arXiv:2302.02535v1 [cs.CV] 6 Feb 2023

PaRot：通过特征解耦和姿态恢复实现逐块旋转不变的网络

Dingxin Zhang*、**Jianhui Yu***、**Chaoyi Zhang**、**Weidong Cai**

澳大利亚悉尼大学计算机科学学院，dzha2344@uni.sydney.edu.au，{jianhui.yu、chaoyi.zhang、tom.cai}@sydney.edu.au

摘要

近年来，人们对点云分析的关注推动了面向三维模型的深度学习方法快速发展。然而，最先进的模型对旋转并不鲁棒；旋转在真实应用中仍是一个未知因素，并会损害模型性能。本文提出一种新颖的逐块旋转不变网络（PaRot），通过特征解耦实现旋转不变性，并为具有任意旋转的样本生成一致的预测。具体而言，我们设计了一个孪生训练模块，通过一对旋转，从不同尺度定义的局部块（例如局部几何和全局形状）中解耦出旋转不变性和等变性。然而，解耦得到的不变特征会丢失每个局部块固有的姿态信息。为解决这一问题，我们提出一种旋转不变的几何关系，利用等变信息恢复不同尺度局部块之间的相对姿态。利用姿态信息，我们提出一个分层模块，用于在三维形状学习中实现尺度内和尺度间的特征聚合。此外，我们还引入了一种姿态感知特征传播过程，其中嵌入了旋转不变的相对姿态信息。实验表明，我们的解耦模块能够提取高质量的旋转鲁棒特征，且所提出的轻量级模型在旋转三维物体分类和部件分割任务中取得了有竞争力的结果。项目页面已发布于：https://patchrot.github.io/。

1 引言

近年来，点云分析引起了研究人员的广泛关注。作为一种常见的三维表示形式，点云应用于无人驾驶和三维人脸识别等领域。近期的深度学习模型（Qi 等人，2017a、2017b）在经过良好对齐的点云分类和分割任务中展现出巨大潜力。然而，在真实场景中，三维物体通常会发生旋转，且方向角未知，这会对容易受旋转影响的深度学习模型造成显著影响。因此，使模型具有旋转不变性成为一项重要的研究课题。

使模型对旋转具有不变性成为一项重要的研究课题。

开创性工作尝试通过将形状变换为规范姿态来增强对旋转的鲁棒性（Qi 等人，2017a；Jaderberg 等人，2015），但无法实现对旋转保持一致的不变性。近期工作从局部几何构建旋转不变表示，并将其作为模型输入（Zhang 等人，2019；Kim、Park 和 Han，2020），从而在随机旋转下实现一致的行为。然而，如图1所示，旋转不变表示会丢失内在的位姿信息（即方向和位置）。为解决这一问题，Li 等人（2021b）和 Zhao 等人（2022）从全局三维形状中生成旋转不变特征，并通过简单拼接局部和全局尺度的特征来恢复全局位姿信息，却未充分探索如何更有效地利用位姿信息。此外，这些方法基于对旋转不变的手工设计特征来实现旋转不变性，这可能会限制模型性能。

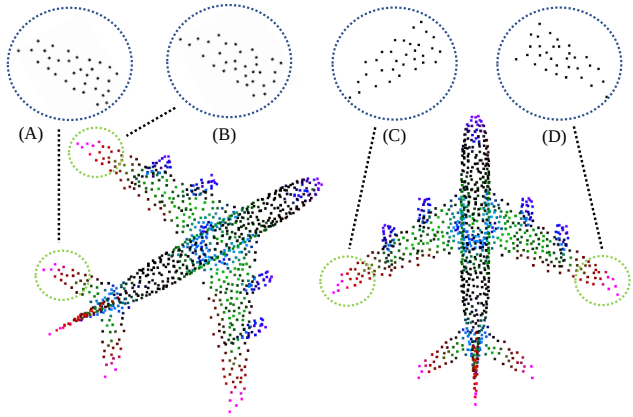


图1：生成逐补丁旋转不变特征时位姿信息丢失的示意图。我们展示了在两种不同方向下，一个飞机实例所学到的逐补丁旋转不变特征。如图所示，对于几何形状完全相同但姿态（即方向和全局位置）不同的区域 B 和区域 C，生成的特征相同。然而，尾部（A）、左翼（B、C）和右翼（D）局部块的特征会非常相似，因此丢失了区分它们差异的信息。

在这项工作中，我们借鉴了 Zhao 等人（2022）的思路，从局部和全局尺度获取旋转不变信息。具体而言，我们从局部几何和全局形状中提取点云局部块

*这些作者贡献相同。

。随后，我们应用特征解耦模块，为每个点云局部块引入成对旋转，通过孪生训练过程提取旋转不变形状内容和旋转等变方向信息。这样便能动态生成不变特征，从而增强特征表示。为恢复位姿信息，我们首先通过计算位置和方向差异，定义两个局部块之间的几何关系。然后，我们提出一种**Pa**tch-wise **Rot**ation-invariant 网络（PaRot），以局部块编码得到的几何关系和旋转不变形状内容为输入，进行尺度内聚合。此外，我们还考虑跨尺度过程，以局部和全局补丁编码得到的几何关系和特征为输入，挖掘全局上下文。此外，我们遵循 PointNet++ （Qi 等人，2017b）进行分割，仅需少量开销即可使用并改进特征传播模块。具体而言，我们提出一种姿态感知特征传播模块，用可学习的过程取代原先静态的、基于距离的特征插值，并利用编码后的几何关系保留旋转不变的相对位姿信息。更重要的是，在不同基准数据集上进行的大量实验展现了我们方法的优越性。

本工作的贡献总结如下：(1) 我们通过引入成对旋转，提出一种孪生训练模块，以解耦局部块级可学习的高质量旋转不变形状内容特征和旋转等变方向特征；(2) 我们定义一种旋转不变几何关系表示，用于恢复局部块之间的相对姿态信息，以指导跨尺度和尺度内学习；(3) 我们设计一种相对位姿感知特征传播方法，以实现更准确的旋转不变分割。

2 相关工作

点云深度学习。 以往针对三维点云的深度学习方法受益于二维卷积神经网络的先进发展。近期研究通过提取逐点特征，直接处理点集数据。PointNet（Qi 等人，2017a）提出了一种开创性的结构，利用多层感知机学习逐点空间特征，并通过最大池化实现置换不变性。后续工作在 PointNet 框架的基础上不断扩展，包括学习局部上下文以抽象不同尺度局部块的几何信息（Qi 等人，2017b；Zhao 等人，2019年；Yu 等人，2021）、开发卷积算子以更好地提取特征（Thomas 等人，2019年；Liu 等人，2019年），以及改进对称函数以促进特征聚合（Xiang 等人，2021；Chen 等人，2022）。然而，大多数方法对旋转敏感，当输入点云发生任意旋转时，其性能会急剧下降。

旋转等变性。 实现旋转鲁棒性的一种方法是确保所有中间层学习到的特征随输入相应地旋转。基于球面卷积的方法（Cohen 等，2018；Esteves 等人，2018；Rao、Lu 和 Zhou，2019）将点云转换到球谐域，并应用球面卷积来提取近似旋转等变的特征。基于张量场的网络（Poulenard和Guibas，2021年；Zhao 等，2020；Deng 等（2021））接收并输出能够保持严格旋转等变性的张量场特征。为实现旋转不变性，这些方法需要额外操作，将高层等变特征转换为不变形式，这会在训练过程中造成信息损失。在我们的工作中，等变方向特征用于在层次化几何学习过程中恢复相对位姿信息，从而减少信息损失。

旋转不变性。 另一种方法侧重于学习旋转不变特征。一种常见做法是在数据预处理阶段，将点云的笛卡尔坐标转换为手工设计的旋转不变表示。Zhang 等（2019）、Chen 等（2019）和 Xu 等（2021）在局部块内设计手工特征。这些方法在生成旋转不变特征时消除了局部块的位姿信息，从而损害几何学习过程。为解决这一问题，Zhang 等（2020）、Li 等（2021b）以及 Chen 和 Cong（2022）在手工设计表示时考虑了相对位姿信息。相邻补丁之间（Chen 和 Cong，2022），或局部补丁与全局形状之间（Zhang 等，2020；Li 等，2021b；Zhao 等，2022）的相对位姿被嵌入手工设计特征中。在我们的工作中，通过神经网络和任意旋转来提取逐补丁的旋转不变特征。同时，通过预测逐补丁方向来保留位姿信息，并通过计算尺度内与尺度间几何关系来恢复该信息。此外，我们在特征传播过程中嵌入几何关系，以提升分割性能。

孪生训练。 孪生训练能够实现特征解耦，是探索数据变化与相似性的一种高效技术。近期研究采用孪生训练来解耦旋转等变与旋转不变特征（Sun 等人，2021年；Gu 等人，2020年；Chen、Yang 和 Tao，2022年；Sajnani 等人，2022年）。然而，这些方法主要关注配准和重建任务，并将提取的等变特征用于规范化。相比之下，我们的方法利用等变方向矩阵构建几何关系，以恢复相对姿态，从而能够聚合通过尺度内和跨尺度学习获得的旋转不变特征。

3 方法

给定一个点云 $\mathbf{P} \in \mathbb{R}^{N\times 3}$，旋转鲁棒点云模型 $f(\cdot)$ 需要对施加于 $\mathbf{P}$ 的任意旋转 $\mathbf{R} \in \mathbb{R}^{3\times 3}$ 保持不变，并生成一致的预测结果：$f(\mathbf{P}) = f(\mathbf{PR})$。本文提出逐补丁旋转不变网络（PaRot）来实现这一目标。我们首先从形状描述符中解耦局部块级旋转不变性与等变性（第 3.1 节）。随后提取具有旋转不变性的全面几何表示，

