

基于双边结构多尺度动态图的牙齿点云分割

王 桢, 陈 胜

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年4月20日; 录用日期: 2024年5月15日; 发布日期: 2024年5月22日

摘 要

牙齿隐形矫治装置设计的关键就是将牙齿区域从口腔内三维点云模型中实现精准分割, 并在尽量降低人为干预的需求下, 对牙齿区域中的单颗牙齿实现全自动分割。传统分割技术需要依赖于专家的先验知识以及大量的人机交互辅助操作, 分割性能易受到牙齿形状和位置变化的影响, 无法实现全自动分割。因此本文提出了一种Multi-Dynamic Graph of Bilateral Structurest (M-DGB)模型。模型使用双边结构将点云特征信息分别输入坐标边与法向量边, 首先利用特征转换模块获得不同尺度的初始全局特征, 之后利用多尺度动态图模块中的K近邻图与改进的动态图卷积EdgeConv++模仿卷积神经网络渐近地增加感受野的方式, 对局部几何特征实现多层次多尺度学习, 进一步提取增强后的局部特征。最后将先前得到的局部增强特征与全局特征密集融合, 以获取更具有表达能力多属性特征。此外, 改进了一种混合损失函数, 加强牙齿与牙龈的边界分割。将该模型在自制数据集上进行实验, 与现有点云分割模型PointNet、PointNet++、MeshSegNet相比, 分割精度提高, 平均Dice系数为0.972、PPV为0.964、SEN为0.987。

关键词

点云分割, 三维牙齿模型, 双边结构, 动态图卷积, 混合损失函数

Tooth Point Cloud Segmentation Based on Multi-Dynamic Graph of Bilateral Structurest

Zhen Wang, Sheng Chen

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 20th, 2024; accepted: May 15th, 2024; published: May 22nd, 2024

Abstract

The key to the design of the invisible orthodontic device is to accurately segment the tooth area from the 3D point cloud model in the oral cavity, and to achieve automatic segmentation of a single tooth in the tooth area with minimal human intervention. Traditional segmentation technology relies on the prior knowledge of experts and a large number of human-computer interaction auxiliary operations, and the segmentation performance is easily affected by changes in tooth shape and position, and fully automatic segmentation cannot be realized. Therefore, a Multi-Dynamic Graph of Bilateral Structurest (M-DGB) model is proposed. The model uses a bilateral structure to input the point cloud feature information into the coordinate edge and normal vector edge respectively, firstly uses the feature transformation module to obtain the initial global features of different scales, and then uses the K-nearest neighbor graph in the multi-scale dynamic graph module and the improved dynamic graph convolutional EdgeConv++ to imitate the way of convolutional neural network asymptotic increase of the receptive field, realizes hierarchical and multi-scale learning of local geometric features, and further extracts the enhanced local features. Finally, the previously obtained local enhancement features are intensively fused with the global features to obtain multi-attribute features with more expressive ability. In addition, a mixed loss function has been improved to enhance the boundary segmentation of teeth and gums. Compared with the existing point cloud segmentation models PointNet, PointNet++ and MeshSegNet, the segmentation accuracy is improved, and the average Dice coefficient is 0.972, PPV is 0.964, and SEN is 0.987.

Keywords

Point Cloud Segmentation, 3D Tooth Model, Bilateral Structurest, Dynamic Graph Convolution, Mixed Loss Function

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

隐形矫治装置的设计使用口内扫描仪(Intraoral Scanners, IOSs)对患者的口腔内牙齿与牙龈等部分软组织进行实时扫描, 并将扫描得到的口腔内三维点云实时重建, 利用计算机数据处理系统对最终生成的数字化口腔内三维模型, 进行计算机辅助牙齿正畸方案的模拟与设计。而其中的一个关键处理过程就是将牙齿区域从数字化口腔内三维模型中实现精准分割, 并尽量降低人为干预的需求, 通过对分割出的牙齿区域进行单颗牙齿点云的分割。

由于患者的牙齿形状和分布位置因人而异, 想要将牙齿从口腔内三维模型中分割出来并非易事。尤其当患者牙齿存在外观异常, 例如发生严重的牙齿错位和缺失现象时, 实现牙齿的分割任务就更加困难。牙齿错位现象通常表现为邻近牙错位、拥挤和间隙不规则, 错位现象会导致牙齿的空间距离不规则, 甚至出现空间位置上的重叠[1] [2]。而牙齿缺失现象不仅会在分割阶段造成不同种类牙齿的错误分割, 在后续对指定单颗牙齿的提取环节也会造成干扰。

传统口腔内三维模型分割方案直接利用口腔内三维模型的三维坐标、法向量及曲率等基本空间特征实现牙齿分割任务。Kumar [3]通过曲率分析和复杂路径搜索实现牙齿与牙龈区域分割的方法, Yuan [4]

根据两颗相邻牙齿之间融合区域的“谷”状特征,对模型表面的最小曲率进行分析,实现单颗牙齿的分割。Zhao [5]根据三角形网格的曲率值,将要素点连接到要素区域。从区域中提取特征线后,借助人工提供的交互信息获得特征等值线实现分割。Ma [6]在人工标注牙齿轮廓线时,增加了曲率估计和矢量优化的交互式方法辅助标注,使单颗牙齿从模型中分割的过程更为简化。Zou [7]提出了一种需要人工少量标注作为先验信息,利用谐波场在牙齿表面模型上实现牙齿分割的方法。

