



(12) 发明专利申请

(10) 申请公布号 CN 114444613 A

(43) 申请公布日 2022. 05. 06

(21) 申请号 202210127425.3

G06V 10/764 (2022.01)

(22) 申请日 2022.02.11

G06V 10/762 (2022.01)

(71) 申请人 吉林大学

G06V 10/80 (2022.01)

地址 130012 吉林省长春市长春高新技术  
产业开发区前进大街2699号

G06V 10/82 (2022.01)

(72) 发明人 刘振泽 陈金炎 董迪锴 吴闯  
何井全 张渝敏 周亚仑

(74) 专利代理机构 长春吉大专利代理有限责任  
公司 22201

专利代理师 朱世林

(51) Int. Cl.

G06K 9/62 (2022.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

G06T 7/11 (2017.01)

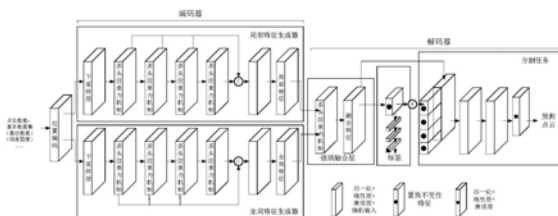
权利要求书3页 说明书7页 附图3页

(54) 发明名称

一种基于3D点云信息的物体分类与物体分割方法

(57) 摘要

本发明涉及一种基于3D点云信息的物体分类与物体分割方法,属于3D点云处理领域,尤其涉及无序点云难处理、点云特征难提取领域。本发明在提取点云特征中,为了提取更多的有效信息,本发明提出并使用了一种全新的联合式全局特征与局部特征结构。在提取高维特征时提出一种新式的提取机制代替传统的对称函数。通过大量实验可以表明我们设计的神经网络可以解决多种点云任务,包括但不限于物体分类、物体分割、场景分割等。本发明的有益效果:所设计的针对点云识别和分割算法主要设计新的Mix-Net 专用神经网络设计,并从下采样算法,特征值提取以及融合机制等多个模块进行创新发明,对于点云识别和分割准确率都有明显提升。



1. 一种基于3D点云信息的物体分类与物体分割方法,包含以下步骤:

步骤一、对于输入点云信息增加位置编码

对于输入点云信息,增加输入相对位置信息,相对位置信息增加在第一次输入层中,以及注意力函数中,经过大量实验可以证明相对位置编码的加入可以提升网络性能,提高最后输出准确率;

步骤二、数据降采样算法

对于输入使用降采样算法,降低整体网络的复杂度,加快网络训练速度,抑制过拟合,以一种基于注意力机制的降采样算法代替最远点距离算法;

步骤三,相对位置信息融合对的多头注意力机制

作为网络中提取特征的主要层,使用相对位置信息融合对的多头注意力层代替卷积层,经过实验的表面多头注意力层增加网络性能,提高最后输出准确率;

步骤四,点云编码器结构设计

设计的点云编码器结构用于将输入点云数据提取出高维特征向量,该特征向量包括点云全局特征信息,点云局部特征信息,并且在融合特征取得点云整体特征时,提供一种新式的特征融合机制代替对称函数;

步骤五,点云解码器结构设计

针对不同的点云处理任务设计不同的解码器结构,相似不相同;

步骤六,设计Mix-Net网络训练方法

根据不同的点云处理任务,针对不同数据集,设计损失函数,设计优化器,以及设计多种训练参数指标。

2. 根据权利要求1所述的一种基于3D点云信息的物体分类与物体分割方法,其特征在于,步骤一所述位置编码的实现过程如下:由于输入点云的容器存在多种储存方法,点与点之间即同一输入不同元素之间很难在网络中直接获得相对位置信息,引入相对位置信息,来增加点与点之间的位置特征提取,位置编码公式如下:

$$\rho = MLP(P_i - \bar{P})$$

其中, $P_i$ 是输入数据或者特征, $\bar{P}$ 是输入数据或者特征的平均值,MLP是常用的多层感知机,在多次实验发现,位置信息的加入在第一层输入与注意力机制模块中对整体网络性能有明显的提升。

3. 根据权利要求1所述的一种基于3D点云信息的物体分类与物体分割方法,其特征在于,步骤二所述数据降采样算法的实现过程如下:

(1) 使用多头注意力机制为基础的算法来代替传统的最远点采样算法

将卷积输出的特征通过多头注意力机制得出注意力得分,通过选取前k个最大注意力得分值达到与最远点采集k个值相似的效果,通过实验证明本发明使用的降采样算法更适合点云,针对多种不同顺序输入可以取得相似特征输出,可以解决点云无序性问题;

(2) 根据取得的k个最大值,在点云输入中匹配确定的序列点,将这些序列点表示为关键点,所述关键点表示点云最关键部分,由关键点可视化可以容易看出,关键是点云的骨架信息;

(3) 根据选取的关键点,使用k-means算法选取每个关键点周边的特征点,取得所有符

合k-means算法的特征点进行特征提取,使用多尺度特征融合方式将多层特征进行融合,输出降采样后的高维特征。

4. 根据权利要求1所述的一种基于3D点云信息的物体分类与物体分割方法,其特征在于,步骤三所述相对位置信息融合对的多头注意力机制的实现过程如下:

(1) 在输入断使用相对距离可以增加网络的整体性能,经过大量实验可以看出,在注意力函数中增加相对距离 $\rho$ 依旧可以增加整体网络性能,将相对距离增加至注意力函数中,公式如下:

$$Attention(Q, K, V) = softmax\left(\frac{Q(K^T + \rho)}{\sqrt{d_k}}\right)(V + \rho)$$