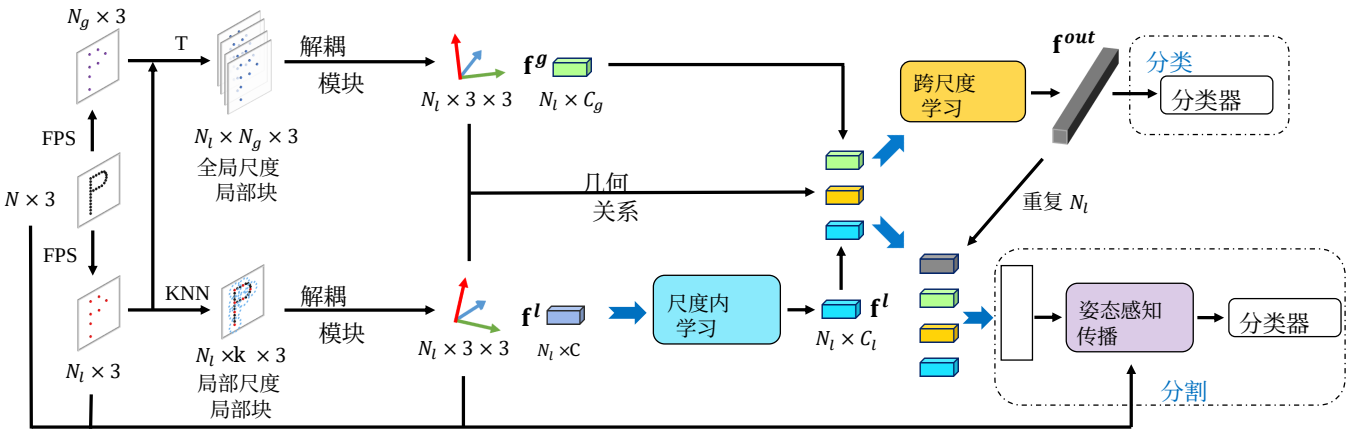


图 2：PaRot 用于三维分类和分割的整体架构，其中 KNN 和 T 分别表示 k 最近邻和 translation。我们通过最远点采样、KNN 和平移操作生成局部尺度补丁和全局尺度补丁。随后，我们分别将不同尺度的局部块解耦为形状内容信息和方向信息。学习到的方向用于确定几何关系，以恢复逐补丁相对姿态，并指导尺度内和尺度间不变特征学习以及姿态感知特征传播。

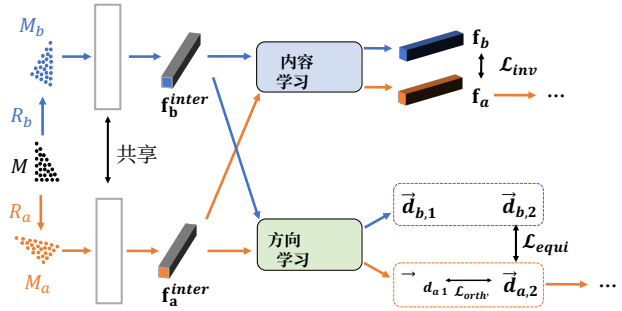


图 3：解耦模块的框架。输入局部块随机旋转两次，然后分别输入模块，作为分支 a 和 b。分支 a 的输出会进行分层编码，而分支 b的输出则用于辅助学习。

涵盖尺度内和尺度间关系，并嵌入几何关系，以保留不同局部块之间的姿态关系（第 3.2 节）。最后，我们提出一种带有几何关系的旋转不变特征传播模块，以在语义点标注过程中保持旋转不变性（第 3.3 节）。

3.1 旋转不变性与等变性解耦

受（Sun 等人，2021年）启发，我们提出了一种基于孪生训练流程的解耦模块，将潜在形状描述符分解为用于研究旋转不变性的旋转不变形状内容，以及用于保留位姿信息的旋转等变形状方向。

旋转不变内容学习。我们讨论了对于任意三维物体，在旋转作用下，其形状内容对随机旋转保持不变，并以局部块为单位提取这类旋转不变内容特征。具体而言，如图 3 所示，我们对输入点云局部块 $\mathbf{M} \in \mathbb{R}^{n\times 3}$引入一对任意旋转 $\mathbf{R}_a$ 和$\mathbf{R}_b$，得到两个随机旋转后的局部块 $\mathbf{M}_a$ 和$\mathbf{M}_b$。我们采用轻量级 PointNet 网络（Qi 等人，2017a），并在$\mathbf{M}_a$ 和$\mathbf{M}_b$ 之间共享该网络，以编码几何信息，得到两个中间形状描述符 $\mathbf{f}_a^{inter}$和$\mathbf{f}_b^{inter}$。随后，我们应用多层感知机（MLP），从潜在形状描述符中解耦出旋转不变形状内容$\mathbf{f}_a$和$\mathbf{f}_b$，并通过最小化$\mathbf{f}_a$与$\mathbf{f}_b$之间的特征距离来实现旋转不变性。因此，我们定义了一个旋转不变 损失函数$\mathcal{L}_{inv}$，以确保这两个特征在旋转下具有较高的相似性，其表示为:

$$\mathcal{L}_{inv} = \|\mathbf{f}_a - \mathbf{f}_b\|_2^2, \tag{1}$$

其中，$\|\cdot\|_2^2$ 表示 L2范数。

旋转等变方向学习。我们在此讨论的一个显而易见的直觉是，对于任何 3D 局部块，都存在一种固有方向，它对于恢复姿态至关重要，并由旋转矩阵表示，且天然具有旋转等变性。为此，我们借鉴 FS-Net（Chen等人（2021））的思路，从潜在形状描述符中提取方向向量。具体而言，如$\vec{\ }$ $\vec{\ }$图 4所示，根据潜在形状描述符 $\mathbf{F}$，为每个分支预测两个相互垂直的方向向量$\mathbf{d}_1$ 和 $\mathbf{d}_2$；当对两个分支执行相同操作时，我们省略用于标记孪生分支的下标（即 a 和 b）。为确保等变性，我们引入一个旋转等变 损失函数 $\mathcal{L}_{equi}$，用于约束学习到的方向向量，如下所示:

$$\mathcal{L}_{equi} = \left\|\vec{\mathbf{d}}_{a,1} - \mathbf{R}_b^{-1}\mathbf{R}_a\vec{\mathbf{d}}_{b,1}\right\|_2^2 + \left\|\vec{\mathbf{d}}_{a,2} - \mathbf{R}_b^{-1}\mathbf{R}_a\vec{\mathbf{d}}_{b,2}\right\|_2^2, \tag{2}$$

其中$\mathbf{R}_b^{-1}\,\mathbf{R}_a$ 是一个旋转矩阵，用于将$\mathbf{M}_a$ 变换为$\mathbf{M}_b$。通过这种方式，方向向量被约束为仅包含

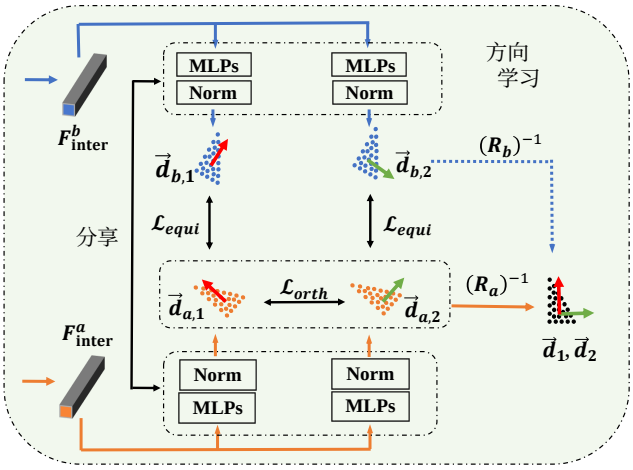


图4：旋转等变方向学习示意图。分支 a 的输出方向经旋转$\mathbf{R}_a^{-1}$转回后，用作原始局部块$\mathbf{M}$的预测方向。

局部图像块的方向信息，随后用于姿态信息嵌入。此外，为了保留$\mathbf{d}_1$ 和$\mathbf{d}_2$ 作为旋转矩阵列向量时的正交性，我们设计了一个损失函数 $\mathcal{L}_{orth}$，表示如下：

$$\mathcal{L}_{orth} = \left\| \vec{\mathbf{d}}_1^{\top} \vec{\mathbf{d}}_2 \right\|_2^2. \tag{3}$$

需要说明的是，我们的方向矩阵是基于$\mathbf{M}$的随机旋转姿态学习得到的，这与初始输入$\mathbf{M}$的姿态不同。由于$\mathbf{R}_a$和$\mathbf{R}_b$ 仅用于解耦，我们需要消除它们的影响，并通过将预测方向向量变换回去，得到$\mathbf{M}$ 的初始方向$\mathbf{O}$：$[\vec{\mathbf{d}}_1, \vec{\mathbf{d}}_2] = \mathbf{R}_a^{-1}[\vec{\mathbf{d}}_{a,1}, \vec{\mathbf{d}}_{a,2}] = \mathbf{R}_b^{-1}[\vec{\mathbf{d}}_{b,1}, \vec{\mathbf{d}}_{b,2}]$。

As $\vec{\mathbf{d}}_1$ 和$\vec{\mathbf{d}}_2$ 确保不平行后，我们可以简单地定义第三个方向向量$\mathbf{d}_3 = \mathbf{d}_1 \times \mathbf{d}_2$，它与$\mathbf{d}_1$ 和$\mathbf{d}_2$都正交。因此，我们将针对初始输入$\mathbf{O}$ 学习得到的方向矩阵表示为方向向量的拼接：$\mathbf{O} = [\vec{\mathbf{d}}_1, \vec{\mathbf{d}}_2, \vec{\mathbf{d}}_3]$,，其中 $[\cdot]$ 表示拼接。最后，由于两个分支共享所有可学习的多层感知机，我们学习到的旋转不变形状内容和方向相同。遵循孪生训练流程的常规实现方式（Sun 等人，2021年），我们在后续模型推理中使用分支 a 的输出，并且仅为分支 a 计算 $\mathcal{L}_{orth}$。

3.2 利用几何关系进行尺度内与尺度间旋转不变性学习

为进行旋转不变性学习，我们提出了一种几何关系嵌入网络，通过跨局部块内和跨尺度学习聚合旋转不变特征。其中，利用局部块方向矩阵和全局绝对位置计算局部块之间的几何关系，以恢复相对位姿信息。

