

基于改进Mask R-CNN的牙齿识别与分割

赵庶旭, 罗庆, 王小龙

兰州交通大学电子与信息工程学院, 甘肃 兰州 730070

【摘要】针对当前的研究方法在牙齿全景X光片上提取的信息较为单一,而未曾考虑将牙齿的类别信息与形状位置信息融合提取的问题,提出一种实例分割方法同时实现牙齿识别与分割。主要通过融合跳跃结构和SE(Squeeze and Excitation)模块对Mask R-CNN实例分割模型中的分割分支进行改进,并以牙齿功能与FDI牙位两种类别编码方式,采用400张牙齿全景X光片数据进行实验仿真。实验结果表明改进后的模型相比于其他模型,可以同时有效地进行牙齿分类和分割,实现牙齿类别、形状、位置信息的融合提取,改善了Mask R-CNN实例分割模型在分割分支中语义信息提取不足的问题。

【关键词】牙齿识别; 牙齿分割; Mask R-CNN; 跳跃结构; SE模块

【中图分类号】R318; T391

【文献标志码】A

【文章编号】1005-202X(2021)10-1229-08

Teeth recognition and segmentation based on improved Mask R-CNN

ZHAO Shuxu, LUO Qing, WANG Xiaolong

School of Electronics and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

Abstract: In view of the fact that the information extracted from dental panoramic X-ray image by current research methods is relatively fewer, and the problem of without considering the fusion extraction of classification information and shape and position information of teeth, an instance segmentation method is proposed to realize the teeth recognition and segmentation simultaneously. The segmentation branch of Mask R-CNN instance segmentation model is improved by fusing skip-connection structure and SE (Squeeze and Excitation) module. Two teeth coding methods of tooth function and FDI tooth position were used for the experimental simulation on 400 dental panoramic X-ray images. The experimental results show that compared with other models, the improved model can effectively complete teeth classification and segmentation at the same time, realize the fusion extraction of tooth category, shape and position, and overcome the problem of insufficient semantic information extraction in the segmentation branch of Mask R-CNN instance segmentation model.

Keywords: tooth recognition; tooth segmentation; Mask R-CNN; skip-connection structure; SE module

前言

牙病一直是困扰我国居民的一大问题。截止到2018年,我国仅有16.73万个口腔医生,每百万人牙医数量只有100多位,相比于发达国家的每百万人牙医数量高达500~1 000位,中国牙科医生十分紧缺^[1-2]。而在欧美等发达国家,尽管其牙科医生数量多于中国且私人牙科诊所普遍,但是其治疗成本过高。面对全球医疗资源短缺的现状以及实际诊疗的需要,利用计算机图形图像处理技术构建智能诊断系统逐渐成为研究热点^[3]。

牙齿全景X光片是诊断牙病的常用辅助工具,使用计算机技术实现对牙齿全景X光片中牙齿形状、位置、类别等基础信息的提取能够大量减少医生的工作量。在目前的牙科X线图像自动化分析研究中^[4],主要是针对牙齿、病灶等目标的分割与识别^[5-8]。如Hasan等^[9]利用梯度矢量流(Gradient Vector Flow, GVF)Snake从全景牙科X线图像中自动分割颌骨。Choi等^[10]采用变分法和卷积神经网络结合的方法,检测牙周受损牙。Patil等^[11]提出利用主成分分析(Principal Component Analysis, PCA)降维和神经网络的龋齿检测方法,用于牙科X射线图像的分析。其中Jader等^[6]提出使用Mask R-CNN用于分割牙齿,详细讨论了牙齿X光片中完整与缺牙情况的检测分割效果,但是其分割精度不高且将所有的牙齿都归为一类,忽略了不同牙齿(如门牙与后槽牙)之间的语义差别,无法做到各个牙位牙齿之间的区分。但

【收稿日期】2021-05-10

【基金项目】国家自然科学基金(6206020135)。

【作者简介】赵庶旭, 博士后, 教授, 研究方向: 计算机视觉、边缘计算,
E-mail: zhaosx@lztu.edu.cn

在实际诊断过程中,对牙齿进行编号是牙病诊断过程中的必要步骤,口腔医生往往会根据病历上牙齿部位的记录来快速了解牙齿情况,同时也方便医生为之后的诊疗工作做病历记录。Chen等^[12]首次提出使用Faster R-CNN来检测和编号牙科图像中的牙齿,为牙齿全景X光片的分析提供了新的方向,但是其只提取了牙齿的位置与类别信息,忽略了牙齿形状及语义信息的重要性。

考虑到这些牙齿分割和牙齿识别分类的深度学习方法只能单一的提取牙齿的语义信息或者类别信息,当需要更全面的牙齿信息时,只能通过多次提取,这在实际应用过程中需要耗费大量的时间与精力,缺乏高效性。而在神经网络学习过程中对于牙齿的分割和识别有大量特征参数可以共享,将这

两个任务融合,可以实现对牙齿信息的融合提取。因此本文利用改进的Mask R-CNN同时完成对牙齿全景X光片的牙齿分割与牙齿分类,并通过实验进行验证,充分证明Mask R-CNN应用于牙齿分割和牙齿检测的有效性。

