

doi: 10.3969/j.issn.1674-1242.2025.01.003

基于局部窗口迭代的两阶段三维脊柱实例分割方法

袁孟

(上海理工大学健康科学与工程学院, 上海 200093)

【摘要】在医学影像领域,准确快速的脊柱实例分割对于早期诊断和治疗规划至关重要。针对以往脊柱实例分割方法存在的效率低和分割性能欠佳问题,该文提出了一个创新的基于局部窗口迭代的两阶段三维脊柱实例分割方法。在第一阶段通过 nn-UNet 模型进行全脊柱的语义分割,以粗略识别出所有相关结构的位置和形态,为第二阶段提供必要的上下文信息。在第二阶段则使用一个创新的基于局部窗口迭代的实例分割方法,基于语义分割结果获得的脊柱单元质心进行窗口划分,并对所有窗口迭代进行实例分割以精确分割每个脊椎单元,最终经过后处理重构获取完整的脊柱实例分割结果。该文创新性地实例分割任务定义为分割局部图像块的中心脊柱单元及上下相邻的脊柱单元,通过将复杂的实例分割任务简单化,有效降低了模型的训练难度,提高了分割准确率。实验结果表明,该文所提方法在 SPIDER 数据集和 CSI2014 数据集上的表现超过了以往的脊柱实例分割方法,展示了其在处理复杂脊柱结构中的优越性和潜力。

【关键词】深度学习; 三维脊柱实例分割; 窗口迭代; 多尺度注意力融合

【中图分类号】 TP391.41

【文献标志码】 A

文章编号: 1674-1242 (2025) 01-0017-08

Two-Stage 3D Spine Instance Segmentation Method Based on Local Window Iteration

YUAN Meng

(School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

【Abstract】In the field of medical imaging, accurate and rapid spinal instance segmentation is essential for early diagnosis and treatment planning. Aiming at addressing the problems of low efficiency and poor performance of previous spinal instance segmentation methods, an innovative two-stage 3D spinal instance segmentation method based on local window iteration is proposed in this paper. In the first stage, semantic segmentation of the whole spine is carried out by nn-UNet model to roughly identify the positions and shapes of all relevant structures, providing necessary contextual information for the instance segmentation stage. In the second stage, an innovative case segmentation method based on local window iteration is used to partition the window through the centroid of the spinal unit obtained based on the semantic segmentation results, and all window iteration instances are segmented to accurately segment each spinal unit. Finally, the complete spinal instance segmentation results are obtained through post-processing reconstruction. In this paper, the instance segmentation task is innovatively defined as the segmentation center and the upper and lower adjacent spinal units. By simplifying the complex instance segmentation task, the training difficulty of the model is

收稿日期: 2024-01-06。

作者简介: 袁孟 (2000—), 男, 硕士研究生, 从事医学图像分割研究。

reduced effectively and the segmentation accuracy is improved. The experimental results show that the proposed method outperforms previous spinal segmentation methods on SPIDER and CSI2014 data sets, demonstrating its superiority and potential in dealing with complex spinal structures.

【Key words】 Deep Learning; 3D Spinal Instance Segmentation; Window Iteration; Multi-Scale Attention Fusion

0 引言

脊柱作为人体最重要的结构之一，支撑着人体的整个骨骼系统，同时保护脊髓不受伤害。随着现代化社会生活节奏的加快，年轻人需要长期伏案工作，老年人长期经受超负荷工作，容易患腰椎间盘突出、脊柱侧弯等常见的脊柱类疾病^[1-3]。因此，及时且早期发现脊柱病变，有利于疾病的诊断、治疗规划及疗效评估。随着科学技术的快速发展，医疗影像和计算机开始更加紧密地合作，成像技术变得更加丰富多样。例如，计算机断层扫描（Computed Tomography, CT）和核磁共振成像（Magnetic Resonance Imaging, MRI）^[4]能够快速扫描人体内部的脊柱组织结构。目前，临床上依赖主治医生人工读片，受限于医生的临床经验和专业知识，读片需要耗费大量的精力且存在一定的主观性。

随着人工智能的快速发展，其被应用于脊柱 MRI 领域，通过高精度的脊柱分割，辅助医生更好地理解脊柱的形态学变化，识别出脊柱的畸形、损伤或疾病，从而为患者提供更准确的治疗方案。目前，图像分割算法分为基于颜色和纹理的传统图像分割算法、基于深度学习的图像分割算法。传统图像分割算法又分为阈值法、边缘分割法和聚类分割法等。阈值法^[5]主要依据图像中物体的像素值不同，并设置灰度阈值，从而将图像分割为目标和背景。边缘分割法^[6]计算图像像素的梯度幅度和方向，选出梯度较大的区域作为边缘候选区域，最后采用非极大值抑制和阈值处理分割出最终的边缘形状。聚类分割法^[7]旨在将图像中的像素分组成具有相似特征的多个聚类中心，侧重于根据像素之间的相似性进行划分。

