

基于球面投影的激光点云目标检测<sup>\*</sup>李加定<sup>1</sup> 万若楠<sup>1</sup> 孙小广<sup>1</sup> 邓磊<sup>2</sup>

(1. 广州城市理工学院电子信息工程学院 广州 510800; 2. 中南大学计算机学院 长沙 410083)

**摘要:** 基于深度学习的激光点云目标检测已成为了一个重要的研究领域。本文采用球面投影和 2D 图像的 SOTA 深度学习网络,实现 3D 激光点云目标快速检测。首先,将 KITTI 数据集单帧 3D 点云经球面投影转换成一帧 2D 的 RGB 三通道图像,像平面的像素位置取决于点云的三维坐标,其 R、G、B 3 个通道灰度值取决于点云归一化后的反射强度、距离、高度。其次,分析了不同分辨率下球面投影的重叠分布情况和对图像质量的技术影响。最后,采用语义分割模型 DeepLab-V3+网络,仿真结果表明:该方法在分割精确度和速度方面都具有良好的性能,应用价值较高。

**关键词:** 激光雷达;激光点云;三维目标检测;球面投影

**中图分类号:** TN958.98;TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.8040

Point cloud data processing and target recognition based on  
spherical projectionLi Jiading<sup>1</sup> Wan Ruonan<sup>1</sup> Sun Xiaoguang<sup>1</sup> Deng Lei<sup>2</sup>

(1. School of Electronic and Information Engineering, Guangzhou City University of Technology, Guangzhou 510800, China;

2. School of Computer Science and Engineering, Central South University, Changsha 410083, China)

**Abstract:** Laser point cloud target detection based on deep learning has become an important research field. This article uses a SOTA deep learning network based on spherical projection and 2D images to achieve rapid detection of 3D laser point cloud targets. Firstly, a single frame 3D point cloud from the Semantic KITTI data set is transformed into a 2D RGB three channel image through spherical projection. The pixel position of the image plane depends on the three-dimensional coordinates of the point cloud, and the grayscale values of the R, G, and B channels depend on the normalized reflection intensity, distance, and height of the point cloud. Secondly, the overlapping distribution of spherical projections at different resolutions and their technical impact on image quality were analyzed. Finally, using the semantic segmentation model DeepLab-V3 + network, simulation results show that this method has good performance in segmentation accuracy and speed, and has high application value. This paper presents a method of license plate character recognition based on the combination of Zernike moment and wavelet transformation features.

**Keywords:** LiDAR; point cloud; 3D object detection; spherical projection

## 0 引言

随着深度学习技术的飞速发展,激光雷达点云实时目标检测已成为自动驾驶和机器人领域的研究热点<sup>[1-2]</sup>。点云数据作为三维空间中物体表面信息的精确描述,近年来,国内外研究者积极探索了多种点云处理方法,以期在实时性和准确性之间达到平衡。其中,将 3D 点云转换为 2D 图像并利用深度学习进行处理的策略,因其能够充分利用成熟的图像处理技术和网络模型,受到了广泛关注。

点云目标检测常见方法有:基于原始点云、基于体素、

基于投影视图。其中基于原始点云的处理方法有 PointNet 及改进网络 PointNet++<sup>[3]</sup>,以及升级的 PointRCNN<sup>[4]</sup>和 3DSSD<sup>[5]</sup>,能直接处理点云数据,但计算量大,效率不高。基于体素的方法如 VoxelNet<sup>[6]</sup>和 SE-SSD<sup>[7]</sup>等,将点云数据划分为一系列小的立方体或柱体来有序存储与降采样,虽然特征提取高效,但存在信息丢失,高分辨率时占用内存大,运算效率受影响。相比之下,基于视图方法中的球面投影具有点分布均衡、信息损失率低以及图像可辨识度高等优势,特别是随着多模态点云处理技术的发展<sup>[8-9]</sup>,点云投影视图与相机 RGB 图像的融合成为研究热点,高质量的投

影视图为其必备条件<sup>[10-11]</sup>。然而,在投影过程不可避免出现多个激光扫描点映射到同一像素位置的情况,这对于后续图像分割和检测任务构成挑战<sup>[12]</sup>。因此,如何合理确定这类像素的灰度值,减少信息损失提高图像质量,降低冗余信息提升检测效率成为点云数据处理中具有挑战的重要环节。

