

文章编号: 1673-3193(2024)06-0753-11

## 结合通道优先卷积注意力的点云配准网络

张艺旋<sup>1,2,3</sup>, 韩 燮<sup>1,2,3</sup>, 赵 融<sup>1,2,3</sup>, 王 松<sup>1,2,3</sup>, 程 郴<sup>4</sup>

- 中北大学 计算机科学与技术学院, 山西 太原 030051;
- 中北大学 机器视觉与虚拟现实山西省重点实验室, 山西 太原 030051;
- 中北大学 山西省视觉处理及智能机器人工程研究中心, 山西 太原 030051;
- 中国辐射防护研究院, 山西 太原 030006)

**摘要:** 点云配准网络在处理大规模点云和捕获局部细节特征时存在局限性, 导致其对点云重叠区域配准的精度不足, 本文提出了一种新的点云配准网络CR-RORNet来解决此问题。该网络结合了通道优先卷积注意力和ResPointNet模块, 克服了现有方法在应对复杂场景和不规则点云时的不足, 提升了复杂点云模型的配准效果。首先, 在粗配准阶段设计了ResPointNet模块, 通过引入残差连接机制强化了点云模型全局特征和多层次特征的提取与融合。其次, 在动态图卷积神经网络中将通道优先卷积注意力机制CPCA (Channel Prior Convolutional Attention) 与跨阶段梯度聚合机制进行结合, CPCA机制利用通道先验信息加强了网络对重要特征通道和区域的关注, 在处理点云重叠部分时, 能有效增强网络模型对点云局部细节特征的捕捉能力并抑制低置信度区域的影响, 从而显著提升配准的效果; 跨阶段梯度聚合机制融合了点云模型不同深度层次的梯度信息, 确保在处理微小零部件或大范围场景点云模型时, 网络能充分理解点云的结构和局部细节, 并使学习到的特征具有良好的表达力, 从而实现了复杂场景下点云数据的高精度配准。实验表明, CR-RORNet在自采数据集上的表现优于其他点云配准方法, 相比基线RORNet, CR-RORNet在 $RMSE(t)$ 误差降低了39.5%, 在 $MSE(R)$ 误差降低了5.1%。公开数据集ModelNet40中的实验结果表明, 该网络具有良好的泛化性能。

**关键词:** 点云配准; 深度学习; 注意力机制; 残差网络

**中图分类号:** TP391

**文献标识码:** A

**doi:** 10.3969/j.issn.1673-3193.2024.06.004

**引用格式:** 张艺旋, 韩燮, 赵融, 等. 结合通道优先卷积注意力的点云配准网络[J]. 中北大学学报(自然科学版), 2024, 45(6): 753-763.

ZHANG Yixuan, HAN Xie, ZHAO Rong, et al. Point cloud registration network combining channel priority convolution attention[J]. Journal of North University of China(Natural Science Edition), 2024, 45(6): 753-763.

## Point Cloud Registration Network Combining Channel Priority Convolution Attention

ZHANG Yixuan<sup>1,2,3</sup>, HAN Xie<sup>1,2,3</sup>, ZHAO Rong<sup>1,2,3</sup>, WANG Song<sup>1,2,3</sup>, CHENG Chen<sup>4</sup>

- School of Computer Science and Technology, North University of China, Taiyuan 030051, China;
- Shanxi Provincial Key Laboratory of Machine Vision and Virtual Reality, North University of China, Taiyuan 030051, China;
- Shanxi Province's Vision Information Processing and Intelligent Robot Engineering Research Center, North University of China, Taiyuan 030051, China;
- China Institute for Radiation Protection, Taiyuan 030006, China)

**收稿日期:** 2024-04-10

**基金项目:** 国家自然科学基金资助项目(62272426); 山西省科技重大专项计划“揭榜挂帅”项目(202201150401021)

**作者简介:** 张艺旋(1998-), 男, 硕士生, 主要从事点云配准的研究。

**通信作者:** 韩 燮(1964-), 女, 教授, 博士, 主要从事虚拟仿真与可视化技术和智能信息处理的研究。E-mail: hanxie@nuc.edu.cn。

**Abstract:** The limitations of point cloud registration networks in processing large-scale point clouds and capturing local detail features, resulted in insufficient registration accuracy for overlapping areas of point clouds, so a new point cloud registration network CR-RORNet was proposed. This network combined channel first convolutional attention and ResPointNet module, and overcame the shortcomings of existing methods in dealing with complex scenes and irregular point clouds, and improved the registration performance for point clouds with significant initial pose differences. Firstly, in the coarse registration stage, the ResPointNet module was designed to enhance the extraction and fusion of global and multi-level features of the point cloud model by introducing residual connection mechanism. Secondly, in the dynamic graph convolutional neural network, channel prior convolutional attention CPCA and cross stage gradient aggregation mechanism were combined. The CPCA mechanism utilized channel prior information to strengthen the network's attention to important feature channels and regions. When dealing with overlapping parts of point clouds, it could effectively enhance the network model's ability to capture local detail features of point clouds and suppress the influence of low confidence regions, significantly improving the registration effect; the cross stage gradient aggregation mechanism integrated gradient information from different depth levels of point cloud models, and ensured that the network could fully understand the structure and local details of the point cloud when dealing with small components or largescale scene point cloud models. And the learned features had good expressive power, so high-precision registration of point cloud data in complex scenes was realized. Experimental results show that CR-RORNet performs better than other point cloud registration methods on self collected datasets. Compared to baseline RORNet, CR-RORNet reduces  $RMSE(t)$  error by 39.5% and  $MSE(R)$  error by 5.1%. Experiment results on the publicly available dataset ModelNet40 show that the network has good generalization performance.