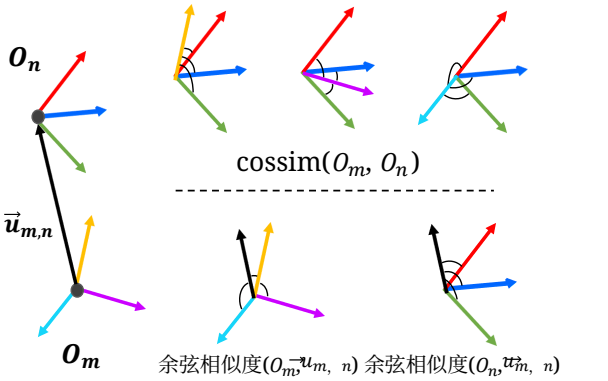


图5：几何关系表示中使用的角度。对于两个局部块$\mathbf{M}_m$ 和 $\mathbf{M}_n$，我们用不同的颜色表示其方向矩阵 $\mathbf{O}_m$和$\mathbf{O}_n$ 中的不同向量。

几何关系表示。 局部块的位姿信息对于几何特征学习至关重要，但由于方向和全局位置对于旋转具有等变性，因此无法将其编码为旋转不变的形式。为了保留这类位姿信息，我们必须考虑不同局部块之间具有旋转不变性的几何属性，从而维持所学习到的旋转不变性。为实现这一目标，我们采用相对几何关系，即局部块之间的相对角度和距离，在特征学习过程中保留原始三维形状的位姿信息。具体而言，给定两个局部块$\mathbf{M}_m$ 和$\mathbf{M}_n$ ，以及对应的参考点$\mathbf{p}_m$ 和$\mathbf{p}_n$，根据我们的解耦模块可以学习它们的方向矩阵$\mathbf{O}_m = [\vec{\mathbf{d}}_1^m, \vec{\mathbf{d}}_2^m, \vec{\mathbf{d}}_3^m]$和$\mathbf{O}_n = [\vec{\mathbf{d}}_1^n, \vec{\mathbf{d}}_2^n, \vec{\mathbf{d}}_3^n]$。因此，两个局部块之间的几何关系 $\mathcal{G}$（$\mathbf{M}_m$,$\mathbf{M}_n$）定义为：

$$\begin{aligned}\mathcal{G}(\mathbf{M}_m, \mathbf{M}_n) = [&\text{dist}(\mathbf{p}_m, \mathbf{p}_n), \text{cossim}(\mathbf{O}_m, \mathbf{O}_n),\\ &\text{cossim}(\mathbf{O}_m, [\vec{\mathbf{u}}_{mn}]),\\ &\text{cossim}(\mathbf{O}_n, [\vec{\mathbf{u}}_{mn}])],\end{aligned} \tag{4}$$

其中，dist(·)计算两点之间的欧几里得距离，cossim(·)计算第一个矩阵的所有列向量与第二个矩阵的所有列向量之间的余弦相似度，而 $\vec{\mathbf{u}}_{mn} = \mathbf{p}_n - \mathbf{p}_m$ 是两个局部块参考点之间的向量，方向从 $\mathbf{p}_n$ 指向 $\mathbf{p}_m$。如图5所示， $\mathcal{G}$ ($\mathbf{M}_m$,$\mathbf{M}_n$)包含1个距离参数和 $9+3+3=15$ 个角度参数。

局部尺度特征解耦。为生成局部块，我们从原始点集 $\mathbf{P}$ 中通过最远点采样选取 N_ℓ 个点，作为局部查询点集$\mathbf{Q} \in \mathbb{R}^{N_\ell \times 3}$。对于$\mathbf{Q}$ 中的任意点$\mathbf{q}_i \in$ ，将其作为参考点，在$\mathbf{P}$上执行 k-近邻搜索，从而得到共 N_ℓ 个局部块。随后，我们为每个局部块生成一个局部参考坐标系，并对局部块进行逐补丁平移，以确保局部块$\mathbf{M}_i^\ell$ 的参考点$\mathbf{q}_i$ 与参考坐标系原点重合：

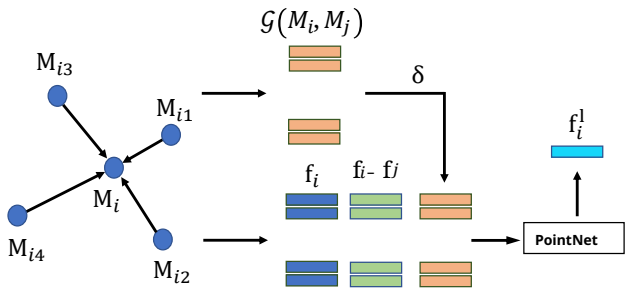


图6：尺度内边缘卷积中嵌入几何关系的示意图。

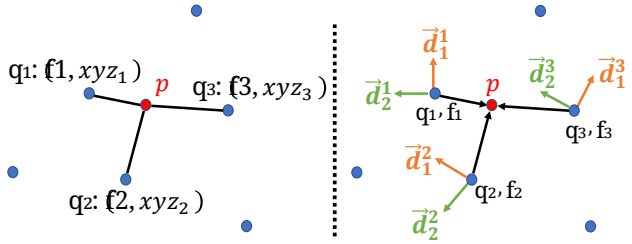


图7：左：典型的特征传播，即基于距离的插值。右：相对位姿感知的特征传播过程。为简洁起见，未显示每个局部块的第三方向$\mathbf{d}_3$。

点：

$$\mathbf{M}_i^\ell = [q_j - q_i]_{j:j\in\mathcal{N}_\mathbf{P}(i)}, \tag{5}$$

其中，$\mathcal{N}_\mathbf{P}(i)$ 表示$\mathbf{q}_i$ 在$\mathbf{P}$中的邻近点。随后，我们以局部块为单位进行形状特征解耦，提取旋转不变的内容特征$\mathbf{f}_i$ 以及旋转等变的方向矩阵$\mathbf{O}_i$。

尺度内相对位姿感知特征学习。为捕捉较大感受野内局部块的几何信息，我们对$\mathbf{f}_i$ 应用边缘卷积（Wang 等人，2019），以进一步编码并细化旋转不变特征。然而，如前所述，解耦后的旋转不变形状特征不包含位姿信息，仅对形状内容特征进行卷积会导致几何信息损失。如图6所示，我们通过提出一种几何关系编码操作来解决这一问题，该操作采用线性映射函数 $\delta(\cdot)$，并将编码嵌入卷积中，聚合局部几何特征$\mathbf{f}_i^\ell$，如下所示：

$$\mathbf{f}_i^\ell = \mathcal{A}_{j:j\in\mathcal{N}(i)}\Big(\mathrm{MLP}\left([\mathbf{f}_j, \mathbf{f}_j - \mathbf{f}_i, \delta(\mathcal{G}(\mathbf{M}_i, \mathbf{M}_j))]\right)\Big), \tag{6}$$

其中 $\mathcal{A}(\cdot)$ 表示最大池化。通过这种方式，我们显式地将不同局部块之间的相对位姿信息嵌入旋转不变性学习过程中，从而获得更具区分性的特征表示。

全局尺度特征解耦。对于全局尺度特征解耦，我们将$\mathbf{P}$ 下采样为新的点集 $\mathbf{G} \in \mathbb{R}^{N_g\times 3}$，其中包含 N_g 个点。我们进一步使用局部参考点 $\mathbf{q}_i \in \mathbf{Q}$ 生成 N_ℓ 个相应的全局补丁。每个全局补丁$\mathbf{M}_i^g$ 都将包含$\mathbf{G}$ 中的所有点，并执行与公式 5 中相同的平移操作，使参考点居中并生成参考坐标系。这样，我们便将全局上下文信息和相应的局部补丁位置信息整合到每个全局补丁中。随后，我们将$\mathbf{M}_i^g$ 解耦为全局尺度旋转不变特征$\mathbf{f}_i^g$ 和方向矩阵$\mathbf{O}_i^g$，其中包含全局补丁的形状信息以及参考点的位置信息。这些信息可用于增强旋转不变表示的特征区分度。此外，值得一提的是，无需对全局尺度补丁执行尺度内学习，

因为全局补丁覆盖所有点，在保留全局上下文的同时，也保留了输入形状。

尺度间旋转不变学习。由于局部分支和全局分支使用相同的下采样点集$\mathbf{Q}$作为参考点集，我们将讨论如何融合局部尺度特征$\mathbf{f}_i^\ell$和全局尺度特征$\mathbf{f}_i^g$，以结合局部几何与全局上下文信息。为此，我们将$\mathbf{M}_i^\ell$ 和$\mathbf{M}_i^g$ 之间的几何关系，以及$\mathbf{f}_i^\ell$和$\mathbf{f}_i^g$作为模块输入，并直接输出一个相对位姿感知的融合特征：

$$\mathbf{f}_i^{out} = \mathcal{A}_{i:i\in\mathbf{Q}}\Big(\mathrm{MLP}\left([\mathbf{f}_i^\ell, \mathbf{f}_i^g, \delta(\mathcal{G}(\mathbf{M}_i^\ell, \mathbf{M}_i^g))]\right)\Big). \tag{7}$$

计算几何关系 $\mathcal{G}(\mathbf{M}_i^\ell, \mathbf{M}_i^g)$ 时，我们将原点 $\vec{\mathbf{0}}$ 定义为全局局部块的参考点；否则，$\mathbf{M}_i^\ell$和$\mathbf{M}_i^g$ 的参考点将相同。

3.3 旋转不变的姿态感知特征传播模块

与 PointNet++ （Qi 等人，2017b）提出的典型方法不同（见图 7 左图）：该方法基于点之间的距离和最近邻 k 的点的绝对位置，将特征从下采样点传播至原始点。我们提出了一种姿态感知特征传播模块，该模块摒弃基于距离的插值，并利用相对几何关系进行姿态信息嵌入。如图 7 右图所示，给定一个特征为 $\mathbf{f}_j^q$ 的点 $\mathbf{q}_j \in \mathbf{Q}$，以及一个特征为 $\mathbf{f}_i^p$ 的点 $\mathbf{p}_i \in \mathbf{P}$，我们按照第 3.2 节的方法，利用 $\mathbf{q}_j$ 与 $\mathbf{p}_i$ 之间的相对几何关系，为特征传播模块嵌入信息。具体而言，定义方向向量 $\vec{\mathbf{u}}_{ij:j\in\mathcal{N}_\mathbf{Q}(i)} = \mathbf{q}_j - \mathbf{p}_i$，其方向由 $\mathbf{p}_i$ 指向 $\mathbf{q}_j$，其中 $\mathcal{N}_\mathbf{Q}(i)$ 是 $\mathbf{p}_i$ 在 Q 中的邻域。由于 $\mathbf{q}_j$ 关联着一个旋转等变方向矩阵 $\mathbf{O}_j$，我们因此按照公式 4，以相同方式定义 $\mathbf{p}_i$ 与其邻居 $\mathbf{q}_j$ 之间的几何关系：

$$\mathcal{G}(\mathbf{p}_i, \mathbf{q}_j) = \left[\mathrm{dist}(\mathbf{p}_i, \mathbf{q}_j), \mathrm{cossim}(\mathbf{O}_j, [\vec{\mathbf{u}}_{ij}]_{j\in\mathcal{N}_\mathbf{Q}(i)})\right]. \tag{8}$$

