

## 基于CA-SegResNet的CT图像标识椎骨自动分割

朱中旗<sup>1</sup>, 高晓隆<sup>2</sup>, 李英豪<sup>1</sup>, 杨光<sup>1</sup>, 郝利国<sup>3</sup>, 汪红志<sup>1</sup>

1. 华东师范大学物理与电子科学学院上海市磁共振重点实验室, 上海 200062; 2. 富锦市中医医院影像科, 黑龙江 富锦 156100;  
3. 齐齐哈尔医学院医学技术学院分子影像研究室, 黑龙江 齐齐哈尔 161006

**【摘要】**针对脊椎计算机断层扫描(CT)图像标识椎骨分割问题,提出一种融合三维坐标注意力机制的三维医学图像分割网络CA-SegResNet。该网络通过深度残差卷积神经网络提取图像特征,将编码器每层输出的特征图与解码器每层的输入相融合,然后引入三维坐标注意力模块捕获通道间关系以及方向和位置信息,建立起不同空间方向的长距离依赖关系,实现标识椎骨的精准分割。在对105个病例的颈椎标识椎骨(第7节颈椎)和胸椎标识椎骨(第12节胸椎)的分割任务中,CA-SegResNet在测试集上的分割平均Dice系数(DSC)分别为0.9345和0.9189,平均豪斯多夫距离(HD)为7和8 mm。与U-Net相比,平均DSC分别提高0.0145和0.0463,平均HD分别缩小176和388 mm。实验结果表明,该算法能够对CT图像标识脊椎进行精准分割。

**【关键词】**深度学习; 计算机断层扫描; 椎骨分割; 分割网络; 坐标注意力

**【中图分类号】**R318; TP391.41

**【文献标志码】**A

**【文章编号】**1005-202X(2024)11-1349-08

### Automatic segmentation of identified vertebral bones from CT images using CA-SegResNet

ZHU Zhongqi<sup>1</sup>, GAO Xiaolong<sup>2</sup>, LI Yinghao<sup>1</sup>, YANG Guang<sup>1</sup>, HAO Ligu<sup>3</sup>, WANG Hongzhi<sup>1</sup>

1. Shanghai Key Laboratory of Magnetic Resonance, School of Physics and Electronic Science, East China Normal University, Shanghai 200062, China; 2. Department of Imaging, Fujin Chinese Medicine Hospital, Fujin 156100, China; 3. Laboratory of Molecular Imaging, School of Medical Technology, Qiqihar Medical University, Qiqihar 161006, China

**Abstract:** A three-dimensional (3D) medical image segmentation network (CA-SegResNet) which incorporates a 3D coordinate attention mechanism is proposed to address the issue of segmenting identified vertebral bones from spinal computed tomography (CT) images. The network extracts image features through a deep residual convolutional neural network and fuses the feature maps from each encoder layer with the input of the corresponding decoder layer. Subsequently, a 3D coordinate attention module is introduced to capture inter-channel relationships as well as directional and positional information, establishing long-range dependencies across different spatial directions, thereby enabling precise segmentation of the identified vertebral bones. For the segmentation tasks involving the identified cervical vertebra (the 7th cervical vertebra) and the identified thoracic vertebra (the 12th thoracic vertebra) across 105 cases, CA-SegResNet achieves average Dice similarity coefficients (DSC) of 0.9345 and 0.9189 on the test set, with average Hausdorff distances (HD) of 7 and 8 mm. Compared with U-Net results, the average DSC is improved by 0.0145 and 0.0463, while average HD is reduced by 176 and 388 mm. The results demonstrate that the network can realize the precise segmentation of identified vertebral bones from CT images.

**Keywords:** deep learning; computed tomography; vertebral segmentation; segmentation network; coordinate attention

**【收稿日期】**2024-05-07

**【基金项目】**国家自然科学基金(61731009)

**【作者简介】**朱中旗, 硕士研究生, 研究方向: 人工智能在医学影像中的应用, E-mail: arcmosin@163.com; 高晓隆, 硕士, 主治医师, 研究方向: 医学影像诊断、医学影像AI技术, E-mail: 185977790@qq.com (朱中旗和高晓隆为共同第一作者)

**【通信作者】**郝利国, 硕士, 教授, 研究方向: 影像医学与核医学, E-mail: haoliguo@qmu.edu.cn; 汪红志, 博士, 高级工程师, CCF会员, 研究方向: 医学影像设备与技术、医学影像人工智能技术, E-mail: hzwang@phy.ecnu.edu.cn

### 前言

因脊椎疾病发病率逐年增高,并呈现出逐渐年轻化的趋势,脊椎疾病正在成为当今社会的重要问题<sup>[1]</sup>。CT影像检查具有图像清晰、无影像重叠、密度分辨率高等优点,使其成为临床诊断中首选的影像检查手段<sup>[2]</sup>。对脊椎的计算机断层扫描(Computed Tomography, CT)图像的感兴趣区进行精准分割,是各种脊椎疾病诊断与治疗的重要步骤。在影像科医

生临床阅片诊断时,往往需要满足快速进行脊柱分段、定位不同椎骨或者病灶区域等需求,所以利用少量关键椎骨的标注,即标识椎骨的标注,使用计算机算法实现对标识椎骨的快速分割,可以帮助影像科医生更准确高效地诊断。然而,脊椎图像椎体边缘成像模糊、成像伪影、相邻椎骨结构相似以及病理造成的变化等,都造成脊椎椎体分割的困难。

