

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2024)12-3786-14

论文引用格式: Zhou X W, Zhu Y, Ge J Y, Pang Q J, Wei R and Gu M. 2024. Dental model segmentation network with fine-grained receptive fields and multiscale fusion. Journal of Image and Graphics, 29(12):3786-3799(周新文, 朱洋, 葛峻沂, 潘钱家, 魏然, 顾敏. 2024. 具有细粒度感受野的多尺度融合口腔模型分割. 中国图象图形学报, 29(12):3786-3799)[DOI:10.11834/jig.230769]

具有细粒度感受野的多尺度融合口腔模型分割

周新文^{1,2}, 朱洋¹, 葛峻沂², 潘钱家², 魏然^{2,3}, 顾敏^{2*}

1. 常州大学计算机与人工智能学院, 常州 213164; 2. 苏州大学附属第三医院\常州市第一人民医院, 常州 213003;

3. 常州大学医学与健康工程学院, 常州 213164

摘要: **目的** 从口内扫描点云模型上精确分割牙齿是计算机辅助牙科治疗中重要的任务, 但存在手动执行耗时且烦琐的问题。近年来, 计算机视觉领域涌现出一些端到端的方法实现三维形状分割。然而, 大多数方法没有注意到口腔分割需要网络具有更加细粒度的感受野, 因此分割精度仍然受到限制。为了解决该问题, 设计了一个端到端的具有细粒度感受野的全自动牙齿分割网络——TRNet, 用于在未加工的口内扫描点云模型上自动分割牙齿。**方法** 首先, TRNet使用了具有细粒度感受野的编码器, 其基于多尺度融合从不同的尺度提取到更全面的口腔模型特征, 并通过更适合口腔模型分割的细粒度分组查询半径以及具有相对坐标归一化的特征提取层来提升分割性能。其次, TRNet采用了基于层级连接的特征嵌入方式, 网络学习到口腔模型中由各个局部区域到覆盖更大范围空间的关键特征, 特征提取更全面, 提升了网络的分割精度。同时, TRNet使用了基于软性注意力机制的特征融合方式, 使网络更好地从融合特征中关注到口腔模型的关键信息。**结果** 使用由口内扫描仪获取的患者口内扫描点云模型数据集评估了TRNet。经过5折交叉验证的实验结果中, TRNet的总体准确率(overall accuracy, OA)达到了97.015±0.096%, 平均交并比(mean intersection over union, mIoU)达到了92.691±0.454%, 显著优于现有方法。**结论** 实验结果表明, 提出的具有细粒度感受野的多尺度融合口腔分割模型在口内扫描点云模型上取得了较好表现, 提高了网络对于口腔模型的分割能力, 使点云分割结果更准确。

关键词: 自动口腔模型分割; 点云; 细粒度感受野; 多尺度特征融合; 坐标归一化; 软性注意力机制

Dental model segmentation network with fine-grained receptive fields and multiscale fusion

Zhou Xinwen^{1,2}, Zhu Yang¹, Ge Junyi², Pang Qianjia², Wei Ran^{2,3}, Gu Min^{2*}

1. School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou 213164, China;

2. The Third Affiliated Hospital of Soochow University, The First People's Hospital of Changzhou, Changzhou 213003, China;

3. School of Medical and Health Engineering, Changzhou University, Changzhou 213164, China

Abstract: **Objective** Dental computer-aided therapy relies on the use of dental models to aid dentists in their practice.

收稿日期: 2023-10-27; 修回日期: 2024-01-19; 预印本日期: 2024-01-26

* 通信作者: 顾敏 gumin106@163.com

基金项目: 教育部产学合作协同育人项目(220605181012316); 江苏省产学研合作项目(BY20230265); 江苏省研究生科研与实践创新计划项目(YPC22020142, YPC23020155); 江苏省“双创博士”项目(JSSCBS20210908, JSSCBS20210896); 常州市科技项目(CJ20235041); 常州市卫健委青年人才科技项目(QN202310, QN202355); 常州市“十四五”卫生健康高层次人才培养工程(2022260); 2023年苏州大学研究生教育改革成果奖培育项目(KY20231517); 常州大学大学生课外创新创业基金项目(QZX22020187); 常州大学优秀本科毕业设计(论文)培育项目(QZX23020036)

Supported by: University-Industry Collaborative Education Program (220605181012316); Jiangsu Industry-University-Research Cooperation Program (BY20230265); Jiangsu Province Graduate Student Research and Practice Innovation Program (YPC22020142, YPC23020155)

One of the most fundamental tasks in dental computer-aided therapy is the automated segmentation of teeth using point cloud data obtained from intra-oral scanners (IOS). The precise segmentation of each individual tooth in this procedure provides vital information for a variety of subsequent tasks. These segmented dental models facilitate customized treatment planning and modeling, thus providing extensive assistance in carrying out further treatments. However, the automated segmentation of individual teeth from dental models faces three significant challenges. First, the indistinct boundary between teeth and gums poses difficulties in segmentation based solely on geometric features. Second, certain factors, such as occlusion during scanning, can lead to suboptimal results, particularly in posterior dental regions, thereby further complicating the segmentation process. Lastly, teeth often exhibit complex anomalies in patients, including crowding, missing teeth, and misalignment issues, which further complicate the task of accurate segmentation. To address these challenges, two conventional methods are proposed for segmenting teeth in images obtained from IOS scanners. The first method employs a projection-based approach, wherein a 3D dental scan image is initially projected into a 2D space, segmentation is then performed in a 2D space, and the result is remapped back into the 3D space. The second method adopts a geometry-based approach and typically utilizes geometric attributes, such as surface curvature, geodesic information, harmonic fields, and other geometric properties, to distinguish tooth structures. However, these methods are not fully automated and rely on domain-specific knowledge and experience. Moreover, the predefined low-level attributes used by these methods lack robustness when dealing with the complex appearance of patients' teeth. Considering the impactful application of convolutional neural networks (CNN) in computer vision and medical image processing, several deep learning methods rooted in CNN have been introduced. Some of these methods directly extract translation-invariant depth geometric features from 3D point cloud data but suffer from a lack of necessary receptive field for fine-grained tasks, such as dental model segmentation. Moreover, the network structure exhibits redundancy and neglects the crucial details of dental models. To address these issues, a fully automatic tooth segmentation network called TRNet is proposed in this paper, which can automatically segment teeth on unprocessed intra-oral scanned point cloud models.

Method In the proposed end-to-end 3D point cloud-based multi-scale fusion dental model segmentation method, an encoder with a fine-grained receptive field is employed to address those challenges posed by the small size of each tooth within the dental model and the lack of distinct features between the teeth and gums. Each tooth within the dental model is relatively small in comparison to the entire dental model, and the boundaries between the teeth and gums lack distinct features. Consequently, a fine-grained receptive field is essential for extracting features from this model. The network adopts a small radius for querying the neighborhood, thus narrowing the receptive field and enabling the network to focus on detailed features. Additionally, downsampling can lead to the uneven density of the point cloud, thereby causing the network trained on sparse point clouds to struggle in recognizing fine-grained local structures. Multiscale feature fusion coding is implemented to address these issues. Given that the encoder uses a small query radius to create a fine-grained receptive field, the relative coordinates become relatively small. Consequently, the network needs to learn large weights to operate on these relative coordinates, thereby introducing further challenges in network optimization. TRNet normalizes the relative coordinates in the feature extraction layer to facilitate network optimization and enhance segmentation performance. The network also employs a highly efficient decoder. Previous segmentation methods often utilize the U-Net structure, which incorporates jump connections for multi-level feature aggregation between the input features of the cascaded decoder and the outputs of the corresponding layer encoder. However, this top-down propagation is considered inefficient for feature aggregation. The decoding approach used by TRNet directly combines the features outputted from all cascade encoders, thereby allowing the network to learn the importance of each cascade. The discrepancies in scales or dimensions of the features represented by fused information in the network may also introduce unwanted bias during the fusion process. To address these issues and ensure that the network focuses on crucial information within the fused features, a soft attention mechanism is incorporated into the fusion process. Specifically, a soft attention operation is performed on the newly combined features after their connection, thereby enabling the network to adaptively balance the discrepancies of different scales or levels in the propagated features.

