

# 基于改进的PointNet++模型的多光谱 LiDAR数据分类方法

景庄伟, 丁荣莉, 何恒翔, 李丰, 谷岳

上海航天电子技术研究所, 上海

收稿日期: 2023年11月14日; 录用日期: 2024年1月22日; 发布日期: 2024年1月30日

## 摘要

多光谱激光雷达(LiDAR)系统可同时并快速获取大范围空间目标地物的光谱强度信息和空间几何信息, 为三维点云分类、语义分割、目标检测等研究提供新的数据源。然而, 由于多光谱点云数据分布的不规则性以及数据量巨大等特性, 使得地物特征的提取过程充满挑战。本文通过将通道注意力机制(SE-Block)和修正后的焦点损失函数嵌入至PointNet++网络中, 提出了一种改进的PointNet++网络架构。PointNet++网络从不均匀采样的点中提取局部特征, 并通过多尺度分组表示点之间的局部几何关系。将SE-Block嵌入至PointNet++网络中, 通过显式地建模通道之间的相互依赖关系, 自适应地重新校准通道方面的特征响应, 从而强调重要通道并抑制不利于预测的无用通道, 提高特征的显著性, 以便更好地进行点云分类。另外, 本文在改进的网络架构基础上利用修正后的焦点损失函数解决了多光谱LiDAR点云数据中类别不均匀分布的问题。本文提出的改进的PointNet++网络架构在托伯莫里港口数据集上进行了评估, 获得的总体精度、mIoU、F<sub>1</sub>-score和Kappa系数分别为95.21%、62.59%、73.58%、0.918。与5个已建立的深度神经网络模型的比较实验证实, 本文提出的改进的PointNet++网络架构在多光谱LiDAR点云分类任务中具有良好的性能。

## 关键词

点云分类, 通道注意力机制, PointNet++模型, 多光谱LiDAR数据

# Multispectral LiDAR Data Classification Method Based on an Improved PointNet++ Model

Zhuangwei Jing, Rongli Ding, Hengxiang He, Feng Li, Yue Gu

Shanghai Aerospace Electronic Technology Institute, Shanghai

Received: Nov. 14<sup>th</sup>, 2023; accepted: Jan. 22<sup>nd</sup>, 2024; published: Jan. 30<sup>th</sup>, 2024

文章引用: 景庄伟, 丁荣莉, 何恒翔, 李丰, 谷岳. 基于改进的 PointNet++模型的多光谱 LiDAR 数据分类方法[J]. 测绘科学技术, 2024, 12(1): 64-76. DOI: 10.12677/gst.2024.121009

## Abstract

A multispectral light detection and ranging (LiDAR) system can simultaneously and quickly collect spectral intensity information and spatial geometric data of a large range of space objects, which provides a new data source for the research of 3D point cloud classification, semantic segmentation and object detection. However, due to the irregularly distributed property of multispectral point cloud data and massive data volume, the extraction process of land cover is full of challenges. In this paper, we propose an improved PointNet++ network architecture by embedding the Squeeze and Excitation Block (SE-Block) and a modified focal loss function into the PointNet++ network. PointNet++ network extracts local features from unevenly sampled points and represents local geometrical relationships among the points through multi-scale grouping. SE-Block is embedded into PointNet++ network, which explicitly models the interdependence between channels and adaptively recalibrates the feature response in terms of channels, emphasizing important channels and suppressing useless channels that are not conducive to prediction, improving the saliency of features for better point cloud classification. In addition, based on the improved network architecture, this article utilizes the modified focal loss function to solve the problem of uneven distribution of categories in multispectral LiDAR point cloud data. The improved PointNet++ network architecture proposed in this paper has been evaluated on the Tobermory Port dataset and achieved an overall accuracy, a mean Intersection over Union (mIoU), an F<sub>1</sub>-score, and a Kappa coefficient of 95.21%, 62.59%, 73.58% and 0.918, respectively. Comparative studies with five established deep neural network models confirm that the improved PointNet++ network architecture proposed in this paper has good performance in multispectral LiDAR point cloud classification tasks.

## Keywords

Point Cloud Classification, Squeeze and Excitation Block, PointNet++ Model, Multispectral LiDAR Data

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

机载单通道激光雷达能够快速获取大规模和高精度的地球表面信息,并形成三维点云数据,在地物分类、地形测绘、城市规划、环境监测等许多领域中得到了广泛的应用[1] [2]。基于三维高度信息和单波段红外强度信息,点云分类目前已成为遥感、摄影测量和计算机视觉领域的重要研究方向[3] [4] [5]。然而,机载单通道激光雷达获取的点云数据缺乏丰富的光谱信息,在细粒度点云分类(例如,草、裸土和道路)中的结果并不理想,通常只能粗略地分为几个类别(例如,建筑物、树木和地面)。因此,激光雷达数据通常与光学图像数据相结合,以更好地实现点云数据的分类[6] [7] [8] [9] [10]。尽管激光雷达数据与图像数据的融合提高了点云分类的精度,但研究过程中存在的一些问题(例如,几何配准、目标遮挡等),仍然未得到有效的解决。

多光谱激光雷达系统除了获取三维空间数据外,同时还收集了不同波段强度的信息,在不进行数据融合的情况下提供了更全面、一致的光谱信息,为目标提供多属性特征。目前的研究表明,多光谱点云数据在更细的尺度上可实现更好的地物分类结果[11]。近年来,多光谱点云数据被越来越多地应用于分类研究中,根据数据的处理类型,这些研究主要可以分为两大类:1) 基于二维多光谱特征图像的方法,2) 基

于三维多光谱点云的方法。基于二维多光谱特征图像的方法是分别栅格化 Titan 多光谱激光雷达点云数据的光谱信息和高度信息,并在图像上对土地利用类型进行分类。该分类方法将三维数据转换为二维数据,降低了数据处理的复杂性,适用于大面积土地覆盖分类工作。至今,在该情况下学者们已进行了大量的研究工作,可以获得 90% 以上的分类准确率[12]-[18]。然而,点云栅格化为图像的过程中会造成空间信息和光谱信息的丢失,一些物体(例如,植被覆盖下的地面、电线杆、围栏等)的转换出现错误,从而导致分类的不完整,甚至出现不正确的分类结果。此外,已有研究指出,基于点的分类效果优于基于图像的分类效果[19] [20]。因此,本研究将基于多光谱 LiDAR 点云进行三维点云场景的分类研究。

