

基于 Feature-RNet 的三维大场景点云分类框架

雷根华¹, 王 蕾^{1,2}, 张志勇¹

(1. 东华理工大学 信息工程学院, 江西 南昌 330013;

2. 江西省核地学数据科学与系统工程技术研究中心, 江西 南昌 330013)

摘 要:随着大场景三维点云应用在越来越多的领域中,近些年对激光点云大场景下的分类研究不断深入,各种分类模型层出不穷,在大场景点云分类任务中表现优异,但是依然存在训练时间长、计算复杂以及分类精度低等问题。针对分类精度低这一问题,提出一种基于 Feature-RNet 的三维大场景点云分类框架模型来实现点云分类工作。该框架模型不直接以三维点云数据作为输入,而是以通过 KNN 方法提取三维点云的二维特征和三维特征构建的特征图像作为输入,避免了网络框架对三维点云数据直接处理的不适应性;该模型设计的 RNet 框架结构利用了残差模块,并对其进行变型以达到提高分类精度的效果。采用公开的 Oakland 三维数据集对 Feature-RNet 框架模型进行训练,与现有的其他深度学习分类框架相比,提出的 Feature-RNet 框架模型在分类精度上有较大的提升,总体分类准确率能达到 97.7%。

关键词:点云特征图像;RNet 网络框架;大场景点云分类;Oakland 数据集;深度学习

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2022)06-0085-07

doi:10.3969/j.issn.1673-629X.2022.06.015

Cloud Classification Framework of 3D Large-scale Scene Based on Feature-RNet

LEI Gen-hua¹, WANG Lei^{1,2}, ZHANG Zhi-yong¹

(1. School of Information Engineering, East China University of Technology, Nanchang 330013, China;

2. Jiangxi Engineering Technology Research Center of Nuclear Geoscience Data Science and
System, Nanchang 330013, China)

Abstract: With the application of large scene 3D point cloud in more and more fields, the classification research of laser point cloud in large scene has been deepened in recent years, and various classification models emerge in endlessly and perform well in the large-scale scenic spot cloud classification task, but there are still problems such as long training time, complex calculation and low classification accuracy. Aiming at such problem of low classification accuracy, a three-dimensional scenic spot cloud classification framework model based on Feature-RNet is proposed to realize point cloud classification. The framework model does not directly take the three-dimensional point cloud data as the input, but takes the two-dimensional features of the three-dimensional point cloud extracted by KNN method and the feature image constructed by the three-dimensional features as the input, which avoids the inadaptability of the network framework to the direct processing of the three-dimensional point cloud data. The framework structure of RNet designed by the model uses the residual module and modifies it to improve the classification accuracy. The open Oakland 3D data set is used to train the Feature-RNet framework model. Compared with other existing deep learning classification frameworks, the proposed Feature-RNet framework model has a great improvement in classification accuracy, and the overall classification accuracy can reach 97.7%.

Key words: point cloud feature image; Rnet network framework; cloud classification of large-scale scene; Oakland dataset; deep learning

0 引言

点云是海量点的集合,在同一空间参考系下表达目标空间分布和目标表面特性,具有不规则且无序等特点。三维点云数据的获取方式主要是通过三维激光

扫描仪进行数据采集得到。激光雷达测量(light detection and ranging, LiDAR),是一种光学遥感技术,由激光发射部分、接收部分、信号处理部分三部分组成,能够获取坐标(X, Y, Z)、强度、颜色等物体信息,

收稿日期:2021-08-16

修回日期:2021-12-16

基金项目:国家自然科学基金资助项目(61561003, 61761003);江西省核地学数据科学与系统工程技术研究中心基金(JETRCNGDSS201902)

作者简介:雷根华(1996-),女(畚),硕士,研究方向为大场景点云分类;王 蕾,博士,副教授,硕导,研究方向为视觉信息获取与处理、计算机视觉,计算机图形图像技术等。

通过 LiDAR 采集到的点云数据格式多种多样。LiDAR 可以单独工作获取小场景的三维点云数据,而对于获取复杂大场景下或者大尺寸物体的点云数据,单独的 LiDAR 往往不能满足要求,需与 GNSS、camera 等其他仪器联合使用,从而获取更丰富、完善的物体信息。

点云分类一般包含从局部和全局两个不同的方面提取点特征和通过学习到的特征将每个点划分为预定义的语义类别两个步骤^[1]。早期的点云分类方法研究主要通过手工设计^[2-3]的规则来提取点特征,然后使用基于机器学习的分类器预测每个点的语义标签^[1]。如高斯混合模型^[4-5]、支持向量机(SVM)^[6-7]、AdaBoost^[8-9]等。由于手工设计规则提取的特征表达能力较弱,使得模型在处理复杂场景时往往无法达到令人满意的效果,与此同时,有关分类器的选择和组合十分复杂,严重依赖研究人员的专业,极大地降低了此类方法应用的灵活性。

