

DOI:10.3969/j.issn.1005-202X.2025.03.005

医学影像物理

基于深度学习对CBCT图像采用多视角投影的牙齿分类

刘沐然^{1,2}, 谭敏慧², 张煜¹

1. 南方医科大学生物医学工程学院, 广东 广州 510515; 2. 上海科技大学生物医学工程学院, 上海 201210

【摘要】目的:针对当前基于锥形束计算机断层扫描(CBCT)图像的牙齿分类方法过于依赖精确分割, 缺少对牙齿形态和位置信息的综合利用, 提出一种基于多视角投影和Transformer架构的牙齿分类方法, 可对全年龄层的CBCT图像中的牙齿(包括儿童病例)进行准确的52分类。**方法:**通过引入多视角投影, 结合Transformer架构, 融合语义分割和实例分割, 由粗至细进行牙齿分类任务, 增强对牙齿空间位置信息的利用。采用国际牙科联盟(FDI)两位数牙位标记法对CBCT图像中的牙齿进行分类, 并对多视角融合效果进行评估。**结果:**改进后的方法能够准确区分恒牙与乳牙, 同时有效地进行牙齿编号, 牙齿层面的分类准确率达到0.982。**结论:**基于多视角投影与Transformer架构的牙齿分类方法实现对牙齿类别与位置信息的有效融合, 提高牙齿分类的精度, 为个性化治疗方案的制定提供更为精确的基础。

【关键词】深度学习; 锥形束计算机断层扫描; 牙齿识别; 多视角投影

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2025)03-0313-07

Classification of teeth in CBCT images using deep learning with multi-view projection

LIU Muran^{1,2}, TAN Minhui², ZHANG Yu¹

1. School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China; 2. School of Biomedical Engineering, ShanghaiTech University, Shanghai 201210, China

Abstract: Objective To address the issue that current methods for classifying teeth in cone beam computed tomography (CBCT) images overly rely on precise segmentation and lack utilization of tooth morphology and positional information, a tooth classification method based on multi-view projection and Transformer architecture is proposed for accurately classifying teeth from CBCT images across all age groups, including pediatric cases, into 52 categories. **Methods** The coarse-to-fine tooth classification task was accomplished after enhancing the utilization of spatial positional information of the teeth by incorporating multi-view projection, integrating Transformer architecture, and combining semantic segmentation with instance segmentation. The two-digit notation system of the Federation Dentaire Internationale was adopted for classifying the teeth in CBCT images, and evaluating the effectiveness of multi-view fusion. **Results** The improved method effectively classified and numbered both permanent and deciduous teeth, achieving a tooth-level classification accuracy of 0.982. **Conclusion** The tooth classification method based on multi-view projection and Transformer architecture successfully integrates tooth category and positional information, and improves the accuracies of tooth classification and numbering, providing a more precise foundation for the formulation of personalized treatment schemes.

Keywords: deep learning; cone beam computed tomography; tooth identification; multi-view projection

前言

随着人们生活质量的提升, 口腔健康逐渐为人们所重视。在日常生活中, 不健康的饮食习惯往往

会造成各种口腔疾病, 其中龋齿和牙周疾病最为常见^[1-2]。然而, 口腔疾病的诊断与治疗常常依赖于医生基于图像等信息对病灶进行检测和识别, 通常需要有经验的医生耗费一定的时间与精力^[3]。因此, 为加快诊断治疗进程, 节约时间, 降低成本, 应用机器学习或模式识别等方法辅助牙科领域的诊断与治疗已成为当下研究热点^[3-5]。

在各种成像方法中, 锥形束计算机断层扫描(CBCT)成像因其分辨率高, 辐射剂量相对较低, 空间信息丰富, 广泛应用在牙科领域当中^[6-7]。基于牙齿CBCT进行的分割分类任务有助于专业人士准确

【收稿日期】2024-11-12

【基金项目】国家自然科学基金(82472056); 广东省自然科学基金(2024A1515012004)

【作者简介】刘沐然, 硕士研究生, 研究方向: 生物医学成像与图像处理, E-mail: liumuran2022@163.com

【通信作者】张煜, 博士, 教授, 研究方向: 生物医学工程、医学影像智能处理与分析, E-mail: yuzhang@smu.edu.cn

识别牙齿的分布位置和形态结构^[8],为精细化诊断,设计个性化治疗方案提供新的解决方法。对牙齿进行精确编号有助于确认牙位,记录病情,是整个就医过程的基础之一^[9]。例如,在正畸治疗中,精确的牙位识别对于正畸方案及托槽的设计具有重要意义^[10-11]。然而,目前大部分研究主要针对牙齿的精细分割设计方案,将牙齿的编号当作附属任务,导致牙齿编号不够精确,为后续的人工检查增加了额外的工作量。本研究旨在对全年龄阶段的病例进行牙齿的52分类,基于牙齿CBCT图像,准确分类并标记不同年龄段的恒牙与乳牙,并探究多视角投影方式对牙齿分类准确度的影响。