1 Mask R-CNN 算法原理

Mask R-CNN是由He等^[13]提出的一种实例分割架构。实例分割是指从图像中用目标检测方法框出不同实例,再用语义分割方法在不同实例区域内进行逐像素标记。Mask R-CNN在目标检测网络Faster R-CNN^[14]基础上增加了一条分割分支使其能对每个检测出的目标进行语义分割,即实现所谓的实例分割。其网络结构图如图1所示。

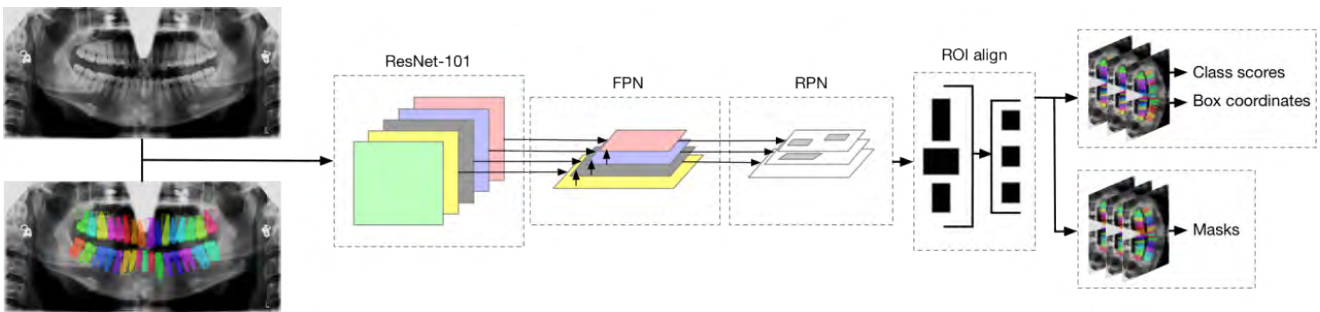


图1 Mask R-CNN网络结构图
Fig.1 Mask R-CNN network architecture

Mask R-CNN 的网络结构主要分为以下4个阶段。阶段1:残差网络(ResNet)和特征金字塔网络(Feature Pyramid Networks, FPN)作为特征提取器,对输入的图像进行特征提取。阶段2:得到的特征图通过区域生成网络(Region Proposal Network, RPN)提取可能存在的目标区域(Region of Interest, ROI)。阶段3:这些 ROI 输入到ROI Align,并通过双线性插值的方式被映射成固定维数的特征向量。阶段4:映射后的特征分别输入3条分支,通过全连接层进行分类和包围框回归,通过全卷积层进行语义分割。

1.1 RPN

RPN^[14]输入特征提取网络的特征图,输出目标候选区域矩形框集合。RPN预先在原图上生成若干个边界框,并判断其中最有可能包含目标的边界框作为候选区域输出。其原理如图2所示,核心为锚点(anchor)机制。首先用3×3的滑窗在FPN输出的特征图上进行遍历,当前位置的滑窗中心在原始图像像素空间的映射点即为锚点,以此锚点为中心,设定不同尺寸(128×128、256×256、512×512)和长宽比例(1:1、1:2、2:1)的锚点框共9种,根据已知的锚点位

置和锚点框的尺寸,便可以得到原始图像中对应区域的尺寸及坐标,这个区域就是预设的候选框。将卷积得到的256维特征向量经过两个全连接层(cls layer、reg layer)分支进行分类和边框回归,再利用非极大值抑制得到最有可能包含目标的区域。

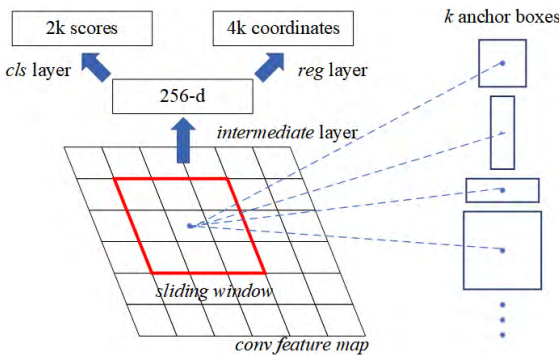


图2 RPN原理图
Fig.2 Schematic diagram of region proposal network

1.2 ROI Align

在Faster R-CNN中ROI Pooling两次量化会造成

区域不匹配问题,为此Mask R-CNN提出了RoI Align的方法来取代ROI Pooling^[13]。ROI Align取消了量化操作,使用双线性内插的方法获得坐标为浮点数的像素点上的图像数值,从而将整个特征聚集过程转化为一个连续的操作。其原理如图3所示,图中虚线框为5×5的特征图,实线框为映射到特征图上的ROI区域。要对该ROI区域做2×2的池化操作,首先把该ROI区域划分4个2×2的区域,然后在每个小区域中选择4个采样点和距离该采样点最近的4个特征点的像素值(图中黑色小方格的4个角点1、2、3、4);通过双线性插值的方法得到每个采样点的像素值;最后计算每个小区的池化值,生成ROI区域的2×2大小的特征图。

