

DOI: 10.13382/j.jemi.B2205143

基于椭球拟合的点云旋转不变网络*

黎书玉 张旭东 孙锐 范之国

(合肥工业大学计算机与信息学院 合肥 230601)

摘要:点云携带丰富的几何信息,在计算机视觉领域具有独特优势。现有基于深度学习的三维模型分类与分割方法能有效识别固定视角下的物体,但在实际应用中,物体方向未知,使得点云描述存在旋转变换问题,极大影响网络的识别精度。针对点云的旋转性问题,提出一种轻量级的基于椭球拟合的旋转不变网络(point cloud rotation invariant network based on ellipsoid fitting, EFRI-N)。设计前置网络模块提取点云的旋转不变特征,包括椭球拟合和特征编码两个部分。通过椭球拟合算法标识原始点云的方向得到旋转不变坐标系,再将原始特征映射到该坐标系中,利用空间信息和角度信息进行编码得到点云的旋转不变特征;为了获取更丰富的几何信息,在分类分割网络中加入多层级的特征连接增强特征传播及复用,提高模型表征能力。采用国际知名公共数据集 ModelNet40 和 ShapeNet Parts 进行分类、分割实验,结果表明,该方法在处理旋转点云的任务中优于主流算法,网络识别精度提升了 1%~62.63%不等,并且网络的计算量和参数量都有着数量级的优势。满足单目标场景下对点云旋转不变性的使用要求,具有良好的应用价值。

关键词:点云;深度学习;分类;分割;旋转不变;轻量级

中图分类号: TN91; TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Point cloud rotation invariant network based on ellipsoid fitting

Li Shuyu Zhang Xudong Sun Rui Fan Zhiguo

(School of Computer and Information, Hefei University of Technology, Hefei 230601, China)

Abstract: Point clouds have unique advantages due to its rich geometric information in computer vision field. Most of the existing point cloud classification and segmentation methods based on deep learning can identify the objects with canonical orientations. In real applications, there are problems of rotation transformation. In this paper, we propose a lightweight framework EFRI-N, namely, rotation invariant network of point cloud based on ellipsoid fitting, focusing on pointset rotation problems. We design a pre-network module to extract the rotation-invariant features. The ellipsoid fitting algorithm is used to identify the direction of the point clouds and obtain the rotation-invariant coordinate. Then the original features are mapped to the coordinate, and the rotation-invariant features were obtained by encoding the spatial and angular information. In order to obtain richer geometric information, multi-level feature connection is added to the network to enhance feature propagation and reuse. The classification and segmentation experiments are carried out by using the famous public datasets ModelNet40 and ShapeNet Parts. The results show that this method demonstrates better performance than state-of-the-art methods in the task of processing rotating point cloud, and the network is improved by 1%~62.63%. Moreover, the computation amount and the number of parameters of the network have an order of magnitude advantage. It can meet the requirements of rotation invariance of point cloud in single object scenario and has good application value.

Keywords: point clouds; deep learning; classification; segmentation; rotation invariant; lightweight

收稿日期: 2022-01-21 Received Date: 2022-01-21

* 基金项目: 国家自然科学基金(61876057, 61971177)、安徽省重点研发计划-科技强警专项(202004d07020012)、安徽省科技重大专项(202103a06020010)项目资助

0 引言

随着图形采集技术的发展^[1],3D 传感器成本越来越低。同时 3D 数据可以提供丰富的几何、形状、比例信息,这极大的促进了人们对 3D 数据的研究。近年来,点云作为描述三维模型的一种常用格式被广泛应用于计算机视觉、自动驾驶等领域^[2],这引起了人们越来越多的关注。

