

基于改进 U-net 的牙齿锥形束 CT 图像重建研究

刘昊鑫^{1,2,3,4}, 赵源萌^{1,2,3,4*}, 张存林^{1,2,3,4}, 朱凤霞^{1,2,3,4}, 杨墨轩^{1,2,3,4}¹首都师范大学物理系, 北京 100048²太赫兹光电子学教育部重点实验室, 北京 100048;³太赫兹波谱与成像北京市重点实验室, 北京 100048;⁴北京成像理论与技术高精尖创新中心, 北京 100048

摘要 锥形束计算机断层扫描(CT)成像技术在口腔疾病诊断中发挥了重要作用。如何从复杂的原始扫描图像中获取牙齿的准确信息,成为口腔医学的一个重要研究问题。引入空间注意力机制,提出一种基于改进 U-net 网络的分割算法,结合等值面提取算法,实现对牙齿锥形束计算机断层扫描图像的准确分割和三维重建。首先对图像进行分割,得到只保留牙齿信息的图像,再对结果进行三维重建,创建出牙齿的三维模型。实验结果表明,该方法能够有效地提取牙齿信息,有助于对口腔疾病特别是牙齿疾病的诊断和治疗。

关键词 成像系统; 锥形束计算机断层扫描; 神经网络; 三维重建; 空间注意力机制

中图分类号 T391

文献标志码 A

DOI: 10.3788/CJL202249.2407207

1 引言

近年来随着信息技术和生物医学工程突飞猛进的发展,锥形束计算机断层成像(CBCT)技术被越来越多地应用于牙齿疾病的诊断与分析^[1]。口腔 CBCT 原始图像中除了牙齿以外还包含其他人体组织(如牙颌骨等人体骨骼组织),会影响医生对病人病情的判断,另外,采集图像时引入的噪声也会降低 CBCT 图像应用效果,故而原始 CBCT 数据不经处理难以直接使用^[2]。为获取病人口腔真实状况,口腔医生常用人工手段对 CBCT 图像进行处理,即通过手工操作分割出牙齿的图像,这样做可以完成牙齿图像的分割提取,但工作量大、效率低、人力成本高。本文结合图像分割及三维重建算法,从口腔 CBCT 图像中自动化、智能化地获取牙齿的三维模型,能帮助医生高效地利用 CBCT 技术完成诊断和治疗。

在以往的研究中,为从复杂的口腔 CBCT 图像中获取准确的牙齿图像,通常采用两种方法:基于 Snakes 算法的水平集算法^[3]和基于神经网络的算法。王立新等^[4]提出加入了几何活动轮廓模型的基于水平集算法的区域自适应算法,有力地提高了对牙根区域的处理效果。Miki 等^[5]提出用神经网络对牙齿类型进行标记,再从原始图像手动修剪每颗牙齿的图像,并将修剪后的二维图像用于学习。崔志明(Cui)等^[6]将 Mask R-CNN 网络和 Girdhar 等^[7]提出的 3D RPN 网络结合起来,提出了一种组合神经网络 ToothNet,实现了

口腔 CBCT 图像的牙齿分割。这种网络把 CBCT 图像当做一个整体的三维数据集,有效分割整个数据集,但将图像视作一个整体容易受到牙齿填充物的严重干扰。

当前牙齿图像分割面临两个核心问题。首先是人类口腔的环境复杂,各种人体组织特别是牙颌骨会严重影响算法对牙齿的分割,这个问题在牙齿根部尤为明显。其次是人体的牙齿填充物会对算法的运行产生严重干扰。因此,相比于其他分割算法,应用于牙齿图像分割的算法对分割准确度的要求更高,算法应具有在各种干扰中从灰度相近的物体中将目标准确分割的能力。常规的图像分割算法如 U-net 算法、Mask R-CNN 算法等难以达到要求的分割准确率。

在人体组织的成像和三维重建方面,朱珊珊等^[8]采用光学相干层析技术实现了人体皮肤的成像;刘浩等^[9]采用光学相干层析技术实现了牙齿的成像;黄硕等^[10]提出利用梯度光图像对人脸进行三维重建;董帅等^[11]提出三维数字图像相关方法对口腔进行测量,完成了对人体口腔的建模。

本文在深度学习框架 U-net 网络的基础上进行改进,提出一种 CBCT 图像的牙齿分割和三维重建方法。通过制作口腔 CBCT 图像数据集,并经编程和训练,实现牙齿图像的三维重建。

2 改进 U-net 网络及图像三维重建原理

2.1 人工神经网络与全卷积神经网络

人工神经网络(ANN)是由对人类神经结构进行

收稿日期: 2022-08-16; 修回日期: 2022-09-27; 录用日期: 2022-10-21

通信作者: *zhao.yuanmeng@cnu.edu.cn

模仿,通过大量人工神经元相连接建立的网络,一般由输入层、输出层和隐藏层组成,其各个节点通过 tanh 函数或 ReLU 函数等激活函数进行连接。它对人工智能的产生和发展起到了奠基作用^[12-13]。在 ANN 的基础上人们发明了卷积神经网络(CNN)。卷积神经网络是典型的以反向传播(BP)为主的神经网络,其具有对图像空间信息的获取能力。

1998 年 LeCun 等^[14]提出了 LeNet 网络,被认为是第一个 CNN 的实现。Girshick 等^[15]在 2014 年首次用 CNN 网络实现了目标检测的功能。之后人们在 CNN 的基础上发展了很多改进算法,尤其是全卷积神经网络(FCN),在图像分割领域做出了巨大的贡献。

2015 年 Long 等^[16]在 CNN 的基础上提出了 FCN。这种全新的神经网络采用反卷积对特征图进行采样以解决语义分割问题,用卷积层取代了原本的全连接层,使其可对图像中的目标物体做到像素点级别的精确识别。它可实现端到端的预测,能够适应任