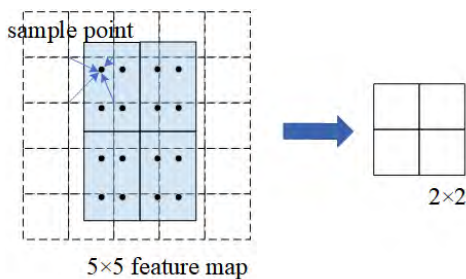


图3 ROI Align原理图
Fig.3 Schematic diagram of ROI Align

1.3 Mask分支

对于Mask分支,是Mask R-CNN在Faster R-CNN基础上增加的分支用于生成检测目标的掩码^[13]。Mask分支的输入来自于经过ROI Align处理后的ROI。对于每一个ROI,Mask分支有K个 $m \times m$ 维度的输出,对这些大小为 $m \times m$ 的Mask进行编码,得到该ROI K个类别的概率值,由此实现实例分割的目的。如图4所示,Mask分支对其进行4次卷积和1次反卷积操作,并使用了像素级sigmoid激活函数,最后得到 $K \times m \times m$ 维度的输出,其中K为检测目标的类别数目,m为特征图的尺寸。由于Mask分支根据分类分支所预测的类别标签来选择输出的掩码,其允许网络为每一类生成一个掩码,不同类之间不存在竞争,这使得分类和掩码生成分解开来,可以提高实例分割的效果。

然而Mask R-CNN作为两阶段检测模型,实行先检测后分割策略,其分割效果受检测结果的制约。对于RPN提取出的ROI特征图,Mask分支采用全卷积操作提取语义信息,这对于局部语义信息有较好的敏感度,但是忽略了上下文信息。

2 基于改进Mask R-CNN的牙齿识别与分割

在医疗图像的临床使用中,既需要整副图的全

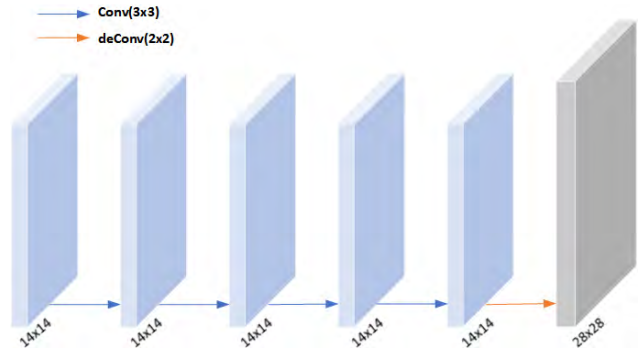


图4 Mask分支
Fig.4 Mask branch

局信息也需要某些特点区域的局部信息,对医学图像各个尺度特征信息的融合,可以一定程度上增大信息量。U-net模型在医学图像分割上取得了优越的效果^[15],其通过跳跃连接更好的融合了上下文语义信息。由此提出利用跳跃连接结构融入多尺度注意力信息对Mask R-CNN分割分支进行改进,弥补Mask R-CNN在掩码分支上所缺乏的深层特征。

2.1 跳跃连接

跳跃连接(Skip-connection)最早由Ronneberger等^[16]提出,用于语义分割的全卷积网络(Fully Convolution Network, FCN)中。之后Shelhamer等^[15]在跳跃连接基础上,提出了用于医学图像语义分割的U-net架构。FCN和U-net架构的不同是,FCN使用求和运算进行特征融合,而U-net将特征进行拼接。图5为密集跳跃连接结构图。

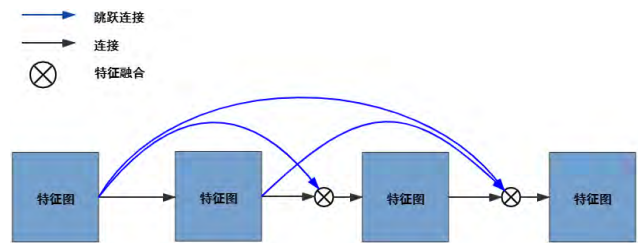


图5 跳跃连接结构
Fig.5 Skip-connection structure

文中跳跃连接的特征融合方式是特征图在通道维度上的拼接,计算公式如式(1)所示:

$$\begin{cases} F(h,w,c) = W(h,w,a) + V(h,w,b) \\ c = a + b \end{cases} \quad (1)$$

其中, $W(h,w,a)$ 和 $V(h,w,b)$ 分别来自不同层的特征图, $F(h,w,c)$ 为拼接之后的特征图, h 和 w 为特征图的长宽, a 、 b 及 c 均为特征图的通道数。这种跳跃连接结构结合了低级特征图中的特征,避免了直接在高级特征图上进行学习,使得最终得到的特征图

既包含了高层特征,又包含了很多低层特征,实现了多尺度下特征的融合。

2.2 SE 模块

虽然跳跃连接更好的融合了上下文语义信息,有效地提取出更多牙齿细节信息,但**低层特征中亮度不均衡、对比度较低仍会对牙齿的细粒度分割造成干扰**。通过引入**注意力机制 SE (Squeeze and Excitation)**模块^[17]来捕获高级语义信息,根据特征图像的值对各特征通道进行加权,提升重要特征的权重,降低不重要特征的权重,从而提升特征提取的效果,提高模型的分割精度。SE 模块的核心是压缩 (Squeeze) 和激励 (Excitation),结构示意图如图 6 所示。