Result A dataset comprising dental models taken from numerous patients with irregular tooth shapes, such as crowding, misalignment, and underdeveloped teeth, was compiled. To establish the labeled values, an experienced dentist meticulously segmented and annotated these models. The dataset was then randomly divided into two subsets, with 146 models allocated for training

and 20 models reserved for testing. Data augmentation techniques, such as random panning and scaling, were employed to enhance the diversity of the training set. In each iteration, intra-oral scan images were shifted by a randomly selected value within the range of $[-0.1, 0.1]$ and scaled by a randomly chosen magnification within the range of $[0.8, 1.25]$, thereby generating new training data. Experimental results from a 5-fold cross-validation reveal that TRNet achieved an overall accuracy (OA) of $97.015 \pm 0.096\%$ and a mean intersection over union (mIoU) of $92.691 \pm 0.454\%$, significantly outperforming the existing methods. **Conclusion** An end-to-end deep learning network called TRNet is introduced in this paper for the automatic segmentation of teeth in 3D dental images acquired from intra-oral scanners. An encoder with fine-grained receptive fields was also implemented to enhance the local feature extraction capabilities essential for dental model segmentation. Additionally, a decoder based on hierarchical connections was employed to allow the network to decode efficiently by learning the significance of each level. This refinement significantly improves the precision of dental model segmentation. A soft attention mechanism was also integrated into the feature fusion process to enable the network to focus on key information within dental model features. Experimental results indicate that TRNet shows excellent performance on intra-oral scanned point cloud models and enhances the ability of the network to segment dental models, thereby improving the accuracy of point cloud segmentation results.

Key words: automatic dental model segmentation; point cloud; fine-grained receptive fields; multiscale feature fusion; coordinate normalization; soft attention mechanism

0 引言

口腔计算机辅助治疗利用口腔模型作为输入,协助牙医进行治疗(Zhou等,2019)。其最基本的任务是对口内扫描仪(intra-oral scanners, IOS)获取的口内扫描点云模型数据执行牙齿的自动分割。在这一任务中,每颗牙齿的准确分割为后续多种任务提供了至关重要的信息。基于这些分割得到的口腔模型可以进行特定的治疗规划和建模等工作,为进一步的治疗提供有力支持。

从口腔模型中稳健地自动分割每个牙齿面临3方面的困难。1)牙齿和牙龈之间的边界不明显,这使得难以基于几何特征进行分割;2)在扫描过程中,由于咬合等原因,可能会导致口腔后牙区扫描结果不利于口腔模型分割;3)患者的牙齿通常有复杂的异常状况,如牙齿拥挤、缺失以及错位的问题,会造成分割困难。

为了应对这些挑战,目前已有研究者提出了两种传统方法,用于在IOS扫描生成的图像上进行牙齿的分割。第1种方法是基于投影的方式(Kondo等,2004; Wongwaen和Sinthanayothin,2010),首先将三维口腔扫描图像投影到二维空间中,然后在二维空间中执行分割,并将分割结果重新映射回三维空间。第2种方法是基于几何的方法,通常利用表面曲率(Yuan等,2010)、测地信息(Sinthanayothin和

Tharanont,2008)和谐波场(Zou等,2015)等几何属性来区分牙齿结构。然而,这些方法通常并非完全自动化的,且依赖特定领域的知识和经验。此外,这些方法所使用的预定义低级属性在应对患者复杂的牙齿外观时缺乏鲁棒性。

鉴于卷积神经网络(convolutional neural network, CNN)在计算机视觉和医学图像处理领域的成功应用,一些基于CNN的深度学习方法也被提出用于牙齿分割。例如,Xu等人(2019)引入了一种多级框架,将三维特征重新映射为二维图像块,通过对牙齿表面的网格单元进行卷积神经网络分类来实现分割,从而取得了出色的分割性能。然而,这种方法在将输入的三维特征局部关联为图像时,未充分考虑空间单元特征的无序性,导致不同的重建方式会生成不同的重建结果。另一方面,Mask-MCNet(Zanjani等,2019a)采用基于体积锚的区域建议网络来进行牙齿分割,但其中的提案生成模块可能会降低分辨率,并需耗费大量内存资源。

PointNet(Qi等,2017a)的问世开创了三维点云特征学习的新纪元。它通过直接从三维点云数据中学习平移不变的深度几何特征,在分类和分割任务上取得了成功。然而,PointNet虽然能够完成分类和分割任务,却难以捕捉细粒度分割所需的局部几何背景。随后,一些先进方法如PointNet++(Qi等,2017b)和PointCNN(Li等,2018)相继提出,进一步提升了三维点云处理的性能。PointNet++专注于建

模局部上下文信息,而PointCNN从加权和排列的角度进行空间变换,从而提升细节分割的精度。但是这些方法缺失了分割口腔模型这种细粒度工作所需要的感受野,网络结构也存在一些冗余性,同时网络一定程度上缺失了对于口腔模型关键信息的关注。

为了应对上述问题,本文提出了端到端的具有细粒度感受野的多尺度特征融合牙齿分割网络,本文将命名为TRNet,其通过细粒度感受野编码器提取口腔模型细粒度特征,在基于层级连接的解码器中使网络中不同层级的口腔模型关键特征相互补充,同时通过基于软性注意力机制的特征融合方式使网络能够更多地关注到口腔模型的关键信息。TRNet主要有3个关键部分。

1)使用了一个具有细粒度感受野的编码器,其基于具有细粒度查询半径的多尺度融合编码器从不同的尺度提取到更全面的口腔模型特征,同时缩小特征提取时的感受野,以应对口腔模型形状特性中的边缘几何特征不明显的情况。在此基础上,对下采样参数进行调整以适应网络的输入点数,且归一化相对坐标解决原相对坐标较小导致网络优化困难的问题,提升分割性能。

2)使用了一个更高效的解码器,其直接组合各层级传播得到的口腔模型特征信息,让网络学习每个层级的重要性,以应对口腔模型结构的牙齿小牙龈大的复杂情况,不同层级上的口腔模型关键特征相互补充,使提取到的特征更为完善,进一步提升分割精度。

3)对于网络中的特征融合方式,加入了软性注意力机制,通过注意力权重使网络关注到各个融合特征中的关键信息,将特征权重更多地分配给更重要的特征。

1 研究现状

1.1 口腔模型分割

1.1.1 非深度学习方法

许多基于预先选择的几何特性的非深度学习传统方法已经被提出用于分割3D口腔模型,大致可以分为:基于曲面曲率的方法、基于曲面轮廓线的方法、基于谐波场的方法以及基于投影的非深度学习方法。

基于曲率的方法通常利用负曲率特征将曲面划

分为不同部分。Yuan等人(2010)计算了牙齿表面的最小曲率,并提取牙齿与牙龈的边界来进行分类。Zhao等人(2005)提出了一种基于三角形网格曲率值的交互式分割方法。Kumar等人(2011)开发了一个系统,用户可以通过直观的滑块设置特定曲率阈值。然而,这些基于曲率的方法对于口腔面的变化非常敏感,一些微小的变化可能会导致分割准确性下降。

基于曲面轮廓线的方法允许在分割过程中进行人机交互,因此在生成牙齿边界时更可靠。具体来说,在这些方法(Ma和Li, 2010; Sinthanayothin和Tharanont, 2008)中,用户可以在口腔模型上手动选择牙齿边界标志,然后算法根据测地信息连接每对相邻点。虽然实现了更好的性能,但是这些方法要求用户多次平移或旋转3D模型,以仔细选择特定的标记,这既烦琐又耗时,不利于处理大量的口腔模型分割工作。

基于谐波场的方法(Zou等, 2015)允许用户先标注有限数量的表面点作为先验,然后使用谐波场来分割牙齿。与前面的方法相比,该方法以最少的交互提供了更高效和更智能的用户界面,但由于需要专业领域的人工操作,这种半自动化方法的有效性在很大程度上取决于操作员的专业知识。

基于投影的方法通过将口腔模型的3D图像投影为2D图像,从而有效地分割3D口腔模型。Yamany和El-Bialy(1999)利用曲率和表面法线信息构建了二维图像表示,并提取了高低曲率结构作为分割依据。Kondo等人(2004)提出在平面视图和全景视图图像上检测牙齿特征。然而,这些方法通常在口腔模型存在多处咬合不正的情况时表现不佳。