**Key words:** point cloud registration; deep learning; attention mechanism; residual network

## 0 引言

随着三维采集技术的日益成熟,高质量三维点云数据在机器导航<sup>[1]</sup>、智慧城市<sup>[2]</sup>与无人驾驶<sup>[3]</sup>等新兴领域的需求越来越大。如何获得高质量点云数据,提高点云配准精度至关重要。点云配准的目标是找到一个刚性变换,定义一对点云之间的对应关系,将不同视角或位置获取的点云数据进行对齐和整合,为点云分割<sup>[4-5]</sup>、目标检测<sup>[6-7]</sup>等场景提供更完整的三维信息。

许多基于深度学习的点云配准网络只考虑小数据模型且只在初始位姿基本一致的情况下进行实验,具有较大的局限性。同时,当点云数据重叠区域提取效果不佳时难以获取准确的配对信息来达到精确配准的目的;另外,越来越多的点云特征提取网络在对点云模型进行特征提取时,过于关注全局特征而忽略了细节特征导致配准精度不高。

如图1所示,实际采集的雕像点云模型噪声和异常点多,由于现有的深度学习网络多采用标准数据集为训练集,对大规模点云模型配准的泛

化性较差,导致对真实数据处理的效果不佳。

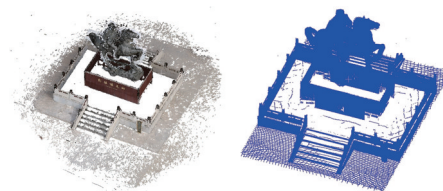


图1 待配准点云模型

Fig. 1 Point cloud model to be registered

针对现有点云配准网络存在的以上问题,本文提出一种新型的点云配准网络 CR-RORNet (CPCA and ResPointNet based on Partial-to-Partial Registration Network With Reliable Overlapping Representations)。首先,在粗配准环节中设计了点残差模块 ResPointNet,以进行多层次的点云特征提取和融合。其次,在动态图卷积神经网络 DGCNN<sup>[8]</sup>中引入通道优先卷积注意力 CPCA<sup>[9]</sup>和跨阶段梯度聚合机制<sup>[10]</sup>,其中,CPCA利用通道先验信息增强有用通道和重要区域的聚焦,可以有效提高特征提取的性能,而跨阶段梯度聚合机制通过多尺度卷积捕捉不同尺度的特征来增加特征的多样性。CR-RORNet通过提取高

质量特征弥补点云重叠估计时的误差,可以提高点云配准效果,如图 2 所示。

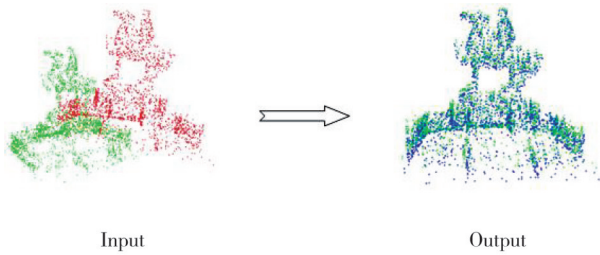


图 2 点云配准结果

Fig. 2 Result of point cloud registration

### 1 相关工作

近年来,随着深度学习理论的快速发展,深度神经网络被广泛应用于3D数据处理。PointNet<sup>[11]</sup>和PointNet++<sup>[12]</sup>的出现解决了过去神经网络无法直接处理无序点云的问题,因此利用深度学习方法进行点云配准受到了广大学者的关注。

现有的基于深度学习的点云配准可分为基于无对应点的方法和基于对应点的方法<sup>[13]</sup>。

基于无对应点的方法利用神经网络提取全局高维特征,通过回归分析直接计算出刚性变换参数。PointLK<sup>[14]</sup>在PointNet基础上融合了Lucas-Kanade(LK)算法,把点映射到高维空间得到多维特征向量,再使用LK算法进行点云配准,取得了良好效果,但是该网络使用微分理论求导计算,计算效率较低。为了解决计算效率低的问题。PCRNet<sup>[15]</sup>使用了数据驱动的深度特征匹配层来输出变换参数,提高了算法的效率和鲁棒性。但

在实际应用中,由于采集到的点云模型具有噪声和异常点,这类方法通常效果不佳。

当前,主流的基于对应点的方法大多基于PointNet和PointNet++。DeepVCP<sup>[16]</sup>基于Mini-PointNet的匹配特征提取模块,从关键点提取高层次的特征表示,生成对应点对进行有效的特征匹配。3DRegNet<sup>[17]</sup>在PointNet基础上结合了深层ResNet提取特征,在分类模块输出离群点的概率,在配准模块输出刚体变换参数,该方法在高度重叠的点云上实现了较高精确度的配准,但在部分重叠的数据上效果不佳。RORNet<sup>[18]</sup>为了减少重叠估计误差对配准的影响,采用了一种表示生成网络RGN(Representation Generation Network)模块对相似度矩阵进行下采样以保留可靠的点,但其在重叠区域的特征提取过程中忽略了局部细节特征,在大规模点云配准中精度有所降低。

综上所述,为了解决以上问题,本文设计了CR-RORNet网络,通过加入ResPoint-Net模块,在DGCNN网络中引入通道优先卷积注意力机制和跨阶段梯度聚合机制,对现有网络存在的问题进行改进,从而使点云配准技术更好地应用在复杂的真实点云模型中。