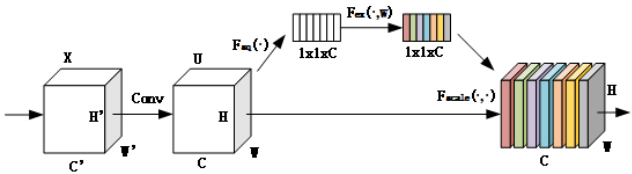


图6 SE 模块
Fig.6 Squeeze and excitation module

在图6中,首先特征 X 经过卷积将其通道数由 C' 变为 C ,将特征图 U 传递给压缩操作。压缩操作使用全局平均池化将每个特征通道都压缩成一个实数,将感受野扩展到全局范围。压缩计算过程如式(2)所示:

$$z = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i,j) \tag{2}$$

其中, u_c 为通过卷积后得到的特征图, c 为 U 的通道数, $H \times W$ 为 U 的空间维度。接着,激励操作捕获压缩后的实数列信息,使用两个全连接层增加模块的非线性。先经过第一个全连接层降维,再通过整流线性单元 ReLU 激活,然后经过第二个全连接层升维,最后经过 sigmoid 激活函数,整个过程如下:

$$s = \sigma \left[W_2 \delta \left(W_1 z \right) \right] \tag{3}$$

其中, δ 为非线性激活函数 ReLU, W_1 和 W_2 分别为两个全连接层的参数, σ 为 sigmoid 函数。最后对原特征进行加权,用原特征逐通道乘以激励操作获得的通道重要度系数,得到具有注意力信息的特征:

$$\tilde{x}_k = s_k \bullet u_k, \quad k = 1, 2, \dots, C \tag{4}$$

2.3 网络结构

改进的 Mask 分支利用卷积层进行下采样编码,反卷积层进行上采样。其中上采样的计算过程如式(5)所示:

$$y_{ij} = \sum_{a,\beta=0}^k \left| 1 - \alpha - \left\{ i/f \right\} \right| \left| 1 - \beta - \left\{ j/f \right\} \right| x_{\left\lfloor i/f \right\rfloor + a, \left\lfloor j/f \right\rfloor + \beta} \tag{5}$$

其中, k 为反卷积核大小, f 为上采样因子即为步长。在上采样的同时通过跳跃连接与 SE 模块将不同尺度的浅层高分辨率特征输入到反卷积层。改进的 Mask 分支网络结构如图 7 所示。

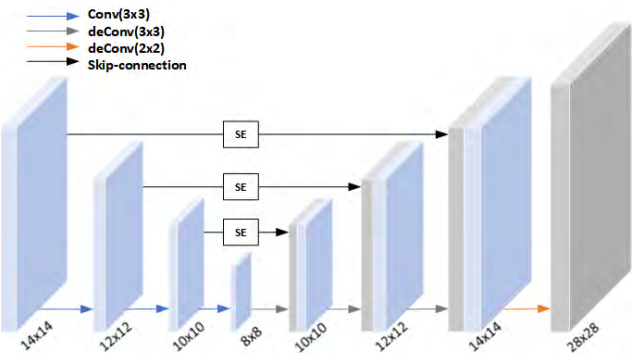


图7 改进后的Mask分支
Fig.7 Improved Mask branch

其中1、2、3层均为卷积层,卷积核大小为3,步长为1。每次卷积之后都跟随了批归一化(BN)层以及ReLU激活函数。4、5、6、7层为反卷积层,其中4、5、6层的卷积核大小为3,步长为1,第7层的卷积核大小为2,步长为2。输入14×14的特征图,经过3个卷积层得到8×8的特征图,经过反卷积层得到28×28的特征图。对于反卷积层的输入,编码器网络和解码器网络对称的层提供跳跃连接,将编码器网络每一层卷积操作的结果经过SE模块之后和解码器网络上采样的结果进行拼接。最后通过sigmoid层生成二进制的分割掩码。这样充分利用了不同尺度特征图所包含的信息,提高特征利用率,使得分割分支在更大的感受野上获取更丰富的细节特征,提高了对目标的细粒度分割效果。

3 实验内容

3.1 实验数据

实验数据使用了包含400张牙齿全景X光片的数据集,该数据来源于临床诊断,并经过脱敏处理,数据的使用获得患者同意。其中每张图片的尺寸为1 024×2 161,其标签数据由牙科医生利用VIA工具所标注存于CSV文件中,经过数据清洗整理后笔者将其格式转换为Mask R-CNN模型需要的JSON格式标签数据,该JSON文件中包含各颗牙齿的轮廓坐标、世界口腔联盟(Federation Dentaire Internationale, FDI)牙位编号等信息。

3.2 损失函数设计

本文采用多任务加权损失函数,通过学习不断减小损失函数的值,直到获得全局最优解。损失函

数的公式如式(6)所示:

$$L = \alpha \cdot L_{\text{cls}} + \beta \cdot L_{\text{box}} + \chi \cdot L_{\text{mask}}$$

(6)

其中, L_{cls} 为分类误差, L_{box} 为包围框误差, L_{mask} 为分割误差。原始的损失函数中, 分类、包围框回归及分割这3条分支并行处理, 具有相同的权重^[13]。然而不同的任务具有不同的收敛速度, 分割任务属于语义级别的分类, 其学习难度远高于目标级别的分类。因此采用加权损失函数, 平衡各项任务的学习。

3.3 类别编码设计

目前的牙齿分割方法中将所有牙齿划为一类, 这意味着不同牙齿之间也会相互学习, 从而忽略了它们之间不同的形态特征^[5, 18-19]。例如, 切牙与尖牙的牙冠呈楔形, 而前磨牙与磨牙的牙冠呈立方形; 上颌磨牙牙根为3根, 下颌磨牙为双根, 而其他牙齿多为单根。

