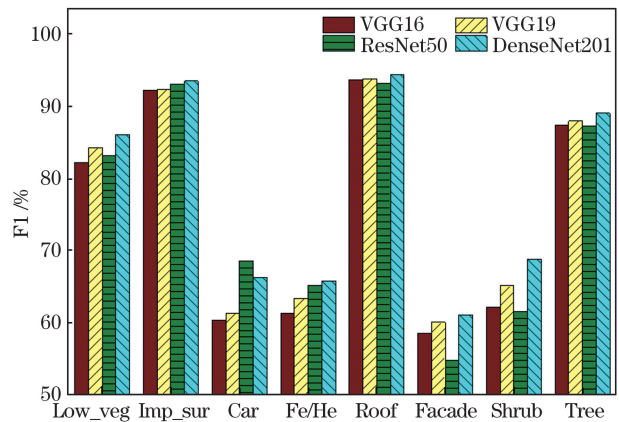


图 9 不同预训练模型分类时的 F1 分数

Fig. 9 F1 scores when classifying different pre-training models

3.2.3 图割优化策略对分类精度的影响

本方法采用图割优化策略对分类结果进行优化

时,需要设计适合的邻域构建无向图模型。其中,邻近点个数 K 的设定与点云密度相关,且会影响点云分类精度。因此,选择不同的 K 进行对比实验,从而确定 K 的最优值。同时将优化结果与未优化结果进行对比,以验证图割优化对分类精度影响,实验结果如表 1 所示。可以发现,未进行图割优化时的分类效果最差,总体分类精度及平均 F_1 分数最低。在进行图割优化的分类结果中,采用 4 个邻近点构建无向图模型的图割优化效果最佳,总体分类精度及平均 F_1 分数最高,可分别达到 89.91%和 82.45%。因此,

实验用 4 个邻近点构建无向图模型,最终的分类结果如图 10 所示。其中,矩形框为对应部分的放大图,以显示分类细节。相比未进行图割优化的分类结果,用 4 个邻近点构建无向图模型的图割优化结果在总体分类精度和平均 F_1 分数上分别提高了 2.15 和 4.32 个百分点。在汽车、栅栏、外墙及灌木类别中的优化效果更明显, F_1 分数分别提高了 13.67、6.01、6.40 及 3.29 个百分点。这表明对于一些点云密度较小、初始分类精度较低的点,在考虑上下文信息的情况下,分类精度可以得到较大程度的改善。

表 1 不同 K 对分类结果的影响

Table 1 Influence of different K on the classification results

unit: %

K	F1 score								OA	Avg F_1
	Low_veg	Imp_sur	Car	Fe/He	Roof	Facade	Shrub	Tree		
0	86.11	93.51	66.24	65.77	94.46	61.09	68.75	89.08	87.76	78.13
1	86.50	93.64	68.49	67.12	94.69	62.05	69.72	89.43	88.18	78.96
2	87.14	94.15	75.89	68.96	95.30	64.57	71.19	89.99	89.12	80.90
3	87.15	93.99	78.44	71.00	95.76	65.91	72.60	90.69	89.60	81.94
4	87.42	94.09	79.91	71.78	96.06	67.49	72.04	90.80	89.91	82.45
5	87.15	93.69	79.62	70.66	96.24	67.23	70.73	90.69	89.76	82.00
6	86.58	93.19	73.60	70.22	96.03	65.22	68.44	90.53	89.27	80.48