随后，通过 $\delta(\cdot)$ 对几何关系 $\mathcal{G}(\mathbf{p}_i, \mathbf{q}_j)$ 进行编码，并将其与邻近点特征 $\mathbf{f}_j^q$ 以及来自原始点的跳跃连接点特征 $\mathbf{f}_i^p$ 拼接。融合后的

旋转敏感	输入	z/z	z/SO3	SO3/SO3
PointNet（Qi 等人，2017a）	pc	85.9	19.6	74.7
PointNet++（Qi 等人，2017b）	pc	89.3	28.6	85.0
PointNet++(Qi 等人，2017b)	pc+n	91.8	18.4	77.4
DGCNN（Wang 等人，2019年）	pc	**92.2**	20.6	81.1
旋转鲁棒	输入	z/z	z/SO3	SO3/SO3
球面卷积神经网络（Esteves 等人，2018年）	体素	88.9	76.9	86.9
SFCNN（Rao、Lu 和 Zhou，2019年）	pc	91.4	84.8	90.1
RI-Conv（Zhang 等人，2019年）	pc	86.5	86.4	86.4
ClusterNet（Chen 等人，2019年）	pc	87.1	87.1	87.1
RI-GCN（Kim、Park 和 Han，2020）	pc	89.5	89.5	89.5
GCANet（Zhang等，2020年）	pc	89.0	89.1	89.2
RIF（Li 等（2021b））	pc	89.4	89.4	89.3
SGMNet（Xu 等（2021））	pc	90.0	90.0	90.0
TFN（Poulenard和Guibas，2021年）	pc	87.6	87.6	87.6
Li 等（2021a）	pc	90.2	90.2	90.2
VN-DGCNN（Deng 等（2021））	pc	89.5	89.5	90.2
OrientedMP（Luo 等人，2022）	pc	88.4	88.4	88.9
ELGANet（Gu 等，2022）	pc	90.3	90.3	90.3
PaRot	pc	90.9	**91.0**	**90.8**

表 1：ModelNet40 上的分类性能。“pc”和“n”分别表示输入数据类型为原始三维坐标和法向量。

特征经过可学习的多层感知机处理，并进一步通过求和进行聚合，以更新原始特征 $\mathbf{f}_i^p$。整个过程可表示如下：

$$\mathbf{f}_i = \sum_{j=1}^{k} \mathrm{MLP}\left([\mathbf{f}_i^p, \mathbf{f}_j^q, \delta(\mathcal{G}(\mathbf{p}_i, \mathbf{q}_j))]\right). \tag{9}$$

由公式 9 可知，与 PointNet++ 不同，我们在传播过程中忽略点$\mathbf{p}_i$ 的旋转敏感全局 xyz 位置，从而确保我们的模块具有旋转不变性。此外，额外引入邻近点特征$\mathbf{f}_j^q$ 进一步提升了特征表示。

4 实验

我们通过三维点云分类和分割任务评估模型，分析其旋转鲁棒性，并与其他具有代表性的方法进行性能比较，同时展示实验结果。此外，我们还分析了所提出模块的效率以及模型的复杂度。我们遵循（Esteves 等，2018）的评估协议：随机将训练和测试对象绕 z 轴旋转（z/z）；将训练对象绕 z 轴旋转，同时对测试对象施加任意旋转（z/SO3）；对训练和测试数据均施加任意旋转（SO3/SO3）。

对于分类，我们将 N_ℓ 设为 256，将 N_g 设为 32。使用 k-NN 生成局部尺度块并搜索用于内部学习的邻居时，我们将局部尺度块的数量 k_ℓ 设为 64，将用于尺度内学习的邻居数量 k_{intra} 设为 32。分割任务的设置相同，但 N_g 和 k_{intra} 分别改为 64 和 16。

4.1 形状分类

我们在合成数据集 ModelNet40（Wu et al. 2015）和真实世界数据集 ScanObjectNN（Uy et al. 2019）上测试模型的分类能力。ModelNet40

方法	z/SO3	SO3/SO3	Δ
PointNet（Qi 等人，2017a）	16.7	54.7	38.0
PointNet++（Qi 等人，2017b）	15.0	47.4	32.4
DGCNN（Wang 等人，2019）	17.7	71.8	54.1
RI-Conv（Zhang 等人，2019）	78.4	78.1	0.3
RI-GCN（Kim、Park 和 Han，2020）	80.5	80.6	0.1
RIF（Li 等（2021b））	79.8	79.9	0.1
Li 等⋆（2021a）	79.3	79.6	0.3
VN-DGCNN⋆（Deng 等（2021））	77.8	76.0	1.8
OrientedMP⋆ (Luo 等人，2022)	76.7	77.2	0.6
PaRot	**82.1**	**82.6**	0.5

表 2：ScanObjectNN$OBJ\ BG$ 数据集在 z/SO3 和 SO3/SO3 设置下的分类准确率。 Δ 表示 z/SO3 与 SO3/SO3 之间的绝对差值。⋆ 表示基于官方实现复现的结果。

是点云分析中最常用的数据集，其中包含从 40 类 CAD 模型中采样得到的 12,311 个预对齐点云形状。官方版本将数据集划分为 9,843 个训练样本和 2,468 个测试样本。ScanObjectNN 包含从 2,902 个真实世界物体扫描得到的 15,000 个不完整物体。为了更好地探究扫描样本对噪声的鲁棒性，我们使用包含背景噪声的 $OBJ\ BG$ 子集进行评估。我们从每个样本中采样 1024 个点作为模型输入。

在训练流程中，我们在解耦模块的每个局部块中引入随机旋转。在测试阶段不进行局部块级旋转，因为最终模型的解耦特征具有旋转鲁棒性，因此局部块级旋转不会影响性能。

我们根据表 1 和表 2 中报告的分类准确率，分别将模型与 ModelNet40 和 ScanObjectNN 上的代表性模型进行比较。结果表明，我们的方法在三种设置下均取得了良好效果，并具有较强的旋转鲁棒性：z/z、z/SO3 和 SO3/SO3 之间的绝对准确率差异不超过 0.5%。

4.2 部件分割

对于形状部件分割任务，我们在 ShapeNetPart 数据集（Yi 等人，2016）上验证了模型。该数据集包含 16,880 个合成样本，其中 14,006 个用于训练，2,874 个用于测试。数据集包含 16个类别，每个类别标注了 2 至 6 个部件，因此总共有 50 个部件标注标签。我们采样 2048 个点作为模型输入。表3报告并比较了 z/SO3 条件下各类别的平均交并比（mIoU）以及 16个类别的平均 mIoU。此外，我们在图 8 中展示了不同方向下不同物体的分割结果。显然，我们的模型能够稳健地分割任意旋转下的点云。

4.3 方法分析

组件研究。我们进行了消融研究，以探究所提出模型中各个模块的有效性

方法	类别平均交并比	aero	bag	cap	car	椅子	耳机	吉他	刀	灯	笔记本电脑	摩托车	马克杯	手枪	火箭	滑板	桌子
PointNet（Qi 等人，2017a）	37.8	40.4	48.1	46.3	24.5	45.1	39.4	29.2	42.6	52.7	36.7	21.2	55.0	29.7	26.6	32.1	35.8
PointNet++（Qi 等人，2017b）	48.3	51.3	66.0	50.8	25.2	66.7	27.7	29.7	65.6	59.7	70.1	17.2	67.3	49.9	23.4	43.8	57.6
DGCNN（Wang 等人，2019年）	37.4	37.0	50.2	38.5	24.1	43.9	32.3	23.7	48.6	54.8	28.7	17.8	74.4	25.2	24.1	43.1	32.3
RI-Conv（Zhang 等人，2019年）	75.3	80.6	80.0	70.8	68.8	86.8	70.3	87.3	84.7	77.8	80.6	57.4	91.2	71.5	52.3	66.5	78.4
GCANet（Zhang等，2020年）	77.2	80.9	**82.6**	81.0	70.2	88.4	70.6	87.1	**87.2**	**81.8**	78.9	58.7	91.0	77.9	52.3	66.8	80.3
TFN（Poulenard和Guibas，2021年）	76.7	80.9	75.2	81.9	73.8	89.0	61.0	90.8	83.0	76.9	80.2	58.5	92.8	76.3	54.0	**74.5**	79.1
Li 等（2021a）	74.1	81.9	58.2	77.0	71.8	**89.6**	64.2	89.1	85.9	80.7	**84.7**	46.8	89.1	73.2	45.6	66.5	**81.0**
VN-DGCNN* (Deng 等（2021）)	75.3	81.1	74.8	72.9	73.8	87.8	55.9	**91.4**	83.8	80.2	84.4	44.5	92.8	74.6	**57.2**	70.2	78.9
PaRot	**79.2**	**82.7**	79.2	**82.3**	**75.3**	89.4	**73.9**	91.1	85.6	81.0	79.5	**65.3**	**93.9**	**79.2**	55.0	72.4	79.5