椎骨自动分割的方法主要有传统方法和基于深度学习的方法两大类。传统方法有分水岭分割<sup>[3]</sup>、基于形状模型分割<sup>[4]</sup>、基于形状先验水平集分割<sup>[5-6]</sup>、基于图谱分割<sup>[7-8]</sup>等,但这些方法计算耗时长,分割精度低,泛用性差,难以应用于临床。随着深度学习算法在图像处理领域的发展,基于神经网络的图像分割方法成为医学图像分割的主流。Janssens等<sup>[9]</sup>提出一种基于级联的全卷积神经网络(Fully Convolutional Networks, FCN),通过一个3D回归FCN将分割网络的注意力转移到腰椎区域,然后通过一个类U-Net网络对每个腰椎进行分割。师文博等<sup>[10]</sup>提出一种双通道语义分割算法DAU-Net,通过空间通道与语义通道分别学习空间信息特征与图像语义特征,并在解码器端对两类特征进行融合,获取脊柱X光图像中更精准的分割边界。Nazir等<sup>[11]</sup>提出ECSU-Net,该网络通过将椎骨图像的3个断面分别用U-Net做分割,然后通过K-Means聚类识别不同椎骨,最后将3个断面的分割结果融合在一起。黄昆等<sup>[12]</sup>提出DD-Net网络用于椎骨分割,该网络采用单编码双解码结构,并加入双重特征融合模块和高效注意力模块,来增强网络的特征提取能力。刘晶等<sup>[13]</sup>提出一种基于3D U-Net框架的脊柱CT图像改进分割网络,通过融合SE残差单元、椎骨边缘分割模型与改进混合通道-空间注意力机制,在保证脊柱分割精度和分割效率有效提高的同时具有较好的泛化性与鲁棒性。以上工作在椎骨的2D与3D图像分割方面都做出不错的成绩,但它们大多只是针对脊椎公开数据集以及整个脊柱进行探索,未讨论针对标识椎骨的分割方法。

对整个脊柱进行分割需要依赖大量的专业医生标注,而通过关键标识椎骨的分割结果,快速进行脊柱分段,椎骨定位与病灶划分,在节省大量勾画成本的同时也能很好地满足临床需求。本研究针对标识椎骨即第7节颈椎(C7)和第12节胸椎(T12)进行分割,基于SegResNet<sup>[14]</sup>网络结构做出改进,提出CA-SegResNet模型。该模型在SegResNet网络的基础上,增加网络深度来提升网络多尺度特征提取能力,同时为了解决椎骨间相似度高的分割难点,使用三维坐标注意力机制对通道间关系进行建模,捕捉椎骨形状的细节特征,并建立长程空间依赖关系,最

终可以实现对标识椎骨的精准自动分割。

## 1 数据与实验设计

### 1.1 实验数据

本实验采用的数据集来自齐齐哈尔医学院第一附属医院和第三附属医院,共105例颈椎和105例胸椎CT样本。单个颈椎病例样本在横断面上约有300张2D图像,单个胸椎病例样本在横断面上约有700张2D图像,样本的数据格式统一为DICOM格式,总计约105 000张图像。数据集中颈椎样本标注第7节颈椎,胸椎样本标注第12节胸椎,这两块标识椎骨即颈椎与胸椎、胸椎与腰椎的分界椎骨。所有样本均由7名具有3~16年影像学经验的放射科医生使用ITK-SNAP软件(版本3.8.0)进行标注。

### 1.2 实验环境

本实验在操作系统为Ubuntu 20.04 LST的云计算平台MATGO上进行,内存86 G, GPU为GeForce RTX 2080 Ti。所有程序由Python3.8, Pytorch1.11和医学图像深度学习框架MONAI实现。实验参数方面,采用Adam优化器,学习率为0.003,动量为0.9,训练最大迭代次数(epoch)为600。训练过程中使用验证集对学习过程进行监控,并通过验证集的表现选择最佳的模型。

### 1.3 实验流程

本文的实验流程如图1所示,包括7个步骤:(1)采用阈值截取法,调整CT图像的窗宽窗位为最小值40,最大值300。使用双线性(bilinear)下采样和邻近值(nearest)插值上采样的方法,统一所有图像的体素间隔(spacing)为1 mm×1 mm×1 mm;(2)将实验数据按8:1:1的比例拆分成训练集、验证集、测试集;(3)对训练集数据进行随机翻转、随机小角度旋转和随机缩放来扩充训练数据量,这种数据增广方式并不是硬性增加训练样本的总量,而是在训练每个epoch时,对每个样本按50%的概率进行上述增广操作;(4)通过Min-Max函数将所有数据的图像体素数值缩放到(0, 1)的范围;(5)为了降低GPU计算开销,将训练集数据每个图像按正负样本比为4:1的比例,随机裁剪为5个patch为96×96×96的子图;(6)使用CA-SegResNet对训练集数据进行训练,同时通过验证集结果对训练过程进行监控来获取最佳模型权重,由于验证集没有经过随机裁剪,故采用滑动窗口方式输入网络,滑动窗口大小为96×96×96,滑动重叠度为25%;(7)使用最佳模型对测试集数据进行推理预测,测试集数据输入网络的方式采用与验证集数据相同的滑动窗口的方式,并对测试集推理结果进行评估。