传统图像分割算法具有计算简单、易于实现和计算效率较高的优点，并且不依赖大量的训练数据。但是，传统图像分割算法容易对噪声敏感，导致分割精度降低，需要手动调参，具有一定的主观性和

人工消耗。此外，传统图像分割方法难以处理高度复杂或重叠的场景。

基于深度学习的图像分割方法^[8,9]通过卷积核的自适应学习能力和端到端的训练方式，显著地改善了图像分割的质量和效果，并成为目前医学图像处理的主流算法。Hornig 等^[10]首先使用水平和垂直强度的投影直方图定义脊柱存在的区域，然后采用渐进阈值方法检测椎体的具体位置，最后使用 U-Net、Dense-Net 和 Residual-Net 3 种不同的算法进行脊柱分割。Dubost 等^[11]使用两个 U-Net 分割脊柱区域和脊柱中心线，Tran 等^[12]提出了基于 U-Net 变体的 MB-Net 算法，用于脊柱和骨头的语义分割。Cheng 等^[13]在 U-Net 的基础上提出了多尺度聚合模块，用于捕获不同尺度的特征信息，并在特征提取时引入了注意力机制。Lessmann 等^[14]提出了迭代分割方法，使用全卷积神经网络依次分割和标记脊柱骨，对所有脊柱骨进行标记，从而实现不同椎骨的分割。

上述算法针对脊柱分割任务取得了一定的效果，但是仍然存在一些不足和缺点。首先，上述算法基于 2D 脊柱图像进行语义分割，但医学图像通常以 3D 图像的形式存在，导致 2D 分割模型在空间提取能力上存在不足，并且无法捕获脊柱复杂结构的全部细节。其次，上述算法侧重于脊柱关节的语义分割，将所有脊柱关节识别成一个类别，但实际情况中脊柱存在不同情况的病变，需要对不同的椎骨进行分割并分类。最后，大多数算法^[12-14]对全图像进行滑动窗口迭代分割，然后将其结合再进行处理，从而实现实例级脊柱分割，但是容易忽略全局信息，导致出现欠分割和错误分割，并且分割效率较低。

为了解决上述问题，本文提出了一个基于 nn-UNet 模型和局部窗口迭代策略的两阶段三维脊柱实例分割方法。所提方法分为两个阶段：全脊柱语

义分割阶段和脊柱实例分割阶段。在全脊柱语义分割阶段，利用 nn-UNet 模型对整个脊柱进行粗略的分割，快速识别出脊柱的大致形态和位置，为下一阶段提供必要的上下文信息。在实例分割阶段，本文提出了基于局部窗口迭代的实例分割策略。首先，计算语义分割结果中所有脊柱单元的质心，并围绕这些质心切割固定大小的图像块。然后，通过改进的基于多尺度 3D-UNet 的实例分割模型对所有的图像块进行实例分割，获取每个图像块的实例分割结果。最后，通过基于 Dice 排序的融合重构获取最终的脊柱实例分割结果。

本文的创新点在于提出了一个两阶段全脊柱实例分割框架，有效提高了分割精度并减少了由于结构重叠和复杂性带来的分割错误。同时，本文将复杂的全局脊柱实例分割任务简化为分割局部图像块的中心脊柱单元及上下相邻的脊柱单元，有效降低了模型的训练难度并提高了分割准确率。此外，本文提出了一个基于多尺度注意力融合的 U-Net 实例分割模型，通过多尺度注意力融合单元，使模型能够更有效地集中处理图像中的关键特征，增强其对三维空间信息的捕获能力。

1 方法设计与实现

1.1 两阶段三维脊柱实例分割方法

本文提出的两阶段三维脊柱实例分割方法由全脊柱语义分割阶段和基于局部窗口迭代的脊柱实例分割阶段组成，如图 1 所示。在全脊柱语义分割阶段，通过 nn-UNet^[15] 语义分割模型对完整的 3D 脊柱图

像进行语义分割，获得全脊柱掩码图像，以识别出图像中所有相关结构的大致位置和形态信息。在基于局部窗口迭代的脊柱实例分割阶段，首先计算第一阶段中获取的全脊柱掩码图像中每个脊椎主体的质心位置，并围绕每个质心裁切一个固定大小窗口的图像块。然后通过 3D-UNet 实例分割模型分割每个图像块中的顶部脊柱单元、中心脊柱单元和底部脊柱单元 3 个实例。最后基于所有窗口图像块的实例分割结果进行合并重构，以获取完整的实例级脊柱分割结果。