何尺寸的输入图像,分割准确且鲁棒性好。

2.2 U-net 网络及其改进

U-net 网络是 2015 年由 Ronneberger 等^[17]在 FCN 的基础上提出的一种典型的语义分割网络,其结构如图 1 所示。U-net 被大量应用于医学图像语义分割领域,如杨柳等^[18]采用改进后的 U-net 网络实现了对细胞的自动聚焦相衬成像,李大湘等^[19]采用改进后的 U-net 网络对视网膜图像中的微小血管进行了有效分割,谢远志等^[20]用其完成了对乳腺肿块图像的分割工作。

从图 1 可知,U-net 网络包含两条路径,即左侧的收缩路径和右侧的扩张路径。数据先通过收缩路径进行特征的提取,再通过扩张路径得到输出结果。U-net 的收缩路径和常规的卷积神经网络结构相同,由两个重复的 3×3 卷积组成,每个卷积后面跟着一个修正线性单元(ReLU)和一个步长为 2 的 2×2 最大池化以达到下采样的目的,每个下采样中特征通道数增加一倍。

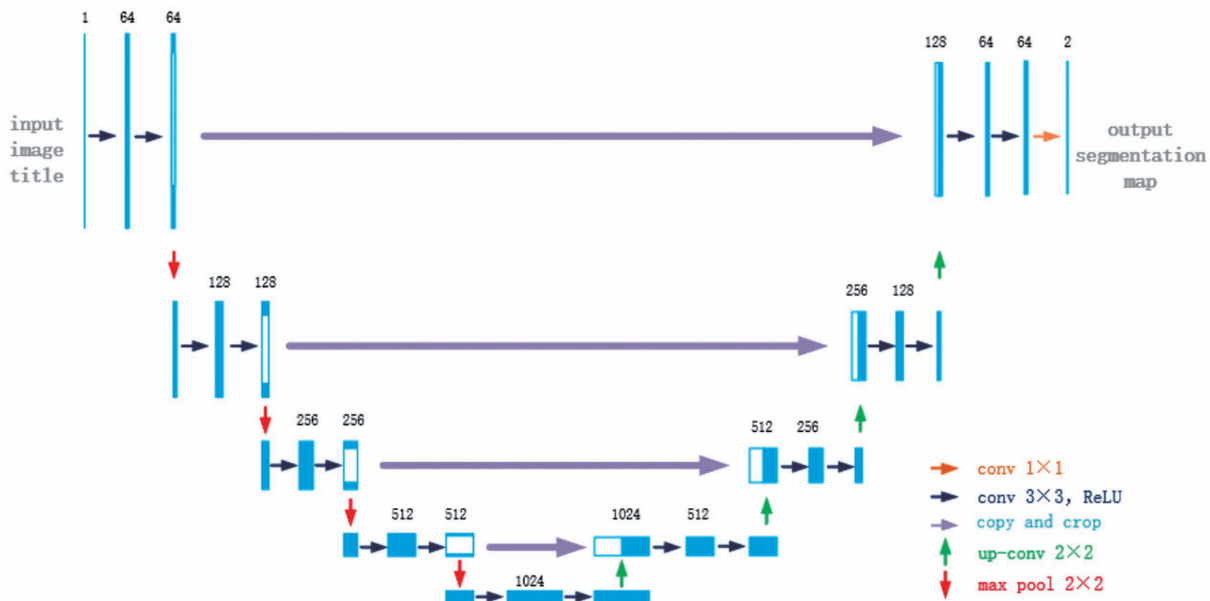


图 1 U-net 网络结构示意图^[17]

Fig. 1 Schematic diagram of U-net network^[17]

在扩张路径上,U-net 网络配置了大量的特征图通道,能让最终复原的分割图像保留丰富的输入信息。每一个步骤包含一个特征图上采样和一个 2×2 的上卷积,使通道数减半,然后是复制和裁剪,把收缩路径中相同层的特征图经过裁剪之后拼接在当前层(由于左侧路径中的特征图比右侧对应路径中的特征图要大一些,因此需要裁剪之后才能做拼接),接着是两个 3×3 卷积和修正线性单元。在最后一个层,使用 1×1 卷积将 64 元素的特征向量映射到不同的类别。这样的 U-net 网络共有 23 个卷积层。这种结构设计可以有效提取图像的细节信息,故有利于其在医学图像语义分割方面的应用。但处理人体口腔图像这种复杂图像时,常规 U-net 的识别效果

已难以满足要求。

为了提高算法的识别效果,本文改进了原始的 U-net 网络,调整了 U-net 网络参数并引入了空间注意力机制^[21]。空间注意力机制是一种模拟人类的思维进行设计的优化方案,其核心思想是依据贡献值对资源进行再分配。本文引入该机制是为了改善神经网络的权重问题。由于研究目的是分割出牙齿,故引入空间注意力机制以改变图像不同区域的权重值,相关性高的区域会拥有更高的权重,使任务需要的目标区域更为突出,算法有更好的识别效果。图 2 为空间注意力机制模块结构图。空间注意力机制可以有效增强牙齿的特征并抑制不重要的特征,从而提高网络的识别能力。

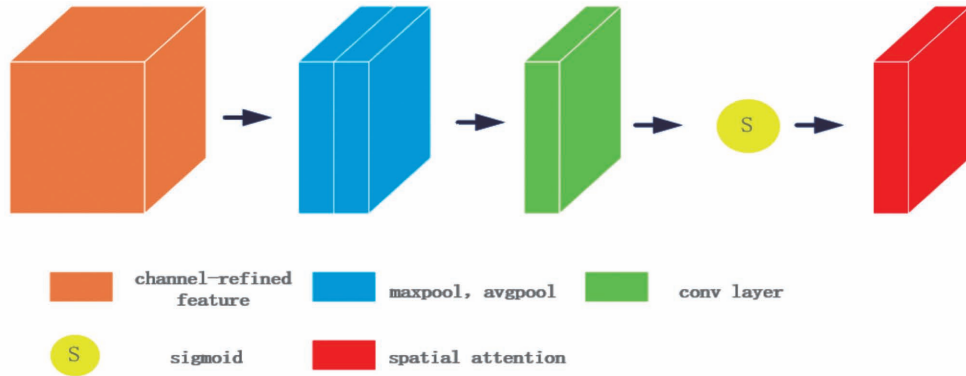


图 2 空间注意力机制模块示意图^[21]