深度学习口腔内三维模型分割方案通过神经网络从口腔内三维模型的基本空间特征中提取更深的高级空间特征实现牙齿分割任务。Xu [8]通过标记每个网格面,提取三维模型表面图像作为二维卷积神经网络(Convolutional Neural Network, CNN)的输入实现分割。Tian [9]提出一种基于稀疏体系八叉树的口腔内三维模型分割方案。Lian [10]在 PointNet [11]的基础上进行改进,提出分割模型 MeshSegNet 直接使用原始网格数据(mesh)作为输入实现牙齿的分割。Cui [12]在 PointNet++ [13]基础上,提出了一种口腔内三维模型两阶段分割网络。

本文提出 M-DGB 模型使用网格形式的点云作为初始输入,相比原始点云拥有更多的空间信息与特征信息,针对牙齿点云中每个待分割网格,进一步利用其局部特征进行了建模,但考虑到每个待分割网格只占据了整个牙齿表面的非常小的一部分,仅单独使用顶点坐标、单元位置坐标、法向量作为输入或混合作为输入进行特征提取实现局部结构建模,容易造成多种空间特征的混淆,使网络难以提取易于识别的特征,进一步忽略空间特征之间的互相影响,不能充分体现局部区域待分割网格的具体分布状态以及网格之间的特征相关性。

本文提出 M-DGB 模型所使用的双边结构是语义分割中常用的结构之一,双边结构中的边在实际应用中等同于轻量化的 backbone 结构,是目前图像处理领域较为热门的研究趋势。双边结构在实现语义分割时主要通过共享浅层网络,从而减少繁冗的特征提取。其次通过高效的多尺度模块,在局部特征提取环节获得更高效的语义信息。在双边结构的核心在于轻量化提升,能够在提高精度的同时确保网络尽量少的消耗时间。针对可能存在分辨率不同引起的特征不对齐问题,仅简单地相加或拼接操作忽略了不同特征之间的差异性,进而会造成分割性能的降低,特征融合模块针对该问题进行优化,如 BiSeNet [14]模型中在特征融合模块增强了特征融合信息解决了特征不对齐的问题。双边结构在 BiSeNet、DDRNet [15]的应用中获得了较为理想的效果,体现了双边结构的实施的有效性与合理性。

因此,为了解决坐标信息与法向量信息出现特征混淆的情况,本文使用双边结构对输入网格的坐标信息与法向量信息分别建模,之后进一步使用 K 近邻图与改进的动态图卷积 EdgeConv++ [16]模仿卷积神经网络渐近地增加感受野的方式,对上下文特征实现多层次多尺度学习,进一步提取增强后的局部特征。最后将先前得到的局部增强特征与全局特征密集融合,以得到更具有表达能力的多属性特征。此外,改进了一种混合损失函数,加强注边界分割。

2. 数据预处理

本文所使用初始输入包含口腔内三维模型中所有网格的初始坐标信息和法向量信息两部分,定义口腔内三维模型中所有网格的集合 $M = \{m_i | i = 1, 2, \dots, N\}$, 其中 N 为网络输入的原始牙齿表面的数量。三角形网格 $m_i \in M$ 所具有的初始特征信息,包含三个顶点坐标组成的初始特征向量 $P_{\text{mesh}} \in R^9$ 。为了丰富网格的初始特征信息,充分利用网格局部拓扑连接的特性,需要对每个网格 m_i 预处理,从而进一步提取更深层次的空间特征信息。

根据三角形网格的法向量特征 $N_{\text{mesh}} \in R^3$, 与其对应的三个顶点的法向量特征 $N_{\text{vertex}} \in R^9$ 。计算得到三角形网格 m_i 的质心坐标值,使其沿每个轴根据三维模型表面上所有网格的最大和最小坐标进行归一化,定义归一化后得到的质心坐标集合为 $P_{\text{centroid}} = \{P_{i_{\text{centroid}}} | i = 1, 2, \dots, N\}$, 归一化后质心 $P_{i_{\text{centroid}}} \in R^3$ 坐标计算

见公式。

$$P_{i\text{centroid}} = \left(\frac{x_i^1 + x_i^2 + x_i^3}{3} - x_{\min}, \frac{y_i^1 + y_i^2 + y_i^3}{3} - y_{\min}, \frac{z_i^1 + z_i^2 + z_i^3}{3} - z_{\min} \right) \quad (1)$$

其中, x_i, y_i, z_i 分别为网格 m_i 的顶点坐标, x_{\max}, x_{\min} 为 x 轴上所有网格的最大最小坐标。

在完成 N 个网格的数据预处理操作后, 沿每个维度进行标准分数归一化。

$$(x', y', z') = \left(\frac{x - \mu_x}{\sigma_x}, \frac{y - \mu_y}{\sigma_y}, \frac{z - \mu_z}{\sigma_z} \right) \quad (2)$$

