



(12) 发明专利申请

(10) 申请公布号 CN 114445280 A

(43) 申请公布日 2022.05.06

(21) 申请号 202210071460.8

(22) 申请日 2022.01.21

(71) 申请人 太原科技大学

地址 030024 山西省太原市万柏林区窰流路66号

(72) 发明人 杨亚坤 王安红

(74) 专利代理机构 太原中正和专利代理事务所  
(普通合伙) 14116

专利代理师 焦进宇

(51) Int. Cl.

G06T 3/40 (2006.01)

G06T 17/00 (2006.01)

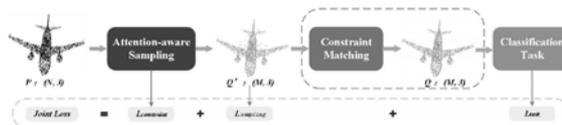
权利要求书2页 说明书8页 附图5页

(54) 发明名称

一种基于注意力机制的点云下采样方法

(57) 摘要

一种基于注意力机制的点云下采样方法,属于点云下采样领域,本发明通过包括输入嵌入模块和注意模块在内的注意感知采样模块实现点云下采样过程,输入嵌入模块用于提取点云的全局和局部特征,注意模块用于模拟面向应用任务的采样过程,得益于注意力机制,本发明可以选择原始点云中的关键点作为采样点,从而进行分类任务,另外本发明设计了约束匹配模块,在推理阶段将采样点匹配为原始点云的子集,我们构造了一个包含任务损失、采样损失和约束损失的联合损失函数用于网络端到端训练,在ModelNet40数据集上的大量实验表明,我们的AS-Net网络在点云分类任务上取得了良好的性能,特别是当下采样点数较少时,结果明显优于其他参考方法。



1. 一种基于注意力机制的点云下采样方法,其特征为:网络结构为:给定一个包含N个点的点云P,该网络通过注意感知采样模块生成采样点云Q',然后通过约束匹配模块来匹配这些点以获得Q,以保证Q是原始点云P的子集,最后,将Q应用于后续的分类任务;

注意感知采样模块包括两个模块:输入嵌入模块和注意模块;

通过输入嵌入模块将原始点云P映射到高维特征空间,该模块可以表征点之间的语义亲和关系,以作为各种点云处理任务的基础,首先学习与原始点云P相关的d维嵌入特征 $k \in \mathbb{R}^{N \times d}$ ,同时,给定采样数目M,也可以获得高维特征 $q \in \mathbb{R}^{M \times d}$ ;

通过输入嵌入模块,获得与原始点云P相关的d维嵌入特征k,以及与采样大小M相关的高维特征q后,根据注意力机制,q和k之间使用MLP学习相似性,并获得与原始点云对应的采样点云的注意图,此注意图包括M行、N列,第i行表示第i个采样点与原始点云中所有N个不同点之间的相关性;然后通过softmax函数对该注意图进行归一化后,得到的采样图更接近理想的采样矩阵;最后,执行原始点云P和采样图S的加权和,以获得采样点云Q';

约束匹配模块

在约束匹配模块中同时约束采样映射图的行和列,我们选择此行中的最大数字并强制它为1,并检查是否采样映射图的每一列仅包含一个元素1,如果一列同时包含两个元素1,则其中一行中第二大数字的位置被指定为1。

2. 根据权利要求1所述的一种基于注意力机制的点云下采样方法,其特征为:所述输入嵌入模块中,基于点云特征提取分为基于点的方法和基于体素的方法;

基于点的分支提取每个点的特征具体步骤如下:

首先,多层感知机MLP用于提取每个点的逐点特征;

然后,将其分为两个通道,一个是得到 $d_1$ 维逐点特征,另一个是通过最大池操作获得整个点云的 $d_2$ 维全局特征;

最后,相加两个通道,即可获得点分支的特征。

3. 根据权利要求1所述的一种基于注意力机制的点云下采样方法,其特征为:所述基于体素的方法,具体步骤如下:

步骤1),根据不同的分辨率,将支路分为三个通道,每个通道首先对三维坐标进行归一化,以减小不同点云的尺度差异,便于后续的网络学习和处理;

步骤2),通过平均坐标落入体素网格的所有特征,将归一化点云转换为体素网格,将点转换为体素网格后,利用三维体积卷积对特征进行聚集,得到点云的局部特征;

步骤3),将通过体素化获得的特征转换回点云域以获得 $d_3$ 维局部特征。

4. 根据权利要求3所述的一种基于注意力机制的点云下采样方法,其特征为:所述步骤2)中使用三线性插值将体素网格转换为点。

5. 根据权利要求1所述的一种基于注意力机制的点云下采样方法,其特征为:该采样方法中端到端网络联合训练损失函数为

$$L_{AS-Net}(P, Q) = L_{task} + \alpha L_{sampling} + \beta L_{constraint}$$

其中,

$$任务损失 L_{task}(Q) = L_T(f_T(Q), y^*)$$

$$采样损失 L_{sampling}(P, Q) = L_1 + L_2, L_1 = \frac{1}{M} \sum_{q \in Q} \min_{p \in P} \|q - p\|_2^2, L_2 = \frac{1}{N} \sum_{p \in P} \min_{q \in Q} \|p - q\|_2^2$$

其中,  $f_T$  是分类网络,  $f_T(Q)$  是预测的分类标签,  $y^*$  是标签真值,  $L_T$  使用交叉熵损失函数,  $L_{\text{sampling}}$  中  $L_1$  确保采样点  $Q'$  尽可能接近原始点云  $P$ ,  $L_2$  确保采样点  $Q'$  均匀分布在原始点云  $P$  中

$$\text{约束损失 } L_{\text{constraint}} = \min \|SS^T - I_M\|_2^2 + \min \|S1_N - 1_M\|_2^2$$

其中,  $S$  为采样图,  $I_M$  是  $M \times M$  的单位矩阵,  $1_N = [1, 1, \dots, 1]^T \in \mathbb{R}^{N \times 1}$ ,  $1_M = [1, 1, \dots, 1]^T \in \mathbb{R}^{M \times 1}$ 。

