

✧ 医学物理与工程学

UNet fusing patch attention for segmenting knee cartilage on MRI

WANG Xiang, SHI Cao*, YUAN Zhengyi

(School of Information Science and Technology, Qingdao University of
Science and Technology, Qingdao 266000, China)

[Abstract] **Objective** To construct a UNet fusing patch attention (PA-UNet) model, and to observe its value for segmenting knee cartilage on MRI. **Methods** Slice and preprocessing were performed on knee MRI selected from Osteoarthritis Initiative-Zuse Institute Berlin dataset. Taken UNet as the backbone network, a PA-UNet model was constructed based on patch attention mechanism. The effect of PA-UNet model and other models for segmenting both femoral cartilage and tibial cartilage were compared by subjective and objective evaluations. Ablation experiments based on UNet, UNet based on SE with layers 2—4 (UNet+SE), +UNet, ++UNet, +++UNet, +U-Net+, ++U-Net++ and PA-UNet models were performed to observe the effect of models for segmenting knee cartilage. **Results** PA-UNet could accurately segment femoral and tibial cartilage in all simple, medium and difficult samples, which had better segmenting effect on small structures than SegNet, UNet and DeepLabv3+ models. The Dice similarity coefficient (DSC) and intersection over union of PA-UNet model for segmenting femoral and tibial cartilage were both higher, while Hausdorff distance of PA-UNet model was lower than those of UNet, DeepLabv3+, SA-UNet, RA UNet and SegNet models. DSC of PA-UNet model for segmenting femoral cartilage and tibial cartilage was 88.97% and 82.72%, respectively, both higher than those of UNet, UNet+SE, +UNet, ++UNet, +++UNet, +U-Net+ and ++U-Net++ models. **Conclusion** PA-UNet could segment knee cartilage completely on MRI, especially for small structures.

[Keywords] knee joint; cartilage; deep learning; magnetic resonance imaging; attention mechanism

DOI:10.13929/j.issn.1003-3289.2024.05.027

融合分区注意力 UNet 模型用于分割 MRI 中的膝关节软骨

王 翔, 史 操*, 袁正一

(青岛科技大学信息科学技术学院, 山东 青岛 266000)

[摘要] **目的** 构建融合分区注意力的 UNet(PA-UNet)模型, 观察其分割 MRI 中的膝关节软骨的价值。 **方法** 对来源于 Osteoarthritis Initiative-Zuse Institute Berlin 数据集的膝关节 MRI 进行切片及预处理, 以 UNet 为骨干网络构建基于分区注意力机制的 PA-UNet 模型, 通过主、客观评价比较该模型与其他模型分割股骨软骨及胫骨软骨的效果; 分别以基于 UNet、基于 SE(第 2~4 层)的 UNet(UNet+SE)、+UNet、++UNet、+++UNet、+UNet+、++UNet++ 及 PA-UNet 模型的消融实验观察各模型分割膝关节软骨的效果。 **结果** PA-UNet 可准确分割低难度、中等难度及困难样本中的股骨及胫骨软骨, 其分割细小结构效果优于 SegNet、UNet 及 DeepLabv3+ 模型; 其分割股骨软骨及胫骨软骨的戴斯相似系数(DSC)及交并比均高于, 而豪斯多夫距离均低于 UNet、DeepLabv3+、SA-UNet、RA UNet 及 SegNet 模型。以 PA-UNet 模型分割股骨软骨及胫骨软骨的 DSC 分别为 88.97% 及 82.72%, 均高于 UNet、UNet+SE、+UNet、++UNet、+++UNet、+UNet+ 及 ++UNet++ 模型。 **结论** PA-UNet 可完整分割 MRI 中的膝关节软骨, 尤其对细小结构的分割效果良好。