Fig. 2 Schematic diagram of spatial attention mechanism module^[21]

空间注意力机制的具体实现,是对反馈特征图进行最大池化和平均池化操作得到结果 X_{\max} 和 X_{avg} ,并将结果进行拼接,得到复合信号 X_{concat} 。通过 1×1 卷积对 X_{concat} 信号进行一维化,经过 Sigmoid 激活函数进行变换得到权重图。最终输出信号 X_{out} 为输入信号 X_{in} 经过权重图影响后的结果。上述过程表示为

$$X_{\text{concat}} = C(X_{\max} + X_{\text{avg}}), \quad (1)$$

$$X_{\text{out}} = X_{\text{in}} \cdot \delta [\text{conv}(X_{\text{concat}})], \quad (2)$$

式中: C 表示拼接操作; conv 表示卷积。

加入空间注意力机制模块并调整了网络参数后,经过改进的 U-net 网络的结构如图 3 所示。将 U-net 网络输入设置为单通道,分类数为 1,用以适应识别牙齿 CBCT 图像的要求,同时将所有训练用 CBCT 图像的分辨率统一为 $256 \text{ pixel} \times 256 \text{ pixel}$ 。

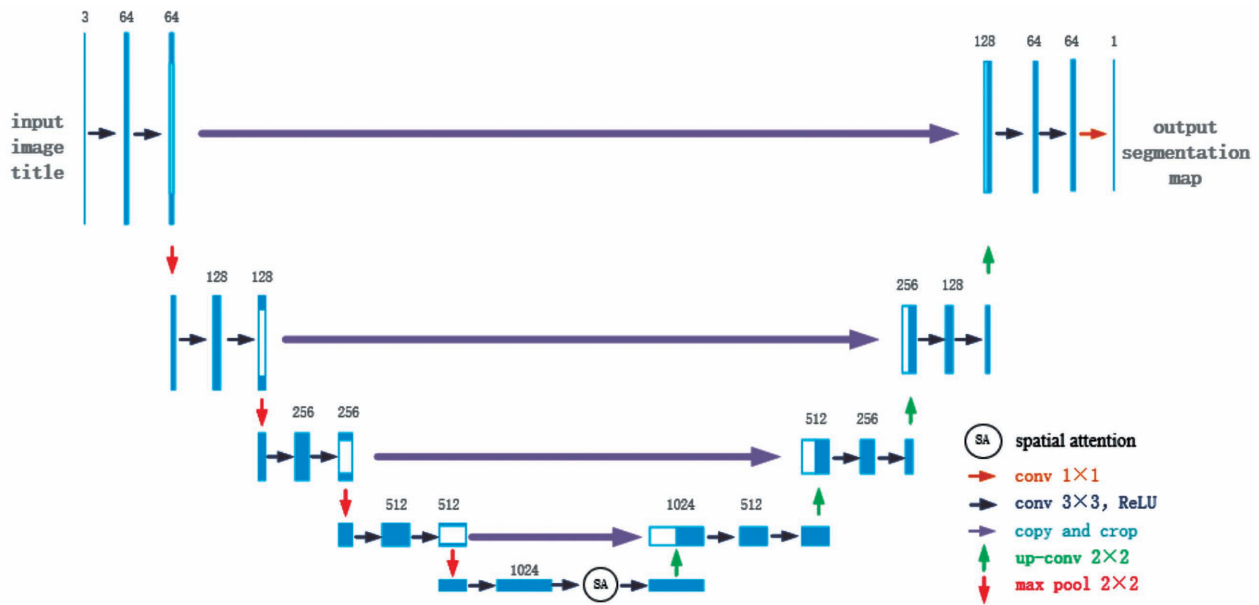


图 3 改进 U-net 网络结构示意图

Fig. 3 Schematic diagram of improved U-net network

2.3 等值面提取三维重建算法

为得到牙齿的三维模型,本文对分割后的 CBCT 图像进行三维重建。常规重建方法主要有体绘制和面绘制两种,其中体绘制法重建效果好,但计算量很大,对设备的性能要求较高;面绘制法具有运行速度快、效率高的优点,医学图像重建采用较多。本文采用 Lorensen 等^[22]提出的经典面绘制重建算法——等值面提取算法,该算法也被称为移动立方体算法。

如图 4 所示,等值面提取算法的中心思想是通过线性插值对空间数据进行逼近,进而对三维数据进行处理。本文用一个由 8 pixel 创建、来自 2 个相邻切片

的体素对表面进行定位。将由三维数据组成的数据集的每个栅格单元看作一个体素,为了便于处理,规定每个体素有 8 个顶点,即立方体的 8 个顶点,对于大于等值面的顶点值规定为正值,反之规定为负值。因此,一个体素与等值面的关系达到 2^8 即 256 种。我们可以建立一个表,将这种关系精确地反映出来。

体素顶点值大于和小于等值面的情况是完全对称的,因此可以把体素与等值面的关系数量减少一半,到 2^7 即 128 种。再利用立方体的旋转对称特性,可以进一步减少到 14 种情形,如图 4 所示。然后对每个体素单元的边建立索引。

