

结合物理与几何特性的机载 LiDAR 数据分类方法

赵毅强^{1,2}, 张琦^{1,2}, 刘长龙^{3,4}, 武唯康^{3,4}, 李尧^{1,2*}

- (1. 天津大学微电子学院, 天津 300072;
2. 天津大学天津市成像与感知微电子技术重点实验室, 天津 300072;
3. 中国电子科技集团公司第五十四研究所微系统中心, 河北石家庄 050081;
4. 通信软件与专用集成电路设计国家工程研究中心, 河北石家庄 050081)

摘要: 机载 LiDAR 数据分类是根据数据特征为每个点指定类别标签。针对现有方法忽略全波形与点云在物理特性上的关联、缺乏对邻域几何和语义相关性的深入挖掘, 从而导致捕获局部结构能力不足的问题, 搭建了结合目标物理与几何特性的分类方法, 实现了由全波形和点云组成的机载 LiDAR 数据端到端分类。首先, 构建了特征融合模块, 提取了全波形时序特征和点云几何特征, 依据两种数据物理意义上的关联, 通过双秩矩阵实现了全波形与点云特征级融合。其次, 构建了邻域特征增强模块, 挖掘点对相关性, 增强对局部几何结构的学习。最后基于层次化编解码结构搭建了分类网络。该网络在机载 LiDAR 数据集上测试, 达到平均精度 0.96、平均召回率 0.90、平均 F1 分数 0.92, 证明了网络的有效性。

关键词: 机载激光雷达; 数据分类; 特征融合; 全波形; 三维点云; 物理特性
中图分类号: TN958.98; TP7 **文献标志码:** A **DOI:** 10.3788/IRLA20230212

0 引言

全波形激光雷达 (Light Detection and Ranging, LiDAR) 可以获取三维点云及全波形, 被广泛应用于地形测绘^[1-2]、电力线检测^[3-4]及海陆分类岸线提取^[5-6]等领域。机载激光雷达数据分类是为每个点指定一个类别标签, 是激光雷达数据后处理的重要环节。研究表明, 全波形形状与目标物理特性密切相关, 例如, 倾斜表面的光路变长导致回波脉宽展宽^[7], 目标的粗糙度或颜色不同导致反射率不同进而影响回波功率^[8], 激光的穿透性为波形带来丰富的垂直结构信息等。而点云描述了目标的三维形状, 却无法反映上述全波形的特性。因此, 在分类时充分考虑全波形蕴含的目标物理特性和点云几何特征, 有助于提升数据分类精度。

目前, 典型的分类方法大多专注于点云数据^[9-13]。这些方法通过多层感知机和卷积等深度学习算法, 提

取了点云的几何结构特征, 但是却忽略了全波形中丰富的目标物理特性和垂直结构信息, 且对邻域内点间的几何和语义相关性的挖掘也不够充分。近年来, 一些研究人员也提出了同时考虑到点云与全波形的分类方法, 其中 Zoriz S 等^[14]将通过卷积生成的类概率向量与高度信息串联, 投影到二维图像进行分类, 但是该方法会丢失空间信息, 从而会降低数据分辨率。Shinohara T 等^[15]提出 FWNNet (Full-waveform Network) 和 FWNNet2^[16]直接将点云和全波形串联进行卷积, 没有考虑到全波形的时序性和目标的物理特性, 无法提取点云与全波形的相关性。并且 FWNNet2^[16]通过复杂的决策融合得到分类结果, 无法全面考虑两部分特征。

针对上述问题, 文中搭建了一种基于目标物理特性和几何特征的分类算法。首先, 文中构建了特征融合 (Feature Fusion, FF) 模块, 通过全波形特征提取 (Full-waveform Feature Extraction, FW) 模块对相邻和

收稿日期: 2023-04-10; 修订日期: 2023-05-20

基金项目: 广西创新驱动发展专项 (2018AA13005)

导师简介: 赵毅强, 男, 教授, 博士, 主要从事集成电路设计和红外成像与感知方面的研究。

通讯作者: 李尧, 男, 博士生, 主要从事混合信号集成电路芯片设计方面的研究。

长距时间位置特征进行提取,充分挖掘全波形的时序性,进一步地,根据全波形和点云在目标物理特性上的关联,构建了双低秩矩阵以学习全波形和点云几何特征间的相关性,实现了特征级融合。其次,构建了邻域特征增强 (Local Feature Enhancemen, LE) 模块,其中邻域相关性挖掘 (Local Interrelation mining, LI) 模块通过局部全连接结构提取点对间的几何和语义相关性,挖掘了邻域的几何结构信息,并通过注意力池化聚合邻域特征。最后,将构建的 FF 和 LE 模块嵌入层次化的编解码结构,以融合多感受野的特征,从而构建了基于特征级融合的端到端分类网络。

1 方法原理

1.1 基于目标物理特性相关性的全波形与点云特征融合

根据激光雷达能量公式,假设入射能量均匀散射成固体角为 Ω 的圆锥,接收激光能量^[17]可表示为:

$$P_r = \frac{P_t D_r^2}{4\pi R^4 \alpha^2} \sigma \quad (1)$$

$$\sigma = \frac{4\pi}{\Omega} \rho A_s \quad (2)$$

式中: P_t 为传输功率; D_r 为接收器孔径; R 为目标距离; α 为激光光束发散角; σ 为后向散射截面; A_s 为目标散射面积; ρ 为目标反射率。由上述公式可知,激光散射能量与目标反射率成正比,记 $\sigma_i = \frac{4\pi}{\Omega} A_s$, 则有:

$$P_r = \frac{P_t D_r^2 \rho}{4\pi R^4 \alpha^2} \sigma_i \quad (3)$$

由于发射波形呈高斯函数状,假设目标的后向散射各向同性,即可采用广义高斯函数描述目标回波。又因为机载激光雷达地物目标高度变化比飞行测量高度小得多,为简化问题,记 $S = \frac{P_t D_r^2}{4\pi R^4 \alpha^2} \sigma_i$, 可认为 S 近似不变。为分析回波特性,将接收面分割为微元,每个微元近似看作小平面,可将回波看作光斑内所有 N 个微元的回波集合,即:

$$f(t) = S \sum_{i=1}^N \rho_i \exp\left(-\frac{|t-t_i|^2}{2s_i^2}\right) \quad (4)$$

式中: s_i 为组分波形标准差; t_i 为回波位置,与目标距离有关,即与 $\frac{2R}{c}$ 有关, c 为激光传播速度。

当激光垂直照在平面上,如图 1(a) 所示, α 为激

光发散角, H 为激光发射点与地面垂直距离, B 、 E 为光斑边缘点, L 为光斑中心点。从图中可以看出, AL 段光路最短, AB 和 AE 段光路最长,因此可计算出回波间延时 Δt_1 :

$$\Delta t_1 = \frac{2(AB-H)}{c} = \frac{2(H/\cos\alpha-H)}{c} \quad (5)$$

当激光照在倾斜表面时,如图 1(b) 所示, θ 为斜面法向量与竖直方向夹角。从图中明显可以看出, AE 段长度大于 H 且大于图 1(a) 中的 AE 段,光路最长,假设 $AB \sim AL$ 间最短光路为 H' ,则可得到回波间延时 Δt_2 :

$$\Delta t_2 = \frac{2(AE-H')}{c} \quad (6)$$

$$AE = H(1 + \tan\alpha \tan\theta) \quad (7)$$

由此可以看出,当 H 相同时,倾斜平面上最长光路 AE 大于水平平面最长光路,最短光路 H' 小于等于水平平面最短光路 H 。因此在倾斜平面上回波间的延时更大,进而造成整体回波脉宽展宽,见图 1(b)。