其中, $Q, K, V$ 是输入根据各自变化矩阵得到的query, key, value矩阵, softmax为激活函数, $\rho$ 为相对距离,本发明在K的转置处增加一个相对距离值,在V处增加一个相对距离值,设计的注意力函数在输入中提取更加充分的元素与元素之间的信息,更加适合点云处理;

(2) 引入多头注意力机制:将输入通过多个独立的注意力层,最后每一层输出结果级联起来得到多头注意力层的输出,通过实现并行运行,在处理输入数据中有明显速度优势,多头注意力机制的公式如下:

$$Q_i = XW^{Q_i}, K_i = XW^{K_i}, V_i = XW^{V_i},$$

$$F_i = Attention(Q_i, K_i, V_i), i = 1, \dots, h,$$

$$MultiHead(Q, K, V) = Concat(F_1, F_2, \dots, F_h) W^0$$

其中, $X$ 代表输入, $W^{Q_i} \in \mathbb{R}^{d_{model} \times d_k}$ ,  $W^{K_i} \in \mathbb{R}^{d_{model} \times d_k}$  和  $W^{V_i} \in \mathbb{R}^{d_{model} \times d_v}$  是三组不同的线性矩阵, $h$ 代表多头注意力机制的个数,本发明使用 $h=4$ ,  $W^0 \in \mathbb{R}^{hd_v \times d_{model}}$ ,代表输出特征矩阵, $F_i$ 代表每一个多头注意力层最终输出的值,Concat代表级联,多头注意力机制将输入分为 $h$ 个独立的注意力机制层,并且并行运算每一个注意力机制层。

5. 根据权利要求1所述的一种基于3D点云信息的物体分类与物体分割方法,其特征在于,步骤四所述点云编码器结构设计的实现过程如下:

(1) 对于输入,一个点云样本的输入是 $P \in \mathbb{R}^{N \times d}$ ,其中 $N$ 代表输入个数,每一个输入点你云均有 $d$ 维向量,常见 $d$ 为3或6, $d$ 为3时,点云只包含位置信息, $d$ 为6时,点云包含位置信息和法向量信息, $F_p$ 是通过相对位置编码层的输出, $F_e$ 是 $F_p$ 经过下采样层的输出特征向量, $F_e \in \mathbb{R}^{N \times d_e}$ , $F_o$ 为 $F_e$ 经过4层多头注意力机制层的输出, $F_o \in \mathbb{R}^{N \times d_o}$ , $F_o$ 输出特征向量公式如下:

$$MSA(A) = Attention(A, A, A),$$

$$F_1 = MSA(F_e),$$

$$F_i = MSA(F_{i-1}), i = 2, 3, 4,$$

$$F_o = Concat(F_1, F_2, \dots, F_4) \cdot W^0,$$

其中,每一个MSA( $F_i$ )代表第 $i$ 层注意力层输出特征,每层的输出维度都与输入维度一致, $W_o$ 是线性层的权重,在网络训练过程中会变化:

(2) 为了提取全局特征 $F_{global}$ 与局部特征 $F_{local}$ 的点云特征,设计一个新式的特征融合层代替了传统的对称函数,新提出的融合层依旧可以解决点云的无序性问题,并在很大程度上减少点云因为降维而损失的特征,新的融合层公式如下:

$MCA(A, B) = \text{Attention}(A, B, B)$ ,

$F_{lg} = MCA(F_{local}, F_{global})$ 。

## 一种基于3D点云信息的物体分类与物体分割方法

### 技术领域

[0001] 本发明涉及3D点云处理领域,尤其涉及无序点云难处理、点云特征难提取领域,提出一种基于3D点云信息的物体分类与物体分割方法。

### 背景技术

[0002] 3D点云技术在目前智能驾驶,虚拟现实、机器人领域都有应用。主要目的是从3D点云中提取有效特征去描述点云的类别或者分割部分。处理3D点云主要需要注意:点云的顺序无序性和点云的置换无序性。顺序无序性是:点云在容器中可以有N种表征形式,但无论那种表征方式依旧描述同一个物体。置换无序性是点云在度量空间中经过旋转平移的简单操作,在标准坐标系下,表征点云的容器值随之改变,但依旧表示同一个目标物体。在上述两种情况下,容器的多种表征依旧可以提取类似或者相同的特征来表征点云,是点云处理的难点。目前现有技术只是针对点云整体特征提取,对于涉及高效的端对端神经网络依旧困难。与现有的神经网络相比,在物体分类、物体分割任务中准确率均有显著提升。

### 发明内容

[0003] 为克服现有技术的不足,本发明提供一种基于3D点云信息的物体分类与物体分割方法,旨在提升点云识别和分割准确率,减少因为降维或者体素变化带来的信息损失;并可广泛用于多种传感感知设备。

[0004] 下面结合附图对本发明做进一步的说明:

[0005] 所述一种基于3D点云信息的物体分类与物体分割方法,包含以下步骤:

[0006] 步骤一、对于输入点云信息增加位置编码。对于输入点云信息,增加输入相对位置信息。相对位置信息增加在第一次输入层中,以及注意力函数中。经过大量实验可以证明相对位置编码的加入可以提升网络性能,提高最后输出准确率。

[0007] 步骤二、数据降采样算法。对于输入使用降采样算法,可以降低整体网络的复杂度,加快网络训练速度,还可以抑制过拟合。本发明提出一种新式的基于注意力机制的降采样算法代替传统的最远点距离算法。

[0008] 步骤三,相对位置信息融合对的多头注意力机制。作为网络中提取特征的主要层,我们使用相对位置信息融合对的多头注意力层代替传统的卷积层。经过实验可以表面多头注意力层可以增加网络性能,提高最后输出准确率。

