



(12) 发明专利

(10) 授权公告号 CN 113988164 B

(45) 授权公告日 2023. 08. 08

(21) 申请号 202111227890.6

G06V 10/80 (2022.01)

(22) 申请日 2021.10.21

G06V 10/82 (2022.01)

(65) 同一申请的已公布的文献号

G06N 3/0464 (2023.01)

申请公布号 CN 113988164 A

G06N 3/08 (2023.01)

(43) 申请公布日 2022.01.28

(56) 对比文件

(73) 专利权人 电子科技大学

CN 112035746 A, 2020.12.04

地址 611731 四川省成都市高新区(西区)

CN 113486963 A, 2021.10.08

西源大道2006号

CN 112561796 A, 2021.03.26

(72) 发明人 朱大勇 罗光春 赵太银 陈爱国

WO 2021164469 A1, 2021.08.26

潘海涛 曹申健

CN 112818999 A, 2021.05.18

WO 2021139069 A1, 2021.07.15

(74) 专利代理机构 成都希盛知识产权代理有限公司

CN 111860666 A, 2020.10.30

51226

Guo M H等.Pct: Point cloud

专利代理师 陈泽斌

transformer.《Computational Visual Media》

.2021,第7卷187-199.

(51) Int. Cl.

审查员 乔君

G06V 10/764 (2022.01)

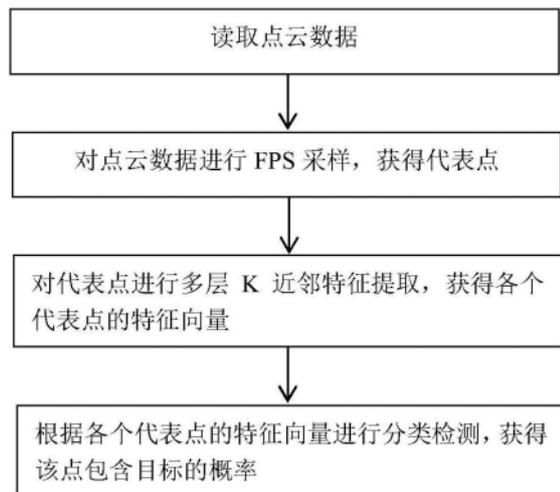
权利要求书2页 说明书5页 附图2页

(54) 发明名称

一种面向代表点自注意力机制的轻量级点云目标检测方法

(57) 摘要

本发明涉及基于点云数据的目标检测技术,其公开了一种面向代表点自注意力机制的轻量级点云目标检测方法,减少模型参数量,提高模型的训练收敛速率和检测准确率。该方法包括以下步骤:S1、读取点云数据;S2、对点云数据进行FPS采样,获得代表点;S3、对代表点进行多层K近邻特征提取,获得各个代表点的特征向量;S4、根据各个代表点的特征向量进行分类检测,获得该点包含目标的概率。



1. 一种面向代表点自注意力机制的轻量级点云目标检测方法,其特征在于,包括以下步骤:

S1、读取点云数据;

S2、对点云数据进行FPS采样,获得代表点;

S3、对代表点进行多层K近邻特征提取,获得各个代表点的特征向量;

S4、根据各个代表点的特征向量进行分类检测,获得该点包含目标的概率;

步骤S3中,在对代表点进行多层K近邻特征提取过程中,每一层K近邻特征提取包括以下步骤:

S31、依次以每个代表点为中心点,采用K最近邻算法选取一定数量的临近点,获得包含临近点特征的代表点特征向量;

S32、将包含临近点特征的代表点特征向量输入到多层感知神经网络中进行学习,获得带有通道特征的代表点特征向量;在获得带有通道特征的代表点特征向量后,对所述带有通道特征的代表点特征向量进行自注意力机制处理,获得加入自注意力掩码之后的代表点特征向量;

S33、进行最大池化处理,获得表征各代表点的特征向量,作为当前层K近邻特征提取的结果;对加入自注意力掩码之后的代表点特征向量进行最大池化处理,获得表征各代表点的特征向量;