为此,本文提出了一种基于球面投影和 2D 图像深度学习策略的 3D 激光点云目标检测方法。将 3D 激光点云经球面投影转换成 2D 图像,并详细分析了投影重叠的比例和分布情况。针对重叠像素的灰度值确定问题,提出了多种方案并进行了实验对比,最终选择了最优方案以生成高质量的图像和标签文件。在此基础上,利用 DeepLab-V3+ 网络<sup>[13]</sup>进行训练和分割检测,实现了高精度的点云目标检测。

## 1 基于球面投影的点云处理方案

### 1.1 激光点云的基本特征

以 64 线车载激光雷达<sup>[14]</sup>为例,每个激光器均配备发射器和接收器,并以固定角度定向发出激光。激光器旋转一周生成一帧点云,每个扫描点均详细记录了反射点的坐标、距离及反射强度等众多信息。激光器在垂直方向的视场由上到下为 $[Fov\_Up, Fov\_Down]$ ,水平面的视场为 $[0^\circ, 360^\circ]$ ,如图 1 所示。

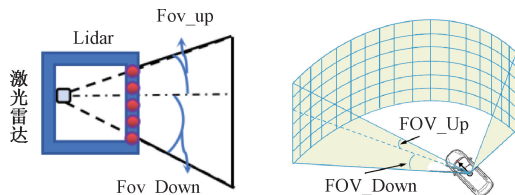


图 1 点云扫描过程示意图

激光雷达扫描一周形成一帧点云数据,坐标中心为激光雷达,各扫描点到激光器的距离和方位角都不同,投影到环球面上,展开的二维图像如图 2 所示。

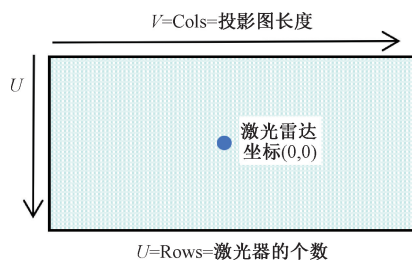


图 2 点云投影展开示意图

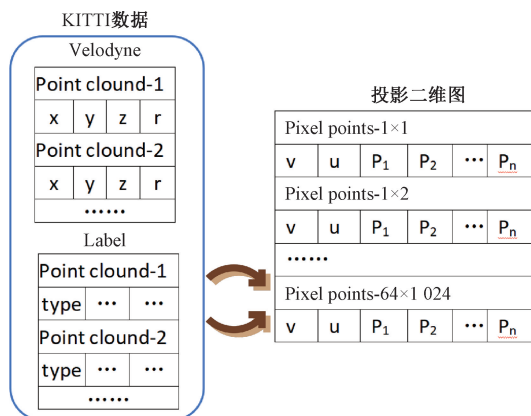


图 3 数据结构的变换

### 1.2 点云球面投影技术方案

如图 3 所示,以 KITTI 数据集为例,有两个文档 Velodyne 和 Label。其中, $r$  为反射强度,标签值  $type \in (0,100)$ ,共 15 个不同类别。扫描点  $P(x,y,z)$  投影后对应二维视图中的坐标值 $(u,v)$ ,且存在多个  $P$  点投影换算得到相同像素坐标值 $(u,v)$ ,需要将这些重叠点的反射强度  $r$  和标签值  $type$  按照一定规则换算成为像素点 $(u,v)$  的灰度值  $I$  和标签值  $G$ ,获得带标签的图像,构成神经网络的训练数据,如图 4 所示。投影运算包括坐标变换和信息转换,其中信息转换包括点云的反射强度  $r$ 、标签值  $type$ ,以及图 4 中未列出的距离  $R$  和高度  $z$  等信息。