[关键词] 膝关节; 软骨; 深度学习; 磁共振成像; 注意力机制

[第一作者] 王翔(1998—), 男, 山东泰安人, 在读硕士。研究方向: 医学图像处理算法。E-mail: wxiang0110@163.com

[通信作者] 史操, 青岛科技大学信息科学技术学院, 266000。E-mail: caoshi@yeah.net

[收稿日期] 2023-10-27 **[修回日期]** 2024-02-26

[中图分类号] R684; R445 [文献标识码] A [文章编号] 1003-3289(2024)05-0764-05

膝关节软骨对于支撑及保护膝关节至关重要,准确定位膝关节软骨为分析膝关节骨性关节炎的重要预处理步骤;但膝关节解剖复杂,分割图像中的较小结构易受多种因素影响。随着深度学习算法的迅速发展,基于膝关节 MRI 的卷积神经网络已在提高分割膝关节软骨准确性方面展现出巨大潜力。本研究参考文献^[1-3]方法,以 UNet 为骨干网络,引入新的分区注意力模块,并采用特征聚合模块对来自编码器和解码器的特征进行增强、选择及融合,构建基于分区注意力机制的 UNet(PA-UNet)模型,观察其分割 MRI 中的膝关节软骨的价值。

1 资料与方法

1.1 数据集 采用来源于 Osteoarthritis Initiative (https://pubdata. zib. de)-Zuse Institute Berlin (OAI-ZIB)数据集共 507 幅膝关节三维 MRI^[4],均为以 Siemens 3.0T Trio MR 仪采集的三维双回波稳态水激发序列图像,其中每个三维体素均含 160 片分辨率为 384×384 的切片。采用统计形状模型生成股骨软骨、胫骨软骨、股骨及胫骨轮廓,并经柏林 Zuse 研究所专家手动调整。

1.2 数据预处理 将 MRI 保存为 PNG 格式,选取 40~120 位次切片,将膝关节之外的背景标记为 0、股骨软骨标记为 1、胫骨软骨标记为 2。

1.3 网络结构 PA-UNet 由分区注意力模块和特征聚合模块组成。以 UNet 编码器处理 MRI 特征后,融合解码器间与分区注意力模块特征,之后经激活函数获得预测结果。见图 1。

1.3.1 分区注意力模块 向分区注意力模块输入特征图 $F \in R^{C \times H \times W}$ (C、H 及 W 分别代表输入特征图的

通道数、高度及宽度)后,模块输出与输入大小相同的特征图。分区注意力模块主要由分区通道注意力和分区区间注意力组成。分区通道注意力模块:将以编码器获得的特征图 F 分为 $n \times n$ 个分区,向每个分区施加 squeeze 和 excitation(即 SE)操作,获取加权后的分区特征 P ;设超参数分区大小为 32,使所有分区排列回归初始位置,进而获取加权后的特征图 F' 。分区区间注意力:沿通道维度将特征图 F 与原始图像进行拼接,经 1×1 卷积层 $Conv_s$ 后获得输入特征图 F_{in} ;采用 3×3 卷积层 $Conv_j$ 于 F_{in} 中提取分区特征图 F_j^m (j 表示 F^m 中第 j 个位置分区),以 ReLU 为激活函数;以卷积核大小为 $ps \times ps$ 的平均池化层对特征图 F_j^m 进行池化操作(ps 代表分区大小),聚合空间信息,生成空间特征 $E_{in} \in R^{1 \times \frac{H}{ps} \times \frac{W}{ps}}$, $E_{in} = AvgPool(\sigma(Conv_{1 \times 1}(F_{in})))$ 。通过多层感知机(multi-layer perceptron, MLP)将 E_{in} 重新分配至每个分区以获得 E_{out} ;沿通道和空间维度扩展 E_{out} ,并与 F_{in} 相乘,获得逐分区增强的特征 F'' 。见图 2。

1.3.2 特征聚合模块 特征聚合模块由特征增强模块和特征选择模块组成;后者由 1×1 的卷积层 $Conv_{1 \times 1}$ 和 sigmoid 激活函数 σ 各 1 个共同组成,共包含 3 个分支;第 1 个由含 2 个 3×3 卷积层 $Conv_{3 \times 3}$ 的 Base Block 及 batch normalization (BN)组成、第 2 个不经任何处理、第 3 个经 1 个特征增强模块处理。编码器特征 T 和解码器特征 D 经特征增强模块后生成 T' 和 D' ,引入特征选择模块可使其拼接并输出 2 个注意力特征图 P 和 Q ,以之指导特征融合,最终计算加权特征的逐元素之和获得融合后特征图 O ^[5]。见公式(1)~(3)及图 3。