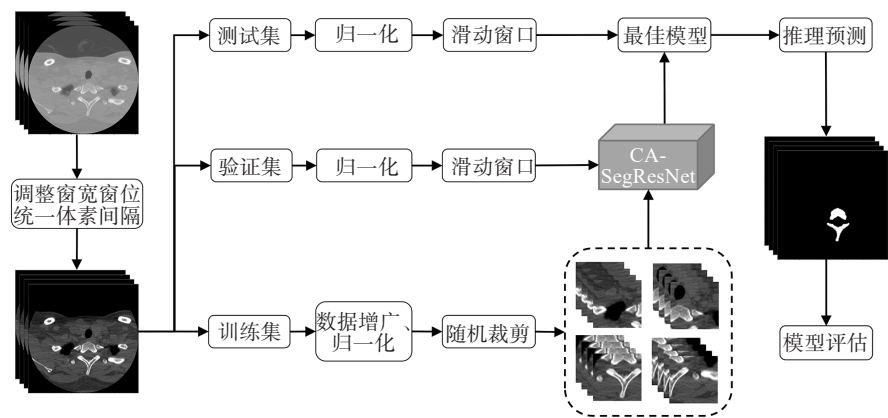


图1 实验流程  
Figure 1 Experiment procedure

1.4 CA-SegResNet 网络

本文方法基于SegResNet网络进行改进,加深网络深度同时结合基于坐标注意力改进的三维坐标注意力机制(3D Coordinate Attention)<sup>[15]</sup>,提出CA-SegResNet,其结构如图2所示。CA-SegResNet使用5层残差卷积

神经网络对脊椎CT图像进行多尺度特征提取,并通过跳跃连接结构将底层信息与高层信息融合,来避免梯度消失的问题。三维坐标注意力模块则可以建立不同空间方向的长距离依赖关系,有效地捕捉目标区域的形状和细节。

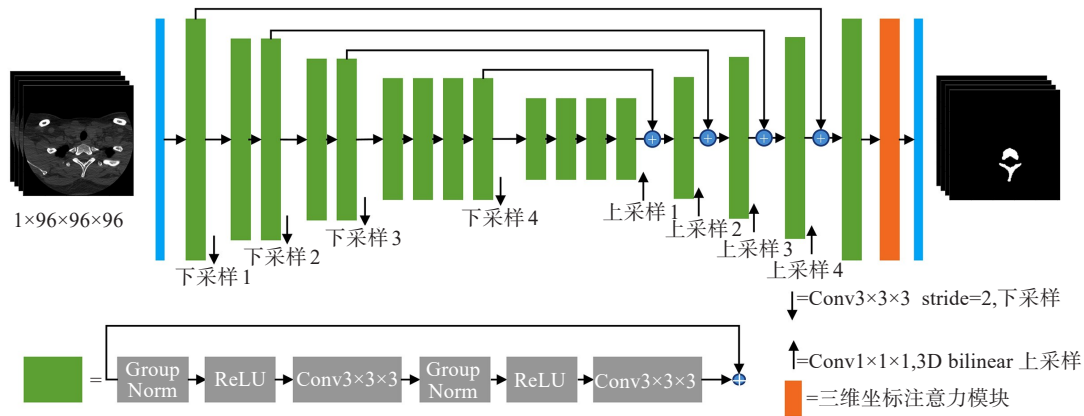


图2 CA-SegResNet网络结构  
Figure 2 CA-SegResNet structure

**1.4.1 编码器结构** 在编码器中,使用分割残差卷积块(SegResBlock, SRBlock)作为基础卷积单元<sup>[14]</sup>,其中每个块由两个带有分组归一化(Group Norm)<sup>[16]</sup>和ReLU<sup>[17]</sup>的卷积组成,并在输入和输出之间加入残差连接<sup>[18]</sup>。编码器具有5个网络层级,分别包含1、2、2、4、4个SRBlock单元。每个层级之间采用大小为3×3×3,步长为2的卷积操作作为下采样方式。编码器每一层都会将输入的特征图通道数增加一倍,初始层会输出通道数为8的特征图,最后一层会输出通道数为128的特征图。

**1.4.2 解码器结构** 解码器具有4个网络层级,每个网络层的输入图像都会通过大小为1×1×1、步长为2

的转置卷积进行上采样,将上一层的特征图分辨率扩大一倍,通道数减小一倍。然后,将输入图像与对应编码器层输出的特征图进行融合,最后通过1个SRBlock单元进行卷积运算,得到输出的图像。

**1.4.3 三维坐标注意力机制** 注意力机制已广泛地应用在深度学习的诸多领域。基于注意力机制的结构模型不仅能够记录信息间的位置关系,还能依据信息的权重去度量不同信息特征的重要性<sup>[19]</sup>。三维坐标注意力是本文基于坐标注意力机制提出的三维改进型,利用图像中3个方向的空间信息来控制该位置的注意力权重,将图像中的每个位置表示为其空间坐标,在此基础上使用多个全连接层来计算该位置的注意力权重,其