为了充分利用多光谱激光雷达的空间光谱信息,需要合理有效地进行特征提取。目前,对于多光谱激光雷达数据的基于点的分类研究较少[21]-[27]。在多光谱激光雷达点云数据上,Wichmann 等人[21]开创了基于三维多光谱点云的方法,他们重点研究了特定目标的光谱特征,探索了多光谱 LiDAR 点云在地物分类中的潜力。此外,一些多光谱点云分类研究表明,基于三维多光谱激光雷达点云的方法优于基于二维多光谱特征图像的方法,OA 在[22]中提高了 10%,在[19]中提高了 3.8%。袁鹏飞等人[23]结合多光谱 LiDAR 点云数据中的光谱信息高度信息,提取了基于强度、密度和平坦度等统计特征,并采用随机森林分类器实现了点云中道路面点云和非道路面点云的有效提取。景庄伟等人[24]基于多光 LiDAR 点云数据的多通道和多次散射回波的典型特性,生成了感兴趣的热力图,根据热力图特征值和 nDSM 辅助数据对感兴趣地物进行了分类。再采用 CNN 模型学习结果与面向对象影像分析 OBIA 分类方法相结合对分类结果进行精化。Wang 等人[18]提出了一种基于张量表示的三维地物测绘分类模型,该模型利用多光谱点及其邻域点将每个点表示为二阶张量,并引入 TMDE 算法获取低维空间和光谱特征用于后续分类,表明空间邻域信息在分类中具有关键作用。上述研究仅使用 MS-LiDAR 直接提供的高度和强度特征,简单的指数特征(例如,植被指数和水指数)和几何特征(例如,线性和平面度)等低级特征对点云进行分类。近期的研究已经从主要的光谱强度特征转向开发更深层的特征挖掘,依据提取的深层特征信息学习推理出原始点云中点的标签,从而实现高效的分类。Li 等人[25]使用 GGM (Graph Geometric Moments)卷积显式地编码点集的局部几何结构,提出了一种可在多光谱 LiDAR 点云中高效分类出建筑物的网络架构。Briechele 等人[26]利用 PointNet++ 深度神经网络,结合基于无人机的激光雷达数据和多光谱图像,对多树种和直立枯树进行了分类。Shi 等人[27]采用邻域点选取的方法提取了多尺度的空域-光谱邻域特征,并对这些特征进行了级联,得到了多光谱 LiDAR 数据的多尺度邻域空域-光谱特征。由此可见,特征的提取和选择在点云分类中起着重要的作用,但是目前缺少简单的方法来预先确定最佳特征数和最合适的特征,以确保鲁棒的点云分类精度。因此,为了进一步提高点云分类精度,本文将探索直接在三维多光谱激光雷达点云上进行点云分类的深度学习方法。

基于深度学习的方法通常在规则结构化的数据(例如,图像)上进行处理[28] [29] [30] [31] [32]。本文中,我们提出了一个改进的 PointNet++ 架构,该架构将 SE-Block (Squeeze and Excitation) [33]集成至 PointNet++ 中,通过显式地建模通道之间的相互依赖关系来重新校准通道方面的特征响应,能够充分顾及地物空间几何结构复杂性及地物光谱反映差异性,得到更为稳健的地物特征表达效果。另外,本文在改进的网络架构基础上利用修正后的焦点损失函数解决了多光谱 LiDAR 点云数据中类别不均匀分布的问题。本文从 2015 年 Optech Titan 采集的多光谱 LiDAR 数据中选择了 14 个代表性的区域,并手动进行标注,构建了托伯莫里港口数据集。本文将提出的改进的 PointNet++ 网络架构在托伯莫里港口数据集上进行了评估,同时设计了对比实验,研究多光谱 LiDAR 数据的可行性和所提出网络架构在点云分类任务中的优越性。

## 2. 研究方法

### 2.1. 研究区域及数据集

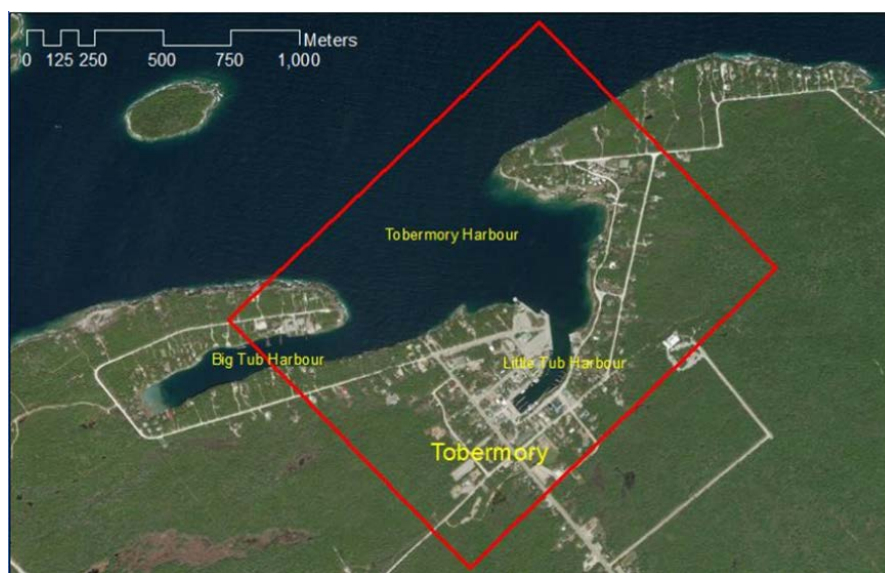
本文研究区域是位于加拿大安大略省北布鲁斯半岛市里的一个小镇——托伯莫里,该城镇中心区域

的经纬度为(45.24°N, 81.68°W)。托伯莫里平均高程约为 186 m, 地势较为平坦。图 1 为本文数据集的研究区域, 该区域主要包括森林区域、水域、居民区和商业区。通过谷歌地图以及目视检查数据集发现, 该研究区主要由大量的树木、开阔的草地、山形屋顶建筑、沥青道路、混凝土码头、水体和一些电力线组成, 为点云分类研究提供了足够多样的类别。本文使用的数据集由 Optech 公司生产的 Titan 多光谱 LiDAR 传感器于 2015 年 4 月采集得到。飞机在高于地面 457 m 的高度飞行, 速度为 259 公里/小时。Titan 传感器的扫描角度为 $\pm 20^\circ$ , 该系统以 625 kHz 的脉冲重复频率和  $40^\circ$  的扫描角(FOV)运行, 通过三个通道获取 LiDAR 点云数据, 平均点密度约为  $58 \text{ pts/m}^2$ 。表 1 中详细介绍了数据收集参数。Titan 多光谱 LiDAR 传感器包含了 1550 nm (C1)、1064 nm (C2)和 532 nm (C3)三个通道, 每个通道分别代表中红外(通道 1)、近红外(通道 2)和可见光(通道 3)波段。该数据集共采集了 10 条航带数据, 覆盖范围大约 25 平方公里。