### 2 新型点云配准网络

#### 2.1 网络整体架构

本文使用RORNet网络的重叠估计模块DOEN,在两个输入模型中提取重叠区域,之后重点关注表示生成网络。

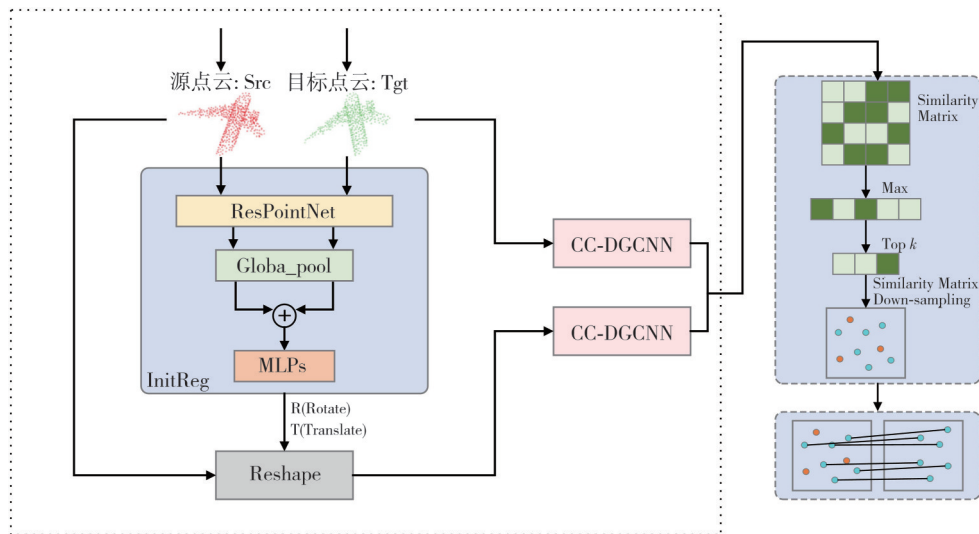


图 3 CR-RORNet网络结构图

Fig. 3 CR-RORNet model structure

如图3虚线框内所示,首先,设计了一个粗配准网络InitReg,并在该网络中提出了一个新的Res-PointNet模块,增强了神经网络对点云整体模型全局特征和多层次特征的提取与融合,增强了点云粗配准时的精度,InitReg网络输出一个粗略的刚性变换参数,使用该参数变换其中一个输入模型的位姿,有效提高了后续点云精配准的效果;其次,在特征提取阶段使用基于CPCA机制和RepNCSP模块改进的CC-DGCNN网络,采用多尺度的深度可分离卷积和动态注意力权重,使网络更高效地捕捉和聚合不同空间尺度下的重要特征,在通道维度上计算注意力权重,根据其包含特征的重要性动态分配权重系数;结合了跨阶段梯度聚合机制,跨层次融合了点云的特征编码,增强了点云的特征提取效果。最后,使用提取到的特征计算相似度矩阵,为了选取可靠的点,取相似度矩阵中每一行最大的值组成最大值向量,取最大值向量中前 $k$ 个最大值,即找到可靠的对应点集,利用SVD估计出旋转矩阵和平移向量,完成点云配准。

## 2.2 粗配准模块

如图4所示,InitReg模块用于实现点云模型的粗配准。

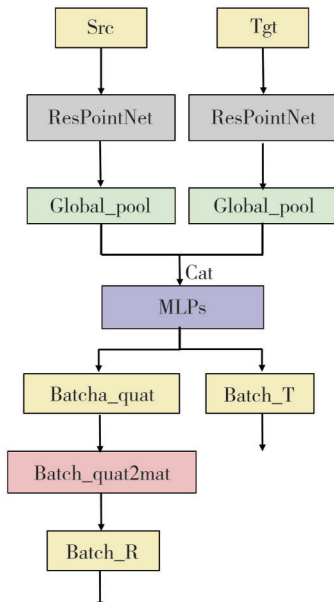


图4 InitReg网络结构图

Fig. 4 InitReg model structure

该模块使用ResPointNet作为编码器,MLP作为解码器,接收两个点云模型作为网络的输入。将输入的源点云和目标点云分别记为

$$\begin{aligned} S &\in \mathbf{R}^{B \times 3 \times N_s}, \\ T &\in \mathbf{R}^{B \times 3 \times N_t}, \end{aligned} \quad (1)$$

式中: $B$ 为批次大小; $N_s, N_t$ 分别为源点云和目标点云的点数。使用ResPointNet编码器分别对 $S$ 和 $T$ 进行特征提取,可得

$$\begin{aligned} S_{\text{emb}} &= E(S) \in \mathbf{R}^{B \times D \times N_s}, \\ T_{\text{emb}} &= E(T) \in \mathbf{R}^{B \times D \times N_t}, \end{aligned} \quad (2)$$

式中: $D$ 为编码器输出的维度。

针对常规编码器结构的缺点,使用基于PointNet架构引入残差连接的ResPointNet用于提取点云的特征表示,如图5所示。

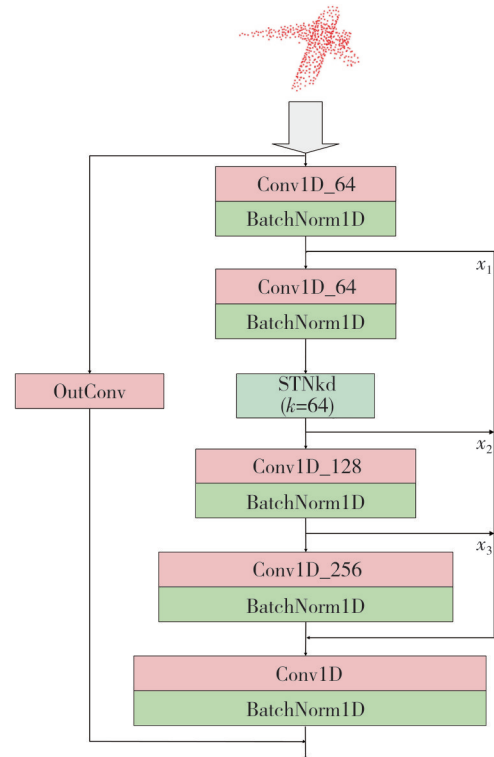


图5 ResPointNet网络结构图

Fig. 5 ResPointNet model structure

该网络通过多个卷积层和归一化层提取点云的逐点特征,并利用空间变换网络对特征进行局部对齐,将不同层次的特征连接并通过最后一层卷积以及残差连接来保留原始输入的信息,从而得到最终的特征表示。将输入点云记为 $X \in \mathbf{R}^{B \times 3 \times N}$ ,对输入 $X$ 应用两次卷积和批量归一化,可得

$$\begin{aligned} X_1 &= \text{ReLU}(\text{BN}(\text{Conv1}(X))) \in \mathbf{R}^{B \times 64 \times N}, \\ X_2 &= \text{ReLU}(\text{BN}(\text{Conv2}(X_1))) \in \mathbf{R}^{B \times 64 \times N}. \end{aligned} \quad (3)$$

对 $X_2$ 应用空间变换网络STN得到变换矩阵并对 $X_2$ 进行维度交换,与变换矩阵进行矩阵乘法实现点云的空间变换,即

$$\begin{aligned} T &= \text{STN}(X_2) \in \mathbf{R}^{B \times N \times N}, \\ X_2^{\epsilon} &= (X_2 \cdot T)^T \in \mathbf{R}^{B \times 64 \times N}. \end{aligned} \quad (4)$$

对变换后得到的  $X_2^\epsilon$  应用连续两次卷积和批量归一化, 可得

$$\begin{aligned} X_3 &= \text{ReLU}(\text{BN}(\text{Conv3}(X_2^\epsilon))) \in \mathbf{R}^{B \times 128 \times N}, \\ X_4 &= \text{ReLU}(\text{BN}(\text{Conv4}(X_3))) \in \mathbf{R}^{B \times 256 \times N}. \end{aligned} \quad (5)$$