在点云分类的深度学习方法中,2006 年,卷积神经网络(CNN)的提出引起了广泛关注。2015 年,Su 等人提出了多视图卷积神经网络(MVCNN)^[10],通过设置不同位置、不同角度的视点获取多个视角的二维视图,以此作为 MVCNN 的输入。MVCNN 在分类任务中取得了不错的成绩,但是由于无法动态选择视图等因素,导致分类精度受到影响。Maturana 等人提出的 VoxNet^[11]是一种最早将非结构化的点云数据体素化为网格数据并应用于 3D 卷积神经网络的模型,在一定程度上解决了点云非结构化的问题,但是计算所需内存过大。Barnea 和 Filin^[12]等提出将输入的点云数据构造成深度图的模式,利用深度图的特点使用均值平移算法对点云数据进行分类分割,最后将划分出的特征子集进行联合以达到良好的分类效果。Barnea^[13]改进了 Filin^[12]的分类算法,通过使用迭代的思想对其网络框架进行改进,能更好地得到分割边界和分割区域,但不能准确得到分割目标对象。

基于深度学习的点云分类方法大多都是基于二维图像设计的。pointNet^[14]及其变体的提出给研究人员提供了新的思路。文献[15]提出了一种基于多维特

征矩阵和 PointNet (Feature-pointNet 框架)的深度神经网络模型,以特征图像代替三维点云数据作为 PointNet 的输入,由于上述改变,文献[15]在后面的 pointNet 框架中并没有使用 T-Net 矩阵。该方法取得了不错的分类效果,但是在参数和运行时间效率上还有待提高。

基于文献[15]中的 Feature-pointNet 网络框架,该文设计出一种基于 Feature-RNet 的网络框架。在该方法中,通过借鉴点云特征图像的获取方法(特征图像的生成方法与文献[15]中的方法一致)作为 RNet 框架的输入;其次,设计 RNet 网络框架模型,以达到提升分类精度的目的;最后使用公开的 Oakland 3D 数据集进行 RNet 的分类模型训练。

1 研究基础

现今许多领域的发展都离不开三维大场景点云分类,例如自动驾驶等,为此许多研究人员对点云做了大量的研究工作。R. B. Rusu^[16]提出 FPFH 的方法,通过寻找点云内部或者不同点云之间的关联关系进而采取有效措施来加快 FPFH 的计算速度,但是其中点云法线的选取好坏对最终结果影响较大。孙杰^[17]提出一种能够有效运行在大场景点云数据集上的面向对象的方法,该方法能够获取 LIDAR 点云数据的多种特征,最终达到一个准确的分类效果。Riegler^[18]提出 OCNet,由于输入数据中的稀疏性的这一特性从而构建不平衡八叉树模型来划分网络空间,是一种高分辨率的 3D 卷积神经网络。Hu^[19]提出了一个点到图的框架,通过提取窗口内的相邻点并将其转换为图像,将点的分类问题转换成图像的分类问题,该方法在 DTM 生成任务(二值分类)中取得了不错的成绩,但是该方法在其他类型的点云分类任务中仍然存在一些问题。文献[15]改进 Hu 的方法,提出了一种基于点的特征图像生成方法,通过提取每个点的二维和三维特征从而构建特征图像,再训练 pointNet 网络框架模型并且使用该模型得到最终分类结果。

