ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 40,No. 2,Mar. 2025,pp. 469－484
DOI:10. 16337/j. 1004-9037. 2025. 02. 015

http://sjcj. nuaa. edu. cn
E-mail:sjcj@nuaa. edu. cn
Tel/Fax:＋86-025-84892742

基于MSC-LSAM的多尺度交叉超声医学图像分割方法

王朝欣，杨汶汶，戎　泽，李铮昱，王　行，马　磊

(南通大学信息科学技术学院,南通 226019)

摘　要：脑卒中是全球范围内致死致残率最高的疾病之一,颈动脉狭窄和心脏病变是缺血性脑卒中的重要致病因素。超声(Ultrasound,US)是检查由颈动脉狭窄和心脏病变引起的缺血性脑卒中的常用影像学手段,但超声图像噪声多、边界模糊,具有较高的分割难度。本文提出MSC-LSAM算法,一种多尺度交叉的双编码器超声图像分割网络,旨在实现颈动脉腔体和心脏腔体的快速、准确分割,辅助医生完成疾病诊断。MSC-LSAM在编码器部分并行了分割一切模型(Segment anything model,SAM)的视觉编码器和UNet编码器,在解码器部分采用UNet解码器。本研究首先冻结了预训练的SAM视觉编码器,并在Transformer层中引入高效的适配器(Adapter)块,被称可学习的分割一切模型(Learnable SAM,LSAM)。LSAM在拥有较低参数量的同时,保留学习能力和高度泛化性。然后,在UNet全局网络引入多尺度交叉注意力(Multi-scale cross-axial attention,MCA),实现多尺度特征的交叉融合,有效提升边缘分割能力,抑制模型过拟合。最后,通过高效通道注意力(Efficient channel attention,ECA)实现双编码器多尺度特征的高效融合,减少模型误分割。结果表明,本研究提出的MSC-LSAM在心脏超声公开数据集CAMUS和颈动脉超声自建数据集CAUS上均取得了良好的效果。CAMUS的两心腔(2CH)和四心腔(4CH)数据集分割的平均Dice相似系数(Dice similarity coefficient,DSC)分别达到0.927和0.934;CAUS数据集的平均DSC达到0.917。MSC-LSAM在颈动脉腔体和心脏腔体超声图像分割任务上获得了良好的分割准确度,高于主流分割算法,具有良好的应用前景。

关键词：缺血性脑卒中;超声图像分割;分割一切模型;多尺度交叉注意力;高效通道注意力

中图分类号：R318;TP391.41　　**文献标志码**：A

Multi-scale Crossed Algorithm for Ultrasound Medical Image Segmentation Based on MSC-LSAM

WANG Zhaoxin, YANG Wenwen, RONG Ze, LI Zhengyu, WANG Xing, MA Lei

(School of Information Science and Technology, Nantong University, Nantong 226019, China)

Abstract: Stroke is one of the leading causes of death and disability around the world. Carotid artery stenosis (CAS) and cardiac lesions are important contributing factors to ischemic stroke, and ultrasound imaging has shown great potential in diagnosing ischemic stroke caused by CAS and cardiac lesions. But ultrasound images present significant segmentation challenges due to noise and blurred boundaries. To address this issue, the MSC-LSAM algorithm, a multi-scale crossed dual encoder network for ultrasound image segmentation is proposed. It aims to achieve rapid and accurate segmentation of carotid and cardiac