表3：在 z/SO3 条件下，ShapeNetPart 数据集的逐类分割结果及类别平均交并比。其中，类别平均交并比表示16个类别的平均交并比。

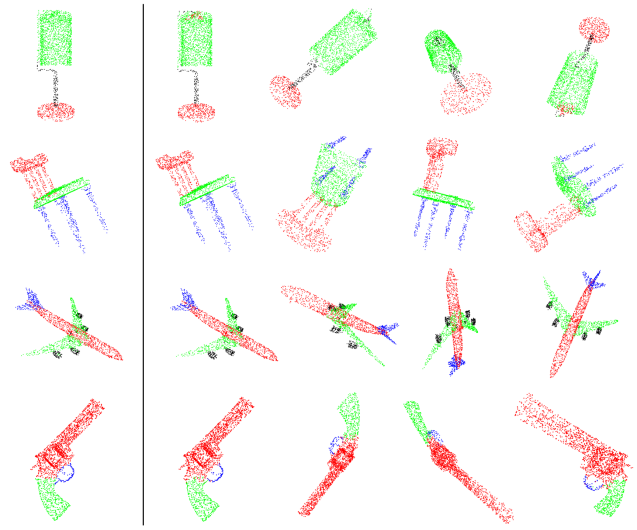


图8：ShapeNet 部件分割结果的可视化。最左侧列为选定方向下的真实标注，左起第二列为模型在所选方向下的分割结果，其余各列为随机旋转样本的结果。

并在表 4 中报告结果。我们还将几何表示$\mathcal{G}(M_m,M_n)$拆分为两部分：cossim(O_m,O_n)和[dist(p_m,p_n),cossim(O_m,[$\vec{u}_{mn}$]),cossim(O_n,[$\vec{u}_{mn}$])]，以研究用于姿态恢复的特征。其中，cossim(O_m, O_n)仅包含相对方向信息，缺少位置信息；而[dist(p_m,p_n),cossim(O_m,[$\vec{u}_{mn}$]),cossim(O_n,[$\vec{u}_{mn}$])]能够检测相对位置，但对于绕向量 $\vec{u}_{mn}$的旋转存在歧义。

模型 A 是我们的基线，仅提取局部尺度旋转不变形状内容特征用于分类。当我们采用尺度内学习模块（A→B），但不进行相对姿态恢复操作时，模型会遭遇姿态信息丢失问题；不过，它能够聚合来自更大尺度局部块的特征，并略微提高准确率。该模型与 RI-Conv 和 ClusterNet 的原理及性能相似。尺度间学习模块旨在利用全局上下文。然而，在添加尺度间学习模块时，

模型	局部	内部-	之间-	姿态恢复		z/SO3
				方向。	位置	
A	✓					84.3
B	✓	✓				87.4
C	✓		✓			70.9
D	✓	✓	✓			58.6
E	✓		✓	✓	✓	89.6
F	✓	✓		✓	✓	89.5
G	✓	✓	✓	✓		90.4
H	✓	✓	✓		✓	90.5
I	✓	✓	✓	✓	✓	91.0

表4：ModelNet40 在 z/SO3 下的组件研究结果。

若不采用几何关系表示（A→C），模型性能便会下降，因为网络无法发现逐块特征与全局上下文之间的关系。使用两个学习模块（B→D）时，性能下降幅度大于模型 A→C，因为 B 比 A 更深，也更容易受到深层添加的噪声特征影响。姿态恢复策略能够保留姿态信息、引导学习过程，并持续提升性能（B→F、C→E、D→I）。

此外，模型 G、H 和 I 表明，几何表示的两个部分均可恢复部分姿态信息并提高准确率，而完整表示能够恢复更多信息，并取得最佳性能。

特征传播分析。为了验证所提出的姿态感知几何关系嵌入式特征传播方法，我们在 z/SO3 ShapeNet Part 分割任务中，将所有实例上的平均交并比（I.mIoU）和各类别的平均平均交并比（C.mIoU）作为指标，与 PointNet++ 中由 Qi 等人，2017b 提出的典型插值策略进行比较。PointNet++ 中嵌入的 xyz 坐标位置被移除，否则会破坏旋转不变性。如表 5 所示，我们提出的传播方法优于插值策略。此外，插值策略根据距离倒数对最近邻局部块的特征进行加权平均；距离较远的局部块对生成新特征的贡献很小，增加邻居数量也无法提升性能。在我们的方法中，目标点与邻居局部块之间的几何关系将分别进行编码，因此，性能可以

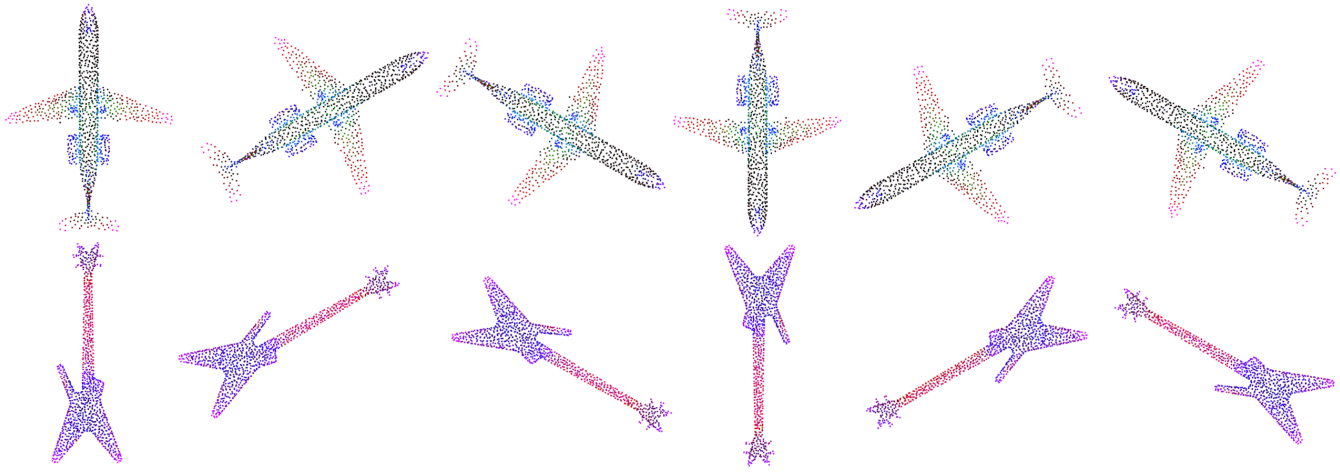


图 9：解耦后的 旋转不变特征的可视化。从左到右，每个样本都绕 z 轴旋转 60° is.

邻居数量	插值		姿态感知	
	实例平均交并比	类别平均交并比	实例平均交并比	类别平均交并比
3	81.1	76.6	82.1	77.6
5	80.9	76.7	82.5	78.3
7	80.8	76.8	82.8	78.1
9	80.8	76.7	83.0	79.1
11	80.8	76.7	82.9	**79.2**
13	80.6	76.4	**83.2**	79.0

表 5：传播分析。ShapeNetPart 上 z/SO(3) 部件分割的实例平均交并比和类别平均交并比结果。最左列表示用于特征传播的邻居局部块数量。

方法	参数。	浮点运算次数	Acc.
PointNet++（Qi 等人，2017b）	1.41M	863M	85.0
DGCNN（Wang 等人，2019年）	1.72M	2449M	81.1
RI-GCN（Kim、Park 和 Han，2020）	4.19M	1237M	89.5
RIF（Li 等（2021b））	2.36M	6535M	89.4
Li 等（2021a）	2.91M	3747M	90.2
VN-DGCNN（Deng 等（2021））	2.77M	3183M	89.5
PaRot 训练	1.55M	2091M	91.0
PaRot 测试版	1.55M	1431M	91.0

表 6：ModelNet40 上的模型大小、计算复杂度和 z/SO3 准确率对比。测试版本移除了不必要的辅助网络模块。

通过增加邻近查询的数量进一步提升。

模型复杂度。 得益于高效的层次结构和几何关系编码技术，我们的旋转鲁棒模型以较小的模型规模和较低的计算成本实现了高性能。在开发过程中，我们避免计算动态图（Wang 等人，2019），并平衡网络的宽度和深度。此外，测试时无需采用孪生过程，因此可以将其移除，以进一步降低计算复杂度。如表 6 所示，与 VN-DGCNN 相比，所提出的模型仅使用 55% 的参数量和 45% 的浮点运算次数，准确率却提高了 1.5%。

解耦旋转不变特征可视化。为了考察我们的解耦模块在任意方向下为局部块提取一致特征的效果，我们对两个不同物体（即飞机和吉他）的特征响应进行可视化，这两个物体均绕 z 轴旋转。具体而言，我们选取从局部分支提取的解耦旋转不变特征中的三个通道，并将其作为 RGB 通道进行可视化。我们以 P 中的所有点为参考点，生成 1024 个局部尺度块，并使用保存的最佳分割模型进行解耦。图 9 中的结果表明，对于形状内容不同的局部块，学习到的特征各不相同，同时对旋转保持不变。

5 结论

在本文中，我们提出了 PaRot，这是一种用于三维点云识别的新型旋转不变学习模型。给定一个点云，我们通过在孪生训练流程中引入成对旋转，为局部尺度块和全局尺度图块解耦旋转不变的形状内容特征与旋转等变方向。为了在保持学习旋转不变性的同时恢复旋转敏感的姿态信息，我们利用逐块方向矩阵计算几何关系，以表示局部块之间的相对姿态。几何关系用于指导尺度内和尺度间的特征聚合。基于利用几何关系恢复姿态信息的相同思想，我们进一步设计了一种旋转不变的特征传播方法，从而提高了模型的分割准确率。大量实验表明了我们模型的有效性和高效性。