文献[15]中采用的框架流程如图 1 所示。

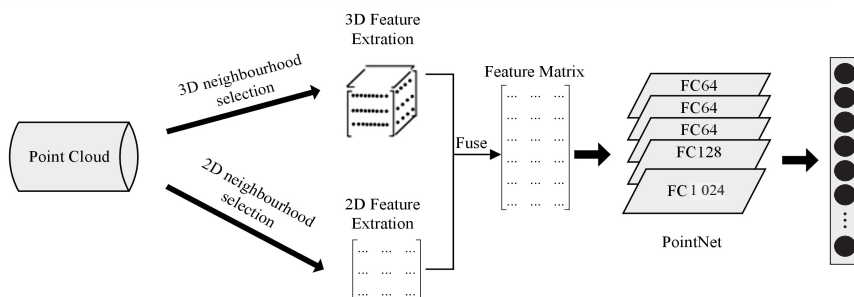


图 1 Feature-pointNet 网络框架流程

首先将三维点云数据集分别在二维和三维两个层面上进行特征提取。在三维层面上,采用 KNN 方法在每个点周围选取 100 个点计算 17 个特征值,得到的三维特征矩阵为 $[L_\lambda, N_x, N_y, N_z, P_\lambda, S_\lambda, M_x, M_y, M_z, O_\lambda, A_\lambda, E_\lambda, T_\lambda, C_\lambda, D, Q, V]$, 其中的 $[N_x, N_y, N_z]^T$ 、 $[M_x, M_y, M_z]^T$ 分别代表点的最大分布方向、点的最小分布方向。在二维层面上,即三维点云在 x, y, z 三个不同方向上的二维平面投影,同三维平面上一致,通过 KNN 的方法在每个点周围选取 100 个点作为一个计算单位,获取 5 个特征值(三个不同平面共计 15 个特征值),得到的二维特征矩阵为 $[r_k, D_{2D}, R_{\lambda, 2D}, S_{2D}, EVratio_{2D}]$, 从左到右分别代表最优邻域半径、局部点密度、特征值比率、特征值和、高度方差。将得到的二维特征和三维特征进行融合构建点云的特征矩阵,作为 pointNet 框架的输入。需要注意的是,不同于传统的 pointNet,文献[15]采用的是 pointNet 中的 vanilla 版本,由两个网络层 $MLP(64, 64)$ 、 $MLP(64, 128, 1024)$ 以及一个全连接网络层 $MLP(512, 256, C)$ 构成。与传统的 pointNet 框架相比, vanilla 舍弃了旨在保持点云不变性的 T-Net 矩阵。

在文献[15]的 Feature-pointNet 网络框架模型中,提出的点云特征提取方法能够有效地提取点云数据特征,并融合为点云特征图像作为框架输入,为以后的研究提供了新的方式且分类效率更高。正是由于对点云特征转换为点云特征图像的这一转变,使得应用在图像分类的网络框架能够适用于大场景三维点云分类任务中。该文正是借鉴其方法,将处理得到的点云特征图像作为框架输入进行分类训练。另外,在分析比较文献[15]及其他典型的分类框架后,设计出一种 RNet 网络框架,该框架与 pointNet (vanilla 版本)相似,采用的都是 1×1 的卷积核,不同之处在于 RNet 由网络层 $MLP(64, 64)$ 、 $MLP(64, 64, 1024)$ 以及全连接网络层 $(512, 256, C)$ 组成,同时在框架中多次添加 \oplus 操作,将不同层次的经过卷积提取的三维点云特征相加形成新的特征,作为下一个层次的输入,使得特征多样化,从一定程度上避免了点云有效信息丢失,提升了分类精度。

2 网络框架

工作流程如图 2 所示。

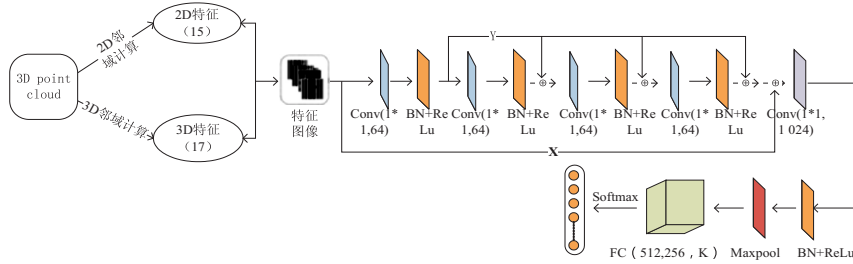


图 2 Feature_RNet 网络框架

首先,完成特征图像的生成,即对三维点云数据集进行特征提取,所用方法与文献[15]一致(在上一章节已经详细介绍了,这里就不进行赘述)。然后将特征值进行融合得到特征图像,再将特征图像作为 RNet 网络框架的输入。对于提出的 RNet 网络框架的设计,灵感来源于 pointNet 论文里的 vanilla 版本(即没有使用 T-Net 矩阵)的框架模型,主要区别在于:(1)没有

卷积核大小为 1×1 ,通道数为 128 的卷积层;(2)多次添加 \oplus 操作,实现特征相加,从而得到新的特征矩阵作为下一层次的输入,能够更有效地获取特征信息,进而提升分类精度。