当光斑内存在多个目标时,从图 1(c) 可以看出,激光照射在相距足够远的目标上,回波会呈现多个波峰,而照射在相距较近的目标上,反射的多个回波叠加会引起脉宽展宽。事实上,当激光照在多目标上,目标的反射率不同会造成回波振幅不同,同时根据公式 (3) 可知,回波能量与目标距离的四次方成反比,当目标相距较远时,回波的振幅也不同,下文对实验数据集地面和建筑类别全波形的分析也印证了这点。

相较于全波形,点云更直观地描述了目标几何形

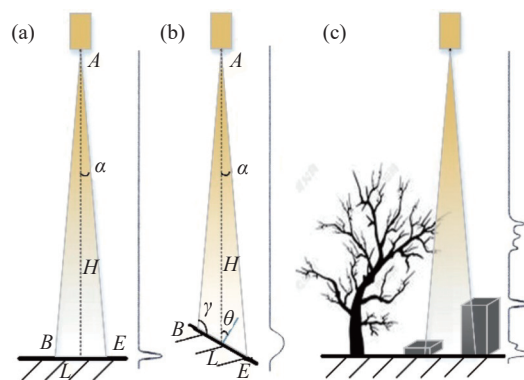


图 1 不同的激光照射情况。(a) 垂直于激光束的平面;(b) 倾斜平面;(c) 多目标

Fig.1 Different LiDAR irradiation situations. (a) Plane perpendicular to LiDAR beam; (b) Inclined plane; (c) Multiple targets

状,根据点云邻域可以提取丰富的几何特征,包括高程均值、方差、点云密度、法向量、曲率、线指数、面指数、球指数等。点云几何特征与全波形形状有着密切的相关性。当激光照在斜面上,根据斜面倾角,由公式(6)~(7)可推出光斑内回波延时,进而推测回波展宽情况,而根据点云邻域的最小二乘平面拟合又可估算该点的法向量,根据法向量坐标可计算得到该点所在面元的倾斜度,该倾斜度与全波形的展宽情况密切相关。激光照射在无法穿透的地面、建筑上形成单波峰回波,而该类点云的邻域高程方差较小,且没有复杂的垂直结构,曲率和球指数较小。由于植被具有缝隙,激光照射在植被上通常产生多回波,而植被点云的分布更复杂,曲率和球指数更大。考虑到点云和全波形所描述目标特性的相关性,通过深度学习进行高维特征的融合,有助于挖掘高维的目标特征,提升分类效果。

现有的结合全波形与三维点云的方法主要存在两方面问题:1) 现有方法大多不具备单独的全波形特征提取模块,而是将全波形看作点云的附加特征,由于点云坐标不具备时序性,采用多层感知机将所有特征维度无序相加,这种方法忽略了全波形的时序性,因此无法充分提取目标物理特性;2) 现有的特征融合方法通常为相加或串联拼接,这两种融合方法无法提取特征间复杂的线性交互关系,从而降低了特征提取效果,并且相加的方法还限制了融合的两组特征的维度。针对上述问题,文中构建了高维特征融合 FF 模块,如图 2 所示。首先,构建了全波形长短距特征提取 FW 模块来提取全波形时序性特征,其次,采用邻域相对位置编码和柱坐标编码,提取点云几何特征,最后,利用双低秩矩阵自适应地学习全波形与点云在高维特征间的相关性,从而提取出更深层的目标物理特性。下文将对模块各部分进行详细说明。

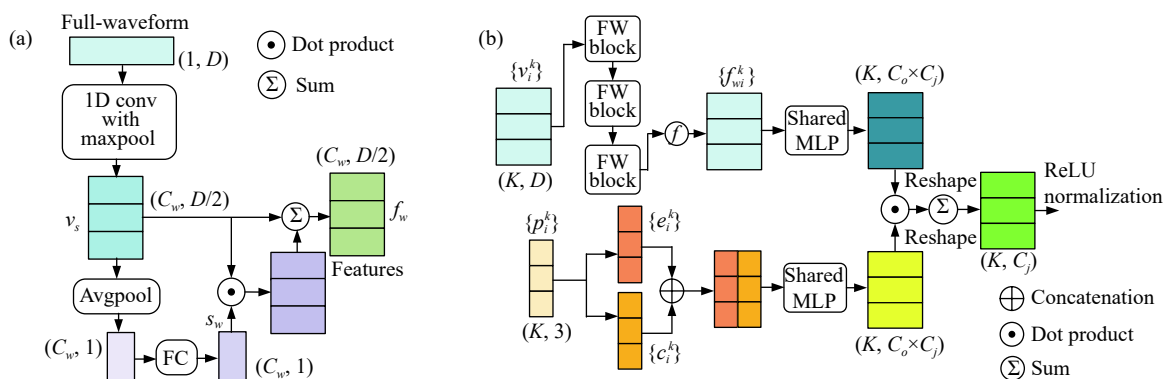


图 2 FW 与 FF 模块结构图。(a) FW 模块结构; (b) FF 模块结构

Fig.2 Structure of FW and FF block. (a) Structure of FW block; (b) Structure of FF block

1.1.1 全波形特征提取模块

全波形作为时间序列,直接在全波形时间维度上进行 1D 卷积可以提取相邻时间步的时序信息,却无法提取长距时间步的长期依赖关系。文中充分考虑了时间位置差异,先后对相邻的短距时序特征和全局的长距时序特征进行提取,构建了全波形特征提取 FW 模块。如图 2(a) 所示,其中 D 是组成全波形的元素数量, C_w 是卷积特征维数。首先,采用卷积核尺寸为 3 的 1D 卷积,提取相邻时间位置的短距特征 v_s ,采用归一化和 ReLU 函数进行非线性处理。进一步地,通过过滤器数量为 2 的最大池化层,实现对输入微小

扰动的近似不变,同时减少参数。接着,为提取长距的全局特征的影响,在时间维度上进行全局平均池化来提取全局特征,采用全连接层自动学习权重 $s_w \in R^{1 \times C_w}$,动态调整短距特征 v_s ,从而得到长距特征。为防止随层数堆叠导致拟合效果变差,将得到的长距特征作为残差与短距特征 v_s 相加。最终输出包含目标物理特性的全波形时序特征 f_w 如下:

$$v_s = \text{MaxPool}(g_c(V)) \quad (8)$$

$$s_w = g_f(\text{AvgPool}(v_s)) \quad (9)$$

$$f_w = s_w \odot v_s + v_s \quad (10)$$

式中: V 表示一个由 D 个元素组成的全波形向量;

g_c 为卷积核大小为 3 的 1D 卷积; g_f 为全连接层; \odot 代表元素级乘法。最后,为了获取不同时间感受野的波形特征,将 FW 模块进行堆叠使用,如图 2(b)左上部分所示,最终得到每个点的全波形特征。