[0009] 步骤四,点云编码器结构设计。本发明设计的点云编码器结构用于将输入点云数据提取出高维特征向量,该特征向量包括点云全局特征信息,点云局部特征信息。并且在融合特征取得点云整体特征时本发明提出一种新式的特征融合机制代替传统的对称函数。

[0010] 步骤五,点云解码器结构设计。本发明针对不同的点云处理任务设计不同的解码器结构,相似不相同。

[0011] 步骤六,设计Mix-Net网络训练方法。主要根据不同的点云处理任务,针对不同数据集,设计损失函数,设计优化器,以及设计多种训练参数指标。

[0012] 进一步地,步骤一所述位置编码的实现过程如下:由于输入点云的容器存在多种储存方法,点与点之间(同一输入不同元素之间)很难在网络中直接获得相对位置信息。所以引入相对位置信息,来增加点与点之间的位置特征提取。本发明的位置编码公式如下:

$$[0013] \quad \rho = MLP(P_i - \bar{P})$$

[0014] 其中, $P_i$ 是输入数据或者特征, $\bar{P}$ 是输入数据或者特征的平均值,MLP是常用的多层感知机。本发明在多次实验发现,位置信息的加入在第一层输入与注意力机制模块中对整体网络性能有明显的提升。

[0015] 进一步地,步骤二所述数据降采样算法的主要实现过程如下:

[0016] (1)、使用多头注意力机制为基础的算法来代替传统的最远点采样算法。将卷积输出的特征通过多头注意力机制得出注意力得分,通过选取前k个最大注意力得分值达到与最远点采集k个值相似的效果。通过实验证明本发明使用的降采样算法更适合点云,针对多种不同顺序输入可以取得相似特征输出,可以解决点云无序性问题。

[0017] (2)、根据取得的k个最大值,在原点云输入中匹配确定的序列点,将这些序列点表示为关键点。关键点表示点云最关键部分,由关键点可视化可以容易看出,关键是点云的骨架信息

[0018] (3)、根据选取的关键点,使用k-means算法选取每个关键点周边的特征点,取得所有符合k-means算法的特征点进行特征提取,使用多尺度特征融合方式将多层特征进行融合。输出降采样后的高维特征。

[0019] 进一步地,步骤三所述相对位置信息融合对的多头注意力机制的实现过程主要如下:

[0020] (1)、在输入断使用相对距离可以增加网络的整体性能,经过大量实验可以看出,在注意力函数中增加相对距离 $\rho$ 依旧可以增加整体网络性能。本发明将相对距离增加至注意力函数中,公式如下:

$$[0021] \quad Attention(Q, K, V) = softmax\left(\frac{Q(K^T + \rho)}{\sqrt{d_k}}\right)(V + \rho)$$

[0022] 其中,Q,K,V是输入根据各自变化矩阵得到的query,key,value矩阵,softmax为激活函数, $\rho$ 为相对距离,本发明在K的转置处增加一个相对距离值,在V处增加一个相对距离值。本发明设计的注意力函数在输入中可以提取更加充分的元素与元素之间的信息,更加适合点云处理。

[0023] (2)、引入多头注意力机制。将输入通过多个独立的注意力层,最后每一层输出结果级联起来得到多头注意力层的输出,该方法可以实现并行运行,在处理输入数据中有明显速度优势。多头注意力机制的公式如下:

$$[0024] \quad Q_i = XW^{Q_i}, K_i = XW^{K_i}, V_i = XW^{V_i},$$

$$[0025] \quad F_i = Attention(Q_i, K_i, V_i), i = 1, \dots, h,$$

$$[0026] \quad MultiHead(Q, K, V) = Concat(F_1, F_2, \dots, F_h) W^0$$

[0027] 其中,X代表输入, $W^{Q_i} \in \mathbb{R}^{d_{model} \times d_k}$ , $W^{K_i} \in \mathbb{R}^{d_{model} \times d_k}$ 和 $W^{V_i} \in \mathbb{R}^{d_{model} \times d_v}$ 是三组不同的线性矩阵,h代表多头注意力机制的个数,本发明使用 $h=4$ , $W^0 \in \mathbb{R}^{hd_v \times d_{model}}$ ,代表

输出特征矩阵。 $F_i$ 代表每一个多头注意力层最终输出的值,Concat代表级联,多头注意力机制将输入分为h个独立的注意力机制层,并且并行运算每一个注意力机制层。

[0028] 进一步地,步骤四所述点云编码器结构设计的实现过程如下:

[0029] (1)、对于输入,一个点云样本的输入是 $P \in \mathbb{R}^{N \times d}$ ,其中N代表输入个数,每一个输入点你云均有d维向量,常见d为3或6,d为3时,点云只包含位置信息,d为6时,点云包含位置信息和法向量信息。 $F_p$ 是通过相对位置编码层的输出, $F_e$ 是 $F_p$ 经过下采样层的输出特征向量, $F_e \in \mathbb{R}^{N \times d_e}$ 。 $F_o$ 为 $F_e$ 经过4层多头注意力机制层的输出, $F_o \in \mathbb{R}^{N \times d_o}$ , $F_o$ 输出特征向量公式如下:

[0030]  $MSA(A) = \text{Attention}(A, A, A)$ ,

[0031]  $F_1 = MSA(F_e)$ ,

[0032]  $F_i = MSA(F_{i-1}), i = 2, 3, 4$ ,

[0033]  $F_o = \text{Concat}(F_1, F_2, \dots, F_4) \cdot W_o^0$ ,

[0034] 其中,每一个MSA( $F_i$ )代表第i层注意力层输出特征,每层的输出维度都与输入维度一致。 $W_o$ 是线性层的权重,在网络训练过程中会变化。