由于点云点数的不确定性及无序性,传统方法通常将点云转换为规则的数据表示,例如体素^[3]、多视图集合^[4-5],这类方法在转换过程中不可避免的造成了形状信息的丢失。区别于处理规则数据的方法,基于点的方法更为直观简洁。Pointnet 系列^[6-7]首次将深度学习技术应用到点云,它使用卷积层将原始坐标映射到高维空间,最后通过最大化池化操作减小点排列的影响并提取全局特征用于分类、分割任务。DGCNN^[8]提出边缘卷积方法,在语义上将点聚合起来并且嵌入已有的点云深度学习框架中,增强对局部特征的提取。SpiderCNN^[9]通过参数化一系列卷积滤波器,将卷积运算从规则数据扩展到不规则点集提取深层语义特征。此类方法在点云学习和形状分析方面取得了显著的成果,但他们默认处理固定视角下的三维点云,在实际应用中,模型方向通常先验未知,所以这些方法难以应用于旋转形状分析。为了解决点云旋转性问题,Stein 等^[10]、Tombari 等^[11]提出通过构造局部参考系将三维模型对齐到正则姿态中提取旋转不变点特征,然而这些方法依赖于点云局部表面的变换,对噪声和点密度很敏感。SO-Net^[12]通过构建自组织映射网络来模拟点云的空间分布,将网络学习到的空间信息和角度信息映射到一组初始节点上替代原始输入,一定程度上提升了旋转不变性,但效果取决于网络的学习能力,难以实现稳定的置换不变性。Spherical CNN^[13]提出一种具有旋转不变的球面卷积操作从连续网格面提取全局特征,虽然验证了球面卷积对网络的适用性,但它只能应用于结构化数据上,无法直接处理不规则点云。PRIN^[14]使用密度感知自适应采样将稀疏点信号转换为球形信号,随后采用球形体素卷积提取每个点的近似旋转不变特征,在一定程度上提高了网络对方向的鲁棒性,但也增加了网络复杂度。

针对现有框架存在的问题,本文提出基于椭球拟合的点云旋转不变网络(EFRI-N)。通过前置网络中的椭球拟合算法为每个形状拟合最佳椭球面,利用椭球面的中心点及半轴方向信息对原始输入进行编码,得到具有强旋转不变性的特征。为了充分利用底层信息,在网络中加入多层级的特征连接增强特征复用并结合全连接层以及池化层学得高维特征用于三维点云的分类和分割任务。本文主要贡献分为如下两点:

1) 提出旋转不变特征提取前置网络,直接处理点云数据并通过椭球拟合算法为每个原始输入确定唯一的方向描述符,经过特征编码得到旋转不变特征 EFRI。相比现有方法,EFRI 具有更强的旋转不变性和通用性。

2) 提出轻量级的旋转不变点云分类分割网络 EFRI-N,使用直观简洁的特征编码方法替代传统算法中复杂的对齐网络模块以及冗余性较高的编码方法,将提取到的旋转不变特征作为网络输入,最后通过多层级的特征连接获取更丰富的几何信息。在增强网络对旋转物体鲁棒性的同时降低网络复杂度。

1 本文方法

点云本质上是一长串点($n \times 3$ 矩阵,其中 n 是点数)。在几何上,点的顺序不影响它在空间中的整体形状,例如相同的点云可以由两个完全不同的矩阵表示,这给基于深度学习处理点云的方法带来很大影响,为了克服点云旋转性问题,本文提出了 EFRI-N,其总结构如图 1 所示。网络先将不同点数和尺度的原始点云采样、归一化进一个单位球内,之后通过前置网络模块中的椭球拟合算法唯一标识物体方向并构造旋转不变坐标系,利用点-轴相对位置不变的特性进一步编码提取稳定的旋转不变特征,最后通过全连接层和多层级连接操作充分学习底层信息实现点云的分类和分割任务。

1.1 采样和归一化

假设点云数据集为 $\mathbf{X} = \{\mathbf{X}_i, i = 1, \dots, p\}$, p 表示总共包含的点云数。每一个点云都用若干个来描述: $\mathbf{X}_i = \{x_{ij}, i = 1, \dots, p, j = 1, \dots, N\}$ 。由于每个物体的尺度以及包含的点数都不相同,同时也为了减小计算量,首先对每个点云进行降采样,与 PointNet 相同,采用最远距离采样法取 $N = 2048$ 个点作为最终的数据输入。然后将所有点云归一化进一个单位球内,归一化方法如式(1)所示:

$$x_{ij} = \frac{x_{ij} - E(\mathbf{X}_i)}{\max\{\|x_{ij} - E(\mathbf{X}_i)\|\}}, x_{ij} \in \mathbf{X}_i \quad (1)$$