其中 x, y, z 变换前的坐标, μ_x 为 x 轴方向总体数据的均值, σ_x 为 x 轴方向总体数据的标准差。

将法向量 N_{mesh} 与顶点法向量 N_{vertex} 拼接成法向量特征矩阵 $F_n = (f_{n1}, f_{n2}, \dots, f_{nN})$, 将初始特征向量 P_{mesh} 与 P_{centroid} 拼接成坐标向量特征矩阵 $F_m = (f_{m1}, f_{m2}, \dots, f_{mN})$ 。共同作为网络的初始输入特征矩阵 F , 其中 $F_m \in R^{N \times 12}$, $F_n \in R^{N \times 12}$, $F \in R^{N \times 24}$ 。

3. M-DGB 模型

3.1. M-DGB 模型结构

M-DGB 网络结构如图 1 所示, 将输入原始数据转换为 $F \in R^{N \times 24}$ 的矩阵, 其中 N 为点云中所含网格数量, 建立从原始数据矩阵 F 到分割预测矩阵 $P \in R^{N \times (c+1)}$ 的非线性映射, 预测矩阵 P 中的每一行表示相对应网格属于不同类别(即 LR1-LR7、LL1-LL7 和 GUM, C 的取值为 14 对应 14 种牙齿类别)的概率。

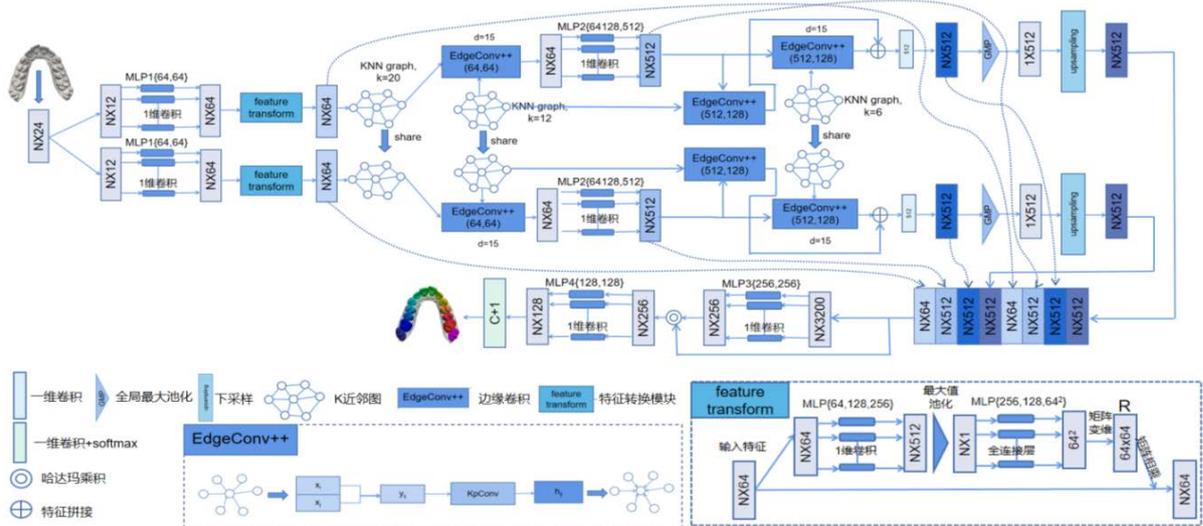


Figure 1. M-DGB structure

图 1. M-DGB 结构

3.2. 特征转换模块

特征转换模块(Feature Transform Module)是由特殊的多层感知机(Multi-Layer Perceptron, MLP)组成, 它预测来自经 MLP1 得到的 $F^0 \in R^{N \times 64}$ 的变换矩阵 $T \in R^{64 \times 64}$, 将所有网格的初始特征 F_0 更新为:

$$\hat{F}_0 = F_0 T \quad (3)$$

如图 2 所示, 特征转换模块对前一部分 MLP1 提取出的 64 维特征映射到一个规范的特征空间中进行对齐, 即在特征层面对输入的网格的特征向量进行变换, 获得的初始特征值也是不同尺度的。使得网格在不同的方向、比例和位置等情况下, 都能被准确识别和分割, 以提高学习到的特征表示关于输入曲面的潜在的几何变换的稳健性。

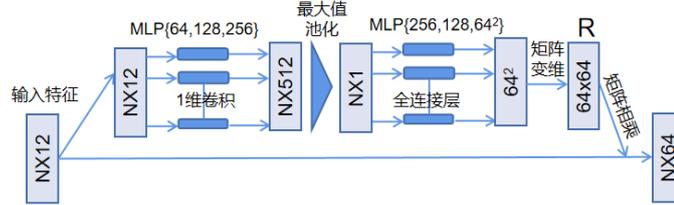


Figure 2. Feature transform module
图 2. 特征转换模块

3.3. 多尺度动态图模块

多尺度动态图模块(Multi-Dynamic Graph Module)由三个不同尺度的动态图子模块组成, 沿着网络的前向路径将输入牙齿点云的局部特征在不同阶段进行整合, 模仿卷积神经网络渐近地增加感受野的方式 [17], 实现多层次多尺度特征学习过程, 显式地捕捉输入牙齿点云的局部几何特征。