2.1 RNet 网络框架

RNet 框架如图 3 所示。

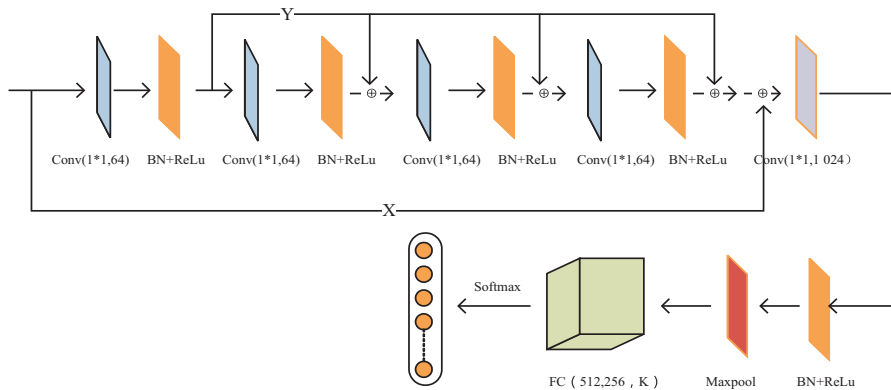


图 3 RNet 网络框架

将特征图像输入到 RNet 网络框架中,通过 RNet 框架中的网络层对输入的特征图像进行特征提取操作,最终完成点云分类任务。由于深度学习网络框架会随着卷积层数的增加而出现过度拟合的情况,为了防止这一情况的出现,RNet 网络模型中加入了 BN 和 ReLu,从而能更有效地获取点云的特征信息,同时,类似残差块的设计能够在一定程度上解决梯度爆炸以及梯度消失等问题。

RNet 网络框架包含 4 个 1×1 卷积核,通道数为 64 的 conv,1 个 1×1 卷积核,通道数为 1 024 的 conv。卷积核部分都采用了 1×1 的卷积核,在卷积神经网络中,相比于大的卷积核而言,在相同情况下,使用小的卷积核使得卷积层计算过程中参数量减少很多。由于在卷积神经网络层对点云特征图像提取特征的过程中部分有效特征信息并未提取到,从而导致最终分类结果精度下降,因此,为了改进这一问题,在 RNet 网络框架中(如图 3 所示), \oplus 实现张量相加,通过叠加输入的方法,使得网络框架将深层特征与浅层特征结合起来,能够获得更为全面的特征信息,从而提高分类精度。

接下来介绍点云特征图像进入到 RNet 框架之后的工作流程。首先通过文献[15]中的特征图像的获取方法,得到了形如 [BatchSize, Channel, Width, Height] 格式的 [BatchSize, 1, 32, 32] 的特征图像,记为 X ,其中,Channel 为通道数,Width,Height 分别为特征图像的宽度和高度。将上述得到的特征图像输入到 RNet 网络框架中,首先经过一个卷积核大小为 1×1 ,输出通道数为 64 的网络层,输出为 [BatchSize, 64, 32, 32] 格式的特征矩阵,将上述输出记为 Y 。再将 Y 作为下一个卷积核大小为 1×1 ,输出通道数为 64 网络层的输入,最终得到 [BatchSize, 64, 32, 32] 格式的特征矩阵,记为 out1。值得注意的是,在第三个卷积核大小为 1×1 ,输出通道数为 64 的网络层时,输入并非为 out1,而是 $Y + \text{out1}$,最终得到 [BatchSize, 64, 32, 32] 格式的特征矩阵,记为 out2。同样的在第四个卷积核大小为 1×1 ,输出通道数为 64 的网络层,输入并非为 out2,而是 $Y + \text{out2}$,最终得到 [BatchSize, 64, 32, 32] 格式的特征矩阵,记为 out3。接下来是一个卷积核大小为 1×1 ,输出通道数为 1 024 的网络层,以 $\text{out3} + X + Y$ 作为输入,最终得到 [BatchSize, 64, 32, 32] 格式的特征矩阵。再经过 Flatten 层之后,输出 (None, 230 400) 的向量,最终经过全连接 FC (512, 256, K) 得出分类结果。

2.2 残差块进一步的应用

深度残差网络 ResNet^[20] 由许多的残差单元组成,传统的残差单元由直接映射部分和残差部分组成,每

个残差块可以用以下公式表示:

$$y_l = h(x_l) + F(x_l, W_l) \quad (1)$$

$$x_{l+1} = f(y_l) \quad (2)$$

其中, x_{l+1} 和 x_l 分别是第 l 个残差单元的输出和输入, F 是残差函数,在文献[20]中, $h(x_l) = x_l$ 是一个恒等映射, f 是一个 ReLU^[21] 函数。