其次沿着通道维度连接上述特征, 可得

$$X_5 = \text{Concat}(X_1, X_2^\epsilon, X_3, X_4, \text{dim} = 1) \in \mathbf{R}^{B \times 512 \times N}. \quad (6)$$

对连接后的特征  $X_5$  应用卷积和批量归一化, 得

$$Y = \text{ReLU}(\text{BN}(\text{Conv5}(X_5))) \in \mathbf{R}^{B \times D \times N}. \quad (7)$$

对原始输入  $X$  应用输出卷积层得到残差分支特征  $Z$ , 将其与主分支特征  $Y$  求和实现残差连接。

$$O = Y + \text{OutConv}(X) \in \mathbf{R}^{B \times D \times N}, \quad (8)$$

输出特征张量  $O$ 。

接着将通过 ResPointNet 得到的点云特征分别进行全局最大池化操作, 以获取每个点云的全局特征。

$$\begin{aligned} s_{\text{glob}} &= \max_{n_s} S_{\text{emb}}[:, :, n_s] \in \mathbf{R}^{B \times D}, \\ t_{\text{glob}} &= \max_{n_t} T_{\text{emb}}[:, :, n_t] \in \mathbf{R}^{B \times D}. \end{aligned} \quad (9)$$

完成两个点云模型的全局特征向量连接, 即

$$g = [s_{\text{glob}}; t_{\text{glob}}] \in \mathbf{R}^{B \times 2D}. \quad (10)$$

将连接好的全局特征输入到解码器网络中, 得到七自由度的相对姿态估计

$$p_{7d} = D(g) \in \mathbf{R}^{B \times 2D}, \quad (11)$$

式中:  $p_{7d}$  包含 3 个平移分量和 4 个旋转分量, 完成点云模型间的粗配准。

ResPointNet 中的空间变换网络 STN 由一系列卷积层和全连接层组成, 该模块从输入的点云特征中学习一个变换矩阵, 用来对点云数据进行仿射变换或相似变换的, 目的是对齐点云中的局部结构, 消除由于视角变化、采集设备误差等因素导致的非一致性。STN 通过输出变换矩阵与单位矩阵相加确保了变换的可逆性和稳定性, 然后将该变换应用于点云特征, 从而实现点云的自适应对齐。ResPointNet 中连接的多个批量归一化层有助于学习鲁棒性较高的特征表示, 而且在 STN 之后直接应用 BN 层, 标准化每一层的输出有效减少了内部神经元之间的耦合, 从而减轻了局部扰动和不规则性的影响。该网络末尾实现了特征连接, 将经过多层卷积和 STN 变换得到的深层特征和原始输入特征经过简单线性变换的结果相连接, 这种处理方式保留了低层详细信息的同时结合了高层的抽象信息, 实现了多尺度特征的有效融合。有助于模型捕捉不同层次的几何和结构信息, 提升了特征的表达能力和模型的识别能力。

在 InitReg 模块的编码器中引入 ResPointNet 结构, 能够自适应地对点云进行局部对齐, 消除了点云内部的局部扰动和不规则性, 通过多尺度特征融合提升了特征的表达能力, 完成了粗配准任务, 从而为后续精配准提供较好的变换初值。

### 2.3 通道优先卷积注意力和跨阶段梯度聚合驱动的特征提取网络

DGCNN 提出了一个动态图边卷积层 EdgeConv, 在网络中每一个点与其近邻构成一个动态图结构。EdgeConv 在这个图结构上进行卷积操作, 对点云中每一对点(中心点与邻居点)执行一次小规模 MLP, 并将 MLP 的输出视为这两点间的特征表示, 然后通过对所有邻居点的特征进行聚合, 生成一个新的中心点特征。这种卷积操作有效地扩展了感受野的同时保持了稀疏性。但是, 基于 CNN 的方法在捕捉全局信息依旧会受到有限感受野的限制。EdgeConv 在处理点云数据时, 通过考虑每个点与其邻居的关系来进行特征提取和更新, 这种方式由于在局部邻域内进行注意力特征处理, 无法有效地建立大范围的特征依赖关系。为了克服以上问题, 本文提出了一种使用 RepNCSP 模块和通道优先卷积注意力机制的动态图边卷积网络(CC-DGCNN), 网络结构如图 6 所示。该方法通过通道混合策略增强特征提取效果, 采用多尺度的深度可分离卷积模块来构成空间注意力, 有助于增强点云重叠部分的特征判别并忽略部分置信度较低的区域, 从而提高了点云配准的效果。

RepNCSP 模块的架构图如图 7 所示, 该模块由 Wang 等<sup>[10]</sup>提出, 其构建了一个包含两个分支的神经网络模型, 其中一个分支通过多次 RepNBottleneck 处理输入特征图, 另一个分支仅通过  $1 \times 1$  卷积处理输入特征图。

该模块使用瓶颈结构, 先压缩通道数再恢复到期望的输出通道数, 实质上实现了特征的压缩与扩张, 在面对点云数据量大、稀疏密不均的特点时, 这种结构能在有效提取核心特征的同时去除冗余信息, 保留关键信息, 通过轻量化特征的融合在不同尺度下提升点云的表达力。同时, 使用了跨阶段局部特征融合, 而 RepNBottleneck 通过多尺度卷积来捕捉不同尺度的特征, 两路分支所提取的特征具有一定的差异性, 同时, 点云数据的复杂性和不规则性要求网络能够捕捉多尺度特征, 而 RepNCSP 中的多尺度卷积模块, 通过不同尺度的卷积, 能捕捉点云中的多层次结构, 这意味着网络能更好地理解局部细

