XXXX DOI:

AMAF-Net:基于自适应多尺度注意力融合网络的冠状动脉精准分割研究

于子腾[1],宋雪桦[1],袁　昕[1],朱海涛[2],曹雄锋[2]
(1.江苏大学,计算机科学与通信工程学院,镇江　212000;2.江苏大学附属医院影像科,镇江　212001)

【摘　要】目的:基于3D-UNet构建一种自适应多尺度注意力融合网络(adaptive multi-attention fusion-network,AMAF-Net),以提升3D-UNet网络的冠状动脉计算机断层扫描血管成像(computed tomography angiography,CTA)图像分割精准性。**方法:**在3D-UNet编码器中引入多尺度膨胀卷积模块(multi-scale dilated convolution,MSDC)和压缩与激励(squeeze-and-excitation,SE)注意力机制;在模型架构的瓶颈层中采用标准3D卷积模块;解码器中加入加性注意力机制(additive attention mechanism,AAM),最终构建AMAF-Net。采用公开自动化冠状动脉分割(automated segmentation of coronary artery,ASOCA)数据集和自制冠状动脉CTA40数据集中的冠状动脉CTA图像作为AMAF-Net及8种经典U-Net网络模型的冠状动脉分割训练、验证和测试数据。上述网络模型的冠状动脉分割性能评价包括指标评价(Dice相似系数、召回率、精确率和平均表面对称距离)、视觉评价和消融实验。不同网络模型的分割性能差异比较采用单因素方差分析及Turkey检验。**结果:**在ASOCA数据集和CTA40数据集中,9种模型间Dice相似系数、召回率、精确率和平均表面对称距离差异皆具有统计学意义(均$P<0.001$)。AMAF-Net的Dice相似系数、精确率和平均表面对称距离最高,召回率相对排名靠前,且各指标与多个模型间差异具有统计学意义(均$P<0.05$)。AMAF-Net在冠状动脉复杂结构区域皆具有最佳的分割效果,且所有分割结果最接近于手动分割金标准真实标签。消融实验中,随着模块数目增加,网络模型的Dice相似系数、召回率、精确率逐渐增加,平均表面对称距离逐渐降低,各模型间所有指标间差异皆具有统计学意义(均$P<0.001$)。**结论:**AMAF-Net的冠状动脉分割精准性得到明显提升,且表现出良好泛化性,具有潜在临床推广应用价值。

【关键词】冠状动脉分割;膨胀卷积;注意力机制;计算机断层扫描血管成像
【中图分类号】R322.1　　**【文献标志码】**A　　**【收稿日期】**2025-04-08

Precise coronary artery segmentation based on an adaptive multi-scale attention fusion-network

Yu Ziteng[1],*Song Xuehua*[1],*Yuan Xin*[1],*Zhu Haitao*[2],*Cao Xiongfeng*[2]
(*1.School of Computer Science and Communication Engineering,Jiangsu University;*
2.Department of Radiology,The Affiliated Hospital of Jiangsu University)

【Abstract】*Objective*: To construct an adaptive multi-scale attention fusion-network (AMAF-Net) based on 3D-UNet, and to improve the precision of the segmentation of coronary computed tomography angiography (CTA) images using 3D-UNet network. ***Methods*:** The multi-scale expanded convolution module (MSDC) and the Squeeze-and-Excitation (SE) attention mechanism were introduced into the 3D-UNet encoder. Standard 3D convolutional modules are adopted in the bottleneck layer of the model architecture. The additive attention mechanism (AAM) is added to the decoder, and finally the AMAF-Net is constructed. The coronary CTA images in the public dataset automated segmentation of coronary artery (ASOCA) and/or the self-made dataset CTA40 were used as the training, validation, and test data for coronary artery segmentation using AMAF-Net and eight classical U-Net network models. The performance of the above network models in coronary artery segmentation was assessed based on Dice similarity coefficient, recall, precision, average surface symmetry distance, visual evaluation, and ablation experiment. A one-way analysis of variance and the Turkey test were used to compare the segmentation performance of different network models. ***Results*:** In the ASOCA and CTA40 datasets, there were significant differences in Dice similarity coefficient, recall, precision, and aver-

作者介绍:于子腾,Email:2222308005@stmail.ujs.edu.cn,研究方向:人工智能,医学图像处理。
通信作者:曹雄锋,Email:caoxiongfeng512@163.com。
基金项目:江苏省社会发展面上资助项目(编号:BE2023759);镇江市重点研发计划社会发展资助项目(编号:SH2023017)。

age surface symmetry distance between the nine models (all $P<0.001$). AMAF-Net had the highest Dice similarity coefficient, precision, and average surface symmetry distance and a relatively high recall, and there were significant differences in each indicator between AMAF-Net and various other models (all $P<0.05$). AMAF-Net had the best segmentation performance in the negative complex structure region of the coronary artery, and all segmentation results were the closest to the gold standard of manual segmentation true label. In the ablation experiment, with the increase in the number of modules, the network model showed gradual increases in Dice similarity coefficient, recall, and precision and a gradual reduction in average surface symmetry distance, and there were significant differences in all indicators among the models (all $P<0.001$). ***Conclusion***: There is a significant improvement in the accuracy of coronary artery segmentation based on AMAF-Net, with good generalization and a potential clinical application value.

【Key words】coronary artery segmentation; dilated convolution; attention mechanism; computed tomography angiography

冠心病是全球发病率和死亡率最高的心血管疾病，其早期精准诊断对改善患者预后至关重要。冠状动脉计算机断层扫描血管成像（coronary computed tomography angiography，CCTA）借非侵入性、高空间分辨率及三维可视化优势，已成为临床评估冠状动脉解剖结构、检测狭窄与斑块的核心成像技术[1-2]。然而，CCTA图像中冠状动脉的自动化分割仍面临严峻挑战：①冠状动脉拓扑结构高度复杂，直径<2 mm的末梢分支与周围心肌组织对比度低，难以有效提取细微结构特征。②冠状动脉个体解剖变异、钙化斑块、血管扭曲导致形态学特征异质性显著，进一步增加分割难度。③部分容积效应与图像噪声易造成细小血管伪影，影响边界精准界定。因此，设计一种冠状动脉分割精度高、鲁棒性更强的算法尤为重要。