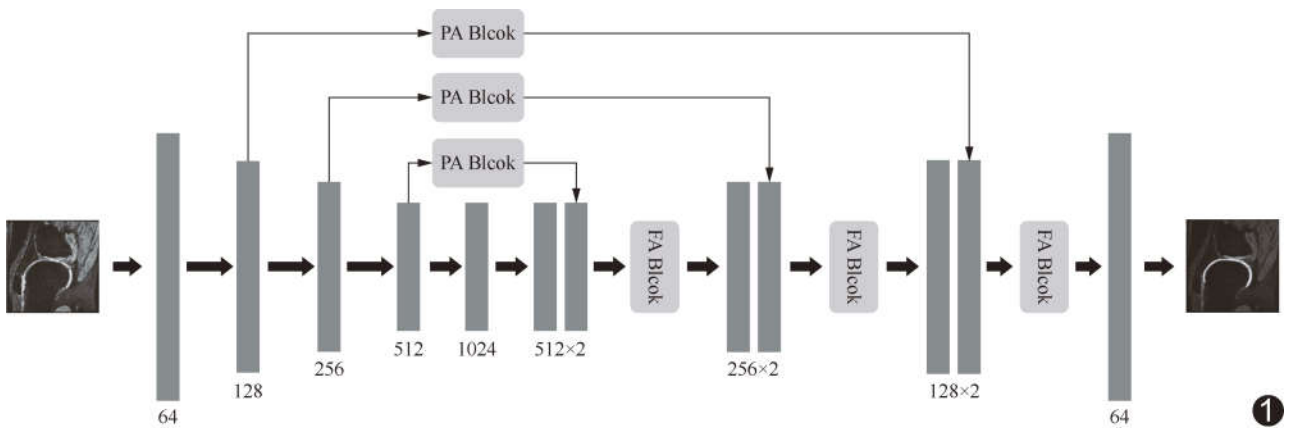


图 1 PA-UNet 结构示意图(PA Block 代表分区注意力模块、FA Block 代表特征聚合模块)

