

王子民¹ 周悦¹ 关挺强¹ 郭欣¹ 胡巍² 王茂发¹

基于改进 U2-Net 网络的多裂肌 MRI 图像分割算法

摘要

针对腰间盘突出患者 MRI 图像多裂肌病变部位分割精度较低的问题,提出一种改进的 U2-Net 网络的新模型,目标是使得编码和解码的子网络通过一系列嵌套的跳跃路径来相互连接,重新设计 U2-Net 模型中 RSU-7、RSU-6、RSU-5、RSU-4 中间的跳跃连接,RSU-4F 部分不变,用来降低编码解码子网络中特征图的语义缺失.为了提取到高质量的多裂肌特征,加入通道注意力模块,通过学习每个通道的权重,使网络能够更好地关注对任务有贡献的通道,从而提升模型的性能.为验证模型的有效性,在多裂肌 MRI 图像数据集上进行实验,发现相较于 U-Net、U2-Net、U-Net++ 网络结构,骰子系数(Dice)、豪斯多夫距离(HD)以及均交并比(MIoU) 3 个指标均有优化.实验结果表明,本文提出的算法对于多裂肌的 MRI 图像分割有较好的效果,能够辅助医生对病情做出判断.

关键词

磁共振成像(MRI);深度学习;医学图像分割;多裂肌;注意力机制;稠密连接;U2-Net

中图分类号 TP391.4

文献标志码 A

收稿日期 2023-07-17

资助项目 国家自然科学基金(61866009,42164002);广西重点研发计划(AB21220037);广西科技计划(基地和人才专项)(桂科AD20325004);桂林市科学研究与技术开发项目(20210227-2);国家自然科学基金青年基金(41504037)

作者简介

王子民,男,博士,副教授,主要从事智能信息处理、模式识别等研究.worthyman@guet.edu.cn

1 桂林电子科技大学 计算机与信息安全学院,桂林,541004

2 柳州市人民医院,柳州,545006

0 引言

腰椎间盘突出症(Lumbar Disc Herniation)是临床常见病和多发病,好发于成年人.至少 95%的腰椎间盘突出症发生于 L_{4,5}和 L₅-S₁^[1].国外相关研究显示腰椎间盘突出症发病率大约 2%~3%,而 35 岁以上的男性发病率约 4.8%,35 岁以上女性发病率约 2.5%^[2].既往研究表明,腰痛与多裂肌萎缩密切相关^[3].腰椎间盘突出症是一种以腰椎疼痛和腿部疼痛为特征的疾病,由于纤维环破裂后动脉粥样硬化核突出导致神经根部压缩^[4].由于腰骶部为人体主要承重部位,频繁参与各种旋转等大幅度动作,容易导致多裂肌疲劳和功能性下降^[5],诱发腰椎间盘突出和复发.肌肉萎缩在影像学上主要表现为肌肉体积缩小和肌内脂肪浸润增加^[6].磁共振成像(Magnetic Resonance Imaging, MRI)是影像学诊断进入三维空间的标志,对软组织的分辨率高,可以准确看到椎间盘的突出位置、突出方向和突出大小,可以提供更准确的影像学诊断信息,在临床上使用较多,是影像学中诊断腰间盘突出的首选方法^[7].然而,由于 MRI 图像的对比度较低,医生需要手动描绘多裂肌的横截面积和脂肪浸润情况,而不同的医生可能会导致阅片结果的差异.因此,运用深度学习技术对多裂肌病变区域,即多裂肌脂肪浸润面积进行分割,可有效提升诊断准确性,为医生的诊疗工作提供更加便捷的条件.

目前,脊椎 CT 图像分割方法主要有基于传统方法和基于深度学习方法两类.传统脊椎 CT 图像分割方法,例如:Aslan 等^[8]使用统计水平集方法分割脊椎,但该方法需要利用期望最大化算法进行初始化和参数估计;Rasoulia 等^[9]提出一种基于配准的脊椎 CT 图像分割技术,并建立了相应统计的多脊椎形状+姿势模型,该模型可以捕获脊椎形状和姿态的变化,同时减少配准参数的数量;Lim 等^[10]提出一种改进水平集的脊椎分割框架,该框架通过 Willmore 流提取局部几何形状特征,并利用先验形状核密度估计器获取所需的先验知识,从而能够处理因噪声污染导致信息缺失的图像.近年来,深度学习^[11]应用广泛,基于深度卷积神经网络(CNN)的方法^[12-13]已成功地打破传统手工提取特征方法的瓶颈,特别是全卷积神经网络^[14](Fully Convolutional Network, FCN)、金字塔场景解析网络^[15]、DeepLab^[16]、U-Net^[17]及 U-Net 的变体^[18-20],这些基于深度学习的神经网络因优异的性能在医学图像处理中取得了不错的成果^[21-22].在针对脊椎图像分割方法研

究上, Zhang 等^[23]提出一种基于 PointNet++ 的颈椎分割方法. 该方法需对 CT 图像进行预处理, 以及对 PointNet++ 分割结果做后处理, 这虽然提高了分割的准确度, 但是仅使用了点云边缘信息, 强度信息利用率较低, 导致对一些精细的结构分割不全. Kuok 等^[24]提出一种基于全卷积密集网络的胸腰椎分割方法. 该方法将脊椎分割结果和椎间平面检测结果相结合重建出了脊椎的三维模型, 但是由于数据集较少, 导致网络泛化性能不够. 为获取更多的细节信息, 许多学者研究了不同形式的注意力模块, 例如: Ramachandran 等^[25]利用一种自注意模块取代空间卷积, 以提高性能并减少计算工作量; Hu 等^[26]提出采用压缩和激励注意力模块构建卷积特征通道之间的相互依赖性, 用以提高网络的特征表示能力; Oktay 等^[27]提出 Attention U-Net, 该网络设计了注意力门控机制(Attention Gate, AG), 然后将 AG 模块替换 U-Net 的跳跃连接, 有效地抑制了网络对无关区域的学习, 同时使网络关注学习与任务有关区域; Zhou 等^[28]提出 U-Net++, 该网络在 U-Net 基础上设计了嵌套结构和密集跳跃连接, 解决了不同场景应用对网络深度的要求, 同时使解码器更好地融合多尺度信息; Huang 等^[29]提出 U-Net3+, 它采用全尺度跳跃连接将不同尺度特征图的语义信息相结合, 同时通过深度监督从全尺度聚合的特征图中学习层次表示; Qin 等^[30]提出的 U2-Net 网络已经在显著性目标检测、人像分割、自然场景分割等任务上取得了显著的成果, 并成为图像分割领域的重要参考模型之一. 相比于以上介绍的网络, U2-Net^[30]具有更深的网络结构和更多的卷积模块, 它能够更好地处理复杂的图像场景, 提取更丰富的特征信息, 并且性能更优. 但上述模型普遍存在边界预测不准确、图像细节分割不够明确的问题.