1.1.2 点云几何特征提取及与全波形特征的融合

首先编码点云邻域的几何特征。通过 k 近邻算法得到每个采样点 p_i 的 k 邻域点 $N_k(p_i)$, 计算邻域点 $p_i^k \sim p_i$ 的相对位置和距离。由于邻域对中心点的影响与距离呈负相关,因此用相对距离的负指数来编码,得到邻域相对位置特征:

$$e_i^k = p_i \oplus p_i^k \oplus (p_i - p_i^k) \oplus \exp(-dis_i^k) \quad (11)$$

$$dis_i^k = \sqrt{x_i^{k2} + y_i^{k2} + z_i^{k2}} \quad (12)$$

式中: \oplus 代表特征维度上的串联; (x_i^k, y_i^k, z_i^k) 为笛卡尔坐标系下的相对坐标。由于真实场景中同一类物体的方向可能不同,增加相对中心点 z 轴旋转不变的柱坐标编码 c_i^k , 可表示为:

$$c_i^k = \phi_i^k \oplus disxy_i^k \quad (13)$$

$$\phi_i^k = \arctan\left(\frac{y_i^k}{x_i^k}\right) \quad (14)$$

$$disxy_i^k = \sqrt{x_i^{k2} + y_i^{k2}} \quad (15)$$

将柱坐标编码与笛卡尔坐标编码串联,通过多层感知机进行特征映射,最终得到邻域点 p_i^k 的几何特征。

其次,对邻域点的全波形 v_i^k , 采用 1.1.1 节所述方法得到全波形特征 f_i^k 。之后提取全波形与点云几何特征在高维特征上的关联。由前面的分析可知,全波形包含了丰富的目标物理特性,其形状与点云几何特征密切相关,而传统相加的融合方法限制了融合的两组特征的维度,串联的融合方法无法提取特征间复杂的线性交互关系。针对该问题,文中采用映射矩阵 $G \in R^{m \times n}$ (其中 n, m 为两特征的维度) 概括全波形和点云几何特征在每个特征维度上的相互映射,学习特征间的线性关系。对于任一点 p , 其几何特征 $g \in R^n$, 全波形特征 $f_w \in R^m$, 构建映射矩阵 G , 将线性融合特征 $f \in R^l$ 可表示为:

$$f_j = g \times G \times f_w \quad (16)$$

式中: \times 为矩阵乘法; f_j 为 f 的第 j 维特征。图 3 说明了

映射矩阵 G 学习两特征间的线性关系的原理。根据矩阵乘法原理,映射矩阵 G 的每一列都与特征向量 g 的所有维度一一对应,如图 3 中红色框所示,每列与特征 g 对应相乘并相加得到该列的结果,映射矩阵 G 的所有列都会得到一个单独的计算结果,因此每列的参数可看作是向量 g 不同维度上的相关性;同理,每一行与特征向量 f_w 的所有维度一一对应,如图 3 中的蓝色框所示,每行可看作是向量 f_w 不同维度上的相关性,因此映射矩阵可以有效地学习不同特征维度间的两两交互关系,通过与映射矩阵相乘,可得到一个点云与全波形的线性融合特征 f_j 。

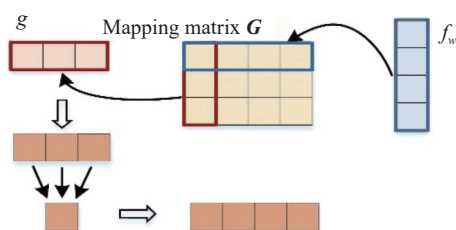


图 3 映射矩阵作用原理

Fig.3 Principle of mapping matrix

为丰富特征表达,通常会求得多个维度的融合特征,因此需要多个不同的映射矩阵,但这会引入大量的参数,增大计算开销,甚至造成过拟合。为减少参数,将矩阵 G 拆分为两个相乘的低秩矩阵 $W \in R^{m \times o}$ 和 $U \in R^{o \times n}$ (其中 $o \ll n, m$), 则 f_j 可表示为:

$$f_j = \sum (g \times W_j) \odot (f_w \times U_j) \quad (17)$$

式中: $W_j \in R^{m \times o}$; $U_j \in R^{o \times n}$ (其中 $o \ll n, m$); \sum 代表对结果的所有元素求和。 W_j 和 U_j 学习两个特征在每个维度上的相互映射,可用多层感知机学习参数,将其分别与特征 g 和 f_w 相乘,得到 $S \in R^{1 \times o}$, $T \in R^{1 \times o}$, 则 S 与 T 各自融合了两特征的所有维度。为实现两特征间的交互,将 S 与 T 进行元素级乘法并融合所有维度,得到第 j 维的融合特征 f_j 。为快速计算 f 的所有维度,同时计算 f 的 j 维特征的全部映射,得到矩阵 $S \in R^{1 \times (j \times o)}$ 和 $T \in R^{1 \times (j \times o)}$, 先将矩阵恢复到三维,之后再对矩阵进行元素级乘法,得到融合特征 $f \in R^l$, 如图 2(b) 所示,图中 C_j 表示特征维度, C_o 表示映射矩阵维度。由于过程中引入了元素乘法,输出神经元大小可能变化很大,为防止陷入局部最小,最后对融合后的特征进行归一化处理。

1.2 局部邻域特征增强模块

由于机载激光雷达飞行高度较高,采样点更稀疏,因此加强对邻域内上下文信息的挖掘有助于提升分类精度。以往的研究大多专注于邻域与中心点的相对位置,忽略了对其他点间相关性的挖掘。针对这一问题,文中构建了邻域特征增强 LE 模块。首先,搭建了邻域相关性挖掘 LI 模块,通过邻域全连接结构,提取点对间几何和语义相关性。之后根据每个点的特征,采用注意力池化有针对性地聚合邻域。

1.2.1 邻域相关性挖掘

为提取邻域中每个点对其他点的影响,将邻域点两两相连,构成如图 4(a) 所示的全连接结构,其中每个点对 (p_i^j, p_i^k) , 对应特征 (f_i^j, f_i^k) , 代表第 i 点的第 j 个邻域点和第 k 个邻域点形成的点对结构。LI 模块结构如图 4(b) 左上部分所示,图中 C 代表特征维度。邻域点往往是由同一目标或相近目标扫描得到的,因此在几何形状上和语义特征上相关。为提取点对的几