残差单元包含具有“快捷连接”的前馈神经网络。快捷连接^[22-24]是跳过一个或多个层的连接。快捷连接只是执行标识映射,它们的输出被添加到堆叠层的输出中。标识快捷方式连接既不增加额外的参数,也不增加计算复杂性。直接将 x 从浅层神经网络传送到神经网络的深层,通过这种方式能够有助于解决梯度消失和梯度爆炸问题,使得在训练更深的网络模型时,能保证良好的鲁棒性。

如图 3 所示,RNet 框架模型在残差模块结构的基础上,做出了一些变型——构造 4 个 \oplus 操作(\oplus 代表张量的叠加),与传统残差块不同的是 addition 是在 ReLu 之前,而在 RNet 模型中 addition 表现在激活函数 ReLu 之后。从提取特征的角度上来说,将浅层提取的全局特征传送到网络的更深层,与该更深层提取的全局特征进行结合,在不增加参数量的前提下,提高了最终分类结果精度。

2.3 特征图像的生成

在文献[15]的 Feature-PointNet 网络框架模型中,利用构造的特征图像作为框架的输入,以三维特征与二维特征相结合的方式表示整体点云特征,以达到提高分类准确率的效果。在此基础上,该文利用文献[15]的特征图像构造方法,设计出特征图像作为输入并对 RNet 框架进行训练。

三维特征组合为: $[L_\lambda, N_x, N_y, N_z, P_\lambda, S_\lambda, M_x, M_y, M_z, O_\lambda, A_\lambda, E_\lambda, T_\lambda, C_\lambda, D, Q, V]$

二维特征组合为: $[r_k, D_{2D}, R_{\lambda, 2D}, S_{2D}, EVratio_{2D}]$

对于三维数据点 $X_i(x_i, y_i, z_i)$,在点 X_i 周围选取 n 个点,则中心点值计算如公式 3 所示,再得到 $M = (X_1 - \bar{X}, \dots, X_n - \bar{X})$,通过 M 可以计算得出协方差矩阵 C (如公式 4 所示)

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (3)$$

$$C = \frac{1}{n} M^T M \quad (4)$$

从特征矩阵中,可以计算出 $\lambda_1, \lambda_2, \lambda_3$,通过这些特征值可以计算出 Linearity: $L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_3}$, Planarity:

$$P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1}, \text{ Omnivariance: } O_\lambda = \sqrt[3]{\lambda_1 \lambda_2 \lambda_3},$$

$$\text{Eigenentropy: } E_\lambda = - \sum_{i=1}^3 e_i \ln(e_i)。$$

对于二维数据点 $P_k(x_k, y_k)$ 作为点 P 的第 k 个近邻数据点, 则 r_k (最优邻域半径) 的计算公式如下:

$$r_k = \sqrt{(x - x_k)^2 + (y - y_k)^2} \quad (5)$$

特征值比率 $R_\lambda = \frac{\lambda_2}{\lambda_1}$, λ_1, λ_2 为二维协方差矩阵的特征值。

3 实验结果

3.1 数据集

该文采用 Oakland^[25] 数据集作为实验数据集, 该数据的获取是通过配备了侧视 SICK LMS 激光扫描仪的 Navlab11 在宾夕法尼亚州匹兹堡奥克兰的 CMU 校园周围收集的。该数据集包含 160 万个 3D 点云数据, 共 44 个标签。Oakland 数据集包含了训练数据集、测试数据集和验证数据集, 其中训练数据集和测试数据集经过了过滤和标记, 将 44 个标签重新映射成了 5 个标签, 这 5 个标签分别为 scatter_misc (植被)、utility_pole (柱状物)、load_bearing (地面)、default_wire (电线)、facade (立面), 如表 1 所示。

表 1 Oakland 3D 数据集

| 标签名称 | 训练数据集/个 | 测试数据集/个 |
|--------------|---------|---------|
| scatter_misc | 14 441 | 9 278 |
| default_wire | 2 571 | 481 |
| utility_pole | 1 086 | 368 |
| load_bearing | 14 121 | 71 863 |
| facade | 4 713 | 7 821 |
| 合计 | 36 932 | 89 811 |

GML 数据集^[26] 是通过 ALTM2050 这一系统采集获取的三维点云数据集, 该系统提供了 GML_A 数据集和 GML_B 数据集两个不同的数据集。由于 GML 数据集过大, 总共包含约 2077 万个 3D 数据点, 而实