1.2 全脊柱语义分割

全脊柱语义分割阶段主要负责通过三维脊柱分割模型对原脊柱图像进行分割，以获得全脊柱分割掩码图像。在该阶段，首先将所有图像重新采样到相同的分辨率，确保每个图像的方向和大小一致，以便进行后续处理和分割。然后将预处理过的原始脊柱图像输入训练好的脊柱语义分割模型中，分割得到全脊柱分割掩码图像。本文采用 nn-UNet^[15] 模型作为该阶段的语义分割模型，这是一个自适应的网络结构选择和训练策略，专门为医学图像分割任务设计，能够根据数据的特点自动调整网络架构和训练参数，从而优化分割性能。同时，在训练阶段，由于 MRI 图像的尺寸可能非常大，本文采用基于小图像块 (Patch) 的方法进行训练和推理，即将大图像分割成多个重叠的小图像块，每次并行处理多个小图像块，然后将结果拼接成完整的脊柱分割掩码图像，这一策略有效减少了对计算资源的需求。

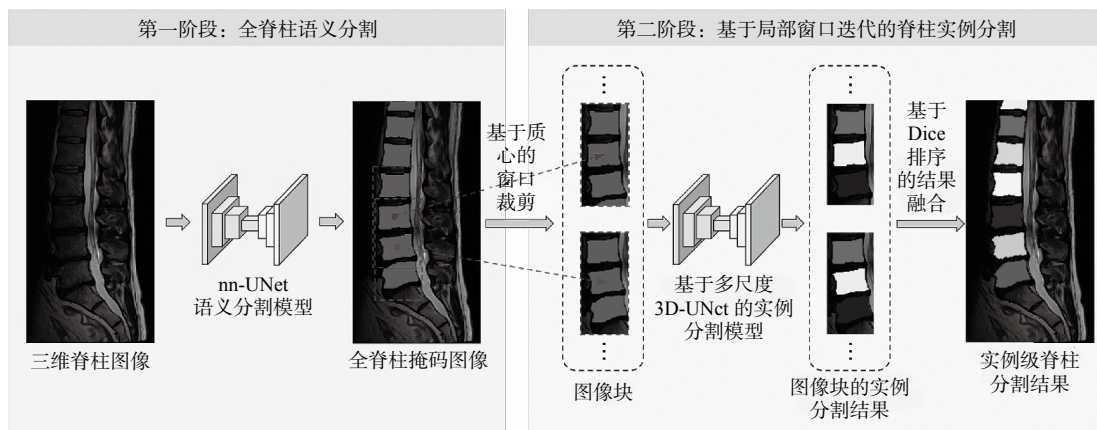


图 1 基于局部窗口迭代的二维脊柱实例分割方法结构

Fig.1 Structure of two-stage 3D spinal instance segmentation method based on local window iteration

通过这一阶段的处理，本文所提方法能够有效地从复杂的MRI图像中提取出完整脊柱的粗略结构信息，为后续的细致实例分割提供了基础。该阶段的过程可由如下公式表示。

$$\begin{cases} [P_0, P_1, \dots, P_n] = \text{Split}(I) \\ \text{Mask}_i = \text{nn} - \text{UNet}(P_i), \quad i \in 0, 1, \dots, n \\ \text{Mask} = \text{Reconstruct}([\text{Mask}_0, \text{Mask}_1, \dots, \text{Mask}_n]) \end{cases}$$

式中， I 表示原始脊柱图像； P_i 和 Mask_i 分别表示第*i*个小图像块和其对应的脊柱语义分割掩码图像。 Mask 表示重构之后的完整脊柱语义分割掩码图像； Split 和 Reconstruct 分别表示图像分块操作与图像块重构操作。

1.3 基于局部窗口迭代的脊柱实例分割

在本文提出的两阶段分割方法中，第二阶段的实例分割旨在将第一阶段得到的语义分割结果进一步细分为具体的实例，即识别图像中每个独立脊椎的具体位置和边界。具体地，在实例分割前，首先通过计算第一阶段获得的语义分割掩码图像中每个脊椎结构的质心，确定每个脊椎的大致中心位置，用于确定后续处理的窗口裁剪区域。基于此，围绕每个脊椎的中心位置，切割出固定大小窗口的脊柱掩码图像块。其中，每个窗口的大小均被设计为足以包含至少3个脊椎（中心脊柱单元及其相邻的上下脊柱单元），以便模型在足够的上下文中进行准确的分割。然后对每个窗口应用训练好的基于多尺度3D-UNet的实例分割模型，得到窗口内各脊椎实例的分割结果。由于每个脊椎会在多个窗口中被重复识别和分割，因此每个实例将有多个预测结果。最后通过计算不同图像块的预测结果之间的Dice分数并按照由高到低的顺序排序，确定每个脊柱单元实例，将多个窗口中的预测结果融合为单一的、连贯的脊椎实例分割图。值得注意的是，本文通过将实例分割任务定义为分割中心脊柱单元及其相邻的上下脊柱单元，使模型能够更有效地利用局部上下文信息，有助于模型学习和理解脊椎结构之间的相对位置与关系，从而提高分割的准确性。基于局部窗口迭代的脊柱实例分割阶段的过程可由如下公式表示。

$$\begin{cases} \{MP_0, MP_1, \dots, MP_m\} = \text{SSplit}(\text{Mask}) \\ \text{Inst}M_i = \text{3D-UNet}(MP_i), \quad i \in 0, 1, \dots, m \\ \text{Inst}M = \text{DiceFusion}([\text{Inst}M_0, \text{Inst}M_1, \dots, \text{Inst}M_m]) \end{cases}$$