1.1.2 深度学习方法

随着技术发展,领域内涌现出多种基于深度学习的方法用于从口腔模型中分割牙齿。Xu等人(2019)提出将预定义的手工特征转化为2D图像进行分类。Tian等人(2019)则首次采用稀疏八叉树对口内扫描图像进行体素化,然后使用3D卷积神经网络实现牙齿分割。然而,这些方法将口腔扫描图像转换为网格结构,未充分考虑三维几何数据的无序性,可能在体素化过程中引入量化误差。另外,Zanjani等人(2019a)提出了一种基于Mask R-CNN(He等, 2017)的网络,用于分割从口内扫描图像提取的3D点云,但提案生成模块可能导致分辨率下降。之后,

Zanjani 等人(2019b)提出了基于生成对抗网络的口腔模型分割网络,利用判别器从真实标签和预测标签的统计量中提取信息,提升分割网络的预测性能。Lian 等人(2019)通过整合多图约束学习模块,逐层提取细粒度的局部几何特征,但该方法在内存利用方面效率不高。基于此,Wu 等人(2022)在这一工作的基础上,将图形约束学习模块中的对称平均池化替换为边缘卷积操作,显著减少了计算复杂度。Cui 等人(2021)提出了一种基于 PointNet++ 的两阶段牙齿分割方法,首先检测所有牙齿的质心,然后对每颗牙齿进行分割。然而,两阶段方法使训练复杂,而且生成牙齿质心的训练需要大量口腔模型样本。Zhao 等人(2021)提出了一种基于图卷积的方法,利用一系列图注意力卷积层和全局结构分支提取局部几何特征和全局特征,以全面学习分割任务的信息。在前述工作的基础上,Zhao 等人(2022)还提出了双流网络框架 TSGCNet (two-stream graph convolutional network),分别从口腔模型的坐标和法向量独立学习特征表示,有效减少了不同几何输入类型的混淆。Duan 和 Chen (2023)在 TSGCNet 的基础上提出了 SGTNet (semantics-based feature learning with graph-Transformer network),其通过加入图 Transformer 以及自适应特征聚合,进一步提升分割性能。Hao 等人(2022)提出了双模块系统,包含深度学习模块和金丝雀模块,前者生成精细的分割输出,后者评估输出准确性并进行人工校正。Qiu 等人(2022)提出通过估计牙弓来辅助生成牙齿质心,提高分割准确率,同时减少分割训练所需的标签注释,从而降低数据准备成本。

1.2 三维点云分割

三维理解在计算机视觉中扮演着重要的角色。在不同的输入数据中,3D 点云表示由于其灵活性越来越受欢迎,并且已应用于深度学习架构中,用于各种场景的分割任务。这些方法主要分为两类:基于图的方法和基于点的方法。

1.2.1 基于图的方法

图卷积因其灵活性而取得了巨大的成功,许多基于图卷积神经网络(graph convolutional network, GCN)的方法已经用于三维形状识别和分割(Liang 等,2020;Wang 等,2019a)。这些方法通常将三维数据表示为图,其中点或单元之间的空间关系被建模为图的边,然后使用基于空间的图卷积来汇聚每个

节点的局部信息(Wang 等,2019b)。然而,这种方法仅仅对相关点之间的连接性进行建模,可能无法准确地区分特征相似但不同的点。在口腔模型中,可能会造成相邻牙齿以及牙齿与牙槽骨之间的分割不准确。

1.2.2 基于点的方法

基于点的方法旨在使用深度学习架构直接处理三维几何数据。PointNet(Qi 等,2017a)通过多层感知机和对称函数,将非结构化的点云作为输入,从不规则的点云中学习平移不变的几何特征。尽管 PointNet 在点分割和表面分割任务中取得了令人满意的成果,但由于其结构独立地学习每个单元的特征,容易忽略三维形状上的局部空间关系,难以捕捉细粒度分割所需的局部上下文。为了克服这个限制,PointNet++(Qi 等,2017b)构建了一个层次结构,通过递归应用 PointNet 来利用三维形状上的局部空间关系。这种方法能够学习带有递增上下文尺度的局部特征,已在多个分割和分类任务上表现出色。其他方法,如 GS-Net(geometry sharing network)(Xu 等,2020)通过整合几何共享模块进一步扩展了 PointNet++ 的功能,Jiang 等人(2019)通过融合边缘分支来增强局部信息传递。类似地,PointCNN(Li 等,2018)通过 χ -变换重新组织无序点,以执行一般的卷积操作,同样取得了优异的结果。此外,邱云飞和牛佳璐(2023)通过融合原型对齐算法并计算查询集点云特征与支持集原型的欧氏距离实现点云分割,减少了对样本数的依赖。

然而,这些深度学习方法并没有考虑到口腔模型细粒度特征提取所需的感受野,且往往结构冗余,未很好地提取到口腔模型的关键特征。TRNet 通过多尺度融合和多层级连接使不同尺度以及不同层级的信息进行相互补充,通过穿插于其间的细粒度感受野和软性注意力机制进行分离性的精细度分割,两者相结合,能够更好地分割和分类牙齿。

2 本文方法

2.1 网络结构

TRNet 将通过采样得到的口腔模型点云的三维坐标作为输入,为每个点输出唯一的标签。网络结构如图 1 所示,图中展示了牙齿分割任务中特征是如何在网络中编码和传播,其中每个滤波器下方的

数字代表其输入和输出的通道数。TRNet网络包含具有细粒度感受野的特征编码器、高效的特征解码器以及带有注意力机制的特征融合方式。为了从全部的点集中提取特征,网络编码器使用了一个分层结构,其层次结构由3个特征编码器组成,该结构将来自各个小区域的特征组合成覆盖更大范围空间的抽象特征。每个特征编码器包括采样层、分组层和特征提取层。在特征解码器中,采用了基于距离

的插值和融合所有层级编码器得到的特征进行解码的策略,将特征传播至原始点集,最后为每个点输出最终的标签值。

本文方法采用了具有更适合口腔模型分割的细粒度感受野的特征编码器以及更高效的特征解码器,它们都为分割网络带来了更精确的分割结果。同时,对于各个的特征融合加入软性注意力机制,以进一步提升网络的分割性能。