[0035] (2)、为了提取全局特征 $F_{\text{global}}$ 与局部特征 $F_{\text{local}}$ 的点云特征,本发明设计一个新式的特征融合层代替了传统的对称函数。新提出的融合层依旧可以解决点云的无序性问题,并在很大程度上减少点云因为降维而损失的特征,新的融合层公式如下:

[0036]  $MCA(A, B) = \text{Attention}(A, B, B)$ ,

[0037]  $F_{1g} = MCA(F_{\text{local}}, F_{\text{global}})$ 。

[0038] 本发明的有益效果:本发明专利所设计的针对点云识别和分割算法主要设计新的Mix-Net专用神经网络设计,并从下采样算法,特征值提取以及融合机制等多个模块进行创新发明,对于点云识别和分割准确率都有明显提升。并且该网络为端对端网络不需要增加其他辅助模块来调整输入,并且直接针对输入点云直接进行处理,减少因为降维或者体素变化带来的信息损失。已经在常用数据集上,常用雷达设备,常用RGB-D摄像机做了充分实验,本发明所提出的算法可广泛用于多种传感感知设备。

## 附图说明

[0039] 图1 Mix-Net网络示意图

[0040] 图2 Mix-Net网络总体结构流程图

[0041] 图3 Mix-Net网络下采样模块

[0042] 图4 Mix-Net网络在分类任务中的表现

[0043] 图5 Mix-Net网络在分割任务中对于飞机模型分割,左图为人标签分割,右图为网络输出结果

[0044] 图6 Mix-Net网络在分割任务中对于水杯模型分割,左图为人标签分割,右图为网络输出结果

[0045] 图7 Mix-Net网络在分割任务中对于椅子模型分割,左图为人标签分割,右图为网络输出结果

## 具体实施方式

[0046] 下面结合附图对本发明的实施方式作出详细的说明,使得发明目的、技术方案、特征等更加易于理解,显然,所描述的实施例仅仅是本发明一部分实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。实施例仅用于解释本发明,而不是对本发明的限定。

[0047] 本发明提出一种基于3D点云信息的物体分类与物体分割方法,主要结构流程图如图1、图2,完整方法可由以下步骤说明:

[0048] 步骤一、输入增加位置编码。位置编码时自然语言处理常用的方法。可以增加输入特征之间的物理联系,有助于提升网络整体性能,提升分类与分割任务的准确率。在输入过程中需要增加位置信息,位置信息一般可以分为相对位置与绝对位置。在点云处理中,输入数据已经包含了明确的三维位置信息。但是由于输入点云的容器存在多种储存方法,点与点之间(同一输入不同元素之间)很难在网络中直接获得相对位置信息。所以引入相对位置信息,来增加点与点之间的位置特征提取。本发明的相对位置编码公式如下:

$$[0049] \quad \rho = MLP(P_i - \bar{P})$$

[0050] 其中, $P_i$ 是输入数据或者特征, $\bar{P}$ 是输入数据或者特征的平均值,MLP是常用的多层感知机。本发明在多次实验发现,位置信息的加入在第一层输入与注意力机制模块中对整体网络性能有明显的提升。

[0051] 步骤二、数据降采样算法。降采样算法是将输入数据或特征降数据输出,可以方便后续网络针对强特征信息,降低整体网络的复杂度,降低网络整体参数,加快网络训练过程。整体结构图如图3表示,具体实现由以下步骤:

[0052] 1、针对输入点云使用卷积处理。卷积网络容易提取低维特征信息相比较于注意力机制,通过两层卷积层,每一个卷积层包含一维卷积、批量归一化层和ReLU激活函数层。

[0053] 2、使用多头注意力机制为基础的算法来代替传统的最远点采样算法。将卷积输出的特征通过多头注意力机制得出注意力得分,通过选取前k个最大注意力得分值达到与最远点采集k个值相似的效果。通过实验证明本发明使用的降采样算法更适合点云,针对多种不同顺序输入可以取得相似特征输出,可以解决点云无序性问题。

[0054] 3、根据取得的k个最大值,在原点云输入中匹配确定的序列点,将这些序列点表示为关键点。关键点表示点云最关键部分,由关键点可视化可以容易看出,关键点是点云的骨架信息。

[0055] 4、根据选取的关键点,使用k-means算法选取每个关键点周边的特征点,取得所有符合k-means算法的特征点进行特征提取,使用多尺度特征融合方式将多层特征进行融合。输出降采样后的高维特征。

[0056] 步骤三、相对位置信息融合的多头注意力机制。具体实现由以下步骤:

[0057] 1、注意力机制,是运用在自然语言处理领域,针对输入语句计算各个单词之间的相关度。 $Q, K, V$ 是输入语句根据各自变化矩阵得到的query, key, value矩阵。针对输入  $F_{inx} \in \mathbb{R}^{N \times d_x}$ ,  $F_{iny} \in \mathbb{R}^{N \times d_y}$ , 当  $d_x = d_y$  时为自注意力机制; 当  $d_x \neq d_y$  不同时, 为交叉注意力机制, 注意力机制计算公式如下:

$$[0058] \quad Q = F_{inx} \cdot W^Q, K = F_{iny} \cdot W^K, V = F_{iny} \cdot W^V$$



[0059] 其中,  $W^Q, W^K, W^V$  都是线性矩阵, 并且  $d_q, d_k, d_v$  分别是 query, key, value 的维度。可以看出 query 是通过  $F_{inx}$  决定, key, value 是由  $F_{iny}$  决定。本发明用  $d_q = \frac{d_k}{4}$ , 可以增加注意力机制的计算效率。

[0060] 一个注意力机制层, 主要是根据 Q 和相关 K 的乘积, 经过特定激活函数乘以对应的 V 得到输出。常用的注意力函数公式如下:

$$[0061] \quad Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

[0062] 其中, 注意力权重使用 Q 与对应 K 的转置进行向量相乘, 并除以  $d_k$ , 可以防止整体乘积得到的值过大。使用 softmax 函数得到最终的注意力权重。最后注意力权重与对应的 V 相乘得到注意力输出。