式中， Mask 表示第一阶段获取的完整脊柱语义分割掩码图像； MP_i 和 $\text{Inst}M_i$ 分别表示第*i*个脊柱掩码图像块和其对应的脊椎实例分割图； SSplit 和 DiceFusion 分别表示单个脊柱掩码图像块的分割处理过程与基于Dice排序的结果融合操作； $\text{Inst}M$ 表示最终的脊椎实例分割结果。

1.4 基于多尺度3D-UNet的实例分割模型

在基于局部窗口迭代的脊柱实例分割中，本文提出了一个基于多尺度3D-UNet的实例分割模型，该模型在3D-UNet^[16]的基础上进行优化改进。3D-UNet通过采用三维卷积层替换二维卷积层，有效地深入挖掘了沿体积深度方向的空间信息，这对理解复杂的组织结构至关重要。为了使网络更专注于有价值的特征，抑制非关键信息，本文提出了多尺度3D-UNet。相较于3D-UNet，多尺度3D-UNet加入了一个关键的多尺度注意力融合单元（Multi-scale Attention Fusion Unit, MAFU），以根据不同的分辨率自适应地选择特征。多尺度3D-UNet网络结构如图2所示。多尺度3D-UNet由具有跳跃连接的编码器、解码器和MAFU组成。MAFU结构如图3所示。与U-Net相似，MAFU的编码器部分由4个分辨率层级构成，每层都包含两个卷积块。每个卷积块包含一个 $3 \times 3 \times 3$ 卷积，随后是批量归一化（BatchNorm）和一个ReLU激活函数。这种设计能有效地提取和整合体积数据中的空间特征，捕捉到关键的局部上下文信息，同时增强模型对于复杂空间结构表征的学习能力。同层级的两次卷积后是一个步长为2的 $2 \times 2 \times 2$ 最大池化层（Max Pooling）。解码器部分则开始于一个 $2 \times 2 \times 2$ 的转置卷积，其步长为2，用于将特征图的大小增加一倍以逐步恢复分辨率。随后将上采样的特征图与经过MAFU的跳跃连接特征进行拼接，以帮助解码器恢复在特征提取过程中丢失的细节信息。在这之后，再次应用两个 $3 \times 3 \times 3$ 卷积以进一步细化通过上卷积增大的特征图。