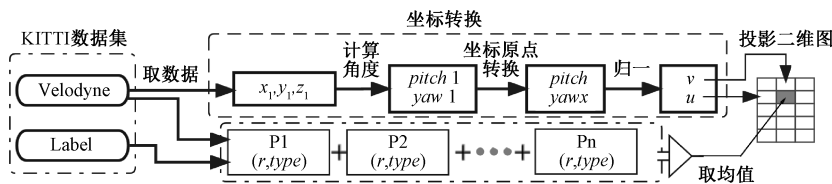


图 4 球面投影运算过程

### 1.3 3D 点云投影到 2D 平面的坐标转换

以一个扫描点  $Point(x_1, y_1, z_1)$  投影变换到二维图像的某个像素点 $(u_1, v_1)$  为例,如图 5 所示。该扫描点在三维坐标中的偏航角(yaw)和俯仰角(pitch)的值来表示该点对应 2D 像平面的坐标位置 $(u, v)$ 。 $x$  轴为车辆前进方向,图 5 右侧图是二维图数据结构保存形式示意图, $n$  是为了保存重叠扫描点的信息,避免投影降维过程信息丢失。

图 5 中①-计算点云中的点  $Point(x_1, y_1, z_1)$  在三维直角坐标系中的俯仰角(pitch1)和偏航角(yaw1):

$$pitch1 = \sin^{-1}(z_1/R) = \arcsin(z_1/R) \quad (1)$$

$$yaw1 = \tan^{-1}(y_1/x_1) = \arctan(y_1/x_1)$$

其中,  $yaw$  可换成另外一种形式:

$$yaw1 = \tan^{-1}(y_1/x_1) = \arcsin(y_1/\sqrt{x_1^2 + y_1^2}) \quad (2)$$

由 3D 点云数据计算得到的 $(yaw1, pitch1)$ 将构成 2D

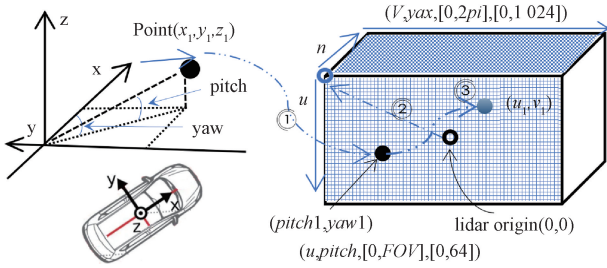


图5 点云球面投影过程示意图

图的坐标值。

图5中②-坐标原点变换:

一般来说,图像的坐标原点位于图像左上角,而前面得到的 $(yaw, pitch)$ 的坐标原点在投影图的中心位置,故需要进行坐标原点变换,假设变换前的坐标为 $(yaw_1, pitch_1)$ ,变换后的坐标为 $(yaw, pitch)$ ,变换关系为:

$$\begin{aligned} pitch &= FOV\_Up - pitch_1 \\ yaw &= yaw_0 + \pi \end{aligned} \quad (3)$$

偏转角  $yaw$  如果由 KITTI 数据集本身提供的偏航角度  $yaw_0 \in (-\pi, \pi)$  转成  $yaw \in (0, 2\pi)$  则由上式得出。本文中未采用  $yaw_0$ , 是由  $(x_1, y_1, z_1)$  坐标值转换而来, 由于  $\arctan(y/x) \in (-\frac{\pi}{2}, \frac{\pi}{2})$ , 目标  $yaw \in (0, 2\pi)$ , 需考虑  $(x, y)$  所在象限区间, 作如下转换:

$$yaw = \begin{cases} 0 - \arctan(y/x), \\ \pi - \arctan(y/x), \\ \pi - \arctan(y/x), \\ 2\pi - \arctan(y/x), \end{cases} \quad (4)$$

将  $Point(x_1, y_1, z_1)$  代入上式即可获得。

图5中③-图像规范化和缩放:

不同类型的激光雷达投影图像的大小不同,必要的缩放与规范化处理可以让投影图尽可能适应更多的激光扫描点。

1) 二维投影图像中  $u$  值的计算

投影图像的行数或竖直方向上像素点个数应等于激光雷达在竖直方向激光器的个数,记为  $row\_scale$ 。从而保障每个激光器产生的扫描点投影后在二维图竖直方向不重叠。投影图像素点在竖直方向坐标值以俯仰角 ( $pitch$ ) 来划分,并将其归一化。

$$normalized\_pitch = \frac{Fov\_Up - pitch_1}{Fov\_Up - Fov\_Down} \quad (5)$$

由于  $Fov\_Down < 0$ ,

故:  $Fov = Fov\_Up + \text{abs}(Fov\_Down)$

将式(1)和(5)结合,得到扫描点投影后在二维图像中竖直方向的坐标值  $u$ 。

$$u = \frac{Fov\_Up - \arcsin \frac{z}{R} \times 180/\pi}{Fov} \times row\_scale \quad (6)$$