验硬件环境远远达不到能够处理 GML 数据集的条件, 同时, GML_A 数据集相比于 GML_B 数据集分布更为合理, 能够更好地体现网络模型分类效果, 因此, 采用了 GML_A 数据集作为实验数据集。在应用到框架之前, 对 GML_A 数据集进行了预处理 (Space 采样), 目的是适配现有的硬件环境以及通过减少数据集数据大小从而减少在特征提取计算上的时间耗费。经过了预处理之后的 GML_A 数据集每个类别数据如表 2 所示。

表 2 经过预处理之后的 GML_A 数据集大小

| 类别 | 训练数据 | 测试数据 |
|----------|---------|---------|
| Ground | 118 545 | 81 269 |
| Building | 2 713 | 13 145 |
| Car | 1 619 | 0 |
| Tree | 57 255 | 69 037 |
| 合计 | 180 132 | 163 451 |

3.2 实验

实验是在 Intel i7-4790、NVIDIA RTX 2070、8G 内存, 在 Window10 和 Python3.7 下搭建 CUDA 10.0、CUDNN7.6.4、Pytorch 0.6 的环境下进行的, 学习率初始化为 0.001, 训练总轮数为 120, 使用 weight_decay = 1e-8, 第一次估计的指数衰减率为 0.9, 第二次估计的指数衰减率为 0.99 的 Adam 优化方法来训练 RNet 网络框架。通过提取点云特征计算特征图像, 而在提取特征值的过程中, 是以 100 个点为单位计算特征值, 以此为固定值, 通过修改后面的 RNet 网络框架, 与不同的神经网络框架下的准确率以及花费时间进行对比。

对比其他方法在 Oakland 3D 数据集上进行分类任务的精度, 以及为了验证图 3 中 \oplus (张量加 addition) 的作用, 对比了有 \oplus 操作的 RNet 网络框架以及去除 \oplus 操作的 RNet 框架在 Oakland 数据集上的分类精度, 如表 3 所示, 其中 OA 代表总体分类精度。

表 3 不同方法分类精度对比

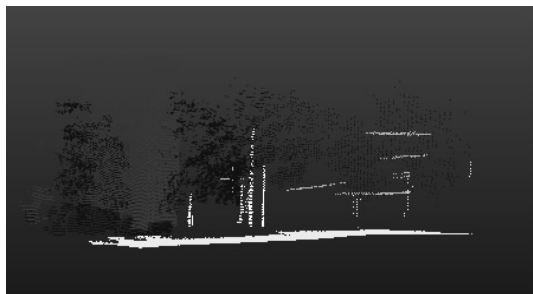
%

| 方法 \ 标签 | utility_pole | scatter_misc | default_wire | load_bearing | Façade | OA |
|---------------------|--------------|--------------|--------------|--------------|--------|-------|
| 方法[27] | 22.3 | 90.7 | 5.3 | 99.6 | 87.6 | 93.53 |
| 方法[28] | 68.42 | 80.68 | 92.93 | 98.37 | 71.13 | 94.75 |
| 方法[29] | 70.11 | 80.55 | 93.08 | 98.22 | 70.95 | 94.68 |
| 方法[30] | 59.7 | 92.0 | 10.7 | 99.9 | 94.6 | 95.5 |
| RNet(无 \oplus 操作) | 0 | 99.85 | 6.65 | 98.97 | 17.69 | 91.08 |
| RNet(有 \oplus 操作) | 0 | 99.56 | 9.14 | 98.77 | 95.98 | 97.7 |

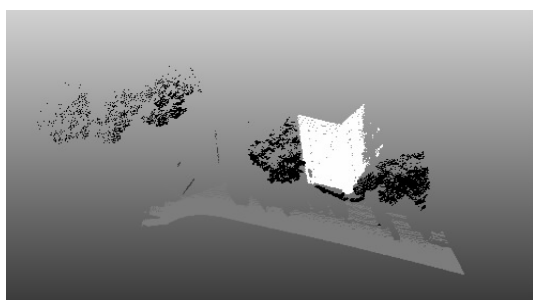
由表 3 可知, 在 scatter_misc 和 facade 这两大类别的分类精度表现优异, 在总体分类精度上, 所提出的 RNet 网络框架达到了 97.7%, 要优于其他方法, 并且有 \oplus 操作可以提升框架的整体分类精度。但是在

utility_pole 跟 default_wire 这两类上表现的差强人意, 造成这个结果的原因主要有: (1) 这两个类别的点数数据量较少; (2) 在进行特征提取的过程中, 存在点的覆盖问题。

图 4(a)为 Oakland 原始测试数据集可视化图,图 4(b)为在 Oakland 数据集上进行分类结果可视化。从图 4 的可视化可知,类别 utility_pole 跟 default_wire 的分类效果不好,后续还要对框架进行优化。