结构如图3所示。在这个过程中,相邻位置之间的空间关系会被编码到注意力权重中,从而使得模型能够更好地理解图像中的结构信息。

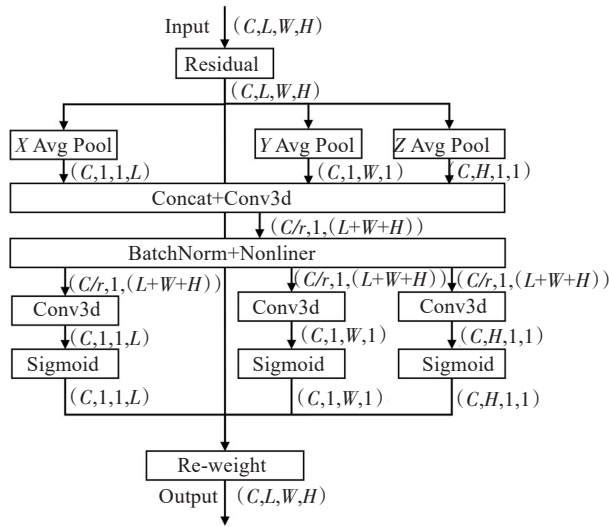


图3 三维坐标注意力机制

Figure 3 Three-dimensional coordinate attention mechanism

三维坐标注意力模块由坐标信息嵌入和坐标注意力生成两部分组成。

坐标信息嵌入。将给定的输入 $X$ ,通过图中的 $X$  Avg Pool、 $Y$  Avg Pool和 $Z$  Avg Pool这3个大小为 $(H, 1, 1)$ 、 $(1, W, 1)$ 和 $(1, 1, L)$ 的池化核,分解为3个方向成对应的一维特征编码,从而将坐标信息嵌入到输入序列中。 $z_c^l(l)$ 、 $z_c^w(w)$ 和 $z_c^h(h)$ 分别表示3个方向上的注意力一维编码, $L$ 、 $W$ 、 $H$ 为特征图的长、宽、高, $x_c(i, j, k)$ 为特征图在 $(i, j, k)$ 位置的值。其数学表达式如式(1)~式(3)所示:

$$z_c^l(l) = \frac{1}{W \times H} \sum_{\substack{0 \leq j < W \\ 0 \leq k < H}} x_c(l, j, k) \quad (1)$$

$$z_c^w(w) = \frac{1}{L \times H} \sum_{\substack{0 \leq i < L \\ 0 \leq k < H}} x_c(i, w, k) \quad (2)$$

$$z_c^h(h) = \frac{1}{W \times L} \sum_{\substack{0 \leq i < L \\ 0 \leq j < W}} x_c(i, j, h) \quad (3)$$

坐标注意力生成。首先将上阶段生成的3个方向的注意力一维编码进行拼接,然后通过 $1 \times 1 \times 1$ 的卷积转换函数 $F_l$ 生成一个对3个方向上的空间信息进行编码的中间特征图 $f$ ,然后再沿着空间维度将 $f$ 拆分成3个独立的张量 $f^l$ 、 $f^w$ 和 $f^h$ ,最后通过3个 $1 \times 1 \times 1$ 的卷积转换函数 $F_l$ 、 $F_w$ 和 $F_h$ 分别将 $f^l$ 、 $f^w$ 和 $f^h$ 变换为与原始输入 $X$ 具有相同通道数的张量,计算过程如式(4)~式(7)所示:

$$f = \delta(F_l([z_c^l, z_c^w, z_c^h])) \quad (4)$$

$$g_c^l(i) = \sigma(F_l(f^l)) \quad (5)$$

$$g_c^w(j) = \sigma(F_w(f^w)) \quad (6)$$

$$g_c^h(k) = \sigma(F_h(f^h)) \quad (7)$$

其中, $\delta$ 是一种非线性激活函数, $f \in R^{C/r \times (L+W+H)}$ 是中间特征图,用于编码3个方向的空间信息, $r$ 是控制块大小的缩减率, $\sigma$ 为Sigmoid函数。坐标注意力模块的最终输出如式(8)所示:

$$y_c(i, j, k) = x_c(i, j, k) \times g_c^l(i) \times g_c^w(j) \times g_c^h(k) \quad (8)$$

相较于通道注意力<sup>[20]</sup>模块只关注重新加权不同通道重要性和空间注意力<sup>[21]</sup>模块只关注重新加权单一维度空间重要性不同,三维坐标注意力模块不仅能获取不同通道重要性的信息,还能获取不同维度空间重要性的信息,从而使得模型更准确地定位感兴趣区的分割位置。

**1.4.4 损失函数** 本文所使用的损失函数为Dice loss函数,其表达式如式(9)所示:

$$\text{Dice loss} = 1 - \frac{2TP}{FN + 2TP + FP} \quad (9)$$

其中,TP、FP、TN、FN分别为分割任务预测像素的真阳性、假阳性、真阴性、假阴性的数量。

### 1.5 评价指标

本文方法用的评价指标包含Dice系数(Dice Similarity Coefficient, DSC),灵敏度(Sensitivity)和豪斯多夫距离(Hausdorff Distance, HD),定义式如式(10)~式(12)所示:

$$\text{DSC} = \frac{2TP}{FN + 2TP + FP} \quad (10)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (11)$$

$$\text{HD} = \max \left\{ \max_{g \in G} \min_{p \in P} \|g - p\|, \max_{p \in P} \min_{g \in G} \|p - g\| \right\} \quad (12)$$

其中, $G$ 、 $P$ 分别为真值(ground truth)和预测值(prediction)的表面点积, $g$ 和 $p$ 属于 $G$ 和 $P$ 的子集。

本文使用HD95,即计算标记图与预测图点集之间的距离,然后按距离从小到大的顺序对这些距离进行排序,取前95%的距离值,并选取其中的最大值作为HD95。HD95可以有效地评价出标记图与预测图边缘轮廓分割效果的同时减小离群值对HD计算的影响。DSC用于评估标记图与预测图之间内部填充的重叠程度,灵敏度代表分割方法正确分割ROI像素的概率。

## 2 实验结果

### 2.1 不同网络模型对比实验

在相同的实验条件下,本文使用U-Net<sup>[22]</sup>、Residual U-Net<sup>[23]</sup>、Attention U-Net<sup>[24]</sup>、Vnet<sup>[25]</sup>和

CA-SegResNet 进行标识椎骨分割对照实验, 选用的各网络模型所使用的超参数与激活函数均与其原文献相同, 不同的模型所采用的数据预处理方式相同, 表 1 为实验使用测试集数据得到的实验结果。