在构建 K 近邻图时分别将三个动态图子模块的 K 值选取为 6、12、20, 定义使用 K 近邻算法所构建的图结构 G 为:

$$G = (V, E) \quad (4)$$

其中 $V = \{m_i | i = 1, 2, \dots, K\}$ 表示 K 近邻图的中心节点集合, 其中的每个节点都代表着一个网格。

$E = \{e_{ij} = (e_{i1}, e_{i2}, \dots, e_{ij}) | j = 1, 2, \dots, K\}$ 表示节点之间的边缘特征向量的集合。

由于法向量信息与牙齿点云中的相邻空间结构无关, 在法向量边中不针对法向量构建 K 近邻图, 而是直接共享坐标边构建的 K 近邻图, 从而避免法向量边构建图出现中心节点与邻居节点随机排列在空间中, 造成局部特征建模的混淆。

边缘卷积 EdgeConv 所定义的边缘特征 e_{ij} , 不仅对中心点 x_i 的全局信息进行提取, 也对中心点与邻域点的局部信息实现了提取, 如公式(1)所示。

$$e_{ij} = g_{\Theta}(x_i, x_j) = \bar{g}_{\Theta}(x_i, x_j - x_i) \quad (5)$$

式中: $g_{\Theta}: R^F \times R^F \rightarrow R^F$ 是基于非对称的边的特征计算函数, 具有可学习参数 Θ 。

如图 3 所示, 改进得到的 EdgeConv++, 计算 x_i 与邻居节点 x_{ij} 的相对距离 y_{ij} , 通过核点卷积 [18] (Kernel Point Convolution, KPConv) 后得到增强后的局部特征 k_{ij} , 然后将全局特征与局部特征拼接形成新的边缘特征 n_{ij} 。

$$n_{ij} = g_{\Theta}(x_i, k_{ij}) = g_{\Theta}(x_i, KPConv(y_{ij})) = \bar{g}_{\Theta}(x_i, KPConv(x_{ij} - x_i)) \quad (6)$$

对于任意的 y_{ij} , 其核函数 $KPConv(y_{ij})$ 为:

$$KPConv(y_{ij}) = \sum_{k < K'} h(y_{ij}, \tilde{x}_k) w_k = \sum_{k < K'} \left[\max \left(0, 1 - \frac{\|y_{ij} - \tilde{x}_k\|}{\beta} \right) \right] w_k \quad (7)$$

其中 K' 为核心节点的总数, 核函数 $KPConv(y_{ij})$ 由权重矩阵 w_k 及其对应的权重系数 $h(y_{ij}, \tilde{x}_k)$ 加权求和得到。 $h(y_{ij}, \tilde{x}_k)$ 权重系数由核心节点 \tilde{x}_k 和 y_{ij} 之间相对距离计算得到, 相对距离与权重系数成反比, 权重系数的范围为 $[0, 1]$ 。 β 是核心节点的影响距离, 超过这个距离的节点不会进行相乘, 从而实现任意核心节点之间有重叠的影响区域。

$KPConv(y_{ij})$ 相当于是面向 x_{ij} 计算得到的权重矩阵, 利用它对 y_{ij} 进行特征变换获得增强后的局部特征 k_{ij} 。 k_{ij} 既包含了中心节点与邻域节点的局部几何信息, 也包含了邻域节点与核心节点之间的位置信息。

$$k_{ij} = KPConv(y_{ij})y_{ij} \tag{8}$$

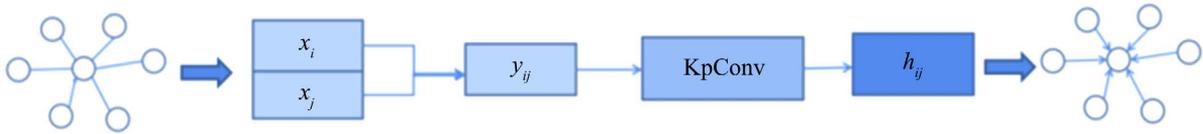


Figure 3. Dynamic graph module
图 3. 动态图模块

3.4. 密集融合模块

密集融合模块用于将局部特征密集到全局特征, 以获得更全面的特征描述。全局最大池化(Global Max Pooling, GMP)应用于第三个动态图模块后产生平移翻转不变的全局特征, 实现对整个牙齿的全局特征进行建模, 然后, 使用密集融合策略来密集将来自特征转换模块, 第一个动态图模块, 第二个动态图模块的局部特征连接到全局特征, 然后通过 MLP3 模块与 MLP4 模块生成维度为 $N \times 128$ 的特征矩阵。最后使用具有 softmax 激活的 1D Conv 层来预测 $N \times (C + 1)$ 概率矩阵, 其中 $C = 14$ (即 14 类牙齿)。