为了根据不同的分辨率自适应地选择特征，模

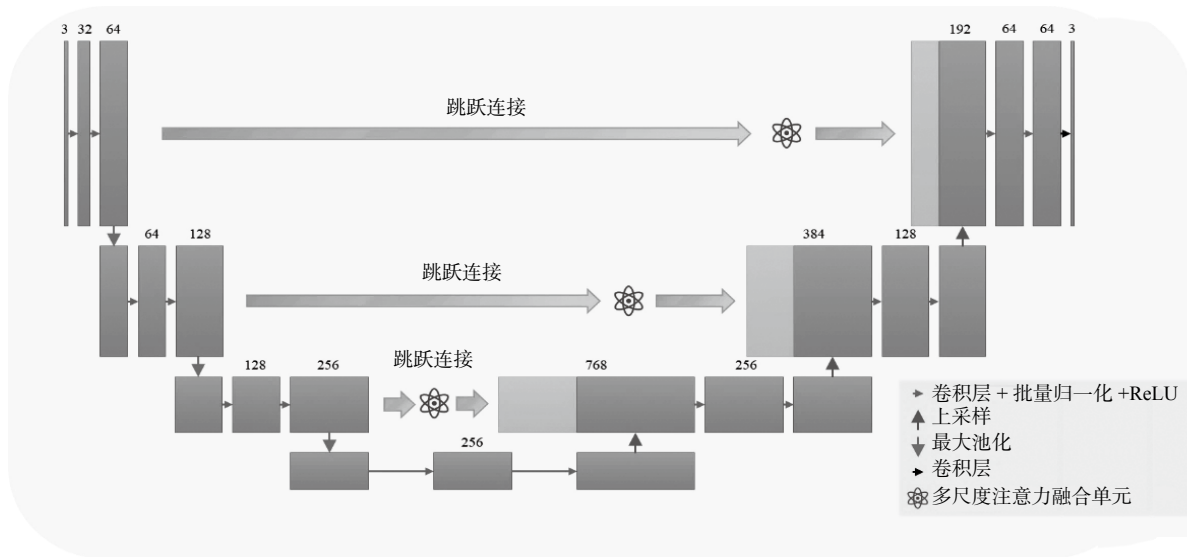


图2 多尺度 3D-UNet 网络结构
Fig.2 Multi-scale 3D-UNet network structure

型采用了3种具有不同感受野的卷积来提取特征，分别为 $1 \times 1 \times 1$ 逐点卷积、标准的 $3 \times 3 \times 3$ 卷积和扩张率为2的 $3 \times 3 \times 3$ 扩张卷积。逐点卷积专注于提取不同维度的特征，扩张卷积则扩大了感受野以融入更广泛的上下文信息。当跳跃连接传递的特征 f_{in} 传入MAFU后，这3种卷积并行工作以提取多层次的空间特征，并通过批量归一化提高网络训练的稳定性。随后，将这3个不同尺度的特征拼接，并采用ReLU激活函数统一处理以确保非线性的引入。接着，利用另一个 $1 \times 1 \times 1$ 逐点卷积来对拼接特征进行筛选，突出有利于任务的特征并抑制无关或干扰信息。通过Sigmoid函数激活后，将得到的特征与 f_{in} 逐元素相乘，这一注意力加权步骤使特征的注意力集中于最有价值的部分。最后，将强化的

特征与 f_{in} 进行残差连接，找回丢失的细节，同时融合关键特征，得到最终输出的特征 f_{out} 。这种多尺度自适应特征融合的方法充分发挥了不同感受野特征的优势，并通过筛选和融合机制有效提升了网络对复杂数据的表达能力和识别精度。MAFU的计算公式为

$$f_{concat} = \sigma_1(\text{Concat}\{\text{BN}\{\text{Point wiseConv}(f_{in})\}, \text{BN}\{\text{OrdinaryConv}(f_{in})\}, \text{BN}\{\text{DilationConv}(f_{in})\}\}) \quad (1)$$

$$f_{out} = f_{in} \sigma_2(\text{Point wiseConv}(f_{concat})) + f_{in} \quad (2)$$

式中， f_{in} 表示跳跃连接输入特征； f_{concat} 表示拼接后的特征； f_{out} 表示MAFU的输出特征， σ_1 和 σ_2 分别表示ReLU、Sigmoid激活操作。

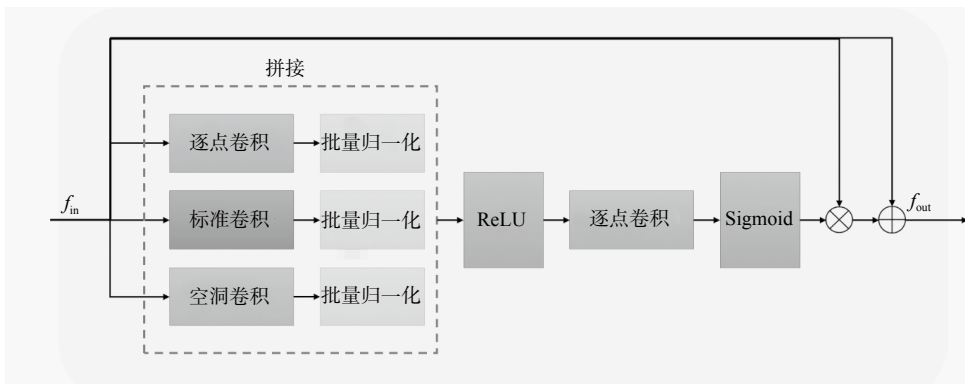


图3 MAFU 结构
Fig.3 Multi-scale Attention Fusion Unit

2 实验

为验证本文提出的基于局部窗口迭代的两阶段三维脊柱实例分割方法的有效性,在公开数据集上对本文所提方法与其他现有的图像分割方法进行了详细的对比分析,对比的图像分割方法具体如下。

(1) DDS: Zhou 等^[17]提出的基于深度监督的判别全卷积网络,对显著的形状变化和低对比度医学图像的语义分割任务具有较高的鲁棒性。

(2) nn-UNet: Graaf 等^[18]提出的迭代数据标注方法,在数据集的一小部分上训练 nn-UNet 和基准分割算法,对剩余部分进行半自动分割。

(3) MB-Net: Ye 等^[19]提出的一种基于自适应多尺度特征聚合和边界感知的侧脑室分割卷积神经网络,有效解决了磁共振图像上目标区域明显不同、边界模糊及形状不规则的问题。

(4) R2U-Net: Alom 等^[20]提出的一种基于 U-Net 模型的循环残差卷积神经网络,通过循环残差卷积层更有效地训练深度架构,确保分割任务具有更好的特征表示,在医学图像分割中具有更好的性能。