近年来，基于深度学习的U-Net网络[3-4]及其扩展模型[脑肿瘤分割Transformer网络（Brain Tumor Segmentation Using Transformer，TransBTS）[5]、体积医学图像分割全卷积神经网络（Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation，VNet）[6]、图像分割编码器-解码器卷积网络（Convolutional Encoder-Decoder Architecture for Image Segmentation，SegNet）[7]、综合注意力卷积神经网络（Comprehensive Attention Convolutional Neural Networks，CA-Net）[8]、曲线结构深度学习分割网络（Deep Learning Segmentation of Curvilinear Structures，CS^{2}Net）[9]、特征融合校正3D卷积神经网络（Feature-Fusion-and-Rectification 3D-UNet，3D-FFR-UNet）[10]、密集连接体积卷积神经网络（Densely-Connected Volumetric ConvNets，DenseVoxNet）[11]]采用编码器-解码器结构，非常适合捕获冠状动脉局部细节信息，在CCTA图像分割中应用尤为广泛。尤其是扩展模型在U-Net基础上融合了不同的模块，一定程度上改善了冠状动脉图像分割分割效果。然而，现有用于CCTA图像分割的U-Net网络模型仍存诸多不足：①多尺度特征融合机制不完善，难以平衡主干血管与末梢分支的上下文信息；②全局语义与局部细节的协同建模不足，导致钙化区域及低对比度边界的分割不完全；③跨设备、多中心数据的泛化性能受限，制约临床大规模应用。本研究在传统3D-UNet框架基础上改进并构建了一种新型自适应多注意力融合网络，包括：①在编码器集成多尺度膨胀卷积模块和压缩与激励（squeeze-and-excitation，SE）注意力机制，通过扩展感受野与动态通道加权，提升模型对不同尺度冠状动脉血管结构的建模能力；②在解码器设计加性注意力机制，优化特征重建过程中的血管结构拓扑连续性，降低冠状动脉局部细节漏分割率；③在网络模型跳跃连接处设计自适应多重注意力融合模块，弥补编码器与解码器在语义层次和空间细节上的差异。此外，本研究在公开数据集ASOCA与自制数据集CTA40中验证并对比了AMAF-Net模型与现有U-Net扩展模型（TransBTS、VNet、SegNet、CA-Net、CS^{2}Net、3D-FFR-UNet）的多中心冠状动脉分割效果，旨在为改善U-Net网络的CCTA图像分割精准性及泛化性提供参考依据。