3.5. 混合损失函数

医学分割任务大多采用 Generalized Dice Loss 优化模型参数, 然而它更多的关注全局信息, 与牙齿点云的所有网格有关, 但对于单颗牙齿区域占比较小, 牙龈区域占比较多的情况, 在预测时会导致模型训练波动较大, 其缺少对边界的检测, 边缘信息获取不充分。Lovasz Softmax Loss [19]用于多类别图像分割, 它是基于 lovasz 扩展的概念和 softmax 函数的组合而来的, 可以解决传统的交叉熵损失函数在处理不平衡数据时的问题, 能够更好地处理类别不平衡和边界模糊的情况。为了使模型训练重点关注边界信息, 提出改进的加权复合损失函数:

$$L_{Loss} = \alpha L_{Dice} + \beta L_{Lovasz} \tag{9}$$

其中, α 、 β 是两个函数的权重系数, 通过训练测试, 令 $\alpha = 0.45$, $\alpha + \beta = 1$ 。

其中, Generalized Dice Loss 的公式如下:

$$L_{Dice} = 1 - \frac{2|P \cap T| + smooth}{|P| + |T| + smooth} = 1 - \frac{2 \sum_i^N p_i g_i + smooth}{\sum_i^N p_i^2 + \sum_i^N g_i^2 + smooth} \tag{10}$$

T 分别代表模型预测结果与真实标记结果; smooth 平滑因子取值为 0.00001, 防止分母预测为 0; N 为特征图中像素点的个数; p_i 是模型预测结果图中第 i 个网格的值, 范围在 $[0, 1]$, g_i 是真实标记结果中第 i 个网格的值, 取值为 0 或 1。

Lovasz Softmax Loss 的具体实现原理如下:

$$\bar{\Delta}: m \in R^p \mapsto \sum_{i=1}^p m_i g_i(m) \quad (11)$$

$$g_i(m) = \Delta(\{\pi_1, \dots, \pi_i\}) - \Delta(\{\pi_1, \dots, \pi_{i-1}\}) \quad (12)$$

函数 $\bar{\Delta}$ 的定义域为连续的 RP , P 为像素数量, m 为误差项, π 是误差项按降序排列的下标号, 例如 $m_{\pi_1} \geq m_{\pi_2} \dots \geq m_{\pi_p}$, 函数 $\bar{\Delta}$ 是连续的, 进行求导后的导数为 $g(m)$ 。

4. 实验

4.1. 数据集

面向口腔内三维模型的分割任务, 暂无可供使用的开源数据集, 为方便评估本研究设计的口腔内三维模型分割方案的性能, 制作了一个精标注的口腔内三维模型数据集。考虑所使用数据量对模型分割效率的影响以及标注所需时间成本, 选取 50 名不同的原始下颌牙齿表面, 由高精度口内扫描仪直接采集和重建生成口腔内三维模型, 每个模型的初始表面具有的三维网格数目大约在 300,000 个左右。为了进一步降低冗余的数据量对模型分割效率的影响, 同时保持模型网格数目的统一性, 在维持口腔内三维模型真实拓扑结构的前提下, 对每个口腔内三维模型的网格数目通过下采样统一为 15,000 个网格(即 $N = 15,000$)。

在下采样网格表面, 采用国际牙科联合会(International Dental Federation, FDI)发布的牙齿命名规则如表 1 所示, 将三维牙齿模型的分割类别定义为 15 种($C = 14 + 1$), 牙龈标记为 GUM, 如图 4 所示, 分割类别具体编号为:

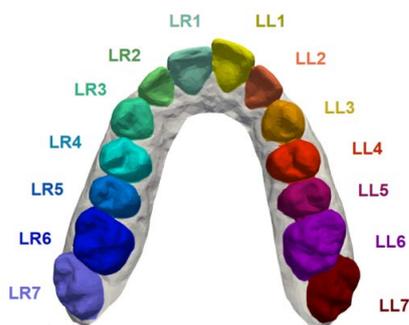


Figure 4. Tooth labeled number

图 4. 牙齿标注编号

Table 1. Tooth labeled number table

表 1. 牙齿标注编号表

牙齿类别	左编号	右编号
前牙	LL1	LR1
侧切牙	LL2	LR2
颌尖牙	LL3	LR3
第一双尖牙	LL4	LR4
第二双尖牙	LL5	LR5
第一白齿	LL6	LR6
第二白齿	LL7	LR7

4.2. 实验环境

网络在深度学习框架 Pytorch-Lightning 上实施, 实验所使用硬件设备的 GPU 型号为 NVIDIA GTX3090, 采用 Generalized Dice Loss、lovasz softmax loss 加权结合和 ADAM 优化器的 AMSGrad 变体训练网络。batch_size 设置为 4, 训练次数设置为 200 epoch, 每次迭代步长为 20, 使用 StepLR 作为更新学习率的优化器, 初始学习率设置为 0.001, 网络训练每重复 20 次, 学习率减少至原有的一半, 学习率最低降至 0.00001。

4.3. 评价指标

对于分割精度的评价指标主要采用 Dice 相似系数(the Dice similarity coefficient, DSC), 真阳性预测率(the positive prediction value, PPV), 敏感度(the sensitivity, SEN)来判定。

Dice 相似系数是一种集合相似度度量指标, 通常用于度量两个样本的相似度, Dice 相似系数值的范围为[0, 1], 值越大表示越相似。

$$\text{Dice}(P, T) = \frac{2 \times |P \cap T|}{|P| + |T|} \Leftrightarrow \text{Dice} = \frac{2TP}{FP + 2TP + FN}$$