2) 投影二维图像中  $v$  值的计算

图像长度 ( $col\_scale$ ) 的设定跟激光器旋转一周会产生

多少个扫描点有关,分辨角  $\Delta\theta = 2\pi/col\_scale$ , 扫描一周形成的点越多图像分辨率越高。归一化后的偏航角  $yaw$  与图像的列数 ( $col\_scale$ ) 相乘就得到了图像尺寸  $v$  方向的值:

$$v = \frac{yaw}{2\pi} \times col\_scale \quad (7)$$

借助式(6)和(7),可将点云在三维直角坐标位置参数  $(x_1, y_1, z_1)$  换算成球面投影图像中的像素点坐标位置  $(u_1, v_1)$ 。而这一点的其他信息如反射强度  $r$  和标签值  $G$  也随之迁移到对应的像素点上。

#### 1.4 点云球面投影重叠信息融合

球面投影过程中,由于存在多个 3D 扫描点对应同一个 2D 像素点的情况,这些扫描点的信息如反射强度  $r$ 、距离  $R$ 、高度  $z$  等信息在必要时需要合并成同一个像素点对应的值。

投影过程可定义为一个函数,输入是点云  $P_k$ , 包括坐标值  $(x, y, z)$ 、反射强度  $r$ , 标签  $type$ , 输出为图像  $T_{(u,v)}$ , 若某个像素点  $[u, v]$  对应多个扫描点,需要将对应式(8)保存下来以备后面反向映射回点云数据。

$$\begin{cases} P_k = (x, y, z, r, type) \\ T_{(u,v)} = [u, v, P_1, P_2 \dots P_n] \end{cases} \quad (8)$$

以像素点  $[u, v]$  的灰度值  $I$  由反射强度  $r$  来定义为例,融合方案可以为重叠扫描点反射强度  $r$  的平均值、最大值、距离最近扫描点对应值等方案,并择优得到出最佳方案。

## 2 投影 2D 图像分析

为评估本文设计的点云处理方案的性能,使用 KITTI 标准数据集。从数据集导入数据为  $(x, y, z, r, type)$ , 其中  $x \in (-80, 80)$ ,  $y \in (-55, 52)$ ,  $z \in (-11.3, 2.9)$ 。对于 Velodyne 公司 HDL 64-E 型激光雷达 (high-definition logic 64-bit evolutionary, HDL 64-E), 激光器在垂直方向的视场参考值  $Fov\_Up = 2.5^\circ$ ;  $Fov\_Down = -24.8^\circ$ , 由此而来的俯仰角  $pitch \in (-0.0349, 0.4328)$  rad, 偏转角  $yaw \in (0, 2\pi)$  rad。HDL 64-E 的竖直角分辨率为  $0.4^\circ$ , 水平角分辨率为  $0.08^\circ \sim 0.35^\circ$ 。因此,在最坏的情况下,至少可以得到(竖直方向  $25.2/0.4 = 68$ 、水平方向  $360/0.35 = 1028$ )  $64 \times 1024$  个点。在计算机视觉中,图像大小习惯设置为 2 的倍数,取投影图  $col\_scale = 64$ ,  $row\_scale = 1024$  或  $2048$ 。

将点云融合后的反射强度  $r$  作为图像灰度值  $I$ , 其中一帧点云球面投影图如图 6 所示,与普通相机拍摄的图像有差别,但人眼可辨识其中物体的轮廓、形状、类别等信息,那么这种图像也能被训练好的深度学习网络给出良好的分割和分类结果。