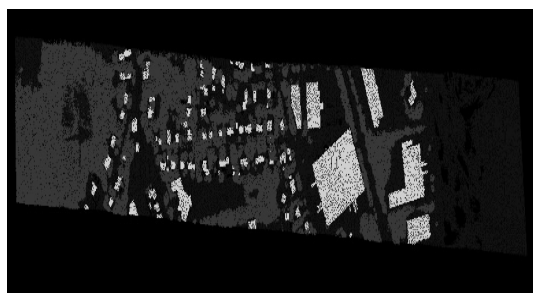
(a) Oakland 原始测试数据集可视化



(b) 结果可视化

图 4 可视化效果

收集了一些应用 GML_A 数据集的实验研究,与文中方法进行比较,如表 4 所示,表现比较好的是 Tree 和 Ground 这两个类别,分别达到了 91.77%、78.4% 的准确率,但是在 Building 和 Car 类别上的精度差强人意,造成这种情况的原因有数据集该类别的数据少等,后续要改进参数以达到更好的分类效果。结果可视化如图 5(b)所示。



(a) 经过预处理(Space 采样)之后的 GML_A 测试数据集



(b) 结果可视化

图 5 可视化效果

表 4 不同方法的对比精度 %

| 文献 | Tree | Car | Building | Ground |
|----------|-------|------|----------|--------|
| 文献 1[31] | 95.8 | 0.00 | 46.5 | 78.2 |
| 文献 2[32] | 83.2 | 0.00 | 33.56 | 44.82 |
| 文献 3[33] | 83.2 | 10.1 | 96.3 | 90.8 |
| ours | 91.77 | 0.00 | 0.00 | 78.4 |

4 结束语

由于近些年科技的发展与进步,大场景三维点云的应用领域越来越广泛,但是大场景点云分类中还存在很多问题,如由于大场景点云数据量庞大导致相关工作所需计算量巨大,占用内存过大,训练时间也过长,以及分类精度低等问题。该文提出了一种基于 Feature-RNet 的大场景三维点云分类框架,在 Oakland 3D 数据集上取得了不错的分类精度,但是在数据量少的点云类别上效果不尽如人意。同时在 GML_A 数据集上取得了不错的成绩,但是依然存在很大的进步空间,后续将不断优化框架以达到更好的分类准确率。在未来的工作中,将继续研究如何提升深度学习框架本身的计算性能,在保证不错的分类精度的同时能使用更少的训练时间,同时针对数据量少的点云也能取得不错的分类效果,并且将该模型框架应用到不同的大场景点云数据集中,以便提高框架的通用性。

参考文献:

- [1] 文 沛,程英蕾,余旺盛. 基于深度学习的点云分类方法综述[J]. 激光与光电子学进展,2021,58(16):41-67.
- [2] WANG Z,ZHANG L,FANG T,et al. A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification[J]. IEEE Transactions on Geoscience & Remote Sensing,2015,53(5):2409-2425.
- [3] YANG Bisheng,ZHEN Dong,YUAN Liu,et al. Computing multiple aggregation levels and contextual features for road facilities recognition using mobile laser scanning data[J]. ISPRS Journal of Photogrammetry and Remote Sensing,2017,126:180-194.
- [4] LALONDE J F,UNNIKRISHNAN R,VANDAPEL N,et al. Scale selection for classification of point-sampled 3D surfaces[C]//Fifth international conference on 3d digital imaging & modeling. Ottawa, ON, Canada: IEEE, 2005: 285-292.
- [5] LALONDE J F,VANDAPEL N,HUBER D,et al. Natural terrain classification using three-dimensional lidar data for ground robot mobility[J]. Journal of Field Robotics,2006,23(10):839-861.
- [6] MOUNTRAKIS G, IM J, OGOLE C. Support vector machines in remote sensing: a review [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2010, 66(3): 247-