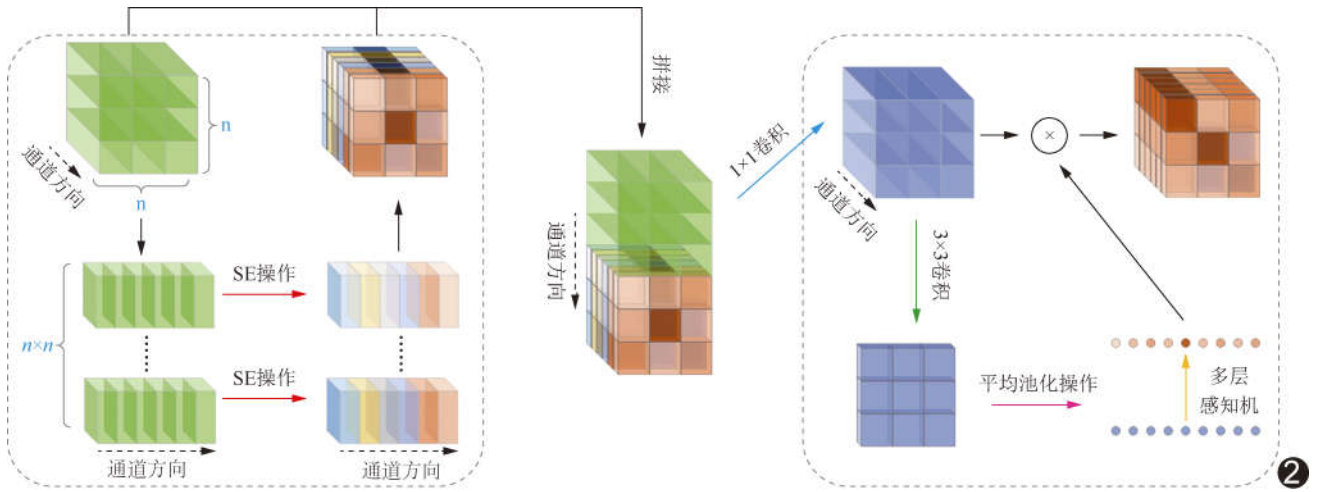


图 2 分区注意力模块结构示意图

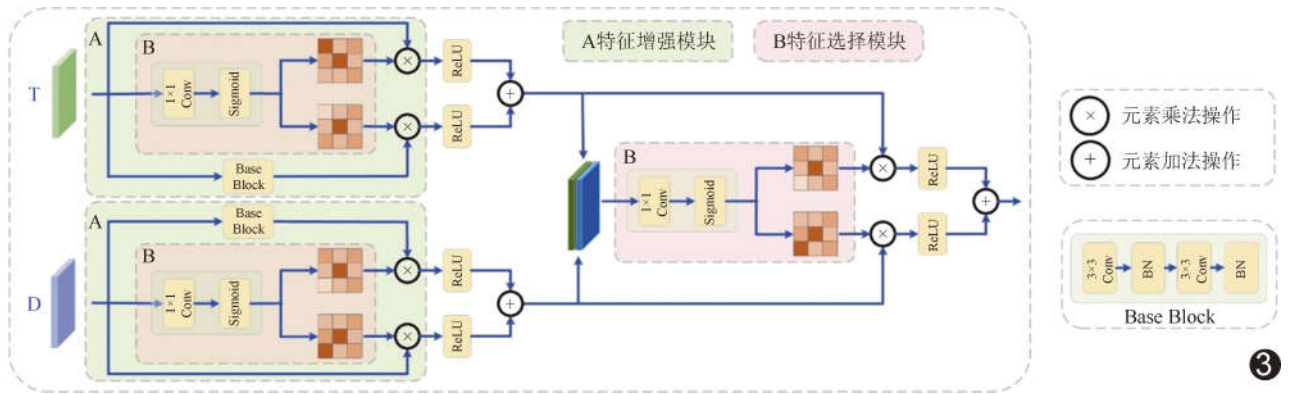


图 3 特征聚合模块示意图

$$T = \eta(P \otimes T) + \eta(Q \otimes \text{Con } v_{3 \times 3}(\text{Con } v_{3 \times 3}(T))) \quad (1)$$

$$[P, Q] = \sigma(\text{Con } v_{1 \times 1}([T, D'])) \quad (2)$$

$$O = \eta(P \otimes T) + \eta(Q \otimes D') \quad (3)$$

其中 η 代表激活函数 ReLU

1.4 训练网络 以 Adam 算法优化目标函数, 设学习率初始值 initial_lr 为 0.02, 根据 $lr = \text{initial_lr} \times \gamma^{\frac{\text{epoch}}{\text{step_size}}}$ 进行衰减(初始值 γ 及 step_size 分别设为 0.7 及 4)^[6-7]。由于模型在 60 轮内收敛, 故将 epoch 设为 60 以确保训练有效。

以 Focal Loss^[4] 及 Dice Loss^[8] 组成混合损失函数, 前者可基于交叉熵损失引入模块化因子 $(1-p_i)^{\gamma}$ 解决训练中的类别不平衡, 后者可将像素级别分类问题转化为类别级别距离问题。将原图像 (x^s) 输入编码器 (E) 和像素级分类器 (C) 以生成预测映射, 与标签 (y^s) 通过总损失优化分割网络, 见公式 (4)。

$$L_{\text{seg}} = L_{\text{focal}}(C_i(E(x^s)), y^s) +$$

$$L_{\text{dice}}(C_i(E(x^s)), y^s) \quad (4)$$

1.5 评估分割效果

1.5.1 主观评价 分别选取低难度、中等难度及困难膝关节 MRI, 观察以 PA-UNet 与 SegNet、UNet 及 DeepLabv3+ 模型分割细小结构的效果。