#### 2.1 球面投影降维损失与重叠分析

以其中一帧点云为例,约 124 600 个点,由于  $124\ 600 > 64 \times 1024$ , 映射到 2D 图像必然出现两个以上的扫描点对



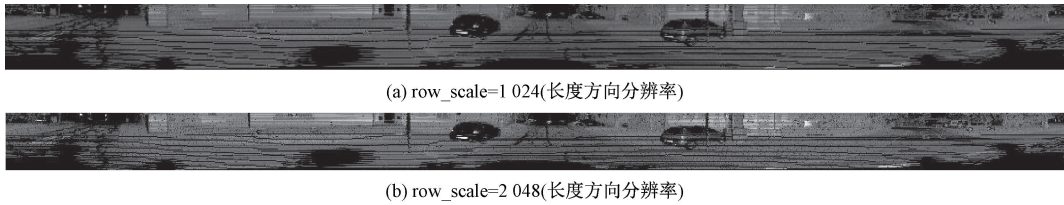


图 6 点云球面投影灰度图

应同一个像素点。计算表明:2D 图像  $64 \times 1\,024$  个像素点中有  $N=52\,225$  个像素点至少有  $n=1$  个激光扫描点,面积占比 79.69%, $n=2$  个点投影到 2D 图像的同一个像素点共有  $N=48\,967$  个,面积占比 74.72%,详细重叠情况如表 1 和重叠分布图 7 所示。

表 1 点云投影映射重叠情况

长度方向分辨率 1 024				长度方向分辨率 2 048			
$n$	$N$	面积占比%	数量占比%	$n$	$N$	面积占比%	数量占比%
1	52 225	79.69	41.9	1	99 794	76.14	80.09
2	48 967	74.72	39.3	2	21 333	16.28	17.12
3	14 094	21.51	11.3	3	3 072	2.34	2.46
4	6 401	9.77	5.1	4	374	0.29	0.30
5	2 100	3.20	1.6	5	3	0	0
6	608	0.93	0.5				
7	160	0.24	0.1				
8	37	0.06	0.03				

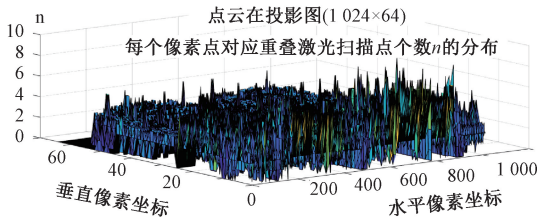


图 7 row\_scale=1 024 时扫描点重叠分布情况

由上述数据分析可知:

- 1) 扫描点在图像分布区域广且相对均衡,2D 图像能充分反映点云所携带的信息。
- 2) row\_scale=1 024 时,投影重叠比例高,集中在 2~3 个激光扫描点对应 1 个像素点。
- 3) row\_scale=2 048 时,投影重叠会得到明显的改善,但网络训练时间会呈几何级数增加。

2.2 投影二维图分辨率取值

在图 6 中,row\_scale=1 024 或 2 048 形成了两幅图,在未对图像滤波降噪的情况下,利用德国 idealo 服务商的开源图像质量评价工具<sup>[15]</sup>来分析对比,这里侧重于技术评价而非美学评价,即侧重图像质量。结果为 row\_scale=2 048 得分为 3.434,而 row\_scale=1 024 得分为 3.686。

另外考虑到不增加网络训练时间,二维图水平分辨率取 1 024 为最佳选择。

2.3 投影图像素灰度值

如果使用反射强度  $r$  来区分投影图的灰度值,考虑到重叠,图像灰度值需按一定规则进行融合。假设点  $P_1$ 、 $P_2$ 、…… $P_n$  投影到同一个像素点,每个点都有一个反射强度  $r$ ,确定像素点灰度值  $I$  的常见方法有:

1) 取平均值

$$I = \frac{\text{average}[r_{(u,v)}]}{|\max(r) - \min(r)|} \times 255 \quad (9)$$

2) 取最大值

$$I = \frac{\max[r_{(u,v)}]}{|\max(r) - \min(r)|} \times 255 \quad (10)$$

3) 只取距离最近  $R_{\min}$  点

$$I = \frac{r[R_{\min(u,v)}]}{|\max(r) - \min(r)|} \times 255 \quad (11)$$

其中,  $|\max(r) - \min(r)|$  是一帧点云所有点的反射强度  $r$ ,  $r_{(u,v)}$  是像素点  $(u,v)$  对应的扫描点反射强度组成的一组数,  $R_{\min(u,v)}$  是指像素点  $(u,v)$  对应的扫描点中与激光雷达距离最近的一点。3 种方式图像效果对比如图 8 所示。