其中, $E(\mathbf{X}_i)$ 为点云的中心点坐标。

1.2 旋转不变特征提取前置网络

为了描述物体方向,首先为每个输入形状拟合椭球面。传统方法通常在均匀分布的球体里提取点云特征,但球体具有极强的对称性,无法直接标识物体的朝向。而椭球面具有 3 个半轴长度不等的性质,有助于点云形状在三维坐标系中的方向定位。同时,最小二乘法^[15]是解决曲面拟合问题最常用的方法,因此本文提出将椭球拟合算法应用于点云的方向描述,通过最小二乘法计算目标函数和点云数据之间的最小均方误差找到每个形状匹配的最佳椭球面,无论物体如何旋转,椭球面与物体的

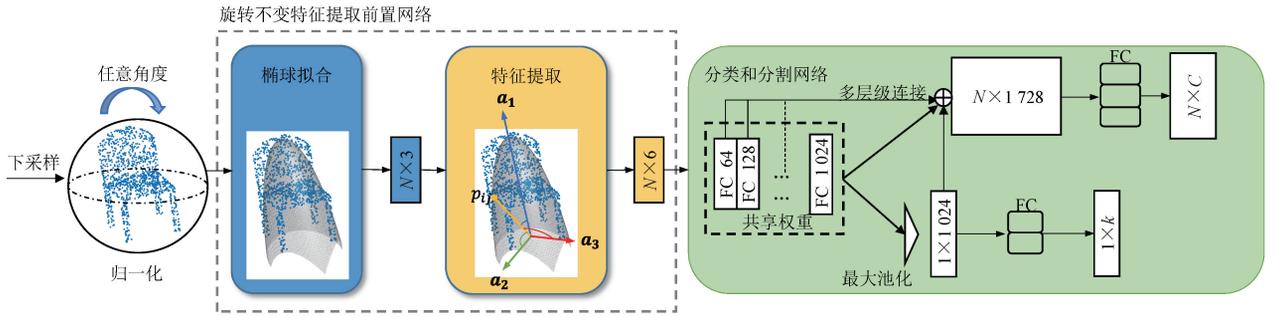


图 1 EFRI-N 整体框架

Fig. 1 EFRI-N overall framework

相对位置关系始终保持不变,这确保了椭球面的旋转不变性,为下一步旋转不变特征的提取奠定基础。通过计算得到椭球面中心点 O 和 3 个半轴方向 a_1, a_2, a_3 信息(相互正交且相交于点 O)。对于不同角度的输入,总能找到唯一一个最近似输入分布的椭球面,不论物体怎么旋转,点与所选轴之间的相对位置关系保持固定,可视化结果如图 2(a)所示。

在已知椭球面的基础上,引入特征编码模块。在点与轴相对位置关系保持不变的前提下利用空间信息和角度信息重新定义网络输入,从而得到形状的旋转不变特征。本方法克服了传统球面卷积方法计算冗余、网络复

杂的问题,能快速得到稳定的旋转不变特征,消除角度问题给特征表示带来的影响。

在由椭球面中心点和 3 个半轴方向构成的新坐标系中,首先利用相对位置 $Op_{ij} = (p_{ij} - O) \in R^3$ 更新每点的坐标信息,再提取每个点相对于 3 个轴的角度特征 $\theta_1, \theta_2, \theta_3$,编码过程及其结果如图 3(b)、(c)所示。对于第 i 个点云中的第 j 个点来说,它的特征表示具体如下:

$$f_{ij} : (Op_{ij}, \theta_1, \theta_2, \theta_3) \in R^6 \quad (2)$$

则第 i 个点云的特征可以表示为:

$$F_i = \{f_{ij}, j = 1, \dots, N\} \in R^{N \times 6} \quad (3)$$

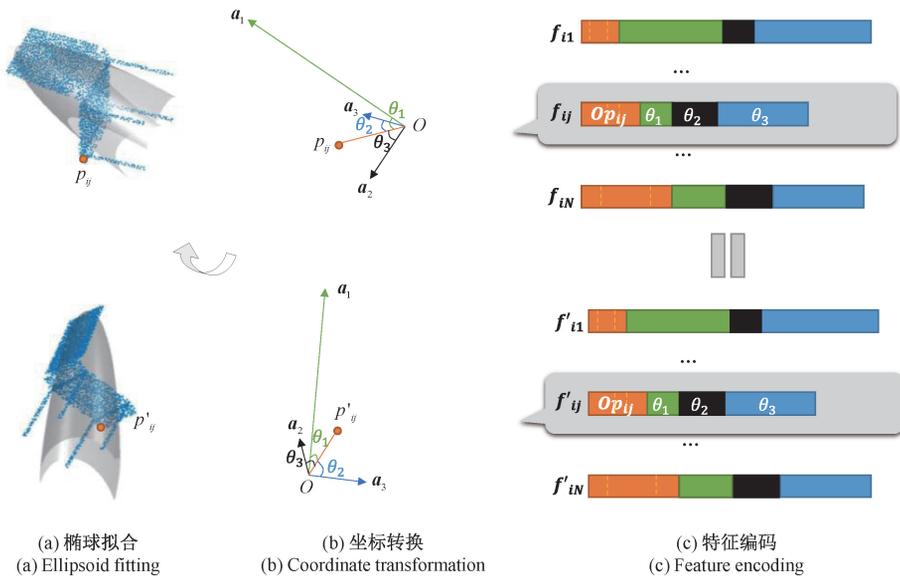


图 2 旋转不变特征提取示意图

Fig. 2 The process of rotation invariant feature extraction

1.3 轻量级的点云旋转不变分类分割网络

在旋转不变特征基础上,提出一种轻量级的点云旋转不变分类分割网络(EFRI-N),具体网络结构如图 3 所

示,使用通用结构 PointNet 作为骨干网络。与不同的是,本文设计的椭球拟合算法提取的特征本身存在强旋转不变性,因此去除了空间对齐网络 T-Net^[16] 模块,省去

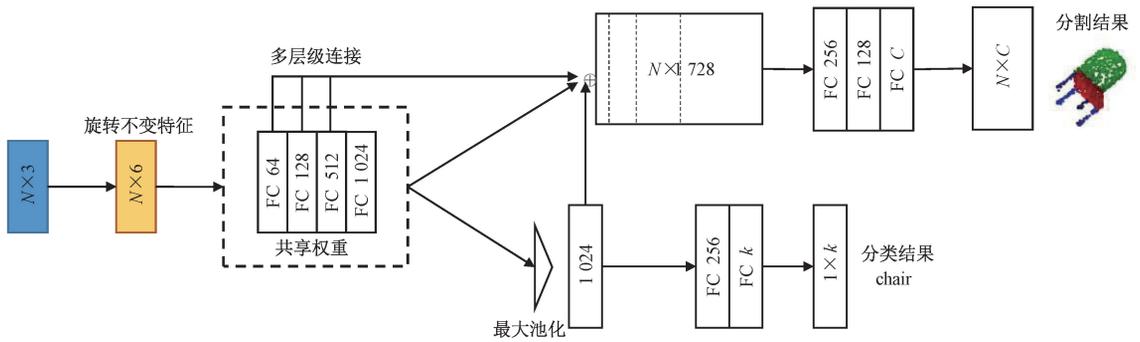


图 3 旋转不变分类分割网络

Fig. 3 Rotation invariant classification and segmentation network

了在输入层和特征层的对齐操作,大大降低了网络复杂度。另外采用多层级的特征连接充分利用底层信息以提高识别精度。

针对不同尺度和点数的原始数据,通过采样归一化方法得到输入为 $N \times 3$ 的点云矩阵;利用椭球拟合算法和特征提取模块对输入进行编码得到 $N \times 6$ 的旋转不变特征。在分类任务中,特征经过几个共享权重的全连接层和最大池化层学习全局特征,最终得到一个 $1 \times k$ 的类别分数;对于部件分割任务,利用多层级连接将低维特征和高维特征结合得到更具有表征力的几何特征,并配置全连接层和 ReLU 层提高模型的表达能力最后输出 $N \times C$ 的特征矩阵,表示点云中每个点对应的分割结果。