1.5.2 客观评价 对比分割结果与相应标签, 获得戴斯相似系数 (Dice similarity coefficient, DSC)、交并比 (intersection over union, IoU) 及豪斯多夫距离 (Hausdorff distance, HD) 见公式 (5)。分别以 UNet^[3]、DeepLabv3+^[9]、SA-UNet^[10]、Residual attention UNet (RA UNet)^[11] 及 SegNet 模型^[12] 为参考, 观察 PA-UNet 模型测试阶段的 DSC、IoU 及 HD。

$$\begin{cases} HD = \max(h(|A|, |B|), h(|B|, |A|)) \\ h(|A|, |B|) = \max_{a \in |A|} \{ \min_{b \in B} ||a-b|| \} \\ h(|B|, |A|) = \max_{b \in |B|} \{ \min_{a \in A} ||b-a|| \} \end{cases} \quad (5)$$

1.5.3 消融实验 分别设计基于 UNet、基于 SE (第

2~4 层)的 UNet(UNet+SE)、+UNet、++UNet、+++UNet、+UNet+、++UNet+(1) + 及 PA-UNet 模型的消融实验,观察其分割股骨软骨及膝关节软骨的效果;前缀“+”代表于 UNet 中引入分区注意力模块,后缀“+”代表于 UNet 中引入特征聚合模块,“+”符号的数量代表模块数量。

2 结果

2.1 主观评价 各模型分割简单样本效果未见明显差异。PA-UNet 模型对于中等难度样本可完整分割细小结构,在端点处表现良好,其分割效果优于 SegNet、UNet 及 DeepLabv3+ 模型;对困难样本,PA-UNet 模型可检出目标并完整分割股骨软骨及胫骨软骨,而其他模型均存在欠分割。见图 4。

2.2 客观评价 PA-UNet 模型分割股骨软骨及胫骨软骨的 DSC 及 IoU 均高于,而其 HD 均低于 UNet、DeepLabv3+、SA-UNet、RA UNet 及 SegNet。见表 1。

表 1 PA-UNet 与其他参考模型客观评价指标比较

模型	股骨软骨			胫骨软骨		
	DSC (%)	IoU (%)	HD (mm)	DSC (%)	IoU (%)	HD (mm)
PA-UNet	88.84	77.92	9.11	82.73	72.42	11.74
UNet	86.72	77.85	10.12	81.32	70.85	12.65
DeepLabv3+	87.35	77.42	10.56	81.69	71.33	13.57
SA-UNet	87.28	77.34	10.33	81.54	70.64	12.93
RA UNet	86.65	76.98	11.41	81.13	70.95	13.21
SegNet	86.41	76.75	12.39	80.25	69.32	15.30

2.3 消融实验 以 PA-UNet 模型分割股骨软骨的 DSC 为 88.97%,高于 UNet、UNet+SE、+UNet、++UNet、+++UNet、+UNet+ 及 +UNet++ 模型 (DSC 分别为 86.72%、87.35%、86.86%、88.27%、88.51%、87.29%及 88.74%);以 PA-UNet 模型分割胫骨软骨的 DSC 为 82.72%,高于其他模型 (81.34%、81.67%、81.41%、81.93%、82.34%、