敏感度表示所有牙齿区域的网格, 被分割为牙齿区域网格的概率。

$$\text{SEN} = \frac{TP}{TP + FN}$$

真阳性预测率表示所有牙齿区域的网格, 被正确分割为牙齿区域网格的概率。

$$\text{PPV} = \frac{TP}{TP + FP}$$

TP 表示正确分割出的牙齿区域的网格数、 TN 表示正确被分割至牙龈区域的网格数、 FP 表示牙齿区域被分割为牙龈区域的网格数、 FN 表示牙龈区域被分割为牙齿区域的网格数。

5. 实验结果与分析

5.1. 对比实验

在相同的实验环境下, 将 M-DGB 的分割性能与 PointNet, PointNet++以及 MeshSegNet 进行了对比。PointNet 是第一个可以直接将点云数据作为输入的端到端网络, 有效地解决了点云数据存在的无序性, 平移不变性以及点与点之间内部空间关系等问题。PointNet++通过对三维空间数据进行局部区域构建的方式来提取局部特征信息, 并在一定程度上提升了网络的学习能力, 但由于其仅简单地使用最大值池化进行局部特征聚合, 依旧会丢失了其他细节信息。MeshSegNet 采用邻居矩阵将 PointNet 的特征提取过程进行改进, 通过对局部区域内的网格分配权重 0 和 1, 来体现局部区域的网格的连接关系。

三重交叉实验下的平均 Dice 相似系数, 真阳性预测率, 敏感度分割准确率指标对比如表 2 所示。从图 5 所示分割结果与表 2 所示评价指标可知, 由于 PointNet 直接将牙齿点云全部输入网络, 仅实现了对全局特征的提取, 而缺少对局部特征的建模, 从而在 DSC、PPV 与 SEN 三项指标上均显著低于其他模型, PointNet 对牙齿与牙龈区域边界的分割与相邻牙齿的分割并不理想, 因此可得出局部特征对牙齿点云分割精度具有较大影响。PointNet++与 PointNet 相比在三项分割指标有一定提升, 然而由于其对牙齿点云局部特征的提取与建模方式较为简单, 在分割精度的提升上有限制。MeshSegNet 相比前面两种模型拥有较理想的分割表现, 但它仅仅通过对局部区域内的网格分配权重 0 和 1, 来体现局部区域的网格的连接关系, 而没有对牙齿点云局部区域的具体结构进行利用, 因此在相邻牙齿区域的分割还存在不足。如图

5 所示 M-DGB 在牙齿与牙龈区域边界以及相邻牙齿区域分割表现突出, 在三项评价指标上都显著高于其他三种分割模型, 从而进一步证明了 M-DGB 所提出的双边结构与多尺度动态图卷积在牙齿点云分割的有效实施。

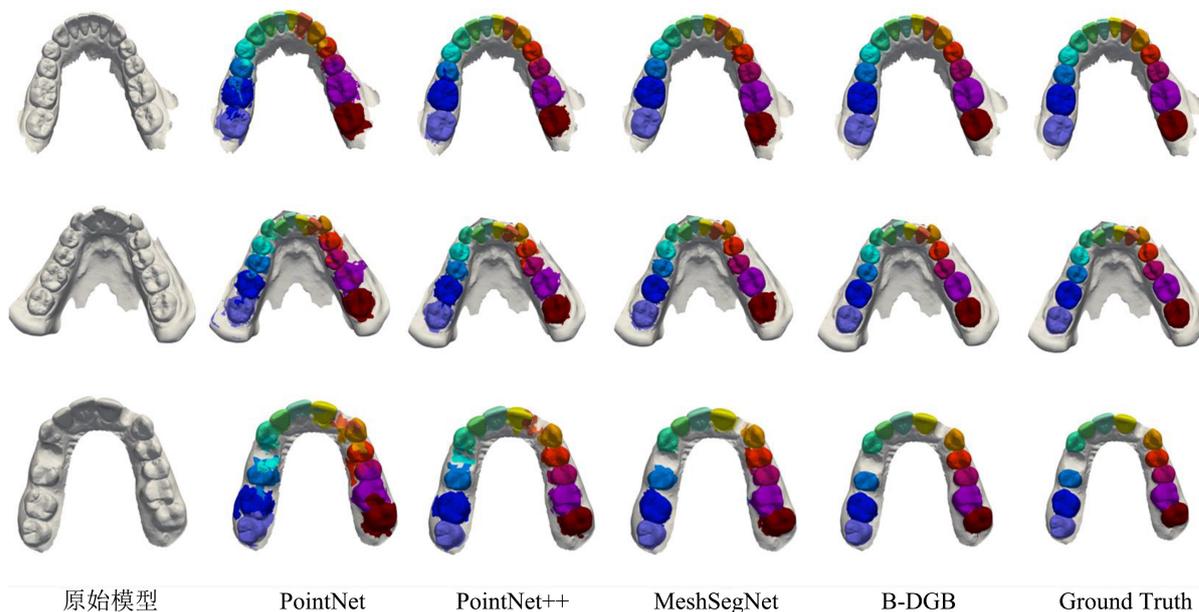


Figure 5. Comparison experiment segmentation result

图 5. 对比实验分割结果

Table 2. Comparison experiment evaluating indicators

表 2. 对比实验评价指标

模型	DSC	PPV	SEN
PointNet	0.815	0.751	0.863
PointNet++	0.873	0.826	0.895
MeshSegNet	0.947	0.908	0.929
B-DGB	0.964	0.972	0.987