2 实验结果与分析

2.1 实验配置和评估指标

分别采用 ModelNet40^[17] 和 ShapeNet Parts^[18] 数据集对提出的旋转不变网络模型 EFRI-N 分类、分割有效性进行验证。实验使用的 PC 机硬件配置为 Intel (R) Core (TM) i7-6850K CPU @ 3.60 GHz, GPU GEFORCE GTX TITAN, 软件环境由 Ubuntu 16.04, cuda9.0 和 cudnn 7.05 组成,深度学习框架基于 Pytorch 1.6 框架和 Python 3.6 实现。EFRI-N 采用共享权重的全连接层,批处理大小为 32,使用 Adam^[19] 优化器训练。

使用准确率作为分类和分割实验的评价指标。假设有 m 个测试样本,网络对 n 个测试样本作出了正确的判断,则准确率表示为:

$$Acc = \frac{n}{m} \times 100\% \quad (4)$$

使用浮点运算数 FLOPs 衡量算法的时间复杂度,全连接层的 FLOPs 计算方法为:

$$FLOPs = [I + (I - 1) + 1] \times O \quad (5)$$

其中, I 表示输入神经元个数, O 表示输出神经元个

数,括号中 I 代表乘法运算量, $I - 1$ 代表加法运算量, $+1$ 表示考虑偏置。使用参数量衡量算法的空间复杂度即计算机内存消耗程度,全连接层的参数量计算方法为:

$$parameter = (N_{in} + 1) \times N_{out} \quad (6)$$

式中: N_{in} 为输入特征通道数, N_{out} 为输出特征通道数。

2.2 分类实验

实验选取斯坦福大学提供的点云基准数据集 ModelNet40 来测试 EFRI-N 的分类性能。该数据集包含 12 308 个形状,共有 40 个类别,其中 9 840 个形状用于训练,其余用于测试,每个点云数据都由三维坐标信息和标签信息组成。

选取 PointNet 系列、Kd-Network^[20]、Spherical CNN 等主流算法与本文方法作比较,实验结果如表 1 所示,第 3 列 NR/NR 表示在训练和测试中都使用不经过任何处理的点云数据,第 4 列 NR/AR 表示使用原始数据进行训练,使用旋转任意角度的数据进行测试,最终对于旋转测试集,本文方法获得了最高精度,并且大幅度优于其他方法。对比骨干网络 PointNet,本方法总体识别准确率提升了 58.65%。其中 PRIN 对旋转物体也表现出较好的鲁棒性,但综合性能仍差于本文方法。EFRI-N 能表现出较好的性能,原因在于前置网络模块能有效提取旋转不变特征,多层级的特征连接也使得网络学习到更加丰富的特征描述符。

表 1 在 ModelNet40 上的分类实验结果

Table 1 Classification results on ModelNet40 dataset

方法	输入	NR/NR	NR/AR
PointNet		88.45	12.47
PointNet++		89.82	21.35
Kd-Network		86.20	8.49
Spherical CNN	Point	81.73	55.62
SO-Net	(x, y, z)	93.64	9.64
PRIN		80.13	69.85
Ours		86.13	71.12

此外,网络去除了传统算法中的空间对齐网络模块并采取直观简洁的特征编码方法,结合简单的全连接层在一个 GPU 上用较短的时间对旋转不变特征进行训练和测试,在旋转测试集中,识别精度远超广泛应用于点云分析网络的 PointNet 系列,对比其他方法也有着明显优势。在普通测试集中可以达到与其他方法相当的性能,但计算效率仍然最高,其中参数量仅有 0.2 M,浮点运算数 FLOPs 为 1.8×10^6 ,对比其他方法都有着数量级别的优势,具体统计如表 2 所示。

表 2 计算量 FLOPs 与参数量统计结果

Table 2 Statistics result of FLOPs and parameter

方法	FLOPs	参数量
PointNet	4.4×10^8	3.5×10^6
PointNet++	1.6×10^9	1.7×10^6
KD-Networks	$> 10^9$	$> 10^7$
Spherical CNNs	2.9×10^7	5.0×10^5
SO-Net	$> 10^8$	$> 10^6$
PRIN	$> 10^8$	4×10^6
Ours	1.8×10^6	2×10^5