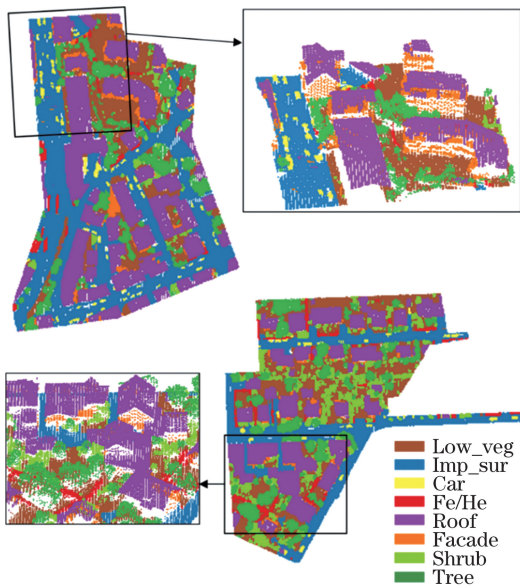


图 10 $K=4$ 时的分类结果

Fig. 10 Classification results when $K=4$

为了直观体现图割优化对于错误点的纠正,对优化前后的错误分类结果图进行对比,结果如图 11 所示。其中,阴影为错误分类点。可以发现,图 11(b)中校正了图 11(a)中很多分类错误的噪声点,特别是图中矩形框和椭圆框标注的区域。

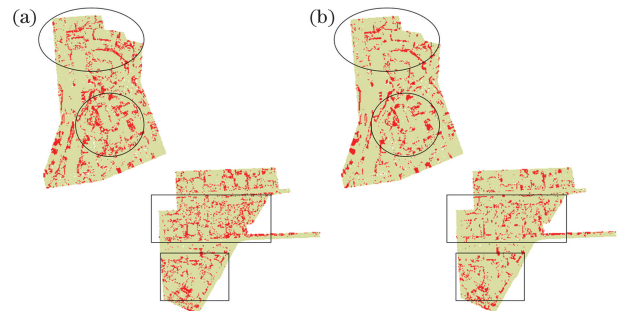


图 11 错误分类结果的对比。(a)图割优化前的错误分类结果;(b)图割优化后的错误分类结果

Fig. 11 Comparison of the misclassification results.

(a) Misclassification result before graph-cuts optimization; (b) misclassification result after graph-cuts optimization

3.3 不同分类方法的对比分析

为了验证本方法在分类任务中的优势,将本方法的分类结果与 ISPRS 网站上已上传的分类结果以及其他迁移学习点云分类方法的分类结果进行了对比分析。

3.3.1 与 ISPRS 竞赛方法的对比分析

将本方法与 ISPRS 竞赛中排名前 8 的方法 (UM^[23]、WhuY2、WhuY3^[24]、LUH^[25]、BIJ_W^[26]、

RIT_1^[14]、NANJ2^[27]、WhuY4^[28]) 进行对比, 结果如表 2 所示。可以发现, 除了低矮植被以及未分类的电力线, 本方法的分类精度均是最高的。特别是在汽车、栅栏、外墙、灌木以及树木的分类结果中远高于其他分类方法, 相比第 2 名算法的 F1 分数, 分别提高了 5.2、18.1、11.2、16.1 以及 7.7 个百分点, 总体精度提高了 4.7 个百分点。原因是相比 WhuY2、WhuY3、NANJ2、WhuY4 使用特征图进行点云分类的方法, 本

方法考虑了三维信息的特征图生成方法, 同时采用 DenseNet201 的预训练模型对特征图进行深度特征提取, 可从浅层特征图中挖掘更有效的特征, 提高分类精度; 相比 UM、LUM 采用传统机器学习的方法, 本方法采用迁移学习及深度学习的方式进行点云分类; 相比 RIT_1 直接对点云进行三维深度学习, 本方法采用 DenseNet201 的预训练模型对点特征图进行训练, 可以提取更具有表征能力的深度特征。

表 2 不同方法的 F1 分数及总体分类精度

Table 2 F1 scores and overall classification accuracy of different methods

unit: %

Method	F1 score									OA
	Pow	Low_veg	Imp_surf	Car	Fe/He	Roof	Facade	Shrub	Tree	
UM	46.1	79.0	89.1	47.7	5.2	92.0	52.7	40.9	77.9	80.8
WhuY2	31.9	80.0	88.9	40.8	24.5	93.1	49.4	41.1	77.3	81.0
WhuY3	37.1	81.4	90.1	63.4	23.9	93.4	47.5	39.9	78.0	82.3
LUH	59.6	77.5	91.1	73.1	34.0	94.2	56.3	46.6	83.1	81.6
BIJ_W	13.8	78.5	90.5	56.4	36.3	92.2	53.2	43.3	78.4	81.5
RIT_1	37.5	77.9	91.5	73.4	18.0	94.0	49.3	45.9	82.5	81.6
NANJ2	62.0	88.8	91.2	66.7	40.7	93.6	42.6	55.9	82.6	85.2
WhuY4	42.5	82.7	91.4	74.7	53.7	94.3	53.1	47.9	82.8	84.9
Ours	-	87.4	94.1	79.9	71.8	96.1	67.5	72.0	90.8	89.9