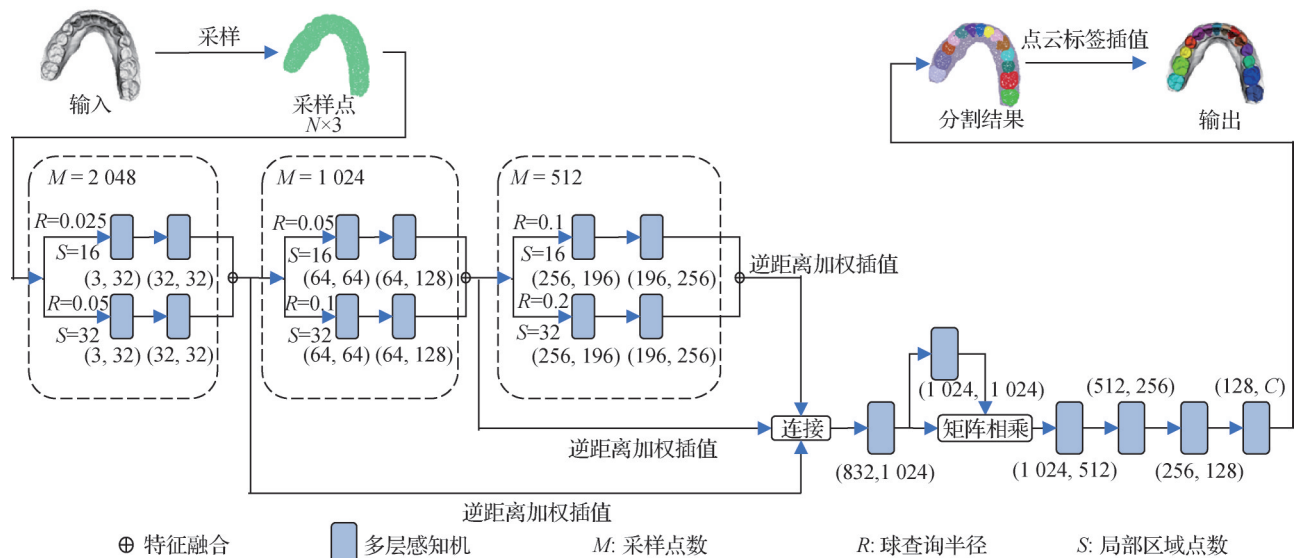


图1 用于3D口腔模型分割的TRNet网络框架

Fig. 1 The TRNet network framework for 3D dental model segmentation

2.2 具有细粒度感受野的编码器

本文方法中的编码器经过最远点采样后,通过查询半径进行分组,得到以每个采样点为中心、相应查询半径下的局部特征。查询半径的大小决定了分组后每个采样点周围局部特征的半径大小,也就是网络提取局部特征时的感受野。口腔模型中,每个牙齿相较于整个口腔模型非常小,且牙齿与牙龈边缘也不具有非常明显的特征,因此需要细粒度的感受野来进行口腔模型的特征提取。本文通过设定特征分组的相关参数(Cui等,2021),缩小感受野,使网络能够关注到更详细的特征。在此基础上,为了使网络进一步适应其输入点数,也对编码器中的采样点数参数进行了调整。同时,下采样可能会带来点云密度不均匀的问题,稀疏点云训练的模型可能无法识别细粒度的局部结构,因此采用了多尺度特征融合的编码方式,通过不同尺度上的特征相互补充,不仅能够补偿下采样点云的密度不均,也从不同尺度上带来了更全面的口腔模型特征。网络的特征提

取层设计如图1所示,每个特征提取层采用两个多层感知机(multilayer perceptron, MLP)模块。然而,较小的分组半径使得特征分组后的相对坐标较小,导致网络在MLP进行特征提取时需要学习一个更大的权值作用于该相对坐标,这不利于网络训练时的优化。因此,通过设置的查询半径,对特征分组后的局部相对坐标进行归一化,扩大特征分组得到的相对坐标,最终特征提取过程可表示为

$$y = h_{\theta} \left[x; \frac{\Delta p}{r} \right] \quad (1)$$

式中, x, y 分别代表输入和输出的特征, Δp 是分组过程中的相对坐标, r 为查询半径, h_{θ} 即代表以连接输入特征 x 和相对坐标 Δp 作为输入的多层感知机。经过归一化相对坐标后,网络无需再学习更大的权值作用于相对坐标,也就不再过度依赖于优化技术,为模型提供了更多最优解的可能性,因此网络性能也得到了进一步提升。

2.3 基于层级连接的特征嵌入

对于特征解码,为了得到更好的上采样结果,应该对不同级别的特征进行更好的聚合。网络中较低的层级对应于较小尺度的局部特征,而较高的层级对应于较大尺度的局部特征。先前的一些分割方法使用了 U-Net 结构,其采用跳跃连接进行级联解码器的输入特征与对应层级编码器输出的特征进行多级特征聚合,但是这种自上而下的传播方式对于聚合特征并不是很有效。本文的网络采用直接将所有层级编码器输出的口腔模型特征组合起来的解码方式,让网络学习每个层级的重要性,以应对口腔模型结构的牙齿小牙龈大的复杂情况。其中,越高级的特征对应越趋向于模型全局的特征;越低级的特征对应越趋向于模型局部的特征,不同层级上的口腔模型关键特征相互补充,提取得到更为完善的特征。网络通过组合来自不同层的多尺度单元特征,从而产生编码不同层级的局部信息特征矩阵。然而,由于分层特征提取时,网络逐步对每层编码器的输入进行二次下采样,在各层局部特征的特征点数需要统一为初始输入的点数,本文通过逆距离加权插值的方法(Qi 等, 2017b),将特征的通道数统一为初始输入点数 M ,并进行连接操作。一个插值点 $\mathbf{x}$ 的特征被计算为

$$f(\mathbf{x}) = \frac{\sum_{i=1}^3 w_i(\mathbf{x}) f(\mathbf{x}_i)}{\sum_{i=1}^3 w_i(\mathbf{x})} \quad (2)$$

式中, $w_i(\mathbf{x}) = 1/d(\mathbf{x}, \mathbf{x}_i)$ 为逆距离权重, $\mathbf{x}_i$ 为距离最近的 3 个邻居。网络通过连接逆距离加权插值的特征(即 f^1 、 f^2 和 f^3)得到组合特征后,在此基础上将 MLP 运用于组合特征以学习该特征视图的更高级的 1 024 维特征表示 $\mathbf{F}$,此操作可表示为

$$\mathbf{F} = \text{MLP}(f^1 \oplus f^2 \oplus f^3) \quad (3)$$

最终,将多层感知机应用于该视图表示,得到 $M \times C$ 的矩阵,其中每行表示每个点属于 C 个不同分割类别的概率。

2.4 基于软性注意力的特征融合

在编码器部分,每个编码器都具有两个不同查询半径的流,融合它们的输出可以使网络能够从不同尺度更全面地理解口腔模型的结构。在解码器部分,同样地,融合各个编码层提取的局部信息特征矩阵可以使网络从不同的层级理解口腔模型的特征。

这些融合的信息所表示的特征的尺度或维度的不同可能会导致融合过程中存在不必要的偏差。为了解决这个问题,使网络更好地关注融合特征中的关键信息,我们在融合过程中加入了软性注意力机制,即在特征连接后对形成的新的组合特征进行注意力运算,使网络自适应地平衡不同尺度或不同层级传播的特征,其可以表示为

$$\mathbf{F}_{\text{out}} = \beta_{F_{\text{in}}} \times \mathbf{F}_{\text{in}} \quad (4)$$

式中, $\beta_{F_{\text{in}}} = \text{MLP}_{\text{att}}(\mathbf{F}_{\text{in}})$ 是注意力机制的权重,其可以由对输入特征 $\mathbf{F}_{\text{in}}$ 通过一个多层感知机 MLP_{att} 计算得到,输出 $\beta_{F_{\text{in}}}$ 的大小与 $\mathbf{F}_{\text{in}}$ 相同。通过带有注意力机制的特征融合方式,本文网络调整提取的特征中关键信息的占比,进而关注到口腔模型分割所需的关键信息,提高网络的性能。

3 实验与分析

本节根据从医院口腔科收集的包括上下颌的数据集评估 TRNet。数据集、评价指标、实现细节以及实验评估详细描述如下。

3.1 数据集与评价指标

实验所使用的精确可靠的三维数字口腔模型数据均由苏州大学附属第三医院口腔科的专业口腔医师通过 iTero 数字化口内扫描仪(iTero Element Flex, 美国)扫描患者获得。数据集由同一位口腔医师对数据进行标签标注工作,以确定单颗牙齿边界。每个口腔模型分为 15 个部分和类别,即从左右中切牙至左右第二磨牙共 14 个牙齿,以及牙龈。在 TRNet 中,牙齿表面的每个点都由三维输入进行描述,即三维空间坐标。数据集包括 166 个口腔模型,其中包括许多牙齿形状异常的病例,如牙齿拥挤、错位,以及牙齿未发育完全。我们将数据集随机均衡地划分为 5 份,通过 5 折交叉验证对分割效果进行了量化评价(平均值±标准差),实验结果如表 1。为了解决网络模型在数据加载时下采样的有效特征丢失问题,本文的网络在每次迭代时都会对输入数据重新下采样。此外,还通过随机平移和随机缩放来对训练集进行数据增强。具体来说,在每次迭代中,每个口腔模型都在 $[-0.1, 0.1]$ 之间随机选取的值中进行位移,并以 $[0.8, 1.25]$ 之间随机选取的倍率进行缩放,形成新的数据进行训练。

表 1 在 5 折交叉验证下现有深度学习方法和本文 TRNet 方法的分割结果
Table 1 The segmentation results quantified under 5-fold cross-validation for existing deep learning methods and TRNet