S34、将当前层K近邻特征提取的结果与上一层K近邻特征提取的结果进行拼接,获得拼接后的向量,将所述拼接后的向量作为最终的代表点的特征向量。

2. 如权利要求1所述的一种面向代表点自注意力机制的轻量级点云目标检测方法,其特征在于,步骤S1还包括:在读取点云数据时,对读取的点云的数量进行一致性采样。

3. 如权利要求1所述的一种面向代表点自注意力机制的轻量级点云目标检测方法,其特征在于,步骤S32中,所述多层感知神经网络包括三个一维卷积核,每个卷积核的大小为 1×4 ,三个卷积核的输出通道数分别为16,32,64。

4. 如权利要求1所述的一种面向代表点自注意力机制的轻量级点云目标检测方法,其特征在于,所述对带有通道特征的代表点特征向量进行自注意力机制处理,具体包括:

S321、基于带有通道特征的代表点特征向量,对各代表点的相邻点的特征向量分别进行最大池化和平均池化,对应获得相邻点每一维度特征的最大值以及平均值;

S322、对各代表点的相邻点每一维度特征的最大值以及平均值经过一维卷积处理,得到最大池化操作和平均池化操作的特征掩码;

S323、对最大池化操作和平均池化操作的特征掩码进行叠加并采用激活函数进行激活,获得第一注意力掩码;

S323、将第一注意力掩码与代表点的相邻点的特征向量相乘,获得带有注意力权值的特征向量;

S324、对带有注意力权值的特征向量分别进行最大池化和平均池化,并对最大池化和平均池化的结果进行叠加;

S325、对叠加后的结果经过一维卷积处理,并采用激活函数进行激活,获得第二注意力掩码;

S326、将第二注意力掩码与带有注意力权值的特征向量进行相乘,获得加入自注意力

掩码之后的代表点特征向量。

一种面向代表点自注意力机制的轻量级点云目标检测方法

技术领域

[0001] 本发明涉及基于点云数据的目标检测技术,具体涉及一种面向代表点自注意力机制的轻量级点云目标检测方法。

背景技术

[0002] 三维计算机视觉技术日新月异,发展迅猛。目标检测作为计算机视觉中一个重要的分支,在社会生活和工业领域应用前景广泛,无论在学界还是工业界都是一个热门领域。由于点云数据不易受光照、尘雾等因素的影响,而传统图像可能会受困于天气等视觉干扰的因素,因此点云数据在三维场景的目标检测任务中得到了广泛的应用,通过使用点云数据获得三维空间的信息,提升了三维目标检测的效果,具有广泛的应用前景和应用价值。

[0003] 在三维目标检测的过程中,出现了众多的方法。其中部分方法是通过二维图像去得到三维目标检测结果,例如Gs3d利用二维图像潜在的三维信息去得到检测目标的三维边框,这种直接使用二维图像得到三维检测框的方法由于缺乏深度信息,因此往往难以达到令人满意的效果。Frustum pointnet方法则是先通过二维图像生成检测到的二维边框(bounding box)边框,然后再以视锥体的方式去除那些不属于二维边框内的点云,再将框内的点云送入编码器进行三维边框估计,从而得到最终的检测结果。该类方法的性能表现过于依赖于二维图像的目标检测结果。

[0004] 为了从点云数据中学习到目标检测任务所需的几何特征表达,MV3D通过对点云进行投影,得到二维平面上的数据表示,之后再利于二维卷积进行特征提取,最后利用特征表达得到三维检测结果。Voxelnet先对点云数据进行处理,将数据分为同等大小的网格区域,每个区域进行各自的特征提取,然后使用三维卷积对特征进一步的学习,最后生成三维检测结果。Shi等人提出的PointRCNN的思想将目标检测任务看作一个分割的方式进行处理,然后得到每一个点云属于前景点的概率,然后得到三维边框,再通过精细化的处理得到最终的检测结果。