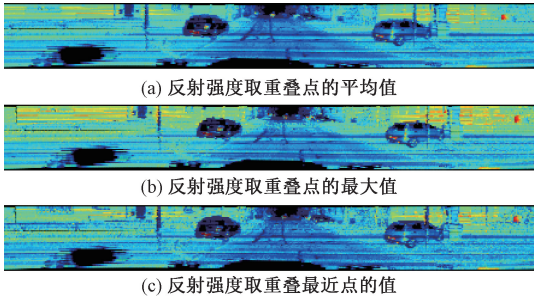


图 8 反射强度取不同值为像素灰度值的图像对比

若人眼识别,3 幅图相差不大,取平均值的图像稍微平滑,轮廓边沿更清晰,也更容易被网络所识别。表 2 中的图像质量分值更加定量地反映取均值是最优选择。

表 2 不同反射强度的图像质量对比

平均值	最大值	最近点
2.562 1	2.437 9	2.442 7

2.4 球面投影 RGB 彩色图

前面将反射强度  $r$  作为像素灰度值得到了投影后的灰度图,但是 Deeplab v3+神经网络输入为 RGB 图,另外



彩色图像也更利于人眼审美和特征识别,有利于提高语义分割精度。将距离  $R$ 、高度  $z$ 、反射强度  $r$  分别对应 RGB

颜色 3 个通道,得到如图 9 所示效果,与普通相机图像相比,还保留了深度和高度信息。

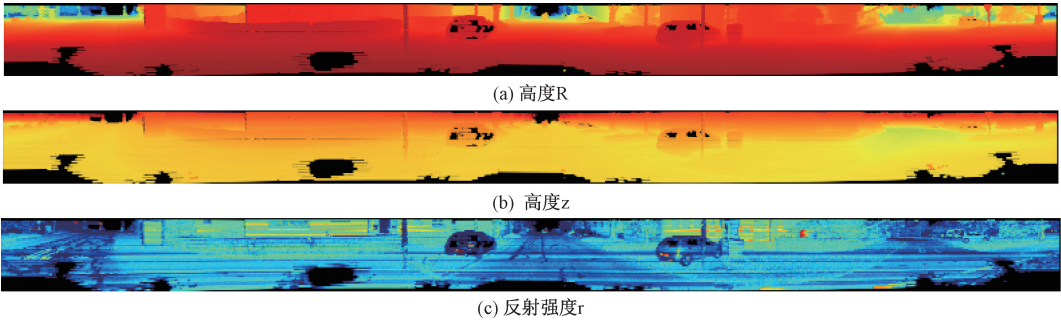


图 9 点云球面投影 RGB 三通道图

2.5 投影图的标签

考虑到投影点重叠,从原始点云  $type$  值到投影图像素级标签值  $G$  的转换,取重复出现高的  $type$  值作为  $G$  值,同频率则取距离最近点的  $type$  值,这样得到整个投影图



图 10 根据标签值分颜色投影图

总体而言,点云球面投影成 2D 图像,由数据可见点分布均衡、信息损失率低、图像可辨识度高。当  $row\_scale = 1\ 024$ ,重叠点云反射强度的平均值经过归一化作为 2D 图像灰度值时,图像质量最佳。

3 投影图语义分割网络训练与评估

3.1 语义分割网络

KITTI 数据集对点云中的每个点进行语义分类,赋予标签值,共有 15 个类别如 Car、Tree、Pedestrian 等。采用图像分割的卷积神经网络 (convolutional neural network, CNN) 为 Deeplab v3+,该网络具有从预训练的 Resnet-18 网络初始化的权重,在对 CamVid 图像数据集训练和识别

的标签值。图 10 是根据不同  $G$  值以不同颜色呈现的区域分类图像效果。可知,能较好区分地面、汽车等物体,这种标签值融合的方式是可行的。

中有良好的效果<sup>[16]</sup>。在使用中,将前面球面投影得到的二维图和标签数据 (如图 9 和 10 所示) 来替代 CamVid 数据集。为了使训练更容易,将原 CamVid 中的 32 个原始分类组为 11 个类,指定这些类,并与点云的 15 个类建立对应关系。准备训练集、验证集和测试集,Matlab 中,在 Resnet-18 基础上创建网络,迁移数据,平衡类的权重,选择具有动量的随机梯度下降 (stochastic gradient descent with momentum, SGDM) 的优化算法进行训练,最后检测训练效果。