本文主要工作包括:

- 1) 基于 U2-Net^[30] 网络, 引入稠密连接模块, 重新设计 RSU-7 层、RSU-6 层、RSU-5 层、RSU-4 层的跳跃连接, 以此抓取不同层次的特征, 获得不同层次的感受野;
- 2) 引入通道注意力机制模块, 每个特征赋予不同的权重, 从而关注重要的特征并抑制不重要的特征, 提高分割精度.

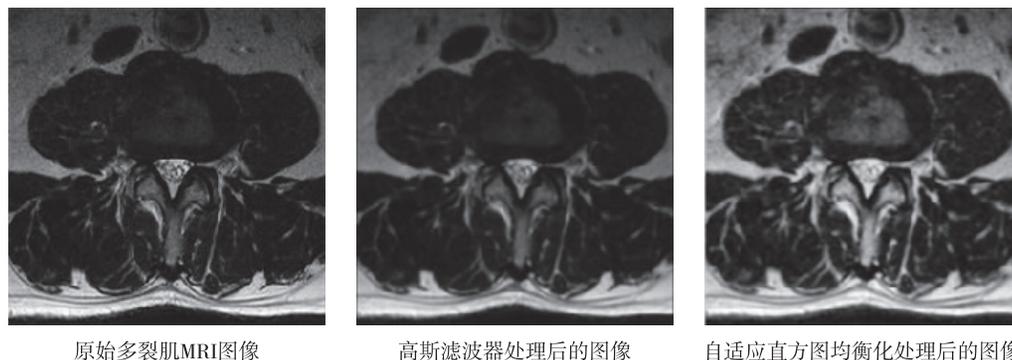
1 本文方法

1.1 实验数据介绍

本研究的实验数据由合作医院提供. 研究使用的初始数据集包括 310 幅腰椎间盘突出患者的多裂肌 MRI 图像, 这些图像均呈现了多裂肌在腰椎间盘突出患者中的形态. 多裂肌的脂肪浸润由两名外科医生使用 labelme 图像标注软件标注, 并由一名脊柱专家进行检查. 在扩充数据并根据患者分配选择数据集后, 984 个数据集被用作训练集、108 个被用作验证集、152 个被用作测试集.

1.2 数据预处理和数据扩增

在成像过程中, MRI 图像或多或少会受到设备电压、信号强度和受试者衣着的影响. 这些因素使得图像信噪比下降, 导致最终结果不理想. 因此, 选择高斯滤波器来降低图像噪声. 由于分割出的多发性畸形区域与其他区域的灰度差异较小, 对分割的准确性略有影响, 因此, 对图像进行自适应直方图平滑处理, 增加多发性畸形区域与其他区域的对比度, 这样可以减少无关区域对目标区域分割的影响. 此外, 由于数据集样本量少, 对数据进行数据扩增, 具体的方法包括: 旋转操作设置为顺时针、逆时针旋转 30° 和 60°; 镜像操作设置为水平翻转、垂直旋转. 数据预处理后的结果如图 1 所示.



原始多裂肌MRI图像

高斯滤波器处理后的图像

自适应直方图均衡化处理后的图像

图 1 数据预处理后的结果

Fig. 1 Results after data pre-processing

1.3 SE-U2Net++算法

U2-Net 网络的结构类似于 U-Net 网络,它是一种 U 型网络结构,具有编码和解码的能力.U2-Net 网络模型具备多层特征提取的能力,且其网络结构更为复杂,能够更加深入地挖掘图像信息的多层次结构.由于采用了残差结构,因此降低了对计算和内存的需求,进而提高了系统的效率和可靠性.

本模型(SE-U2Net++)在 U2-Net 整体编解码网络结构中嵌套了编解码的子网络,在此基础上重新设计 U2-Net 模型中 RSU-7、RSU-6、RSU-5、RSU-4 中间的跳跃连接,RSU-4F 部分不变.这些重新设计的跳跃连接部分即为稠密连接模块,可以降低编码和解码过程中的语义缺失.此外,为了提取到高质量

的多裂肌特征,还加入了通道注意力模块,通过使用注意机制来增加表现力、聚焦重要特征、抑制无关特征,有效地保持网络信息的流动性,增强了表达能力.本文提出的分割模型如图 2 所示.

1.3.1 稠密连接模块

在深度学习中,长连接(Long Connections)和短连接(Short Connections)是两种常见的连接方式.长连接可以帮助缓解梯度消失或梯度爆炸的问题,特别是在深层网络中,由于信息可以跨层级传播,梯度可以更轻松地通过网络传播到较浅的层,从而使得网络更容易训练.此外,长连接可以促进特征的重用和信息的传递.通过将前面层的特征与后面层的特征相加或连接,可以使得网络更加关注重要的特征,