[0063] 在输入使用相对距离可以增加网络的整体性能, 经过大量实验可以看出, 在注意力函数中增加相对距离  $\rho$  依旧可以增加整体网络性能。本发明将相对距离增加到注意力函数中, 公式如下:

$$[0064] \quad Attention(Q, K, V) = softmax\left(\frac{Q(K^T + \rho)}{\sqrt{d_k}}\right)(V + \rho)$$
 其中, 本发明在 K 的转置处增加一个相对距离值, 在 V 处增加一个相对距离值。本发明设计的注意力函数在输入中可以提取更加充分的元素与元素之间的信息, 更加适合点云处理。

[0065] 2、多头注意力机制

[0066] 由于受到输入向量或者特征过大的限制, 单头的注意力机制可能提取的特征不够充分。为了解决这个问题, 本发明使用多头注意力机制, 该方法可以加快网络整体训练速度, 较小过拟合。主要思想为: 将输入通过多个独立的注意力层, 最后每一层输出结果级联起来得到多头注意力层的输出, 该方法可以实现并行运行, 在处理输入数据中有明显速度优势。多头注意力机制公式如下:

[0067]  $Q_i = XW^{Q_i}, K_i = XW^{K_i}, V_i = XW^{V_i},$

$$[0068] \quad F_i = Attention(Q_i, K_i, V_i), i = 1, \dots, h,$$

$$[0069] \quad MultiHead(Q, K, V) = Concat(F_1, F_2, \dots, F_h)W^0$$

[0070] 其中, h 代表多头注意力机制的个数, 本发明使用  $h=4$ 。其中  $W^0 \in \mathbb{R}^{hd_v \times d_{model}}$ , 代表输出特征矩阵。  $F_i$  代表每一个多头注意力层输出的值。

$W^{Q_i} \in \mathbb{R}^{d_{model} \times d_k}, W^{K_i} \in \mathbb{R}^{d_{model} \times d_k}$  和  $W^{V_i} \in \mathbb{R}^{d_{model} \times d_v}$  是三组不同的线性矩阵。最终, 多头注意力机制将输入分为 h 个独立的注意力机制层, 并且并行运算每一个注意力机制层。

[0071] 步骤四、点云编码器结构。Mix-Net 的总体结构图为图 2。Mix-Net 的目标是通过网络编码器结构来取得一个新的高维向量来表征点云。该高维向量对于不同输入有一定的鲁棒性, 对于输入点云旋转和平移以及点云的无序输入都有泛性效果。该高维特征向量可以作为大多数点云处理任务的关键向量。本发明中, 该特征向量定义为点云特征向量, 结合局部特征向量和全局特征向量。

[0072] 本发明设计的编码器结构, 针对输入点云增加相对位置信息, 将通过降采样层得

到较为明显的点云特征数据,降低整体网络的复杂性,使用多尺度特征融合算法,并且使用了注意力机制为分数指标的新式降采样算法。根据输出的降采样特征,分别提取点云全局特征与点云局部特征。其中本发明设置全局特征向量提取的维度为512,局部特征向量的维度为128。经过4层多头注意力机制,并且级联每个多头注意力机制的输出,在每个多头注意力层中还增加残差结构,数据归一化层和ReLu激活层,可以加快网络收敛速度。得到高维的全局特征向量 $F_{\text{global}}$ 和局部特征向量 $F_{\text{local}}$ 。

[0073] 1、对于输入,一个点云样本的输入是 $P \in \mathbb{R}^{N \times d}$ ,其中N代表输入个数,每一个输入点云均有d维向量,常见d为3或6,d为3时,点云只包含位置信息,d为6时,点云包含位置信息和法向量信息。 $F_p$ 是通过相对位置编码层的输出, $F_e$ 是 $F_p$ 经过下采样层的输出特征向量, $F_e \in \mathbb{R}^{N \times d_e}$ 。 $F_o$ 为 $F_e$ 经过4层多头注意力机制层的输出, $F_o \in \mathbb{R}^{N \times d_o}$ , $F_o$ 输出特征向量公式如下:

[0074]  $MSA(A) = \text{Attention}(A, A, A)$ ,

[0075]  $F_1 = MSA(F_e)$ ,

[0076]  $F_i = MSA(F_{i-1})$ ,  $i = 2, 3, 4$ ,

[0077]  $F_o = (F_1 \oplus F_2 \oplus F_3 \oplus F_4) \cdot W^o$ ,

[0078] 其中,每一个MSA( $F_i$ )代表第i层注意力层输出特征,每层的输出维度都与输入维度一致。 $W_o$ 是线性层的权重,在网络训练过程中会变化。

[0079] 2、为了提取全局特征 $F_{\text{global}}$ 与局部特征 $F_{\text{local}}$ 的点云特征,本发明设计一个新式的特征融合层代替了传统的对称函数。新提出的融合层依旧可以解决点云的无序性问题,还可以在很大程度上减少点云因为降维而损失的特征,新的融合层公式如下:

[0080]  $MCA(A, B) = \text{Attention}(A, B, B)$ ,

[0081]  $F_{1g} = MCA(F_{\text{local}}, F_{\text{global}})$ ,

[0082] 步骤五、点云解码器结构。针对不同点云处理任务,由于输出的维度不同,在解码器部分的结构也是不同。

[0083] 1、物体分类任务。分类网络显示在整体流程图解码器部分,见图2。将高维特征 $F_{1g}$ 转化为 $N_c$ 种物体目录分数(比例如,飞机,桌子,椅子),并使用数字打标,例如用0代表飞机类别,1代表桌子类别,2代表椅子类别,一共存在 $N_c = 40$ 种物体类别。本发明将高维特征放入预设的分类解码器中,其中包含一个卷积网络,两个级联的线性、归一化、ReLu激活函数组合层,并且加入dropout层防止过拟合。最后,使用线性函数预测最后的分类分数 $S \in \mathbb{R}^{N_c}$ 。本次预测分数最大的序列对应的类别则为本次神经网络预测类别。