[0005] 这些方法存在以下不足:(1)某些点存在重复计算的问题,同时增加的模型参数量,并且需要更长的时间和占用更多的计算资源进行训练。(2)在选取代表点的过程中可能会忽略掉代表点与周围相邻点的特征联系,存在代表点局部特征信息缺失的问题。(3)由于点云数据具有稀疏性的特点,当点云数据在前向传播的过程中会愈发稀疏,导致其特征表现能力不足。

发明内容

[0006] 本发明所要解决的技术问题是:提出一种面向代表点自注意力机制的轻量级点云目标检测方法,减少模型参数量,提高模型的训练收敛速率和检测准确率。

[0007] 本发明解决上述技术问题采用的技术方案是:

[0008] 一种面向代表点自注意力机制的轻量级点云目标检测方法,包括以下步骤:

[0009] S1、读取点云数据;

- [0010] S2、对点云数据进行FPS采样,获得代表点;
- [0011] S3、对代表点进行多层K近邻特征提取,获得各个代表点的特征向量;
- [0012] S4、根据各个代表点的特征向量进行分类检测,获得该点包含目标的概率。
- [0013] 作为进一步优化,步骤S1还包括:在读取点云数据时,对读取的点云的数量进行一致性采样。
- [0014] 作为进一步优化,步骤S3中,在对代表点进行多层K近邻特征提取过程中,每一层K近邻特征提取包括以下步骤:
- [0015] S31、依次以每个代表点为中心点,采用K最近邻算法选取一定数量的临近点,获得包含临近点特征的代表点特征向量;
- [0016] S32、将包含临近点特征的代表点特征向量输入到多层感知神经网络中进行学习,获得带有通道特征的代表点特征向量;
- [0017] S33、进行最大池化处理,获得表征各代表点的特征向量,作为当前层K近邻特征提取的结果。
- [0018] 作为进一步优化,步骤S32中,所述多层感知神经网络包括三个一维卷积核,每个卷积核的大小为 1×4 ,三个卷积核的输出通道数分别为16,32,64。
- [0019] 作为进一步优化,步骤S32还包括:在获得带有通道特征的代表点特征向量后,对所述带有通道特征的代表点特征向量进行自注意力机制处理,获得加入自注意力掩码之后的代表点特征向量;
- [0020] 步骤S33中,对加入自注意力掩码之后的代表点特征向量进行最大池化处理,获得表征各代表点的特征向量。
- [0021] 作为进一步优化,所述对带有通道特征的代表点特征向量进行自注意力机制处理,具体包括:
- [0022] S321、基于带有通道特征的代表点特征向量,对各代表点的相邻点的特征向量分别进行最大池化和平均池化,对应获得相邻点每一维度特征的最大值以及平均值;
- [0023] S322、对各代表点的相邻点每一维度特征的最大值以及平均值经过一维卷积处理,得到最大池化操作和平均池化操作的特征掩码;
- [0024] S323、对最大池化操作和平均池化操作的特征掩码进行叠加并采用激活函数进行激活,获得第一注意力掩码;
- [0025] S323、将第一注意力掩码与代表点的相邻点的特征向量相乘,获得带有注意力权值的特征向量;
- [0026] S324、对带有注意力权值的特征向量分别进行最大池化和平均池化,并对最大池化和平均池化的结果进行叠加;
- [0027] S325、对叠加后的结果经过一维卷积处理,并采用激活函数进行激活,获得第二注意力掩码;
- [0028] S326、将第二注意力掩码与带有注意力权值的特征向量进行相乘,获得加入自注意力掩码之后的代表点特征向量。
- [0029] 作为进一步优化,步骤S3中,在对代表点进行多层K近邻特征提取过程中,每一层K近邻特征提取步骤还包括:
- [0030] S34、将当前层K近邻特征提取的结果与上一层K近邻特征提取的结果进行拼接,获

得拼接后的向量,将所述拼接后的向量作为最终的代表点的特征向量。

[0031] 本发明的有益效果是:

[0032] (1) 采用适用于弱性能硬件环境下的K近邻局部特征提取的方法进行点云特征提取,降低了计算资源的浪费,能够防止同一个点在不同半径下被多次计算,同时K近邻也对模型轻量化带来了极大的优势,模型的参数量大大降低,加快了模型的收敛速度,在保证模型较好精度的情况下,实现了模型的轻量化;

[0033] (2) 引入自注意力机制对输入中重要的代表点加以关注,得到代表点与周围相邻点的相似性和差异性。当模型在上采样时,对自注意力机制加强的特征进行连接,使得结果更关注于相似性更高的代表点,从而提高代表点对模型的表达能力;

[0034] (3) 通过将前一层K近邻特征提取的输出和当前层K近邻特征提取的输出进行叠加,在特征表达能力上具有一种互补的能力,从而促进网络具有更好的表现性能,提高对目标的检测准确性。

附图说明

[0035] 图1为本发明的面向代表点自注意力机制的轻量级点云目标检测原理架构图;

[0036] 图2为本发明实施例中的面向代表点自注意力机制的轻量级点云目标检测方法流程图;

[0037] 图3为本发明实施例中的首次K近邻特征提取及特征跳跃式连接示意图。

具体实施方式

[0038] 本发明旨在提出一种面向代表点自注意力机制的轻量级点云目标检测方法,减少模型参数量,提高模型的训练收敛速率和检测准确率。其核心思想是:(1) 针对点云特征提取参数量大的问题,本发明提出了一种适用于弱性能硬件环境下的K近邻局部特征提取方法,该方法降低了计算资源的浪费,防止同一个的点在不同半径下被多次计算,同时也对模型轻量化带来了极大的优势,模型的参数量大大降低,加快了模型的收敛速度,在保证模型较好精度的情况下,实现了模型的轻量化。(2) 通过引入自注意力机制,对代表点及其周围点加以关注,对非代表点进行抑制,以解决FPS采样的随机性导致代表点局部特征损失的问题,即通过面向点云数据的自注意力机制对每一个代表点与其周围相邻点的特征的重要程度进行学习,学习出一组权值,对相邻点的特征值进行有效地增强或者抑制,提高代表点对模型表达的影响力。(3) 为了防止特征在前向传播的过程中特征信息的丢失,受到深度残差网络中跳跃连接思想的启发,设计了一种跳跃连接的方法。该方法将前一个K近邻特征提取模块的输出和当前K近邻特征提取模块的输出进行叠加,使得上一层的特征和下一层的特征叠加,在特征表达能力上具有一种互补的能力,从而促进网络具有更好的表现性能。

[0039] 基于上述核心思想,本发明设计的面向代表点自注意力机制的轻量级点云目标检测原理架构如图1所示,其采用多个级联且跳跃式连接的K近邻局部特征提取模块作为特征提取网络,进行点云特征提取,获得代表点特征向量表达,然后再经过分类检测得到点云数据中的前景点,并生成包含前景点的三维边框。

[0040] 实施例:

[0041] 本实施例中的面向代表点自注意力机制的轻量级点云目标检测方法流程如图2所

示,其包括以下步骤:

[0042] S1、读取点云数据;

[0043] 本步骤中,单个场景下的点云数据是由一个无序的点云集合组成, $\{x_1, x_2, \dots, x_n\}$, 每一个点的信息是由三维坐标和反射率组成,由于每一个场景包含的点云数量是不同的,为了保证多个场景输入数据量的一致性,因此,在读取点云数据时需要点对点云数量做一致性采样。

[0044] S2、对点云数据进行FPS采样,获得代表点;本步骤中,对读取的点云数据使用最远距离采样(Farthest Point Sampling,FPS),得到8192(实验所得最佳取值)个的点云,作为当前采样层的代表点,得到 $8192 \times (3+1)$ 维特征向量。

[0045] S3、对代表点进行多层K近邻特征提取,获得各个代表点的特征向量;

[0046] 考虑到目标检测效果和目标检测速度的平衡,本实施例中采用如图1所示的4个级联并且跳跃式连接的K近邻特征提取模块进行特征提取,从而形成4层K近邻特征提取网络,每个模块的采样数依次为8192、4096、1024、256。