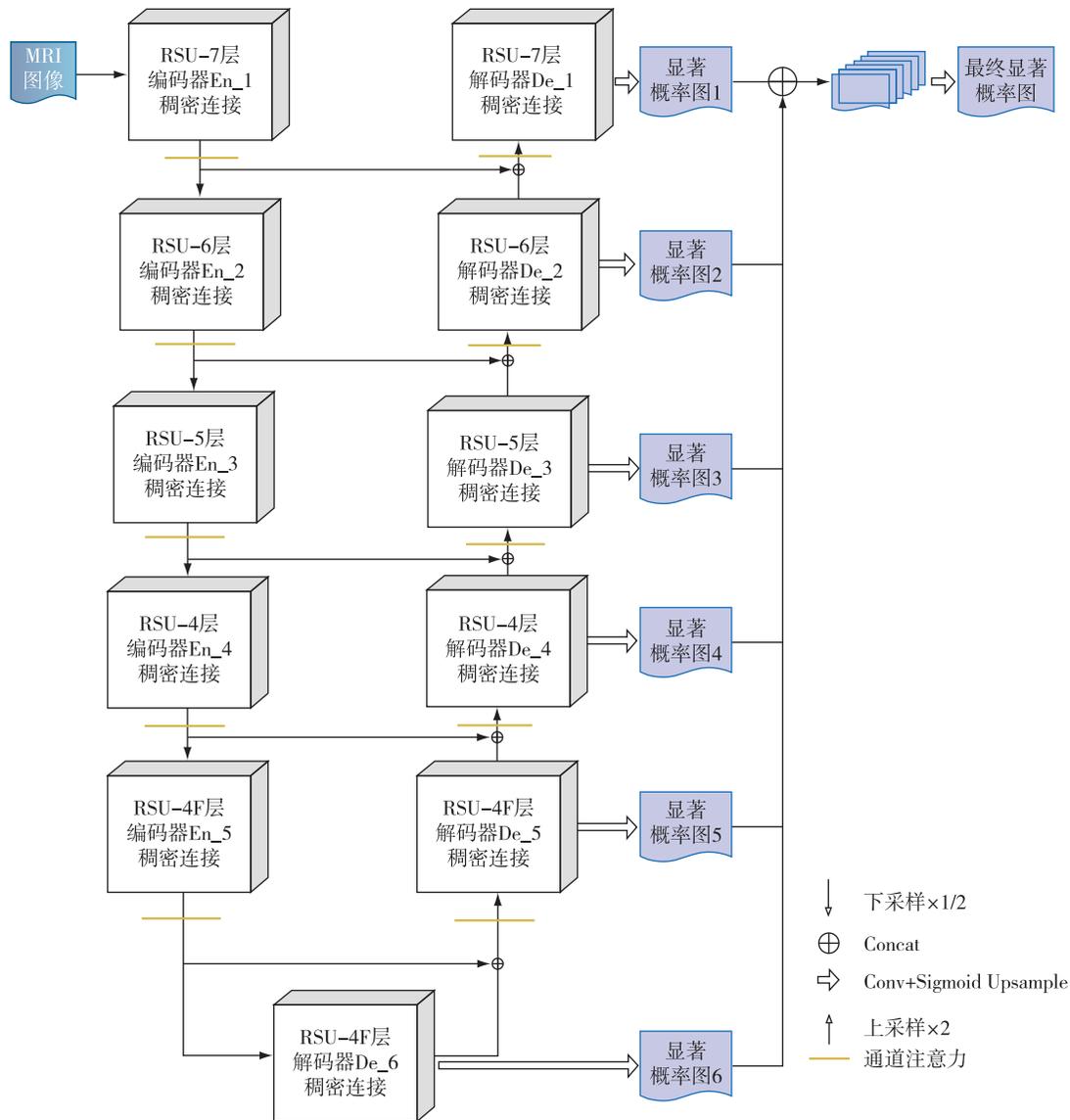


图 2 分割模型

Fig. 2 The proposed MRI image segmentation model

并避免信息丢失,这对于处理细节信息和保留上下文信息都是有益的.而短连接可以实现多层特征的融合,从而允许网络同时利用不同层级的特征信息.这样的融合机制可以提高网络的表示能力,增强对不同尺度和抽象级别的特征的理解.同时,短连接可以传递低层特征的细节信息到高层,这对于任务中需要处理细粒度结构或纹理的情况非常有用.通过保留和传递细节信息,网络可以更好地还原原始输入的细节.因此,长连接和短连接共同作用可以提高网络的性能和表达能力.

本文通过将嵌套的、密集的长短跳跃连接相结合的网络结构来构建稠密连接模块.因此,相较于U2-Net的朴素跳跃连接方式,本文所提出的稠密连接模块能够更加精准地捕捉不同层次的特征,并且对于不同大小的目标对象表现出了不同的敏感度.因此,通过特征叠加的方式,稠密连接模块能够将不同层次的特征融合在一起,使本文网络拥有不同大小的感受野.RSU-7层、RSU-6层、RSU-5层、RSU-4层重新设计后的稠密连接如图3所示:黑色部分呈

现出U2-Net结构的原始形态,由编码器下采样、解码器上采样以及黑色连线的跳跃连接3个组成部分构成;粉色部分表示在跳跃连接上的稠密卷积块,包括卷积和上采样2部分;淡蓝色连线部分就是重新设计后的跳跃连接,为子网络提供了跳跃连接.

1.3.2 通道注意力模块

通道注意力机制(Channel Attention Mechanism)是一种用于增强深度学习模型对不同通道的关注程度的注意力机制.它通过自动学习每个通道的重要性,调整通道的特征表示,以提升模型在计算机视觉任务中的性能.在通道注意力模块中,特征的每一个通道都代表着一个专门的检测器,因此,通道注意力关注的是什么样的特征是有意义的.

对于一张输入的特征图A,它的尺寸大小为 x_{ji} .为汇总空间特征,采用全局平均池化和最大池化两种方式来分别利用不同的信息,可以通过给每个特征赋予不同的权重,从而关注重要的特征而抑制不重要的特征.对于输入的特征图A通过式(1)得到通