为了对比本方法与 ISPRS 竞赛中总体分类精度最高的 NANJ2 方法的分类效果, 用训练好的模型对未参与训练的测试数据集进行测试, 结果如表 3 所示。可以发现, 本方法的总体分类精度及平均 F1 分数均高于 NANJ2 方法, 分别提高了 1.0 和 1.6 个百分点。除汽车外, 本方法对其他所有类别

的 F1 分数均高于 NANJ2 方法。值得注意的是, NANJ2 方法是在每个类别中随机取 4000 个点作为训练样本, 并采用 CNN 进行特征图深度特征提取, 同时对分类结果进行后处理。相比 NANJ2 方法, 本方法在小样本基础上的分类精度更高, 进一步验证了本方法的优越性。

表 3 本方法与 NANJ2 方法的分类结果

Table 3 Classification results of our method and NANJ2 method

unit: %

Method	F1 score								OA	Avg F_1
	Low_veg	Imp_surf	Car	Fe/He	Roof	Facade	Shrub	Tree		
NANJ2	86.9	93.1	69.5	70.9	95.2	73.0	65.3	92.1	89.1	80.8
Ours	88.6	93.8	68.7	75.5	95.6	74.0	68.8	94.3	90.1	82.4

3.3.2 不同迁移学习方法的对比分析

为了验证本方法在同类型方法中的优越性, 将本方法与其他基于迁移学习的点云分类方法进行对比, 包括 DRN^[15] 以及 DRN-1^[29], 所有方法均采用相同的样本选择策略及测试数据, 结果如表 4 所示。可以发现, 本方法的总体分类精度及平均 F1 分数均高于其他方法。除不透水表面及树木 2 个类别外, 本方法的 F1 分数最高。特别是对汽车及栅栏

的分类结果, 相比 DRN-1 方法, 本方法的 F1 分数分别提高了 37.7 及 6.1 个百分点。这表明相比提取点云几何特征生成多视角特征图并采用 ResNet50 预训练模型、用 CNN 进行特征提取及分类的方法, 本方法采用的光谱特征更稳定, 能更好地表达点云的三维空间特征, 且本方法采用的 DenseNet201 预训练模型及 FCN 可以提取更有效的深度特征, 从而获得更准确的分类结果。

表 4 不同迁移学习方法的分类结果

Table 4 Classification results of different transfer learning methods

unit: %

Method	F1 score								OA	Avg F_1
	Low_veg	Imp_surf	Car	Fe/He	Roof	Facade	Shrub	Tree		
DRN	83.3	92.5	52.4	62.5	95.2	62.8	64.9	88.7	86.8	75.3
DRN-1	87.0	94.3	42.2	65.7	96.0	66.0	71.7	91.8	89.5	78.2
Ours	87.4	94.1	79.9	71.8	96.1	67.5	72.0	90.8	89.9	82.4

3.3.3 分类效率的对比

除了分类精度的对比外,实验还对比了不同方法的分类效率。直接基于点云的深度学习分类方法 RIT_1 需要全部的训练数据集(约 75 万个点)作为训练样本、基于特征图的点云分类方法 WhuY3 需要约 10 万个样本点、基于迁移学习的点云分类方法^[3]需要约 14 万个样本点,而本方法仅采用 16000 个样本点(约占数据集的 1.4%)就可以完成点云的高精度分类。文献[24,27-28]中的分类方法需要 10 h 以上的训练时间,而本方法可以在 20 min 内完成整个训练过程。这表明本方法可以在较小样本的情况下,获得较高的点云分类精度。