**Table 1.** Specifications of the Titan multispectral LiDAR system

**表 1.** Titan 多光谱 LiDAR 系统的规格

参数(单位)	C1	C2	C3
波长(nm)	1550	1064	532
前向角( $^\circ$ )	3.5	0	7
波段	SWIR	NIR	GREEN
扫描视场角( $^\circ$ )	$\pm 20$	$\pm 20$	$\pm 20$
脉冲重复频率(kHz)	50~300	50~300	50~300
扫描频率(Hz)	210	210	210
飞行高度(m)	1000~1100	1000~1100	1000~1100
平均点间距(m)	0.5	0.5	0.5
平均点密度(points/m <sup>2</sup> )	3.6	3.6	3.6



**Figure 1.** Study area of the Tobermory Port dataset

**图 1.** 托伯莫里港口数据集的研究区域

## 2.2. 数据预处理

为了构建可供深度神经网络架构输入的数据集,需要进行两步数据预处理操作(即数据融合和数据注释),如图 2 所示,数据融合旨在合并 532 nm、1024 nm 和 1550 nm 波长进入一个单点云,其中每个点都包含坐标信息和三个波长强度值信息:  $[LRI_{\lambda_{1550}}, LRI_{\lambda_{1064}}, LRI_{\lambda_{532}}]$ 。数据注释旨在将多光谱点云数据分类为六个类别(例如,道路、草地、树木、水体、建筑物、电力线),并获得可用于深度学习模型训练的点云数据集。

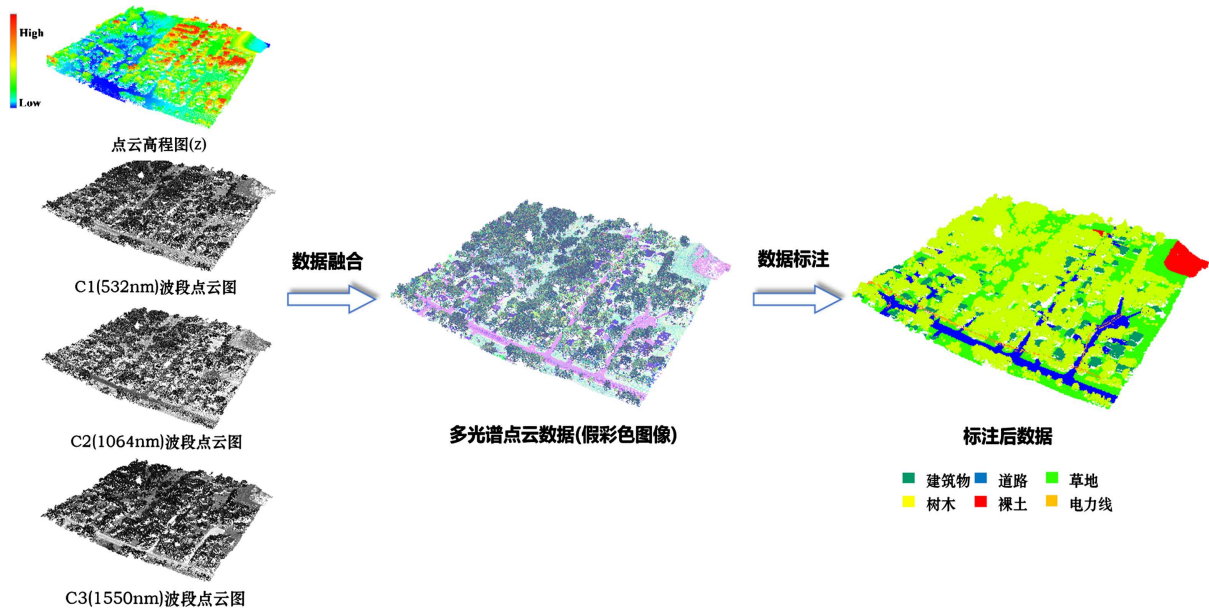


Figure 2. Preprocessing process of multispectral LiDAR data

图 2. 多光谱 LiDAR 数据的预处理过程

### 2.2.1. 数据融合

Titan 多光谱 LiDAR 数据由三个独立的点云组成,对应 532 nm、1024 nm 和 1550 nm 三个激光波长,为了充分利用三个波段的反射特性并提高点云的点密度,首先将三个独立的点云合并为一个单一的高密度多光谱点云,每个点包含其自身的强度和其他两个指定的其他波段的强度。由于三个通道采集的激光束偏离最低点方向(+3.5°、0°、+7°)各不相同,导致不同通道采集到同一目标物体的点可能不会完全重合在同一个位置。因此,本研究采用三维空间连接技术解决不完全重合的问题,在三个点云中,每个单波段点云依次作为参考,通过最近邻搜索算法对参考点云的每个点进行处理,从而在另外两个波段的点云中找到其邻域点,然后通过双线性插值方法从另外两个波段的邻域点上获得强度值,接着将合并后的点云中重复的点进行移除,最后得到的多光谱 LiDAR 点云数据包含坐标信息和三个波长强度值信息:  $[x, y, z, LRI_{\lambda_{1550}}, LRI_{\lambda_{1064}}, LRI_{\lambda_{532}}]$ 。

### 2.2.2. 数据标注

经过数据融合后,得到了可用的多光谱 LiDAR 数据,我们从融合后的多光谱 LiDAR 点云数据集中选择了 14 个代表性的测试区域,其中包括六个类别(即道路、草地、树木、水体、建筑物、电力线)。14 个区域约由 1017 万个点组成。作为一种有监督的方法,我们必须在将每个选定的训练和测试区域输入网络模型之前手动标记它们。本研究通过 CloudCompare 软件,手动逐点标记六个类别。每个选定区

域保存在单独的文件中, 其中每个点包含七个属性: 坐标 $(x, y, z)$ 、三个波长 $(LRI_{\lambda_{1550}}, LRI_{\lambda_{1064}}, LRI_{\lambda_{632}})$ 和类别标签。

## 2.3. 深度神经网络

### 2.3.1. 训练样本生成

受 RandLA-Net [34]的启发, 我们提出了一种结合最远点采样和  $k$  最近邻点(FPS-KNN)的样本生成方法, 该方法可有效保持场景的完整性, 保证原始场景中的每一个点都被标记, 从而生成符合神经网络输入需求的训练样本和测试样本。FPS-KNN 生成的样本既满足标准卷积神经网络的输入数据形式要求, 又能获得场景的全覆盖。图 3 显示了 FPS-KNN 方法的数据处理工作流程。我们将输入的数据定义为点集  $S = \{p_1, \dots, p_n | p_i = (x, y, z, f)\}$ , 其中  $x, y, z$  为坐标值,  $f$  为特征向量, 如颜色、强度值等。对于给定场景, 我们复制一个相同的点集  $S_1$  作为评估点集。首先将点集  $S_1$  归一化至  $(-1, 1)$  之间的值, 得到一个新的点集, 这里定义为点集  $S_2$ ; 然后从  $S_2$  中任选一点作为种子点  $a$ , 在  $S_2$  中寻找其  $k$  个最近邻点,  $k$  值的设置根据样本点的个数来确定(例如: 如果每个样本包含 4096 个点, 则  $k$  的值配置为 4096), 并从点集中将当前种子点  $a$  和其  $k$  个邻近点复制一份, 种子点及其  $k$  个邻近点保存为一个样本, 并从评估点集中删除; 接着根据最远点采样原则, 在点集  $S_1$  中计算每个点与种子点  $a$  的距离, 选择最远点  $a_1$  作为下一个种子点; 然后计算点集中其它点与已选种子点  $a, a_1$  的距离, 其中最远距离的点  $a_2$  作为下一个样本的种子点; 最后整合输出样本, 检查是否覆盖整个场景, 若完全覆盖, 则整合所有输出样本, 若没有完全覆盖, 则继续重复上述种子点选取操作。这样可以从给定场景中获得大量固定点数的样本, 这些样本可以直接输入标准深度神经网络中。同时, 我们可以确保场景中的每个点都包含在一些样本中, 这意味着场景的完全覆盖。我们还注意到一些样本不可避免地重叠。对于重叠部分内的点, 我们选择预测最多的标签作为其最终预测标签。