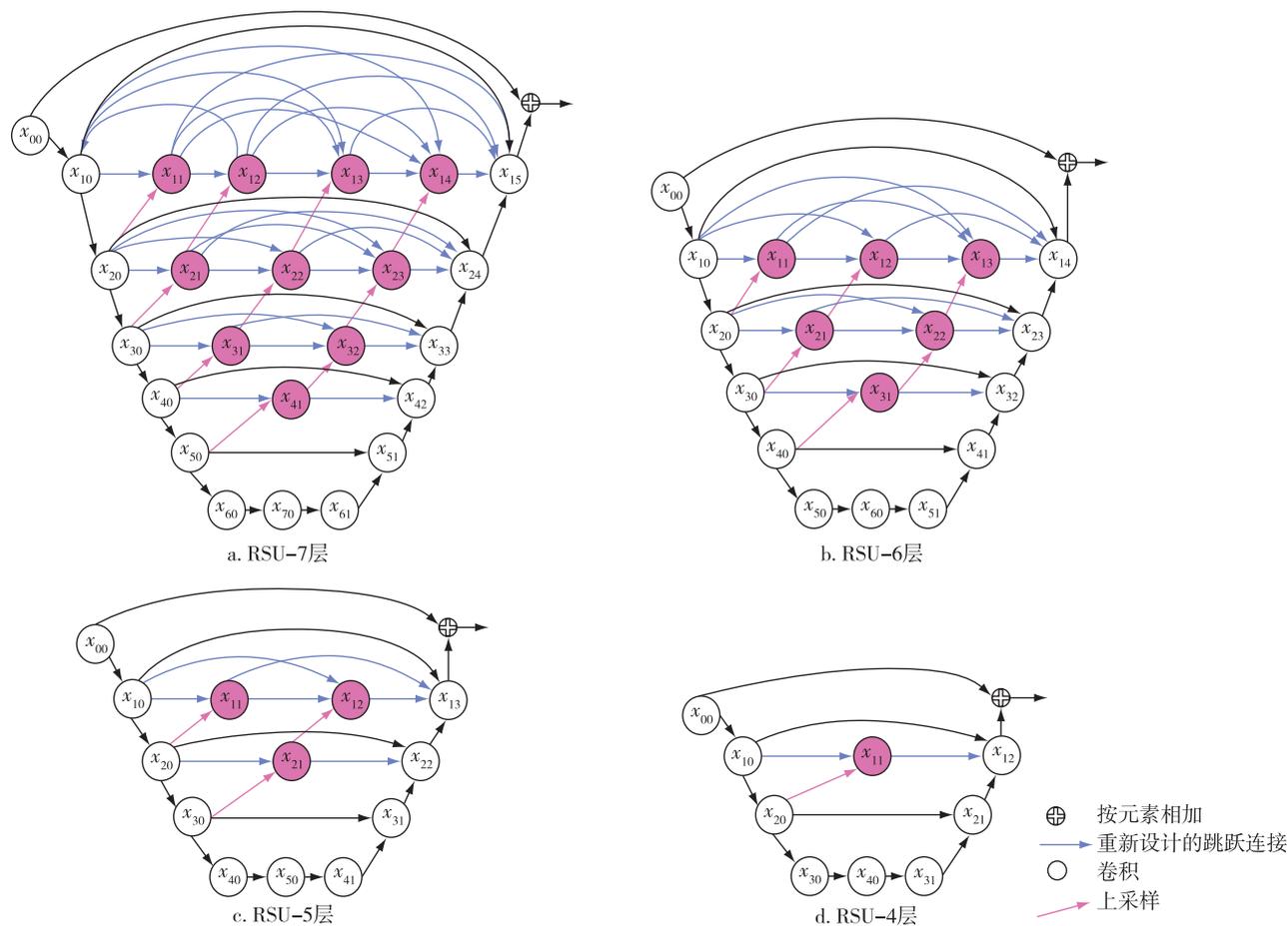


图3 稠密连接模块

Fig. 3 Thick link module

道注意力矩阵 \mathbf{X} , 它的尺寸大小为 $C \times C$. 其中 x_{ji} 表征的是通道 i 对通道 j 的影响. A 重组后的特征矩阵与 \mathbf{X} 进行相乘后, 再进行重组并分配一个 β 权重, 然后与原始的 A 进行相加得到最终输出的通道注意力机制特征图 E , 它的尺寸大小为 $C \times H \times W$, 即式(2).

$$x_{ji} = \frac{\exp(A_i \times A_j)}{\sum_{i=1}^c \exp(A_i \times A_j)}, \quad (1)$$

$$E_j = \beta \sum_{i=1}^c (x_{ji} \cdot A_i) + A_j. \quad (2)$$

通道注意力机制可以增强网络对不同通道的表示能力, 使网络能够自动学习到每个通道的重要性, 从而增强对特定特征的关注能力. 通过调整通道的特征表示, 通道注意力可以帮助模型更好地区分和捕捉不同特征之间的信息, 从而提升模型在计算机视觉任务中的性能. 通道注意力还可以在在一定程度上减少网络参数和计算量, 通过自适应地调整通道的权重, 使得模型更加紧凑, 仅保留对任务有用的通道信息, 减少冗余和不必要的计算, 从而降低模型的存储和计算成本. 通道注意力流程如图 4 所示.

1.4 损失函数

本文所采用的损失函数为 BCELoss(), 该函数是计算目标值和预测值之间的二进制交叉熵损失函数, 其函数计算方法如下:

$$l_n = -\mathbf{w}_n \cdot [\mathbf{y}_n \cdot \ln \mathbf{x}_n + (1 - \mathbf{y}_n) \cdot \ln(1 - \mathbf{x}_n)]. \quad (3)$$

其中: \mathbf{w}_n 表示权重矩阵; \mathbf{x}_n 表示预测值矩阵(输入矩阵被激活函数处理后的结果); \mathbf{y}_n 表示目标值矩阵.

2 实验结果分析

2.1 参数设置

本文所采用的实验环境为: 深度学习框架 PyTorch 1.1, Python 3.6, Windows 10, 处理器 Intel(R) Core(TM) i7-6500Ux CPU @ 3.80 GHz, 2.50 GHz. 使用 Adam 优化算法, 初始学习率为 0.001, 用于计算

梯度以及梯度平方的运行平均值的系数分别为 0.9、0.999. 将多裂肌 MRI 图像按照 4:1 的比例分为训练集和测试集, BatchSize 设为 4, 共计进行 150 个 epoch.

2.2 评价指标

为了量化本模型的分割结果, 本文将从骰子系数(Dice)、豪斯多夫距离(Hausdorff Distance, HD)以及均交并比(Mean Intersection over Union, MIoU) 3 个维度来评价分割结果的优劣性. Dice 系数是一个综合性指标, 可以从像素层面衡量分割结果中正确分割、错误分割以及欠分割程度, 但是对于分割结果中的离群值无法表征, 因此引入了 HD, 用于评价病灶标签和分割结果边缘的差异性. 与此同时, 引入 MIoU 评价指标, 来判断预测像素的准确率. Dice、MIoU 数值越大, HD 数值越小时, 表示该模型性能越好. 它们的计算公式如式(4)~(8)所示. 其中, k 为类别数, i 表示真实值, X 和 Y 分别表示预测集和 Ground-Truth 像素集. TP、FN、FP、TN 如表 1 所示.

$$c_{\text{Dice}} = \frac{2n_{\text{TP}}}{2n_{\text{TP}} + n_{\text{FP}} + n_{\text{FN}}} = \frac{2|X \cap Y|}{|X| + |Y|}, \quad (4)$$

$$D_{\text{H}}(X, Y) = \max\{h(X, Y), h(Y, X)\}, \quad (5)$$

$$h(X, Y) = \max_{x \in X} \min_{y \in Y} \|x - y\|, \quad (6)$$

$$h(Y, X) = \max_{y \in Y} \min_{x \in X} \|y - x\|, \quad (7)$$

$$i_{\text{MIoU}} = \frac{1}{k+1} \sum_{i=0}^k \frac{n_{\text{TP}}}{n_{\text{TP}} + n_{\text{FP}} + n_{\text{FN}}}. \quad (8)$$