为区分不同牙齿, 本文提出了两种牙齿类别编码方法, 其表示方法分别如式(7)、(8)所示:

$$V_k = \begin{cases} 1, & j = 1, 2 \\ 2, & j = 3 \\ 3, & j = 4, 5 \\ 4, & j = 6, 7, 8 \end{cases}$$

(7)

$$V_{ij} = (i - 1) \times 8 + j$$

(8)

编码算法的伪代码如下所示。

Algorithm Teeth Coding

Input: label of tooth

1: **for** each label **do**:

2: Compute $i = \text{label} / 10, j = \text{label} \% 10$

3: Compute $V_{ij} = (i - 1) * 8 + j$

4: **if** ($j = 1$ or 2):

5: $V_k = 1$;

6: **else if** ($j = 3$):

7: $V_k = 2$;

8: **else if** ($j = 4$ or 5):

9: $V_k = 3$;

10: **else** :

11: $V_k = 4$;

Output: V_{ij}, V_k

其中, V_{ij} 依据 FDI 牙位表示法^[20]进行编号, V_k 依据牙齿的形态特点和功能特性进行编号^[21], i 代表牙齿所在的象限, j 代表牙齿的位置, 具体分布如图8所示。

3.4 实验设置

实验超参数设置如下: 学习率为0.001, 批处理大小为100, 迭代次数为35。实验中80%的数据被用于训练集, 剩下20%用于测试。考虑到牙齿分割是像素级的分类任务, 其对数据量的要求比牙位识别高。由于数据量较少, 我们无法从头训练整个深度学习

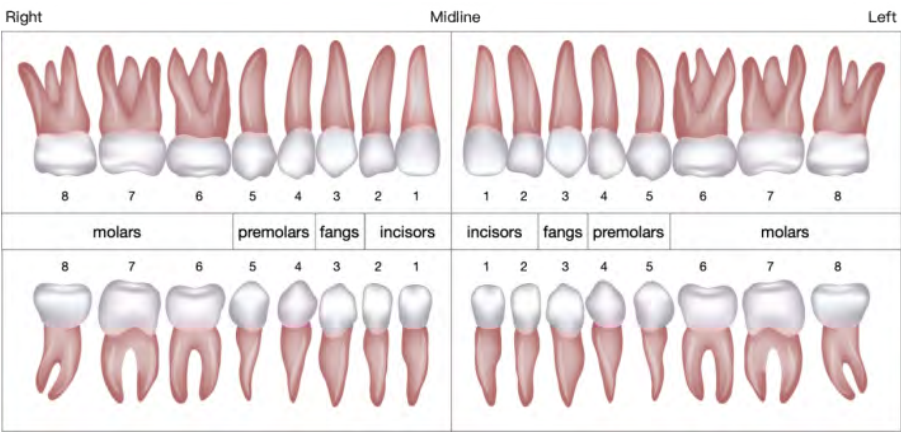


图8 牙齿分布示意图

Fig.8 Tooth number indicates

网络, 对此我们采用迁移学习的方法来提升网络的训练效果。迁移学习是指将某个领域或任务上学习到的知识或模式应用到不同但相关的领域或问题中^[22]。在我们所使用的模型中, 主干网络特征提取部分采用迁移学习的思想导入了在MS COCO数据集^[23]上的预训练权重, 并运用已有的数据对网络头部进行微调。之后分别采用公式(7)、(8)两种类别表示方法 V_k 和 V_{ij} 进行实验, 并记为实验一(Mask R-CNN-1)、实验二(Mask R-CNN-2)。

4 实验结果与分析

4.1 评价指标

为评估实验结果, 我们采用了交并比 (Intersection Over Union, IOU) 阈值为0.50时的精度值和召回率^[24], 即IOU值大于0.5时为正样本, 反之为负样本。其中精准率 (precision) 反映了查准率, 它表示的是预测为正的样本中有多少是真正的正样本, 而召回率 (recall) 反映了查全率, 它表示样本中的正例有多少被预测正确, 计

算公式如式(9)、(10)所示:

$$\text{precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})}$$

(9)

$$\text{recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})}$$

(10)