方法	所有部分		各个类(IoU)					/%
	OA	mIoU	切牙	尖牙	前磨牙	磨牙	牙龈	
PointNet	88.527±0.754	74.669±2.021	71.525±2.819	72.319±2.371	71.121±2.944	71.290±3.101	87.088±0.696	
PointNet++	91.435±0.297	79.709±0.807	72.309±0.498	81.668±0.821	74.378±2.398	79.106±0.990	91.083±0.201	
DGCNN	92.611±0.269	82.638±1.132	77.588±1.754	80.074±1.871	81.098±1.437	83.531±1.404	90.898±0.628	
TSGCNet	94.587±0.403	87.493±1.377	81.186±2.811	87.342±1.734	89.224±2.527	86.753±1.431	92.960±0.547	
SGTNet	96.094±0.584	90.825±1.870	86.326±3.296	91.320±1.751	91.838±2.279	90.048±3.087	94.993±0.145	
TRNet	97.015±0.096	92.691±0.454	89.298±0.985	91.183±0.484	92.470±0.470	93.918±0.391	95.986±0.262	

注:加粗字体表示各列最优结果。

实验在一台配备有一个英特尔 XEON CPU 和 3 块 Nvidia RTX4090 GPU(每块显存 24 GB)的工作站上进行,使用的是 CUDA 11.4 和 CUDNN 8.0.5,操作系统为 Ubuntu。所有代码都是基于深度学习框架 PyTorch 1.10.0 开发的。在实现阶段,训练整个框架 251 个周期,将小批量大小设置为 8,选用初始学习率为 0.001 的 Adam 优化器。

3.2 与其他方法比较

将本文的框架与几种相关的现有分割方法进行比较,包括 PointNet (Qi 等, 2017a)、PointNet++ (Qi 等, 2017b)、DGCNN (Wang 等, 2019b)、TSGCNet (Zhao 等, 2022)和 SGTNet (Duan 和 Chen, 2023)。这几种方法在牙齿分割任务中取得了不错的性能。为了进行公平比较,我们使用相同的数据集设置训练这些方法和 TRNet。下文简要概述这些方法。

1) PointNet: 以和 TRNet 类似的方式训练 PointNet, 设置下采样点数 $N = 20\,000$ 进行训练, 训练后的网络处理牙齿表面, 以输出 $N \times 4$ 矩阵, 其中前 3 列为牙齿表面点的三维坐标, 第 4 列为牙齿标签。将小批量大小设置为 8, 对 PointNet 进行了 251 个周期的训练。

2) PointNet++: 该方法在 PointNet 的基础上, 通过对相邻点之间的空间关系进行分层建模, 进一步优化性能。实验中网络模型采用的是其多尺度分组的版本。训练测试集的划分也与所有实验保持一致。在训练期间, 小批量大小设置为 16, 网络训练了 251 个周期。

3) DGCNN: 该方法将图卷积的思想运用到了三维点云上, 通过描述点与其相邻点之间的关系得到

边缘特征, 然后通过边缘特征显式地构造局部图, 并且在网络的每一层之后动态更新。实验中将小批量大小设置为 2, 网络同样训练了 251 个周期。

4) TSGCNet: 该方法采用两个图形学习流, 分别从坐标和法向量中提取更具辨别性的高级几何表示, 并将从不同的流学习到的特征表示进一步融合, 用于预测。实验中将小批量大小设置为 2, 网络训练了 251 个周期。

5) SGTNet: 该方法在 TSGCNet 的基础上, 加入图 Transformer 以充分利用语义信息, 更准确地掌握局部和非局部依赖关系, 并对跨域特征进行自适应特征聚合, 进一步优化性能。同样, 实验中将小批量大小设置为 2, 网络的迭代次数设置为 251。

3.2.1 定量评估

表 1 列出了几种方法的分割结果, 其中包含几个主要的评价指标。与 PointNet 相比, 其他几种方法产生了更好的结果。TRNet 使总体准确率 (overall accuracy, OA) 和平均交并比 (mean intersection over union, mIoU) 分别提高了 8.488% 和 18.022%, 这证明了对于局部特征的建模对于在三维点云上进行更精确的分割是必要的。与 PointNet++ 相比, 本文方法的 OA 和 mIoU 也分别有了 5.580% 和 12.982% 的提升。同样与 DGCNN 相比, 本文方法的 OA 和 mIoU 分别提升了 4.404% 和 10.053%, 这是因为 DGCNN 对相关点之间的边进行建模, 容易忽略特征相近点的区分, TRNet 是直接对点进行更精确的建模, 受到的影响较少。与 TSGCNet 相比, 本文方法的 OA 和 mIoU 分别有 2.428% 和 5.198% 的提升, 同样与 SGTNet 相比, 本文方法的 OA 和 mIoU 分别提升

了0.921%和1.866%,证明多尺度特征融合结构对口腔模型分割的重要性,其可以使分割网络具有更好的分割性能。在分割口腔模型的不同部分时,多数情况下TRNet都比其他现有方法获得了更高的准确率和IoU值,这表明TRNet在处理不同的牙齿时具有更强的泛化能力。

3.2.2 定性评估

图2中将本文自动分割代表性的示例结果进行直观比较。第1列是从唇向角度观察牙齿分割的结果;第2列展示了牙齿小、牙冠特征不明显时的分割情况;第3列展示了口腔深处扫描结果复杂的情况,

易影响到第二磨牙的分割;第4列代表了牙齿拥挤而造成错位的情况。可以看到,在这些情况中,其他现有方法都明显出现了将牙龈或其他牙齿的一部分识别为目标牙齿的情况,相较而言,TRNet整体性能都优于现有方法,能够更胜任一些有难度的牙齿分割任务。图2所示的视觉评估结果与先前的定量评估结果一致,这进一步表明了TRNet在牙齿分割任务中的实用性。