近年来飞速发展的深度学习,为基于图像数据分割识别牙齿注入了新的动力^[12]。如Tuzoff等^[13]首次将深度学习的方法应用于牙齿全景片,使用Fast R-CNN进行牙齿区域探测,应用启发式算法提高牙齿标号的精确度。然而,基于牙齿全景X光片进行牙齿的分割与识别,由于图像性质,缺失的深度信息一定程度上限制了结果的准确性^[13-15]。因此,有部分学者针对于牙齿的3D CBCT图像开展了研究。Chen等^[16]使用多任务3D全卷积神经网络,结合基于标记符的分水岭方法解决弱边缘难以区分的问题,以分割识别单颗牙齿。Cui等^[17]在精确的分割结果上,将牙齿标号作为附属任务,根据牙齿空间质心信息结合多层感知机进行32分类。以上方法虽然可以较好地地区分单颗牙齿,然而其分类准确性依赖于牙齿精确分割所提供的特征信息,以作为分类器的输入进行分类,也就导致分割过程中牙齿特征的提取直接影响了分类结果。为了解决这一问题,Miki等^[18]使用深度卷积神经网络对CBCT牙齿区域进行7分类,其中卷积层和池化层的稳健性,使得牙齿的分类无需依赖精确的分割结果,但该方法忽略了智齿的识别,也无法给出准确的牙位标定信息。

上述方法将牙齿分类作为附属任务,受限于较小的数据量,导致分类精度较低,且仅对恒牙进行分类,因此主要适用于成人病例。为补充现有研究中

未充分覆盖的儿童病例和更广泛的分类应用,本文旨在获取更丰富的牙齿特征信息,实现所有类别牙齿上的高精度分类。在医学图像处理领域,通常采用多视角观察图像以提高性能,这为牙齿分类精度的提升提供一种有效的解决方案。对3D图像数据从矢状面、冠状面和水平面3个视角分别提取组织或器官的特征信息^[8,19-20]。例如,在多器官分割任务中,Xia等^[20]对三维医学影像数据进行旋转和排列,通过不同的深度神经网络获得不同视角的预测结果,并最大化视角间的相似性,采用协同训练的方式提升各个视角的分割性能。另外采用置信度分数获得不同视角的权重以计算伪标签,与真实值做监督训练。针对牙齿分割识别,李孜政等^[8]采用多视角融合的方式,从水平面、矢状面不同视角的数据特点分别提供分割指导和标签信息以还原CBCT数据,改善了牙齿轮廓边缘和咬合面分割不精的问题。在对结节的探测和分割中,Setio等^[21]不局限于传统的3个视角,而是从9个不同的视角提取组织的二维补丁,通过多视角卷积神经网络降低假阳性,最后融合不同视角的输出以区分探测部分是否存在结节。

本文基于深度学习方法,将牙齿的52分类作为主要任务,利用多视角投影结合Transformer架构对CBCT图像进行处理^[22],通过扩充数据和丰富位置信息,实现对全年龄层的恒牙与乳牙的准确分类。该方法有效解决了因数据稀缺导致的牙齿分类精度低的问题,使牙齿分类任务独立于牙齿分割。此外,还探讨不同视角投影对分类结果准确度的影响。精确的牙位标定不仅帮助医生制定个性化治疗方案和确定手术部位,还能准确记录并传达患者信息,确保医疗信息的准确性和一致性。

1 方法

针对现有研究存在的问题,笔者提出一个针对牙齿分类任务的二阶段由粗到细的深度学习网络框架(图1)。其中,牙齿分类的标签基于国际牙科联盟(Federation Dentaire Internationale, FDI)两位数牙位

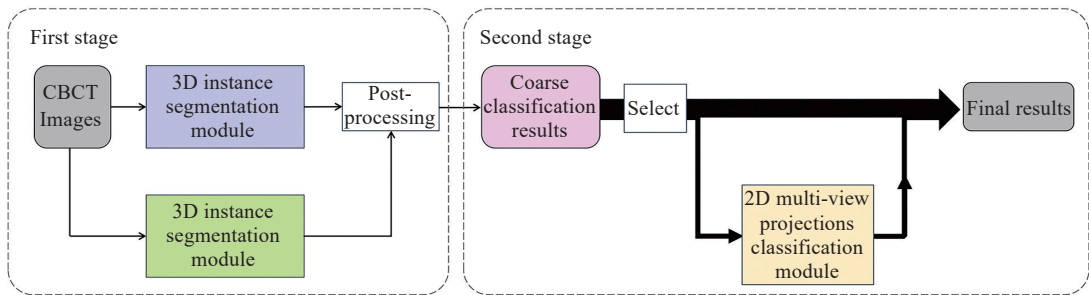


图1 本文提出的网络架构示意图

Figure 1 Schematic diagram of the proposed network architecture

标定系统,将口腔划分为4个象限(右上、左上、左下、右下),每个象限包含13个位置,恒牙和乳牙在各自象限内采用不同编号规则,共计52个编号(如图2所示^[23]),实现对不同生长阶段牙齿的完整标定。