节和整体几何,在提高对齐精度连接后增加了特征的多样性,从而提升了网络的表达能力。

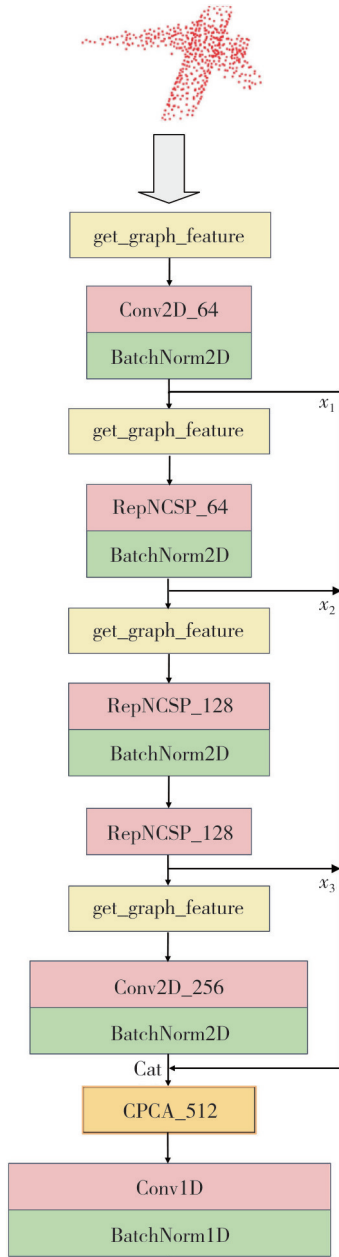


图6 CC-DGCNN网络结构图  
Fig. 6 CC-DGCNN model structure

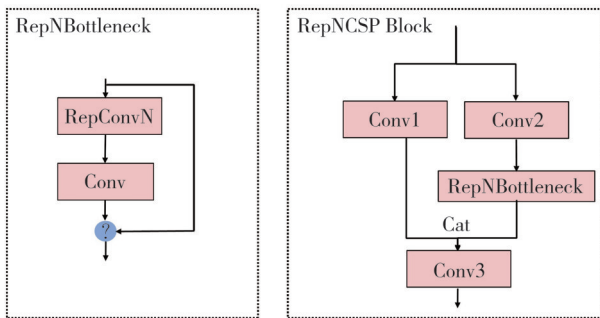


图7 RepNCSP网络结构图  
Fig. 7 RepNCSP model structure

CPCA为通道优先卷积注意力模块,网络结构如图8所示。

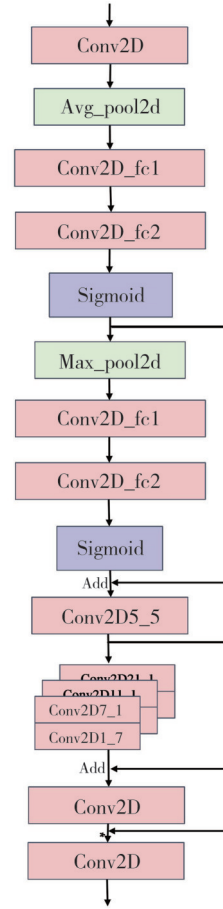


图8 CPCA网络结构图

Fig. 8 CPCA model structure

CPCA的输入为点云模型的特征  $X \in \mathbb{R}^{512 \times H \times W}$ , 特征图通过全局平均池化和最大化得到  $X_{avg}, X_{max} \in \mathbb{R}^{512 \times 1 \times 1}$ , 将所得结果通过一个全连接层后得到  $A_{avg}, A_{max}$ 。通道注意力权重为

$$A = A_{avg} + A_{max} \quad (12)$$

接着对  $X$  应用 GELU 激活函数得到新的特征图  $X'$ , 并与计算得到的通道注意力权重  $A$  相乘得到  $X''$  来实现通道注意力机制的应用。之后使用多个方向性卷积层(共7层,卷积核大小分别为  $5 \times 5, 1 \times 7, 7 \times 1, 1 \times 11, 11 \times 1, 1 \times 21$  及  $21 \times 1$ ), 捕捉输入特征图在不同方向和尺度上的特征。各个方向性卷积的输出相加得到中间特征图  $Z$ , 然后通过一个  $1 \times 1$  卷积层得到空间注意力权重  $S$ , 与  $X''$  相乘得到  $O$  来实现空间注意力机制的应用。最后经过输出卷积得到最终的特征图  $\hat{O}$ 。

CPCA 机制使用全局平均池化和最大池化对输入的点云特征图的各个通道进行池化操作,生成通道注意力权重,实现了通道特征加权,应用分组卷

积提取不同方向和尺度的空间特征,将空间特征合并后生成空间注意力权重,与原始输入特征图相乘实现了空间特征加权,最后将加权后的值输出得到最终特征。这种设计根据输入点云模型的全局信息来动态调整每个通道的权重,实现了对输入特征的自适应加权,通过增强网络对重要特征通道和空间位置的关注来提高对关键特征的识别能力,显著提升了特征提取效果。CPCA 中的通道注意力机制通过通道先验信息能够自动发现并强调那些对点云配准起决定性作用的特征通道,例如强度或深度,通过加权这些通道,网络能更有效地集中于那些对配准关键的特征,并忽略不相关信息或噪声信息。同时,CPCA 利用多尺度深度可分离卷积模块,从不同尺度捕捉特征,融合局部细节和整体形状,从而提升配准精度。

CC-DGCNN 通过构建动态图卷积神经网络模拟出类似图像的空间注意力机制,通过关注点与其邻域内的关系,CPCA 机制帮助网络捕捉局部细节和结构信息,这对于重叠区域的配准尤为重要,通过这种方式,网络即使在处理有大量噪声环境的点云也能更好地理解点云的几何结构。同时,CPCA 机制有效融合了不同尺度和方向的局部信息,提高了捕捉点云中复杂形状和细微结构的能力,特别是在处理噪声较大或分布不均的点云时,可提高模型的鲁棒性。通过堆叠多个 RepNCSP 模块形成深层次特征金字塔来保留低层次的细节信息和高层次的抽象信息,提高了模型在不同粒度上理解和建模点云信息的能力。

相比其他主流特征提取网络,CR-RORNet 显著提高了复杂场景的建模能力,在点云配准的粗配准阶段,设计 ResPointNet 模块,提高了网络对深部特征的学习,尤其是在处理大规模点云时,能够更稳定地提取和融合全局特征和多层次特征,从而增强对大规模数据结构的理解能力。CPCA 机制在处理大规模点云模型和真实点云模型时能动态地强调对配准任务贡献较大的特征通道和空间位置,并抑制无关干扰信息,这样可以帮助网络在大量的数据点中快速定位到关键的局部细节,提高了配准精度,从而进一步提升了特征表达的针对性和有效性。RepNCSP 模块的跨尺度的信息整合能够帮助网络在不同分辨率下捕捉到点云的结构特性,即使在点云数据非常庞大且细节繁多的情况下也能保持良好的特征表达和配准效果。