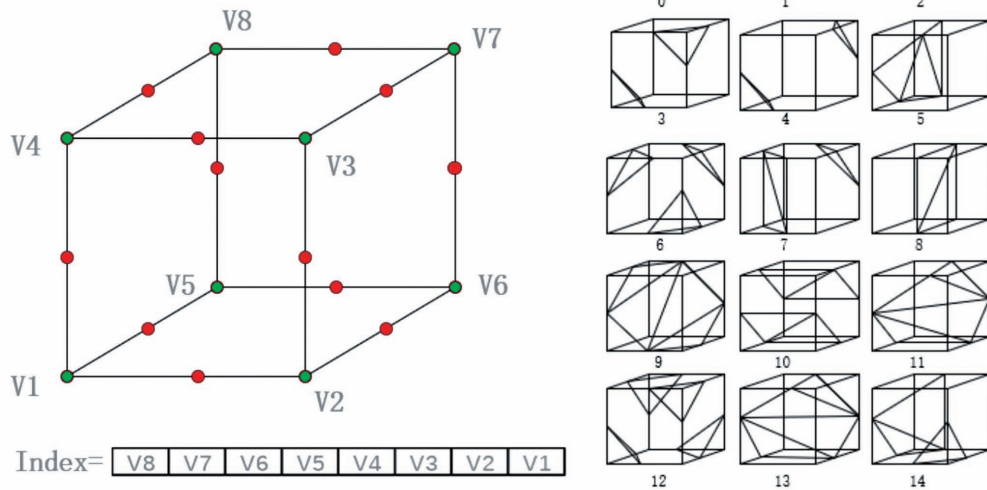


图 4 等值面提取的体素单元顶点和边的命名及 14 种拓扑结构^[22]

Fig. 4 Names of vertices and edges of voxel elements from isosurface extraction and 14 topological structures^[22]

采用相同方法记录边的关系,再由线性插值对交点进行确认,从而生成局部三角。再将全部的三角连接起来,就可得到重建后的三维模型了。

总的说来,本文三维重建方法共有 7 个步骤:(1)将分割后的图片读入内存中;(2)按照顺序创建体素;(3)对比体素与等值面的顶点大小,计算出索引;(4)使用索引从表中查找对应参数;(5)重复第 3 步,对边进行计算;(6)根据顶点和边的情况,生成局部三角;(7)组合三角并绘制等值面图像。

3 CBCT 图像的牙齿图像分割与三维重建

3.1 数据集的制作

为了对数据进行训练,通过与口腔医院合作,在征得病人同意的情况下,采用 CBCT 医学图像扫描仪获取了 10 名病人的牙齿扫描 DICOM 格式的数据。为了便于之后的训练,本文将分辨率统一处理为 256 pixel×256 pixel。然后,采用 Labelme 软件对训练图片进行标记。根据形态学对 10 名病人的口腔 CBCT 断层扫描图像进行手工标记,定义 Label 为 Tooth,选取每 4 frame 标注一次,得到 500 frame 图片的数据。图 5 为采用 Labelme 软件对数据进行标注的具体操作。

3.2 训练模型

本文实验的深度学习平台采用 PyTorch 1.7.0,采用 NVIDIA GTX2080ti 图形处理器(GPU)进行训练,使用了 10 组数据共计 500 张图片,其中 8 组数据用于训练,2 组数据用于测试。为了方便训练,将图片从 RGB 格式转换为灰度图。网络训练过程中输入的初始学习率设定为 0.00001,权重系数设为 0.0005,批尺寸设置为 1;采用 RMSprop 优化算法,权重衰减

为 0.0000001,超参数为 0.9。

RMSprop 算法是在 AdaGrad 算法的基础上改进而来的,原理为先确定一个全局学习率,每次学习中将全局学习率不断地除以经过衰减系数控制的历史梯度平方和的平方根,从而保证每个参数的学习率不同。RMSprop 算法可用下式表示:

$$\begin{cases} S_w = \beta S_w + (1 - \beta) dw^2 \\ S_b = \beta S_b + (1 - \beta) db^2 \\ w = w - \alpha \frac{dw}{\sqrt{S_w} + \epsilon} \\ b = b - \alpha \frac{db}{\sqrt{S_b} + \epsilon} \end{cases}, \quad (3)$$

式中: S_w 和 S_b 表示权重 w 和偏置值 b 在 $t-1$ 轮迭代中的梯度动量; dw 和 db 表示权重 w 和偏置值 b 在 $t-1$ 轮迭代中的微分; α 为学习率; β 为超参数; ϵ 为防止分母为零的参数项。