1 资料与方法

1.1 AMAF-Net网络模型构建

AMAF-Net网络模型具体结构如图1所示，包括编码器（下采样路径）、瓶颈层以及解码器（上采样路径）。

1.1.1 编码器构建 在编码器部分，每一层均嵌入多尺度膨

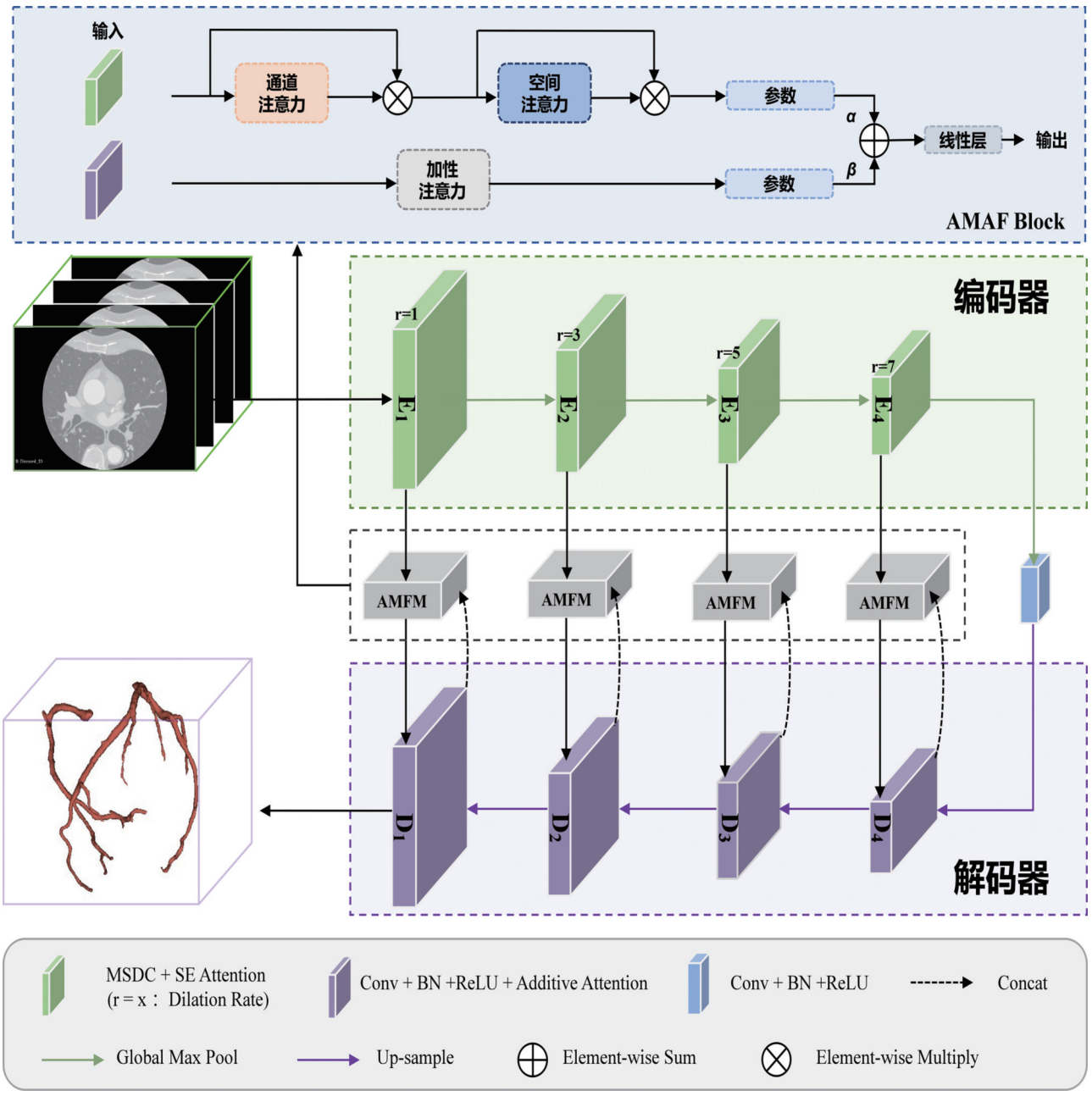


图1　用于冠状动脉分割的AMAF-Net结构示意图

胀卷积模块和SE注意力机制。MSDC模块由多个不同膨胀率的并行膨胀卷积、通道融合卷积以及ReLU激活函数组成，可以充分捕捉多尺度特征，其具体构成，见图2所示。通过使用不同膨胀率的卷积来扩大感受野，有效增强对冠状动脉不同解剖尺度特征的感知能力。MSDC模块每个膨胀卷积的卷积核大小均设置为3×3×3，这种设计能够在三维空间内捕获更全面的结构特征。所有膨胀卷积的步长均设置为1，以确保输出特征图的空间尺寸与输入特征图一致。填充大小设置为与膨胀率r相等的值。为了捕获不同尺度的特征，MSDC模块在每层设计了多个并行膨胀卷积，每个卷积使用不同的膨胀率r。第一层设置r=1等价于标准卷积，主要用于保留局部细节特征和低级特征表示，确保基础信息不丢失。第二层设置r=1，3表示同时捕获局部和中尺度特征，逐步增强模型的感受野。第三层设置r=1，3，5进一步扩展感受野，同时保持对中尺度和大尺度特征的感知。最后一层设置r=1，3，5，7最大限度整合局部细节和全局上下文信息，为瓶颈层提供丰富的多尺度特征表示。每个膨胀卷积层的输出特征在通道维度上进行拼接，形成多尺度特征表示，记为F_{concat}：

$$F_{concat} = Concat(F_{r=1}, F_{r=3}, F_{r=5}, F_{r=7}) \quad (1)$$

式中，$F_{r=x}$表示膨胀率为r=x的卷积层输出，$Concat$表示特征拼接。

拼接后的特征通过1×1×1卷积进行通道融合，进一步优化特征表示并减少冗余，表示为F_{fused}：

$$F_{fused} = Conv_{1\times1\times1}(F_{concat}) \quad (2)$$

融合后的特征通过ReLU激活函数引入非线性，生成最终的特征表示为F_{output}：

$$F_{output} = ReLU(F_{fused}) \quad (3)$$

SE注意力机制模块由通道注意力和空间注意力模块构成，可以自适应调整通道权重，提高模型判别性能，其具体构成，见图3所示。通道注意力模块是先对输入特征进行全局平均池化，提取通道间的全局语义信息；然后通过2层全连接层对特征进行非线性变换，生成每个通道的权重；最后用Sigmoid激活函数对权重进行归一化，并对原始特征进行通道加权。空间注意力模块是先对特征图在通道维度上进行全局最大池化和全局平均池化，提取重要的空间信息；然后将2种池化特征进行拼接后，通过一个7×7×7的卷积操作生成空间权重；最后用Sigmoid激活函数将权重应用到原始特征上，完成空间加权。编码器的注意力特征融合了通道和空间信息，这样能同时增强关键通道和关键空间位置的特征表达能力，为解码和融合提供更多的全局信息支持。

1.1.2 瓶颈层构建 瓶颈层介于在编码器与解码器之间，采用一组3D卷积、批归一化与ReLU激活组成的标准卷积模块，作为编码器与解码器之间的过渡层，为后续的空间结构重建提供语义信息支撑。

1.1.3 解码器构建 解码器部分由多个解码模块组成，采用基于加性注意力机制模块对上采样特征进行加权，增强目标区域表达，并抑制无关特征，从而提高图像分割精度。具体结构，见图4所示。输入解码器的特征图$F_{dec} \in \mathbb{R}^{B\times C\times D\times H\times W}$首先在空间维度$(D,H,W)$上展平为二维矩阵$F'_{dec} \in \mathbb{R}^{B\times(D\cdot H\cdot W)\times C}$。其中，B是批量大小，C是通道数。通过2个独立的可学习线性变换矩阵$W_q, W_k \in \mathbb{R}^{C\times d}$，分别将$F'_{dec}$映射为查询矩阵$Q$和键矩阵$K$，分别表示为$Q,K \in \mathbb{R}^{B\times(D\cdot H\cdot W)\times d}$。其中，$d$为查询和密钥的嵌入维度。查询矩阵$Q$和键矩阵$K$的生成过程提取了解码器特征的低维表示，同时保留了全局的空间上下文信息。为了计算注意力权重，查询矩阵Q与可学习参数向量$w_\alpha \in \mathbb{R}^d$进行点积操作并缩放，生成全局查询注意力向量$\alpha \in \mathbb{R}^{B\times(D\cdot H\cdot W)}$：