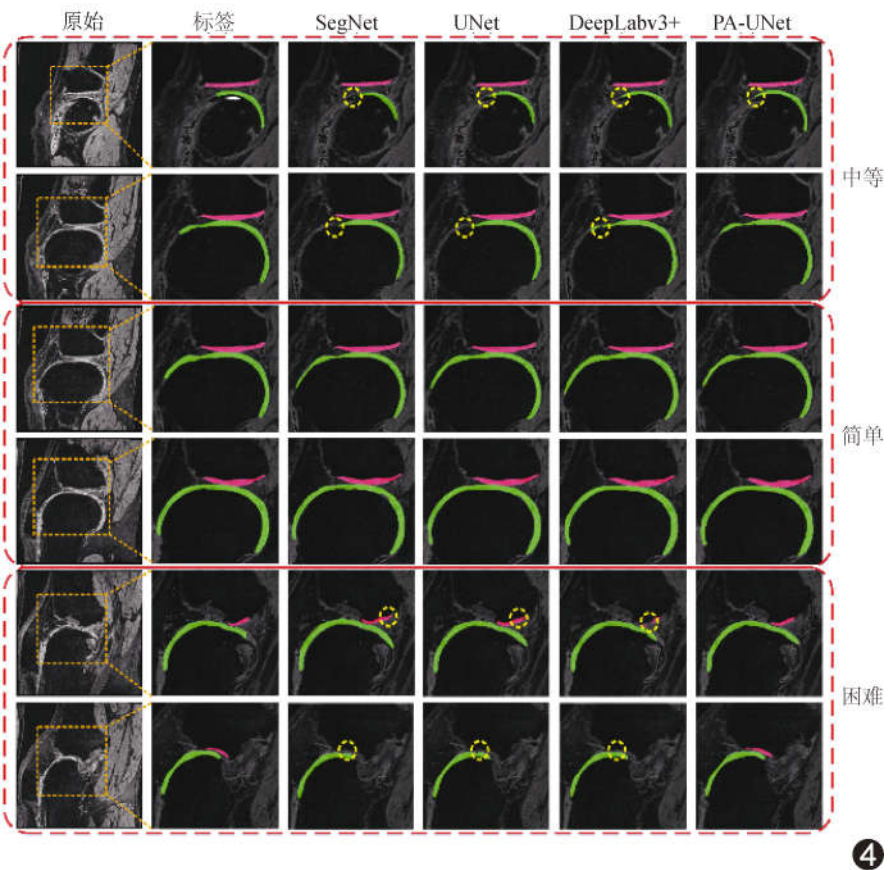


图 4 各模型分割 MRI 中的膝关节软骨示意图 (绿线和红线分别代表胫骨软骨及股骨软骨分割结果,黄圈代表不同模型存在分割差异区域)

81.79%及 82.41%)。分区大小为 32 时,PA-UNet 模型分割股骨软骨及胫骨软骨的 DSC 最高,见表 2。

表 2 不同分区大小 PA-UNet 模型分割股骨软骨及胫骨软骨的 DSC 值(%)

分割部位	16×16	32×32	64×64
股骨软骨	87.02	88.83	86.43
胫骨软骨	81.26	82.73	80.65

3 讨论

基于编码解码架构的 UNet 模型已成为分割医学图像的常用方法,通过引入跳跃连接使高层与低层特征于解码阶段融合,可弥补下采样操作引起的信息损失。引入跳跃连接可有效传递细节信息及上下文信息,提高分割精度及准确性;如 KILIÇ 等^[13]提出的基于 UNet++ 编码解码结构的脊柱分割方法,即通过跳跃连接机制将编码器与解码器中的多层特征进行融合,以增强分割脊柱结构的准确性。

利用注意力机制可通过提取医学图像中的多尺度信息对模型加以训练,亦为分割医学图像的常用方法。

引入注意力机制有助于关注局部信息并结合全局依赖关系抑制无效信息,突显对分割有益的特征,显著提升分割医学图像的性能。

分割膝关节软骨过程中,编码膝关节空间及语义信息、预测得分图均十分重要^[14];但软骨大小、形状存在差异,以膝关节为整体进行图像处理常存在局部信息丢失等问题^[6]。本研究提出 PA-UNet,即在跳跃连接中引入注意力模块,同时对传统通道及空间注意力进行改良,以有效解决因软骨大小、形状差异所致欠分割问题,提高分割准确性和鲁棒性;此法更关注于输入特征中提取的每个局部分区,并通过分区通道注意力模块和分区区间注意力模块增强语义信息及空间细节。不同于传统通道注意力模型,PA-UNet 中的分区通道注意力模块将输入的膝关节 MRI 分割为不同的“块”,以区分“块”级不同通道的重要性,有助于减少冗余特征信息、凸显与膝关节软骨相关的关键通道,避免池化操作所致上下文信息丢失。此外,分区区间注意力模块用于分析特征分区间的联系,可进一步提升分割复杂形状、大小差异较大软骨区域的精确度。