表 1 分类结果的混淆矩阵

Table 1 Confusion matrix for classification results

预测值	实际值	
	正样本	负样本
正样本	TP	FP
负样本	FN	TN

2.3 实验结果与分析

为验证实验结果准确性, 共进行 15 次重复性实

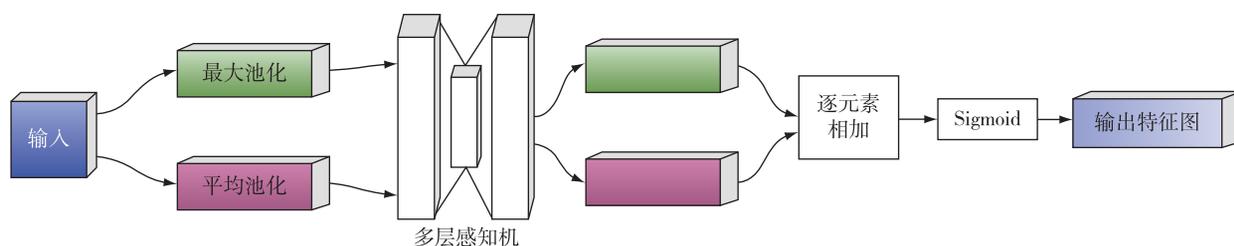


图 4 通道注意力模块

Fig. 4 Channel attention module

验. 所获得的 MIoU、HD、Dice 评价指标的均值、方差、中值已被记录在表 2 中. 根据表 2 所示: MIoU、Dice、HD 的平均值分别为 0.657、0.784 和 3.821; 方差均小于 0.01, 表明模型具有较高的可靠性和稳定性. 通过综合 3 项指标的平均值和方差, 可以得出结论: 所提出的模型表现出较高的稳定性和分割效果, 能够成功地完成多裂肌 MRI 图像的分割任务.

表 2 评价指标统计结果

Table 2 Statistical results of evaluation indicators

	MIoU	HD	Dice
均值	0.657	3.821	0.784
方差	0.009	0.001	0.002
中值	0.649	3.854	0.812

图 5 为使用改进的 U2-Net 分割模型对多裂肌 MRI 图像实现的分割效果. 图 5 中第 1 列为输入到分割网络的原始多裂肌 MRI 图像, 第 2 列为 Ground-Truth(应与分割结果进行对照的正确数据标签), 第

3 列为本文模型的分割结果. 对 Ground-Truth 和本文模型的分割结果进行对比可得, 本文所提出的模型有较好的分割效果.

2.4 多裂肌 MRI 数据集实验结果

2.4.1 本文模型与传统机器学习方法对比

为了验证本文所提出的模型相对于多裂肌 MRI 图像分割的性能提升, 将本文模型与全局阈值分割方法(Global Threshold Segmentation Method)^[31]、自适应阈值分割方法(Adaptive Threshold Segmentation Method)^[32]、边缘检测分割法(Edge Detection Segmentation Method)^[33] 3 种传统机器学习方法进行对比研究, 结果如图 6 所示.

与深度学习方法相比, 传统机器学习方法在处理复杂图像和分割任务时存在一定的限制. 全局阈值分割方法和自适应阈值分割方法适用于目标灰度分布均匀、变化小, 目标和背景灰度差异较明显的图像, 简单易实现且效率高. 然而, 该类方法通常只考虑像素自身的灰度值, 未考虑图像的语义、空间等特征信息, 且

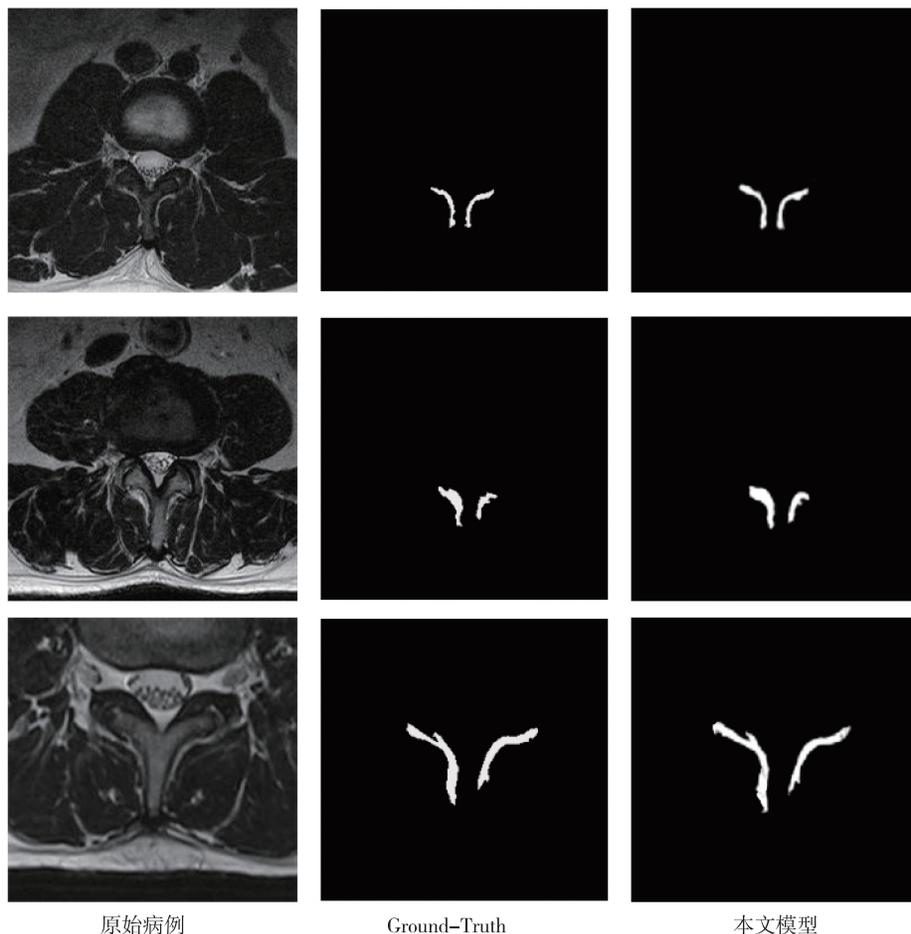


图 5 典型病例的分割结果

Fig. 5 Segmentation results for a typical case

易受噪声影响,对于复杂的 MRI 图像,阈值分割的效果并不理想.基于边缘的图像分割方法对像素灰度值具有明显突变的图片处理效果较好,但该算法易受噪声影响,同样不适于分割复杂 MRI 图像^[34].鉴于 MRI 图像中的病变区域往往呈现复杂多样的形态和纹理,不同病例之间的病变特征可能存在多样性.传统机器学习算法通常依赖于手工设计的特征,而这在 MRI 图像病变分割任务中是一项具有挑战性的工作.手工设计的特征往往难以充分反映病变区域的复杂信息,难以捕捉其内在的复杂性和多样性,从而导致在不同病变的准确分割上存在困难.

综上,本文方法基于深度学习,可以自动学习特征表示,具有更强大的表征能力,能够从原始数据中学习高级抽象特征,在复杂图像分割任务中表现更加优越.

2.4.2 本文模型与经典网络对比

为了验证本文所提出的模型相对于多裂肌 MRI 图像分割的性能提升,进行 U2-Net^[30]、U-Net^[17]、U-Net++^[28]、CE-Net^[35] 的对比研究,分割效果如表 3 所示.