其中, TP、FN 和 FP 分别表示真阳性、假阴性和假阳

性。真阳性是指预测为正,实际为正;假阴性是指预测为负,实际为正;假阳性是指预测为正,实际为负。

4.2 结果与分析

图9为实例分割结果,可以看出牙齿的位置、类别及语义信息都被提取。为方便与其他方法相对比,文中将分割结果与分类结果分开进行对比讨论。

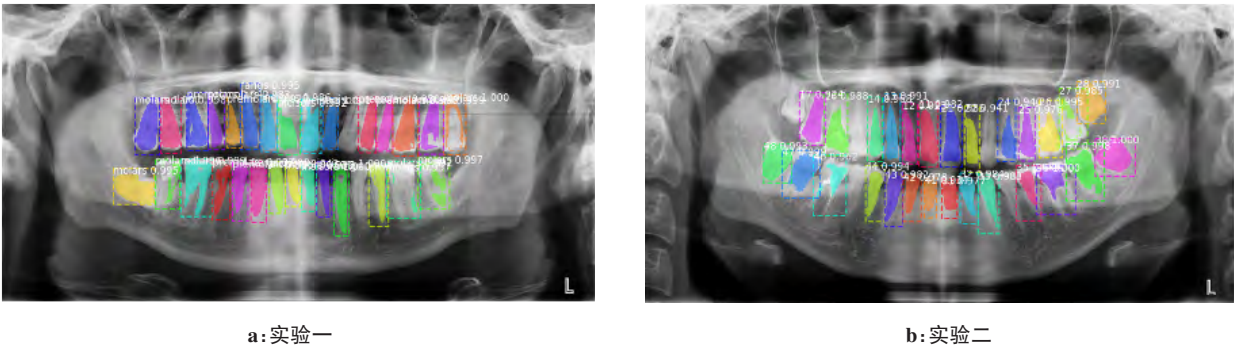


图9 实例分割结果
Fig.9 Results of instance segmentation

4.2.1 分割结果 两组类别编码方式的实验结果与U-net^[18]和传统图像处理方法及原始Mask R-CNN^[6]方法结果对比如表1所示,其中区域增长^[5]为传统图像处理方法的分割结果;Mask R-CNN-1与Mask R-CNN-2分别为实验一与实验二的分割结果。可以看出深度学习方法相比于传统图像处理方法有着明显优势,在精准率上传统Mask R-CNN表现最优,在召回率上Mask R-CNN-1表现最优。相比于U-net,Mask R-CNN-1在精准率和召回率上提升了4%和1%,Mask R-CNN-2在精准率上提升了2%。然而在Mask R-CNN-1与Mask R-CNN-2的对比中,Mask R-CNN-2的精准率和召回率都远低于Mask R-CNN-1。

表1 牙齿分割结果对比
Tab.1 Comparison of teeth segmentation results

方法	精准率	召回率
区域增长	0.355	0.634
U-net	0.920	0.930
Mask R-CNN	0.980	0.840
Mask R-CNN-1	0.968	0.939
Mask R-CNN-2	0.943	0.844

图10为实验一与实验二的分割结果。在Mask R-CNN-1中,每颗牙齿都生成了掩膜,但是存在掩膜覆盖不完整的情况,如图中的33、34号牙。在Mask R-CNN-2中,存在较多的牙齿并没有生成掩膜,如18、34、38号牙。造成这种结果的原因是Mask R-CNN-1

将牙齿分为4类,同1类下的不同实例之间相互贡献损失,而Mask R-CNN-2将牙齿分为32类,每1类下只有1个实例,无法相互贡献损失,因此实验二的分割效果不如实验一。

4.2.2 分类结果 两组实验结果与Faster R-CNN^[12]方法的结果对比如表2所示,其中Mask R-CNN-1与Mask R-CNN-2分别为本文实验一与实验二的分类结果。表中显示Mask R-CNN-1表现最优,能准确识别出牙齿并根据功能分类,但是其只做了四分类,无法做到牙位识别的效果。Faster R-CNN与Mask R-CNN-2都是对牙齿做32分类,能够有效识别出不同的牙位,但Mask R-CNN-2精度相比于Faster R-CNN有所提升。另外,在Mask R-CNN-1与Mask R-CNN-2的对比中,Mask R-CNN-2的精准率和召回率都远低于Mask R-CNN-1。

图11为实验一与实验二的目标检测结果,表3对应为图11中目标检测包围框上方的文字即目标分类标签及其分数值。从图11和表3可以看出在Mask R-CNN-1中的查全率较高,每颗牙齿都被检测且都正确分类,识别分数均值能达到0.93以上。Mask R-CNN-2中存在没有检测到的目标,如图中14、24、27、34牙位的牙齿并没有包围框;且检测到的目标分数比Mask R-CNN-1要低,均值在0.95左右。

综合两种实验方案,实验一在牙齿分割任务中表现优于实验二,但是其牙齿分类程度没有实验二精细,无法做到牙位识别的效果。实验二能较好的完成牙位识别任务,却也因其各个牙位之间的分割损失值不能