参考文献

Chen, C.；Li, G.；Xu, R.；Chen, T.；Wang, M.；以及 Lin, L. 2019年。ClusterNet：具有严格旋转不变表示的点云分析深度层次聚类网络。发表于*CVPR*。Chen, J.；Kakillioglu, B.；Ren, H.；以及 Velipasalar, S. 2022。既然可以回收，为什么要丢弃？：用于三维点云分析的回收最大池化模块。发表于*CVPR*。Chen, R.；以及 Cong, Y. 2022。问题出在姿态上：通过姿态感知卷积实现无歧义的三维旋转不变学习。发表于*CVPR*。Chen, W.；Jia, X.；Chang, H. J.；Duan, J.；Shen, L.；以及 Leonardis, A. 2021。FS-Net：采用解耦旋转机制的类别级六维物体姿态估计快速基于形状的网络。发表于*CVPR*。Chen, Z.；Yang, F.；以及 Tao, W. 2022。DetarNet：通过孪生网络解耦平移与旋转，实现点云配准。发表于*AAAI*。Cohen, T. S.；Geiger, M.；K¨ohler, J.；以及 Welling, M. 2018。球面卷积神经网络。发表于*ICLR*。Deng, C.；Litany, O.；Duan, Y.；Poulenard, A.；Tagliasacchi, A.；以及 Guibas, L. J. 2021。向量神经元：SO(3)等变网络的通用框架。发表于*ICCV*。Esteves, C.；Allen-Blanchette, C.；Makadia, A.；以及 Daniilidis, K. 2018。利用球面卷积神经网络学习 SO(3) 等变表示。发表于*ECCV*。Gu, J.；Ma, W.；Manivasagam, S.；Zeng, W.；Wang, Z.；Xiong, Y.；Su, H.；以及 Urtasun, R. 2020。野外弱监督三维形状补全。发表于*ECCV*。Gu, R.；Wu, Q.；Li, Y.；Kang, W.；Ng, W.；以及 Wang, Z. 2022。面向旋转不变点云表示的增强局部与全局学习。发表于*MultiMedia*。Jaderberg, M.；Simonyan, K.；Zisserman, A.；等。2015。空间变换网络。发表于*NeurIPS*。Kim, S.；Park, J.；以及 Han, B. 2020。面向三维点云的旋转不变局部到全局表示学习。发表于*NeurIPS*。Li, F.；Fujiwara, K.；Okura, F.；以及 Matsushita, Y. 2021a。深入探究旋转不变深度点云分析。发表于*ICCV*。Li, X.；Li, R.；Chen, G.；Fu, C.；Cohen-Or, D.；以及 Heng, P. 2021b。深度点云分析的旋转不变框架。发表于 *TVCG*。Liu, Y.；Fan, B.；Xiang, S.；以及 Pan, C. 2019年。用于点云分析的关系形状卷积神经网络。发表于*CVPR*。Luo, S.；Li, J.；Guan, J.；Su, Y.；Cheng, C.；Peng, J.；以及 Ma, J. 2022。通过学习用于消息传递的方向进行等变点云分析。发表于*CVPR*。Poulenard, A.；以及 Guibas, L. J. 2021。张量场网络旋转等变非线性的函数方法。发表于*CVPR*。Qi, C. R.；Su, H.；Mo, K.；以及 Guibas, L. J. 2017a。Point-Net：基于点集的三维分类与分割深度学习

. 在*CVPR* .

Qi, C. R.; Yi, L.; Su, H.; and Guibas, L. J. 2017b. Point-Net++: 度量空间中点集的深度层次特征学习。发表于*NeurIPS*。Rao, Y.; Lu, J.; and Zhou, J. 2019年. 用于点云识别的球形分形卷积神经网络。发表于*CVPR*。Sajnani, R.; Poulenard, A.; Jain, J.; Dua, R.; Guibas, L. J.; and Sridhar, S. 2022. ConDor：部分形状三维姿态的自监督规范化。发表于*CVPR*。Sun, W.; Tagliasacchi, A.; Deng, B.; Sabour, S.; Yazdani, S.; Hinton, G. E.; and Yi, K. M. 2021. 规范胶囊网络：规范姿态下的自监督胶囊。发表于*NeurIPS*。Thomas, H.; Qi, C. R.; Deschaud, J.; Marcotegui, B.; Goulette, F.; and Guibas, L. J. 2019年. KPConv：面向点云的灵活可变形卷积。发表于*ICCV*。Uy, M. A.; Pham, Q.; Hua, B.; Nguyen, D. T.; and Yeung, S. 2019年. 重新审视点云分类：真实世界数据上的新基准数据集和分类模型。发表于*ICCV*。Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S. E.; Bronstein, M. M.; and Solomon, J. M. 2019年. 用于点云学习的动态图卷积神经网络。发表于*ACM ToG*。Wu, Z.; Song, S.; Khosla, A.; Yu, F.; Zhang, L.; Tang, X.; and Xiao, J. 2015. 3D ShapeNets：体积形状的深度表征。发表于*CVPR*。Xiang, T.; Zhang, C.; Song, Y.; Yu, J.; and Cai, W. 2021. 云中漫步：学习用于点云形状分析的曲线。发表于*ICCV*。Xu, J.; Tang, X.; Zhu, Y.; Sun, J.; and Pu, S. 2021. SGM-Net：通过排序格拉姆矩阵学习旋转不变的点云表征。发表于*ICCV*。Yi, L.; Kim, V. G.; Ceylan, D.; Shen, I.; Yan, M.; Su, H.; Lu, C.; Huang, Q.; Sheffer, A.; and Guibas, L. J. 2016. 用于三维形状集合区域标注的可扩展主动框架。发表于*ACM ToG*。Yu, J.; Zhang, C.; Wang, H.; Zhang, D.; Song, Y.; Xiang, T.; Liu, D.; and Cai, W. 2021. 3D Medical Point Transformer：将卷积引入注意力网络以进行医学点云分析。发表于*CoRR*。Zhang, Z.; Hua, B.; Chen, W.; Tian, Y.; and Yeung, S. 2020. 用于三维点云理解的全局上下文感知卷积。发表于*3DV*。Zhang, Z.; Hua, B.; Rosen, D. W.; and Yeung, S. 2019年. 用于三维点云深度学习的旋转不变卷积。发表于*3DV*。Zhao, C.; Yang, J.; Xiong, X.; Zhu, A.; Cao, Z.; and Li, X. 2022. 旋转不变点云分析：局部几何与全局拓扑的交汇。发表于*PR*。Zhao, H.; Jiang, L.; Fu, C.; and Jia, J. 2019年. PointWeb：增强点云处理的局部邻域特征。发表于*CVPR*。Zhao, Y.; Birdal, T.; Lenssen, J. E.; Menegatti, E.; Guibas, L. J.; and Tombari, F. 2020. 用于三维点云的四元数等变胶囊网络。发表于*ECCV*。

arXiv:2302.02535v1 [cs.CV] 6 Feb 2023

PaRot：通过特征解耦和姿态恢复实现逐块旋转不变的网络 **补充材料**

Dingxin Zhang*, **Jianhui Yu***, **ChaoyiZhang,WeidongCai**悉尼
大学计算机科学学院，澳大利亚 dzha2344@uni.sydney.edu.au，{
jianhui.yu, chaoyi.zhang,tom.cai}@sydney.edu.au

在补充材料中，我们首先在第1节介绍 PaRot 架构的训练流程和网络结构细节。随后，我们在第2节讨论解耦模块中提出的三种损失函数的作用。第3节展示跨尺度学习恢复的姿态特征。最后，我们在第4节提供超参数设置的更多消融研究结果，并在第5节介绍与分割相关的实验结果。

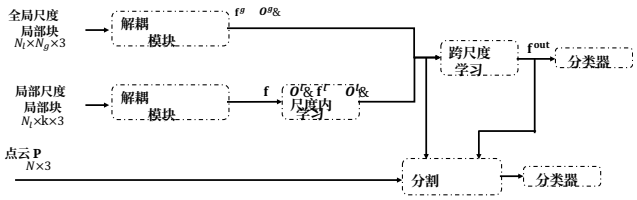


图1：PaRot 的整体结构。

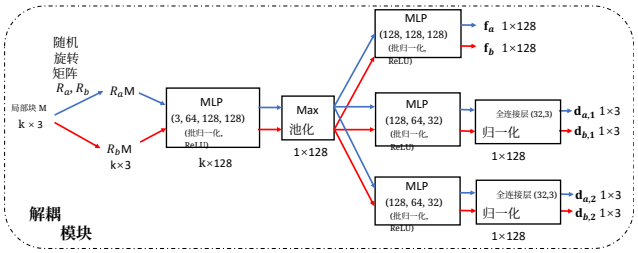


图 2：所提出的解耦模块的详细架构。两个架构相似的解耦模块分别用于处理局部尺度块和全局尺度图块。

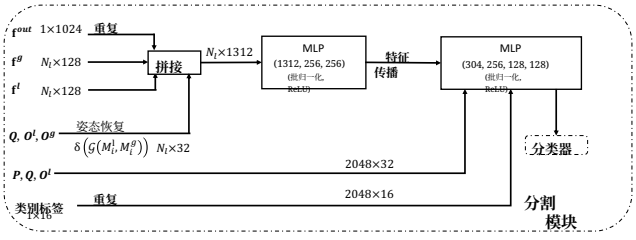


图 3：PaRot 分割模块的详细架构。

1 实现细节

1.1 实验设置

模型使用 PyTorch 在 Nvidia RTX3090 上进行评估。关于局部块生成的设置已在正文中介绍。

我们的总损失函数 $\mathcal{L}_{total}$ 定义如下：

$$\mathcal{L}_{total} = \mathcal{L}_{cls} + \alpha_\ell \mathcal{L}_{equi_\ell} + \mathcal{L}_{orth_\ell} + \beta_\ell \mathcal{L}_{inv_\ell} \\ \alpha_g \mathcal{L}_{equi_g} + \mathcal{L}_{orth_g} + \beta_g \mathcal{L}_{inv_g}, \quad (1)$$

其中，$\mathcal{L}_{cls}$ 是交叉熵分类损失，下标 ℓ 和 g 分别表示局部尺度和全局尺度解耦模块对应的损失函数。此外，α_ℓ、α_g、β_ℓ和 β_g 是用于调整局部和全局尺度下不同损失函数贡献的加权参数。我们分别将 α_ℓ、α_g、β_ℓ和 β_g设为0.2、0.1、0和0。关于为何将不变损失函数 L_{inv} 设为0，详见第2节。

对于分类任务，在训练期间，将输入点云随机缩放至 [0.67至1.5]的范围内进行数据增强，训练轮数为250，批量大小为32。使用Adam优化器，学习率初始化为1e-3，并通过余弦退火调度器调整至1e-5。动量和权重衰减分别设为0.9和1e-6。对于分割任务，实验设置与分类任务相同，但将 N_g 改为64，将 k_{intra} 改为16。我们按照PointNet++ （Qi 等人，2017b）的实现方式，将独热类别标签向量拼接到最后一个特征层。

1.2 模型架构

PaRot模型的整体架构如图1所示。有关解耦模块和

*这些作者贡献相同。

分割模块的详细信息分别见图 2和图 3。尺度内学习模块、尺度间学习模块和特征传播模块的架构相对简单，且已在主要工作中进行了说明，因此我们仅以书面表达式提供相关信息。值得一提的是，所有几何关系编码函数，即 *i.e.*， $\delta(\cdot)$，均由一个输出通道数为32的全连接层、一个批归一化层和一个ReLU组成。