第一阶段的粗分类网络首先对CBCT图像进行实例分割以区分不同的牙齿实例,结合3D语义分割获得牙齿的初步标签;在随后的第二阶段网络中,为进一步提高分类准确度,对第一阶段网络生成的带有初步标签信息的3D牙齿模型进行投影,对投影后的2D图像结合先验信息,利用Transformer^[22]架构获取全局上下文信息,进行更为细致的语义分割;最后将2D图像根据投影时的对应关系反投影回3D模型,获得牙齿的最终分类标签。整个网络实现由粗到细的分类,利用实例分割和语义分割相结合,从多个视角收集信息对牙齿进行分类。为了提高效率,笔者在实验部分探究了视角数量对于分类结果的影响。

1.1 阶段网络

传统的实例分割方法能够精确区分不同的牙齿实例,使每颗牙齿在图像中独立分割,但无法提供标签信息^[24];而传统的语义分割方法可以获得像素(体素)层面的标签信息,但无法精确捕捉牙齿形状和边界。因此,笔者提出将语义分割与实例分割相结合,同时关注像素(体素)层级和实例层级的信息,为每个牙齿实例分配更精确的标签。

具体地,CBCT原始图像首先输入到一个实例分割模块以区分不同牙齿实例^[17]。该模块使用V-Net结构的分层形态引导网络检测表示每颗牙齿,输出牙齿质心和骨架的3D偏移图以及牙齿分割掩码。通过快速聚类方法,每颗牙齿可以通过其质心和骨架结构进行表示。然后,另一个V-Net多任务分割网络基于牙齿质心图、骨架图和感兴趣区域,预测每颗牙齿的体积掩码,实现CBCT图像中的精准牙齿实例分割。

然而,上述工作仅将牙齿作为独立对象进行区分,并未对牙位标签进行仔细考虑。为了实现牙位标定,基于3D语义分割网络^[25]按照FDI牙位标记系统进行52分类(图2)获得体素级标签信息。如图3所示,进行语义分割后的每个体素根据其局部特征被分配一个牙齿标签,但这种密集预测往往会在同一牙齿区域内的不同体素间产生多个不同的标签。此外,儿童病例的稀缺性也使乳牙的区分度较低。因此,笔者考虑使用体素级投票策略来统一每颗牙齿区域内所有体素的标签。具体来说,首先通过实例分割模块提取单颗牙齿实例区域S,并将其对应到语义分割后的牙齿区域R。随后,提取R区域内的所有标签信息进行统计,选择体素占比最大的标签作为最终标签,将该标签映射回对应的牙齿实例区域S,生成第一阶段网络牙齿粗分类结果。