2.3 分割实验

为了进一步验证 EFRI-N 方法的有效性,在常用分割

表 3 在 ShapeNet Parts 上的分割实验结果

Table 3 Segmentation results on ShapeNet Parts dataset

Method	Arbitrary Rotation																No Rotation			
	avg. inst.	avg. cls.	air plane	bag	cap	car	chair	car phone	guitar	knife	lamp	laptop	motor bike	mug	pistol	rocket	skate board	table	avg. inst.	avg. cls.
PointNet	31.3	29.38	19.9	46.25	43.27	20.81	27.04	15.63	34.72	34.64	42.1	36.4	19.25	49.88	33.3	22.07	25.71	29.74	83.15	78.95
PointNet++	36.66	35	21.9	51.7	40.06	23.13	43.03	9.65	38.51	40.91	45.56	41.75	18.18	53.42	42.19	28.51	38.92	36.57	84.63	81.52
SO-Net	26.21	14.37	21.08	8.46	1.87	11.78	27.81	11.99	8.34	15.01	43.98	1.81	7.05	8.78	4.41	6.38	16.1	34.98	84.83	81.16
DGCNN	43.79	30.87	24.84	51.29	36.69	20.33	30.07	27.86	38	45.5	42.29	34.84	20.51	48.74	26.25	26.88	26.95	28.85	85.15	82.33
SpiderCNN	31.81	35.46	22.28	53.07	54.2	22.57	28.86	23.17	35.85	42.72	44.09	55.44	19.23	48.93	28.65	25.61	31.36	31.32	85.33	82.4
Ours	59.33	52.04	15.36	58.52	54.38	39.81	55.18	29.63	79.34	81.24	74.5	68.87	20.89	48.63	53.1	15.27	61.47	76.45	82.31	76.59

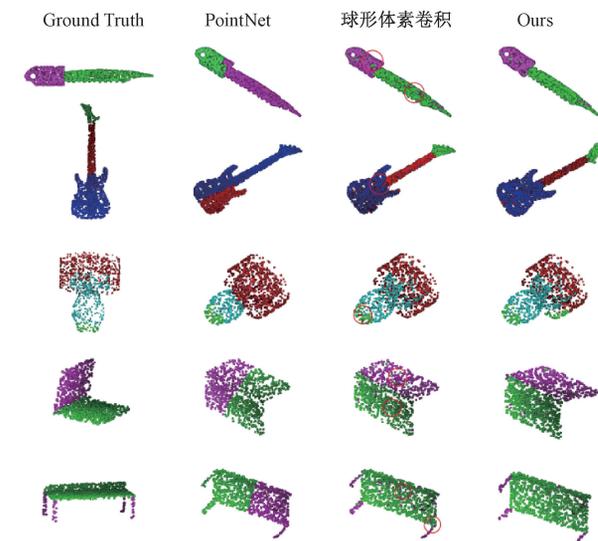


图 4 可视化分割结果

Fig. 4 Visualization of segmentation results

数据集 ShapeNet Parts 上进行实验,该数据集包括 16 881 个形状,共有 16 个类别、50 个零件标签。每个形状包含若干个点,每个点都有对应的标签,因此细颗粒度的分割实验对于旋转物体来说更具有挑战性。与分类实验类似,用不经过处理的数据集训练,并在非旋转和旋转测试集中进行评估。对比了一些具有局部旋转不变性的网络结构,结果如表 3 所示。PointNet 等经典方法在处理对齐点云数据中表现良好,但使用旋转数据集输入时,性能大幅度下降,而本文方法不仅可以完成分割任务,并且在处理旋转物体时有明显优势,充分验证了本算法的有效性。