何形状相关性,用相对距离的负指数与点对坐标串联,来表征点对间的相互影响。为提取语义相关性,计算点对语义特征差异 Δf_{jk} 及特征乘积,则点对间的相互关系可表示为:

$$\Delta g_{jk} = p_i^j \oplus p_i^k \oplus \exp(-dis_{jk}) \quad (18)$$

$$dis_{jk} = \sqrt{(x_i^j - x_i^k)^2 + (y_i^j - y_i^k)^2 + (z_i^j - z_i^k)^2} \quad (19)$$

$$\Delta f_{jk} = \exp(-\text{mean}(\text{abs}(f_i^j - f_i^k))) \quad (20)$$

$$\Delta F_{jk} = \text{MLP}(\Delta g_{jk} \oplus \Delta f_{jk} \oplus (f_i^j \times f_i^k)) \quad (21)$$

进一步地,计算邻域点对相关性分数,与对应点特征相乘并聚合,则邻域中的每个点都根据点对差异动态更新了对该点的影响,融合了与其他所有点间的相互关系,从而得到增强的局部上下文特征如下:

$$F_i^k = f_i^k + \sum_{j=0}^{j=K} \text{Softmax}(\Delta F_{jk}) f_{jk} \quad (22)$$

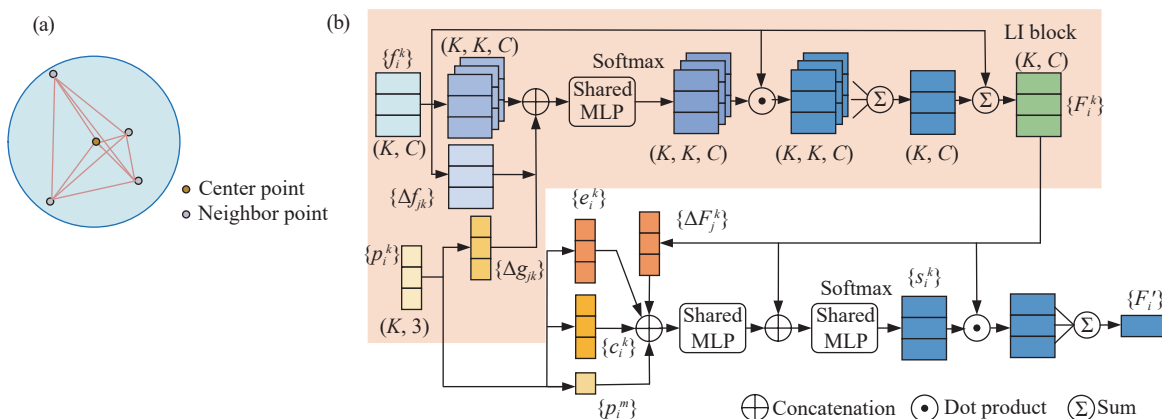


图 4 局部邻域特征增强。(a) 局部邻域全连接结构;(b) LE 模块结构

Fig.4 Local neighborhood feature enhancement. (a) Local neighborhood fully connected structure; (b) Structure of LE block

1.2.2 注意力池化

为有效聚合邻域,文中根据邻域几何和语义特征进行注意力池化,采用下文所述方式提取邻域几何特征,并实现对 z 轴的旋转不变。对采样中心及邻域点语义特征 F_i 和 F_i^k , 用负指数表征语义特征差异:

$$\Delta F_i^k = \exp(-\text{mean}(\text{abs}(F_i - F_i^k))) \quad (23)$$

同时计算邻域质心 p_i^m , 以从 $p_i \sim p_i^m$ 的方向值表征邻域分布情况,通过使用平均值降低下采样来引入随

机性。对拼接的特征计算注意力得分并聚合,得到体现邻域几何结构和语义关系的中心注意力特征:

$$s_i^k = \text{Softmax}(\text{MLP}(\text{MLP}(g_i^k \oplus p_i^m \oplus \Delta F_i^k) \oplus F_i^k)) \quad (24)$$

$$F_i' = \sum_{j=0}^{j=K} s_i^j \times F_i^j \quad (25)$$

式中: s_i^k 表示邻域点 p_i^k 对采样中心点 p_i 的注意权重。

1.3 全局特征增强

为了补充全局上下文,提升分类准确性,文中构

建全局特征增强 GE 模块。对于解码器输出特征 $F \in R^{N \times C}$, 其中 N 是场景中采样点数, C 是特征通道数, 首先通过平均池化层提取全局特征, 其次用带有激活函数的全连接层进一步学习全局场景特征, 将其作为全局特征权重 s_g , 最后与解码器输出特征 F 相乘, 根据全局场景动态调整解码器输出特征。得到的全局增强特征可表示为:

$$Z = F + s_g \odot F \quad (26)$$

1.4 整体架构

文中搭建了一种基于目标物理特性和几何特征的分类方法。该方法根据全波形与点云反映的目标物理特性, 考虑到二者特征间的相关性, 构建了高维特征融合 FF 模块, 利用双低秩矩阵提取高维特征间相关性, 进一步挖掘目标深层物理特性。同时还构建了邻域特征增强 LE 模块, 构建了邻域全连接结构, 充

分挖掘邻域点对间几何和语义相关性, 增强了局部结构信息。最后将模块嵌入层次化编码-解码器, 构建了全波形机载激光雷达数据的端到端分类网络。图 5 显示了整体架构, 网络将 $N \times (3+D)$ 的二维矩阵输入到编解码结构中, 其中 D 是全波形维度, N 为输入采样点数。在编码阶段采取最远点采样, 以最大限度保留场景空间结构, 每层通过 LE 模块编码, 逐层学习空间上下文特征。需要注意的是, 在第一层采用 LE 模块提取特征前先使用特征融合 FF 模块, 融合全波形和点云几何特征, 如图 5 中砖红色矩形所示, 之后再对融合后的特征进行进一步提取。在解码阶段采用最近邻插值上采样, 为合并低级别信息、增加不同尺度的感受野, 将相同维度的下采样与上采样块的特征相连, 利用卷积生成密集的特征预测。最后通过 GE 模块补充全局场景特征, 将结果输入到全连接层, 进行标签分类。