表 1 不同网络模型分割结果对比( $\bar{x}\pm s$ )  
Table 1 Performance comparison among different network models (Mean±SD)

椎体	模型	DSC	灵敏度	HD/mm
颈椎	U-Net	0.920 0±0.024 3	0.917 0±0.045 2	183±58
	Residual U-Net	0.922 2±0.012 1	0.905 3±0.028 6	219±30
	Attention U-Net	0.863 1±0.101 6	0.905 3±0.039 2	213±32
	Vnet	0.910 4±0.051 0	0.926 8±0.039 1	133±78
	CA-SegResNet	0.934 5±0.003 9	0.930 2±0.013 6	7±2
胸椎	U-Net	0.872 6±0.050 9	0.877 2±0.052 6	396±132
	Residual U-Net	0.869 3±0.047 2	0.843 6±0.070 6	373±132
	Attention U-Net	0.894 7±0.027 6	0.941 6±0.024 2	134±95
	Vnet	0.821 1±0.087 8	0.765 0±0.130 5	401±131
	CA-SegResNet	0.918 9±0.009 0	0.920 2±0.015 9	8±2

在颈椎标识椎骨的分割实验中, 相比于 U-Net、Residual U-Net、Attention U-Net、Vnet 的分割结果, 本文算法得到的 DSC 分别提升 0.014 5、0.012 6、0.071 4、0.024 1, 灵敏度分别提升 0.013 2、0.024 9、0.024 9、0.003 4。这说明本文算法在图像内部的分割方面相对于传统网络有较大的提升。在 HD 上相对于 U-Net、Residual U-Net、Attention U-Net、Vnet 这 4 种模型分别缩小 176、212、206、126 mm。HD 是一个在图像边缘分割效果对比非常敏感的指标, 其他模型分割产生的孤立区域会使得 HD 对比本文模型产生较大差距, 但这也表明在图像的边缘分割方面, 本文算法相

对于传统网络有明显的提升。在胸椎标识椎骨的分割实验中, 可以看到本文算法的分割结果也同样优于其他网络模型。

使用不同模型对椎骨分割的视觉效果如图 4 和图 5 所示。图 4 和图 5 分别是颈椎和胸椎标识椎骨的横断面切片展示。通过对比可以看到, U-Net、Residual U-Net 和 Attention U-Net 都容易产生分割图像内部的孔洞, Attention U-Net 和 Vnet 在边界的分割效果不够吻合, 而本文算法的分割图像与标注图像几乎完全保持一致。

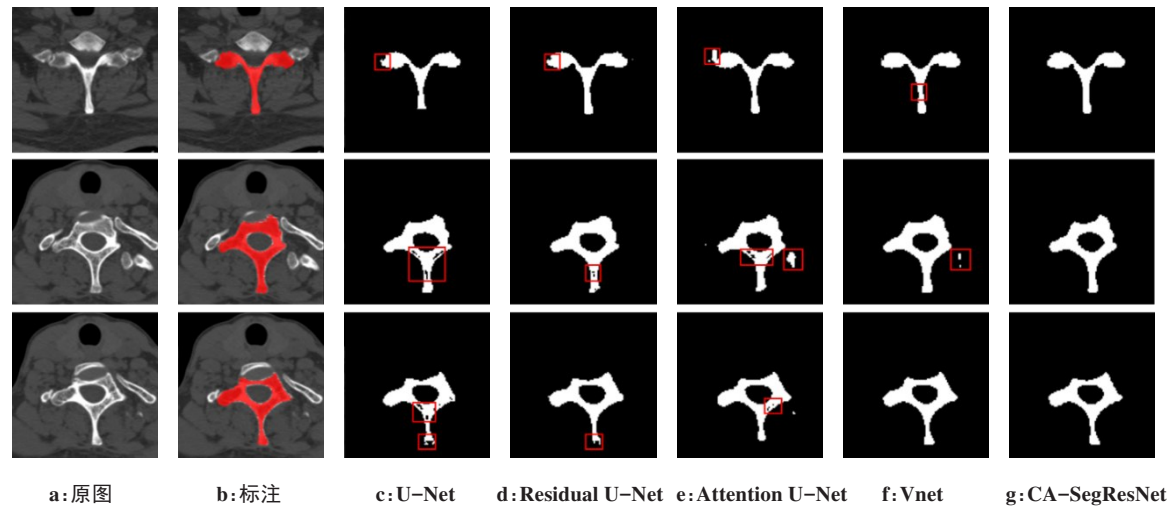


图 4 不同模型对颈椎标识椎骨的分割结果

Figure 4 Segmentation results of different models for the identified cervical vertebra