[0084] 2、物体分割任务。针对物体分割,需要将网络对于输入的每一个点都标记一个标签,例如,桌子面部分,桌子腿部分,我们需要对每一个点都进行预测。一般每一种物体类别分为2-4个部分,使用顺序依次标记,一共计有20种物体类别,50种物体分割区域。本发明使用高维特征 $F_{1g}$ ,为了学习到常见模型的多个部分,我们将标签转换为一维目标向量与高维特征 $F_{1g}$ 进行级联。对于后续的物体分割解码器与物体分类解码器相似,主要使用卷积对高维信息进行提取特征,使用两个级联的线性、归一化、ReLu激活函数组合层进行解码,使用dropout层防止过拟合。最后我们可以预测每一个点的分割分数 $P \in \mathbb{R}^{N \times N_s}$ ,每一个点被确定为最高得分对应的标签部分。

[0085] 步骤六、Mix-Net网络训练方法。

[0086] 1、点云分类任务与点云分割任务数据集。

[0087] 点云分类任务中,我们使用ModelNet40测试集验证我们的Mix-Net网络,该数据集每一个点云样本包含10000个点云信息,并且包含40种物体类别,其中包含9843训练样本以及2468个测试样本。在网络训练部分,我们对于输入点云进行了数据增强,包括范围缩放,缩放比例为 $[0.8, 1.25]$ ,区域平移,平移范围为 $[-0.1, 0.1]$ ,还包括随机损失输入,损失比例为 $[0, 12.5\%]$ 。在实验过程中,我们使用 $N=1024, D=6$  (xyz+normal), local\_dim=256, global\_dim=512, K=50。点云分类任务结果如图4所示,在测试集上分类准确率可达92.5%。

[0088] 点云分割任务中,我们使用非常有挑战性的ShapeNet数据集来验证我们的Mix-Net网络,该数据集包括13998个训练集和2874个训练集。包含16种不同的类别,每种类别包含2到4个部分标签,总共有50个部分标签。分割目标是对于每一个输入点进行目标分类。在网络训练部分,我们对于输入点云进行了数据增强,包括范围缩放,缩放比例为 $[0.8, 1.25]$ ,区域平移,平移范围为 $[-0.1, 0.1]$ ,还包括随机损失输入,损失比例为 $[0, 12.5\%]$ 。在实验过程中,我们使用 $N=1024, D=6$  (xyz+normal), local\_dim=256, global\_dim=512, K=50。点云分割任务部分结果如图5所示,左图为打标点云,右图为神经网络预测点云。

[0089] 2、Mix-Net其他训练细节:针对点云分类任务以及点云分割任务。损失函数均使用交叉熵损失函数,在训练中,使用随机梯度下降方法最为优化方法,并设置动量值为0.9。学习率均设置为 $1e-4$ , batch\_size=32,使用gpu为RTX-3070,使用神经网络架构为pytorch。

[0090] 综上所述,本发明提出针对点云处理的分类与分割算法,能够有效、准确的实现点云的物体分类、物体分割任务。与常用的算法相比, Mix-Net神经网络更加适合点云处理,针对点云的置换无序性与顺序无序性,在新式特征提取结构与新式融合结构都能较好的解决,降低网络的整体参数量,提高整体输出的准确率。

[0091] 以上所述是结合具体实施方式对本发明所作的进一步详细说明,不能认定本发明具体实施仅局限于此。对于本发明专利所属及相关技术领域的技术人员来说,在基于本发明专利技术思路前提下,所作的拓展以及操作方法、结构的替换,都应当落在本发明专利保护范围之内。