5.2. 消融实验

5.2.1. 双边结构有效性

为进一步验证本文所提出的双边结构的有效性, 选择了两种网络结构进行对比: 1) 坐标边, 仅将坐标信息作为输入; 2) 法向量边, 仅将法向量信息作为输入。三种网络结构所得到的分割指标如表 3 所示, 在仅有法向量边时, 由于法向量信息并不能体现牙齿点云网格的空间位置分布, 无法提取到点云网格的结构信息, 该网络在三项分割评价指标中的表现都极差。仅有坐标边时, 坐标信息具有牙齿点云网格的空间位置信息, 该网络能够应对大部分点云网格分割任务, 但由于其没有法向量信息所带来的补充信息, 分割指标相比双边结构仍旧较低。双边结构拥有较为理想的分割性能, 在各项评价指标中表现突出, 因而可以验证双边结构能够更有效地利用坐标边与法向量边提供的空间位置信息。

Table 3. Evaluation indicators for three network structures ablation experiments
表 3. 三种网络结构消融实验评价指标

网络结构	DSC	PPV	SEN
坐标边	0.892	0.753	0.836
法向量边	0.418	0.274	0.305
双边	0.964	0.972	0.987

5.2.2. 多尺度动态图卷积模块 K 值选取

为进一步验证多尺度动态图卷积模块中 K 值的选取对牙齿点云网络分割的影响, 在保证其他条件不改变, 通过更改 K 值的方式来进行消融实验, 如图 6 所示, 当 K 值逐渐增加, 中心节点对应的局部区域随之增加, 其中所拥有的局部特征信息更加充分, 所以对应的分割指标显著提高。然而当 K 值增大到一定程度, 如 K 值选取 25, 对应的分割指标反而减少。这是由于当局部区域增加到一定程度, 会将其他不属于该中心节点的其余节点作为邻居节点, 对这些节点进行特征提取与建模时会产生负面效应。因此, 当 K 值选取 20, 网络可以达到最理想的分割性能。

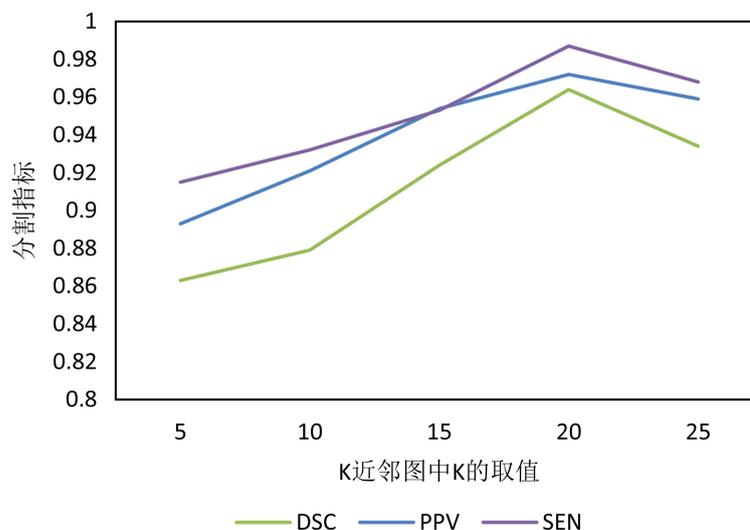


Figure 6. Ablation experiment of K value selection on network segmentation performance

图 6. K 值选取对网络分割性能的消融实验

6. 结论

本文提出的 B-DGB 网络, 使用双边结构对输入网络的坐标信息与法向量信息分别建模, 有效地解决坐标信息与法向量信息直接作为网络输入, 出现特征混淆的情况, 并通过与 PointNet, PointNet++ 以及 MeshSegNet 进行的对比实验, 以及与向量边, 法向量边两种网络结构进行消融实验, 进一步验证了双边结构在解决特征混淆问题的有效实施。通过多尺度动态图子模块, 实现多层次多尺度对局部特征进行提取与建模, 进一步提取增强后的局部特征。最后将先前得到的局部增强特征与全局特征密集融合, 以得到更具有表达能力多属性特征, 有效地解决了现有方法存在的局部特征提取问题。此外, 改进了一种混合损失函数, 对牙齿区域边界分割与相邻牙齿区域的分割实现进一步精细化。将该模型在自制数据集

上进行实验, 与现有点云分割模型 PointNet、PointNet++、MeshSegNet 相比, 分割精度提高, 平均 Dice 系数为 0.972、PPV 为 0.964、SEN 为 0.987。