[0047] 本步骤中各个K近邻特征提取模块的具体提取过程包括:

[0048] S31、依次以每个代表点为中心点,使用K最邻近法(K Near Neighbor,KNN),将上一阶段选取的每一个代表点作为一组,以临近点到中心点的距离作为选取的标准,选择最近点,每组临近点由32个点组成,最终得到 $N_i \times K \times (3+c)$ 的特征向量(包含临近点特征的代表点特征向量),其中, N_i 表示第i层的代表点数量,K为选取的相邻点个数,c为三维坐标信息以外的特征维度数,作为下一步的输入;在如图3所示的首次K近邻特征提取中,最终得到 $8192 \times 32 \times (3+1)$ 的特征向量。

[0049] S32、将包含临近点的代表点特征向量输入到一个共享的多层感知神经网络中进行学习,多层感知机是由三个一维卷积核组成,每个卷积核的大小为 1×4 ,单个卷积核的输出通道数分别为16,32,64,最终得到一组特征通道数为64的特征图,得到 $N_i \times K \times d$ 的特征向量(带有通道特征的代表点特征向量),其中, N_i 表示第i层的代表点数量,K为选取的相邻点个数,d为经过K近邻特征提取后得到特征维度;在如图3所示的首次K近邻特征提取中,最终得到 $8192 \times 32 \times 64$ 的特征向量,每个代表点的临近点的特征包含64维特征,每个代表点含有32个临近点。

[0050] S33、进行最大池化处理,获得 $N_i \times d$ 维的表征各代表点的特征向量,作为当前层K近邻特征提取的结果。

[0051] 在输入点每经过一次K近邻特征提取之后,每一个代表点都会得到一组特征向量,随着网络层数的加深每一个代表点的特征维度也会变大,为了让网络对特征中较为重要的通道加以关注,本实施例引入自注意力机制,对代表点与其周围相邻点的特征进行相关度的学习,同时对特征相似性不高的特征进行削弱,使得网络将注意力放置到特征相似度更高的通道特征中。

[0052] 具体而言,在步骤S32中,在获得带有通道特征的代表点特征向量后,对所述带有通道特征的代表点特征向量进行自注意力机制处理,获得加入自注意力掩码之后的代表点特征向量;实施步骤如下:

[0053] (1)首先,输入由上一步得到的带有通道特征的代表点特征向量 C_i ,其维度为 $N_i \times K \times d$,其中, N_i 表示第i层的代表点数量,K为选取的相邻点个数,d为经过K近邻特征提取后得

到特征维度；

[0054] 在本实施例中，在首次K近邻特征提取中，输入的代表点特征向量的维度为 $8192 \times 32 \times 64$ ，8192表示当前层代表点数量，32为选取的相邻点个数，64为经过K近邻特征提取后得到特征维度。

[0055] (2) 对每个代表点的相邻点的特征向量进行最大池化，得到相邻点中每一维的最大特征值，经过最大池化后的特征维度为 $N_1 \times 1 \times d$ ，在本实施例中，在首次K近邻特征提取中，经过最大池化后获得的特征维度为 $8192 \times 1 \times 64$ 。

[0056] (3) 对每个代表点的相邻点的特征向量进行平均池化，得到相邻点中每一维特征的平均值，经过平均池化后的特征维度为 $N_1 \times 1 \times d$ ，在本实施例中，在首次K近邻特征提取中，经过平均池化后获得的特征维度为 $8192 \times 1 \times 64$ 。

[0057] (4) 再将上述步骤(2)、(3)获得的两种特征向量放入到一维卷积核中，该卷积核的输出通道为64维，得到最大池化和平均池化的特征掩码。

[0058] (5) 再将两种特征掩码对应位置的值相加，对特征掩码进行互补，得到待激活状态的注意力掩码。

[0059] (6) 待激活注意力掩码经过一次激活函数进行激活，其中激活函数选用ReLU，得到第一注意力掩码。

[0060] (7) 将第一注意力掩码与输入特征向量 C_1 相乘得到带有注意力权值的特征向量 C_2 ，该特征向量的维度保持不变，即为 $N_1 \times K \times d$ 。