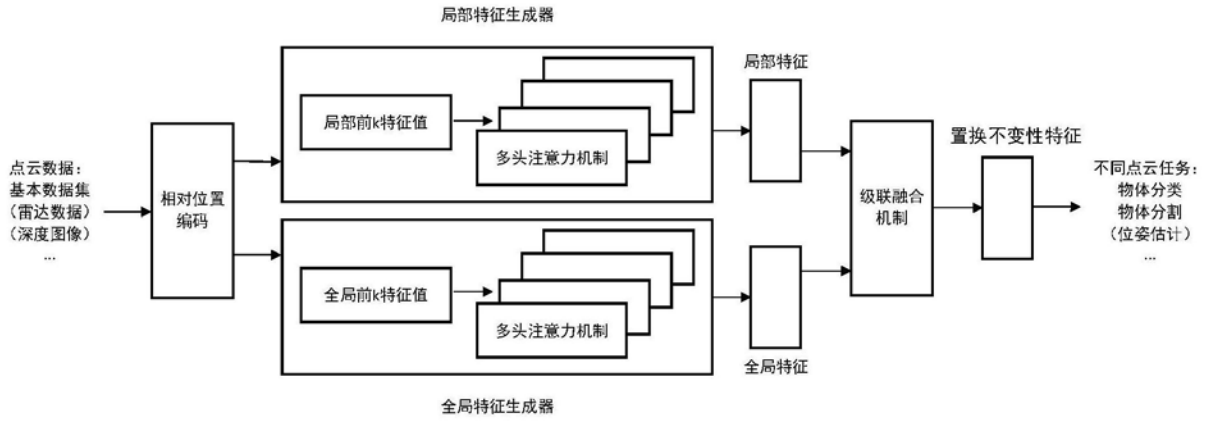


图1

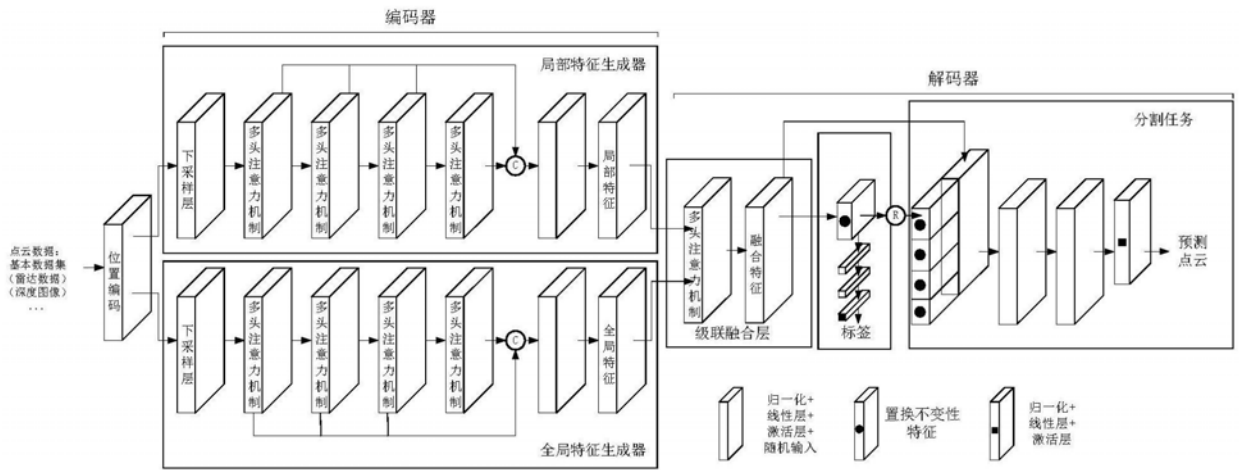


图2

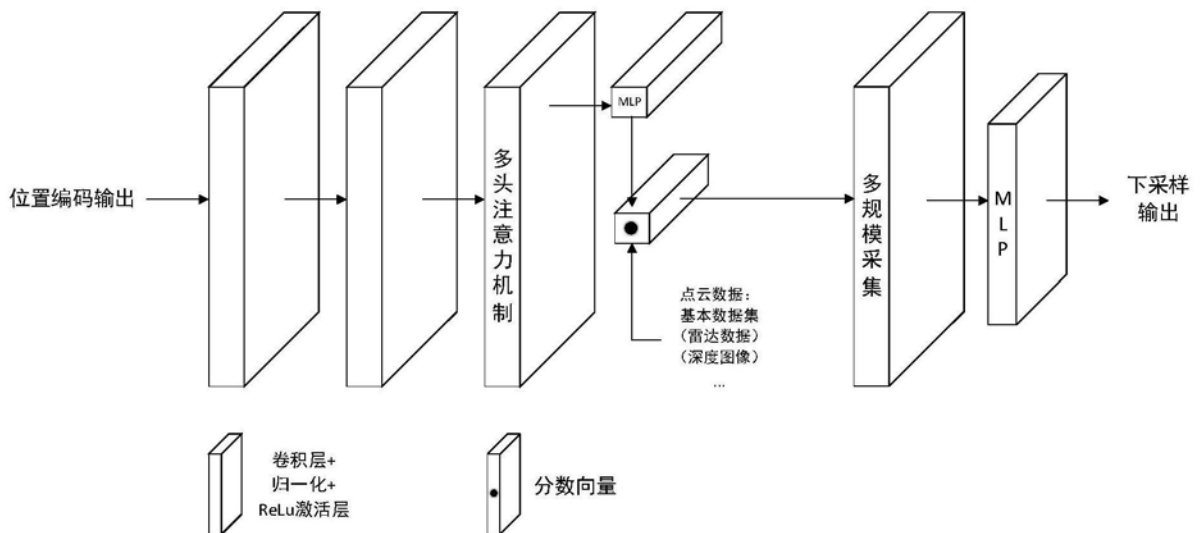


图3

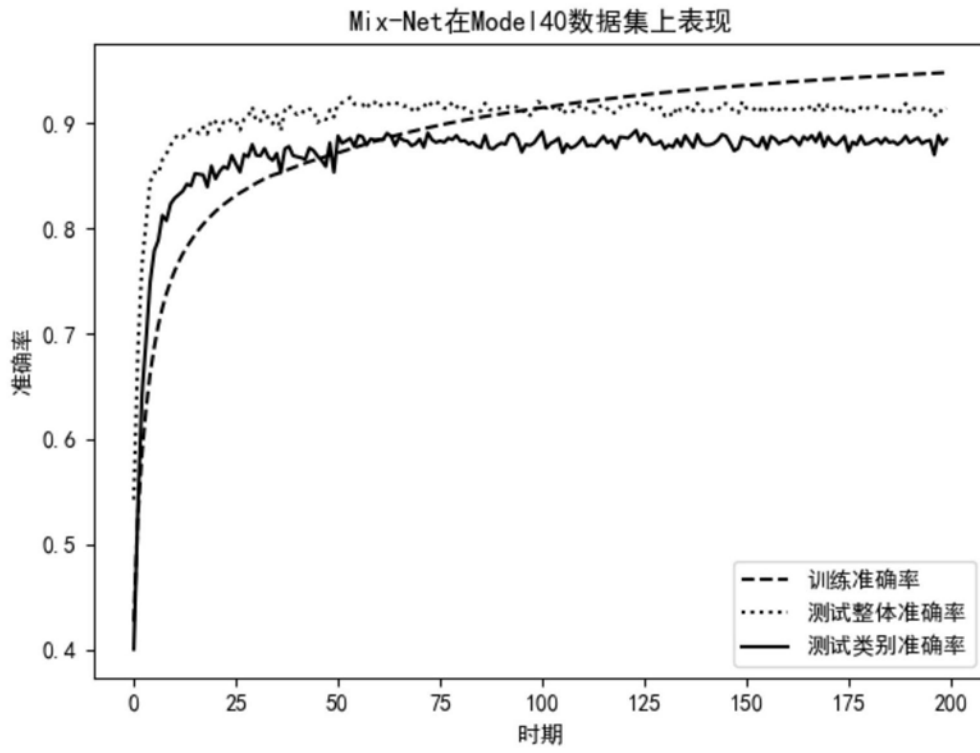