## 一种基于注意力机制的点云下采样方法

### 技术领域

[0001] 本发明属于点云下采样领域,特别是涉及一种基于注意力机制的点云下采样方法。

### 背景技术

[0002] 随着三维传感器的快速发展,近年来获取三维数据变得越来越容易。由于点云能够在三维空间提供丰富的几何、形状和尺度信息,基于点云的各种应用正在增加,包括自动驾驶、虚拟现实、增强现实、机器人等。然而,点云数据量大、格式不规则、具有稀疏性,因此处理和传输难度大。为了克服这一挑战,点云下采样成为简化点云的一种流行且有效的方法,从而节省存储空间,减少传输带宽和通信负载。

[0003] 传统的点云下采样方法包括随机采样(RS),最远点采样(FPS),泊松圆盘采样(PDS)等。它们通过迭代生成均匀分布的样本,从而保留原始点云的几何结构。然而,这些下采样方法只关注于减少几何采样损失,严重降低了后续应用任务的性能。通常,对于三维点云,我们不仅希望简化的点云保持其原始形状,还希望其适用于后续应用任务。这些要求可以通过训练网络同时满足采样损失和任务损失来实现。

[0004] 点云的应用任务通常包括形状分类,目标检测,分割,重建,补全,注册和其他应用。关于点云的深度学习已经吸引了越来越多的关注。最早将点云下采样和应用任务结合的是Dovrat等人提出的名为S-Net的深度学习网络。由于其面向任务的特点,S-Net在各种应用中优于传统算法。后来,Lang等人提出了SampleNet来改进S-Net,方法是引入额外的投影模块,使生成的点更接近原始点云。最近,由Qian提出的MOPS-Net是一个可解释的基于深度学习的网络,它把采样看作是一个约束可微矩阵优化问题。然而,S-Net和SampleNet在特征提取中只依赖于点云的全局特征,不使用高维局部特征,这限制了生成点云的质量。MOPS-Net使用局部信息,但其特征提取模块效率不高,当采样点数目较小时,后续分类任务的性能不是很好。

[0005] 此外,匹配模块是将采样点云映射成为原始点云的子集的重要步骤。S-Net提出了两种匹配方法:地球移动距离(EMD)匹配和最近邻(NN)匹配。SampleNet提出了软投影操作匹配。MOPS网络也是将采样点云的每个点映射到原始点云中最近的点,并使用FPS获得下采样子集。这些匹配方法虽然能取得较好的效果,但过于复杂和耗时。

[0006] 近年来,注意力机制在自然语言处理和计算机视觉中取得了巨大的成功。注意力机制还可以使网络聚焦于点云的重要部分,并捕获丰富的特征表示以简化点云。对于点云分类,Guo等人提出了一个名为Point CloudTransformer的网络来进行点云学习,并证明PCT在形状分类、部件分割等任务方面实现了最先进的性能。对于点云上采样,也引入了注意力模块来从点云中提取特征,例如PU-GAN和PUI-Net。因此,我们将结合注意力机制来完成面向分类任务的点云下采样。

## 发明内容

[0007] 本发明针对现有技术存在的不足,旨在提供一种基于注意力机制的点云下采样方法。本发明的方法通过注意力感知采样模块和约束匹配模块进行下采样,能够得到比其他下采样方法更高的分类精度和更高的效率。

[0008] 为解决上述技术问题,本发明采用的技术方案为:一种基于注意力机制的点云下采样方法,网络结构为:给定一个包含N个点的点云P,该网络通过注意感知采样模块生成采样点云Q',然后通过约束匹配模块来匹配这些点以获得Q,以保证Q是原始点云P的子集,最后,将Q应用于后续的分类任务;

[0009] 注意感知采样模块包括两个模块:输入嵌入模块和注意模块;

[0010] 通过输入嵌入模块将原始点云P映射到高维特征空间,该模块可以表征点之间的语义亲和关系,以作为各种点云处理任务的基础,首先学习与原始点云P相关的d维嵌入特征 $k \in \mathbb{R}^{N \times d}$ ,同时,给定采样数目M,也可以获得高维特征 $q \in \mathbb{R}^{M \times d}$ ;

[0011] 通过输入嵌入模块,获得与原始点云P相关的d维嵌入特征k,以及与采样大小M相关的高维特征q后,根据注意力机制,q和k之间使用MLP学习相似性,并获得与原始点云对应的采样点云的注意图,此注意图包括M行、N列,第i行表示第i个采样点与原始点云中所有N个不同点之间的相关性;然后通过softmax函数对该注意图进行归一化后,得到的采样图更接近理想的采样矩阵;最后,执行原始点云P和采样图S的加权和,以获得采样点云Q';

[0012] 约束匹配模块

[0013] 在约束匹配模块中同时约束采样映射图的行和列,我们选择此行中的最大数字并强制它为1,并检查是否采样映射图的每一列仅包含一个元素1,如果一列同时包含两个元素1,则其中一行中第二大数字的位置被指定为1。

[0014] 进一步的,所述输入嵌入模块中,基于点云特征提取分为基于点的方法和基于体素的方法;

[0015] 基于点的分支提取每个点的特征具体步骤如下:

[0016] 首先,多层感知机MLP用于提取每个点的逐点特征;

[0017] 然后,将其分为两个通道,一个是得到 $d_1$ 维逐点特征,另一个是通过最大池操作获得整个点云的 $d_2$ 维全局特征;

[0018] 最后,相加两个通道,即可获得点分支的特征。

[0019] 进一步的,所述基于体素的方法,具体步骤如下:

[0020] 步骤1),根据不同的分辨率,将支路分为三个通道,每个通道首先对三维坐标进行归一化,以减小不同点云的尺度差异,便于后续的网络学习和处理;

[0021] 步骤2),通过平均坐标落入体素网格的所有特征,将归一化点云转换为体素网格,将点转换为体素网格后,利用三维体积卷积对特征进行聚集,得到点云的局部特征;