4 结 论

提出了一种整合迁移学习与 FCN 的小样本机载 LiDAR 点云高精度分类方法。首先,引入多尺度及多投影特征图对三维点的空间特征进行更准确的表达;然后,采用迁移学习对特征图进行深度特征的提取;最后,通过构建 FCN 及图割优化后处理实现机载 LiDAR 点云的准确分类。实验结果表明,本方法在仅使用每类别 2000 个样本点的情况下,可获得较理想的分类结果,总体精度达到 89.9%,且能显著提高汽车、栅栏等空间特征丰富的地物分类精度。由于整个训练过程是分步进行的,在保存数据及加载数据的过程也会消耗一定的时间。因此,下一步研究还需将整个过程进行整合,以进一步缩短训练时间,提高分类效率。此外,本方法对电力线等稀疏点云的分类效果较差,因此,还需针对该问题进行改进,以提高本方法的适用性。

参 考 文 献

- [1] Zhao Z F. Research on road extraction based on airborne LiDAR point cloud[D]. Kunming: Kunming University of Science and Technology, 2012.
赵振峰. 基于机载 LiDAR 点云的道路提取研究[D]. 昆明: 昆明理工大学, 2012.
- [2] Wang H T, Lei X D, Zhao Z Z. 3D deep learning

classification method for airborne LiDAR point clouds fusing spectral information[J]. Laser & Optoelectronics Progress, 2020, 57(12): 122802.

王宏涛, 雷相达, 赵宗泽. 融合光谱信息的机载 LiDAR 点云三维深度学习分类方法[J]. 激光与光电子学进展, 2020, 57(12): 122802.

- [3] Zhao C, Zhang B M, Yu D H, et al. Airborne LiDAR point cloud classification using transfer learning[J]. Optics and Precision Engineering, 2019, 27(7): 1601-1612.

赵传, 张保明, 余东行, 等. 利用迁移学习的机载激光雷达点云分类[J]. 光学精密工程, 2019, 27(7): 1601-1612.

- [4] Zhang J X, Lin X G, Ning X G. SVM-based classification of segmented airborne LiDAR point clouds in urban areas[J]. Remote Sensing, 2013, 5(8): 3749-3775.

- [5] Shi X S, Cheng Y L, Zhao Z Y, et al. Point cloud classification algorithm based on IPTD and SVM[J]. Laser & Optoelectronics Progress, 2019, 56(16): 161002.

释小松, 程英蕾, 赵中阳, 等. 基于三角网滤波和支持向量机的点云分类算法[J]. 激光与光电子学进展, 2019, 56(16): 161002.

- [6] He M Y, Cheng Y L, Liao X J, et al. Building extraction algorithm by fusing spectral and geometrical features [J]. Laser & Optoelectronics Progress, 2018, 55(4): 042803.

何曼芸, 程英蕾, 廖湘江, 等. 融合光谱特征和几何特征的建筑物提取算法[J]. 激光与光电子学进展, 2018, 55(4): 042803.

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

- [8] Ramiya A M, Nidamanuri R R, Krishnan R. Object-oriented semantic labelling of spectral-spatial LiDAR point cloud for urban land cover classification and buildings detection[J]. Geocarto International, 2016, 31(2): 121-139.