- 259.
- [7] COLGAN M S, BALDECK C A, FERET C A, et al. Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data [J]. *Remote Sensing*, 2012, 4(11): 3462–3480.
 - [8] MIAO X, HEATON J S. A comparison of random forest and Adaboost tree in ecosystem classification in east mojave desert [C]//*Proceedings of the 18th international conference on geoinformatics*. Beijing, China: IEEE, 2010: 1–6.
 - [9] CHAN C W, DESIRÉ P. Evaluation of random forest and adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery [J]. *Remote Sensing of Environment*, 2008, 112(6): 2999–3011.
 - [10] SU H, MAJI S, KALOGERAKIS E, et al. Multi-view convolutional neural networks for 3D shape recognition [C]//*IEEE international conference on computer vision (ICCV)*. Santiago, Chile: IEEE, 2015: 945–953.
 - [11] MATURANA D, SCHERER S. VoxNet: a 3D convolutional neural network for real-time object recognition [C]//*2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. Hamburg, Germany: IEEE, 2015: 922–928.
 - [12] RAU J Y, JHAN J P, HSU Y C. Analysis of oblique aerial images for landcover and point cloud classification in an urban environment [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(3): 1304–1319.
 - [13] BARNEA S, FILIN S. Segmentation of terrestrial laser scanning data using geometry and image information [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2013, 76: 33–48.
 - [14] QI C R, SU H, MO K, et al. PointNet: deep learning on point sets for 3D classification and segmentation [C]//*IEEE conference on computer vision and pattern recognition (CVPR)*. Honolulu, HI, USA: IEEE, 2017: 77–85.
 - [15] 张志勇, 王蕾, 程海霞. 基于 Feature-pointNet 大场景点云分类 [J]. *电脑与信息技术*, 2021, 29(1): 6–9.
 - [16] RUSU R B, BLODOW N, BEETZ M. Fast point feature histograms (FPFH) for 3D registration [C]//*IEEE international conference on robotics & automation*. Kobe, Japan: IEEE, 2009: 3212–3217.
 - [17] 孙杰, 赖祖龙. 利用随机森林的城区机载 LiDAR 数据特征选择与分类 [J]. *武汉大学学报: 信息科学版*, 2014, 39(11): 1310–1313.
 - [18] RIEGLER G, ULUSOY A O, GEIGER A. OctNet: learning deep 3D representations at high resolutions [C]//*IEEE conference on computer vision and pattern recognition (CVPR)*. Honolulu, HI, USA: IEEE, 2017: 6620–6629.
 - [19] HU X, YUAN Y. Deep-learning-based classification for DTM extraction from ALS point cloud [J]. *Remote Sensing*, 2016, 8(9): 730.
 - [20] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. Las Vegas, NV, USA, 2016: 770–778.
 - [21] NAIR V, HINTON G E. Rectified linear units improve restricted Boltzmann machines [C]//*Proceedings of the 27th international conference on machine learning*. Haifa, Israel: [s. n.], 2010: 807–814.
 - [22] BISHOP C M. *Neural networks for pattern recognition* [M]. Oxford: Oxford University Press, 1995.
 - [23] RIPLEY B D. *Pattern recognition and neural networks* [M]. Cambridge: Cambridge University Press, 1996.
 - [24] VENABLES W N, RIPLEY B D. *Modern applied statistics with S-PLUS* [M]. New York, NY: Springer, 1999.
 - [25] MUNOZ D, BAGNELL J A, VANDAPEL N, et al. Contextual classification with functional max-margin Markov networks [C]//*IEEE conference on computer vision and pattern recognition*. Miami, FL, USA: IEEE, 2009: 975–982.
 - [26] SHAPOVALOV R, VELIZHEV A, BARINOVA O. Non-associative Markov networks for 3D point cloud classification [C]//*International archives of the photogrammetry, remote sensing and spatial information sciences*. Saint-Mandé, France: [s. n.], 2010: 103–108.
 - [27] WEINMANN M, JUTZI B, MALLET C. Feature relevance assessment for the semantic interpretation of 3D point cloud data [J]. *Isprs Annals of Photogrammetry*, 2013, II-5/W2: 313–318.
 - [28] HE E, WANG H, CHEN Q, et al. An improved contextual classification method of point cloud [J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(3): 362–370.
 - [29] WANG L, MENG W, XI R, et al. Large-scale 3D point cloud classification based on feature description matrix by CNN [C]//*Proc. 31th int. conf. comput. [s. l.]: Animation Social Agents*, 2018: 43–47.
 - [30] MUNOZ D, VANDAPEL N, HEBERT M. Directional associative Markov network for 3-D point cloud classification [C]//*Fourth international symposium on 3D data processing, visualization and transmission, 3DPVT 2008 - proceedings*. Atlanta, GA, USA: [s. n.], 2008.
 - [31] ZHANG Z, ZHANG L, TONG X, et al. Discriminative-dictionary-learning-based multilevel point-cluster features for ALS point-cloud classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016(12): 7309–7322.
 - [32] ZHANG Z, ZHANG L, TAN Y, et al. Joint discriminative dictionary and classifier learning for ALS point cloud classification [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 56(1): 524–538.
 - [33] HO T K, HULL J J. Decision combination in multiple classifier systems [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1994, 16(1): 66–75.