## 2.4 损失函数及评估指标

由于重叠估计预测为二分类任务,所以使用式(13)的交叉熵损失来进行监督。

$$L_1 = \frac{1}{2N_S} \sum_{i=1}^{N_S} [\hat{m}_i^s \log(m_i^s) + (1 - \hat{m}_i^s) \log(1 - m_i^s)],$$

$$L_2 = \frac{1}{2N_T} \sum_{j=1}^{N_T} [\hat{m}_j^t \log(m_j^t) + (1 - \hat{m}_j^t) \log(1 - m_j^t)],$$

$$l_{\text{olp}} = L_1 + L_2, \quad (13)$$

式中: $\hat{m}$  表示真值的掩码向量; $m^s$  和  $m^t$  分别表示源点云和目标点云的预测重叠掩码; $L_1$  和  $L_2$  分别表示 S 点云和 T 点云的重叠损失估计; $l_{\text{olp}}$  为交叉损失。

对于配准任务,使用式(14)来测量综合生成的点云对和真值之间  $R$  和  $t$  的偏差

$$l_{\text{reg}} = \|R^T R_{\text{gt}} - I\|^2 + \|t - t_{\text{gt}}\|^2, \quad (14)$$

式中: $R_{\text{gt}}$ ,  $t_{\text{gt}}$  表示地面真值; $I$  表示单位矩阵。

最后使用组成损失函数式(15)来优化整体网络模型,组成损失定义为

$$l = \lambda l_{\text{olp}} + (1 - \lambda) l_{\text{reg}}, \quad (15)$$

式中: $\lambda$  表示控制叠加估计损失和配准损失权重的超参数。

对于点云配准任务,使用均方根误差(RMSE),平均绝对误差(MAE)和均方误差(MSE)评估网络估计和真实值之间的旋转误差  $R$  和平移误差  $t$ ,其中  $M$  表示数据的组数, $(x_i, y_i)$  表示一组数据,具体计算公式为

$$MSE = \frac{1}{M} \sum_{i=1}^M |x_i - y_i|^2,$$

$$MAE = \frac{1}{M} \sum_{i=1}^M (x_i - y_i)^2,$$

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M (x_i - y_i)^2}. \quad (16)$$

## 3 实验与结果分析

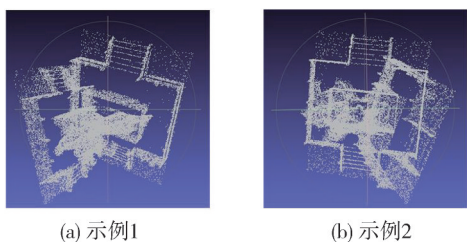
### 3.1 数据集

在 ModelNet40 数据集和自制数据集上进行实验以验证网络结构的有效性。

ModelNet 数据集是一个三维 CAD 模型集合,CAD 模型采用 .off 格式存储。本文使用基于 ModelNet40 制作的 ModelNet40\_ply\_hdf5\_2048 数据集。该数据集包含 12 308 个点云模型,其中 9 840 个用于训练,2 468 个用于测试。每个点云模型被稀疏成 2 048 个点,并采用 h5 文件进行存储。

ModelNet 数据集集中的大部分模型由 CAD 软

件导出,与真实场景中采集的数据存在较大的差距,在几何形状上相似且大部分都是正则化的。鉴于此,本文提供了一个小型点云数据集用来训练和测试本文提供的CR-RORNet网络模型和其他对比实验模型,该数据集包括400个(200对)点云模型,其中150对作为训练集,50对作为测试集。该数据集通过采集中北大学彭德怀雕像的点云,经过旋转、缩放、下采样和切割生成,并采用h5文件进行存储。部分效果图如图9所示。



(a) 示例1 (b) 示例2

图9 自采数据集部分示意图

Fig. 9 Part diagram of own data set

### 3.2 实验环境与参数设置

CR-RORNet基于PyTorch框架实现,运行在搭载两块NVIDIA RTX3090的服务器上,服务器操作系统为Ubuntu20.04。通过RAdam优化器进行参数优化,学习率设置为0.001。在自建数据集上训练了100个epoch,批量大小设置为2,重叠估计比例设置为0.8。

### 3.3 计算复杂性分析

本文对RORNet模型和本文所提出的模型进行计算复杂性分析,从浮点运算次数GFLOPs(Giga Floating-point Operations Per Second)和参数量MPrams来进行模型结构评估。RORNet和CR-RORNet的网络结构类似,分为重叠估计网络和表示生成网络,重叠估计网络部分两个网络模型的浮点运算次数均为46.328 GFLOPs,参数量为10.877 MParams,表示生成网络部分CR-RORNet的浮点运算次数为27.405 GFLOPs,参数量为56.732 MParams,RORNet的浮点运算次数为26.407 GFLOPs,参数量为51.415 MParams。对比浮点运算次数,本文提出的模型比原模型上升了3.78%,基本保持一致,参数量比原模型增加了10.34%。

CR-RORNet在浮点运算次数上相比RORNet仅上升了3.78%,这一轻微的增长表明新模

型在运算效率上与原始模型保持着高度的相似性,没有因性能提升而显著增加计算负担。然而在参数量方面,CR-RORNet增加了10.34%,这意味着相比于RORNet,新模型在模型容量上有较为明显的扩展,引入了更多的可学习参数来捕捉更为复杂的特征表示。

复杂性分析结果表明,CR-RORNet的设计通过相对较小的计算成本增长,实现了模型表达能力的显著增强。

### 3.4 对比实验

为了评估CR-RORNet在点云配准方面的性能,首先将CR-RORNet与当前主流的点云配准网络在自建数据集上进行了比较,包括PointLK<sup>[14]</sup>、PCRNet<sup>[15]</sup>、DCP-v2<sup>[16]</sup>、ReAgent<sup>[19]</sup>、UTOPIC<sup>[20]</sup>和RORNet<sup>[18]</sup>。为了保证对比的公平性,所有的方法都在自建数据集上训练及测试,使用RMSE(R)、RMSE(t)、MAE(R)、MAE(t)、MSE(R)和MSE(t)六个指标进行点云配准性能评价。定量结果如表1所示,最佳配准效果用加粗表示。