[0022] 步骤3),将通过体素化获得的特征转换回点云域以获得 $d_3$ 维局部特征。

[0023] 进一步的,所述步骤2)中使用三线性插值将体素网格转换为点。

[0024] 进一步的,该采样方法中端到端网络联合训练损失函数为

$$[0025] \quad L_{AS-Net}(P, Q) = L_{task} + \alpha L_{sampling} + \beta L_{constraint}$$

[0026] 其中,

$$[0027] \quad \text{任务损失 } L_{task}(Q) = L_T(f_T(Q), y^*)$$

[0028] 采样损失  $L_{\text{sampling}}(P, Q) = L_1 + L_2$ ,  $L_1 = \frac{1}{M} \sum_{q \in Q} \min_{p \in P} \|q - p\|_2^2$ ,  $L_2 = \frac{1}{N} \sum_{p \in P} \min_{q \in Q} \|p - q\|_2^2$

[0029] 其中,  $f_T$  是分类网络,  $f_T(Q)$  是预测的分类标签,  $y^*$  是标签真值,  $L_T$  使用交叉熵损失函数,  $L_{\text{sampling}}$  中  $L_1$  确保采样点  $Q'$  尽可能接近原始点云  $P$ ,  $L_2$  确保采样点  $Q'$  均匀分布在原始点云  $P$  中

[0030] 约束损失  $L_{\text{constraint}} = \min \|SS^T - I_M\|_2^2 + \min \|S1_N - 1_M\|_2^2$

[0031] 其中,  $S$  为采样图,  $I_M$  是  $M \times M$  的单位矩阵,  $1_N = [1, 1, \dots, 1]^T \in \mathbb{R}^{N \times 1}$ ,  $1_M = [1, 1, \dots, 1]^T \in \mathbb{R}^{M \times 1}$ 。

[0032] 本发明的优点及其积极效果是:

[0033] 1、本发明提出了一种新的注意感知采样模块来获取采样点云。它包含两个子模块:输入嵌入模块和注意模块。输入嵌入模块用来提取丰富的点云特征。它可以利用点方法的内存消耗小,以及体素方法良好的内存局部性和数据规则性优点。注意模块生成采样图来模拟实际的下采样,同时解决下采样过程的不可微问题。

[0034] 2、本发明中提出了一种新的约束匹配模块,将采样图转换为理想的匹配图,保证采样点是原始点云的子集,本发明的约束匹配模块简单有效,适用于推理阶段。

[0035] 3、本发明提出了一种结合任务损失、采样损失和采样图约束损失的端到端训练联合损失函数。

## 附图说明

[0036] 图1为本发明的网络结构。

[0037] 图2为注意感知采样模块结构。

[0038] 图3为输入嵌入模块结构。

[0039] 图4为点分支详细结构。

[0040] 图5为体素分支详细结构。

[0041] 图6为注意模块详细结构。

[0042] 图7为按行约束示意图。

[0043] 图8为约束匹配模块中同时约束行和列示意图。

[0044] 图9为不同方法采样点云为生成的点云  $Q'$  时的分类精度对比图。

[0045] 图10为不同方法采样点云为匹配点云  $Q$  时的分类精度对比图。

[0046] 图11为精度与测试时间的对比图。

[0047] 图12为精度与GPU内存消耗的对比图。

[0048] 图13为不同采样大小  $M$  下的匹配后的采样点云  $Q$ 。

## 具体实施方式

[0049] 为使本发明的目的、特征和优点能够更为明显易懂,下面对本发明的具体实施方式做详细的说明。

[0050] 本发明提出了一种面向点云分类任务的深度学习下采样网络。我们的网络如图1所示。给定一个包含  $N$  个点的点云  $P$ , 该网络通过注意感知采样模块生成采样点云  $Q'$ , 然后通

过约束匹配模块来匹配这些点以获得Q,以保证Q是原始点云P的子集。最后将Q应用于后续的分类任务。训练中的联合损失函数是约束损失、采样损失和任务损失的总和。

[0051] 将采样点数目定义为M,将分类任务定义为T。下采样网络的目的是找到一个包含M个点的子集Q,从而使任务网络T的目标函数f最小化:

$$[0052] \quad Q = \arg \min(f(T(Q))), Q \subset P, M < N$$

[0053] 如图2所示,注意感知采样模块中有两个关键模块:输入嵌入模块和注意模块。

[0054] 输入嵌入模块

[0055] 在注意感知采样模块中,首先通过输入嵌入模块将原始点云P映射到高维特征空间,该模块可以表征点之间的语义亲和关系,以作为各种点云处理任务的基础。首先学习与原始点云P相关的d维嵌入特征 $k \in \mathbb{R}^{N \times d}$ ,同时,给定采样数目M,也可以获得高维特征 $q \in \mathbb{R}^{M \times d}$ 。

[0056] 图3给出了输入嵌入模块的具体结构。在输入嵌入模块中,点云特征提取分为基于点的方法和基于体素的方法。通过这两个分支,我们分离了细粒度特征转换和粗粒度邻居聚集两个任务,从而使每个分支都非常高效。与S-Net和SampleNet中使用的PointNet相比,本方法可以提取局部结构特征。与MOPS-Net中提到的PointNet++相比,本方法能够具有良好内存局部性和规律性,减少内存占用。特别是考虑到现实世界中的点云往往疏密不同,仅在单个分辨率下学习局部特征会使网络不健壮,我们在不同分辨率下提取局部特征,实现多尺度特征学习,可以扩展感受野,解决了点云稀疏和不规则表示的影响,使网络在捕捉细节方面更加准确和具有鲁棒性。

[0057] 其中,基于点的分支的详细结构如图4所示。基于点的分支提取每个点的特征,它不聚集邻域信息,并且可以提供非常高的分辨率和较小的内存占用。首先,多层感知机(MLP)用于提取每个点的逐点特征。然后,将其分为两个通道,一个是得到 $d_1$ 维逐点特征,另一个是通过最大池操作获得整个点云的 $d_2$ 维全局特征。相加两个通道,即可获得点分支的特征。这种细粒度、高分辨率的基于点的信息对于补充基于体素的粗粒度信息非常重要。