参考文献

- [1] Le, T., Bui, G. and Duan, Y. (2017) A Multi-View Recurrent Neural Network for 3D Mesh Segmentation. *Computers & Graphics*, **66**, 103-112. <https://doi.org/10.1016/j.cag.2017.05.011>
- [2] Graham, B., Engelcke, M. and Van Der Maaten, L. (2018) 3d Semantic Segmentation with Submanifold Sparse Convolutional Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, 18-22 June 2018, 9224-9232. <https://doi.org/10.1109/CVPR.2018.00961>
- [3] Kumar Y., Janardan R., Larson B., et al. (2011) Improved Segmentation of Teeth in Dental Models. *Computer-Aided Design and Applications*, **8**, 211-224. <https://doi.org/10.3722/cadaps.2011.211-224>
- [4] Yuan, T.Y., Liao, W.H., Dai, N., et al. (2010) Single-Tooth Modeling for 3D Dental Model. *International Journal of Biomedical Imaging*, **2010**, Article ID: 535329. <https://doi.org/10.1155/2010/535329>
- [5] Zhao, M.X., Ma, L.Z., Tan, W.Z., et al. (2006) Interactive Tooth Segmentation of Dental Models. 2005 *IEEE Engineering in Medicine and Biology 27th Annual Conference*, Shanghai, 17-18 January 2006, 654-657. <https://doi.org/10.1109/IEMBS.2005.1616498>
- [6] Ma, Y.Q. and Li, Z.K. (2010) Computer Aided Orthodontics Treatment by Virtual Segmentation and Adjustment. 2010 *International Conference on Image Analysis and Signal Processing*, Xiameng, 9-11 April 2010, 336-339. <https://doi.org/10.1109/IASP.2010.5476100>
- [7] Zou, B.J., Liu, S.J., Liao, S.H., et al. (2015) Interactive Tooth Partition of Dental Mesh Base on Tooth-Target Harmonic Field. *Computers in Biology and Medicine*, **56**, 132-144. <https://doi.org/10.1016/j.compbiomed.2014.10.013>
- [8] Xu, X.J., Liu, C. and Zheng, Y.Y. (2018) 3D Tooth Segmentation and Labeling Using Deep Convolutional Neural Networks. *IEEE Transactions on Visualization and Computer Graphics*, **25**, 2336-2348. <https://doi.org/10.1109/TVCG.2018.2839685>
- [9] Tian, S.K., Dai, N., Zhang, B., et al. (2019) Automatic Classification and Segmentation of Teeth on 3D Dental Model Using Hierarchical Deep Learning Networks. *IEEE Access*, **7**, 84817-84828. <https://doi.org/10.1109/ACCESS.2019.2924262>
- [10] Lian, C.F., Wang, L., Wu, T.H., et al. (2019) Meshsnet: Deep Multi-Scale Mesh Feature Learning for End-to-End Tooth Labeling on 3d Dental Surfaces. In: Shen, D., et al. Eds., *Medical Image Computing and Computer Assisted Intervention*, Springer, Cham, 837-845. https://doi.org/10.1007/978-3-030-32226-7_93
- [11] Qi, C.R., Su, H., Mo, K., et al. (2017) PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Piscataway, 21-26 July 2017, 77-85.
- [12] Cui, Z.M., Li, C.J., Chen, N.L., et al. (2021) TSegNet: An Efficient and Accurate Tooth Segmentation Network on 3D Dental Model. *Medical Image Analysis*, **69**, Article ID: 101949. <https://doi.org/10.1016/j.media.2020.101949>
- [13] Qi, C.R., Yi, L., Su, H., et al. (2017) PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 5105-5114.
- [14] Yu, C.Q., Gao, C.X., Wang, J.B., et al. (2021) BiSeNet V2: Bilateral Network with Guided Aggregation for Real-Time Semantic Segmentation. *International Journal of Computer Vision*, **129**, 3051-3068. <https://doi.org/10.1007/s11263-021-01515-2>
- [15] Zhang, Z.H., Chen, G.L., Wang, X. and Shu, M.C. (2021) DDRNet: Fast Point Cloud Registration Network for Large-Scale Scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, **175**, 184-198. <https://doi.org/10.1016/j.isprsjprs.2021.03.003>
- [16] Defferrard, M., Bresson, X. and Vandergheynst, P. (2016) Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. *Advances in Neural Information Processing Systems*, **29**, 3844-3852.
- [17] Wang, Y., Sun, Y.B., Liu, Z.W., et al. (2018) Dynamic Graph CNN for Learning on Point Clouds. *ACM Transactions on Graphics*, **38**, 146-158. <https://doi.org/10.1145/3326362>
- [18] Thomas, H., Qi, C.R., Deschaud, J.E., et al. (2019) KPConv: Flexible and Deformable Convolution for Point Clouds. 2019 *IEEE/CVF International Conference on Computer Vision*, Seoul, 27 October-2 November 2019, 6410-6419. <https://doi.org/10.1109/ICCV.2019.00651>
- [19] Berman, M., Triki, A.R. and Blaschko, M.B. (2018) The Lovasz-Softmax Loss: A Tractable SI Urrogate for the Optimization of the Intersection-Over-Union Measure in Neural Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 44413-4421. <https://doi.org/10.1109/CVPR.2018.00464>