基金项目:江苏省产学研合作项目(FZ20220736)。
收稿日期:2024-07-23;**修订日期**:2024-09-13

cavities, assisting physicians in disease diagnosis. In the MSC-LSAM, the encoder part parallels a segment anything model (SAM) vision encoder and an UNet encoder, while the decoder part utilizes an UNet decoder. In the SAM image encoder, we froze the pretrained SAM image encoder and introduce efficient adapter blocks in Transformer layers, referred to as learnable SAM (LSAM). LSAM maintains learning capability and high generalization ability while having a low number of parameters. In the global UNet network, we incorporate the multi-scale cross-axial attention (MCA) blocks to achieve cross-fusion of multi-scale features between different axes, effectively enhancing edge segmentation capabilities and suppressing model overfitting. Following the parallel encoders, the efficient channel attention (ECA) block is added to enable integration of multi-scale features from dual encoders, reducing incorrect segmentation caused by feature level mismatches. MSC-LSAM achieves good performance on both the publicly available cardiac ultrasound dataset of CAMUS and the self-constructed carotid artery ultrasound dataset of CAUS. Average dice similarity coefficients (DSCs) for the segmentation of the two-chamber (2CH) and four-chamber (4CH) datasets in CAMUS reach 0.927 and 0.934, respectively; while the average DSC for the CAUS dataset reaches 0.917. MSC-LSAM achieves good segmentation accuracy in tasks of carotid lumen and cardiac chamber ultrasound image segmentation, surpassing mainstream segmentation algorithms, and shows promising application prospects.

Key words: ischemic stroke; ultrasound image segmentation; segment anything model (SAM); multi-scale cross-axial attention (MCA); efficient channel attention (ECA)

引　言

脑卒中是一个全球重点关注的公共卫生问题，具有高发病率、高致死致残率和高复发率等特点，一旦发生将造成不可逆的损伤，严重降低患者的生命质量[1]。缺血性脑卒中约占脑卒中总数的80%，及时识别缺血性脑卒中的风险因素并采取相应的治疗措施可以显著降低其发生率和复发率，具有重要意义[2]。

颈动脉疾病和心脏疾病是引起大动脉粥样硬化(Large artery atherosclerosis，LAA)型和心源性梗死(Cardioembolic，CE)型缺血性脑卒中的两个主要原因[3-4]。LAA型脑卒中主要由颈动脉和颅内大动脉粥样硬化型狭窄引起，颈动脉狭窄、闭塞，引发大脑供血不足，造成缺血性脑卒中[5]；CE型脑卒中主要由心房颤动引起，心房颤动导致心房血流缓慢、停滞，增加了血栓形成的风险，血栓进入循环系统，导致大脑血管堵塞[6]。

超声检查具有成像速度快、无电离辐射及价格较低等优点，临床上多用超声(Ultrasound，US)检测颈动脉狭窄和心脏疾病[7]。心脏超声可以观察患者的心脏结构和功能，心腔容积和直径可以辅助评估患者发生CE型缺血性脑卒中的风险[8]。颈动脉超声可以检测颈动脉狭窄程度和斑块的存在等缺血性脑卒中风险指标，为评估LAA型脑卒中风险提供重要参考[9]。

超声图像分割是一项重要工作，心腔分割和颈动脉管腔分割可以辅助医生定量地评估患者心脏和颈动脉状况。超声图像的手动分割过程繁琐且很大程度上依赖于医学专家的经验[10]。近年来，深度学习技术的发展为超声图像分割带来了新的解决方案[11]。但是，超声图像存在对比度低、噪声多、图像模糊及伪影多等问题，显著影响了分割精度[12]。

针对超声图像分割困难的问题，本文提出一种具有高鲁棒性和泛化能力的算法MSC-LSAM，旨在从复杂的心脏和颈动脉超声影像中，实现心脏腔体和颈动脉管腔的精准分割，为缺血性脑卒中早期诊断提供辅助。

1 相关工作

1.1 医学图像分割

传统的医学图像分割算法主要基于边缘检测法、主动轮廓法和水平集算法等传统算法,这些算法大部分都是半自动方法,需要人工事先划分感兴趣区域,且易受噪声等因素的影响,分割精度较低[13-15]。为了克服传统算法的局限性,研究人员转向机器学习和深度学习算法的研究中。机器学习特征提取能力和复杂结构处理能力不强,在医学图像分割中表现出的效果较差[16]。

卷积神经网络(Convolutional neural network,CNN)的出现使得基于深度学习的医学影像分割性能得到进一步提升。U-Net是一种被广泛应用于医学影像分割任务的卷积神经网络架构,其独特的编码-解码结构以及跳跃连接的设计使其在分割任务中表现出色[11]。ResUNet引入了ResNet的思想,解决了U-Net在处理更深层次网络时可能遇到的梯度消失问题,同时提高了模型对图像细节的捕捉能力[17]。UNet++在UNet种引入了密集跳跃连接和级联连接的概念,使得网络能够更好地利用多尺度和多层次的特征,增强特征融合和信息传递[18]。