3.2 网络训练结果评估

评估经过训练的网络,衡量多个测试图像的准确度,对整个测试集运行 Semanticseg 检验函数,返回基于测试集结果计算各个语义分割指标<sup>[17]</sup>,如表 3 所示。

表 3 测试集语义分割各项整体指标

标签/类别	CamVid			球面投影图		
	Accuracy	IoU	Mean BFScore	Accuracy	IoU	Mean BFScore
Sky	0.931	0.902	0.895	0.906	0.866	0.851
Building	0.785	0.761	0.585	0.746	0.749	0.572
Pole	0.716	0.215	0.514	0.704	0.099	0.502
Road	0.930	0.915	0.767	0.871	0.925	0.75
Pavement	0.885	0.706	0.709	0.873	0.663	0.681
Tree	0.874	0.763	0.708	0.82	0.699	0.672
SignSymbol	0.794	0.393	0.483	0.696	0.381	0.073
Fence	0.815	0.465	0.486	0.779	0.353	0.381
Car	0.910	0.768	0.693	0.861	0.745	0.548
Pedestrian	0.876	0.437	0.608	0.841	0.414	0.594
Bicyclist	0.878	0.608	0.551	0.854	0.474	0.451

可知,本文球面投影整体效果还是令人满意的。具体类别中 Road、Sky 和 Building 类具有较高的 IoU 分值,而 Pedestrian 和 Bicyclist 等类的分数较低。

### 3.3 基于投影图分割的点云

前面语义分割结果是基于投影图,还需要将结果通过式(8)对应关系反向映射回点云。仍然以前面分析的一帧点云为例,获得类别标识值的点为  $N=51\ 955$ ,占原始点云数量  $N=124\ 600$  的 41.7%,两者分割<sup>[18]</sup>效果如图 11 所示。经过投影识别后的点变稀疏,但主要轮廓特征都保留了,表征信息量未明显减少,依然能被较好地分割识别,必将减少后续决策待处理的点云数量,提高运行速度,具有实用价值。

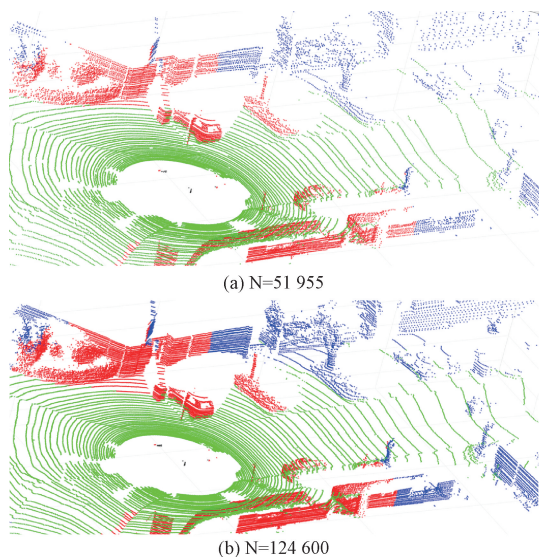


图 11 识别后的点云与原始点云成像对比

经过前面实验分析可见,将 3D 点云球面投影成 2D 图像再运用深度学习网络模型进行目标识别的方案是可行的。识别结果接近相机的 RGB 图像识别效果,且保留了点云的距离和高度尺度信息。另外球面投影能大幅减少了目标识别时的冗余信息,提升处理速度。

## 4 结 论

将 3D 激光点云通过球面投影转换为 2D 图像,是一种高效实用的点云处理方式,其中将点云的反射强度作为像素参考值,得到的 RGB 图像中物体特征分辨最明显,对于投影重叠情况,像素值取重叠点云的反射强度均值得到的图像图像质量最佳。一帧点云投影成像,分辨率选  $64 \times 1\ 024$  像素的图像质量更为理想。由此建构的数据集,在 DeepLab-V3+ 语义分割网络模型的运行表现出良好的分割精确度和速度优势。这种技术方案对于点云的二维视图处理、多视图融合以及与相机图像的融合处理等方面均有参考意义。