相较于经典网络 U2-Net、U-Net、U-Net++、CE-Net,本文算法加入的稠密连接模块可以使模型在信息传递、梯度传播、特征复用、改进细节预测方面取得优势^[28,36],而通道注意力模块则使本模型在减少

参数和计算复杂性、抑制噪声和冗余、增强对重要信息的关注方面取得优势^[37].根据表 3 所示,本文所提出的模型在 Dice、HD、MIoU 3 个指标上均获得了提升.其中:相较于 U2-Net 网络,本文模型的 MIoU 提高了 5.89%,HD 降低 8.10%,Dice 提高 10.58%;相较于 U-Net,本文模型的 MIoU 提高了 16.16%,HD 降低 9.24%,Dice 提高 10.58%;相较于 U-Net++,本文模型的 MIoU 提高了 15.74%,HD 降低 13.18%,Dice 提高 10.42%;相较于 CE-Net,本文模型的 MIoU 提高了 15.33%,HD 降低 12.50%,Dice 提高 14.96%.实验结果表明,本文所提出的模型在多裂肌 MRI 图像的精准分割和病灶定位方面,具备相当的优越性,其高效的诊断能力为医生提供了有力支持.

表 3 对比实验结果

Table 3 Comparison of experimental results

模型	MIoU	HD	Dice
U2-Net ^[30]	0.611	4.158	0.709
U-Net ^[17]	0.557	4.210	0.717
U-Net++ ^[28]	0.559	4.401	0.710
CE-Net ^[35]	0.561	4.367	0.682
本文模型	0.647	3.821	0.784

为了更加直观地展示各种预测方法的差异,将其可视化结果呈现于图 7 中.Input 表示原始图像,而

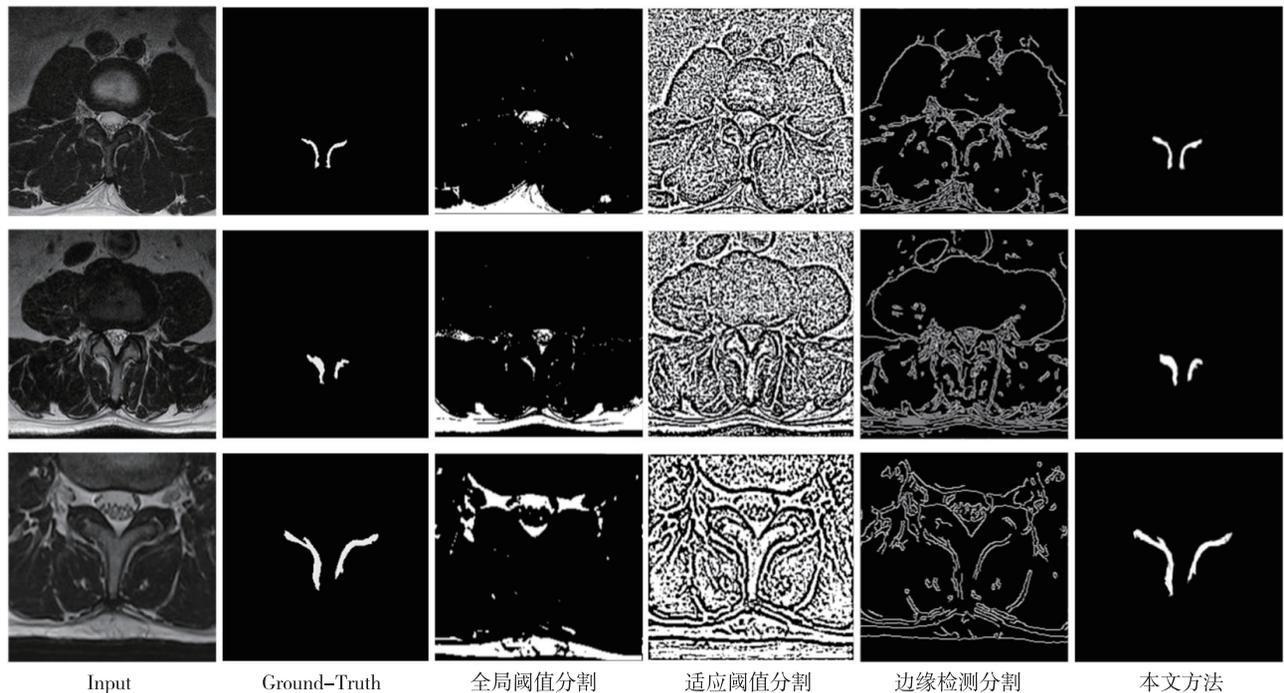


图 6 传统机器学习与本文方法分割结果

Fig. 6 Segmentation results via traditional machine learning methods and the proposed approach