2.1 实验环境及参数设置

本实验训练与测试的硬件环境为:配备 Xeon(R) Platinum 8362 CPU;内存为 48GB;显卡为 24 GB 内存 NVIDIA RTX 3090。软件环境为:Linux 操作系统,深度学习框架 Pytorch2.0, CUDA11.4 加速学习,编程语言为 Python3.10。训练阶段使用 ADAM 算法优化器,共训练 100 个周期,初始学习率为 1×10^{-4} 。

2.2 数据集和评价指标

本实验数据集使用 SPIDER 数据集^[18]和 CSI2014 数据集^[21]。

SPIDER 数据集是一个大型的公开可用的多中心腰椎磁共振成像数据集,其中包含椎骨、椎间盘和椎管的参考分割。该数据集包括 218 名有腰痛病史的患者的 447 个矢状 T1 和 T2 MRI 系列,本实验随机挑选了 179 名患者作为训练集,剩下 39 名患者作为测试集。

CSI2014 数据集是一个包括胸椎和腰椎的规范化 CT 图像与参考分割的数据集,由加州大学欧文分校医学院放射科学系获取,发布自 CSI2014 研讨会分割挑战赛。该数据集采集了 20 名测试人员

的数据,其中 10 名年轻人(16~25 岁)的扫描图作为训练集,5 名年轻人(20~35 岁)及 5 名椎体压缩性骨折患者的扫描图作为测试集。为了评估不同图像分割方法的分割性能,本实验使用了距离相似性系数(Dice Similarity Coefficient, DSC)和平均对称表面距离(Average Symmetric Surface Distance, ASSD)作为评价指标。其中,DSC 越大,表明结果越好;ASSD 越小,表明结果越好。在实例分割时,交并比(Intersection over Union, IoU) ≥ 0.5 的实例被认为分割正确。

2.3 实验结果与分析

表 1 和表 2 展示了不同图像分割方法在 SPIDER 数据集、CSI2014 数据集上的语义分割与实例分割结果。从表中可以清楚地观察到,在执行语义分割任务时,本文所提方法在 DSC 指标上优于其他对比方法,尤其是在 SPIDER 数据集上,本文所提方法的 DSC 数值达到了 0.936 ± 0.024 ,明显优于 DDS 方法和 nn-UNet 方法;在执行实例分割任务时,相较于其他对比方法,本文所提方法在 SPIDER 数据集和 CSI2014 数据集上均取得了最高的评价指标数值,本文所提方法在 SPIDER 数据集和 CSI2014 数据集上的 DSC 数值分别为 0.928 ± 0.019 、 0.923 ± 0.024 ,ASSD 数值分别为 0.258 ± 0.151 和 0.267 ± 0.146 ,显著优于 DDS 方法和 nn-UNet 方法。综上所述,本文所提方法在语义分割任务和实例分割任务上均取得了最优效果,能够很好地适用于医学图像的脊柱分割任务。

表 1 本文所提方法与对比方法在 SPIDER 数据集上的评价指标对比
Tab.1 Comparison of the evaluation indexes of the proposed method and the comparison method on the SPIDER dataset

方法	DSC \uparrow	ASSD \downarrow
语义分割		
MB-Net	0.891 ± 0.051	—
R2U-Net	0.881 ± 0.035	—
DDS	0.921 ± 0.021	—
nn-UNet	0.926 ± 0.029	—
本文所提方法	0.936 ± 0.024	—
实例分割		
DDS	0.903 ± 0.120	0.534 ± 0.355
nn-UNet	0.906 ± 0.104	0.529 ± 0.435
本文所提方法	0.928 ± 0.019	0.258 ± 0.151

表2 本文所提方法与对比方法在CSI2014数据集上的评价指标对比
Tab.2 Comparison of the evaluation indexes of the proposed method and the comparison method on the CSI2014 dataset

方法	DSC ↑	ASSD ↓
语义分割		
MB-Net	0.879 ± 0.022	—
R2U-Net	0.865 ± 0.012	—
DDS	0.911 ± 0.025	—
nn-UNet	0.915 ± 0.039	—
本文所提方法	0.920 ± 0.026	—
实例分割		
DDS	0.892 ± 0.118	0.544 ± 0.259
nn-UNet	0.896 ± 0.104	0.526 ± 0.155
本文所提方法	0.923 ± 0.024	0.267 ± 0.146