图4

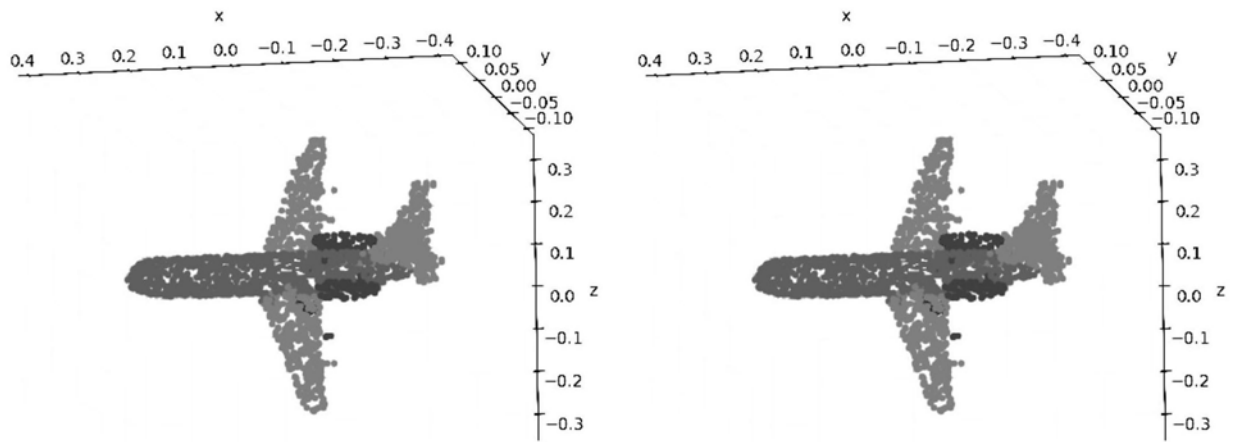


图5

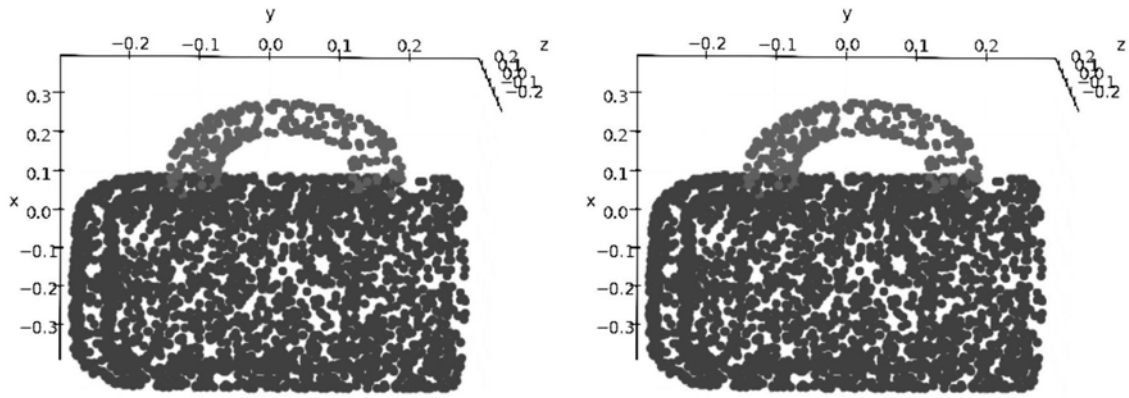


图6

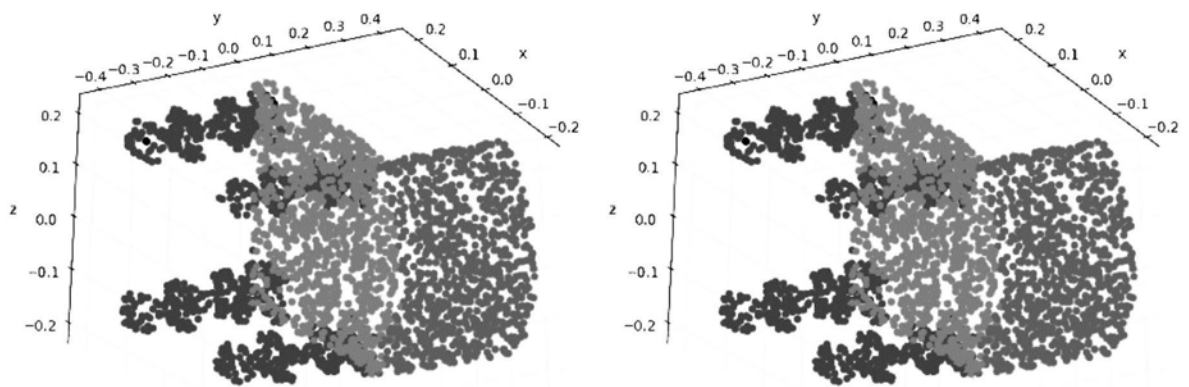


图7