3.3 消融实验

本文进行了详细的消融实验,以评估TRNet关键组成部分的效果。

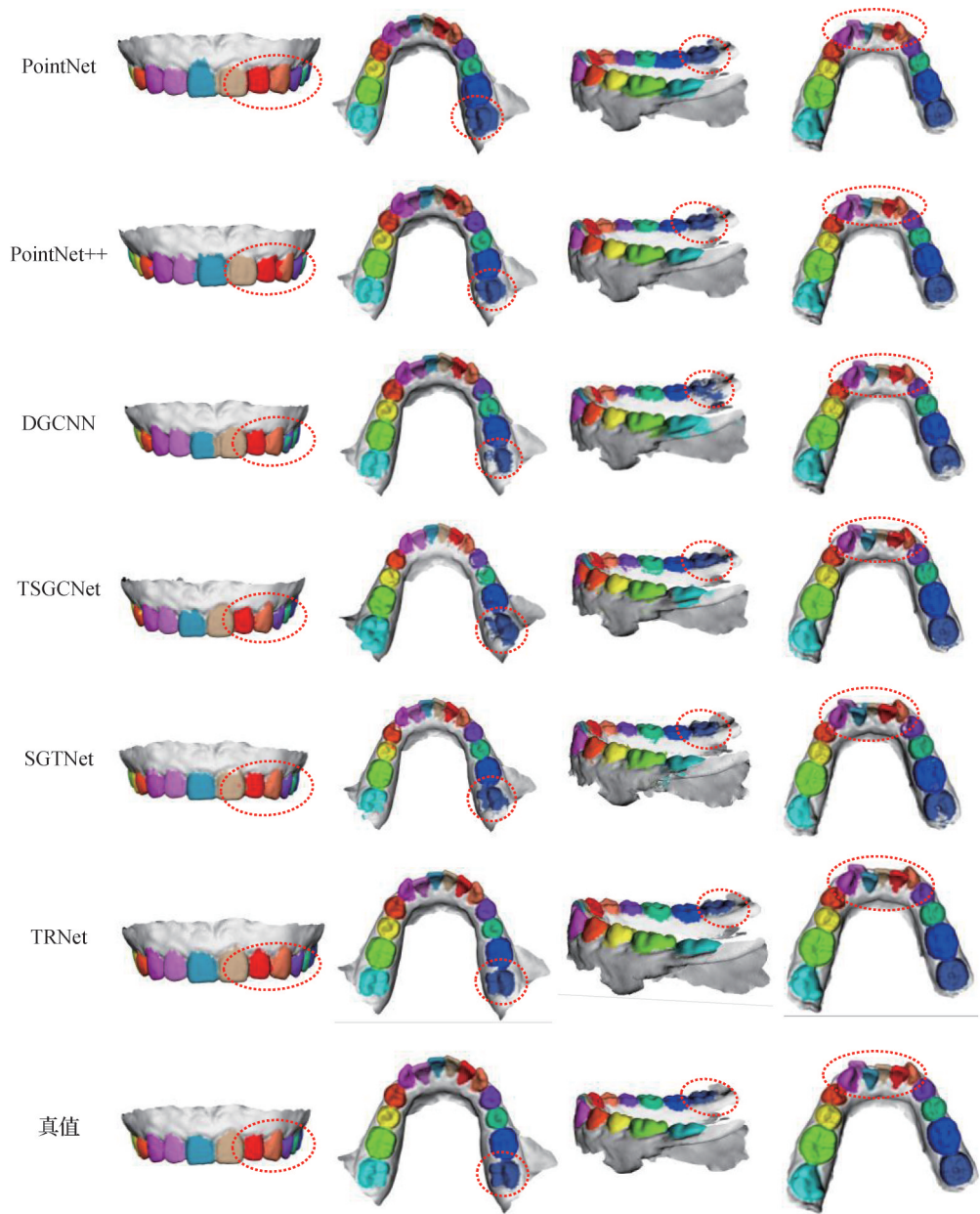


图2 6种方法分割口腔模型的典型实例
Fig. 2 Typical examples of dental models segmentation through six methods

3.3.1 细粒度编码器的效果实验

本文的网络使用了具有细粒度感受野的编码器,具体来说,针对口腔模型的特性,采用了更合适的分组查询半径,提升网络的细粒度感知能力,在此基础上重新调整了下采样层的参数以适应网络的输入点数,并归一化组内的相对坐标降低网络优化难度,提升网络的精度。为了评估本文设计的有效性,移除了特征提取中的相对坐标归一化,并选用默认的查询半径参数设置形成本文方法的变体 TRNet_{cm},然后在 TRNet_{cm}的基础上改变编码器参数设置,使之与 TSegNet 的骨干网络保持一致,形成另一个变体 TRNet_{ts},本文对比了 TRNet_{cm}、TRNet_{ts} 和 TRNet 的分割结果,定量结果见表 2。可以看出,网络在失去细粒度查询半径和相对坐标归一化后,分割精度有所下降,其中 OA 减少了 3.532%, mIoU 减少了

8.145%,口腔的不同部分的分割结果也都有较大的性能降低,而使用 TSegNet 的编码器参数后,TRNet_{ts} 所得到的指标相较于 TRNet_{cm} 有所升高,但相较于 TRNet 仍然有较大的下降。

本文也进行了定性评估,如图 3 所示,图中的口腔模型存在两个中切牙间隙较大的情况,从可视化示例可以看到,TRNet_{cm} 与 TRNet_{ts} 的结果中都出现了部分牙龈被误识别为牙齿,而使用本文的具有细粒度感受野的编码器得到了更接近标签值的结果。由此可以看出,本文所使用的编码器可以提升网络的性能,在口腔模型牙齿不规整的复杂情况下依然能够取得较好的分割结果,原因在于其为网络带来了口腔模型分割所需的更好的局部特征提取能力,并且更易于网络优化,提供了更多最优解的可能性,使网络得到更好的分割性能。

表 2 在 5 折交叉验证下 TRNet、TRNet_{cm} 和 TRNet_{ts} 的分割结果
Table 2 Segmentation results quantified under 5-fold cross-validation for TRNet, TRNet_{cm} and TRNet_{ts}

方法	所有部分		各个类(IoU)				
	OA	mIoU	切牙	尖牙	前磨牙	磨牙	牙龈
TRNet _{cm}	93.483±0.490	84.546±1.662	77.877±2.843	83.657±1.752	83.572±2.039	85.606±1.793	92.021±0.415
TRNet _{ts}	93.748±0.349	85.041±0.773	80.452±0.598	83.232±1.089	83.758±0.900	85.515±1.243	92.252±0.419
TRNet	97.015±0.096	92.691±0.454	89.298±0.985	91.183±0.484	92.470±0.470	93.918±0.391	95.986±0.262

注:加粗字体表示各列最优结果。

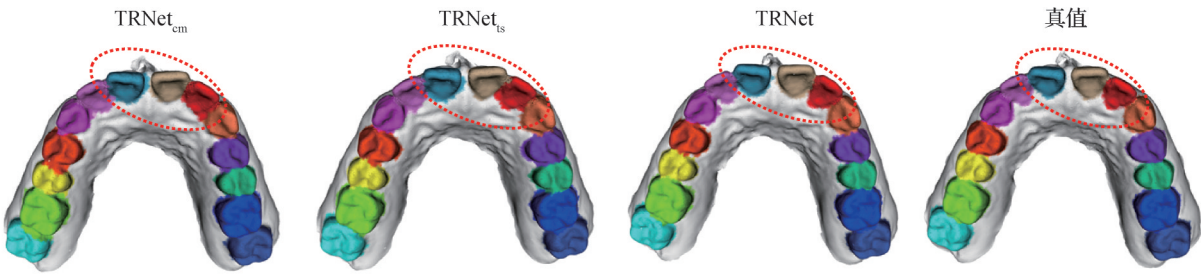


图 3 TRNet、TRNet_{cm} 和 TRNet_{ts} 对比的一个典型实例
Fig. 3 A typical example of comparing TRNet, TRNet_{cm} and TRNet_{ts}

为了进一步验证编码器中的采样点数参数的重新设置和相对坐标归一化对于分割性能的影响,本文也分别进行了消融实验,如表 3 所示,将 TRNet_{ts} 作为 baseline,从结果可以看出,各模块均能提高网络的性能,再次证明了它们的有效性。

3.3.2 特征嵌入的效果实验

本文的网络也使用了能够学习不同层级特征重

要性的特征解码器,该方式直接将所有层级编码器输出的特征组合起来,让网络进行解码。先前其他的一些相关分割方法使用了 U-Net 结构,它们使用跳跃连接来进行级联解码器的输入特征与对应层级编码器输出的特征进行多级特征聚合,然而,这种自上而下的传播方式对于聚合特征并不高效。为了评估本文所使用的特征嵌入的有效性,比较了使用 U-

表3 在5折交叉验证下不同编码器设置的分割结果

Table 3 Segmentation results quantified under 5-fold cross-validation for different encoder settings

					/%
Baseline	采样点数调整	相对坐标归一化	OA	mIoU	
√	-	-	93.748±0.349	85.041±0.773	
√	-	√	95.407±0.130	88.852±0.749	
√	√	-	94.639±0.167	87.206±0.850	
√	√	√	97.015±0.096	92.691±0.454	

注:加粗字体表示各列最优结果,“√”表示使用该模块,“-”表示未使用该模块。

Net结构和本文中使用的特征嵌入方式的分割结果,定量结果见表4。可以看出,相较于使用本文网络使用的特征嵌入方式,使用U-Net结构使得分割结果有了一定的下降,其中OA相差0.547%,mIoU相差1.334%,口腔的不同部分的分割结果也都有较大的性能降低。本文也进行了定性评估,如图4所示,定性评

估结果中,使用U-Net结构导致牙齿与牙龈间以及牙齿间的识别不准确,而本文所使用的解码器得到了更接近标签值的结果,由此可见,本文网络使用的特征解码器更适合口腔模型分割,这是由于其可以使网络学习不同层级的重要性,进而提取所有层次中更为重要的口腔模型特征,提高了网络的精度。

表4 在5折交叉验证下使用U-Net结构和使用本文特征嵌入的分割结果

Table 4 Segmentation results quantified under 5-fold cross-validation for method using U-Net structure and our method

方法	所有部分		各个类IoU					/%
	OA	mIoU	切牙	尖牙	前磨牙	磨牙	牙龈	
U-Net结构	96.468±0.194	91.357±0.710	88.975±1.414	90.414±0.994	91.081±0.673	92.963±0.610	95.348±0.323	
所有层级融合学习的特征嵌入结构	97.015±0.096	92.691±0.454	89.298±0.985	91.183±0.484	92.470±0.470	93.918±0.391	95.986±0.262	