$$a = \frac{Q \times w_a}{\sqrt{d}} \quad (4)$$

通过对α应用Softmax函数，可以将其归一化为注意力权重，确保所有空间位置的权重和为1：

$$a_i = \mathrm{Softmax}(a) \quad (5)$$

然后，将注意力权重α重新加权到查询矩阵Q上，从而

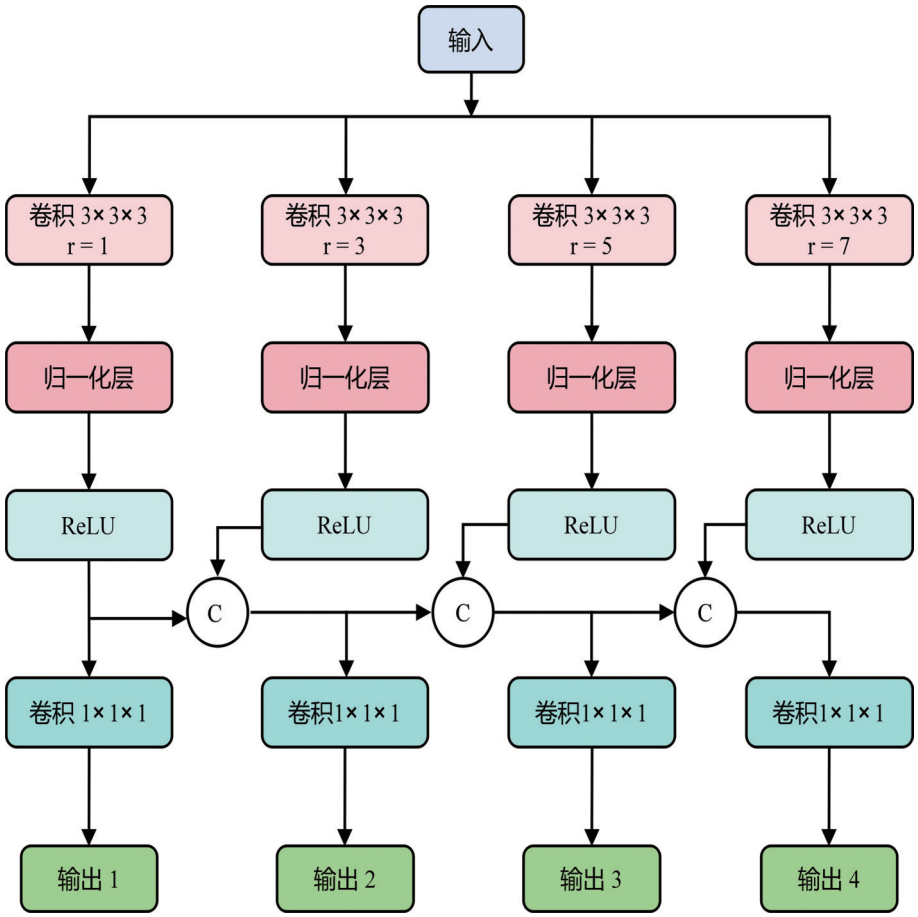


图2 多尺度膨胀卷积模块示意图

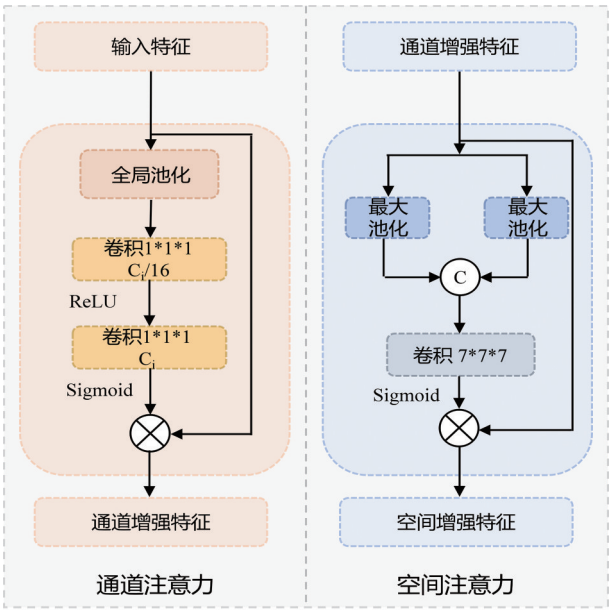


图3 通道注意力和空间注意力结构

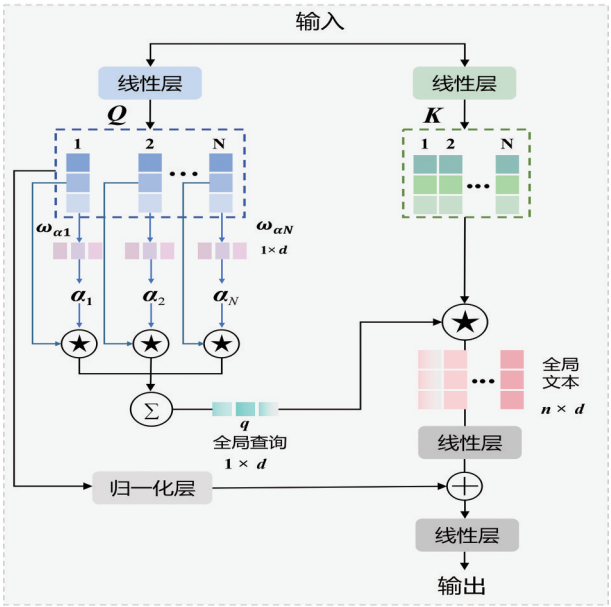


图4 加性注意力机制

生成全局的查询向量$q\in\mathbb{R}^{B\times d}$,表示如下:

$$q=\sum_{i=1}^{D\cdot H\cdot W}\alpha_i\cdot Q_i \tag{6}$$

此时,全局查询向量q集成了所有空间位置的重要信息,突出表征了解码器特征的核心内容。最后将全局查询向量q与键矩阵K进行逐元素相乘,生成权重后的解码特征F_{att}:

$$F_{att}=\mathrm{Proj}(q\odot K) \tag{7}$$

式中,$\odot$表示逐元素相乘,Proj表示后续的线性变换操作,用于将融合后的特征维度还原为与原始解码器特征一致的形状。

1.1.4 损失函数 最终损失函数为联合加权交叉熵(weighted cross entropy,WCE)损失和Dice损失。WCE损失是通过对类别权重的调整来降低背景区域对目标区域学习的干扰。Dice损失通过优化预测区域与真实区域的重叠程度,进一步提高分割结果的一致性。最终损失函数、WCE损失和Dice损失计算公式分别为:

$$L=\lambda L_{Dice}+(1-\lambda)L_{WCE} \tag{8}$$

$$L_{WCE}=-\frac{1}{N}\sum_{i=1}^{N}(\alpha g_i\log p_i+(1-g_i)\log(1-p_i)) \tag{9}$$

$$L_{Dice}=1-\frac{2\sum_{i=1}^{N}p_ig_i+\varepsilon}{\sum_{i=1}^{N}p_i^2+\sum_{i=1}^{K}g_i^2+\varepsilon} \tag{10}$$

$$\alpha=\frac{N-\sum_{i=1}^{N}p_i}{\sum_{i=1}^{N}p_i} \tag{11}$$

式中,λ为权重平衡系数,通过实验设定为0.6,L_{Dice}和L_{WCE}分别表示Dice损失和加权交叉熵损失。α为类别权重因子,用于平衡前景(目标区域)和背景的学习权重。N表示体素的总数,$g_i\in\{0,1\}$和$p_i\in[0,1]$分别为第i个体素的真实值和预测值。ε为一个极小值,用于避免分母为0,在本实验中设置为10^{-6}。