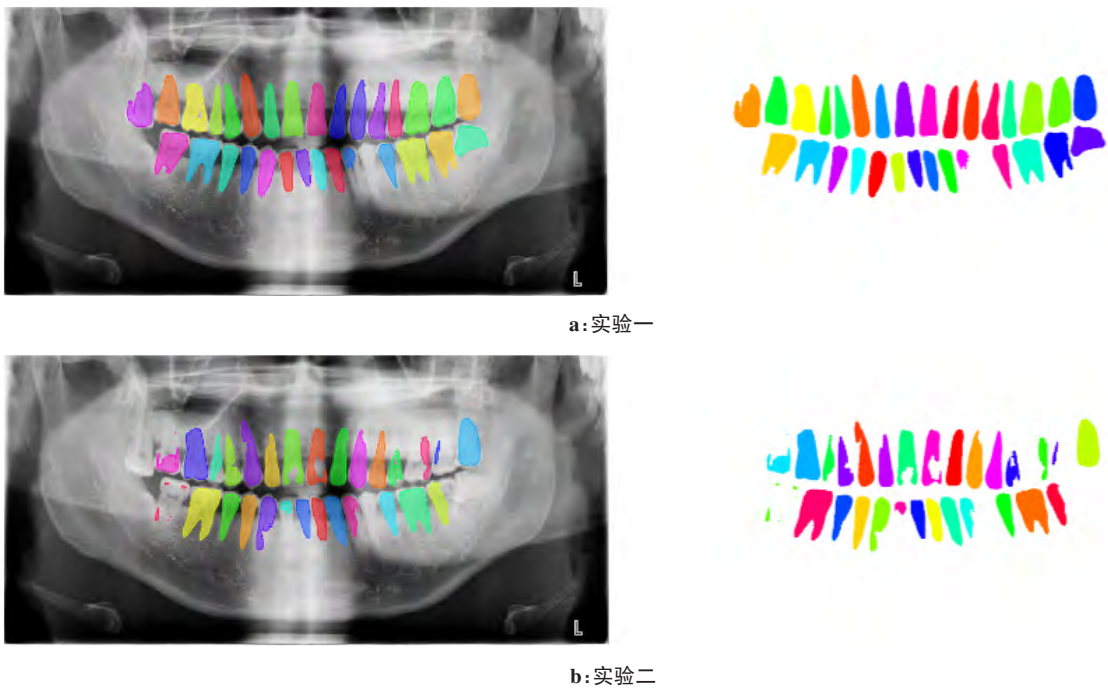


图 10 分割结果

Fig.10 Segmentation results

表 2 牙齿分类结果对比

Tab.2 Comparison of teeth classification results

方法	精准率	召回率
Faster R-CNN	0.920	0.910
Mask R-CNN-1	0.968	0.939
Mask R-CNN-2	0.943	0.844

共享而导致分割效果不好.但是两组实验充分证明了 Mask R-CNN 具备同时实现牙齿分割和牙齿分类的能力,且能保证在牙齿分类与分割的精度达到90%。在过去的研究中^[19],我们针对单颗牙齿应用灰度值统计方法进行纹理分析,可以在一定程度上反映牙齿正常与否,与本文方法相结合,可以为单颗牙齿的病理分析提

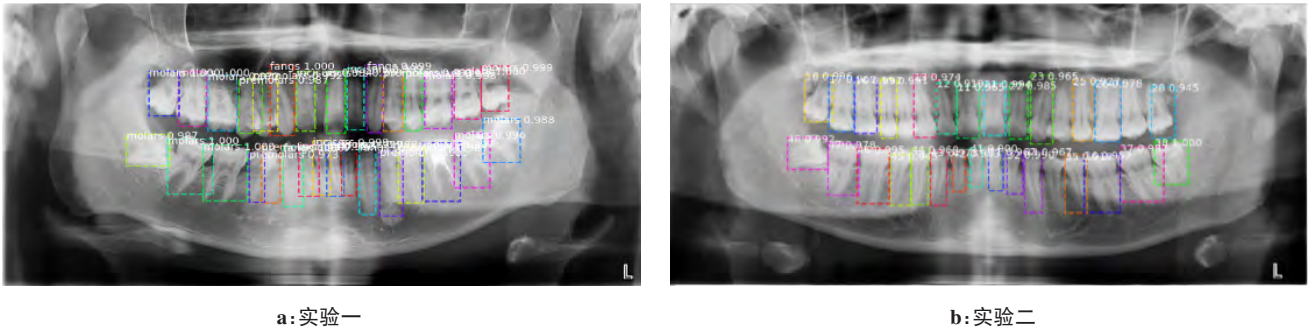


图 11 分类结果
Fig.11 Classification results

供一定的基础,且对于各颗牙齿之后的龋齿、牙周病、根尖炎的疾病诊断分析有重要支撑作用。但是实验中使用的预训练模型是基于 MS COCO 数据集,与文中的牙齿 X 光片数据集相似度较低,在未来的研究中可以考虑采用其他的 X 光片数据集,这可以在一定程度上降低微调的难度,得到更好的实验结果。另外,实际诊疗过程中的牙齿 X 光片往往存在缺牙、残根及种植体等非正常情况,这对于本文的模型来说是一个挑战。在之后的工作中可以考虑加入应用牙齿排列的规则等其他后

处理方法,有针对性的进行分析。

5 结束语

本文基于 Mask R-CNN 模型对牙齿 X 射线全景图中的牙齿进行分割和分类研究,以求较为准确的描绘出 X 射线全景图上的每颗牙齿的形态与位置。从牙齿功能和 FDI 牙位表示两种编码方式的实验表明,使用改进的 Mask R-CNN 能同时进行牙齿分割和牙齿分类任务,并在这两项任务中均达到 90% 以上

表3 检测结果对比

Tab.3 Comparison of detection results

FDI牙位编号	检测类别 (实验一)	检测分值 (实验一)	检测类别 (实验二)	检测分值 (实验二)
11	未检出	--	11	0.965
12	切牙	0.992	12	0.919
13	尖牙	1.000	13	0.974
14	前磨牙	0.992	未检出	--
15	前磨牙	0.987	15	0.991
16	后磨牙	0.972	16	0.997
17	后磨牙	1.000	17	0.969
18	后磨牙	1.000	18	0.996
21	切牙	0.998	21	0.994
22	切牙	0.999	22	0.985
23	尖牙	0.999	23	0.965
24	前磨牙	0.998	未检出	--
25	前磨牙	0.999	25	0.927
26	后磨牙	0.998	26	0.978
27	后磨牙	1.000	未检出	--
28	后磨牙	0.999	28	0.945
31	切牙	0.982	31	0.962
32	切牙	0.999	32	0.990
33	尖牙	0.986	33	0.967
34	前磨牙	0.955	未检出	--
35	前磨牙	0.987	35	0.913
36	后磨牙	0.997	36	0.952
37	后磨牙	0.996	37	0.987
38	后磨牙	0.988	38	1.000
41	切牙	0.937	41	0.900
42	切牙	0.998	42	0.983
43	尖牙	0.964	43	0.976
44	前磨牙	0.952	44	0.968
45	前磨牙	0.973	45	0.973
46	后磨牙	1.000	46	1.000
47	后磨牙	1.000	47	1.000
48	后磨牙	0.987	48	0.992