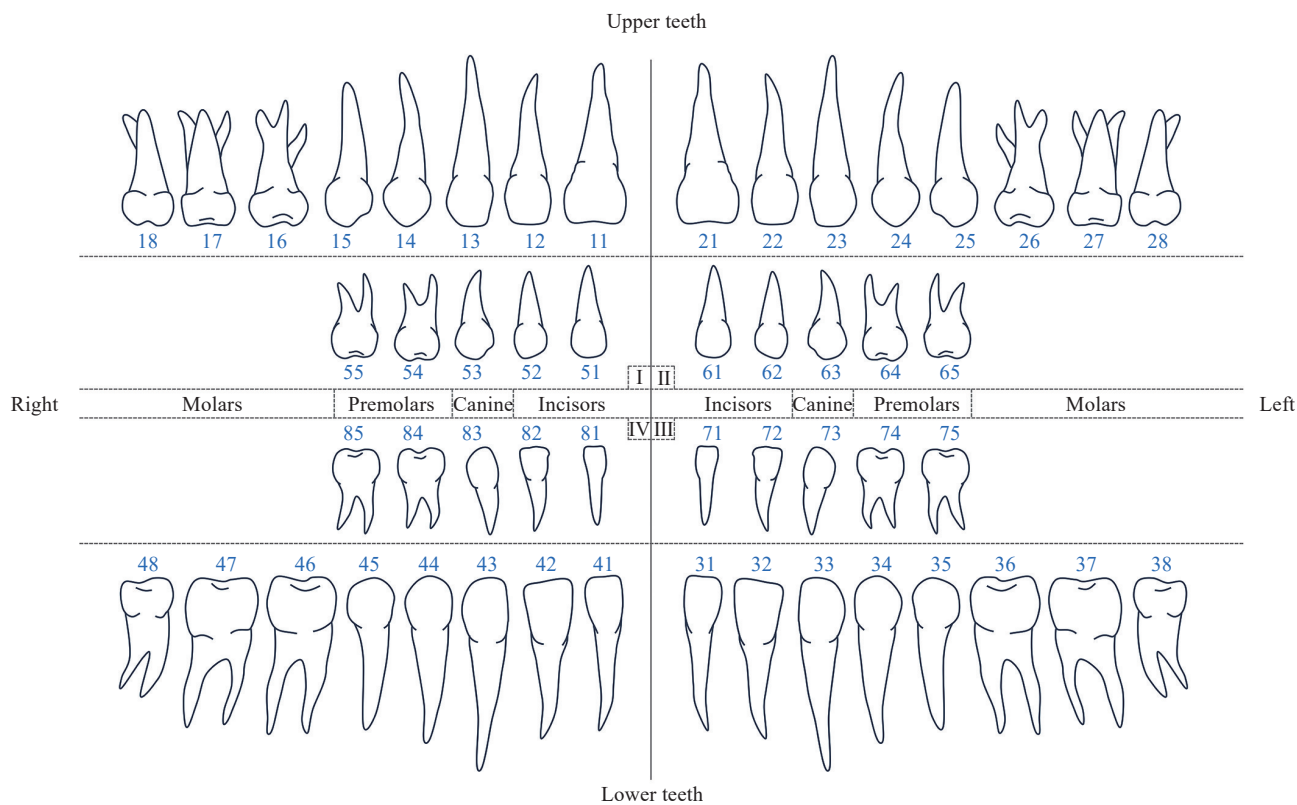


图2 基于FDI的52类牙齿分类编号
Figure 2 FDI-standard 52-class tooth classification and numbering

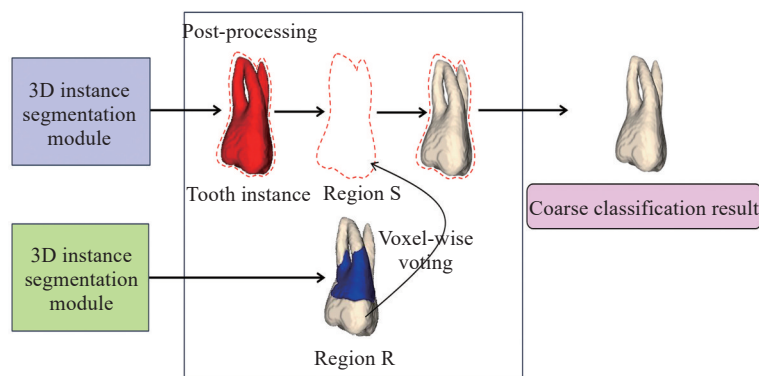


图3 牙齿的后处理示意图

Figure 3 Post-processing flowchart

1.2 二阶段网络

在经过了第一阶段网络的粗分类后,得到了带有牙齿清晰轮廓边界和初步标签的3D牙齿模型。经过检查发现,单纯结合语义分割与实例分割的方法能够准确识别牙齿所在象限,同时区分象限内的切牙、尖牙以及前磨牙(如图2所示),但对磨牙和乳牙的分类准确度较低,这主要是由于智齿和乳牙的数据相对较少。一方面,由于第三磨牙(智齿)的特殊性使得磨牙标定具有一定挑战:不同个体的第一、第二和第三磨牙数量不一,部分人可能缺少一颗或多颗磨牙,同时各磨牙的形状和位置高度相似。另一方面,包含乳牙的儿童CBCT图像较为稀缺。因此模型在这些牙齿类别上的表现相对不佳。基于以上问题,笔者希望通过扩充数据集来实现磨牙和乳牙分类精度的提升。基于粗分类网络获得的3D牙齿模型,采用多视角投影生成多个2D投影图像,使用一个包含Transformer架构的2D语义分割网络^[26],通过多头注意力机制捕获上下文信息,进一步做细粒度的图像语义分割。

为实现细粒度分割,不同于第一阶段网络中对全口牙进行52分类,该部分只对第一阶段网络中预测错误的牙齿分为13个类别,忽略其口腔象限信息,从而在数据稀缺的情况下降低模型的预测难度,其中恒牙标记为1~8,乳牙标记为9~13。第一阶段分类网络预测错误的原因在于单颗牙齿区域内大部分体素被预测成错误标签。为了区分这样的牙齿,提出设置一个阈值 $t \in (0,1)$,具体区分方法为:单颗牙齿区域内,最多体素被预测为同一个标签的占比如果大于 t ,则认为该标签置信度够高,将其作为该颗牙齿的最终预测标签;其占比若小于 t ,则认为该颗牙齿存在预测困难,提取其第一阶段网络输出结果的象限信息,送入第二段网络做进一步分类。在本实验中,将 t 设置为0.7。