在尺度内学习模块中，我们通过在欧几里得空间内进行 k-近邻搜索来实现 Edge-Conv（Wang 等人，2019年），输入为来自两个局部块的旋转不变特征以及编码后的姿态特征。尺度内学习模块中多层感知机各层的通道数为 288、128、128 和 128。尺度间学习模块负责特征聚合和通道扩展，因此多层感知机的通道数设为 288、256、512、1024，并采用 LeakyReLU（0.2）作为激活函数。分类和分割所用的分类器借鉴自 PointNet++ （Qi 等人，2017b）。

2 损失函数

所提出的解耦模块包含三个损失函数，用于约束学习到的特征为旋转不变特征或旋转等变特征。本节通过消融研究，考察主文中的三个损失函数（即公式（1–3））对模型性能的影响。

如表 1 所示，在不施加任何限制时，分类准确率（F）为 90.4%，低于我们最佳模型 B 的准确率 91.0%。图 4（c）和（f）显示，$\mathcal{L}_{equi}$ 与 $\mathcal{L}_{orth}$ 的结合加快了旋转等变向量的学习，并充分促使两个向量彼此垂直，使准确率提高了 0.6%。将模型 C–E 与模型 B 进行比较，可以清楚地看出，单独应用损失函数无法取得最佳结果。此外，我们发现，应用所有损失函数（模型 A）时，模型性能会比模型 B 差。如图 4（a）、（b）、（d）和（e）所示，模型 A 的 $\mathcal{L}_{inv}$ 下降速度快于模型 B，并且在前 30 个训练周期中的表现更好。然而，当训练轮次足够多时，$\mathcal{L}_{inv}$ 会妨碍形状内容特征的学习，导致准确率下降。

为进一步分析实现组合

模型	$\mathcal{L}_{inv}$	$\mathcal{L}_{equi}$	$\mathcal{L}_{orth}$	Acc.
A	✓	✓	✓	90.6
B		✓	✓	**91.0**
C	✓			90.6
D		✓		90.3
E			✓	88.6
F				90.4

表 1：对我们的解缠模块中的损失函数进行消融研究。报告了在 z/SO3 下 ModelNet40 上的结果。

搜索	半径	k_ℓ	z/z	z/SO3	浮点运算次数
球查询	0.2	64	90.3	90.4	1431M
球查询	0.3	64	90.3	90.5	1431M
knn	-	32	90.7	90.6	1220M
knn	-	64	**90.9**	**91.0**	1431M
knn	-	128	90.8	90.6	1852M

表 2：局部尺度块生成的消融研究。在 ModelNet40 上评估，条件为 z/z、z/SO3。

$\mathcal{L}_{equi}$ 与 $\mathcal{L}_{orth}$的效果，我们在图 4 (c) 和 (f) 中可视化了 B 和 F 的等变损失曲线与正交损失曲线。当不使用 $\mathcal{L}_{equi}$ 与 $\mathcal{L}_{orth}$ 时，等变损失仍会缓慢下降，但 $\mathcal{L}_{orth}$ 会持续增大，这意味着学习到的两个方向向量彼此平行，在恢复位姿信息时会导致一些歧义问题。此外，这表明学习局部尺度块的方向矩阵比学习全局尺度块的方向矩阵更困难。

3 恢复姿态特征可视化

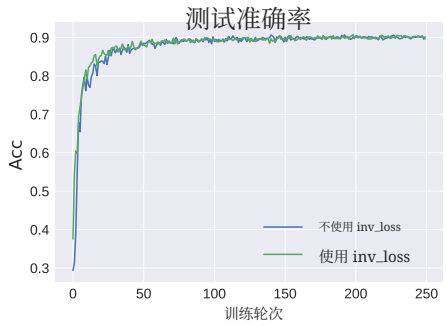
为考察恢复的位姿信息的旋转不变性和有效性，我们在 ShapeNet 部件分割任务中对尺度间学习模块的恢复姿态特征进行可视化。我们遵循论文正文中可视化解耦特征的相同步骤，选取三个通道作为 RGB 值，并从飞机、吉他和手枪类别中各选择一个物体，使其绕 z 轴旋转以进行可视化。我们将 N_ℓ 设为 2048，以利用保存的模型生成稠密结果。

正如论文正文所述，尺度间学习的姿态恢复模块旨在探索逐块特征与全局上下文之间的关系。学习到的特征需要具备旋转不变性，并同时包含相对位置信息和逐补丁方向信息。如图 5 所示，恢复的特征在不同方向下保持一致。此外，靠近物体中心的区域通常呈绿色，而较远的区域呈粉色。另外，受特定局部块方向的影响，一些边缘区域呈蓝色，而一些复杂的局部块（即机翼与机身连接处以及手枪的车轮）则呈红色。

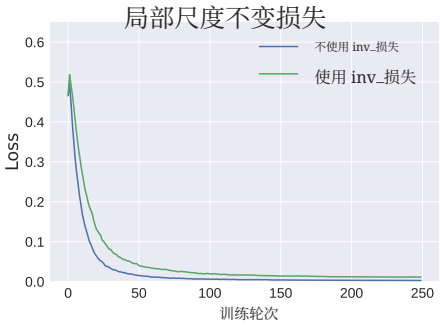
4 消融研究

4.1 局部尺度块的生成

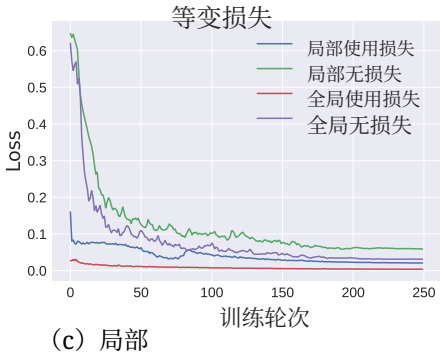
我们研究了两种用于生成局部尺度块的邻域搜索方法：k-NN 搜索和球查询。在表 2 中，我们考察了这两种方法，并报告了提取不同数量邻居时对应的结果。结果发现，采用 k-nn 的模型优于采用球查询方法的模型。这可能是因为球查询方法严格限制了生成的局部块大小，而 k-nn 方法更加灵活，能够从点云的稀疏部分和密集部分生成更好的局部块。当设置 N_ℓ 为 64 时，基于 k-nn 的模型能够取得最佳性能，同时计算成本也较为适中。



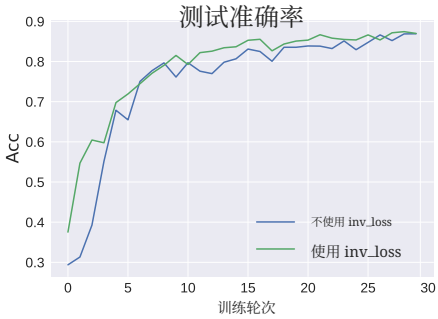
（a）A 和 B 的准确率曲线（使用 L_{inv} 与未使用 L_{inv}）。