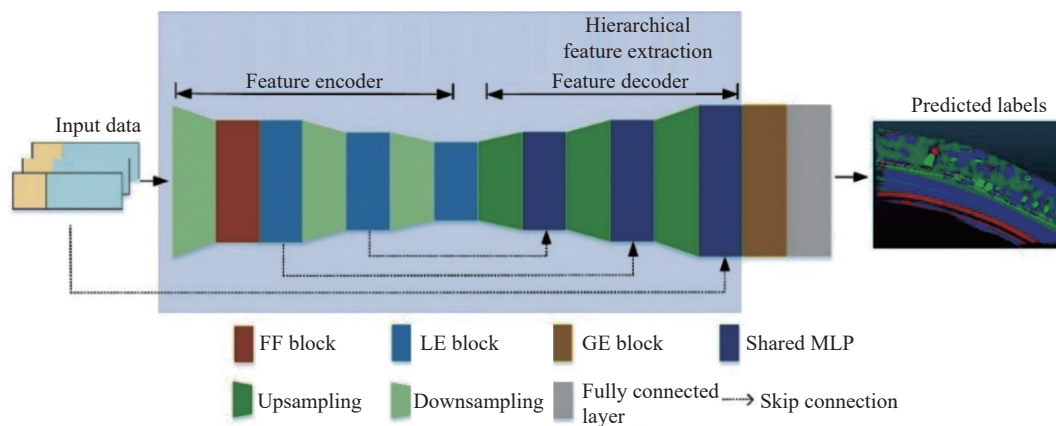


图 5 网络架构

Fig.5 Network architecture

2 实验与分析

2.1 数据集介绍

文中使用 RiegI LMS-Q780 全波形 ALS 获得的数据集^[14]进行训练和测试, RiegI LMS-Q780 是一款多功能高空激光雷达传感器, 它运行在离地面 15500 ft (1 ft= 12 in), 全视野可达 60°, 可用于大规模地形测绘。数据集由超过 980 万个采样点组成, 分六类: 地面、植被、建筑、电力线、传输塔和街道。每个点由 164 个数据组成, 包含三部分: 三维坐标、160 个值的全波形和类标签。图 6 为全波形可视化, 其中植被往

往有多回波, 波峰数量较多且不同点的波峰数量和分布位置不同, 横坐标为相对时间, 纵坐标为波形相对强度。电力线、传输塔也有多回波, 但明显少于植被。地面、建筑和街道回波相似, 多为单波峰, 波峰位置相近且有拖尾。为进一步寻找三类波形差异, 每类随机取 1000 个回波, 统计三类回波峰值及波峰位置分布的区间, 对比发现三类回波峰值位置相近, 峰值分布区间有交叉, 区间上限相近, 区间下限由高到低为建筑、地面和街道。由此可见, 建筑、地面和街道在波形形状上大体相似但仍有微小差别, 但与其他三

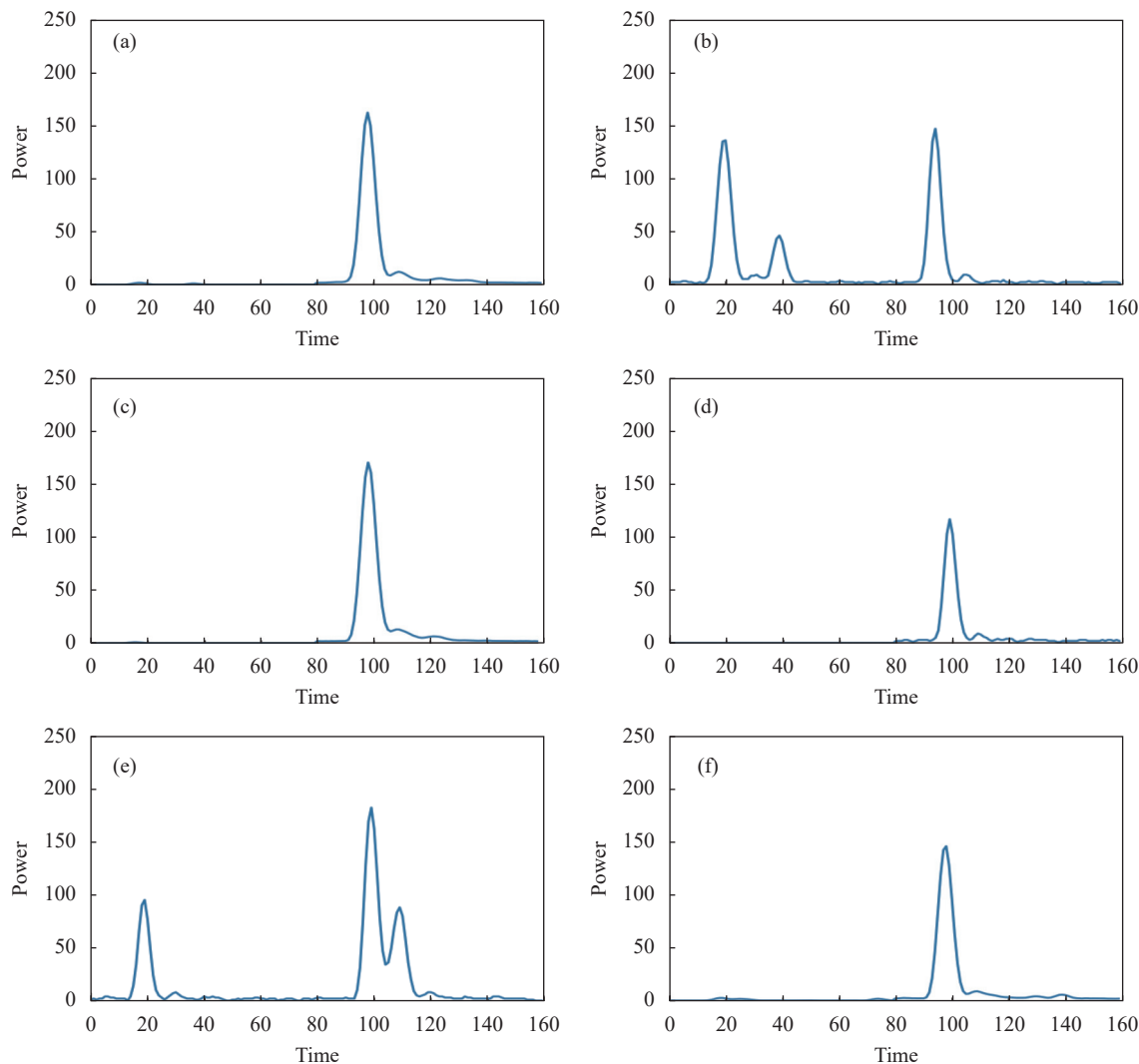


图 6 不同类别采样点的全波形。(a) 地面全波形; (b) 植被全波形; (c) 建筑全波形; (d) 电力线全波形; (e) 传输塔全波形; (f) 街道全波形

Fig.6 Full-waveform of sampling points of different categories. (a) Ground full-waveform; (b) Vegetation full-waveform; (c) Building full-waveform; (d) Power line full-waveform; (e) Transmission tower full-waveform; (f) Street path full-waveform

类相比差异明显。

2.2 预处理、实验细节和评价指标

原始数据集是对场景的连贯扫描,包含数百万采样点,被划分存储到 19 个子集。每个子集包含数十万点,其中一个子集作测试集,仅在测试阶段使用,其余子集作训练集。由于 GPU 内存限制,无法对如此大规模的数据直接进行计算,因此将所有训练子集分割为长宽均为 21.88 m 的样本块,进一步地,减去中心坐标转换为局部坐标系,使样本更具一般性。测试时以相同尺寸分割测试集,便于数据批处理。由于扫描对象的形状和大小不同,数据集在不同类别上的分布不平衡,因此为降低对模型性能的影响,对分割的样本块进行筛选。从表 1 可知,地面和植被点数量明显