此外,还进行了定性分析,将 EFRI-N 与骨干网络 PointNet、球形体素卷积方法的分割结果进行对比,如图 4 所示。不难发现,PointNet 出现大面积错误预测点,很难完成旋转形状的分割任务,球形体素卷积方法在形状过度特征以及边缘处存在较多错误预测点,如小刀刀柄和刀刃、吉他琴颈与琴箱连接部分、台灯底座、电脑键盘以及桌面等位置的画圈处,本文方法的分割结果更接近 Ground Truth,在物体表面均匀分布部分表现出更好的性能,预测错误率更低,能够有效识别出旋转物体的零部件。

2.4 消融实验

为了探讨本文设计的各个模块对点云识别结果的影响,分别构建了不同的网络结构进行对比试验。实验 1) 为一个简单的逐点学习的 mlp 网络,实验 2) 在实验 1) 的基础上增加 T-Net 模块试图将点云对齐到规范姿态,实验 3) 中使用本文提出的旋转不变特征提取前置网络模块替代 T-Net 模块,以对比两种变换网络的优劣。实验 4) 在实验 3) 的基础上增加了多层级的跳跃连接以验证该模块对网络的有效性。消融实验结果如表 4 所示,表中“√”代表有指定模块,“\”则代表无。对比实验 1) ~ 3) 发现,在面对点云的旋转性问题时,本文提出的前置网络模块有着显著优势,分类分割准确率分别提高 50% 和 20% 左右,原因在于椭球拟合本质上是基于数学的计算方法,提取的旋转不变特征更加具有唯一性,受网络影响较小。而 T-Net 网络通过学习的方法得到旋转矩阵来对齐点云,旋转矩阵的有效性不仅与网络结构密切相关,也

更容易受到外界因素的影响。实验 4) 中引入多层级的跳跃连接也使得网络性能提高了约 1%, 说明低层级和高层级的特征融合在点云识别中也发挥着重要作用。

表 4 不同组件的网络模型识别准确率

Table 4 Recognition accuracy of different components

模块名称	组成结构			
	1)	2)	3)	4)
T-Net	\	✓	\	\
旋转不变前置网络	\	\	✓	✓
多层次跳跃链接	\	\	\	✓
分类准确率/%	10.38	12.47	69.73	71.12
分割准确率/%	27.65	31.3	58.04	59.33

为了测试网络对输入点云分辨率的鲁棒性, 设计如下消融实验, 改变输入点云的采样点数, 范围为 256 ~ 2 048, 实验结果如图 5 所示, 随着点数的变化, 模型的识别精度只有微小的波动, 表明尽管点云的空间分布不同, EFRI-N 仍然能够提取有效的全局信息。当点数设置为 2 048 时, 模型获得最佳效果, 原因在于点数越多, 携带的信息越丰富, 为了均衡网络性能和计算量, 本文最终选取 2 048 个点作为网络输入。

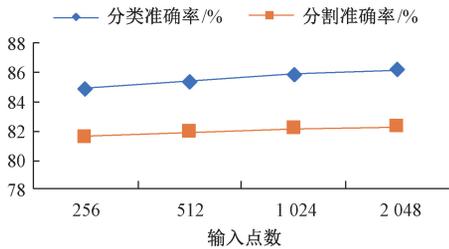


图 5 采样点数的影响

Fig. 5 The effect of the number of sampled points

3 结论

本文针对传统点云网络的旋转性问题提出一种基于椭球拟合的旋转不变网络。利用前置网络中的椭球拟合算法准确标识物体方向, 结合特征编码方法得到稳定的旋转不变特征, 最后通过多层级的特征连接获取更具有鉴别力的几何信息完成分类和分割任务。对比目前先进算法, 本方法能够解决角度未知问题造成的形状识别误差, 很大程度上提高旋转物体的分类、分割精度, 增强网络鲁棒性, 同时也极大的降低了网络复杂度。本算法目前仅针对的单目标场景下的三维模型做到了旋转不变, 接下来将对包含多个三维目标的自然场景展开研究。

参考文献

[1] 郭晓锐, 章智涵, 张弛, 等. 激光点云孤岛数据高精度滤波算法研究[J]. 电子测量技术, 2021, 44(24): 116-121.

GUO X R, ZHANG ZH H, ZHANG CH, et al. Research on high precision filtering algorithm for isolated data of LiDAR point cloud [J]. Journal of Electronic Measurement and Instrumentation, 2021, 44 (24) : 116-121.