第 1 行与第 2 行子图分别为样本 A 的不同切片位置, 第 3 行为样本 B 的切片

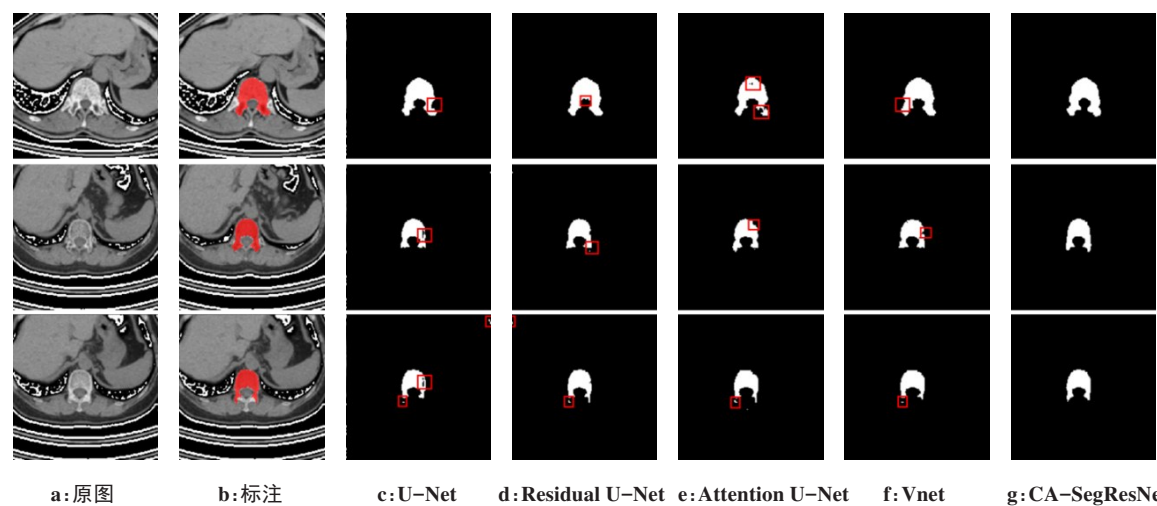


图5 不同模型对胸椎标识椎骨的分割结果

Figure 5 Segmentation results of different models for the identified thoracic vertebra

第1行与第2行子图分别为样本A的不同切片位置,第3行为样本B的切片

2.2 消融实验

为了评估本文算法中增加 SegResNet 模型网络深度和三维坐标注意力模块的有效性,在颈椎标识椎骨的分割实验中,本文设计消融实验来验证实际的提升效果,结果如表 2 所示。其中 Model 1 为原始的 4 层深度的 SegResNet 网络,Model 2 为 5 层深度的 SegResNet 网络,Model 3 为 Model 2 的基础上加上三维坐标注意力模块。从表 2 可以看到,本文算法 (Model 3) 在椎骨方面的效果相对于原始 SegResNet (Model 1) 和仅加深深度的 SegResNet (Model 2),在 DSC、灵敏度和 HD 方面均有明显的提升。为了更加直观地对比消融实验中各网络的分割结果,本文使用 3D Slicer 软件 (version 5.0.3) 来展示消融实验中各网络模型分割结果感兴趣区 3D 对比图。如图 6 所示,可以很明显地看到,原 SegResNet 模型 (Model 1) 无论是在颈椎还是在胸椎标识椎骨的分割结果中,椎骨棘突的部分有较多的分割残缺,在椎体表面也有较多与标注图不贴合的部分。而仅加深深度的 SegResNet (Model 2) 虽然分割结果优于 Model 1,但在胸椎横突与椎体的分割精度上依旧不如 CA-SegResNet (Model 3)。CA-SegResNet 无论是在棘突等突出部位填充,还是椎体、椎孔等表面的分割结果都与标注十分吻合,比原始 SegResNet 和仅加深深度的 SegResNet 都有明显提升,消融实验的实验结果验证了 CA-SegResNet 各模块的有效性。

本实验也尝试过继续加深 CA-SegResNet 的网络深度,发现在 6 层和 7 层的情况下,CA-SegResNet 的分割效果并没有明显提升。

2.3 不同注意力机制对比实验

为了评估本文算法中三维坐标注意力模块对比

表 2 消融实验结果 ( $\bar{x} \pm s$ )

Table 2 Ablation experiment results ( $Mean \pm SD$ )

椎体	模型	DSC	灵敏度	HD/mm
颈椎	Model 1	0.925 4±0.019 8	0.922 0±0.041 9	167±86
	Model 2	0.928 9±0.006 9	0.897 6±0.016 0	39±26
	Model 3	0.934 5±0.003 9	0.930 2±0.013 6	7±2
胸椎	Model 1	0.861 8±0.015	0.867 8±0.105 2	294±174
	Model 2	0.890 3±0.019 2	0.920 1±0.021 8	12±10
	Model 3	0.918 9±0.009 1	0.920 2±0.015 8	8±2

空间注意力模块、通道注意力模块和卷积注意力模块 (Convolutional Block Attention Module, CBAM) 的优异性<sup>[21]</sup>,本文设计不同注意力机制的对比实验,用于验证三维坐标注意力模块的优越性,结果如图 7 所示。

图 7 展示在训练过程中 SegResNet 使用不同注意力模块的训练集 loss 和在验证集 DSC 随 epoch 的变化曲线。loss\_tr\_3D-CA、loss\_tr\_ChannelAtt、loss\_tr\_SpatialAtt 和 loss\_tr\_CBAM 分别代表三维坐标注意力模块,通道注意力模块,空间注意力模块和 CBAM 在训练过程中的训练集 Loss 变化曲线。dice\_val\_3D-CA、dice\_val\_ChannelAtt、dice\_val\_SpatialAtt 和 dice\_val\_CBAM 分别代表三维坐标注意力模块,通道注意力模块,空间注意力模块和 CBAM 在训练过程中的验证集 DSC 变化曲线。从图中可以看出,相比其他注意力机制,三维坐标注意力在训练过程中训练集 loss 具有更快的下降速度,验证集的 DSC 增长更加稳定,最终 DSC 也更高。