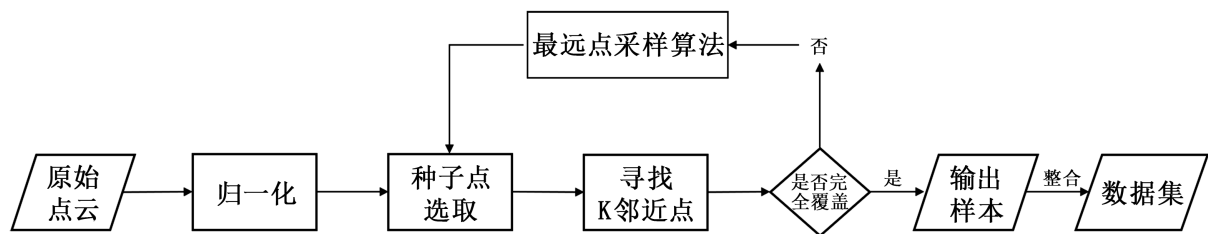


Figure 3. Flow chart of sample generation method based on FPS-KNN

图 3. 基于 FPS-KNN 样本生成方法的流程图

### 2.3.2. 通道注意力机制(SE-Block)

SE-Block 结构的核心主要是利用压缩和激励操作对神经网络中卷积特征通道之间的关系进行建模, 并调整通道的权重, 从而提高网络产生的特征表示的质量。图 4 为 SE-Block 的示意图, 其中  $\mathbf{X}$  为输入数据,  $\tilde{\mathbf{X}}$  为输出数据,  $W' \times H' \times C'$  为输入数据的尺寸( $W'$  表示为宽度,  $H'$  表示为高度,  $C'$  表示为通道数),  $W \times H \times C$  为输出数据的尺寸,  $F_{tr}$  为传统的卷积结构,  $F_{sq}(\mathbf{u}_c)$  表示压缩过程,  $F_{ex}(\mathbf{z}, \mathbf{W})$  表示激励过程。首先进行一个简单的卷积操作  $F_{tr}$ , 将输入数据  $\mathbf{X}$  映射至特征图  $\mathbf{U}$  中, 然后对特征图  $\mathbf{U}$  进行压缩操作  $F_{sq}(\mathbf{u}_c)$ , 其目的是进行空间维度的压缩, 通过在特征图上执行全局平均池化操作来获取当前特征图的全局压缩特征向量, 将每个维度的通道信息变为一个具有全局感受野的实数, 最终得到的输出维度与输入特征通道数相同; 其次是激励操作  $F_{ex}(\mathbf{z}, \mathbf{W})$ , 该操作类似于循环网络中的门机制, 通过两个全连接层和激活函数为压缩过程中输出  $1 \times 1 \times C$  分配每个通道的权重; 最后通过通道级乘法操作  $F_{scale}$  对各通道按权重估值进行调整, 调整后的通道  $\tilde{\mathbf{X}}$  进入下一层网络。

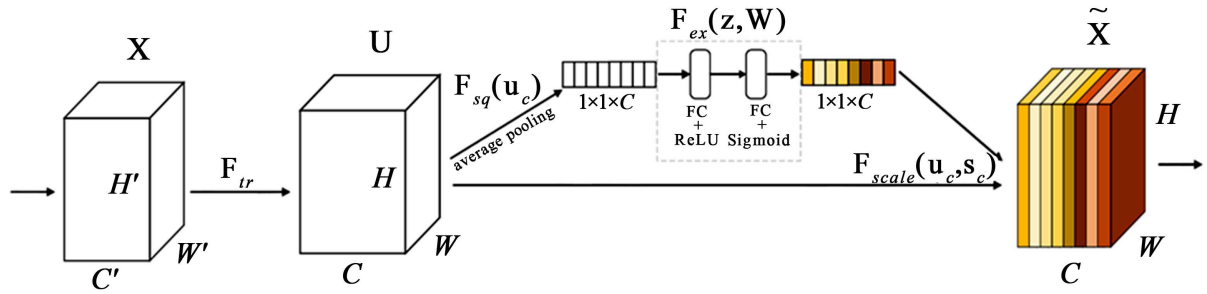


Figure 4. Schematic diagram of SE-Block model  
图 4. SE-Block 模型示意图

### 2.3.3. 修正后的焦点损失函数

多光谱 LiDAR 点云数据中的类别分布不均匀，数据量较小的类别训练效率较低，这种情况下很难提供有用的学习信号，反而会淹没训练过程，导致模型退化。为了避免过采样和欠采样，常用的方法是在小类别数据中增加采样率，从而提高性能。另一种方法是重置权重，使算法密切关注小类别。焦点损失函数(focal loss function)旨在解决类别不均衡问题。它是一个动态缩放的交叉熵损失，通过一个动态缩放因子，可以动态降低训练过程中易区分样本的权重，从而将重心快速聚焦在那些难区分的样本，目前已广泛应用于图像特征识别，很少应用于三维点云处理。焦点损失函数由交叉熵函数修正，写成：

$$FL(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i) \quad (1)$$

其中， $-\log(p_i)$  是交叉熵函数， $\alpha_i$  是平衡正负值重要性的类别权重系数。虽然  $\alpha$  可以平衡正负值的重要性，但它不能区分简单是容易还是困难，因此，我们这里选择使用调节因子  $(1 - p_i)^\gamma$ 。其中， $\gamma$  为聚焦参数， $p_i$  为模块的估计概率。该策略旨在鼓励模型关注少数类别中的点。

焦点损失函数中， $\alpha_i$  参数设置如下：

$$\alpha_i = \begin{cases} \alpha & y = 1 \\ 1 - \alpha & \text{otherwise} \end{cases} \quad (2)$$