的精确率,实现了牙齿语义、位置及类别信息的融合提取。但是在语义信息提取上仍然存在一定的提升空间,该问题可能是由于数据集中影像数量较少造成。在将来的研究中,我们考虑进一步提高实验精准度,并基于实验结果进行智能分析系统研究。

【参考文献】

[1] 王兴.我国口腔医疗行业的现状、机遇与挑战[J].科技与金融,2018(5): 10-13.
WANG X. The present situation, opportunities and challenges of Stomatological medical industry in China[J]. STF Monthly, 2018(5): 10-13.

[2] 冯希平.中国居民口腔健康状况—第四次中国口腔健康流行病学调查报告[C]//2018年中华口腔医学会第十八次口腔预防医学学

术年会,2018: 13.

FENG X P. Residents of oral health, China-the fourth Chinese oral health epidemiological survey report [C]//2018 The 18th annual meeting of The Chinese Dental Association on Oral Preventive Medicine, 2018: 13.

[3] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. Med Image Anal, 2017, 42(9): 60-88.

[4] RAZZAK M I, NAZ S, ZAIB A. Deep learning for medical image processing: overview, challenges and future [J]. arXiv org, 2017 Apr 22.

[5] SILVA G, OLIVEIRA L, PITHON M. Automatic segmenting teeth in X-ray images: trends, a novel data set, benchmarking and future perspectives[J]. Expert Syst Appl, 2018, 107: 15-31.

[6] JADER G, FONTINELE 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. IEEE, 2018: 400-407.

[7] ØLBERG J V, GOODWIN M. Automated dental identification with lowest cost path-based teeth and jaw separation[J]. Scand J Caring Sci, 2016, 2(2): 44-56.

[8] KEERTHANA K M, RAJESHWARI B, KEERTHI S, et al. Classification of tooth type from dental X-ray image using projection profile analysis [C]//2017 International Conference on Signal Processing and Communication (ICSPC). IEEE, 2017: 394-398 .

[9] HASAN M M, ISMAIL W, HASSAN R, et al. Automatic segmentation of jaw from panoramic dental X-ray images using GVF snakes [C]//2016 World Automation Congress (WAC). IEEE, 2016: 1-6.

[10] CHOI J Y, EUN H, KIM C, et al. Boosting proximal dental caries detection via combination of variational methods and convolutional neural network[J]. J Signal Process Sys, 2018, 90(12): 87-97.

[11] PATIL S, KULKARNI V, BHISE A. Caries detection using multidimensional projection and neural network [J]. International Journal of Knowledge-Based and Intelligent Engineering Systems, 2018, 22(3): 155-166.

[12] CHEN H, ZHANG K L, LYU P J, et al. A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films[J]. Sci Rep, 2019, 9: 1-11.

[13] HE K, GEORGIA G, PIOTR D, et al. Mask R-CNN [J] . IEEE T Pattern Anal, 2020, 42(2): 386-397.

[14] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Trans Pattern Anal, 2017, 39(6): 1137-1149.

[15] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Trans Pattern Anal, 2017, 39(4): 640-651.

[16] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015: 234-241.

[17] HU J, SHEN L, SUN G. Squeeze and excitation networks[C]. CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA, June 18-23, 2018.

[18] KOCH T L, PERSLEV M, IGEL C, et al. Accurate segmentation of dental panoramic radiographs with U-Nets [C]//IEEE 16th International Symposium on Biomedical Imaging (ISBI). IEEE, 2019: 15-19.

[19] ZHAO S X, LIU C R, LUO Q. Dental data analysis based on dental X-ray panorama [C]//ISICDM 2019: Proceedings of the Third International Symposium on Image Computing and Digital Medicine, 2019: 133-137.

[20] 韩永战,任菊香.建议统一记录牙齿的部位和名称[J].牙体牙髓牙周病学杂志, 2012, 22(8): 488.
HAN Y Z, REN J X. Suggested to record the location and name of teeth uniformly [J]. Chinese Journal of Conservative Dentistry, 2012, 22(8): 488.

[21] 陈世贤.法医学概论[M].北京:北京大学出版社, 1992: 17-19.
CHEN S X. Introduction to dentistry[M]. Beijing: Peking University Press, 1992: 17-19.

[22] 庄福振,罗平,何清,等.迁移学习研究进展[J].软件学报, 2015, 26(1): 26-39.
ZHUANG F Z, LUO P, HE Q, et al. Research progress of transfer learning[J]. Journal of Software, 2015, 26(1): 26-39

[23] VINYALS O, TOSHEV A L, BENGIO S, et al. Show and tell: lessons learned from the 2015 MSCOCO image captioning challenge [J]. arXiv.org. 10.1109/TPAMI.2016.2587640

[24] 周飞燕,金林鹏,董军.卷积神经网络研究综述[J].计算机学报, 2017, 40(6): 1229-1251.
ZHOU F Y, JIN L P, DONG J. Review of convolutional neural network [J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251.

(编辑:薛泽玲)