1.2 数据集

本研究所涉及的所有网络模型的冠状动脉分割效果测评(包括训练、验证与测试)均采用2个数据集,包括:公开数据集ASOCA,来自于MICCAI 2020冠状动脉自动分割挑战赛公开数据集,包含40例患者的冠状动脉CT血管成像,其中30例用于训练,5例用于验证,5例用于测试。ASOCA数据由GE 64排LightSpeed CT扫描仪采集。扫描范围为气管隆突下1 cm至心底下1 cm。心电图触发轴位扫描模式。扫描参数:探测器宽度为40 mm,Z轴覆盖范围40 mm,管电压100~120 kV,管电流自调制,螺旋扫描模式,扫描野为20~30 cm,旋转速度0.35 s/r,噪声指数为22,采集层厚0.625 mm,重建层厚0.625 mm,层间距0.625 mm,重建类型为标准重建Stnd+ASiR-V60-80%。采用双筒高压注射器双流注射模式,经肘静脉注射非离子碘对比剂70 mL,对比剂为碘佛醇[35 mgI/100 mL,江苏恒瑞医药股份有限公司]。注射流率为5.0 mL/s,对比剂注射结束后以5.0 mL/s流率注射20 mL生理盐水。采用智能追踪技术(Bolus Tracking)触发扫描,感兴趣区置于升主动脉,对比剂追踪激发阈值为150 HU,延时2 s触发,扫描时病人吸气后屏住呼吸。自制数据集CTA40,来自江苏大学附属医院,包含40例患者的冠状动脉CT血管成像,其中30例用于训练,5例用于验证,5例用于测试。CTA40数据集由GE 256排Revolution CT扫描仪采集。扫描范围为气管隆突下1 cm至心底下1 cm。心电图触发轴位扫描模式。扫描参数:探测器宽度为160 mm,扫描野为20~30 cm,管电压70~100 kV,管电流150~370 mA,旋转时间0.28 s/r,噪声指数为22,采集层厚0.625 mm,重建层厚0.625 mm,层

间距0.625 mm，重建类型为标准重建Stnd+ASiR-V 60%~80%。采用双筒高压注射器双流注射模式，经肘静脉注射非离子碘对比剂70 mL，对比剂为碘佛醇（35 mgI/100 mL，江苏恒瑞医药股份有限公司）。注射流率为5 mL/s，对比剂注射结束后以5.0 mL/s流率注射40 mL生理盐水，感兴趣区置于升主动脉，对比剂追踪激发阈值为200 HU，延时2 s触发，扫描时病人吸气后屏住呼吸。

1.3 模型性能测评实验

1.3.1 测评过程 AMAF-Net网络模型训练采用配备NVIDIA GeForce RTX 3090 GPU的Ubuntu18.04系统的PyTorch软件（版本号：2.4.1）。训练数据采用NPY格式的3D体素图像数据，每例患者的CCTA薄层图像（平均约250层，层厚为0.625 mm）加载后，首先进行三维尺度标准化处理，随后采用滑动窗口裁剪策略，窗口尺寸设定为128×160×160体素，窗口滑动步长为96×120×120体素，以确保对整个影像空间结构的充分覆盖及局部结构连续性。为减少冗余背景区域的干扰，本研究仅选择前景体素比例>0.2%的数据块。通过该方法处理后，共计从30例训练集中获得约16 000个有效三维训练样本。本研究选用改进的3D-UNet作为主干网络，输出通道为2，适用于二分类分割任务。模型超参数见表1所示。初始学习率设为1×10^{-3}，并采用余弦退火重启策略动态调整学习率，初始周期长度T_0=10，周期倍增因子T_mult=2，最小学习率η_min=1×10^{-5}，以增强训练的稳定性和收敛速度。训练总轮数为800轮，批大小为2，训练过程中每轮进行验证评估，每隔50轮保存一次模型快照。所有训练日志均写入文件以便追踪，确保实验过程具有良好的可重复性与可验证性。为了对比突出AMAF-Net网络模型的冠状动脉分割性能，本研究对AMAF-Net及现有8种主流冠状动脉分割网络模型（SegNet、3D-UNet、VNet、TransBTS、CA-Net、DenseVoxNet、CS^{2}Net、3D-FFR-UNet）分别进行了相同的数据预处理流程，并使用一致的学习策略进行模型性能测评实验和优化。

表1 网络超参数设置

参数	取值	参数	取值
优化器	Adam	学习率测量	余弦退火重启
初始学习率	1×10^{-3}	最小学习率	1×10^{-5}
权重衰减	1×10^{-8}	训练轮数	800
批量大小	2	随机裁剪尺寸	128×160×160

1.3.2 性能评价 冠状动脉的分割性能评价包括3种方法，包括：指标评价、视觉评价和消融实验。①指标评价。分别在ASOCA数据集和CTA40数据集中测试所有网络模型的冠状动脉分割性能，测试次数为10次，每次测试记录的评价指标参考先前报道文献，包括Dice相似系数、召回率、精确率和平均表面对称距离4个指标[12]，具体计算公式如下：

$$DSC = \frac{2|G \cap P|}{G \cup P} = \frac{2TP}{2TP + FN + FP} \tag{12}$$

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

$$ASSD = \frac{1}{N}\left(\sum_{g \in T(G)} \min_{q \in T(Q)} \|g - q\| + \sum_{q \in T(Q)} \min_{g \in T(G)} \|q - g\| \right) \tag{15}$$