更多,占测试集的 89.8%,因此训练时对包含了全部较少四类的样本块全部选取,对其余样本块随机选取。实验中训练集共被分割为 934 个样本块,选取包含全部较少四类的共 539 个样本,其余样本随机选取,共选取了 600 个训练样本。针对较少四类样本间仍存在的类别不平衡问题,在计算损失函数时添加类别权重,公式如下:

$$\lambda_b = \frac{1}{\ln\left(\beta + K_b \sqrt{\sum_{i=1}^B K_i}\right)} \quad (27)$$

式中: λ_b 为第 b 类的权重; K_b 为第 b 类的点数; B 为类别总数; β 为类平衡系数,文中设置为 1.1。

网络使用 Pytorch 框架实现,用 Adam 优化器在

表 1 数据集分布

Tab.1 Dataset distribution

Label	Class	Train		Test	
		Number	Number	Number	Number
1	Ground	1 787 352	20.4%	193 070	18.1%
2	Vegetation	4 719 634	53.9%	765 327	71.7%
3	Building	1 514 486	17.3%	49 138	4.6%
4	Power line	71 978	0.8%	8 152	0.8%
5	Transmission tower	32 008	0.4%	1 829	0.2%
6	Street path	633 606	7.2%	49 580	4.6%

两块 NVIDIA GeForce RTX 3090 GPU 上进行 200 轮训练。设置初始学习率为 0.001, 每 30 个训练周期后减少 10%。使用精度、召回率和 F1 分数评价分类效果。其中, 精度是衡量过度检测的指标, 召回率是衡量返回多少真正相关结果的指标, F1 分数根据精度和召回率计算得到, 在类别不平衡的情况下更适用。计算公式如下:

$$precision = \frac{TP}{TP + FP} \quad (28)$$

$$recall = \frac{TP}{TP + FN} \quad (29)$$

$$F1\ score = \frac{precision \times recall}{precision + recall} \quad (30)$$

式中: TP、FP 和 FN 分别表示真正例、假正例和假反例的数量。

2.3 测试结果与分析

为定量评价分类性能, 对比现有的融合全波形和点云的深度学习方法, 包括 CNN^[14]、FCN^[14] 和 FW-Net2^[16], 以及典型的基于点云的分类方法, 包括 PointNet^[9]、PointNet ++^[10]、RandLA-Net^[11]、GACNN^[12] 和 GFSAE^[13], 分类结果见表 2。需要注意的是, 在对比方法中, CNN 为仅采用全波形数据进行分类, 为对比文中网络的效果, 其余网络均同时使用了全波形与点云数据进行分类。其中 FCN 先通过 1D 卷积提取全

表 2 不同方法之间的定量比较

Tab.2 Quantitative comparison of different methods

Method	Metric	Ground	Vegetation	Building	Power line	Transmission tower	Street path	Mean
CNN ^[14]	Precision	0.60	0.96	0.07	0.10	0.01	0.18	0.32
	Recall	0.18	0.78	0.11	0.88	0.44	0.60	0.50
	F1 score	0.28	0.86	0.09	0.19	0.02	0.28	0.28
FCN ^[14]	Precision	0.79	0.95	0.99	0.98	0.63	0.67	0.84
	Recall	0.84	0.97	0.93	0.91	0.88	0.69	0.87
	F1 score	0.81	0.96	0.96	0.94	0.73	0.68	0.85
PointNet ^[9]	Precision	0.56	0.97	0.95	0.92	0.61	0.94	0.83
	Recall	0.91	0.85	0.83	0.84	0.48	0.62	0.76
	F1 score	0.69	0.91	0.88	0.88	0.53	0.75	0.77
PointNet++ ^[10]	Precision	0.56	0.98	0.99	0.97	0.52	0.96	0.83
	Recall	0.93	0.85	0.85	0.85	0.66	0.60	0.83
	F1 score	0.70	0.91	0.91	0.91	0.58	0.74	0.80
GFSAE ^[13]	Precision	0.77	0.97	0.93	0.96	0.70	0.83	0.86
	Recall	0.90	0.95	0.93	0.93	0.73	0.62	0.84
	F1 score	0.83	0.96	0.93	0.94	0.71	0.71	0.85
RandLA-Net ^[11]	Precision	0.74	0.98	0.83	0.93	0.81	0.99	0.88
	Recall	0.95	0.93	0.95	0.97	0.67	0.60	0.85
	F1 score	0.83	0.95	0.89	0.95	0.73	0.75	0.85
GACNN ^[12]	Precision	0.76	0.98	0.87	0.95	0.74	0.94	0.87
	Recall	0.95	0.94	0.95	0.98	0.79	0.63	0.87
	F1 score	0.84	0.96	0.91	0.96	0.76	0.75	0.86
FWNet2 ^[16]	Precision	0.80	0.98	0.99	0.99	0.99	0.97	0.95
	Recall	0.94	0.96	0.95	0.99	0.79	0.71	0.89
	F1 score	0.86	0.97	0.97	0.99	0.88	0.82	0.92
Proposed method	Precision	0.83	0.98	0.98	0.99	0.97	0.98	0.96
	Recall	0.95	0.96	0.95	0.99	0.82	0.75	0.90
	F1 score	0.89	0.97	0.96	0.99	0.89	0.85	0.92

波形特征,之后与 z 坐标结合投影到二维图像进行分类;而余下方法将全波形作为点云附加特征,与点云串联拼接。实验结果表明,文中的平均精度、召回率和 $F1$ 分数分别达到 0.96、0.90 和 0.92,6 个类的 $F1$ 分数均大于 0.85,可以有效实现分类。相比于当前最优的 FWNet2,该方法在平均精度和召回率上均提升 0.01,平均 $F1$ 分数也达到最优效果。同时文中方法在地面和街道类的精度、召回率和 $F1$ 分数比 FWNet2 分别提高了 0.03、0.01、0.03 和 0.01、0.04、0.03,植被和传输塔类也达到了最优效果。

为定性观察分类效果,将结果可视化如图 7 所

示,可以看出模型成功为大部分点生成了正确的标签,并提升了准确度,如图中白色圈中所示。观察可见,大部分街道周围分布着地面和植被。其中地面与街道点云形状相似,其全波形多为单波峰,峰值相近但仍有区别。植被全波形大多为多回波,通常最后一个波峰即为地物回波,因此植被全波形包含了该点的垂直空间结构。文中的特征融合模块提取了植被、地面、街道全波形中蕴含的丰富物理特性,自适应地学习其与点云几何特征的相关性,从而挖掘深层的物理特性关联,同时邻域特征增强模块学习了植被融合特征中包含的丰富垂直结构,提升了对街道的识别效果。