[0058] 基于体素的分支的详细结构如图5所示。根据不同的分辨率,将支路分为三个通道,每个通道首先对三维坐标进行归一化,以减小不同点云的尺度差异,便于后续的网络学习和处理。然后,通过平均坐标落入体素网格的所有特征,将归一化点云转换为体素网格。在我们的方法中,体素分支只提取局部特征,并且体素分辨率不需要很大,因此体素化表示不会占用大量内存。将点转换为体素网格后,利用三维体积卷积对特征进行聚集,得到点云的局部特征。由于需要将基于体素分支获得的信息与基于点分支的特征融合,因此将通过体素化获得的特征转换回点云域以获得 $d_3$ 维局部特征。在这里,我们使用三线性插值将体素网格转换为点,这可以确保映射到每个点的特征是不同的。体素化和非体素化只需扫描所有点一次,这降低了内存成本。

[0059] 点云的逐点特征、全局特征和不同分辨率下的多尺度局部特征可以提供丰富的互补信息。我们的网络有效地将这两个分支的多个通道合并在一起,以实现高效、准确、健壮的点云特征学习。

[0060] 注意模块

[0061] 深度学习中的注意力机制类似于人类视觉的注意力机制。它将注意力集中在大量信息中的重要信息上,选择关键信息,而忽略其他不重要信息。我们可以将下采样视为从原

始点云中选择少量重要点的过程。因此可通过将原始点云P与采样矩阵S'相乘来模拟点云下采样,以获得采样点云Q。

$$[0062] \quad Q = S' \times P$$

[0063] 理想的采样矩阵S'仅包含两个元素0和1,其中1表示选择该点作为采样点,而0表示不选择该点。采样矩阵需要满足等式(3)中的约束条件。

$$[0064] \quad S' S'^T = I_M, S' \cdot 1_N = 1_M \quad (3)$$

[0065] 其中, $I_M$ 是 $M \times M$ 的单位矩阵, $1_N = [1, 1 \dots 1]^T \in \mathbb{R}^{N \times 1}$ , $1_M = [1, 1 \dots 1]^T \in \mathbb{R}^{M \times 1}$ 。

[0066] 然而,理想的采样矩阵是离散的、不可微的,这在神经网络中是无法实现端到端训练的。因此,我们提出了一种软采样映射图来逼近理想的采样矩阵,使得采样过程变得连续可微,网络可以实现端到端的训练。在这个采样图中,每个元素不是0或1,而是介于0和1之间的数字,越接近0或1越好。该采样图需要满足等式(4)中的约束条件,该约束条件是在放松等式(3)中的约束条件后获得的。

$$[0067] \quad \|\|SS^T - I_M\|_F \leq \varepsilon, \|\|S1_N - 1_M\|_F \leq \varepsilon \quad (4)$$

[0068] 我们的下采样网络使用注意模块获得采样图。获取采样图的过程如图6所示。

[0069] 通过输入嵌入模块,获得与原始点云P相关的高维特征k,以及与采样大小M相关的高维特征q。根据注意力机制,q和k之间的相关性可以通过三种方法学习:点积、余弦函数和MLP。我们使用MLP学习相似性,并获得与原始点云对应的采样点云的注意图。此注意图包括M行、N列,第i行表示第i个采样点与原始点云中所有N个不同点之间的相关性。注意图中的数字越大,意味着采样点与原始点的相似度越高,就越容易对该位置的原始点进行采样。通过softmax函数对该注意图进行归一化后,得到的采样图更接近理想的采样矩阵。最后,执行原始点云P和采样图S的加权和,以获得采样点云Q'。

[0070] 约束匹配模块

[0071] 由于生成的采样点云Q'不能保证是原始点云P的子集,并且之前的匹配方法过于复杂且计算量大,因此我们提出了一个约束匹配模块,在后处理步骤中将Q'与Q进行匹配。

[0072] 如图7所示,在注意模块中进行softmax操作后,采样图中每行的每个元素都是一个介于0和1之间的数字,总和为1。数字越大,表示此位置的点越容易被选择。我们按行约束它,选择此行中最大的数字并强制它为1,其他数字强制为0。将原始点云P与更改后的匹配图相乘后,我们可以从对应位置为1的原始点云中获取采样点云Q。但是,这其中有一些采样点可能相同,采样点的数量 $M_1$ 可能小于要求的采样尺寸M。

[0073] 因此,如图8所示,我们在约束匹配模块中同时约束采样映射图的行和列。我们选择此行中的最大数字并强制它为1,并检查是否采样映射图的每一列仅包含一个元素1。如果一列同时包含两个元素1,则其中一行中第二大数字的位置被指定为1。这样,我们可以在原始点云中获取不同的采样点Q,并满足所需的下采样尺寸M。我们的约束匹配模块简单有效,适用于推理阶段。

[0074] 联合损失函数

[0075] 在面向分类任务的下采样网络中,通常考虑两类损失函数:任务损失 $L_{task}$ 和采样损失 $L_{sampling}$ 。在我们的网络结构中,我们构造了一个新的联合损失函数,包括约束损失 $L_{constraint}$ 。

[0076] 在注意模块中,通过放松约束条件得到采样图。采样图中的元素介于0和1之间,接

近于0或1,这与实际采样矩阵存在误差,使得下采样不够准确。因此,我们设计了一个损失函数 $L_{\text{constraint}}$ ,根据等式(4)约束采样映射。

$$[0077] \quad L_{\text{constraint}} = \min \|SS^T - I_M\|_2^2 + \min \|S1_N - 1_M\|_2^2$$

[0078] 因此,本发明提出的端到端网络联合训练损耗为:

$$[0079] \quad L_{\text{AS-Net}}(P, Q) = L_{\text{task}} + \alpha L_{\text{sampling}} + \beta L_{\text{constraint}}$$