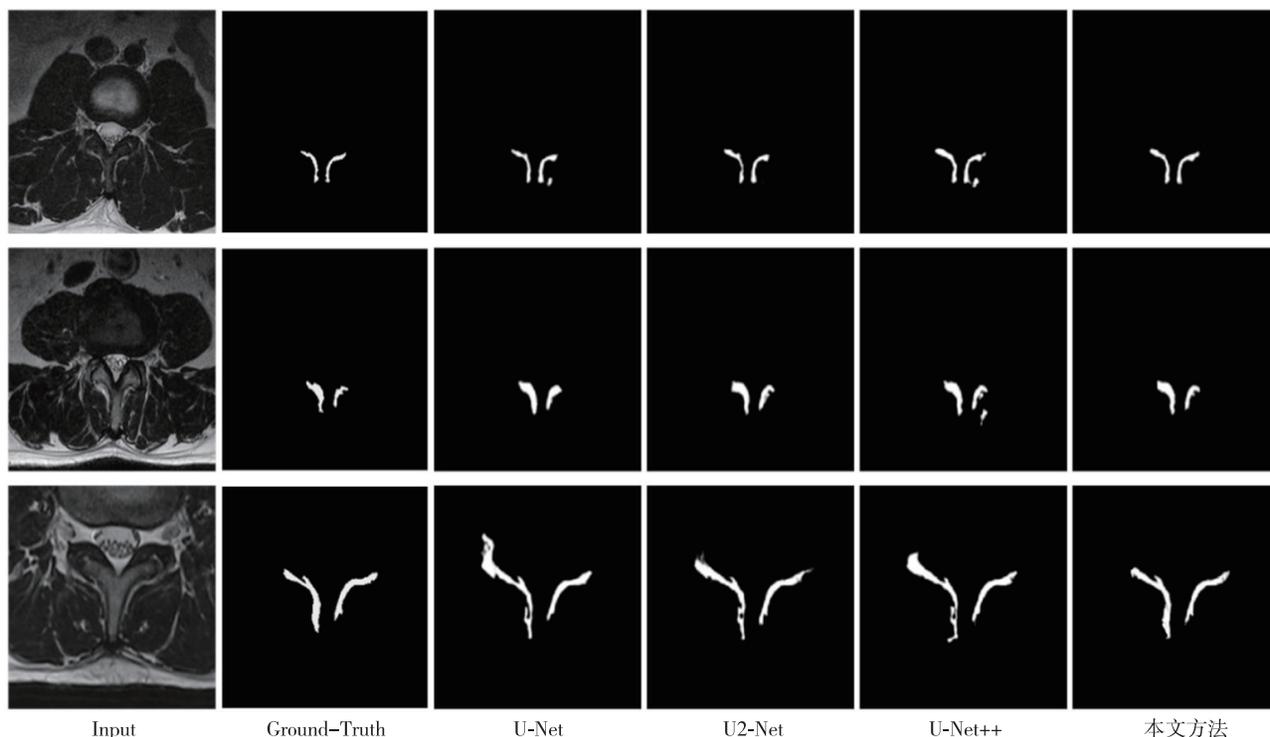


图7 可视化结果

Fig. 7 Visualized results

Ground-Truth 则表示多裂肌中脂肪浸润部分的特征,接着是 U-Net、U2-Net、U-Net++ 以及本文所提出的模型方法对多裂肌脂肪浸润部分进行的预测结果。

2.4.3 消融实验

为了衡量通道注意力机制与稠密连接模块融合对实验结果的影响,设计了模型消融实验,实验结果如表 4 所示。以 U2-Net 为基础网络,进行了 4 项对比实验:1)不添加任何模块;2)添加稠密连接模块,不添加通道注意力模块;3)添加通道注意力模块,不添加稠密连接模块;4)添加稠密连接模块和通道注意力模块。结果表明:相较于原始的 U2-Net 网络,本文模型的 MIoU 提高了 5.89%,HD 降低 8.10%,Dice 提高 10.58%;相较于只增添稠密连接模块,本文模型在 MIoU 提高了 4.69%,HD 降低 6.74%,Dice 提高 6.81%;相较于只增添通道注意力模块,本文模型在 MIoU 提高了 2.21%,HD 降低 2.88%,Dice 提高 0.77%。此结果表明,在 U2-Net 网络基础上添加通道注意力模块,同时对 RSU-7、RSU-6、RSU-5、RSU-4 层添加稠密连接模块可以提高网络性能。

3 结语

为扩大模型的感受野,本文设计了稠密连接模

表 4 消融实验结果

Table 4 Results of ablation experiments

方法	MIoU	HD	Dice
U2-Net	0.611	4.158	0.709
U2-Net+稠密连接	0.618	4.097	0.734
U2-Net+通道注意力	0.633	3.923	0.778
本文模型	0.647	3.821	0.784

块,它将长、短跳跃连接相结合,以捕捉不同层次的特征。同时,还引入了通道注意力模块,以学习不同通道之间的关联和重要性,从而增强对特定特征的关注能力。最终,将这种稠密连接与 U2-Net 网络相结合,设计了 RSU-7、RSU-6、RSU-5、RSU-4 层的跳跃连接,并在每一层的 Encoder 和 Decoder 加入了通道注意力机制。实验结果显示,本文模型显著提升了对腰间盘突出患者多裂肌脂肪浸润部分的分割精度,同时在病灶的定位位置方面表现出较高的准确性,为临床医生提供了客观高效的决策支持,从而有助于对腰间盘突出患者的诊断、治疗方案的制定以及预后等方面做出更加精准的判断。

数据可用性声明

可以向通信作者索取支持本研究结果的数据。

参考文献

References

- [1] Deyo R A, Mirza S K. Herniated lumbar intervertebral disk [J]. *New England Journal of Medicine*, 2016, 374 (18): 1763-1772
- [2] Vialle L R, Vialle E N, Henao J E S, et al. Lumbar disc herniation [J]. *Revista Brasileira de Ortopedia*, 2015, 45 (1): 17-22
- [3] Schryver A, Rivaz H, Rizk A, et al. Ultrasonography of lumbar multifidus muscle in university American football players [J]. *Medicine and Science in Sports and Exercise*, 2020, 52 (7): 1495-1501
- [4] 陈齐齐, 陈立平, 金军, 等. 多裂肌横截面积与经皮椎间孔镜治疗腰椎间盘突出症患者预后的关系 [J]. *实用医学杂志*, 2021, 37 (9): 1136-1140
CHEN Qiqi, CHEN Liping, JIN Jun, et al. Relationship between cross-sectional area of multifidus muscle and prognosis of patients with lumbar disc herniation treated by percutaneous endoscopic transforaminal discectomy [J]. *The Journal of Practical Medicine*, 2021, 37 (9): 1136-1140
- [5] Ward S R, Kim C W, Eng C M, et al. Architectural analysis and intraoperative measurements demonstrate the unique design of the multifidus muscle for lumbar spine stability [J]. *The Journal of Bone and Joint Surgery American Volume*, 2009, 91 (1): 176-185
- [6] 斜大雄, 王君瑞, 陈绍东, 等. 单侧腰痛患者腰椎多裂肌萎缩相关因素研究 [J]. *浙江医学*, 2014, 36 (10): 847-849, 853
DOU Daxiong, WANG Junrui, CHEN Shaodong, et al. Factors related to lumbar multifidus muscle atrophy in patients with unilateral low back pain [J]. *Zhejiang Medical Journal*, 2014, 36 (10): 847-849, 853
- [7] Kalichman L, Hodges P, Li L, et al. Changes in paraspinal muscles and their association with low back pain and spinal degeneration: CT study [J]. *European Spine Journal*, 2010, 19 (7): 1136-1144
- [8] Aslan M S, Farag A A, Arnold B, et al. Segmentation of vertebrae using level sets with expectation maximization algorithm [C] // 2011 IEEE International Symposium on Biomedical Imaging: from Nano to Macro. March 30-April 2, 2011, Chicago, IL, USA. IEEE, 2011: 2010-2013
- [9] Rasoulia A, Rohling R, Abolmaesumi P. Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape pose model [J]. *IEEE Transactions on Medical Imaging*, 2013, 32 (10): 1890-1900
- [10] Lim P H, Bagci U, Bai L. Introducing willmore flow into level set segmentation of spinal vertebrae [J]. *IEEE Transactions on Biomedical Engineering*, 2013, 60 (1): 115-122
- [11] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks [J]. *Science*, 2006, 313 (5786): 504-507
- [12] 许美贤, 郑琰, 周若兰, 等. 基于 BP 神经网络和多元线性回归的辛烷值预测 [J]. *南京信息工程大学学报 (自然科学版)*, 2023, 15 (4): 379-392
- XU Meixian, ZHENG Yan, ZHOU Ruolan, et al. Octane number prediction based on BP neural network and multiple linear regression [J]. *Journal of Nanjing University of Information Science & Technology (Natural Science Edition)*, 2023, 15 (4): 379-392
- [13] 李辉, 闵巍庆, 王致岭, 等. 基于图卷积网络的多标签食品原材料识别 [J]. *南京信息工程大学学报 (自然科学版)*, 2019, 11 (6): 743-750
LI Hui, MIN Weiqing, WANG Zhiling, et al. Multi-label food ingredient recognition via graph convolution network [J]. *Journal of Nanjing University of Information Science & Technology (Natural Science Edition)*, 2019, 11 (6): 743-750
- [14] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39 (4): 640-651
- [15] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 6230-6239
- [16] Zhou Y P, Chang H Y, Lu X L, et al. DenseUNet: improved image classification method using standard convolution and dense transposed convolution [J]. *Knowledge-Based Systems*, 2022, 254: 109658
- [17] 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. Cham: Springer, 2015: 234-241
- [18] Diakogiannis F I, Waldner F, Caccetta P, et al. ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 162: 94-114
- [19] Thomas E, Pawan S J, Kumar S, et al. Multi-res-attention UNet: a CNN model for the segmentation of focal cortical dysplasia lesions from magnetic resonance images [J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25 (5): 1724-1734
- [20] Chen W, Liu B, Peng S, et al. S3D-UNet: separable 3D U-Net for brain tumor segmentation [C] // Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 4th International Workshop, BrainLes 2018, Held in Conjunction with MICCAI 2018. September 16, 2018, Granada, Spain. Springer International Publishing, 2019: 358-368
- [21] 何晓云, 许江淳, 陈文绪. 基于改进 U-Net 网络的眼底血管图像分割研究 [J]. *电子测量与仪器学报*, 2021, 35 (10): 202-208
HE Xiaoyun, XU Jiangchun, CHEN Wenxu. Research on fundus blood vessel image segmentation based on improved U-Net network [J]. *Journal of Electronic Measurement and Instrumentation*, 2021, 35 (10): 202-208
- [22] 赵其杰, 周安稳, 朱俊豪, 等. 基于 PE-Vnet 网络的三维骨骼图像分割方法 [J]. *仪器仪表学报*, 2020, 41 (7): 243-251
ZHAO Qijie, ZHOU Anwen, ZHU Junhao, et al. Segmentation method of three-dimensional bone image based on