图4展示了本文所提方法和DDS方法、nn-UNet方法的实例分割结果。从图中可以清楚地观察到，

DDS方法和nn-UNet方法存在明显的实例分割错误结果，错误结果中不同实例标签的混淆较为明显，主要原因在于这些标签具有相同的语义结构，导致在分割时难以区分。而本文所提方法在所有的分割结果中均未发现此类错误，这得益于本文所提方法利用先进行语义分割再进行局部窗口实例分割的策略，有效避免了实例标签混淆的问题。此外，在原始图像和真实标签中脊柱关节并非呈现规则的矩形形状，易受到其他骨关节的侵蚀。从图4中可以清楚地观察到，相较于其他对比方法，本文所提方法输出的实例分割图中每个脊柱关节的轮廓和真实标签接近，而DDS方法和nn-UNet方法输出的脊柱关节轮廓存在不同程度的扩大与缩小，表明本文所提方法能够准确地定位脊柱关节的区域和大小。

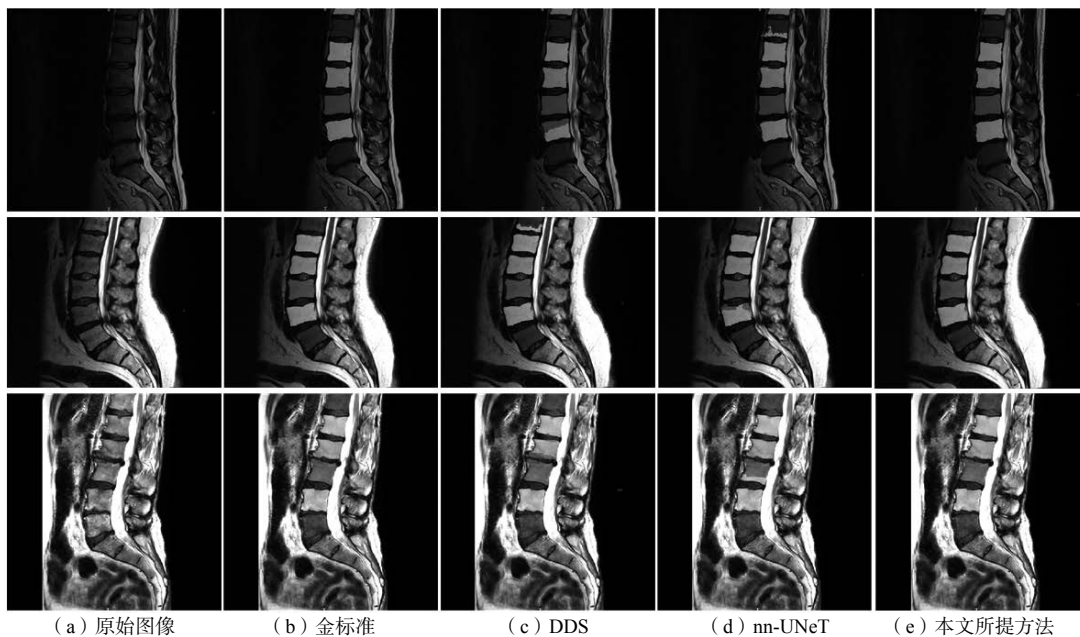


图4 实例分割结果

Fig.4 Instance segmentation result

3 总结

本文提出了一种基于局部窗口迭代的两阶段三维脊柱实例分割方法，旨在解决以往脊柱分割方法的局限性，如分割效率较低、分割精度不足及处理高度复杂或重叠脊柱场景的能力有限。基于局部窗口迭代的两阶段三维脊柱实例分割方法是一个两阶段的分割方法。在第一阶段，使用nn-UNet语义模

型对脊柱图像进行三维语义分割，以有效、快速地定位和识别图像中的脊椎结构，提供丰富的上下文信息以供第二阶段使用。在第二阶段，本文采用了基于局部窗口迭代的实例分割策略，围绕每个脊椎的质心裁切固定大小的图像块，并使用实例分割模型进行精确分割，显著提高了分割精度。其中，本文定义了一个新颖的实例分割任务，即实例分割局

部图像块的中心脊柱单元及上下相邻的脊柱单元，相比以往复杂的全图像实例分割任务，有效地简化了任务的复杂度，从而降低了模型的训练难度并提高了分割准确率。此外，本文提出了基于多尺度 3D-UNet 的实例分割模型，进一步提升了模型对脊椎结构复杂性的理解，使分割过程更加精确和稳定。实验结果证明，本文所提方法在 SPIDER 数据集和 CSI2014 数据集上的表现显著优于现有的图像分割方法，尤其是在实例分割精度和识别质量方面。

本文所提方法先进行脊柱语义分割，再进行实例分割，从而实现了不同类别脊柱的分割和分类。相较于端到端的脊柱分割算法，本文所提方法在实时性和效率方面存在不足。未来将探索新的实例分割策略，进一步简化分割任务，特别是在处理结构高度复杂或重叠的脊柱图像时，提升模型的适用性和稳定性。