完成训练后,使用模型对图像进行分割,再根据等值面提取算法对输入的分割结果进行三维重建,得到牙齿的三维模型。

3.3 分割图片实验结果

以 CBCT 断层扫描图像为输入,以图像分割结果为输出,实际测试算法分割识别的效果。

在实验中采用 100 张图片使用训练模型进行预测,再对比数据集从而对训练模型进行评估,获得其平均交并比(MIoU)参数。MIoU 参数的计算公式为

$$R_{MIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}}, \quad (4)$$

式中: R_{MIoU} 为 MIoU 参数值; k 表示总类别数; i 表示

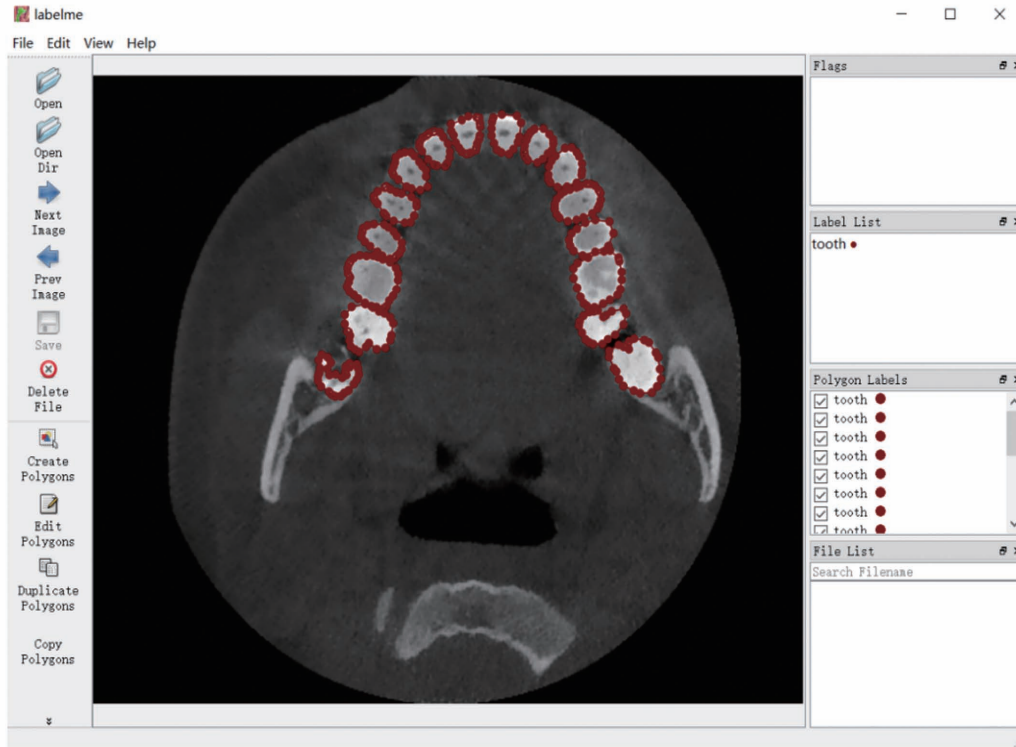


图 5 Labelme 实现对数据集的标注

Fig. 5 Using Labelme to annotate datasets

真实值; j 表示预测值; P_{xy} 表示图像内值 x 被预测为值 y 的像素总个数。

采用常规 U-net 网络训练的模型和采用本文的改进 U-net 网络训练的模型的 MIoU 参数如图 6 所示。

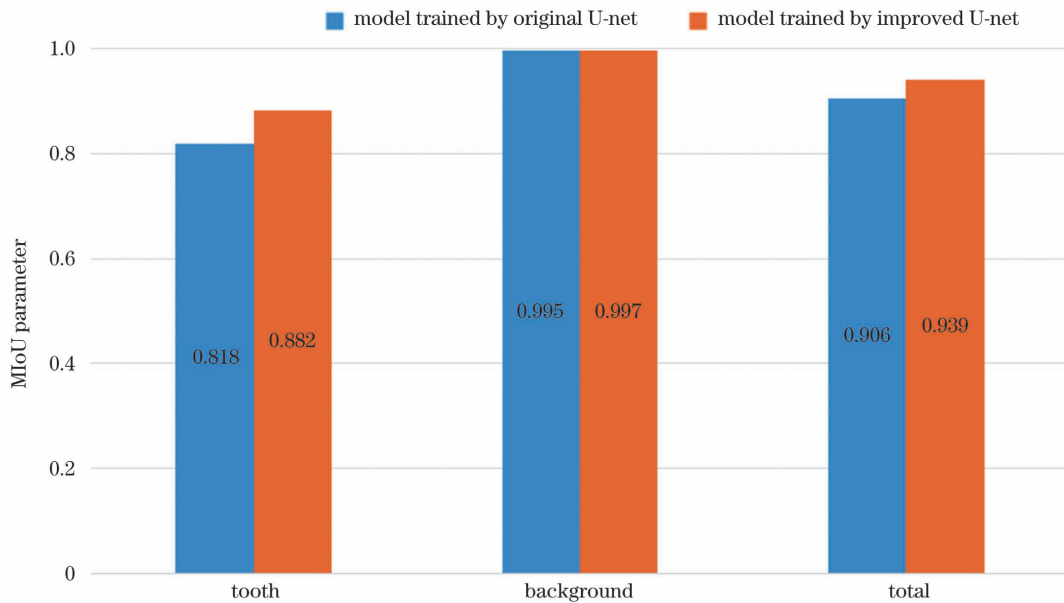


图 6 U-net 模型的 MIoU 参数

Fig. 6 MIoU parameters of U-net models

MIoU 参数可以有效地反映出网络模型对图像的认识精度,其取值为 0~1,数值越大表示模型对目标的识别效果越好。对比本文的改进 U-net 网络和常规的 U-net 网络训练的模型的 MIoU 参数,可以看出改进后的算法对牙齿的识别效果提高了 0.064,对背景的识别效果也提高了 0.002,整体提高了 0.033。通过对 MIoU 参数的定量分析可知改进后的算法有更好

的识别能力。

实际的牙齿分割识别效果如图 7 所示。图 7(a)为待识别图像,图 7(b)为采用常规 U-net 算法训练的模型的识别效果,图 7(c)为采用本文提出的改进 U-net 算法训练的模型的识别效果。