如图4所示,在确定需要进行细分类的牙齿实例

后,对其做特殊标记 T' ,使用开源软件ITK-SNAP (Version 4.0)对3D模型进行可视化。将模型绕Z轴旋转 30° 、 60° 、 90° 、 120° 、 150° 和 180° ,并进行重建,保留可观测到 T' 牙齿的视角集合 P 。在每个视角 P_i 下,沿X轴将模型投影到Y-Z平面,获得2D图像 I_i ,并记录其映射关系。随后,将 I_i 中除了 T' 区域外的其他牙齿的52类标签转化为不含象限信息的13类标签,保留前景牙齿并遮盖高重叠率的背景牙齿,使用2D语义分割网络^[26]对标记为 T' 的牙齿区域进行细粒度分割。其中 $i \in (1,2,\dots,n)$, n 表示能观测到 T' 牙齿区域的视角总数。2D语义分割网络首先对二维图像进行卷积特征提取,随后通过多个Transformer块捕获全局语义信息,并结合前馈网络处理局部细节。特征融合模块逐层融合多尺度特征,最终完成13类分类。由于二维图形中背景占比过大的情况,笔者使用加权的交叉熵损失函数优化模型,以应对类别不平衡的问题,公式如下:

$$L = - \frac{1}{|R|} \sum_{j \in R} w_{y_j} \cdot \log(\hat{y}_j) \quad (1)$$

其中, R 为感兴趣区域的体素集合, y_j 表示第 j 个体素的真实类别标签, w_{y_j} 权重的定义为:

$$w_{y_j} = \frac{\max(\text{class_counts})}{\text{class_counts}[y_j]} \quad (2)$$

通过每个类别标签在训练数据中的频次分配权重,以应对类别不平衡的问题。

最后,对该样本所有投影图像集合 I 的分类结果进行像素级投票,选择 T' 区域中像素级预测占比最大的标签作为细分类结果。结合第一阶段网络的象限信息,转化为FDI牙位标定的52类标签,并根据投影映射关系,将标签信息反投影回三维空间,得到最终的预测结果。

2 实验

2.1 数据集

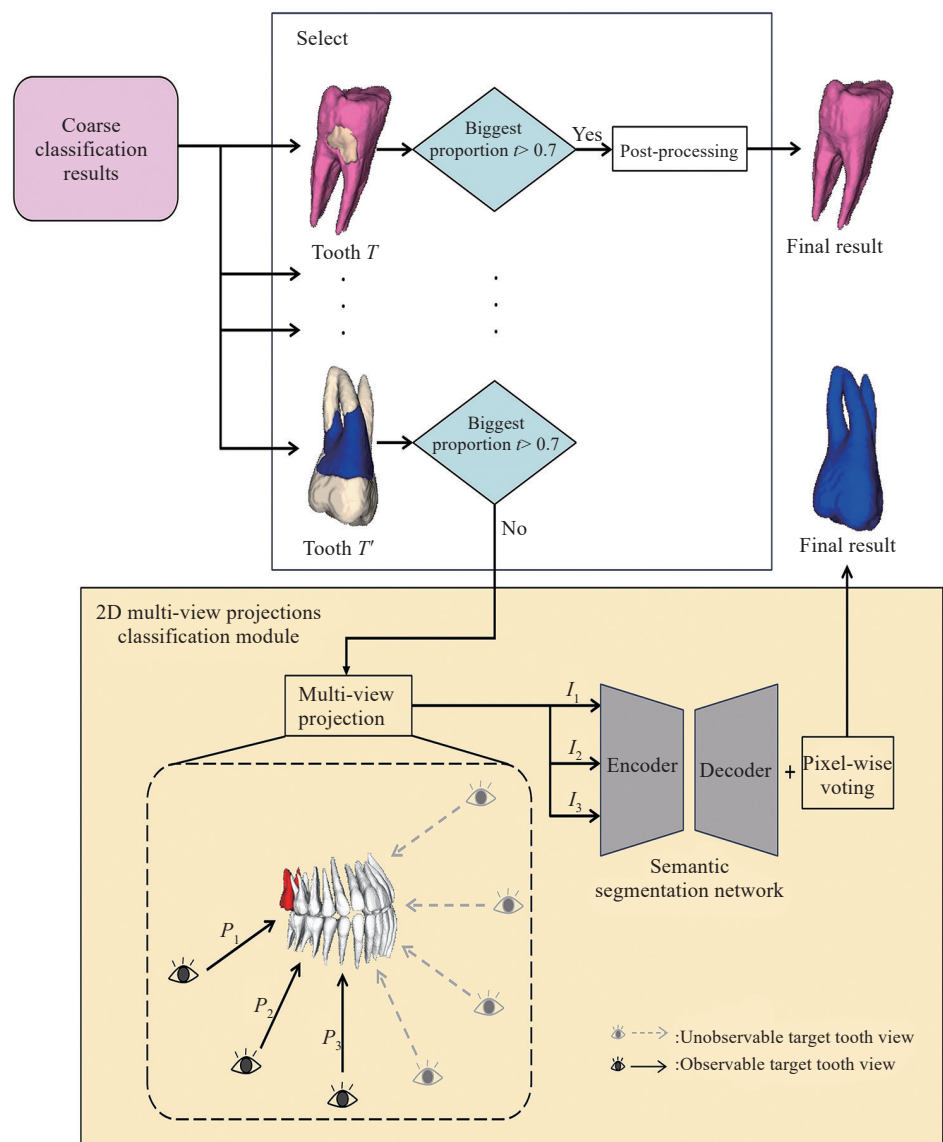


图4 多视角投影的细分类框架

Figure 4 Fine classification framework based on multi-view projection

实验数据来源于上海市第九人民医院,共170例三维牙齿CBCT扫描图像。其中每张CBCT扫描图像对应的标签数据由牙科医生依据FDI牙位标记法,使用ITK-SNAP软件(Version 4.0)进行手工标注,原始扫描数据和对应的标签均以NifTI格式保存,便于后续的处理和分析。在数据集的划分中,按照7:2:1的比例将数据划分为训练集、验证集和测试集。