焦点损失函数中的类别权重系数  $\alpha_i$  为固定值，可以作为超参数处理，也可以用逆向类别频率[35]来设置。然而，多光谱 LiDAR 点云数据中的类别分布很不均匀，对权重因子进行训练并采用逆向类别频率方法的模块受权重大的少数类别的影响较大。同时，超参数方法在很多情况下无法识别。

因此，我们在此基础上进行了修改，该方法根据每个类别中的数量进行计算以获得  $\alpha_i$ 。每个类别的权重系数是通过对其最大值的反百分比的指数函数来计算的，以增加少数类别的值。tanh 函数用于对权重进行归一化：

$$\alpha_i = \frac{e^{c_i} - e^{-c_i}}{e^{c_i} + e^{-c_i}} \quad (3)$$

其中，

$$c_t = \left( \frac{\max(P)}{P_t} \right)^{\frac{1}{3}}, \quad t = 1, 2, 3, \dots, N \quad (4)$$

其中， $N$  表示类别总数， $P_1, P_2, \dots, P_t$  表示每个类别的百分比， $\alpha_i$  是归一化后的权重， $c_t$  表示第  $t$  个类别的权重。修改后的焦点损失函数中的表达式如下所示：

$$FL(p_i) = -\frac{e^{c_i} - e^{-c_i}}{e^{c_i} + e^{-c_i}} (1 - p_i)^\gamma \log(p_i) \quad (5)$$

### 2.3.4. 改进的 PointNet++

本研究的模型基于 PointNet++ 架构[29], 添加了通道注意力机制模块 SE-Block, 并以端到端的方式开发 PointNet++ (见 图 5) 以进行多光谱 LiDAR 点云分类。包含坐标  $(x, y, z)$  以及三个波段  $(LRI_{\lambda_{1550}}, LRI_{\lambda_{1064}}, LRI_{\lambda_{532}})$  的  $N$  个多光谱 LiDAR 点被直接输入至我们的改进的 PointNet++ 网络架构中, 该架构主要由集成抽象模型(set abstraction module)和特征传播模块(feature propagation module)组成。图 5 内容具体展示了其中一个集成抽象模块。具体而言, 集成抽象模块以  $N \times C$  的矩阵作为输入并输出具有  $N_1$  个子采样点的  $N_1 \times C_1$  矩阵, 其中包含  $C_1$  个维度的特征向量, 这些向量包含了局部的上下文信息。如图 5 所示, 集合抽象模块由一个采样层(sampling layer)、一个分组层(grouping layer)和一个通道特征注意力层(channel feature attention layer)组成。首先, 采样层通过迭代 FPS 算法选择  $N_1$  个点作为局部区域的质心。具体而言, 将输入的点数定义为  $\{x_1, x_2, \dots, x_n\}$ , 通过 FPS 选择一个点的子集  $\{x_{i_1}, x_{i_2}, \dots, x_{i_m}\}$ , 使得  $x_{i_j}$  是相对于其余点而言, 距集合  $\{x_{i_1}, x_{i_2}, \dots, x_{i_{j-1}}\}$  最远的点(以公制距离为单位)。随后, 分组层中通过球查询算法(ball query algorithm)搜索  $N_1$  个质心周围的“相邻”点来构造相应的局部区域。对于每个质心, 通过球查询算法, 可以在给定半径的球体范围内找到所有的相邻点, 如果相邻点的数量大于  $K$ , 则从该点集中随机选择  $K$  个点以构建局部区域。实施采样层和分组层之后, 多光谱 LiDAR 点云数据已下采样至  $N_1$  个点集中, 每个点集包含  $K$  个点及其  $C$  个属性。输出的一组点集大小为  $N_1 \times C \times K$ 。随后, 我们通过通道特征注意力模块将这些局部区域编码为具有不同重要性的特征向量。对于每一个点, 我们通过多层感知器(Multi-Layer Perceptrons, MLPs)提取其特征, 随后通过 SE-Block 强调其重要通道而抑制不重要的通道。具体而言, 在 SE-Block 过程中(图 2 中红色部分), 首先使用最大池化层来压缩  $N_1$  点云的每个通道, 然后使用两个 MLP 层和 sigmoid 函数来计算各通道的权重值, 范围为 0 到 1, 权值越高, 通道越重要。最后, 激发具有较高权重值的重要通道。为避免权重值接近 0 时缺少特征的情况, 使用短链接对通道特征注意力层的起始数据和最终数据进行特征连接, 但是, 由于连接的特征尺寸之间存在差异, 因此执行卷积操作以匹配短链接连接过程中的尺寸。随后该模型又重复了一遍上述操作以进行多层特征提取, 最终获得的数据大小为  $N_2 \times C_2$ 。为了将学习到的特征从采样点传播到原始点, 我们通过特征传播模块内的反距离

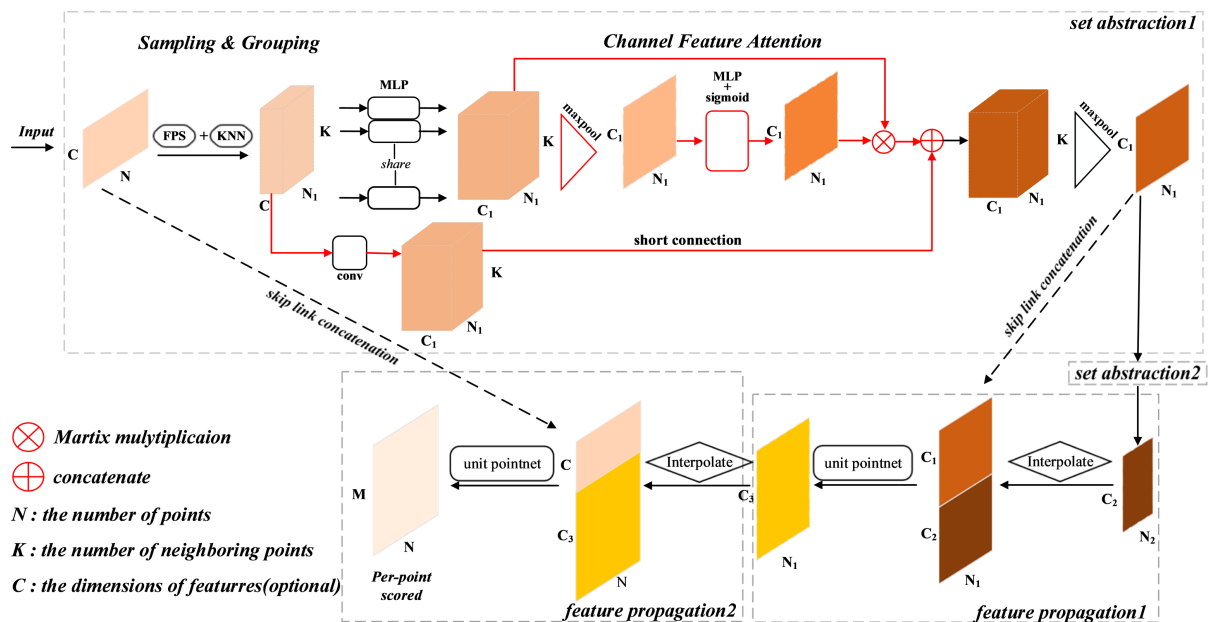


Figure 5. Flow chart of the improved PointNet++ network architecture

图 5. 改进的 PointNet++ 网络架构的流程图



加权插值法进行插值，然后使用跳跃链接将插值后的特征和上述采样层前的点特征进行级联。接下来将级联特征输入至类似于 MLP 的 Unit PointNet 模块中，并应用一些共享的完全连接层和 ReLU 层来更新每个点的特征向量。重复上述过程，直到将特征传播到原始点集为止，以获得每个点的最终语义标签。