- [9] Ni H, Lin X G, Zhang J X. Classification of ALS

- point cloud with improved point cloud segmentation and random forests [J]. *Remote Sensing*, 2017, 9 (3): 288.
- [10] Wei S F. Scene understanding based on point clouds and 2D images [D]. Beijing: University of Chinese Academy of Sciences, 2016.
魏书法. 基于三维点云与二维图像的场景理解 [D]. 北京: 中国科学院大学(中国科学院深圳先进技术研究院), 2016.
- [11] Niemeyer J, Rottensteiner F, Soergel U. Contextual classification of lidar data and building object detection in urban areas [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 87: 152-165.
- [12] Yang S J, Zhang K S, Shao Y S. Classification of airborne LiDAR point cloud data based on multiscale adaptive features [J]. *Acta Optica Sinica*, 2019, 39 (2): 0228001.
杨书娟, 张珂殊, 邵永社. 基于多尺度自适应特征的机载 LiDAR 点云分类 [J]. *光学学报*, 2019, 39(2): 0228001.
- [13] Zhang W M, Qi J B, Wan P, et al. An easy-to-use airborne LiDAR data filtering method based on cloth simulation [J]. *Remote Sensing*, 2016, 8(6): 501.
- [14] Yousefhusien M, Kelbe D J, Ientilucci E J, et al. A multi-scale fully convolutional network for semantic labeling of 3D point clouds [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, 143: 191-204.
- [15] Zhao C, Guo H T, Lu J, et al. ALS point cloud classification with small training data set based on transfer learning [J]. *IEEE Geoscience and Remote Sensing Letters*, 2020, 17(8): 1406-1410.
- [16] Donahue J, Jia Y, Vinyals O, et al. DeCAF: a deep convolutional activation feature for generic visual recognition [C] // International Conference on Machine Learning (ICML), June 21-26, 2014, Beijing, China. Beijing: JMLR. org, 2014: 647-655.
- [17] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [18] Kornblith S, Shlens J, Le Q V. Do better ImageNet models transfer better? [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 2656-2666.
- [19] Zhang C L, Luo J H, Wei X S, et al. In defense of fully connected layers in visual representation transfer [M] // Zeng B, Huang Q M, Saddik A E, et al. *Advances in multimedia information processing-PCM 2017. Lecture notes in computer science*. Cham: Springer, 2018, 10736: 807-817.
- [20] Charles R Q, Hao S, Mo K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 77-85.
- [21] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(9): 1124-1137.
- [22] Nguyen A, Le B. 3D point cloud segmentation: a survey [C] // 2013 6th IEEE Conference on Robotics, Automation and Mechatronics (RAM), November 12-15, 2013, Manila, Philippines. New York: IEEE Press, 2013: 225-230.
- [23] Horvat D, Žalik B, Mongus D. Context-dependent detection of non-linearly distributed points for vegetation classification in airborne LiDAR [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 116: 1-14.
- [24] Yang Z S, Jiang W S, Xu B, et al. A convolutional neural network-based 3D semantic labeling method for ALS point clouds [J]. *Remote Sensing*, 2017, 9 (9): 936.
- [25] Niemeyer J, Rottensteiner F, Soergel U, et al. Hierarchical higher order crf for the classification of airborne lidar point clouds in urban areas [J]. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2016, XLI-B3: 655-662.
- [26] Wang S L, Suo S, Ma W C, et al. Deep parametric continuous convolutional neural networks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 2589-2597.
- [27] Zhao R B, Pang M Y, Wang J D. Classifying airborne LiDAR point clouds via deep features learned by a multi-scale convolutional neural network [J]. *International Journal of Geographical Information Science*, 2018, 32(5): 960-979.
- [28] Yang Z S, Tan B, Pei H K, et al. Segmentation and multi-scale convolutional neural network-based classification of airborne laser scanner data [J]. *Sensors*, 2018, 18(10): 3347.
- [29] Zhao C, Guo H T, Lu J, et al. Airborne LiDAR

point cloud classification based on deep residual network [J]. *Acta Geodaetica et Cartographica Sinica*, 2020, 49(2): 202-213.

赵传, 郭海涛, 卢俊, 等. 基于深度残差网络的机载 LiDAR 点云分类[J]. *测绘学报*, 2020, 49(2): 202-213.

Small-Sample Airborne LiDAR Point Cloud Classification Based on Transfer Learning and Fully Convolutional Network

Lei Xiangda, Wang Hongtao, Zhao Zongze

School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo, Henan 454000, China

Abstract

Significance The ability of airborne light detection and ranging (LiDAR) to obtain high-precision and high-density three-dimensional (3D) point cloud for a large area has been widely used in many fields such as surveying and mapping, forestry, and the electric power industry. In these applications, detecting specific targets such as buildings, trees, and power lines from the LiDAR point cloud is essential, which is usually considered a classification procedure. However, accurately identifying various typical surface objects from discretely and irregularly distributed LiDAR point clouds is challenging.