2.2 数据预处理

在数据预处理过程中,首先对所有的NifTI数据进行统一的重采样,像素间距统一设置为1。接着,将图像的方向统一调整为RAI(即右-前-下坐标系)。随后,笔者找到3D牙齿模型在XYZ轴上的边界,并分别向外扩展:X轴左右各扩展32个体素,Y轴和Z轴左右各扩展16个体素,以裁剪牙齿区域,同时保留适量的上下文信息,加速模型训练过程^[25]。

2.3 模型训练

在对第一阶段网络训练的过程中,直接使用Cui等^[17]提出的HMG-Net框架训练实例分割模型,以获得牙齿的精分割3D模型;为获取初步标签信息,使用Isensee等^[25]提出的nnU-Net框架来训练52类语义分割模型,使用四折交叉验证来评估模型性能。同时考虑到牙齿的对侧相似性,训练时采用不进行镜像增强的训练器“nnUNetTrainerNoMirroring”。

在对第二阶段网络的训练过程中,为了解决数据稀缺的问题,对第一阶段网络的训练数据进行两个操作。首先,将模型绕Z轴旋转,每次的旋转角度为A,沿X轴向Y-Z平面做投影,直到模型旋转至180°。n为旋转后的视角数量,加上0°投影,从一个样本中可以获得n+1个视角的投影数据。每个2D投影图像中的像素值对应于相应的3D图像体素标签。

为了降低模型的训练难度,将像素中 52 类标签转化为不包含象限信息的 13 类标签,同时移除被覆盖率超过 0.5 的背景牙齿。对于每个视角的 2D 投影图像,随机选择 3~6 颗牙齿区域进行标签遮盖。设 $T^{(i)}$ 为第 i 个视角下的图像,遮盖操作可以表示为:

$$T^{(i)} = \text{Mask}(T_i', \{R_1, R_2, \dots, R_m\})$$

(3)

其中, R_1, R_2, \dots, R_m 表示被遮盖的牙齿区域, $m \in [3, 6]$ 是随机选择的牙齿数量。对每个视角进行 5 次遮盖操作,确保每颗牙齿标签至少被遮盖一次。通过上述操作,每个样本生成的 2D 投影图像总数为 $5 \times n$,实现了细分类网络的数据扩充。此外,在输入语义分割网络前对输入图像进行随机旋转 ($\pm 5^\circ$) 和随机裁剪 (大小 545×545 , 裁剪比例为 0.8~1.0, 宽高比为 0.8~1.2), 以增强数据多样性。

2.4 模型测试

为了探究多视角投影对分类结果准确度的影响,进行 4 组实验,对比不做投影时的分类准确度和进行多视角投影的分类准确度。其中分别取 $A=90^\circ$ 、 $A=45^\circ$ 和 $A=30^\circ$, 即保存 $n=3$ 、 $n=5$ 和 $n=7$ 个视角的投影数据训练 2D 语义分割网络。具体来说,对 3 视角模型保存 0° 、 90° 和 180° 的投影数据,对 5 视角模型输入 0° 、 45° 、 90° 、 135° 和 180° 的投影数据,对 7 视角模型输入 0° 、 30° 、 60° 、 90° 、 120° 、 150° 和 180° 的投影数据,其他操作不变。

2.5 实验结果与分析

据表 1 所示,将传统的分类方法如多层感知机 (Multilayer Perceptron, MLP) 和图卷积网络 (Graph Convolutional Network, GCN) 与本文方法做对比 (Our w/o multi-view: 不采用多视角投影模块; Our w/ multi-view: 使用多视角投影模块), 以牙齿分类的准确率为模型表现的评价指标。具体来说,牙齿层面的分类准确率被定义为: 正确预测的牙齿数量占总预测牙齿数量的比例。MLP 与 GCN 方法在成人牙齿数据上的分类准确度较高,说明针对恒牙的分类结果较好;但在儿童牙齿数据上的分类准确度较低,说明仍然难以区分乳牙和未萌出的牙齿。这是因为传统简单的分类器如 MLP, 无法精确提取牙齿的形状信息及位置信息,尤其在儿童数据较少的情况下,使得对乳牙的区分较为困难。本文网络分阶段提取牙齿的特征信息,通过多视角投影结合 Transformer 架构获取全局信息,补充更多的空间位置信息,扩充数据集,从而降低模型训练难度。在儿童数据上的分类准确度相较传统方法有大幅度提升,达到 0.958; 在面对全年龄段牙齿数据时,也能精确区分恒牙与乳牙,分类准确率达到 0.982。