[0080] 这里, $\alpha$ 和 $\beta$ 平衡这三项。

[0081] 总之, $L_{\text{task}}$ 促使网络学习采样点云 $Q'$ ,该点云可以保持分类任务 $T$ 的高性能, $L_{\text{sampling}}$ 和 $L_{\text{constraint}}$ 促使学习的采样点云 $Q'$ 更接近原始点云 $P$ 的子集。

$$[0082] \quad L_{\text{task}}(Q) = L_T(f_T(Q), y^*)$$

$$[0083] \quad L_{\text{sampling}}(P, Q) = L_1 + L_2, \quad L_1 = \frac{1}{M} \sum_{q \in Q} \min_{p \in P} \|q - p\|_2^2, \quad L_2 = \frac{1}{N} \sum_{p \in P} \min_{q \in Q} \|p - q\|_2^2$$

[0084] 其中, $f_T$ 是分类网络, $f_T(Q)$ 是预测的分类标签, $y^*$ 是标签真值, $L_T$ 使用交叉熵损失函数。 $L_{\text{sampling}}$ 中 $L_1$ 确保采样点 $Q'$ 尽可能接近原始点云 $P$ , $L_2$ 确保采样点 $Q'$ 均匀分布在原始点云 $P$ 中。

[0085] 为了更好的验证本发明的算法的有效性。下面通过具体的数据集进行仿真验证。

[0086] 点云分类使用ModelNet40数据集进行实验评估。ModelNet40包含40个类别的12311个3D对象,其中9843个用于训练,2468个用于测试。我们使用配备RTX8000 GPU的计算机来训练我们提出的模型。在训练过程中,我们将批量大小设置为32,代数设置为200,初始学习率为0.01,每20代下降到前一个的一半。体素化分辨率为4,8,16, $\alpha$ 为30, $\beta$ 为10。使用Adam优化器用于训练。为了公平比较,分类网络为PointNet。其他设置与S-Net相同。点云分类的评价指标是总体精度。

[0087] 与其他的方法比较

[0088] 我们比较了现有的三种面向深度任务的下采样方法S-Net、SampleNet、MOPS-Net和本发明的AS-Net。采样大小 $M$ 为8,16,...512。

[0089] 首先,我们比较了采样点云为生成的点云 $Q'$ 时的分类精度。由于下游应用程序是分类任务,因此不需要后期匹配过程来确保采样点必须是原始点云 $P$ 的子集。我们使用深度网络生成的点来表示采样点。实验结果如图9所示。

[0090] 实验结果表明,AS-Net的精度在所有采样大小下都达到了较好的性能。特别是当下采样数目较小时,精度显著提高。这是因为我们的注意感知采样模块结合了注意机制来捕获重要和有用的信息,这更适合于下采样任务。同时,我们的输入嵌入模块可以得到丰富的点云特征,也有助于提高分类精度。

[0091] 接下来,我们比较了采样点云为匹配点云 $Q$ 时的分类精度。在测试时,我们强制采样图的元素为0或1,使采样点云是原始点云 $P$ 的子集。这是一个真正的下采样过程。实验结果如图10所示。

[0092] 这表明我们的AS-Net可以获得比S-Net和SampleNet更好的分类精度。当下采样数目较大时,精度相当于MOPS-Net。当采样数目较小时,其精度明显高于其他方法。这表明我们的约束匹配模块也是非常有效的。

[0093] 消融实验

[0094] 为了说明输入嵌入模块的效率,我们比较了四种特征学习方法,包括PointNet、PointNet++、PVCNN和我们的输入嵌入模块。将其应用于点云分类,并查看其总体精度。此外,我们报告了整个测试数据集的测试时间和GPU内存消耗,以反映效率。实验结果如表1所示。

[0095] 表1各方法分类结果比较

	方法	总体精度	测试时间	GPU 占用
	PointNet	89.2	32s	1147M
[0096]	PointNet++	90.7	75s	5603M
	PVCNN	88.6	17s	2357M
	Ours	90.4	19s	2931M

[0097] 如表1所示,本发明提出的输入嵌入模块的总体精度为90.4%,高于PointNet和PVCNN。在获得与PointNet++相当的精度的情况下,所需的测试时间要短得多,GPU占用率也比PointNet++小得多。

[0098] 精度与测试时间和GPU内存消耗的权衡如图11和图12所示。此方法越靠近图中的左上角,性能越好。我们可以看到,与其他方法相比,我们的方法在精度与测试时间之间取得了明显更好的折衷。在同样的精度下,我们的方法比PointNet++快近3倍。此外,我们的方法实现了精度与GPU内存消耗更高的平衡。与PointNet++相比,我们的方法在同样的精度下节省了近2倍的GPU消耗。美中不足的是,它的GPU内存消耗超过了PointNet。

[0099] 为了更清楚地说明网络中每个特征的有效性,还进行了以下消融实验。实验结果如表2所示。可以看出,全局特征对于点云分类非常重要,而多分辨率局部特征可以进一步提高分类精度。

[0100] 表2消融实验结果

	方法	总体精度
[0101]	点特征+体素特征	88.6
[0102]	点特征+体素特征+全局特征	90.2
	点特征+多分辨率体素特征+全局特征	90.4

#### [0103] 可视化结果

[0104] 为了更直观地演示AS-Net的性能,我们给出了以下可视化结果。图13直观地说明了不同采样大小M下的匹配后的采样点云Q。从该图可以看出,我们的AS-Net可以成功地选择关键部件中的点以及形状轮廓附近的点。当下采样尺寸逐渐变小时,采样点仍然可以分布在语义差异较大的部分。我们可以看到,在飞机中,采样点分布在机翼、尾部和机身的各个部分。在花瓶中,采样点分布在花瓶和鲜花的各个部位。这些观察结果说明我们的AS-Net可以选择更好的采样点,并获得更高的分类精度。

[0105] 本发明提出了一种面向点云分类任务的高效、可解释的端到端深度下采样网络AS-Net。与现有的类似方法相比,本发明AS-Net采用注意感知采样模块来模拟实际的下采

样过程。在输入嵌入模块中,结合点分支和体素分支提取点云特征,节省内存,减少随机访问。然后使用注意模块获取采样图来得到下采样点云,这种注意感知方法允许选择更重要的点。在推理阶段,约束匹配模块强制将采样图的元素变为0或1,使采样点云成为原始点云的子集,然后完成分类任务。通过在ModelNet40数据集上进行实验表明,AS-Net比其他下采样方法具有更高的分类精度和更高的效率。