通过对比可以看出,相对于原始的 U-net 算法,本文算法识别效果更佳。对比两种算法对于第一幅图像

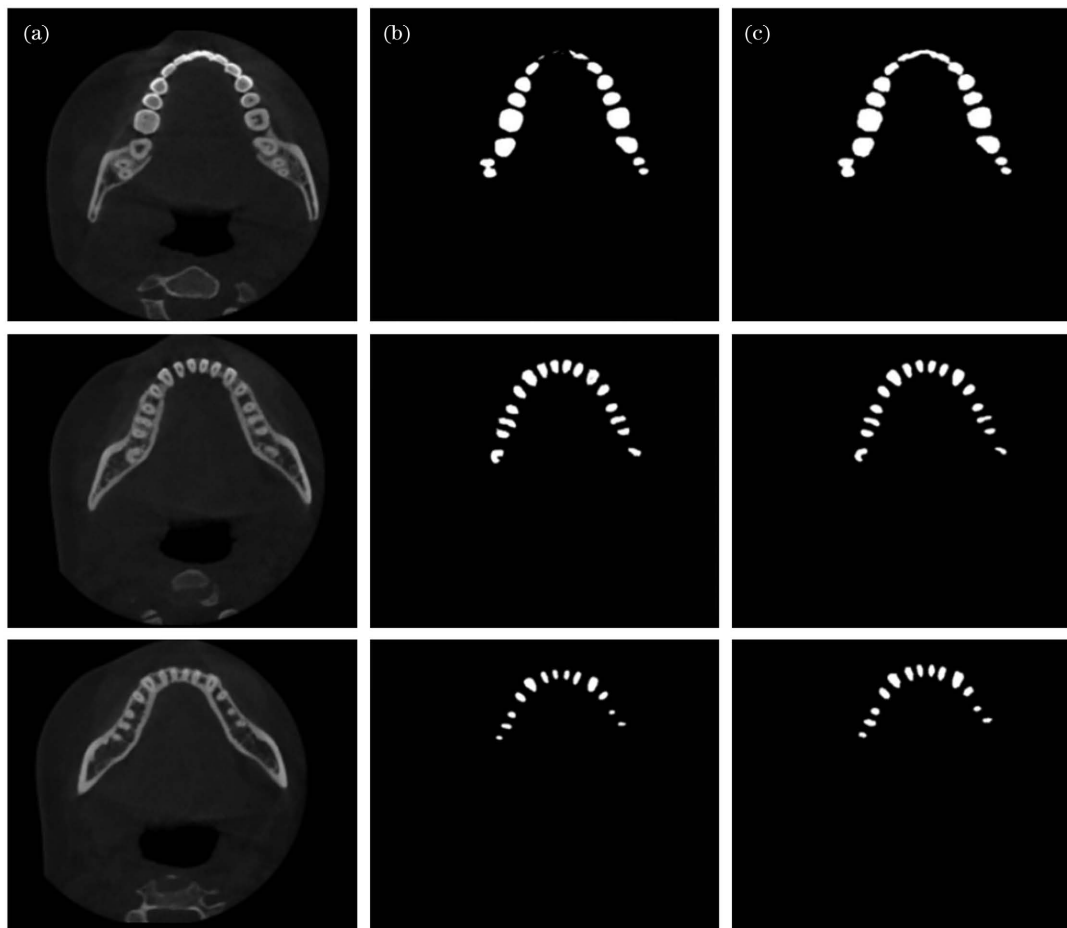


图 7 U-net 模型识别效果。(a)待识别图像;(b)原始 U-net 方法识别效果;(c)改进 U-net 方法识别效果

Fig. 7 Recognition results of U-net models. (a) Images to be recognized; (b) recognition results using original U-net model; (c) recognition results using improved U-net model

的识别效果可见,由于局部亮度影响,原始 U-net 方法缺失了大量图形,而本文的改进算法仍能准确识别。对比另两幅图像的识别效果可见,在图像细节很少的情况下,本文算法识别出的图片中毛刺状伪影更少,证明了本文算法在牙根处也能进行有效识别。

同时本文算法有较强的抗干扰能力。采用患者牙齿含有填充物的图像进行测试,效果如图 8 所示。可见在牙齿填充物的影响下,扫描图像出现了一片亮度极高的区域,但本算法仍然能够有效地实现对该图像

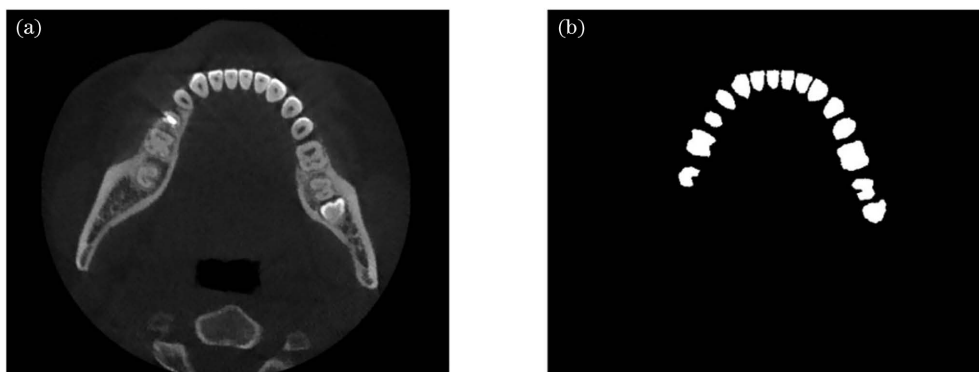


图 8 带有牙齿填充物的图像识别效果。(a)待识别图像;(b)改进 U-net 方法识别效果

Fig. 8 Recognition result of image with dental fillings. (a) Image to be recognized; (b) recognition result using improved U-net model

3.4 三维重建实验结果

采用医学界常用软件 MicroDicom Viewer 进行三维重建开展参照实验。MicroDicom Viewer 三维重

建效果如图 9 所示。

由图 9 可见,由于没有对牙齿进行分割,导致重建出的三维模型包含了所有人体组织,同时重建图像中

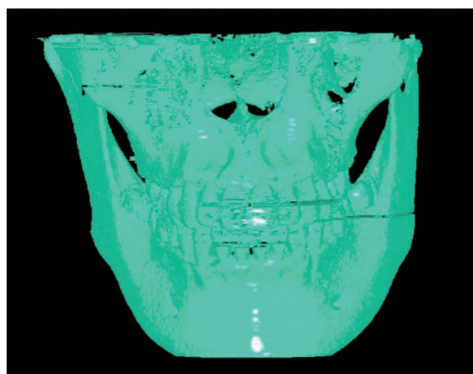


图 9 用 MicroDicom Viewer 对口腔 CBCT 数据进行重建的效果图

Fig. 9 Reconstruction result of oral CBCT data using MicroDicom Viewer

有大量的断层,重建效果并不理想,不利于医生诊断病情。采用本文的改进 U-net 网络,在顺利解决了识别过程中牙颌骨和牙齿图像粘合的问题之后,对分割后的图像进行三维重建,得到的结果如图 10 所示。