此外,笔者对比了不同视角数量对于实验结果

表 1 不同方法在成人、儿童及全年龄段病例牙齿分类准确率对比
Table 1 Comparison of classification accuracy at the tooth level across different methods for adult, pediatric, and all age cases

方法	成人牙齿	儿童牙齿	全年龄段牙齿
MLP	0.931	0.606	0.778
GCN	0.886	0.624	0.763
Our w/o multi-view	0.988	0.945	0.973
Our w/ multi-view	0.993	0.958	0.982

成人牙齿数据指已过换牙期的成人病例,口内只有恒牙;儿童牙齿数据指未换牙或处于换牙期的儿童病例,口内仍然存在乳牙;全年龄段牙齿数据表示成人牙齿数据与儿童牙齿数据的集合

的影响,采用分类准确率和预测成功率作为评价指标。其中预测成功率的定义为: 正确进行细分类的牙齿数量占有需要进行细分类牙齿数量的比例。如表 2 所示,多视角投影对于分类结果有积极作用。在一定范围内,视角数量越多,模型的分类准确性越高。可见,多视角投影的模型在不同视角数量上训练的准确率是随视角增加而提升的,在实际应用于预测需要在细分类的牙齿场景下,从预测准确率可以显著判断多视角的有效性。整体上观测,该二阶段网络在实验数据上的表现也呈逐步攀升的趋势。具体来说,5 个视角与 7 个视角的模型分类准确度相差不大,5 个视角的模型只需更少的训练数据,更快的训练速度,可以达到与 7 个视角几乎相近的准确度。

表 2 细分类网络中不同视角数量对全年龄段牙齿分类结果的分析
Table 2 Analysis of the effects of different numbers of views in the fine classification network on tooth-level classification results across all age groups

方法	训练集分类准确率	预测成功率	分类准确率
Our w/o muti-view	-	-	0.973
Our w/3 views	0.883	0.600	0.975
Our w/5 views	0.944	0.770	0.980
Our w/7 views	0.949	0.801	0.982

训练集分类准确率: 第二阶段网络的训练集分类准确率; 预测成功率: 第二阶段网络中细分类正确的牙齿颗数占需要细分类的牙齿总数的比例; 分类准确率: 该二阶段网络在整体数据测试集上的分类表现; Our w/o muti-view: 不采用多视角; Our w/3 views: 采用 3 个视角; Our w/5 views: 采用 5 个视角; Our w/7 views: 采用 7 个视角

3 结 论

本研究探讨多视角投影对分类结果的积极影响,提出一种由粗至细的二阶段牙齿分类网络。通过结合多视角投影和 Transformer 架构,提供更丰富的空间信息,扩展数据集,并通过实验确定了最优的

投影视角和数量,实现了对全年龄层CBCT牙齿三维图像的高效精确分类和分割。在牙齿模型的7个视角投影下,本文方法在牙齿层面达到了0.982的分类准确度。牙齿的准确标记和分类,能为后续临床诊治提供更好的支持。