式中，Q是预测的分割结果，G是真实标签；TP，TN和FP分别表示分割结果的真阳性、假阴性和假阳性的数量；$T(G)$表示真实标签的表面点集合，$T(Q)$表示分割结果的表面点集合。g和q分别是$T(G)$和$T(Q)$中的点；$\|g - q\|$表示点g和q之间的欧几里得距离；N是表面点的总数，用于对整体距离进行归一化。②视觉评价。从ASOCA数据集中选择3例具有代表性的病例数据进行手动分割及网络模型自动分割，以1名具备5年以上工作经验的影像科医师的手动分割冠状动脉图像结果作为“金标准”真实标签（Ground Truth），对比4种经典模型（CS^{2}Net、3D-UNet、VNet、SegNet）及AMAF-Net模型的冠状动脉分割视觉效果。③消融实验。为了测试AMAF-Net网络各模块的有效性，本研究在ASOCA数据中，以3D-UNet作为基线网络模型，分别测试依次添加MSDC、SE、AAM模块的冠状动脉分割效果，评价指标为Dice相似系数、召回率、精确率和平均表面对称距离。

1.4 统计学分析

采用SPSS 25.0统计软件对数据进行分析。采用Shapiro-wilk检验对所有评价指标进行正态性检验，符合正态分布的评价指标以均数±标准差（$\bar{x} \pm s$）表示。不同网络模型在ASOCA和CTA40数据集中的冠状动脉分割性能评价指标差异比较，以及基于ASOCA数据集的各模块的消融实验结果评价指标差异比较采用方差分析。AMAF-Net与其他模型的冠状动脉分割性能评价指标差异比较，以及基于ASOCA数据集的AMAF-Net完整模块与其它模块的消融实验结果评价指标差异比较采用Turkey检验分析本研究模型。检验水准α=0.005。

2 结 果

2.1 模型指标评价

在ASOCA数据集中的冠状动脉分割性能指标测评结果发现（表2），在9种网络模型中，AMAF-Net的Dice相似系数、精确率和平均表面对称距离最高，9种模型间3个指标差异具有统计学意义（均P<0.001）。AMAF-Net的Dice相似系数与6种模型（SegNet、3D-UNet、VNet、TransBTS、CA-Net、DenseVoxNet）中任意1种差异具有统计学意义（均P<0.05）。AMAF-Net的精确率和平均表面对称距离与其它7种模型（SegNet、3D-UNet、VNet、TransBTS、CA-Net、DenseVoxNet、CS^{2}Net）中任意1种差异皆具有统计学意义（均P<0.05）。9

种模型间召回率差异具有统计学意义(均P<0.001),其中,3D-FFR-UNet最高,AMAF-Net次之,但2者间差异无统计学意义(均P>0.05)。此外,AMAF-Net的召回率与6种模型(SegNet、3D-UNet、VNet、TransBTS、CA-Net、DenseVoxNet)中任意一种差异具有统计学意义(均P<0.05)。

在CTA40数据集中的冠状动脉分割性能指标测评结果发现(表3),在9种网络模型中,AMAF-Net的Dice相似系数、精确率和平均表面对称距离最高,9种模型间3个指标差异具有统计学意义(均P<0.001)。AMAF-Net的Dice相似系数、精确率和平均表面对称距离与其它8种模型中任意1种差异皆具有统计学意义(均P<0.05)。9种模型间召回率差异具有统计学意义(均P<0.001),其中,CS^{2}Net最高,AMAF-Net第3,但2者间差异无统计学意义(均P>0.05)。此外,AMAF-Net的召回率与3种模型(SegNet、3D-UNet、VNet)中任意1种差异具有统计学意义(均P<0.05)。

2.2 视觉评价

视觉评价结果显示(图5),在3个病例冠状动脉分割结果中,与其它4个模型相比,AMAF-Net在冠状动脉负复杂结构区域(黄色、绿色虚线框)皆具有最佳的分割效果,且所有分割结果最接近于手动分割金标准真实标签,而其它模型在存在过度分割或分割不足分割问题。

2.3 消融实验

基于ASOCA数据集的各模块消融实验结果发现(表4),随着模块数目增加,网络模型的Dice相似系数、召回率、精确率逐渐增加,平均表面对称距离逐渐降低,各模型间所有性能评价指标间差异皆具有统计学意义(均P<0.001)。在5个模型中,Model 4(AMAF-Net)具有最高的Dice相似系数和精确率,排名第2的召回率,以及最低的平均表面对称距离。AMAF-Net的Dice相似系数和精确率与其它4个模型间差异皆具有统计学意义(均P<0.05)。AMAF-Net的召回率与基线模型3D-UNet和Model 1差异皆具有统计学意义(均P<0.05),而与Model 2和Model 3差异无统计学意义(均P>0.05)。AMAF-Net的平均表面对称距离与基线模型3D-UNet、Model 1和Model 2差异皆具有统计学意义(均P<0.05),而与Model 3差异无统计学意义(P>0.05)。

表2 不同模型在ASOCA数据集上的冠状动脉分割性能比较

数据集	模型	Dice相似系数(%)	召回率(%)	精确率(%)	平均表面对称距离(mm)
ASOCA	Segnet	80.21 ± 0.16[a]	83.46 ± 0.14[a]	70.16 ± 0.19[a]	1.63 ± 0.01[a]
	3D-UNet	82.54 ± 0.05[a]	86.37 ± 0.10[a]	72.52 ± 0.33[a]	1.37 ± 0.01[a]
	VNet	83.44 ± 0.08[a]	87.39 ± 0.16[a]	73.61 ± 0.14[a]	1.32 ± 0.01[a]
	TransBTS	84.57 ± 0.11[a]	88.39 ± 0.31[a]	78.63 ± 0.39[a]	1.22 ± 0.01[a]
	CA-Net	85.92 ± 0.12[a]	89.13 ± 0.14[a]	79.87 ± 0.14[a]	1.19 ± 0.01[a]
	DenseVoxNet	85.97 ± 0.11[a]	89.33 ± 0.14[a]	78.57 ± 0.27[a]	1.12 ± 0.01[a]
	CS^{2}Net	86.66 ± 0.15	89.42 ± 0.15	80.06 ± 0.24[a]	0.97 ± 0.01[a]
	3D-FFR-UNet	86.24 ± 0.13	**91.23 ± 0.20**	80.75 ± 0.20	0.81 ± 0.01
	AMAF-Net	**87.16 ± 0.10**	90.17 ± 0.21	**81.54 ± 0.13**	**0.76 ± 0.01**
	F	404.948	164.115	290.500	521.766
	P	<0.001	<0.001	<0.001	<0.001