[0106] 上面对本发明的实施例作了详细说明,但是本发明并不限于上述实施例,在本领域普通技术人员所具备的知识范围内,还可以在不脱离本发明宗旨的前提下作出各种变化。

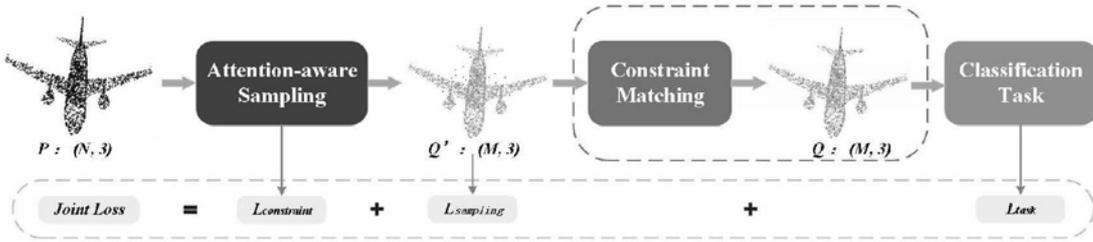


图1

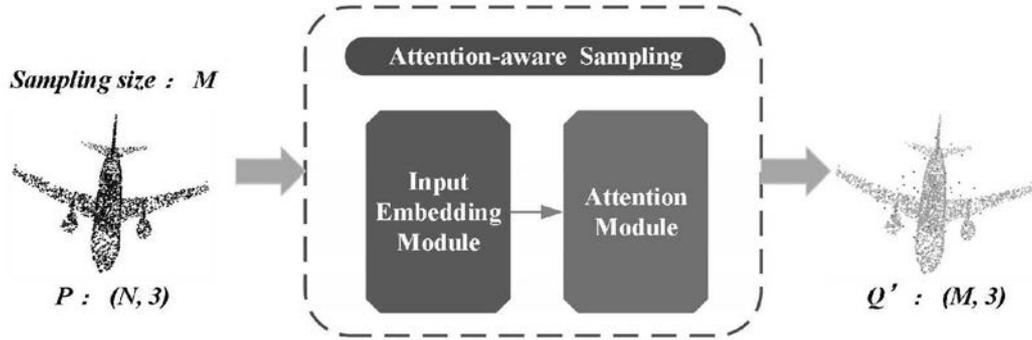


图2

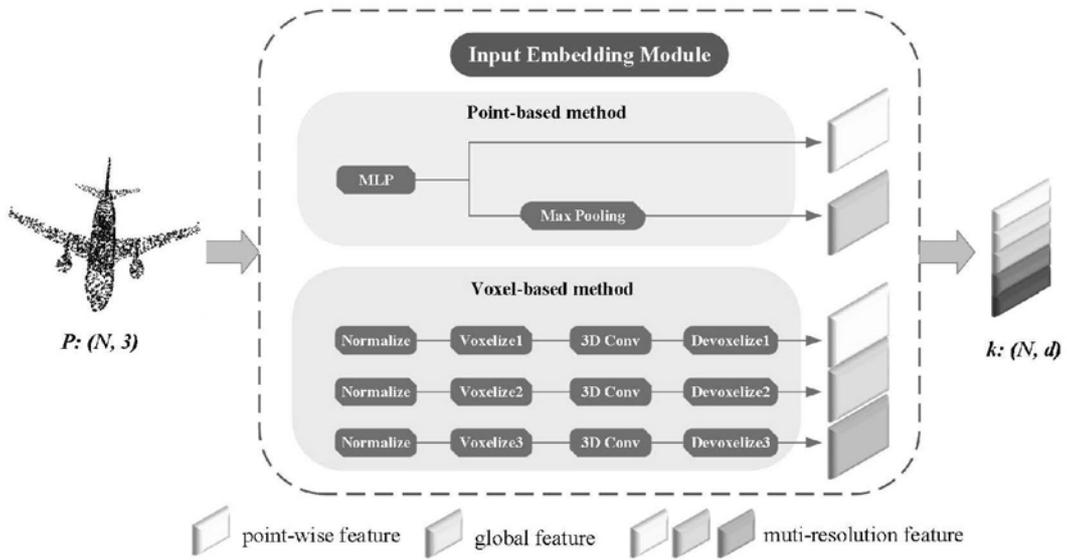


图3