[2] 余洪山, 付强, 孙健, 等. 面向室内移动机器人的改进 3D-NDT 点云配准算法[J]. 仪器仪表学报, 2019, 40(9): 151-161.

YU H SH, FU Q, SUN J, et al. Improved 3D-NDT point cloud registration algorithm for indoor mobile robot [J]. Chinese Journal of Scientific Instrument, 2019, 40(9): 151-161.

[3] 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). IEEE, 2015: 922-928.

[4] SU H, MAJI S, KALOGERAKIS E, et al. Multi-view convolutional neural networks for 3D shape recognition [C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 945-953.

[5] YU T, MENG J, YUAN J. Multi-view harmonized bilinear network for 3D object recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 186-194.

[6] QI C R, SU H, MO K, et al. Pointnet: Deep learning on point sets for 3D classification and segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 652-660.

[7] QI C R, YI L, SU H, et al. Pointnet ++: Deep hierarchical feature learning on point sets in a metric space [J]. ArXiv Preprint, 2017, arXiv:1706.02413.

[8] WANG Y, SUN Y, LIU Z, et al. Dynamic graph cnn for learning on point clouds [J]. Acm Transactions on Graphics (tog), 2019, 38(5): 1-12.

[9] XU Y, FAN T, XU M, et al. Spidercnn: Deep learning on point sets with parameterized convolutional filters [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 87-102.

[10] STEIN F, MEDIONI G. Structural indexing: Efficient 3-D object recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14 (2) : 125-145.

[11] TOMBARI F, SALTI S, DI STEFANO L. Unique signatures of histograms for local surface description [C]. European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2010: 356-369.

[12] ESTEVES C, ALLEN-BLANCHETTE C, MAKADIA A,

- et al. Learning so (3) equivariant representations with spherical cnns [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 52-68.
- [13] COHEN T S, GEIGER M, KÖHLER J, et al. Spherical CNNs[J]. ArXiv Preprint,2018, arXiv:1801.10130.
- [14] YOU Y, LOU Y, LIU Q, et al. Pointwise rotation-invariant network with adaptive sampling and 3D spherical voxel convolution[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12717-12724.
- [15] LEVENBERG K. A method for the solution of certain non-linear problems in least squares [J]. Quarterly of Applied Mathematics, 1944, 2(2): 164-168.
- [16] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks [J]. Advances in Neural Information Processing Systems, 2015, 28: 2017-2025.
- [17] WU Z, SONG S, KHOSLA A, et al. 3D shapenets: A deep representation for volumetric shapes [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1912-1920.
- [18] YI L, KIM V G, CEYLAN D, et al. A scalable active framework for region annotation in 3D shape collections[J]. ACM Transactions on Graphics (ToG), 2016, 35(6): 1-12.
- [19] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. ArXiv Preprint,2014,arXiv:1412.6980.
- [20] KLOKOV R, LEMPITSKY V. Escape from cells: Deep kd-

networks for the recognition of 3D point cloud models[C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 863-872.

作者简介



黎书玉, 2015年于安徽工程大学获得学士学位, 现为合肥工业大学硕士研究生, 主要研究方向为机器视觉。

E-mail: 2019170877@mail.hfut.edu.cn

Li Shuyu received her B. Sc. degree from Anhui Polytechnic University in 2015. Now she is a M. Sc. candidate at Hefei University of Technology. Her main research interest includes machine vision.



张旭东(通信作者), 1989年于合肥工业大学获学士学位, 1992年于合肥工业大学获硕士学位; 2005年于中国科学技术大学获博士学位。现为合肥工业大学教授, 主要研究方向为图像处理、模式识别、智能信息处理。

E-mail: xudong@hfut.edu.cn

Zhang Xudong(Corresponding author) received his B. Sc. degree from Hefei University of Technology in 1989, M. Sc. degree from Hefei University of Technology in 1992 and Ph. D. degree from University of Science and Technology of China in 2005, respectively. Now he is a professor in Hefei University of Technology. His main research interests include image processing, pattern recognition and intelligent information processing.