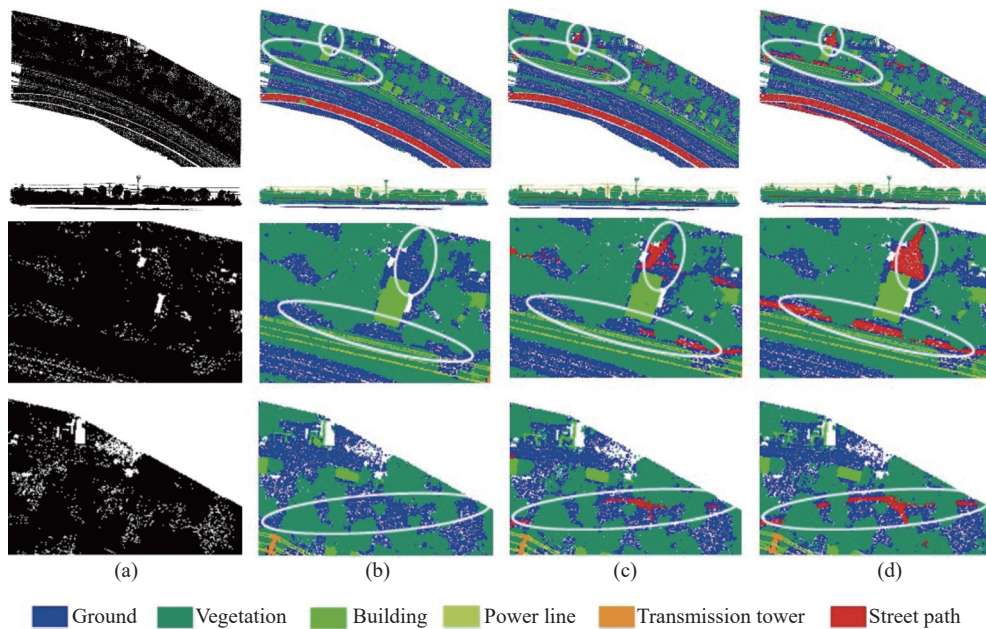


图 7 分类结果可视化。(a) 输入数据; (b) FWNet2 的分类结果; (c) 文中方法的分类结果; (d) 真实类别标签

Fig.7 Visualization of classification results. (a) Input data; (b) Classification results of FWNet2; (c) Classification results of proposed method; (d) Real category label

2.4 消融实验

为验证网络各模块的效果,本节在文献 [14] 的数据集上进行了消融实验研究。

2.4.1 对特征融合模块的消融实验

文中的特征融合模块由两部分组成:全波形特征提取、点云几何特征提取及特征融合。为验证各部分的效果,本节设计并进行了消融实验,如表 3 所示。设计了模型 A 作为对照,该模型采用 1D 卷积提取全波形特征,并以串联方式结合特征。模型 B 在此基础上采用 FW 模块提取全波形时序性特征,平均精度、

召回率和 $F1$ 分数提高了 0.02、0.02 和 0.01,说明 FW 模块通过提取长短距时序信息,得到更丰富的全波形特征,有效地提升了分类性能。模型 C 在模型 A 的基础上以双低秩矩阵的融合方式替代串联融合,提取高维特征间的相关性,平均精度、召回率和 $F1$ 分数提高了 0.04、0.03 和 0.02,表明通过双低秩映射矩阵学习两特征间的关联可提升分类性能。同时,文中方法在模型 B 基础上以双低秩矩阵融合方式代替串联,在模型 C 基础上以堆叠的 FW 模块提取全波形时序特征,分类效果均提高,进一步验证了特征融合模块

表 3 对特征融合模块的消融实验

Tab.3 Ablation experiment on FF block

Model	FW block	Fusion	Mean precision	Mean recall	Mean F1 score
A	×	⊕	0.90	0.86	0.88
B	√	⊕	0.92	0.88	0.89
C	×	√	0.94	0.89	0.90
Proposed method	√	√	0.96	0.90	0.92

对分类性能的提升。

本节还对 FW 模块的堆叠层数进行了消融实验, 以确定最优层数, 实验结果如表 4 所示。可以看出, 最初随着模块堆叠层数的增加, 网络分类能力逐渐增强, 然而当堆叠层数达到四层时, 网络分类性能下降, 因此文中网络选择堆叠三层 FW 模块, 以达到最优的分类效果。

表 4 对 FW 模块堆叠层数的消融实验

Tab.4 Ablation experiment on stacking layers of FW block

Model	Number of layers	Mean precision	Mean recall	Mean F1 score
D	1	0.94	0.89	0.91
E	2	0.95	0.90	0.91
F	3	0.96	0.90	0.92
G	4	0.95	0.89	0.91

2.4.2 对邻域与全局特征增强模块的消融实验

文中的邻域特征增强 LE 模块由两部分构成: 邻域相关性挖掘 LI 模块、注意力池化。为验证 LE 模块各部分以及全局特征增强 GE 模块的效果, 本节进行了消融实验, 如表 5 所示。设计了模型 H 作为对照, 该模型采用多层感知机提取单点特征, 以最大池化聚合邻域。模型 I 在此基础上采用文中的 LI 模块提取邻域点对间的几何和语义相关性, 增强邻域特征, 分类平均精度、召回率和 F1 分数分别提高了 0.03、0.02 和 0.01。模型 J 在模型 H 的基础上, 采用注意力池化替代最大池化, 从而根据邻域几何和语义特征, 有针对性地聚合, 三个指标分别提高了 0.01、0.02 和 0.01, 验证了邻域特征增强模块对分类性能的提升。模型 K 在文中网络的基础上, 去掉 GE 模块, 直接对解码器输出的特征进行分类。文中网络加入 GE 模块后, 三个指标提高了 0.02、0.02 和 0.02, 验证了 GE 模

表 5 对邻域与全局特征增强模块的消融实验

Tab.5 Ablation experiments on LE and GE block

Model	LE		GE	Mean precision	Mean recall	Mean F1 score
	LI block	Attention Pooling				
H	MLP	Max pooling	√	0.91	0.86	0.88
I	√	Max pooling	√	0.94	0.88	0.89
J	MLP	√	√	0.92	0.88	0.89
K	√	√	×	0.94	0.88	0.90
Proposed method	√	√	√	0.96	0.90	0.92

块对分类性能的提升。

3 结 论

文中搭建了一种结合目标物理特性和几何特征深度学习算法, 实现了机载 LiDAR 数据的端到端分类。针对全波形与点云在目标物理特性上的相关性, 构建了特征融合 FF 模块, 提取了全波形时序特征, 通过双低秩矩阵学习全波形与点云几何特征间的相关性, 实现了对高维物理特性的挖掘。通过构建邻域特征增强 LE 模块, 挖掘了邻域点间几何和语义特征的相关性, 加强了对局部结构信息的提取, 并根据邻域特征有针对性地聚合。最后将模块嵌入层次化的编解码结构中, 形成了基于特征级融合的端到端分类网络。在公开数据集上的分类实验表明, 六个类的 F1 分数均大于 0.85, 平均精度、平均召回率和平均 F1 得分分别达到 0.96、0.90 和 0.92, 与已有的分类算法相比, 文中方法有效提升了分类精度。

参考文献:

[1] Zhang K, Ye L J, Wen X, et al. A dual attention neural network for airborne LiDAR point cloud semantic segmentation [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-17.