注:加粗字体表示各列最优结果。

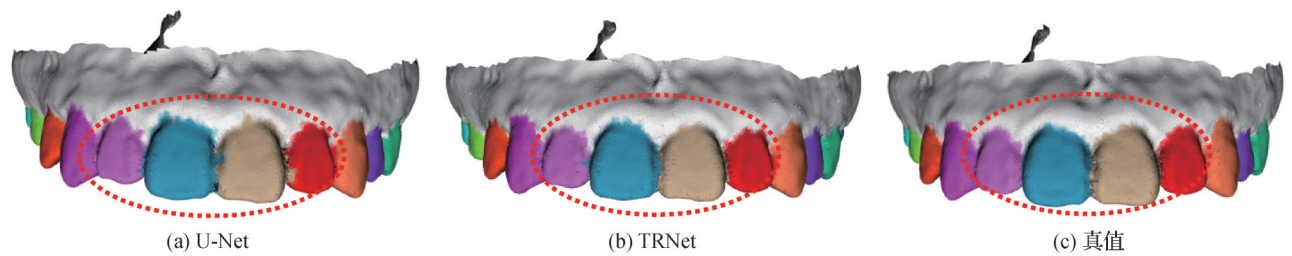


图4 使用U-Net结构和本文方法对比的一个典型实例

Fig. 4 A typical example of comparing method using U-Net structure and our method((a)U-Net;(b)TRNet;(c)ground truth)

3.3.3 软性注意力机制的效果实验

本文的网络在特征融合中加入了注意力机制,具体来说,该方式通过注意力权重调整在提取的特征中口腔模型的关键信息的占比。为了评估该特征融合方式的有效性,将特征融合中的注意力机制移除形成一个变体来进行对比,定量结果见表5。可

以看出,去除了注意力机制后分割结果有了一定的下降,其中OA减少了0.370%,mIoU减少了1.002%,口腔的不同部分的分割结果都有所提升,尤其是切牙部分提升最明显。本文也进行了定性评估,如图5所示,定性评估结果中,相较于没有注意力机制的特征融合方式,具有注意力机制的特征融

合方式分割得到的牙齿边缘更加光滑。由此可见，特征融合的注意力机制的加入使网络的分割性能得到了提升，原因在于其可以使网络在特征提取时关注到口腔模型的关键信息，提升网络性能。

表 5 在 5 折交叉验证下有与无软性注意力机制的特征融合的分割结果
Table 5 Segmentation results quantified under 5-fold cross-validation for feature fusion with and without soft attention mechanism

方法	所有部分		各个类(IoU)					/%
	OA	mIoU	切牙	尖牙	前磨牙	磨牙	牙龈	
无软性注意力机制特征融合	96.645±0.060	91.689±0.168	87.324±0.375	90.853±0.466	91.624±0.499	93.031±0.264	95.613±0.183	
有软性注意力机制特征融合	97.015±0.096	92.691±0.454	89.298±0.985	91.183±0.484	92.470±0.470	93.918±0.391	95.986±0.262	

注：加粗字体表示各列最优结果。

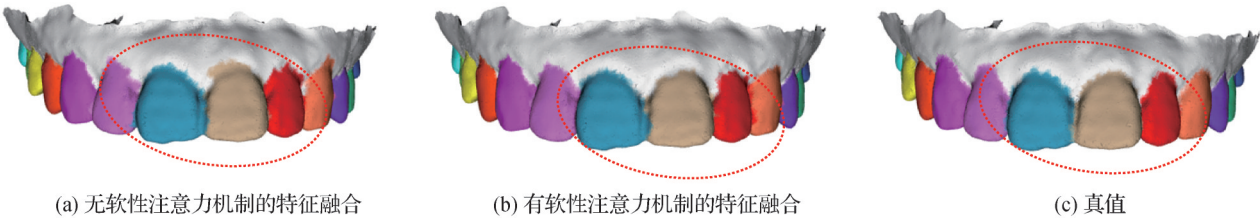


图 5 不同特征融合方式对比的一个典型实例

Fig. 5 A typical example of comparing different feature fusion methods
((a)with soft attention mechanism; (b)withou soft attention mechanism; (c)ground truth)

4 结 论

本文提出了一种端到端的深度学习网络 TRNet，用于对口腔扫描仪获取的口腔三维图像进行牙齿的自动分割。TRNet 使用了 3 个关键部件，首先是具有细粒度感受野的编码器，其采用更适用于口腔模型分割的查询半径来提升分割所需的局部特征提取能力，并归一化相对坐标解决原生相对坐标较小导致网络优化困难的问题。然后是基于层级连接的解码器，其使得网络模型具有更为高效的解码能力，学习到口腔模型中由各个局部区域到覆盖更大范围空间的关键特征，使特征提取更全面，进一步提高口腔模型分割精度。最后是在特征融合过程中加入了注意力机制，其使网络关注到口腔模型特征的关键信息。这样，通过多尺度融合和多层级连接针对口腔模型整体性的分割，通过穿插于其间的细粒度感受野和软性注意力机制进行分离性的精细度分割，两者相结合，由此本文的网络性能得到大幅度的提升，从而实现更精确的牙齿分割。本文在患者数据集上对提出的 TRNet 和其他 5 种现有的方法进行了比较。结

果证明了本文提出方法的优越性，特别是对于分割中具有挑战性的情况。

尽管 TRNet 在三维口腔模型分割任务中取得了领先的性能，但本文方法目前仅针对的是单颌具有 14 颗牙齿的口腔模型，尚未涉及处理存在智齿的特殊情况。同时，本文中也有部分模型分割尚不理想的情况。如图 6 出现两个相邻牙之间分割精确度较低，一个可能的原因是由于扫描的问题使得得到的口腔模型相邻牙部分出现粘连，导致分界处几何特征不明显。这种情况的口腔模型数据量并不多，导致训练阶段很少被网络注意到。

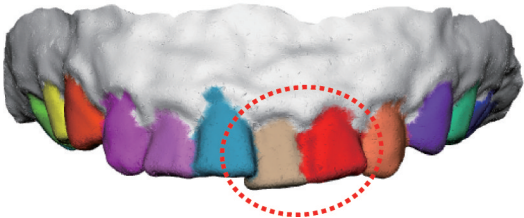


图 6 本文方法分割不佳的一个实例

Fig. 6 An example of the poor segmentation by our method

为了解决这些问题，在未来的模型设计中将考虑智齿的情况，并添加包含智齿的口腔模型作为训练样本，同时也要探索更有效的方法来处理口腔模

型异常的情况,进一步提高边界分割的精度。

参考文献(References)