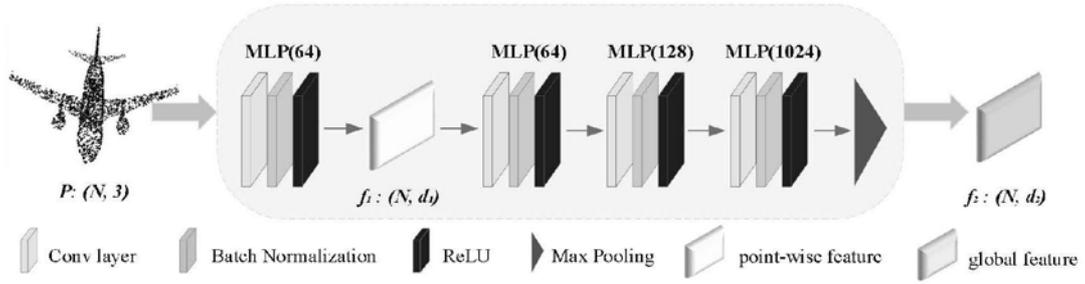


图4

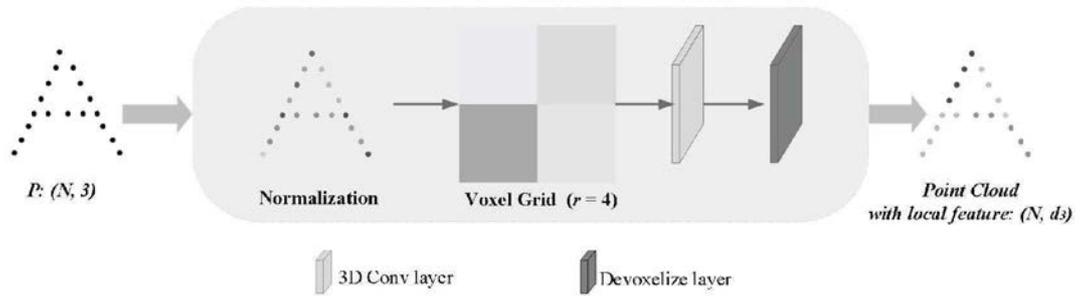


图5

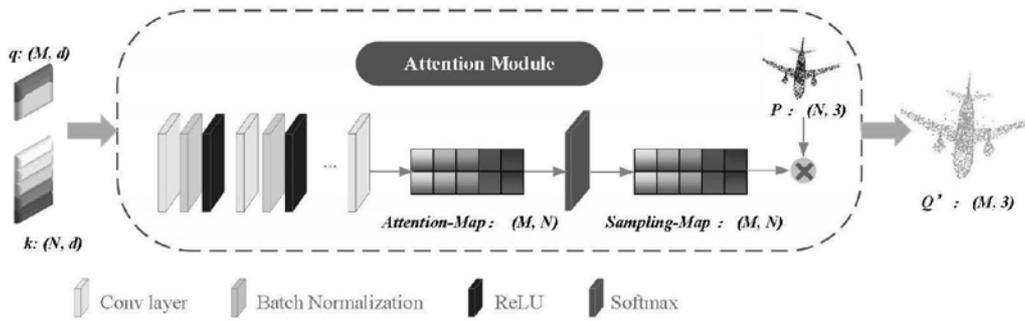


图6

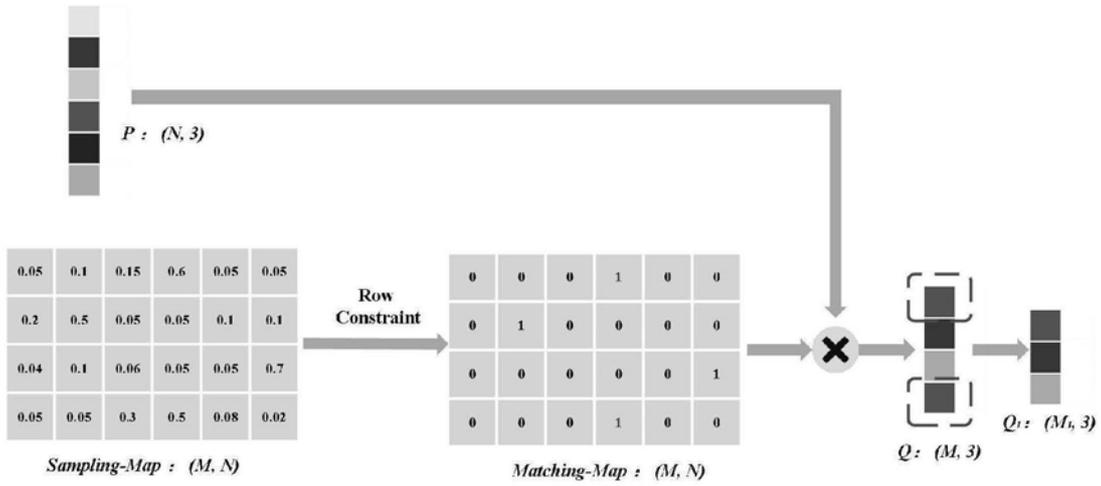


图7

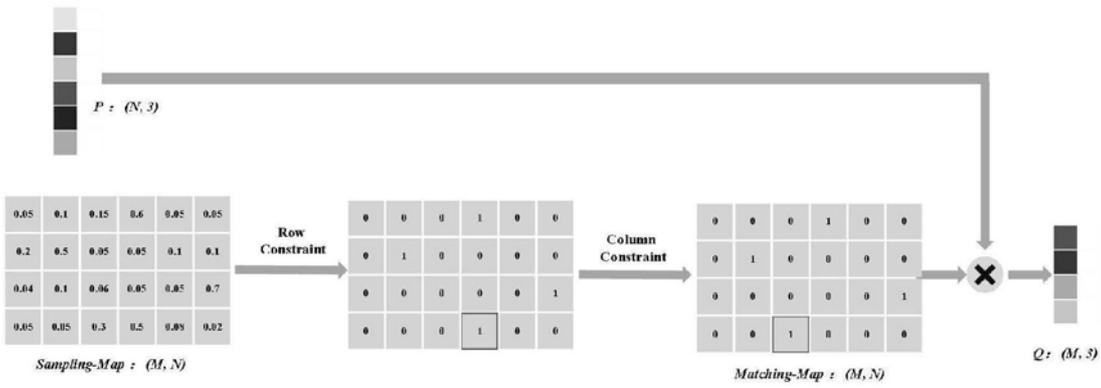


图8