- PE-Vnet network [J]. Chinese Journal of Scientific Instrument, 2020, 41(7): 243-251
- [23] Zhang L, Wang H. A novel segmentation method for cervical vertebrae based on PointNet++ and converge segmentation [J]. Computer Methods and Programs in Biomedicine, 2021, 200: 105798
- [24] Kuok C P, Hsue J Y, Shen T L, et al. An effective CNN approach for vertebrae segmentation from 3D CT images [C]//2018 Pacific Neighborhood Consortium Annual Conference and Joint Meetings (PNC). October 27-30, 2018, San Francisco, CA, USA. IEEE, 2018: 1-6
- [25] Ramachandran P, Parmar N, Vaswani A, et al. Stand-alone self-attention in vision models [C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems, 2019: 68-80
- [26] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141
- [27] Oktay O, Schlemper J, Folgoc L L, et al. Attention U-Net: learning where to look for the pancreas [J]. arXiv e-print, 2018, arXiv: 1804. 03999
- [28] Zhou Z W, Siddiquee M M R, Tajbakhsh N, et al. U-Net++: a nested U-Net architecture for medical image segmentation [C]//4th International Workshop on Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. September 20, 2018, Granada, Spain. Springer International Publishing, 2018: 3-11
- [29] Huang H M, Lin L F, Tong R F, et al. U-Net 3: a full-scale connected UNet for medical image segmentation [C]//2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). May 4-8, 2020, Barcelona, Spain. IEEE, 2020: 1055-1059
- [30] Qin X B, Zhang Z C, Huang C Y, et al. U2-Net: going deeper with nested u-structure for salient object detection [J]. Pattern Recognition, 2020, 106: 107404
- [31] Pun T. A new method for grey-level picture thresholding using the entropy of the histogram [J]. Signal Processing, 1980, 2(3): 223-237
- [32] Niblack W. An introduction to digital image processing [M]. London: Strandberg Publishing Company, 1985
- [33] Khan J F, Bhuiyan S M A, Adhami R R. Image segmentation and shape analysis for road-sign detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(1): 83-96
- [34] 黄鹏, 郑淇, 梁超. 图像分割方法综述 [J]. 武汉大学学报(理学版), 2020, 66(6): 519-531
HUANG Peng, ZHENG Qi, LIANG Chao. Overview of image segmentation methods [J]. Journal of Wuhan University (Natural Science Edition), 2020, 66(6): 519-531
- [35] Gu Z W, Cheng J, Fu H Z, et al. CE-net: context encoder network for 2D medical image segmentation [J]. IEEE Transactions on Medical Imaging, 2019, 38(10): 2281-2292
- [36] Singh D, Kumar V, Kaur M. Densely connected convolutional networks-based COVID-19 screening model [J]. Applied Intelligence, 2021, 51(5): 3044-3051
- [37] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision. Cham: Springer, 2018: 3-19

Segmentation of multifidus muscle MRI images via improved U2-Net

WANG Zimin¹ ZHOU Yue¹ GUAN Tingqiang¹ GUO Xin¹ HU Wei² WANG Maofa¹

¹ School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin 541004, China

² Liuzhou People's Hospital, Liuzhou 545006, China

Abstract To address the low segmentation accuracy of multifidus muscle lesion sites in MRI images of patients with lumbar disc herniation, this paper proposes a new model to improve the U2-Net network with the goal that the encoding and decoding subnetworks are interconnected by a series of nested jump paths. To reduce the semantic missing of feature maps in the encoding and decoding subnetworks, the jump connections in the middle of RSU-7, RSU-6, RSU-5, and RSU-4 in the U2-Net model are redesigned, while the RSU-4F part remains unchanged. In addition, the channel attention module is added to enable the net to focus on channels of higher contribution to task, thus extract high quality multifractal muscle features. The experiments on the multifidus muscle MRI image dataset show that the improved U2-Net outperforms U-Net, U2-Net and U-Net++ network in indicators of Dice, HD and MIoU. It can be concluded that the proposed approach has good performance on MRI image segmentation of multifidus muscle, which can assist doctors to make diagnosis.

Key words magnetic resonance imaging (MRI); deep learning; medical image segmentation; multifidus muscle; attention mechanism; thick link; U2-Net