## 参考文献

- [1] 秦静,王伟滨,邹启杰,等. 基于激光雷达点云的 3D 目标检测算法综述[J]. 计算机科学, 2023, 50(S1): 259-265.
- [2] KUMAR P V, DOGUS K, EGILS A, et al. A survey on deep learning based segmentation, detection and classification for 3D point clouds[J]. Entropy(Basel, Switzerland), 2023, 25(4): 635.
- [3] LIU Z, ZHAO X, HUANG T, et al. TANet: Robust 3D object detection from point clouds with triple attention[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 11677-11684.
- [4] 王庆林,李辉,谢礼志,等. 基于激光雷达点云的车辆目标检测算法改进研究[J]. 电子测量技术, 2023, 46(1): 120-126.
- [5] YANG Z, SUN Y, LIU S, et al. 3DSSD: Point-based 3d single stage object detector[J]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020: 11037-11045.
- [6] XIE J W, ZHONG Z L, GAO R Q, et al. Generative voxelNet: Learning energy-based models for 3D shape synthesis and analysis[J]. IEEE Trans Pattern Anal Mach Intell, 2022, 44(5): 2468-2484.
- [7] ZHAO H Q, ZHANG T S. Fabric surface defect detection using SE-SSDNet [J]. Symmetry, 2022, 14(11): 2373-2373.
- [8] YANG Y G, CHEN F, WU F, et al. Multi-view semantic learning network for point cloud based 3D object detection [J]. Neurocomputing, 2020, 397: 477-485.
- [9] 吴文涛,何赞泽,杜旭,等. 融合相机与激光雷达的目标检测与尺寸测量[J]. 电子测量与仪器学报, 2023, 37(6): 169-177.
- [10] 李研芳,黄影平. 基于激光雷达和相机融合的目标检测[J]. 电子测量技术, 2021, 44(5): 112-117.
- [11] DENG J, ZHOU W, ZHANG Y, et al. From Multi-View to Hollow-3D: Hallucinated Hollow-3D R-CNN for 3D object detection [J]. IEEE Transactions on

- Circuits and Systems for Video Technology, 2021, 31(12):4722-4734.
- [12] WANG L, FAN X, CHEN J, et al. 3D object detection based on sparse convolution neural network and feature fusion for autonomous driving in smart cities [J]. Sustainable Cities and Society, 2020, 54:102002.
- [13] SUNNETCI K M, KABA E, FATMA B C, et al. Comparative parotid gland segmentation by using ResNet-18 and MobileNetV2 based DeepLab v3 + architectures from magnetic resonance images [J]. Concurrency and Computation: Practice and Experience, 2022, 35(1):7405.
- [14] 周治国, 邸顺帆, 冯新. 语义信息增强的 3D 激光 SLAM 技术进展[J]. 仪器仪表学报, 2023, 44(3):209-220.
- [15] HOSSEIN T, PEYMAN M. NIMA: Neural image assessment [J]. IEEE transactions on image processing: a publication of the IEEE Signal Processing Society, 2018, 27(8):3998-4011.
- [16] 李亚舟, 曹江涛, 姬晓飞. 融合改进 Padim 建模和 ResNet 网络的喷涂质量检测算法[J]. 电子测量与仪器学报, 2022, 36(11):91-97.
- [17] WU J Y, HUANG S H, YANG Y X, et al. Evaluation of 3D LiDAR SLAM algorithms based on the KITTI dataset[J]. The Journal of Supercomputing, 2023, 79(14):15760-15772.
- [18] WU T, FU HAO, LIU B K, et al. Detailed analysis on generating the range image for LiDAR point cloud processing[J]. Electronics, 2021, 10(11):1224-1224.

### 作者简介

**李加定**, 硕士, 主要研究方向为计算机视觉、机器人控制。

E-mail: lijid@gcu.edu.cn

**邓磊**(通信作者), 教授, 博士生导师, 主要研究方向为主要研究人工智能与生物信息学, 大数据挖掘、生物计算。

E-mail: leideng@csu.edu.cn