注:加粗值表示对应指标最大值;a,与AMAF-Net比较,P< 0.05

表3 不同模型在CTA40数据集上的冠状动脉分割性能比较

数据集	模型	Dice相似系数(%)	召回率(%)	精确率(%)	平均表面对称距离(mm)
CTA40	Segnet	76.34 ± 0.08[a]	82.93 ± 0.18[a]	71.41 ± 0.23[a]	1.81 ± 0.02[a]
	3D-UNet	79.91 ± 0.11[a]	84.64 ± 0.20[a]	74.62 ± 0.12[a]	1.52 ± 0.02[a]
	VNet	81.69 ± 0.15[a]	86.47 ± 0.32[a]	77.12 ± 0.20[a]	1.49 ± 0.17[a]
	TransBTS	83.13 ± 0.08[a]	87.91 ± 0.11	79.82 ± 0.17[a]	1.35 ± 0.01[a]
	CA-Net	84.22 ± 0.24[a]	88.78 ± 0.15	79.69 ± 0.18[a]	1.21 ± 0.01[a]
	DenseVoxNet	84.76 ± 0.09[a]	89.16 ± 0.19	79.53 ± 0.31[a]	1.12 ± 0.02[a]
	CS^{2}Net	85.32 ± 0.07[a]	**89.98 ± 0.11**	79.15 ± 0.19[a]	1.11 ± 0.01[a]
	3D-FFR-UNet	85.4 ± 0.08[a]	89.37 ± 0.32	82.19 ± 0.27[a]	0.96 ± 0.01[a]
	AMAF-Net	**86.44 ± 0.11**	88.83 ± 0.23	**83.27 ± 0.16**	**0.88 ± 0.01**
	F	666.235	126.368	305.243	447.314
	P	<0.001	<0.001	<0.001	<0.001

注:加粗值表示对应指标最大值;a,与AMAF-Net比较,P< 0.05

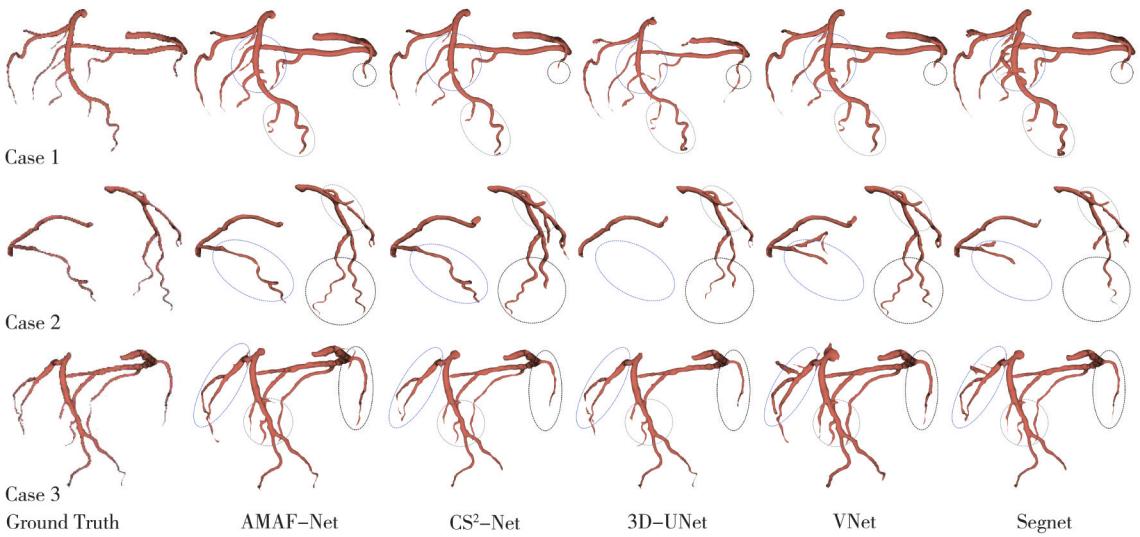


图5 可视化分割结果

表4 基于ASOCA数据集的各模块消融实验结果

模型	模块			Dice相似系数(%)	召回率(%)	精确率(%)	平均表面对称距离(mm)
	MSDC	SE	AAM				
基线	–	–	–	82.54 ± 0.05[a]	86.37 ± 0.10[a]	72.52 ± 0.33[a]	1.37 ± 0.01[a]
Model 1	✓	–	–	84.71 ± 0.11[a]	87.66 ± 0.23[a]	77.25 ± 0.39[a]	1.19 ± 0.04[a]
Model 2	✓	✓	–	85.29 ± 0.14[a]	88.19 ± 0.14	79.33 ± 0.18[a]	1.04 ± 0.02[a]
Model 3	✓	–	✓	85.92 ± 0.11[a]	**89.22 ± 0.39**	80.38 ± 0.16[a]	0.96 ± 0.04
Model 4	✓	✓	✓	**86.44 ± 0.13**	88.83 ± 0.28	**83.27 ± 0.25**	**0.88 ± 0.04**
F				171.797	19.762	211.573	37.143
P				<0.001	<0.001	<0.001	<0.001

注：✓表示使用该模块，–则表示未使用该模块，加粗值表示对应指标最大值；a，与对照组AMAF-Net比较，$P<0.05$

3 讨 论

在冠状动脉分割任务中，模型性能的提升关键在于如何有效融合编码器和解码器的特征，以准确捕捉冠状动脉的复杂结构。本研究提出的AMAF-Net，通过创新性地在3D-UNet中引入自适应多重注意力融合模块，明显优化了特征融合过程，并取得了良好的冠状动脉分割效果。这一发现与现有U-Net改良网络模型所取得的分割效果一致[3-11]。值得注意的是，AMAF-Net模型在编码器中整合了SE机制，以强化特征通道的权重分配，增强模型对关键信息的敏感度。同时，由于冠状动脉血管细小且结构复杂，直接依赖卷积操作较难捕获特征图中远距离位置间的全局依赖关系，导致目标区域的细小血管恢复不完整。因此，本课题组在AMAF-Net模型解码器中设计了AAM，增强了模型的全局空间感知能力，优化了细节信息的表达。以上设计避免了直接使用跳跃连接可能导致的特征冲突或信息冗余问题，从而提升了3D-UNet网络模型的冠状动脉分割性能。