- Cui Z M, Li C J, Chen N L, Wei G D, Chen R N, Zhou Y F, Shen D G and Wang W P. 2021. TSegNet: an efficient and accurate tooth segmentation network on 3D dental model. *Medical Image Analysis*, 69: #101949 [DOI: 10.1016/j.media.2020.101949]
- Duan F and Chen L. 2023. 3D dental mesh segmentation using semantics-based feature learning with graph-transformer//*Proceedings of the 26th International Conference on Medical Image Computing and Computer Assisted Intervention-MICCAI 2023*. Vancouver, Canada: Springer: 456-465 [DOI: 10.1007/978-3-031-43990-2_43]
- Hao J, Liao W, Zhang Y L, Peng J, Zhao Z, Chen Z, Zhou B W, Feng Y, Fang B, Liu Z Z and Zhao Z H. 2022. Toward clinically applicable 3-dimensional tooth segmentation via deep learning. *Journal of Dental Research*, 101(3): 304-311 [DOI: 10.1177/00220345211040459]
- He K M, Gkioxari G, Dollár P and Girshick R. 2017. Mask R-CNN//*Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy: IEEE: 2961-2969 [DOI: 10.1109/ICCV.2017.322]
- Jiang L, Zhao H S, Liu S, Shen X Y, Fu C W and Jia J Y. 2019. Hierarchical point-edge interaction network for point cloud semantic segmentation//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, Korea (South): IEEE: 10433-10441 [DOI: 10.1109/ICCV.2019.01053]
- Kondo T, Ong S H and Foong K W C. 2004. Tooth segmentation of dental study models using range images. *IEEE Transactions on Medical Imaging*, 23(3): 350-362 [DOI: 10.1109/TMI.2004.824235]
- Kumar Y, Janardan R, Larson B and Moon J. 2011. Improved segmentation of teeth in dental models. *Computer-Aided Design and Applications*, 8(2): 211-224 [DOI: 10.3722/cadaps.2011.211-224]
- Li Y Y, Bu R, Sun M C, Wu W, Di X H and Chen B Q. 2018. PointCNN: convolution on X -transformed points//*Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Montréal, Canada: Curran Associates Inc.: #31
- Lian C F, Wang L, Wu T H, Liu M X, Durán F, Ko C C and Shen D G. 2019. MeshsNet: deep multi-scale mesh feature learning for end-to-end tooth labeling on 3D dental surfaces//*Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention-MICCAI 2019*. Shenzhen, China: Springer: 837-845 [DOI: 10.1007/978-3-030-32226-7_93]
- Liang Z D, Yang M, Li H and Wang C X. 2020. 3D instance embedding learning with a structure-aware loss function for point cloud segmentation. *IEEE Robotics and Automation Letters*, 5(3): 4915-4922 [DOI: 10.1109/LRA.2020.3004802]
- Ma Y Q and Li Z K. 2010. Computer aided orthodontics treatment by virtual segmentation and adjustment//*Proceedings of 2010 International Conference on Image Analysis and Signal Processing*. Xiamen, China: IEEE: 336-339 [DOI: 10.1109/IASP. 2010. 5476100]
- Qi C R, Su H, Mo K and Guibas L J. 2017a. Pointnet: deep learning on point sets for 3D classification and segmentation//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE: 652-660 [DOI: 10.1109/CVPR.2017.16]
- Qi C R, Yi L, Su H and Guibas L J. 2017b. 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, USA: Curran Associates Inc.: 5105-5114
- Qiu L D, Ye C J, Chen P, Liu Y B, Han X G and Cui S G. 2022. DArch: dental arch prior-assisted 3D tooth instance segmentation with weak annotations//*Proceedings of 2012 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, USA: IEEE: 20752-20761 [DOI: 10.1109/CVPR52688. 2022.02009]
- Qiu Y F and Niu J L. 2023. Point cloud segmentation algorithm fusing few-shot meta-learning and prototype alignment. *Journal of Image and Graphics*, 28(12): 3884-3896 (邱云飞, 牛佳璐. 2023. 融合小样本元学习和原型对齐的点云分割算法. *中国图象图形学报*, 28(12): 3884-3896) [DOI: 10.11834/jig.220942]
- Sinthanayothin C and Tharanont W. 2008. Orthodontics treatment simulation by teeth segmentation and setup//*Proceedings of the 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*. Krabi, Thailand: IEEE: 81-84 [DOI: 10.1109/ECTICON.2008.4600377]
- Tian S K, Dai N, Zhang B, Yuan F L, Yu Q and Cheng X S. 2019. Automatic classification and segmentation of teeth on 3D dental model using hierarchical deep learning networks. *IEEE Access*, 7: 84817-84828 [DOI: 10.1109/ACCESS.2019.2924262]
- Wang L, Huang Y C, Hou Y L, Zhang S M and Shan J. 2019a. Graph attention convolution for point cloud semantic segmentation//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, USA: IEEE: 10296-10305 [DOI: 10.1109/CVPR.2019.01054]
- Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M and Solomon J M. 2019b. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics*, 38(5): 1-12 [DOI: 10.1145/3326362]
- Wongwan N and Sinthanayothin C. 2010. Computerized algorithm for 3D teeth segmentation//*Proceedings of 2010 International Conference on Electronics and Information Engineering*. Kyoto, Japan: IEEE: 277-280 [DOI: 10.1109/ICEIE.2010.5559877]
- Wu T H, Lian C F, Lee S, Pastewait M, Piers C, Liu J, Wang F, Wang L, Chiu C Y, Wang W C, Jackson C, Chao W L, Shen D G and Ko C C. 2022. Two-stage mesh deep learning for automated tooth segmentation and landmark localization on 3D intraoral scans.

- IEEE Transactions on Medical Imaging, 41 (11) : 3158-3166 [DOI: 10.1109/TMI.2022.3180343]
- Xu M Y, Zhou Z P and Qiao Y. 2020. Geometry sharing network for 3D point cloud classification and segmentation//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA: AAAI: 12500-12507 [DOI: 10.1609/aaai.v34i07.6938]
- Xu X J, Liu C and Zheng Y Y. 2019. 3D tooth segmentation and labeling using deep convolutional neural networks. IEEE Transactions on Visualization and Computer Graphics, 25 (7) : 2336-2348 [DOI: 10.1109/TVCG.2018.2839685]
- Yamany S M and El-Bialy A M. 1999. Efficient free-form surface representation with application in orthodontics//Proceedings of the Three-Dimensional Image Capture and Applications II. San Jose, USA: SPIE: 115-124 [DOI: 10.1117/12.341053]
- Yuan T R, Liao W H, Dai N, Cheng X S and Yu Q. 2010. Single-tooth modeling for 3D dental model. International Journal of Biomedical Imaging, 2010(1): #535329 [DOI: 10.1155/2010/535329]
- Zanjani F G, Moin D A, Claessen F, Cherici T, Parinussa S, Pourtahe-rian A, Zinger S and de With P H N. 2019a. Mask-MCNet: instance segmentation in 3D point cloud of intra-oral scans//Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention-MICCAI 2019. Shenzhen, China: Springer: 128-136 [DOI: 10.1007/978-3-030-32254-0_15]
- Zanjani F G, Moin D A, Verheij B, Claessen F, Cherici T, Tan T and de With P H N. 2019b. Deep learning approach to semantic segmentation in 3D point cloud intra-oral scans of teeth//Proceedings of the 2nd International Conference on Medical Imaging with Deep Learning (MIDL). London, UK: PMLR: 557-571
- Zhao M X, Ma L Z, Tan W Z and Nie D D. 2005. Interactive tooth segmentation of dental models//Proceedings of 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. Shanghai, China: IEEE: 654-657 [DOI: 10.1109/IEMBS.2005.1616498]
- Zhao Y, Zhang L M, Liu Y, Meng D Y, Cui Z M, Gao C Q, Gao X B, Lian C F and Shen D G. 2022. Two-stream graph convolutional network for intra-oral scanner image segmentation. IEEE Transactions on Medical Imaging, 41 (4) : 826-835 [DOI: 10.1109/TMI.2021.3124217]
- Zhao Y, Zhang L M, Yang C S, Tan Y Y, Liu Y, Li P C, Huang T H and Gao C Q. 2021. 3D dental model segmentation with graph attentional convolution network. Pattern Recognition Letters, 152: 79-85 [DOI: 10.1016/j.patrec.2021.09.005]
- Zhou X W, Gan Y Z, Zhao Q F, Xiong J and Xia Z Y. 2019. Simulation of orthodontic force of archwire applied to full dentition using virtual bracket displacement method. International Journal for Numerical Methods in Biomedical Engineering, 35(5): #e3189 [DOI: 10.1002/cnm.3189]
- Zou B J, Liu S J, Liao S H, Ding X and Liang Y. 2015. Interactive tooth partition of dental mesh base on tooth-target harmonic field. Computers in Biology and Medicine, 56: 132-144 [DOI: 10.1016/j.combiomed.2014.10.013]

作者简介

周新文,男,博士,主要研究方向为人工智能和医学图像处理。E-mail: zhouxw@cczu.edu.cn

顾敏,通信作者,女,博士后(在站),副主任医师,主要研究方向为口腔医学与人工智能交叉的多学科研究。

E-mail: gumin106@163.com

朱洋,男,硕士研究生,主要研究方向为计算机视觉和点云分割。E-mail: 944372600@qq.com

葛峻沂,男,硕士研究生,主要研究方向为人工智能图像识别数据库的建立。E-mail: 527605767@qq.com

潘钱家,男,硕士研究生,主要研究方向为口腔医学与人工智能交叉的多学科研究。E-mail: 2743297453@qq.com

魏然,女,讲师,主要研究方向为生物医学工程和医学图像处理。E-mail: weiran@cczu.edu.cn