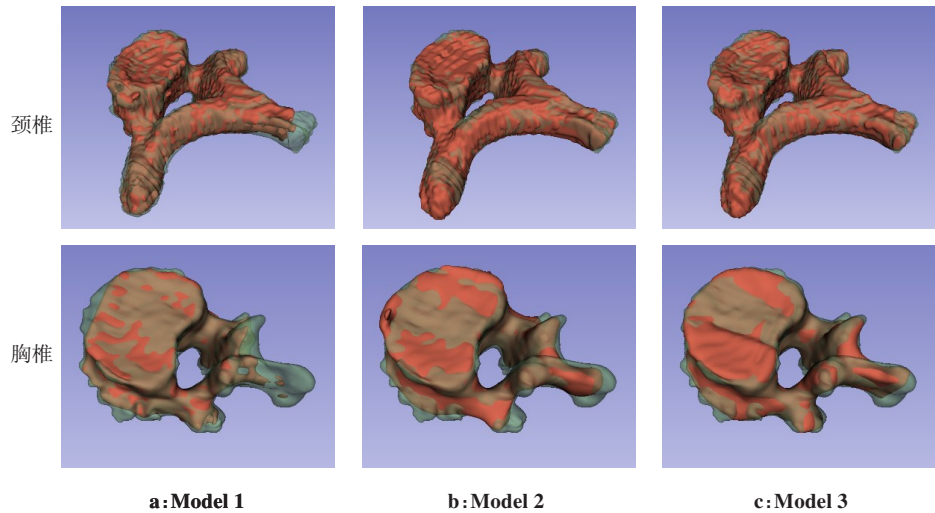


图 6 消融实验中各网络模型分割结果的感兴趣区 3D 可视化对比

Figure 6 Three-dimensional visualizations of regions of interest segmented by different models in the ablation experiment

图中绿色半透明的部分为医生标注, 红色不透明的部分为网络分割结果

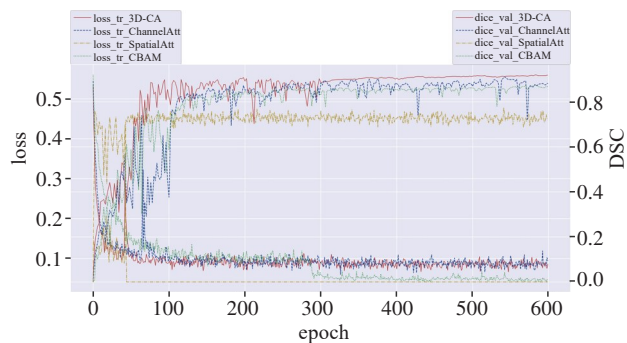


图 7 不同注意力模块训练对比图

Figure 7 Comparison of training with different attention modules

### 3 临床实验

#### 3.1 标识椎骨自动分割软件

本研究设计并开发了标识椎骨自动分割软件来供合作医院的临床实际测试与使用,软件界面如图 8 所示。

该软件具有 CT 图像展示与自动分割两大主要功能,其中自动分割算法所使用的神经网络模型为本文所提出并训练的模型,该软件能够对 CT 脊椎图像的第 7 节颈椎和第 12 节胸椎进行自动分割,并显示 3 个断面的图像与 3D 模型。

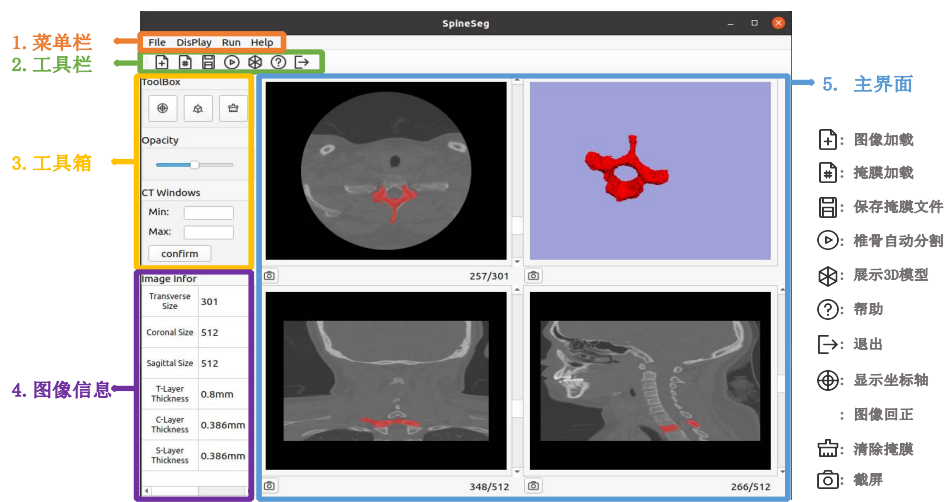


图 8 软件界面及功能

Figure 8 Software interface and functions

#### 3.2 临床实验结果

应用该软件对 5 个临床病例样本进行分割,再经

过有经验的医生定位确认和边界分割复核,分割结果如表 3 所示。结果表明,CA-SegNet 在临床实验

中,颈椎标识椎骨分割平均DSC为0.930 7,胸椎标识椎骨分割平均DSC为0.915 7。所有样本的平均DSC为0.923 2。

表3 临床实验结果(DSC)  
Table 3 Clinical trial results (DSC)

椎体	样本1	样本2	样本3	样本4	样本5	平均
颈椎	0.929 8	0.926 1	0.923 4	0.930 3	0.943 8	0.930 7
胸椎	0.915 6	0.914 7	0.920 4	0.913 4	0.914 6	0.915 7

4 结 语

本文针对标识椎骨分割问题,提出一种基于SegResNet网络的改进型分割算法CA-SegResNet。该算法的网络模型是一种编码解码结构,能够对图像进行多尺度特征提取与融合,并通过三维坐标注意力模块建立起不同空间方向的长距离依赖关系。实验结果表明,相对于原SegResNet网络和U-Net等其他主流分割算法,本文算法具有更高的分割精度。基于颈椎和胸椎标识椎骨的分割结果,医生可以实现脊柱快速分段,定位其他椎骨和病灶区域。相比通过人工标注整个脊柱并进行分割训练,本研究可以节省大量勾画标注的工作量和时间成本。

在未来工作中,将在多种数据集以及不同椎骨实验上验证分割效果,并探索在标识颈椎和标识胸椎的分割结果上,使用半监督学习方法进一步分割整个胸椎区域。

【参考文献】

[1] Parenteau CS, Lau EC, Campbell IC, et al. Prevalence of spine degeneration diagnosis by type, age, gender, and obesity using Medicare data[J]. Sci Rep, 2021, 11(1): 5389.