本研究一系列实验结果表明，AMAF-Net明显提升了U-Net网络的冠状动脉分割效果。在指标评价方面，AMAF-Net在2个不同的数据集(ASOCA和CTA40)上均展现出了优异的分割效果，表现出了更强的精准性和泛化性。相较于其它8种主流冠状动脉分割网络模型，AMAF-Net在Dice相似系数、精确率和平均表面对称距离等核心指标上均有明显提升且排名最高，尽管召回率稍低于排名最高模型[ASOCA：AMAF-Net vs. 3D-FFR-UNet：(90.17 ± 0.21)% vs. (91.23 ± 0.20)%；CTA40：AMAF-Net vs. CS²Net：(88.83 ± 0.23)% vs. (89.98 ± 0.11)%]，但2者间差异无统计学意义。以上结果提示，自适应多重注意力融合模块可以精准结合冠状动脉分割过程中的全局语义信息与局部细节，从而提高AMAF-Net的冠状动脉分割性能。在视觉评价方面，相较于4种经典模型(CS²Net、3D-UNet、VNet、SegNet)，

本研究提出的AMAF-Net视觉评价效果表现更佳，尤其是在处理冠状动脉的复杂形态和细微结构时。具体机制为：AMAF-Net中MSDC模块增强了模型对大范围背景信息的感知能力，注意力机制有效地提升了对关键区域的关注，融合模块在细节捕捉和全局上下文信息整合上得到了明显提升，进而避免了在血管形态复杂的区域出现模糊或漏分的情况，实现了每个冠状动脉分支的高度完整性与连续性。在消融实验方面，实验结果进一步验证了自适应多重注意力融合模块的有效性。随着模块数目的增加，AMAF-Net网络模型的各项性能指标均逐步提升，表明多重注意力融合策略与多尺度卷积在应对冠状动脉复杂形态时具有明显的协同增益效应。这一结果不仅支持了AMAF-Net设计的合理性，也为其在冠状动脉分割任务中的优异表现提供了有力证据。

本研究存在若干局限性：①模型训练与验证主要基于ASOCA和CTA40 2个数据集，尽管已覆盖公开与自建多中心数据，但其在不同设备、扫描协议和种族人群等更广泛场景下的泛化能力仍需进一步验证。②本研究使用的2个数据集提供的CCTA图像训练、验证、测试数据虽然也满足深度学习网络训练的基本数据量要求，但病例数仍相对较少，未来需增加样本量以进一步验证模型的稳定性。③尽管引入了多尺度膨胀卷积和多注意力机制以提升模型的特征表达能力，但在极端解剖形态（如严重钙化、血管断裂、远端狭窄等）中，模型仍可能存在对微小病灶的漏检风险，这主要由于小区域异常在整体特征中的表达不充分，且缺乏结构先验的支持。未来研究中可考虑引入血管中心线先验、基于图神经网络的结构建模、形态连续性约束等方式，进一步提升模型对异常冠脉形态的敏感性与鲁棒性。④虽然AMAF-Net在精度方面实现了性能提升，但相较基础网络模型3D-UNet，其模型结构略为复杂，模型计算负担较重，推理时间较长，内存占用较多。未来可以考虑通过模型剪枝、量化或蒸馏技术等方法来进一步优化模型效率，从而减少计算资源的消耗，提升推理速度，探索在轻量化硬件平台下的部署优化，以促进模型向真实临床场景的迁移与应用。

综上所述，AMAF-Net在冠状动脉分割任务中表现出了良好的精准性和泛化性，为基于U-Net网络的冠状动脉CT血管成像图像的精准分割提供了一种新的网络模型，该模型在特征融合和注意力机制应用方面取得了显著进展，为医学图像处理领域的研究提供了新的思路和方法。随着进一步的优化和验证，AMAF-Net有望在临床应用中实现更广泛的推广，为患者提供更精确的诊疗支持。

参 考 文 献

[1] Maria T, Elias D. Long-Term Coronary Artery Disease Risk Prediction with Machine Learning Models[J]. Sensors, 2023, 23(3): 1193-1193.

[2] Brown JC, Gerhardt TE, Kwon E. Risk Factors for Coronary Artery Disease. In: StatPearls, 2023, StatPearls Publishing: 1-129.

[3] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C] Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015, 2015, Cham. Springer International Publishing: 234-241.

[4] Çiçek Ö, Abdulkadir A, Lienkamp SS, et al. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation[J]. CoRR, 2016, abs/1606.06650.

[5] Wang W, Chen C, Ding M, et al. TransBTS: Multimodal Brain Tumor Segmentation Using Transformer[C] Medical Image Computing and Computer Assisted Intervention-MICCAI 2021, 2021, Cham. Springer International Publishing: 109-119.

[6] Milletari F, Navab N, Ahmadi SA. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation[J]. CoRR, 2016, abs/1606.04797.

[7] 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.

[8] Gu R, Wang GT, Song T, et al. CA-Net: Comprehensive Attention Convolutional Neural Networks for Explainable Medical Image Segmentation[J]. IEEE Transactions on Medical Imaging, 2020, 40(2): 699-711.

[9] Mou L, Zhao YT, Fu HZ, et al. CS^2-Net: Deep Learning Segmentation of Curvilinear Structures in Medical Imaging[J]. Medical Image Analysis, 2021, 67: 101874.

[10] Song AL, Xu LS, Wang L, et al. Automatic Coronary Artery Segmentation of CCTA Images with an Efficient Feature-Fusion-and-Rectification 3D-UNet[J]. IEEE Journal of Biomedical and Health Informatics, 2022, 26(8): 4044-4055.

[11] Yu L, Cheng JZ, Dou Q, et al. Automatic 3D Cardiovascular MR Segmentation with Densely-Connected Volumetric ConvNets[C] Medical Image Computing and Computer-Assisted Intervention-MICCAI 2017, 2017, Cham. Springer International Publishing: 287-295.

[12] Dong C, Dai D, Li Z, et al. A Novel Deep Network with Triangular-Star Spatial-Spectral Fusion Encoding and Entropy-Aware Double Decoding for Coronary Artery Segmentation[J]. Information Fusion, 2024, 112102561-102561.

（责任编辑：周一青）