[0061] (8) 对带有注意力权值的特征向量 C_2 分别进行最大池化和平均池化，得到的最大池化和平均池化的结果的维度均为 $N_1 \times 1 \times d$ ，并对最大池化和平均池化的结果进行叠加，得到特征维度为 $N_1 \times 2 \times d$ ，在本实施例中，在首次K近邻特征提取中，经过叠加后获得的特征维度为 $8192 \times 2 \times 64$ 。

[0062] (9) 将第(8)步得到的特征向量放入到一维卷积核中，卷积核尺寸为 1×2 ，该卷积核的输出通道为64维，得到待激活注意力掩码；

[0063] (10) 对待激活注意力掩码经过一次激活函数进行激活，其中激活函数选用ReLU，得到第二注意力掩码。

[0064] (11) 将第二注意力掩码与带有注意力权值的特征向量 C_2 相乘得到加入自注意力掩码的代表点特征向量 C_3 ，该特征向量的维度保持不变，即为 $N_1 \times K \times d$ 。

[0065] 此外，在步骤S3中，为了让网络有更好的性能，我们使用多级特征跳跃连接对已损失的代表点特征向量进行补充，特征跳跃连接(Skip Connection, SC)通过跳跃连接的方式将采样得到的代表点特征与当前层的输出特征形成融合拼接，对当前层的特征向量进行补充，得到维度更大的特征，作为最终的代表点的特征向量，其维度为 $N_1 \times (d+3+c)$ ，如图3所示。