[2] Jung H. Basic physical principles and clinical applications of computed tomography[J]. Prog Med Phys, 2021, 32(1): 1-17.

[3] Ikhsan IA, Hussain A, Zulkifley MA, et al. Automated vertebrae extraction using watershed segmentation and tree-based modelling approach[J]. J Fiber Bioeng Inform, 2015, 8(3): 547-555.

[4] Suzani A, Rasoulia A, Seitel A, et al. Deep learning for automatic localization, identification, and segmentation of vertebral bodies in volumetric MR images[C]//Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling. Bellingham, WA, USA: SPIE, 2015: 941514.

[5] Huang JY, Jian FZ, Wu H, et al. An improved level set method for vertebra CT image segmentation[J]. Biomed Eng Online, 2013, 12: 48.

[6] Lim PH, Bagci U, Bai L. A robust segmentation framework for spine trauma diagnosis[C]//Proceedings of the Computational Methods and Clinical Applications for Spine Imaging. Cham: Springer International Publishing, 2014: 25-33.

[7] Forsberg D. Atlas-Based registration for accurate segmentation of thoracic and lumbar vertebrae in CT data[M]//Yao JH, Glocker B, Klinder T, et al. Recent advances in computational methods and clinical applications for spine imaging. Cham: Springer International Publishing, 2015: 49-59.

[8] Wang YN, Yao JH, Roth HR, et al. Multi-atlas segmentation with joint label fusion of osteoporotic vertebral compression fractures on CT[C]//Proceedings of the Computational Methods and Clinical Applications for Spine Imaging. Cham: Springer International Publishing, 2016: 74-84.

[9] Janssens R, Zeng GD, Zheng GY. Fully automatic segmentation of lumbar vertebrae from CT images using cascaded 3D fully convolutional networks[C]//2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Piscataway, NJ, USA: IEEE, 2018: 893-897.

[10] 师文博, 杨环, 西永明, 等. 基于自注意力的双通路全脊柱X光图像分割模型[J]. 中国医学物理学杂志, 2022, 39(11): 1385-1392. Shi WB, Yang H, Xi YM, et al. Self-attention based dual pathway network for spine segmentation in X-ray image[J]. Chinese Journal of Medical Physics, 2022, 39(11): 1385-1392.

[11] Nazir A, Cheema MN, Sheng B, et al. ECSU-net: an embedded clustering sliced U-net coupled with fusing strategy for efficient intervertebral disc segmentation and classification[J]. IEEE Trans Image Process, 2022, 31: 880-893.

[12] 黄昆, 张俊华, 普钟. 基于深度学习的脊椎CT图像分割[J]. 电子测量技术, 2022, 45(20): 151-159. Huang K, Zhang JH, Pu Z. Vertebra CT image segmentation based on deep learning[J]. Electronic Measurement Technology, 2022, 45(20): 151-159.

[13] 刘晶, 徐皓, 崔欣欣, 等. 基于多尺度边缘分割与混合注意力机制的脊柱CT图像分割[J]. 中国医学物理学杂志, 2024, 41(4): 463-471. Liu J, Xu H, Cui XX, et al. Spine CT image segmentation based on multi-scale boundary segmentation and hybrid attention mechanism [J]. Chinese Journal of Medical Physics, 2024, 41(4): 463-471.

[14] Myronenko A. 3D MRI brain tumor segmentation using autoencoder regularization[C]//Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Cham: Springer International Publishing, 2019: 311-320.

[15] Hou QB, Zhou DQ, Feng JS. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 13708-13717.

[16] Wu YX, He KM. Group normalization[J]. Int J Comput Vis, 2020, 128 (3): 742-755.

[17] Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines[C]//Proceedings of the 27th International Conference on International Conference on Machine Learning. Madison, WI, USA: Omnipress, 2010: 807-814.

[18] He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2016: 770-778.

[19] 任欢, 王旭光. 注意力机制综述[J]. 计算机应用, 2021, 41(S1): 1-6. Ren H, Wang XG. Review of attention mechanism [J]. Journal of Computer Applications, 2021, 41(S1): 1-6.

[20] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 7132-7141.

[21] Woo S, Park J, Lee JY, et al. CBAM: convolutional block attention module [C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.

[22] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.

[23] Zhang ZX, Liu QJ, Wang YH. Road extraction by deep residual U-Net [J]. IEEE Geosci Remote Sens Lett, 2018, 15(5): 749-753.

[24] Oktay O, Schlemper J, Le Folgoc L, et al. Attention U-net: learning where to look for the pancreas[EB/OL]. (2018-05-20). <https://arxiv.org/abs/1804.03999>.

[25] Milletari F, Navab N, Ahmadi SA. V-net: fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV). Piscataway, NJ, USA: IEEE, 2016: 565-571.

(编辑:薛泽玲)