本研究选取的中等难度样本中,股骨软骨面积相对较小而胫骨软骨与股骨软骨总面积适中,分割此类目标时,问题往往出现于软骨两端而非中心。相比 SegNet、UNet 及 DeepLabv3+ 模型,本研究中,PA-UNet 模型分割股骨软骨及胫骨软骨部分均较优异,且在端点处表现良好;相比常用于分割医学图像且具有不同网络结构或注意力模块的 UNet、DeepLabv3+、SA-UNet、RA UNet 及 SegNet 模型,PA-UNet 模型的客观评价指标、尤其 HD 更优。此外,本研究消融实验结果显示,以 PA-UNet 模型分割股骨软骨及胫骨软骨的 DSC 值分别为 88.97% 及 82.72%,均高于 UNet、UNet+SE、+UNet、++UNet、+++UNet、+UNet+ 及 +UNet++ 模型。

综上所述,PA-UNet 可完整分割 MRI 中的膝关节软骨,尤其分割细小结构效果良好。但本研究仅观察二维图像,未充分利用三维图像的结构信息,有待后续进一步完善。

利益冲突:全体作者声明无利益冲突。

作者贡献:王翔设计方案、数据分析、撰写和修改文章;史操研究实施、审阅文章;袁正一统计分析。

[参考文献]

- [1] SHAN T, YAN J. SCA-Net: A spatial and channel attention network for medical image segmentation[J]. IEEE Access, 2021, 9:160926-160937.
- [2] 陈大千,张凡,郝鹏翼,等.结合多尺度通道注意力和边界增强的 2D 医学图像分割[J].计算机辅助设计与图形学学报,2022,34(11):1742-1752.
- [3] AMER A, LAMBROU T, YE X. MDA-Unet: A multi-scale dilated attention U-net for medical image segmentation [J]. Applied Sciences, 2022,12(7):3676.
- [4] LIU F, ZHOU Z, JANG H, et al. Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging [J]. Magnetic resonance in medicine, 2018,79(4):2379-2391.
- [5] SHI C, XU C, BI H, et al. Lateral feature enhancement network for page object detection [J]. IEEE Transactions on Instrumentation and Measurement, 2022,71:1-10.
- [6] SHI C, XU C, HE J, et al. Graph-based convolution feature aggregation for retinal vessel segmentation [J]. Simulation Modelling Practice and Theory, 2022,121:102653.
- [7] XU C, SHI C, BI H, et al. A page object detection method based on mask R-CNN[J]. IEEE Access, 2021,9:143448-143457.
- [8] XU C, SHI C, CHEN Y. End-to-end dilated convolution network for document image semantic segmentation[J]. Journal of Central South University, 2021,28(6):1765-1774.
- [9] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]. Proceedings of the European conference on computer vision (ECCV). 2018:801-818.
- [10] GUO C, SZEMENYEI M, YI Y, et al. SA-UNet: Spatial attention U-Net for retinal vessel segmentation [C]. 2020 25th international conference on pattern recognition (ICPR). 2021:1236-1242.
- [11] CHEN X, YAO L, ZHANG Y. Residual attention U-Net for automated multi-class segmentation of covid-19 chest CT images[J]. arXiv, 2020. <https://doi.org/10.48550/arXiv.2004.05645>
- [12] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE transactions on pattern analysis and machine intelligence, 2017,39(12):2481-2495.
- [13] KILIÇ R, ALÇI M, GÜNAY E. A SC-CNN-based chaotic masking system with feedback [J]. International Journal of Bifurcation and Chaos, 2004,14(1):245-256.
- [14] CANHUI X, YUTENG L, CAO S, et al. HiM: Hierarchical multimodal network for document layout analysis [J]. Applied Intelligence, 2023,53(20):24314-24326.