【参考文献】

- [1] Jin LJ, Lamster IB, Greenspan JS, et al. Global burden of oral diseases: emerging concepts, management and interplay with systemic health[J]. *Oral Dis*, 2016, 22(7): 609-619.
- [2] 杨云丹,张疆强,黄瑾,等. CBCT在口腔正畸学的应用[J]. *中国医学物理学杂志*, 2017, 34(1): 48-52.
Yang YD, Zhang JT, Huang J, et al. Applications of cone beam CT in orthodontics[J]. *Chinese Journal of Medical Physics*, 2017, 34(1): 48-52.
- [3] 陶天金,贺楚芸,童庆华,等. 人工智能在口腔影像应用的研究进展[J]. *临床口腔医学杂志*, 2024, 40(9): 570-573.
Tao TJ, He CY, Tong QH, et al. Research progress in the application of artificial intelligence in oral imaging[J]. *Journal of Clinical Stomatology*, 2024, 40(9): 570-573.
- [4] Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis[J]. *Med Image Anal*, 2017, 42: 60-88.
- [5] 洪劲超,袁蕊,谢曜爵. 口腔医学数字技术的应用进展[J]. *当代医药论丛*, 2023, 21(1): 34-37.
Hong JC, Yuan R, Xie YJ. Application progress of digital technology in stomatology[J]. *Contemporary Medicine Forum*, 2023, 21(1): 34-37.
- [6] 胡茂承,陈颖,葛云,等. 基于CBCT的种植牙模型定制医用辅助系统[J]. *中国医学物理学杂志*, 2019, 36(11): 1296-1301.
Hu MC, Chen Y, Ge Y, et al. CBCT-based medical auxiliary system for customization of dental implant model[J]. *Chinese Journal of Medical Physics*, 2019, 36(11): 1296-1301.
- [7] Polizzi A, Quinzi V, Ronsaville V, et al. Tooth automatic segmentation from CBCT images: a systematic review[J]. *Clin Oral Investig*, 2023, 27(7): 3363-3378.
- [8] 李孜政,李淳芃,黄晓峰,等. 基于多视角融合的CBCT图像牙齿实例分割算法[J/OL]. *计算机辅助设计与图形学学报*. (2024-08-21). <http://kns.cnki.net/kcms/detail/11.2925.tp.20240820.1909.032.html>.
Li ZZ, Li CP, Huang XF, et al. Tooth instance segmentation algorithm based on multi-view fusion in CBCT images[J/OL]. *Journal of Computer-Aided Design & Computer Graphics*. (2024-08-21). <http://kns.cnki.net/kcms/detail/11.2925.tp.20240820.1909.032.html>.
- [9] 赵庶旭,罗庆,王小龙. 基于改进Mask R-CNN的牙齿识别与分割[J]. *中国医学物理学杂志*, 2021, 38(10): 1229-1236.
Zhao SX, Luo Q, Wang XL. Teeth recognition and segmentation based on improved Mask R-CNN[J]. *Chinese Journal of Medical Physics*, 2021, 38(10): 1229-1236.
- [10] 胡忆戎,宋保龙,李博,等. 无托槽隐形矫治磨牙远中移动不同设计方式上颌牙列的三维受力分析[J]. *中华口腔医学杂志*, 2023, 58(3): 265-270.
Hu YR, Song BL, Li B, et al. Three-dimensional analysis of maxillary dentition during molar distalization with clear aligners under different movement designs: an *in vitro* experiment[J]. *Chinese Journal of Stomatology*, 2023, 58(3): 265-270.
- [11] 李若莹,武建龙,王碧,等. 无托槽隐形矫治器与金属托槽矫治器正畸治疗效果的对比分析[J]. *中国医学物理学杂志*, 2021, 38(5): 620-624.
Li RX, Wu JL, Wang B, et al. Comparison of therapeutic effects of appliances without brackets versus with metal brackets in orthodontic treatment[J]. *Chinese Journal of Medical Physics*, 2021, 38(5): 620-624.
- [12] 刘洪臣. 人工智能口腔医学[J]. *中华口腔医学杂志*, 2020, 55(12): 915-919.
Liu HC. Artificial intelligence stomatology[J]. *Chinese Journal of Stomatology*, 2020, 55(12): 915-919.
- [13] Tuzoff DV, Tuzova LN, Bornstein MM, et al. Tooth detection and numbering in panoramic radiographs using convolutional neural networks[J]. *Dentomaxillofac Radiol*, 2019, 48(4): 20180051.
- [14] Silva B, Pinheiro L, Oliveira L, et al. A study on tooth segmentation and numbering using end-to-end deep neural networks[C]//2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). Piscataway, NJ, USA: IEEE, 2020: 164-171.
- [15] Jader G, Fontineli J, Ruiz M, et al. Deep instance segmentation of teeth in panoramic X-ray images[C]//2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). Piscataway, NJ, USA: IEEE, 2018: 400-407.
- [16] Chen YL, Du HY, Yun ZQ, et al. Automatic segmentation of individual tooth in dental CBCT images from tooth surface map by a Multi-Task FCN[J]. *IEEE Access*, 2020, 8: 97296-97309.
- [17] Cui ZM, Fang Y, Mei LZ, et al. A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images[J]. *Nat Commun*, 2022, 13(1): 2096.
- [18] Miki Y, Muramatsu C, Hayashi T, et al. Classification of teeth in cone-beam CT using deep convolutional neural network[J]. *Comput Biol Med*, 2017, 80: 24-29.
- [19] Prasoon A, Petersen K, Igel C, et al. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2013. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013: 246-253.
- [20] Xia YD, Yang D, Yu ZD, et al. Uncertainty-aware multi-view co-training for semi-supervised medical image segmentation and domain adaptation[J]. *Med Image Anal*, 2020, 65: 101766.
- [21] Setio AA, Ciompi F, Litjens G, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks[J]. *IEEE Trans Med Imaging*, 2016, 35(5): 1160-1169.
- [22] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 6000-6010.
- [23] Turp JC, Alt KW. Designating teeth: the advantages of the FDI's two-digit system[J]. *Quintessence Int*, 1995, 26(7): 501-504.
- [24] Cui ZM, Li CJ, Wang WP. ToothNet: automatic tooth instance segmentation and identification from cone beam CT images[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 6361-6370.
- [25] Isensee F, Jaeger PF, Kohl SA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. *Nat Methods*, 2021, 18(2): 203-211.
- [26] Xie EZ, Wang WH, Yu ZD, et al. SegFormer: simple and efficient design for semantic segmentation with transformers[C]//Proceedings of the 35th International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2021: 12077-12090.

(编辑:薛泽玲)