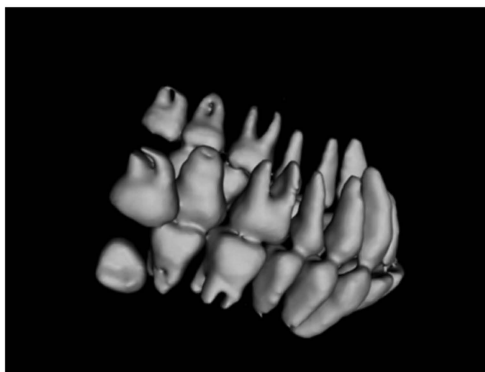


图 10 采用本文算法对口腔 CBCT 数据进行三维重建的效果

Fig. 10 Reconstruction result of oral CBCT data using proposed algorithm

对比图 9 和图 10 可见,图 10 中牙齿被有效地重建出来,能够清晰地看到牙齿的轮廓和相对位置以及牙齿的根部。在原始 CBCT 数据中,牙齿根部往往与牙颌骨靠近,因此难以被分辨出来。另外,图 10 的重建图像轮廓清晰完整,没有出现图 9 中大量存在的断层。

上述实验采用的是一位阻生智齿患者的 CBCT 数据,通过重建可以从侧面清晰地看到病人的病情,明显看出该病人最里侧两颗牙齿的生长方向与正常牙齿完全不同,尤其是下方的阻生齿。实验结果表明本文方法在口腔医疗领域能够取得较好的应用效果。

4 结 论

为解决从口腔 CBCT 图像中分割出牙齿图像并对其进行三维建模的问题,制作了口腔 CBCT 图像数据集,设计了引入注意力机制的改进 U-net 网络并进行训练,结合三维重建算法顺利实现了牙齿图像分割和三维重建。

实验证明,改进后的网络能准确地从原始图像中

识别出牙齿,相比于常规 U-net 网络有更好的识别能力,并能够在牙齿填充物的干扰下准确地进行识别。最终重建出的三维模型相比于 MicroDicom Viewer 软件重建的三维模型有着更为清晰完整的轮廓,能够更加真实地还原病人牙齿状况。

本文方法在合作医院进行试用反馈良好,帮助口腔科医生诊断和治疗的效果显著,有望在口腔医疗健康领域做出更大的贡献。

参 考 文 献

- [1] Lechuga L, Weidlich G A. Cone beam CT vs. fan beam CT: a comparison of image quality and dose delivered between two differing CT imaging modalities[J]. *Cureus*, 2016, 8(9): e778.
- [2] Pavaloiu I B, Goga N, Marin I, et al. Automatic segmentation for 3D dental reconstruction [C] // 2015 6th International Conference on Computing, Communication and Networking Technologies (ICCCNT), July 13-15, 2015, Dallas-Fortworth, TX, USA. New York: IEEE Press, 2015.
- [3] Kass M, Witkin A, Terzopoulos D. Snakes: active contour models[J]. *International Journal of Computer Vision*, 1988, 1(4): 321-331.
- [4] 王立新, 刘新新, 刘希云, 等. 基于区域自适应形变模型的 CT 图像牙齿结构测量方法研究[J]. *生物医学工程学杂志*, 2016, 33(2): 308-314.
Wang L X, Liu X X, Liu X Y, et al. Quantitative segmentation and measurement of tooth from computed tomography image based on regional adaptive deformation model [J]. *Journal of Biomedical Engineering*, 2016, 33(2): 308-314.
- [5] Miki Y, Muramatsu C, Hayashi T, et al. Classification of teeth in cone-beam CT using deep convolutional neural network [J]. *Computers in Biology and Medicine*, 2017, 80: 24-29.
- [6] Cui Z M, Li C J, Wang W P. ToothNet: automatic tooth instance segmentation and identification from cone beam CT images[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 6361-6370.
- [7] Girdhar R, Gkioxari G, Torresani L, et al. Detect-and-track: efficient pose estimation in videos [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 350-359.
- [8] 朱珊珊, 高万荣, 史伟松. 基于图形处理器的人体皮肤组织实时成像谱域相干光断层成像系统[J]. *激光与光电子学进展*, 2018, 55(4): 041701.
Zhu S S, Gao W R, Shi W S. GPU-based Fourier domain optical coherence tomography system for real time imaging of human skin [J]. *Laser & Optoelectronics Progress*, 2018, 55(4): 041701.
- [9] 刘浩, 高万荣, 陈朝良. 手持式牙齿在体谱域光学相干层析成像系统研究[J]. *中国激光*, 2016, 43(2): 0204003.
Liu H, Gao W R, Chen C L. Handheld spectral domain optical coherence tomography for *in vivo* tooth imaging [J]. *Chinese Journal of Lasers*, 2016, 43(2): 0204003.
- [10] 黄硕, 胡勇, 巩彩兰, 等. 基于梯度光图像的高精度三维人脸重建算法[J]. *光学学报*, 2020, 40(4): 0410001.
Huang S, Hu Y, Gong C L, et al. High precision three-dimensional face reconstruction algorithm based on gradient light image [J]. *Acta Optica Sinica*, 2020, 40(4): 0410001.
- [11] 董帅, 戴云彤, 董芃良, 等. 应用多相机三维数字图像相关实现口腔印模三维重构[J]. *光学学报*, 2015, 35(8): 0812006.
Dong S, Dai Y T, Dong E L, et al. Three-dimensional reconstruction of dental impression based on multi-camera three-dimensional digital image correlation method [J]. *Acta Optica Sinica*, 2015, 35(8): 0812006.