表1 在自建数据集上与其他方法的比较

Tab. 1 Comparison with other methods on own dataset

方法	RMSE (R)	RMSE (t)	MAE (R)	MAE (t)	MSE (R)	MSE (t)
PointLK	28.81	0.2970	19.75	0.2320	830.3	0.0882
PCRNet	30.97	0.1559	21.96	0.1275	959.3	0.0243
DCP-v2	32.45	0.2317	23.13	0.1766	1053.5	0.0537
ReAgent	24.15	0.1573	19.79	0.1247	716.1	0.0318
UTOPIC	23.94	0.2175	17.47	0.0941	579.4	0.0216
RORNet	21.96	0.1857	17.01	0.0857	482.51	0.0126
Ours	<b>21.39</b>	<b>0.1123</b>	<b>16.32</b>	0.1117	<b>457.58</b>	0.0344

通过与一些经典的方法以及先进方法进行比较可以发现,CR-RORNet显著降低了点云配准的旋转均方误差,与基线RORNet相比,显著降低了均方根误差。

在ModelNet40上对基于自建数据集训练的模型进行泛化性能的验证,与IDAM<sup>[21]</sup>、PRNet<sup>[22]</sup>、PointLK、PCRNet、DCP-v2、ReAgent、UTOPIC和RORNet<sup>[18]</sup>进行了对比,结果如表2所示,最佳配准效果用加粗表示。由表2可知,CR-RORNet在ModelNet40数据集上的配准效果同样保持较优的竞争力,表明了本文提出的CR-RORNet具有一定的泛化性能,可用于多种点云模型的配准任务。

最后为了验证本文提出模型在公开数据集上的有效性,基于ModelNet40数据进行训练和测试,完成有效性能的验证,并与PointLK、PCRNet、DCP-v2、ReAgent、UTOPIC和RORNet进行了对比,结

果如表 3 所示。

表 2 ModelNet40数据集上各方法泛用性的比较

Tab. 2 Generalization comparison of different methods on ModelNet40

方法	$RMSE(R)$	$RMSE(t)$	$MAE(R)$	$MAE(t)$
IDAM	5.810	0.0742	3.935	0.0507
PRNet	7.743	0.0714	4.708	0.0464
PointLK	5.920	0.0534	4.214	0.0351
PCRNet	4.091	0.0659	2.121	0.0294
DCP-v2	5.619	0.0331	2.497	0.0461
ReAgent	4.132	0.0398	3.5873	0.0317
UTOPIC	3.311	0.0294	1.875	0.0198
RORNet	1.691	<b>0.0216</b>	0.732	<b>0.0102</b>
Ours	<b>1.523</b>	0.0232	<b>0.667</b>	0.0194

由表 3 可知, CR-RORNet 在该数据集上保持了较低的配准误差, 在其他数据数据集上依旧有良好的配准效果。

表 3 ModelNet40数据集上各方法有效性的比较

Tab. 3 Effectiveness comparison of different methods on ModelNet40

方法	$RMSE(R)$	$RMSE(t)$	$MAE(R)$	$MAE(t)$
IDAM	5.194	0.0624	3.935	0.0507
PRNet	7.032	0.0622	4.708	0.0464
PointLK	5.123	0.0413	4.214	0.0351
PCRNet	3.452	0.0569	2.121	0.0294
DCP-v2	4.753	0.0274	2.497	0.0461
ReAgent	3.431	0.0343	3.5873	0.0317
UTOPIC	2.358	0.0187	1.875	0.0198
RORNet	1.119	<b>0.0137</b>	0.732	<b>0.0102</b>
Ours	<b>1.019</b>	0.0146	<b>0.667</b>	0.0194

### 3.5 可视化结果

为了更直观地展示本网络模型的效果。使用可视化的方法展示 CR-RORNet 与其他方法在自建数据集部分场景中的配准效果, 如图 10 和图 11 所示。

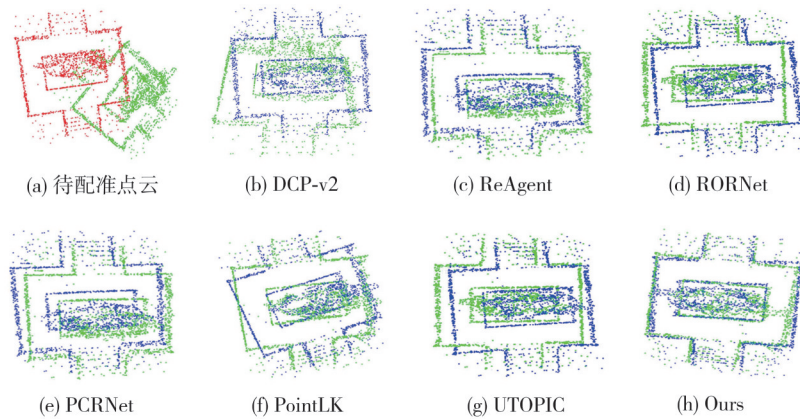


图 10 实验 1 的配准效果对比

Fig. 10 Comparison of registration effect in experiment 1

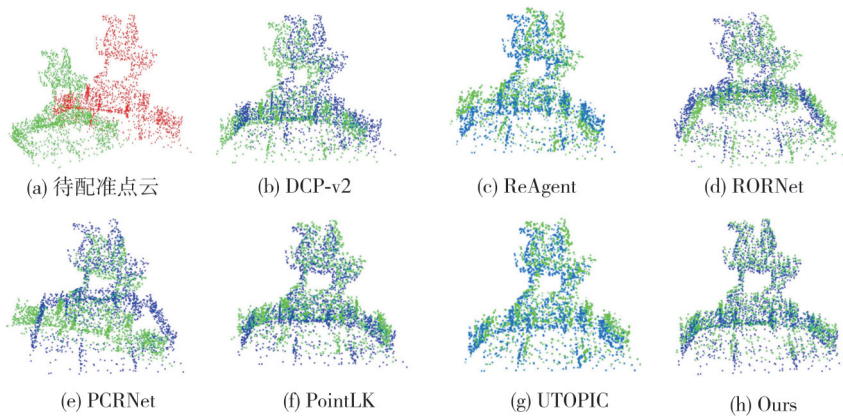


图 11 实验 2 的配准效果对比

Fig. 11 Comparison of registration effect in experiment 2

图 10 中, 每张图代表一个独立的场景, 分别展示了不同方法的点云配准效果。由图 10 和图 11 可知, 对比其他主流点云配准网络, CR-RORNet 网络在点云配准效果上有明显的提升。