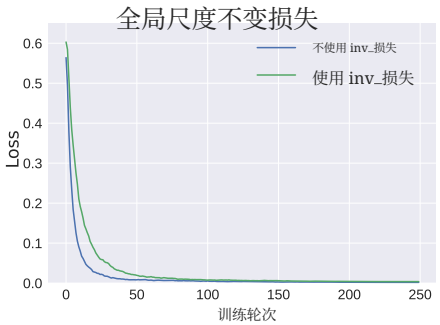
（b）A 和 B 的局部尺度 L_{inv} 曲线。



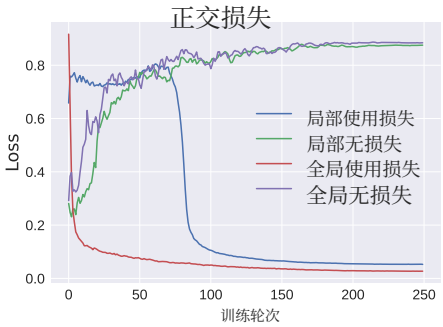
（c）局部和全局尺度补丁的等变损失曲线（B 和 F）。



（d）前 30 个训练周期的准确率曲线（A 和 B（使用 $\mathcal{L}_{inv}$ 与不使用 $\mathcal{L}_{inv}$）。



（e）A 和 B 的全局尺度 $\mathcal{L}_{inv}$ 曲线。



（f）局部块和全局尺度补丁的正交损失曲线 B 和 F 的。

图4：表 1 中模型 A、B 和 F 的准确率与损失曲线。

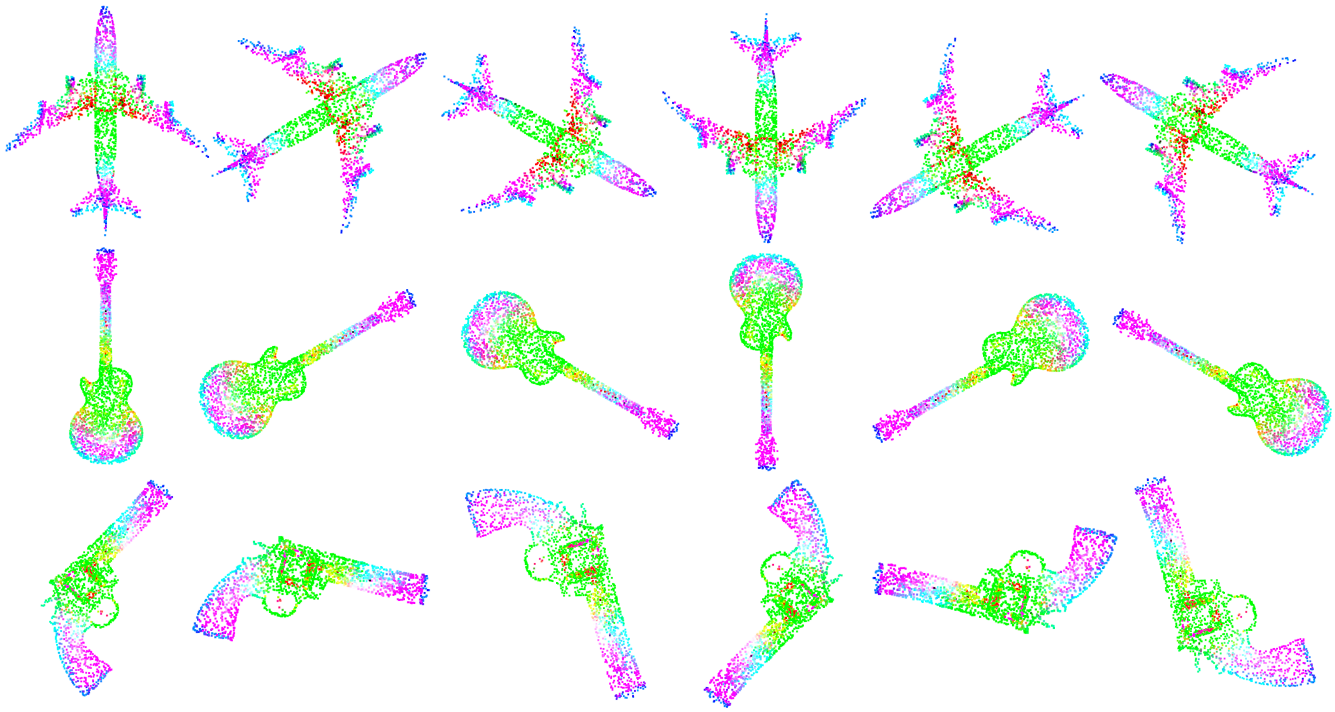


图5：跨尺度学习中恢复的姿态特征的可视化。从左到右，每个样本均绕 z 轴旋转 $60°$。

方法	实例平均交并比	aero	bag	cap	car	椅子	耳机	吉他	刀	灯	笔记本电脑	摩托车	马克杯	手枪	火箭	滑板	桌子
PointNet（Qi 等人，2017a）	74.4	81.6	68.7	74.0	70.3	87.6	68.5	88.9	80.0	74.9	83.6	56.5	77.6	75.2	53.9	69.4	79.9
PointNet++(Qi 等人，2017b)	76.7	79.5	71.6	**87.7**	70.7	88.8	64.9	88.8	78.1	79.2	**94.9**	54.3	92.0	76.4	50.3	68.4	81.0
DGCNN（Wang 等人，2019年）	73.3	77.7	71.8	77.7	55.2	87.3	68.7	88.7	85.5	**81.8**	81.3	36.2	86.0	77.3	51.6	65.3	80.2
RI-Conv（Zhang 等人，2019年）	75.3	80.6	80.2	70.7	68.8	86.8	70.4	87.2	84.3	78.0	80.1	57.3	91.2	71.3	52.1	66.6	78.5
GCANet（Zhang等，2020年）	77.3	81.2	**82.6**	81.6	70.2	88.6	70.6	86.2	**86.6**	81.6	79.6	58.9	90.8	76.8	53.2	67.2	**81.6**
TFN（Poulenard和Guibas，2021年）	78.4	80.3	77.3	82.6	74.7	88.8	**76.3**	90.7	81.7	77.4	82.4	**60.7**	93.2	79.4	54.3	**74.7**	79.6
Li 等（2021）	74.1	81.9	58.2	77.0	71.8	**89.6**	64.2	89.1	85.9	80.7	84.7	46.8	89.1	73.2	45.6	66.5	81.0
VN-DGCNN* （Deng 等（2021））	75.4	81.0	76.1	76.0	71.4	88.1	59.4	91.3	85.0	80.4	85.5	44.7	92.3	74.5	52.4	68.7	78.9
PaRot	**79.5**	**82.9**	82.1	83.2	**75.7**	89.4	76.1	**91.5**	86.1	81.4	80.3	59.3	**94.3**	**79.7**	**57.0**	73.3	79.2

表3：SO3/SO3 条件下 ShapeNet Part 数据集的逐类分割结果及平均交并比。 ⋆表示基于官方实现复现的结果。

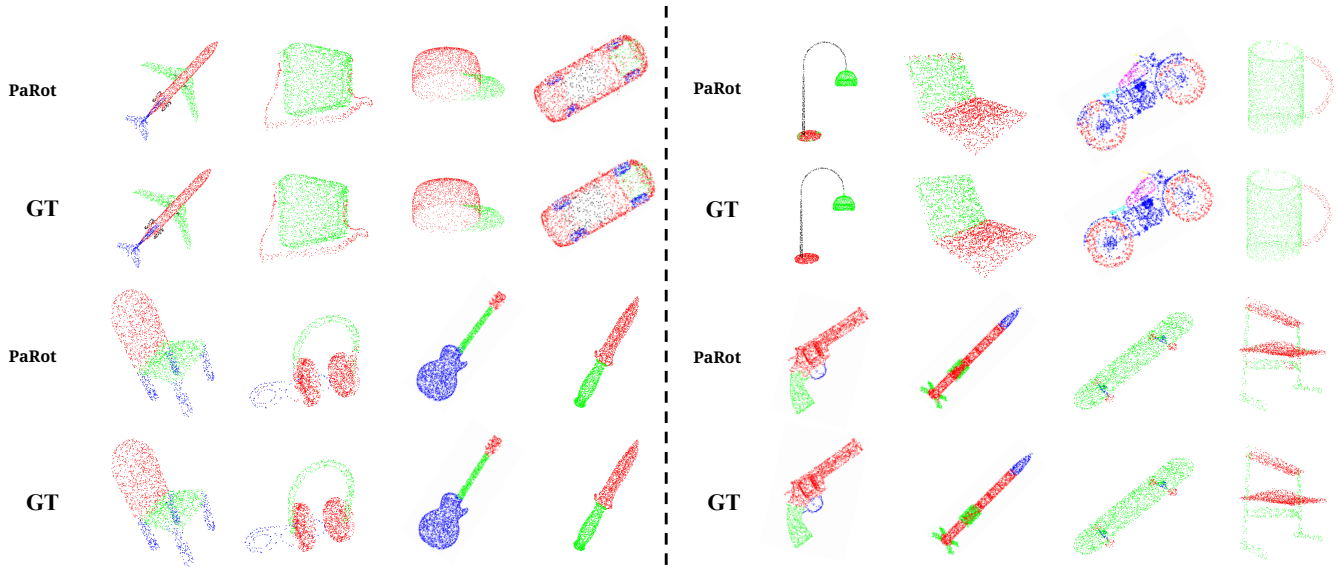


图6：比较 真实标注 (GT) 与 PaRot 为 z/SO3 ShapeNet P 生成的输出 art 分割任务。

N_ℓ	k_{intra}	N_g	z/z	z/SO3	浮点运算次数
64	32	32	90.1	90.0	358M
128	32	32	90.8	90.5	716M
256	32	32	**90.9**	**91.0**	1431M
512	32	32	90.6	90.5	2861M
256	8	32	90.3	90.3	995M
256	16	32	90.5	90.5	1140M
256	64	32	90.6	90.5	2012M
256	32	8	90.7	90.7	1273M
256	32	16	90.7	**91.0**	1325M
256	32	64	90.6	90.5	1641M

表4：对 N_ℓ、 N_g和 k_{intra}进行的消融研究。实验在 ModelNet40 的 z/z、z/SO(3) 设置下进行。

4.2 超参数选择

我们已在第4.1节研究了 k_ℓ ，此外还有三个超参数，即生成局部块的数量 N_ℓ、全局尺度图块中的点数 N_g，以及内部学习中要查询的邻居数量 k_{intra}。为分析这三个超参数的影响，我们在 ModelNet40 上开展了更多实验，结果如表4所示。

对于生成局部块的数量，如果将 N_l 设为较小值，生成的局部块将无法覆盖点云的所有部分，从而导致准确率下降。然而，将 N_l 设为非常大的值，不仅会显著增加计算成本，还会缩小尺度内学习的感受野，损害模型性能。 k_{intra} 的取值也会对计算成本产生很大影响。我们发现，当$N_l = 256$时，将 k_{intra} 设为 32 可获得最佳性能。针对 N_g （为全局尺度补丁采样的点数）进行的消融研究表明，所提出的方法仅使用 8 个点作为全局补丁，就能恢复有效的尺度间位姿信息，并大幅降低计算成本。此外，表4还

表明，在保持高准确率的同时，通过调整超参数，我们还能进一步降低 PaRot 的计算成本。

5 更多分割结果

我们在表3中报告了 ShapeNet Part 的 SO3/SO3 部件分割结果，并以各类别的平均交并比（mIoU）衡量。结果表明，典型的旋转敏感模型在 SO3/SO3 下的表现远优于在 z/SO3 下的表现。通过对训练样本进行旋转增强，典型模型在分割任务中的表现可以超过一些旋转鲁棒方法，尤其是在帽子、灯和笔记本电脑类别上。然而，就各类别平均 mIoU 而言，所提出的 PaRot 方法仍优于这些方法，并且在全部 16个类别上表现均衡。我们还在图6中展示了使用训练好的 z/SO3 模型对每个物体类别各选取一个样本进行可视化的结果。尽管比较真实标注与预测样本时可以发现一些分割错误，但 PaRot 在大多数类别上都能给出准确预测。

参考文献

Deng, C.; Litany, O.; Duan, Y.; Poulenard, A.; Tagliasacchi, A.; and Guibas, L. J. 2021. 向量神经元：SO(3)等变网络的通用框架。In*ICCV*. Li, F.; Fujiwara, K.; Okura, F.; and Matsushita, Y. 2021. 深入探究旋转不变的深度点云分析。In*ICCV*. Poulenard, A.; and Guibas, L. J. 2021. 张量场网络旋转等变非线性的函数方法。In*CVPR*. Qi, C. R.; Su, H.; Mo, K.; and Guibas, L. J. 2017a. PointNet：用于三维分类和分割的点集深度学习。In*CVPR*. Qi, C. R.; Yi, L.; Su, H.; and Guibas, L. J. 2017b. PointNet++：度量空间中点集的深度层次特征学习。In*NeurIPS*. Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S. E.; Bronstein, M. M.; and Solomon, J. M. 2019. 用于点云学习的动态图卷积神经网络。In*ACM ToG*. Zhang, Z.; Hua, B.; Chen, W.; Tian, Y.; and Yeung, S. 2020. 用于三维点云理解的全局上下文感知卷积。In*3DV*. Zhang, Z.; Hua, B.; Rosen, D. W.; and Yeung, S. 2019. 用于三维点云深度学习的旋转不变卷积。In*3DV*。