- [12] Nicholas L. Neural networks for pattern recognition [M]. Oxford: Oxford University Press, 1997.
- [13] Clevert D A, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (ELUs) [EB/OL]. (2015-11-23) [2022-05-06]. <https://arxiv.org/abs/1511.07289>.
- [14] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [15] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 580-587.
- [16] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 3431-3440.
- [17] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation [M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture Notes in Computer Science. Cham: Springer, 2015, 9351: 234-241.
- [18] 杨柳, 王华英, 董昭, 等. 基于改进 U-Net 的自动聚焦相衬技术在细胞成像中的应用 [J]. 中国激光, 2022, 49(15): 1507302.
- Yang L, Wang H Y, Dong Z, et al. Application of auto-focusing technology based on improved U-net in cell imaging [J]. Chinese Journal of Lasers, 2022, 49(15): 1507302.
- [19] 李大湘, 张振. 基于改进 U-Net 视网膜血管图像分割算法 [J]. 光学学报, 2020, 40(10): 1010001.
- Li D X, Zhang Z. Improved U-net segmentation algorithm for the retinal blood vessel images [J]. Acta Optica Sinica, 2020, 40(10): 1010001.
- [20] 谢远志, 闫士举, 魏高峰, 等. 基于 U-Net++ 和对抗性学习网络的乳腺肿块分割 [J]. 激光与光电子学进展, 2022, 59(16): 1617002.
- Xie Y Z, Yan S J, Wei G F, et al. Breast mass segmentation based on U-net++ and adversarial learning network [J]. Laser & Optoelectronics Progress, 2022, 59(16): 1617002.
- [21] Zhong Z L, Lin Z Q, Bidart R, et al. Squeeze-and-attention networks for semantic segmentation [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 13062-13071.
- [22] Lorensen W E, Cline H E. Marching cubes: a high resolution 3D surface construction algorithm [J]. ACM SIGGRAPH Computer Graphics, 1987, 21(4): 163-169.

Study on Tooth Cone Beam CT Image Reconstruction Based on Improved U-net Network

Liu Haoxin^{1,2,3,4}, Zhao Yuanmeng^{1,2,3,4*}, Zhang Cunlin^{1,2,3,4}, Zhu Fengxia^{1,2,3,4},
Yang Moxuan^{1,2,3,4}

¹Department of Physics, Capital Normal University, Beijing 100048, China;

²Key Laboratory of Terahertz Optoelectronics, Ministry of Education, Beijing 100048, China;

³Beijing Key Laboratory for Terahertz Spectroscopy and Imaging, Beijing 100048, China;

⁴Beijing Advanced Innovation Center for Imaging Theory and Technology, Beijing 100048, China

Abstract

Objective Cone beam computed tomography (CBCT) has played an important role in the diagnosis of oral diseases. However, it remains a very important research issue in stomatology to obtain accurate information of teeth from complex original scanning images. Because of the complexity of the human oral environment, various human tissues, especially dental bones, seriously affect the teeth segmentation algorithm. This problem is particularly obvious in the image regions around the roots of the teeth. Moreover, tooth fillings also seriously interfere with the operation of the algorithm. Therefore, compared with other image segmentation algorithms, the current core requirement of teeth image segmentation algorithms is higher segmentation accuracy. The algorithm should have the ability to accurately segment targets out of objects with similar gray levels in various interferences. Conventional image segmentation algorithms such as U-net algorithm and Mask R-CNN algorithm cannot achieve the required segmentation accuracy. The main purpose of this paper is to research and propose a method of teeth image segmentation and three-dimensional (3D) reconstruction from CBCT images. We construct CBCT image dataset from original images and realize 3D teeth image reconstruction through programming and training.

Methods We use a novel segmentation algorithm based on an improved U-net network, for which the spatial attention mechanism is introduced into conventional artificial neural network. Spatial attention mechanism can effectively enhance the features of teeth and suppress unimportant features to improve the recognition ability of the network, which can help us to segment teeth more accurately. Combined with the isosurface extraction algorithm, the accurate segmentation and three-dimensional reconstruction of teeth CBCT images are realized. First, we use Labelme software with the patient data collected from medical institutions to create a human oral CBCT dataset, and then we train the model using the improved U-net network. Second, we use the trained model to segment the images to obtain images that only retain the teeth information. Third, the results are reconstructed to create a three-dimensional model of the teeth by Marching Cubes algorithm.

Results and Discussions We applied the new method proposed in this paper to the CBCT tomographic images, and obtained the image segmentation and reconstruction results. We actually tested the segmentation and recognition effect of the algorithm, and got the MIoU parameters of the model, as well as the image segmentation and reconstruction effect. The MIoU parameters of the improved U-net network and the original U-net network were compared. It could be seen that the improved algorithm has 0.064 higher recognition effect on teeth, 0.002 higher recognition effect on background, and 0.033 higher overall recognition effect. The quantitative analysis of MIoU parameters showed that the improved algorithm has better recognition ability. Compared with the actual teeth image segmentation recognition effect, it could be seen that the newly proposed algorithm in this paper has better recognition effect and can effectively recognize the image under the influence of tooth fillings. Compared with the 3D model reconstructed by MicroDicom Viewer software, it could be seen that the teeth had been effectively reconstructed, and the contour, relative position and the roots of the teeth can be clearly seen. In contrast, the roots of the teeth are difficult to distinguish in the original CBCT data, because they are often close to the jaw. In addition, the reconstructed image of this experiment had clear and complete contour, and there was no fault. The above experiments used the CBCT data of an impacted wisdom tooth patient. After applying our reconstruction algorithm to the data, the patient's oral condition could be clearly seen from the side. It was obvious that the growth direction of the innermost two teeth of the patient is completely different from that of the normal teeth, especially the impacted tooth below. In summary, the experimental results showed that the method proposed in this paper has a good application effect in the field of oral medicine.

Conclusions The experimental results showed that the improved network can accurately recognize teeth from the original image, and has better recognition ability than the original U-net network; besides, the novel network can accurately recognize teeth under the interference of tooth fillings. Compared with the 3D model reconstructed by MicroDicom Viewer software, the final reconstructed 3D model has a clearer and more complete contour, and can more realistically reproduce the patient's tooth conditions. The experimental results also showed that this method can effectively extract tooth information, which is helpful to the diagnosis and treatment of oral diseases, especially tooth diseases. The trial feedback of this method testing in the cooperative hospital is quite terrific, and the effect of helping stomatologists to diagnose and treat is significant. The research findings of this paper are expected to make greater contributions in the field of oral health in the future.

Key words imaging systems; cone beam computed tomography; neural network; three-dimensional reconstruction; spatial attention mechanism