图 10(a) 的示例模型中, 两个待配准点云初

始位姿已基本确定, CR-RORNet 的旋转效果明显好于 DCP-v2 和 PointLK, 平移效果略好于 RORNet 和 PCRNet; 图 11(a) 的示例模型中, 两个待配准点云初始位姿有明显差异, CR-RORNet 的旋转效果明显好于 DCP-v2 和 PCRNet, 平移效果优

于RORNet且基本和PointLK效果持平。

### 3.6 消融实验

CR-RORNet方法中的关键组件包括ResPointNet模块,CC-DGCNN网络中的通道优先卷积注意力机制和深度跨阶段梯度聚合机制,为了验证其优势和有效性,在自建数据集上进行了消融实验,并以RMSE、MAE衡量实验的配准质量。本节比较了ResPointNet,CPCA和RepNCSP三个模块的网络有效性,结果如表4所示。

表4 消融实验中使用不同模块的评价结果

Tab.4 Evaluation results using different modules in ablation experiment

ResPointNet	CPCA	RepNCSP	RMSE (R)	RMSE (t)	MAE (R)	MAE (t)
✓	✓	✓	21.47	0.1441	16.89	0.1005
✓	✓	✓	21.68	0.1697	16.54	0.0981
✓	✓	✓	21.75	0.1573	16.79	0.0994
✓	✓	✓	21.31	0.1121	16.47	0.0979

由表4可知,ResPointNet,CPCA和RepNCSP这3个模块都可以显著提高点云配准的精度,并且3个模块互补可以达到最佳效果。

## 4 结论

针对目前点云配准网络在复杂模型上表现较差、精度不足等问题,基于RORNet提出了一种结合通道优先卷积注意力机制的点云配准网络。首先,在粗配准环节中,通过在PointNet中加入残差连接形成ResPointNet模块,在保证特征提取能力的同时完成对点云模型不同层次的特征融合。然后,在DGCNN网络中结合通道优先卷积注意力机制和深度跨阶段梯度聚合机制,捕捉点云关键结构的局部细节特征,细化点云模型的特征表示,同时通过跨阶段连接实现特征复用,最终提高了点云配准的精度。与PointLK、PCRNet、DCP-v2、ReAgent、UTOPIIC和RORNet等先进的点云配准网络在自建数据集和ModelNet40数据集上进行了实验对比,结果表明CR-RORNet在复杂点云模型的配准上有更好的表现,在公开数据集上具有良好的泛化性能。

### 参考文献:

[1] ZHOU Z, FENG X, DI S, et al. A LiDAR mapping system for robot navigation in dynamic environments

[J]. IEEE Transactions on Intelligent Vehicles, 2023: 1-20.

- [2] ORTEGA S, SANTANA J M, WENDEL J, et al. Generating 3D city models from open LiDAR point clouds: Advancing towards smart city applications [M]. Cham: Springer, 2020.
- [3] PARK J, KIM C, KIM S, et al. PCSCNet: Fast 3D semantic segmentation of LiDAR point cloud for autonomous car using point convolution and sparse convolution network [J]. Expert Systems with Applications, 2023, 212:118815.
- [4] DU Z, YE H, CAO F. A novel local-global graph convolutional method for point cloud semantic segmentation [J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(4): 4798-4812.
- [5] LI S, LIU Y, GALL J. Rethinking 3-D LiDAR point cloud segmentation [J]. IEEE Transactions on Neural Networks and Learning Systems, 2021:1-12.
- [6] MISRA I, GIRDHAR R, JOULIN A. An end-to-end transformer model for 3d object detection [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2886-2897.
- [7] CHEN Y, LI H, GAO R, et al. Boost 3-D object detection via point clouds segmentation and fused 3-D GIoU-L<sub>1</sub> loss [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(2): 762-773.
- [8] PHAN A V, LE NGUYEN M, NGUYEN Y L H, et al. DGCNN: A convolutional neural network over large-scale labeled graphs [J]. Neural Networks, 2018, 108: 533-543.
- [9] HUANG H, CHEN Z, ZOU Y, et al. Channel prior convolutional attention for medical image segmentation [DB/OL]. (2023-06-08) [2024-04-10]. <http://arxiv.org/abs/2306.05196v1>.
- [10] WANG C Y, YE H I H, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information [DB/OL]. (2024-02-21) [2024-04-10]. <http://arxiv.org/abs/2402.13616v2>.
- [11] CHARLES R Q, HAO S, MO K, et al. PointNet: Deep learning on point sets for 3D classification and segmentation [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 77-85.
- [12] QI C R, YI L, SU H, et al. PointNet++: Deep hierarchical feature learning on point sets in a metric space [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 5105-5114.

- [13] ZHANG Z, DAI Y, SUN J. Deep learning based point cloud registration: An overview [J]. *Virtual Reality & Intelligent Hardware*, 2020, 2 (3) : 222-246.
- [14] AOKI Y, GOFORTH H, SRIVATSAN R A, et al. Pointnetlk: Robust & efficient point cloud registration using pointnet [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 7156-7165.
- [15] SARODE V, LI X, GOFORTH H, et al. PCRNet: Point cloud registration network using pointnet encoding[DB/OL]. (2019-08-21)[2024-04-10]. <https://arxiv.org/abs/1908.07906>.
- [16] LU W, WANG G, ZHOU Y, et al. Deepvcv: An end-to-end deep neural network for point cloud registration [C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019: 12-21.
- [17] PAIS G D, RAMALINGAM S, GOVINDU V M, et al. 3DRegNet: A deep neural network for 3D point registration[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 7191-7201.
- [18] WU Y, ZHANG Y, MA W, et al. Rornet: Partial-to-partial registration network with reliable overlapping representations[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023: 1-14.
- [19] BAUER D, PATTEN T, VINCZE M. Reagent: Point cloud registration using imitation and reinforcement learning [C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 14581-14589.
- [20] CHEN Z, CHEN H, GONG L, et al. UTOPIC: Uncertainty-aware overlap prediction network for partial point cloud registration [J]. *Computer Graphics Forum*, 2022, 41(7):87-98.
- [21] LI J, ZHANG C, XU Z, et al. Iterative distance-aware similarity matrix convolution with mutual-supervised point elimination for efficient point cloud registration[C]//*ECCV 2020: 16th European Conference*, 2020: 378-394.
- [22] WANG Y, SOLOMON J M. Prnet: Self-supervised learning for partial-to-partial registration [DB/OL]. (2019-10-27)[2024-04-10]. <https://arxiv.org/abs/1910.12240v2>.