### 3. 实验与分析

#### 3.1. 参数设置

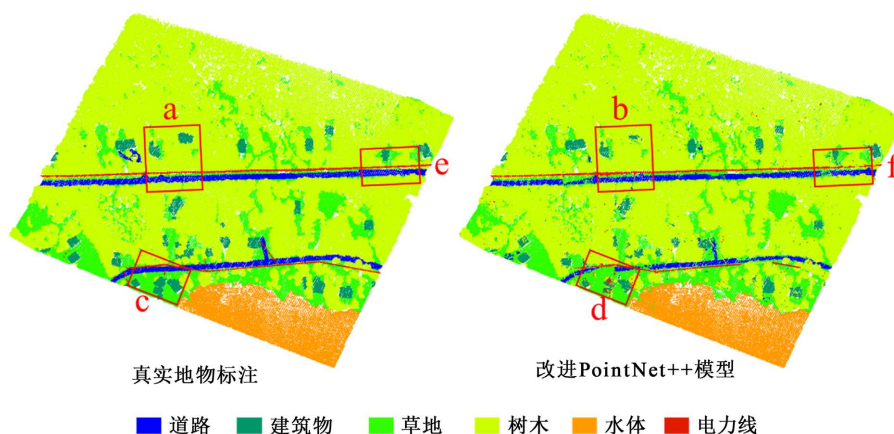
本研究中涉及的采样点个数( $N$ )、Batch size、learning rate、decay rate、max epoch 以及 optimizer 等参数分别设置为 4096、8、0.001、 $10^{-4}$ 、200 以及 Adam。

托伯莫里港口数据集中一共划分了 14 个有效区域(area\_1-area\_14)，其中 area\_1- area\_10 作为训练样本，area\_11- area\_14 作为测试和验证样本。

#### 3.2. 整体性能

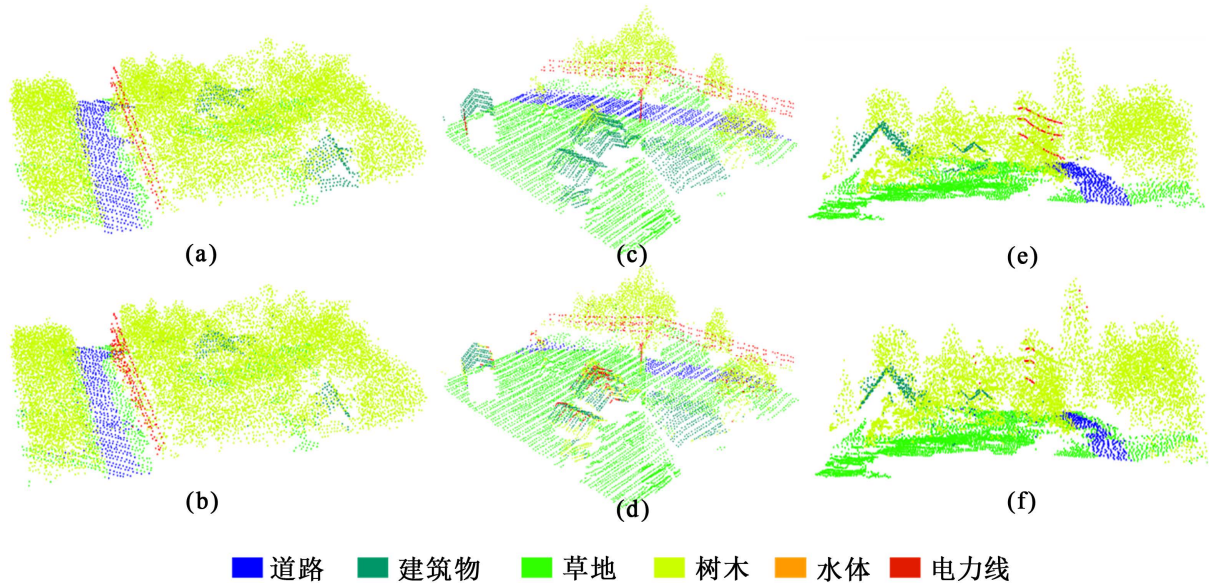
我们选取托伯莫里港口数据集中的 area\_14 作为测试区域并进行分类结果可视化展示。图 6 显示了改进的 PointNet++模型在托伯莫里港口数据集 area\_14 上的分类测试结果与真实标注地物的对比图，在两幅图中分别框出了分类结果差异较大的三个区域并将细节提取在图 7 中。如图 6 分类结果所示，与真实地物标注数据相比，大多数类别(草地、水体、树木和建筑)都被正确分类。虽然水体类和草地类发生了部分混淆，但可视化结果较为可观。从图 7 中可以看出，(d)和(f)这两个区域中，部分的道路被错误地分类成草地类，导致这一现象的原因可能是训练样本预处理中的分块过程破坏了道路的整体结构和连续性，网络难以有效地学习到地物完整的几何结构。此外，从(b)区域不难发现，一些树木点被错误分类为电力线点，以及一些包围在树木点中的建筑物被分类为树木，这可能是由于两个类别之间的混合和缺乏明显的界线造成的。

为了进一步评估本文网络架构的分类能力，我们计算了分类混淆矩阵，如表 2 所示，本文的网络架构在道路、建筑物、草地、水体、树木、电力线这六个类别中均获得了高于 70% 的 OA。特别是树木和水体的分类精度更高，分别为 97.55% 和 96.10%。尽管研究区域包含很少的电力线点，但本文使用修正后的焦点损失函数，动态降低训练过程中易区分样本的权重，将重心快速聚焦在那些难区分的样本，从而更有效地识别了数据中的电力线点。



**Figure 6.** Comparison chart of land cover classification test results of the improved PointNet++ network architecture on area\_14 of the Tobermory Port dataset: Left-hand picture is the real marked ground object in the test area; Right-hand picture is the test visualization results of the improved PointNet++

**图 6.** 改进的 PointNet++模型在托伯莫里港口数据集 area\_14 区域上的地物分类测试结果对比图：左图为测试区域的真实标注地物，右图为改进的 PointNet++的测试可视化结果



**Figure 7.** (a), (c) and (e) are the actual labeled feature details of area\_14; (b), (d) and (f) are the detailed classification results of the improved PointNet++ model