[2] Li Pengcheng, Xu Qing, Xing Shuai, et al. Full-waveform LiDAR data decomposition method based on global convergent LM [J]. *Infrared and Laser Engineering*, 2015, 44(8): 2262-2267. (in Chinese)

[3] Nardinocchi C, Balsi M, Esposito S, et al. Fully automatic point cloud analysis for powerline corridor mapping [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020, 58(12): 8637-8648.

- [4] Li W, Luo Z, Xiao Z, et al. A GCN-based method for extracting power lines and pylons from airborne LiDAR data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 1-14.
- [5] Hu Shanjiang, He Yan, Tao Bangyi, et al. Classification of sea and land waveforms based on deep learning for airborne laser bathymetry [J]. *Infrared and Laser Engineering*, 2019, 48(11): 1113004. (in Chinese)
- [6] Liang G, Zhao X, Zhao J, et al. MVCNN: A deep learning-based ocean-land waveform classification network for single-wavelength LiDAR bathymetry [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023, 16: 656-674.
- [7] Shen Jun, Shang Jianhua, He Yan. Research process of data processing technology of full-waveform airborne laser radars [J]. *Laser Technology*, 2018, 42(3): 295-299. (in Chinese)
- [8] Zhu Shixian. Research on critical technology of intensity imaging of 905 nm pulsed LiDAR[D]. Tianjin: Tianjin University, 2018. (in Chinese)
- [9] Charles R Q, Hao S, Mo K C, et al. Point-Net: deep learning on point sets for 3D classification and segmentation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 77-85.
- [10] Qi C R, Yi L, Su H, et al. Pointnet++: Deep hierarchical feature learning on point sets in a metric space[C]//Advances in Neural Information Processing Systems, 2017: 5099-5108.
- [11] Hu Q, Yang B, Xie L, et al. RandLA-Net: Efficient semantic segmentation of large-scale point clouds [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 11105-11114.
- [12] Cong W, Xiang L, Yao X J, et al. Airborne LiDAR point cloud classification with global-local graph attention convolution neural network [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2021, 173: 181-194.
- [13] Chen Q, Zhang Z, Chen S, et al. A self-attention based global feature enhancing network for semantic segmentation of large-scale urban street-level point clouds [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 113: 102974.
- [14] Zoriz S, Maset E, Fusiello A, et al. Full-waveform airborne LiDAR data classification using convolutional neural networks [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(10): 8255-8261.
- [15] Shinohara T, Xiu H, Matsuoka M, et al. FWNNet: Semantic segmentation for full-waveform LiDAR data using deep learning [J]. *Sensors*, 2020, 20(12): 3568.
- [16] Shinohara T, Xiu H, Matsuoka M, et al. Semantic segmentation for full-waveform LiDAR data using local and hierarchical global feature extraction[C]//Sigspatial'20: 28th International Conference on Advances in Geographic Information Systems, 2020: 640-650.
- [17] Wagner W, Ullrich A, Ducic V, et al. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitizing airborne laser scanner [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2006, 60(2): 100-112.

Airborne LiDAR data classification method combining physical and geometric characteristics

Zhao Yiqiang^{1,2}, Zhang Qi^{1,2}, Liu Changlong^{3,4}, Wu Weikang^{3,4}, Li Yao^{1,2*}

(1. School of Microelectronics, Tianjin University, Tianjin 300072, China;

2. Tianjin Key Laboratory of Imaging and Sensing Microelectronic Technology, Tianjin University, Tianjin 300072, China;

3. Microsystem Center, The 54th Institute of China Electronics Technology Group Corporation, Shijiazhuang 050081, China;

4. National Engineering Research Center for Communication Software and Special Integrated Circuit Design, Shijiazhuang 050081, China)

Abstract:

Objective Full-waveform airborne LiDAR can not only obtain 3D point clouds, but also obtain full-waveform of the target. The classification of full-waveform LiDAR data is to assign a category label to each point based on data characteristics. It is an important part of data post-processing and has significant application value in the fields such as remote sensing and topographic mapping. However, existing methods overlook the correlation

between full-waveform and point clouds geometry in terms of physical characteristics, and lack in-depth exploration of the correlation in neighborhood geometry and semantic features between point pairs. Therefore, the existing methods lack the ability to capture local structural information, which affects the classification effect. To address the above issues, a classification method based on the physical and geometric characteristics of the target is proposed. This method mines the high-dimensional feature association between 3D point clouds and full-waveforms on the high-dimensional physical characteristics of the target, and enhances the learning of local geometric structural information, thereby improving the classification ability of the model.

Methods A classification method based on target physical and geometric features is proposed. Firstly, a high-dimensional feature fusion module is proposed, which extracted the rich physical characteristics from the full-waveform, and extracted the geometric features of the point clouds. Based on the correlation between full-waveforms and point clouds on the physical characteristics of the target, the complex relationship between the two is learned through a dual low-rank matrix, and deeper physical features are excavated (Fig.2). Secondly, a local neighborhood feature enhancement module is designed to enhance the learning of local geometric structures by constructing a fully-connected neighborhood structure, mining the geometric and semantic correlations between neighboring point pairs (Fig.3). Finally, by using the hierarchical encoder-decoder structure, the characteristics of multiple receptive field can be combined. A classification method based on the physical characteristics and geometric characteristics can be constructed to improve the classification ability of the model (Fig.4).

Results and Discussions The proposed method is based on the correlation between point clouds and full-waveforms on the physical characteristics of the target. By fusing the two, the physical meaning of the features is enriched and the learning of local geometric structures is enhanced. From Tab.2, it can be seen that the proposed method successfully generated correct labels for most points, achieving an average accuracy of 0.96, a recall of 0.90, and an $F1$ score of 0.92. Multiple methods such as FCN, GACNN, and FWNet2 are used for testing on the same dataset. Compared with FWNet2, which has the best performance among existing methods, our method has effective improvements in average accuracy, average recall, and average $F1$ score. Among them, compared with FWNet2, the accuracy, recall, and $F1$ score of our method have been improved by 0.02, 0.01, and 0.02 respectively in the ground category, and by 0.02, 0.04, and 0.03 respectively in the street category. The test results are shown (Fig.6).

Conclusions This article proposes a classification algorithm based on the physical and geometric characteristics of the target. This method focuses on the rich physical characteristics and vertical structural information of the target contained in the full-waveform, as well as the correlation between full-waveforms and geometric features of point clouds. The fusion of the two features is achieved using a dual low-rank matrix; And based on the local neighborhood fully-connected structure, the learning of local structural information are strengthened, thus constructing an end-to-end full-waveform airborne LiDAR data classification network. Multiple experiments show that the average accuracy and recall of our classification method are as high as 0.96 and 0.90, respectively, indicating the effectiveness of this method. At the same time, it provides some possibilities for exploration in-depth feature extraction and fusion, based on the correlation in target physical characteristics between full-waveforms and point clouds.

Key words: airborne LiDAR; data classification; feature fusion; full-waveform; 3D point clouds; physical characteristics

Funding projects: Guangxi Innovative Development Grand Grant (2018AA13005)