[0066] S4、根据各个代表点的特征向量进行分类检测，获得该点包含目标的概率。

[0067] 本步骤中，将最终的每个代表点的特征向量输入到一个全连接层中进行检测，该全连接层的输出是一个二分类的概率，结果得到该点含有目标的概率值。

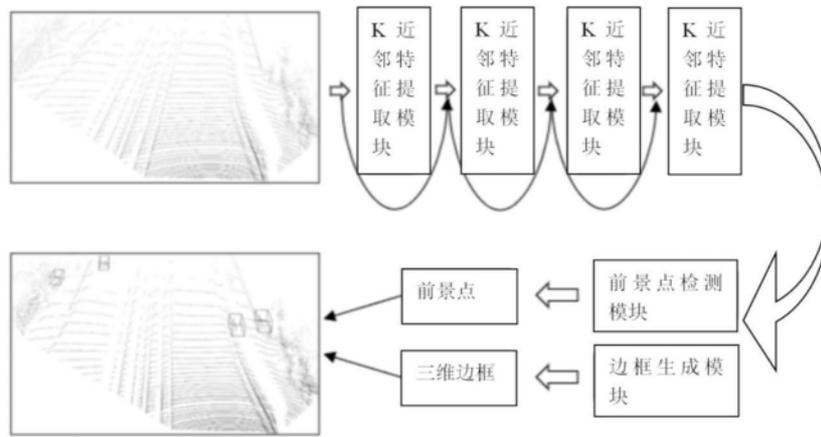


图1

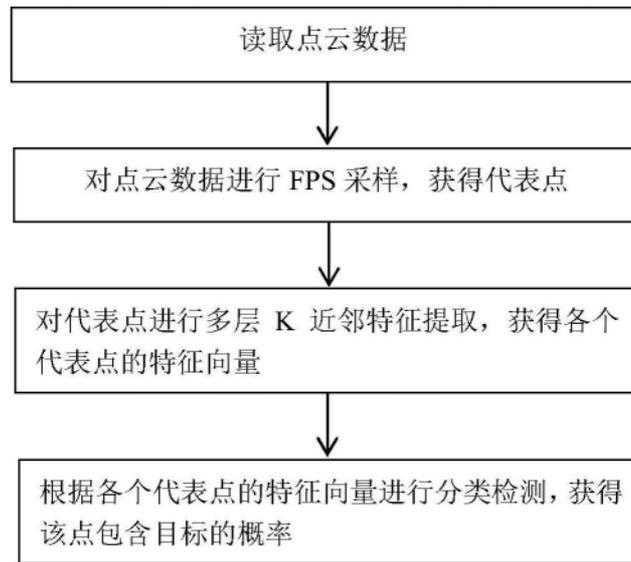


图2

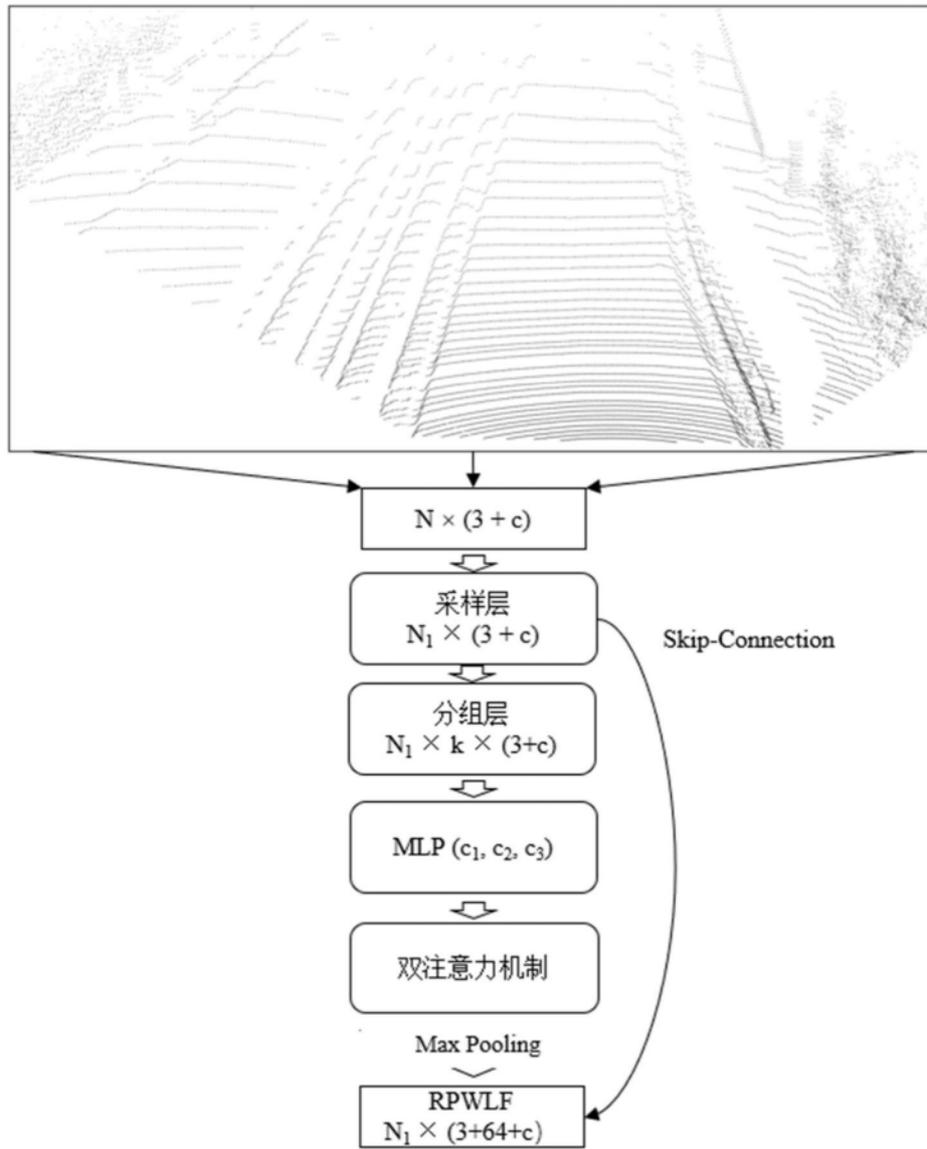


图3