**图 7.** (a)、(c)和(e)为 area\_14 的真实标签地物细节图；(b)、(d)和(f)为改进的 PointNet++模型的细节分类结果图

表 2 中列出了 PointNet ++模型分类结果的混淆矩阵。从表 2 和表 3 的结果对比发现,改进的 PointNet++模型在道路点(占所有点的 2.06%)的分类结果上提升了 20%, 在建筑物点(占所有点的 2.10%)的分类结果上提升了 20%, 在电力线点(占所有点的 0.42%)的分类结果上提升了 14%。

**Table 2.** Confusion matrix of the improved PointNet++ model classification results

**表 2.** 改进的 PointNet ++模型分类结果的混淆矩阵

类别	道路	建筑物	草地	树	水	电力线
道路	<b>28,090</b>	1376	9753	146	0	0
建筑物	728	<b>32,158</b>	1013	6121	0	140
草地	7478	2294	<b>233,853</b>	9239	6201	3
树木	113	14,113	8997	<b>1,090,550</b>	52	4064
水	44	0	17,254	20	<b>426,906</b>	0
电力线	0	143	24	2153	0	<b>5701</b>
OA (%)	71.36	80.07	90.27	97.55	96.10	71.08

**Table 3.** Confusion matrix of the PointNet++ model classification results

**表 3.** PointNet ++模型分类结果的混淆矩阵

类别	道路	建筑物	草地	树	水	电力线
道路	<b>20,057</b>	38	19,082	174	14	0
建筑物	1654	<b>24,136</b>	5981	8284	0	105
草地	7570	805	<b>230,358</b>	10,198	10,137	0

续表

树木	3018	13,030	9442	<b>1,089,742</b>	49	2608
水	188	0	14905	78	<b>429,053</b>	0
电力线	2	110	33	3297	0	<b>4579</b>
OA (%)	50.95	60.10	88.92	97.48	96.58	57.09

### 3.3. 比较实验

为了公平地评估本文改进的 PointNet++ 网络架构的鲁棒性和有效性, 本文将其与其他具有代表性的深度神经网络模型进行了比较, 包括 PointNet [28]、PointNet++ [29]、DGCNN [30]、GACNet [31] 和 RSCNN [32]。尽管这些模型在点云分类和语义分割方面取得了优异的性能, 但它们尚未在包含复杂城市地理特征的大规模多光谱 LiDAR 点云上进行测试。因此, 本文复现了上述五个网络模型, 并在托伯莫里港口数据集上进行了分类实验。实验过程中的基础参数设置一致, 本文改进的 PointNet++ 模型与其他五个网络模型的定量比较结果如表 4 所示。

**Table 4.** Comparison of land-cover classification results of the improved PointNet++ model with other five models (PointNet, DGCNN, RSCNN, GACNet and Our Model) in the Tobermory Port dataset

**表 4.** 改进的 PointNet ++ 模型与其他五个模型(PointNet、DGCNN、RSCNN、GACNet、Our Model)在 Tobermory 港口数据集的地物分类结果比较

模型	OA (%)	mIoU (%)	F <sub>1</sub> -score (%)	Kappa
PointNet	85.54	46.72	58.99	0.780
PointNet++	94.19	61.08	71.53	0.901
DGCNN	95.19	60.85	72.35	<b>0.918</b>
RSCNN	94.48	61.24	72.82	0.905
GACNet	92.7	58.06	69.95	0.899
Our Model	<b>95.21</b>	<b>62.59</b>	<b>73.58</b>	<b>0.918</b>

如表 4 所示, 我们改进的 PointNet++ 模型达到了 OA (95.21%)、mIoU (62.59%) 和 F<sub>1</sub>-score (73.58%) 的最佳精度, 而 DGCNN 在 Kappa 系数上取得了较好的结果。与 PointNet ++ 相比, 我们模型的性能有了显著提高, OA 提高了 1% 以上, F<sub>1</sub>-score 提升了 2%, 尤其是 mIoU 值提高至 62.59%。

## 4. 结论

多光谱 LiDAR 点云数据提供了几何信息和多波段信息, 有助于识别不同的地物。在本研究中, 探讨了深度学习分类方法在多光谱 LiDAR 点云数据中的应用, 在托伯莫里港口数据集上通过结合最远点采样和  $k$  最近邻点(FPS-KNN)的样本生成方法获得了可训练样本; 接着将通道注意力机制模块 Squeeze and Excitation Block (SE-Block) 嵌入至 PointNet++ 模型中, 从而强调重要通道并抑制不利于预测的无用通道, 提高了特征的显著性; 最后, 使用改进的焦点损失函数来处理多光谱 LiDAR 点云数据上不均匀类别分布的问题。实验表明, 本文提出的网络架构获得的总体准确度(OA)、mIoU、F<sub>1</sub>-score 和 Kappa 系数分别为 95.21%、62.59%、73.58% 和 0.918。此外, 与目前已有的五个深度学习模型比较实验证实, 我们提出的方法在多光谱 LiDAR 数据的地物分类中表现出更鲁棒、更准确的性能。