Recently, many deep learning-based methods have achieved good performance in airborne LiDAR point cloud classification. Some scholars have employed convolutional neural networks for point cloud classification by transforming the unevenly distributed point cloud into regular images or voxels. In addition, researchers have applied some deep learning-based methods such as PointNet and PointCNN to the original 3D point cloud. In these methods, a large number of training samples are often required; moreover, they are manually labeled, which is time-consuming and laborious, thus inhibiting their wide use in complex scenarios under different conditions. To solve this problem, a small-sample airborne LiDAR point cloud classification method based on integrating transfer learning (TL) and a fully convolutional network (FCN) is proposed in this study.

Progress The proposed classification method is summarized in Fig. 1. First, a colorful airborne LiDAR point cloud is derived using multispectral images; then, the normalized elevation, intensity values, and normalized difference vegetation index of each 3D point cloud are extracted to derive a three-channel point cloud feature map (Fig. 2). Next, multiscale and multiprojection feature maps are constructed by setting different grid sizes (0.1, 0.3, and 0.5 m) (Fig. 3) and projection directions (X , Y , and Z directions) (Fig. 4), which are fed into a pretrained DenseNet201 model to extract deep features of LiDAR point clouds (Fig. 5). Finally, the global features are extracted by applying a max-pooling operation to the extracted deep features for extending the feature description; the max-pooled features are used as the input for an FCN to achieve an initial classification result (Fig. 6), and then the classification result is improved using a graph-cut optimization algorithm.

Results and Discussions Benchmark datasets provided by the International Society of Photogrammetry and Remote Sensing (ISPRS) are used to verify the proposed method. 1) We used four combinations of features to analyze the effects of multiscale and multiprojection features on classification accuracy: a single feature (S1/P1), multiscale features (MS), multiprojection features (MP), and multiscale and multiprojection features (MSMP) (Fig. 8). Experiment results show that the MSMP features had higher per-class F1 scores and overall accuracy than other feature combinations, which could be explained as the MSMP features being able to provide a more effective expression of airborne LiDAR point clouds. 2) To select more suitable pretrained models for classification, we experimentally validated pretrained VGG16, VGG19, ResNet50, and DenseNet201 models (Fig. 9). The contrast result show that the pretrained DenseNet201 model performed better than other pretrained models owing to its ability to extract deep features with better generalization. 3) To validate the effectiveness of graph-cut optimization, we performed graph-cut optimization by setting different numbers of adjacent points K and then chose the best optimization result with the number of adjacent points $K = 4$ for a comparison with the result without graph-cut optimization (Fig. 11). The contrast result show that the graph-cut optimization algorithm could effectively correct

misclassification points and improve classification accuracy. 4) The proposed method is compared with methods reported on the ISPRS website and TL-based methods, showing better results in overall accuracy and per-class F1 scores, except for powerline and impervious surfaces, than the methods reported on the ISPRS website. It also achieved better overall accuracy, average F1 score, and per-class F1 scores, except for low vegetation and tree, than the TL-based methods.

In general, the aforementioned experiment results show that the overall accuracy of the proposed method is 89.9% when only per-class 2000 points are used as the training sample (approximately 1.4% of the dataset), which demonstrates that the proposed method can acquire high-precision point cloud classification with fewer training samples and in less time.

Conclusions and Prospect In this study, an airborne LiDAR point cloud classification method based on integrating TL and FCN with small samples is proposed. In the proposed method, multiscale and multiprojection feature maps are first introduced to represent the spatial characteristics of 3D points accurately; then, TL is used to extract deep features from feature maps; finally, the deep features are fed into an FCN for the initial classification of LiDAR point clouds, followed by postprocessing with graph-cut optimization to achieve more accurate classification results. Because many steps are involved in the entire classification process—such as processes of saving and loading data—which prolong the classification time, the entire process will be integrated to accelerate training speed and improve data processing efficiency in the future.

Key words remote sensing; airborne LiDAR point cloud; spectral data; transfer learning; small samples; graph-cuts optimization

OCIS codes 280.4788; 280.3640; 100.4996