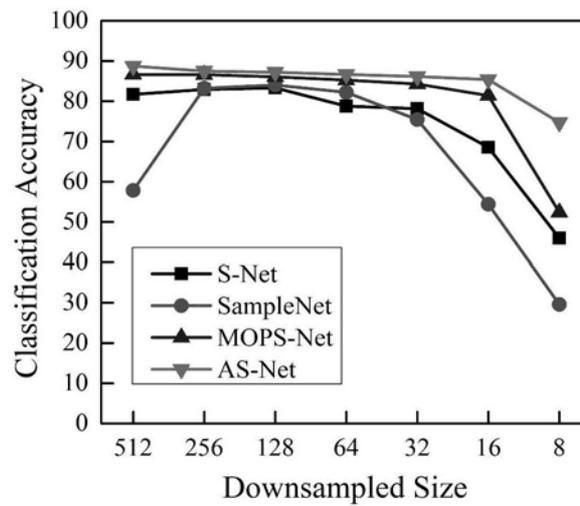


图9

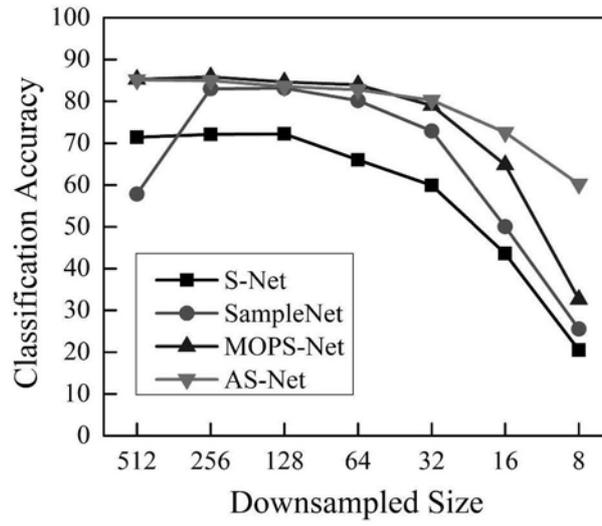


图10

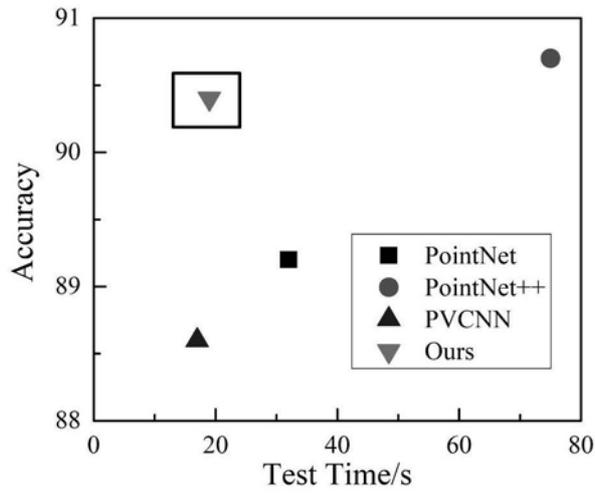


图11

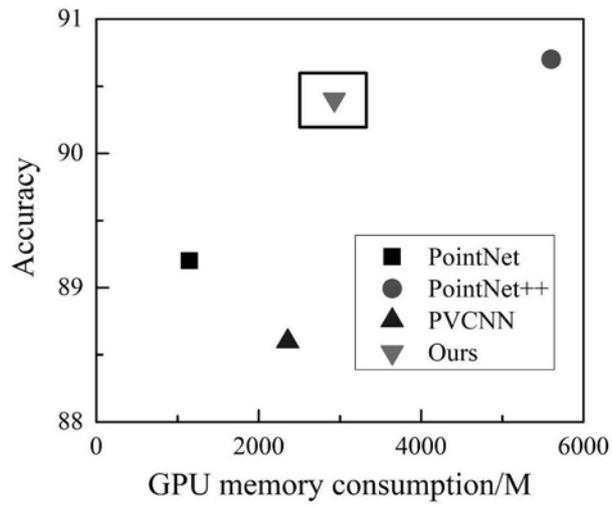


图12

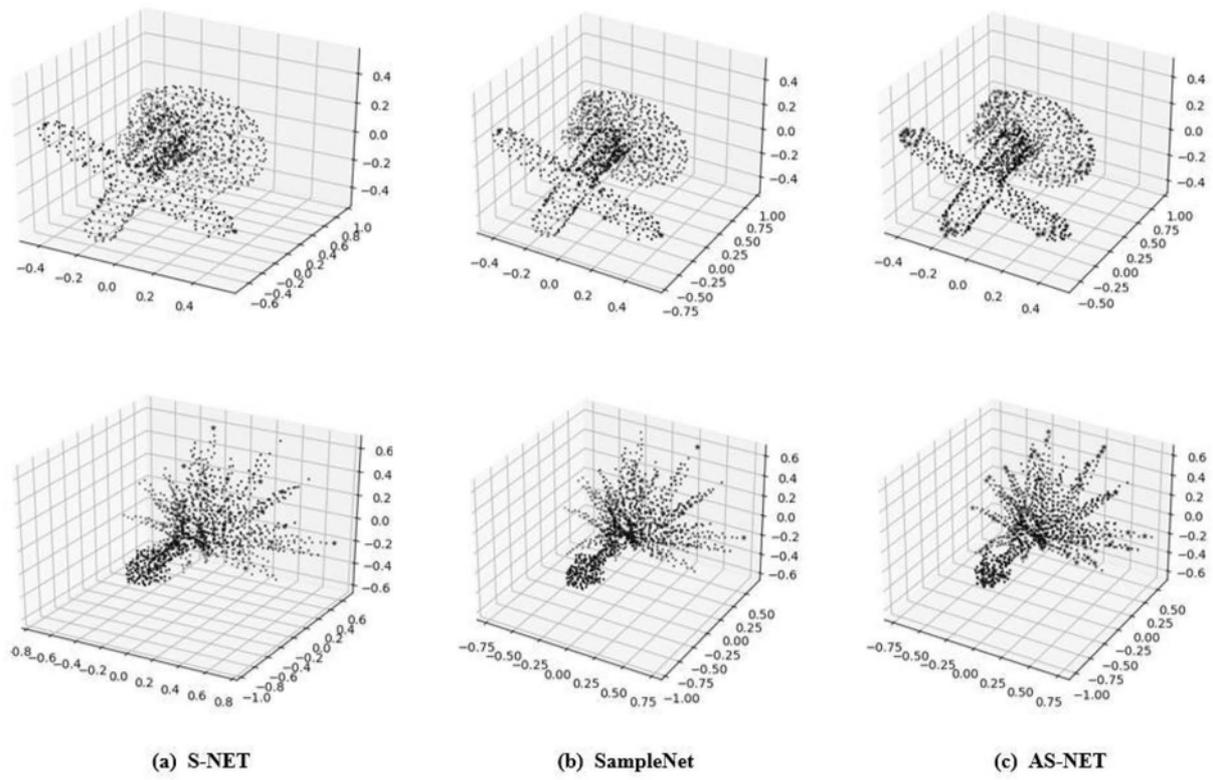


图13