参考文献

- [1] AABY A, RAVN S L, KASCH H, *et al.* The associations of acceptance with quality of life and mental health following spinal cord injury: a systematic review[J]. *Spinal Cord*, 2020(58): 130-148.
- [2] KIGERL K A, ZANE K, ADAMS K, *et al.* The spinal cord-gut-immune axis as a master regulator of health and neurological function after spinal cord injury[J]. *Experimental Neurology*, 2020 (323): 113085.
- [3] CHAMBERS G M, SETTUMBA S N, CAREYK A, *et al.* Prenusinersen economic and health-related quality of life burden of spinal muscular atrophy[J]. *Neurology*, 2020: 1-10.
- [4] GEETHANATH S, VAUGHAN Jr J T. Accessible magnetic resonance imaging: a review[J]. *Journal of Magnetic Resonance Imaging*, 2019, 49(7): 65-77.
- [5] OTSU N. A threshold selection method from gray-level histograms[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, 9(1): 62-66.
- [6] CANNY J. A computational approach to edge detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, 8(6): 679-698.
- [7] HARTIGAN J A, WONG M A. Algorithm as 136: a k-means clustering algorithm[J]. *Journal of the Royal Statistical Society*, 1979, 28(1): 100-108.
- [8] 杨玉聘, 张俊华, 刘云凤. 基于三维循环残差卷积的脊柱 CT 图像分割 [J/OL]. *计算机工程*, 1-11[2024-04-19].
YANG Yudan, ZHANG Junhua, LIU Yunfeng. Segmentation of spinal CT images based on three-dimensional cyclic residual convolution[J/OL]. *Computer Engineering*, 1-11[2024-04-19].
- [9] 郭春麟, 张勇, 刘祎, 等. 脊柱 X 射线图像分割方法 [J]. *国外电子测量技术*, 2022, 41 (7): 23-28.
GUO Chunlin, ZHANG Yong, LIU Yi, *et al.* Spine X-ray image segmentation method[J]. *Foreign Electronic Measurement Technology*, 2022, 41 (7): 23-28.
- [10] HORNG M H, KUOK C P, FU M J, *et al.* Cobb angle measurement of spine from X-ray images using convolutional neural network[J]. *Computational and Mathematical Methods in Medicine*, 2019(1): 6357171.
- [11] DUBOST F, COLLERY B, RENAUDIER A, *et al.* Automated estimation of the spinal curvature via spine centerline extraction with ensembles of cascaded neural networks[C]//*Computational Methods and Clinical Applications for Spine Imaging*. Berlin: Springer, 2020: 88-94.
- [12] TRAN V L, LIN H Y, LIU H W. MBNet: a multi-task deep neural network for semantic segmentation and lumbar vertebra inspection on X-ray images[C]//*Proceedings of the 15th Asian Conference on Computer Vision*. Berlin: Springer, 2020:635-651.
- [13] CHENG M, QIN J, LYU L, *et al.* Multi-scale information aggregation network for spine MRI image segmentation[C]//*Proceedings of the 2022 6th International Conference on Electronic Information Technology and Computer Engineering*. New York: Association for Computing Machinery, 2022: 1377-1382.
- [14] LESSMANN N, VAN GINNEKEN B, DE JONG P A, *et al.* Iterative fully convolutional neural networks for automatic vertebra segmentation and identification[J]. *Medical Image Analysis*, 2019, 53: 142-155.
- [15] ISENSEE F, JAEGER P F, KOHIS A A, *et al.* nn-UNet: a self-configuring method for deep learning-based biomedical image segmentation[J]. *Nature Methods*, 2021, 18(2): 203-211.
- [16] BALLESTAR L M, VILAPLANA V. MRI brain tumor segmentation and uncertainty estimation using 3D-UNet architectures[C]//*International MICCAI Brainlesion Workshop*. Berlin: Springer, 2020:376-390.
- [17] ZHOU S, NIE D, ADELI E, *et al.* Semantic instance segmentation with discriminative deep supervision for medical images[J]. *Medical Image Analysis*, 2022, 82: 102626.
- [18] VAN DER GRAAF J W, VAN HOOFF M L, BUCKENS C F M, *et al.* Lumbar spine segmentation in MR images: a dataset and a public benchmark[J]. *Scientific Data*, 2024, 11(1): 264.
- [19] YE F, WANG Z, ZHU S, *et al.* A novel convolutional neural network based on adaptive multi-scale aggregation and boundary-aware for lateral ventricle segmentation on MR images[C]//*2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. New York: IEEE, 2022: 1431-1435.
- [20] ALOM M Z, HASAN M, YAKOPCIC C, *et al.* Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation[J]. *arXiv Preprint arXiv*: 1802.06955, 2018.
- [21] Yao J, BURNS J E, FORSBERG D, *et al.* A multi-center milestone study of clinical vertebral CT segmentation[J]. *Computerized Medical Imaging and Graphics*, 2016 (49): 16-28.