## 参考文献

- [1] Wen, C., Yang, L., Li, X., Peng, L. and Chi, T. (2020) Directionally Constrained Fully Convolutional Neural Network for Airborne LiDAR Point Cloud Classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, **162**, 50-62. <https://doi.org/10.1016/j.isprsjprs.2020.02.004>
- [2] Li, W., Wang, F. and Xia, G. (2022) A Geometry-Attentional Network for ALS Point Cloud Classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, **164**, 26-40. <https://doi.org/10.1016/j.isprsjprs.2020.03.016>
- [3] Yan, W.Y., Shaker, A. and El-Ashmawy, N. (2015) Urban Land Cover Classification Using Airborne LiDAR Data: A Review. *Remote Sensing of Environment*, **158**, 295-310. <https://doi.org/10.1016/j.rse.2014.11.001>
- [4] 杨必胜, 梁福逊, 黄荣刚. 三维激光扫描点云数据处理研究进展、挑战与趋势[J]. 测绘学报, 2017, 46(10): 1509-1516.
- [5] Blackman, R. and Yuan, F. (200) Detecting Long-Term Urban Forest Cover Change and Impacts of Natural Disasters Using High-Resolution Aerial Images and LiDAR Data. *Remote Sensing*, **12**, Article 1820. <https://doi.org/10.3390/rs12111820>
- [6] 吴杭彬. 融合航空影像的机载激光扫描数据分类与特征提取[J]. 测绘学报, 2011, 40(1): 134.
- [7] 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): 340-347.
- [8] 释小松, 程英蕾, 薛豆豆, 秦先祥. 基于 Point-Net 的多源融合点云地物分类方法[J]. 激光与光电子学进展, 2020, 57(8): 170-178.
- [9] Zhou, K., Ming, D., Lv, X., Fang, J. and Wang, M. (2019) CNN-Based Land Cover Classification Combining Stratified Segmentation and Fusion of Point Cloud and Very High-Spatial Resolution Remote Sensing Image Data. *Remote Sensing*, **11**, Article 2065. <https://doi.org/10.3390/rs11172065>
- [10] 康健, 管海燕, 于永涛, 等. 基于 RFA-LinkNet 模型的高分遥感影像水体提取[J]. 南京信息工程大学学报, 2023, 15(2): 160-168.
- [11] Teo, T.A. and Wu, H.M. (2017) Analysis of Land Cover Classification Using Multi-Wavelength LiDAR System. *Applied Sciences*, **7**, Article 663. <https://doi.org/10.3390/app7070663>
- [12] Fernandez-Diaz, J.C., Carter, W.E., Glennie, C., Shrestha, R.L., Pan, Z., Ekhtari, N., Singhania, A., Hauser, D. and Sartori, M. (2016) Capability Assessment and Performance Metrics for the Titan Multispectral Mapping Lidar. *Remote Sensing*, **8**, Article 936. <https://doi.org/10.3390/rs8110936>
- [13] Morsy, S., Shaker, A., El-Rabbany, A. and LaRocque, P.E. (2016) Airborne Multispectral LiDAR Data for Land-Cover Classification and Land/Water Mapping Using Different Spectral Indexes. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **3**, 217-224. <https://doi.org/10.5194/isprsannals-III-3-217-2016>
- [14] Yu, Y., Guan, H., Li, D., Gu, T., Wang, L., Ma, L. and Li, J. (2019) A Hybrid Capsule Network for Land Cover Classification Using Multispectral LiDAR Data. *IEEE Geoscience and Remote Sensing Letters*, **17**, 1263-1267. <https://doi.org/10.1109/LGRS.2019.2940505>
- [15] Huo, L.Z., Silva, C.A., Klauberg, C., Mohan, M., Zhao, L.J., Tang, P. and Hudak, A.T. (2018) Supervised Spatial Classification of Multispectral LiDAR Data in Urban Areas. *PLOS ONE*, **13**, e0206185. <https://doi.org/10.1371/journal.pone.0206185>
- [16] 潘锁艳, 管海燕. 机载多光谱 LiDAR 数据的地物分类方法[J]. 测绘学报, 2018, 47(2): 198-207.
- [17] Pan, S., Guan, H., Yu, Y., Li, J. and Peng, D. (2019) A Comparative Land-Cover Classification Feature Study of Learning Algorithms: DBM, PCA, and RF Using Multispectral LiDAR Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **12**, 1314-1326. <https://doi.org/10.1109/JSTARS.2019.2899033>
- [18] Wang, Q. and Gu, Y. (2019) A Discriminative Tensor Representation Model for Feature Extraction and Classification of Multispectral LiDAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, **58**, 1568-1586. <https://doi.org/10.1109/TGRS.2019.2947081>
- [19] Morsy, S., Shaker, A. and El-Rabbany, A. (2017) Multispectral LiDAR Data for Land Cover Classification of Urban Areas. *Sensors*, **17**, Article 958. <https://doi.org/10.3390/s17050958>
- [20] Ekhtari, N., Glennie, C. and Fernandez-Diaz, J.C. (2018) Classification of Airborne Multispectral LiDAR Point Clouds for Land Cover Mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **11**, 2068-2078. <https://doi.org/10.1109/JSTARS.2018.2835483>
- [21] Wichmann, V., Bremer, M., Lindenberger, J., Rutzinger, M., Georges, C. and Petrini-Monteferrri, F. (2015) Evaluating the Potential of Multispectral Airborne LiDAR for Topographic Mapping and Land Cover Classification. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **II-3/W5**, 113-119. <https://doi.org/10.5194/isprsannals-II-3-W5-113-2015>

- [22] 赵沛冉, 管海燕, 李迪龙, 等. 利用样本生成方法进行机载多光谱 LiDAR 数据深度学习分类[J]. 测绘通报, 2021(12): 16-21.
- [23] 袁鹏飞, 黄荣刚, 胡平波, 杨必胜. 基于多光谱 LiDAR 数据的道路中心线提取[J]. 地球信息科学学报, 2018, 20(4): 452-461.
- [24] 景庄伟, 管海燕, 臧玉府, 等. 基于深度学习的点云语义分割研究综述[J]. 计算机科学与探索, 2021, 15(1): 1-26.
- [25] Li, D.L., *et al.* (2020) Building Extraction from Airborne Multi-Spectral LiDAR Point Clouds Based on Graph Geometric Moments Convolutional Neural Networks. *Remote Sensing*, **12**, Article 3186.
- [26] Briechle, S., Krzystek, P. and Vosselman, G. (2020) Classification of Tree Species and Standing Dead Trees by Fusing UAV-Based Lidar Data and Multispectral Imagery in the 3D Deep Neural Network PointNet++. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **V-2-2020**, 203-210.
- [27] Shi, S., *et al.* (2021) Land Cover Classification with Multispectral LiDAR Based on Multi-Scale Spatial and Spectral Feature Selection. *Remote Sensing*, **13**, Article 4118.
- [28] Qi, C.R., Su, H., Mo, K. and Guibas, L.J. (2017) PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*, Honolulu, 21-26 July 2017, 77-85.
- [29] Qi, C.R., Yi, L., Su, H. and Guibas, L.J. (2017) PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *Proceedings of the Neural Information Processing Systems (NIPS 2017)*, Long Beach, 4-7 December 2017, 5105-5114.
- [30] Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M. and Solomon, J.M. (2018) Dynamic Graph CNN for Learning on Point Clouds. *ACM Transactions on Graphics*, **38**, 1-12. <https://doi.org/10.1145/3326362>
- [31] Wang, L., Huang, Y., Hou, Y., Zhang, S. and Shan, J. (2019) Graph Attention Convolution for Point Cloud Semantic Segmentation. *Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2019)*, Long Beach, 15-20 June 2019, 10288-10297. <https://doi.org/10.1109/CVPR.2019.01054>
- [32] Liu, Y., Fan, B., Xiang, S. and Pan, C. (2019) Relation-Shape Convolutional Neural Network for Point Cloud Analysis. *Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2019)*, Long Beach, 15-20 June 2019, 8887-8896. <https://doi.org/10.1109/CVPR.2019.00910>
- [33] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. *Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, Salt Lake City, 19-21 June 2018, 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>
- [34] Hu, Q., Yang, B., Xie, L., *et al.* (2020) RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 13-19 June 2020, 11105-11114. <https://doi.org/10.1109/CVPR42600.2020.01112>
- [35] Lin, T.Y., Goyal, P., Girshick, R., He, K. and Dollár, P. (2020) Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 318-327. <https://doi.org/10.1109/TPAMI.2018.2858826>