近期,Transformer和各种注意力机制被广泛运用在医学图像分割领域,使得脑卒中病灶分割算法的性能得到了进一步的提高。Transformer运用自注意力机制,使得神经网络能够捕捉全局信息,在处理长序列数据时表现出色[19]。Swin-Unet运用Swin Transformer模块代替UNet的卷积块,使网络能够更好地捕获图像中的长距离依赖关系和局部特征,提升了分割的准确性、稳定性和鲁棒性[20]。轴向注意力(Axial attention)是对自注意力的改进与升级,将传统的自注意力机制分解为沿多个维度的多个独立的注意力计算,允许模型捕捉长距离的依赖关系,而不会显著增加计算负担。Swin-PSAxialNet将轴向注意力与nn-UNet结合并引入了参数共享机制,使网络在保持较低计算复杂度的同时,取得了较好的分割效果[21]。

多尺度医学图像分割算法通过捕获不同尺度特征,以更好地适应医学图像中的复杂性和变异性。N-Net通过双编码器提取多尺度特征,并引入挤压与激励(Squeeze-and-excitation,SE)模块来实现多尺度信息融合[22]。MCA-UNet引入多尺度交叉跳跃连接(Multi-scale cross skip connection),将不同编码器阶段经过调整尺寸后的特征传递给解码器所有阶段,在分割具有不规则边界和大小变化的病灶时取得了较好的效果[23]。

1.2 分割一切模型

分割一切模型(Segment anything model,SAM)的出现为医学分割提供了方向。SAM由1个视觉编码器、1个灵活的提示编码器和1个快速掩码解码器组成,基于大规模分割数据集SA-1B进行训练,在通用分割任务和Zero-Shot任务中表现出了良好的性能和泛化能力[24]。Mobile SAM是从原始SAM中蒸馏得到的,将原始视觉编码器中的知识迁移到轻量级的SAM中,其大小仅为原始SAM的1/60,但性能相当[25]。

SAM的视觉编码器使用了基于掩码自动编码器(Masked autoencoders,MAE)[26]预训练的Vision Transformer(ViT)[27],并引入了具有高度可扩展性的预训练权重,使模型具有高度泛化性,使得SAM的视觉编码器不需要训练即可在通用分割任务上取得不错的效果[24]。

由于医学图像的复杂性,尽管SAM在通用分割任务中表现出色,但其在医学图像分割中的效果并不理想。MedSAM是在原始SAM的基础上,针对医疗图像进行微调的版本,与未经过微调SAM相比,Dice相似系数(Dice similarity coefficient, DSC)提高了22.51%[28]。Wu等[29]提出Medical SAM Adapter,采用Adapter技术对SAM进行微调,增强其在医疗领域的表现,在特定数据集上分割效果能够超越当前最先进的医学图像分割方法。nnSAM并行了预训练的SAM编码器和UNet编码器,在获得强大

特征提取能力的同时能够更好地适应不同的图像分割任务[30]。

1.3 超声图像分割

针对超声对比度低、噪声多、伪影多和边界模糊等问题，研究者们采用深度学习来实现超声图像的分割。如MFP-UNet将金字塔网络与扩展U-Net模型相结合，增加两个下采样层以提取图像中更密集的细节，并使用Niblack的全局阈值法对心脏超声图像进行预处理以增强对比度[31]；LU-Net运用级联网络，首先定位感兴趣区域(Region of interest, ROI)，然后在感兴趣区域内对左心室进行分割。在实际应用中，ROI的大小必须人为设定，严重影响了模型的检测和分割性能[32]；Liu等[33]在UNet中引入一个金字塔形的局部注意机制来捕捉上下文信息，并设计了一种标签一致性学习机制，以提高左心室边缘像素的分类精度。

Zhou等[34]提出一种基于FCN的颈动脉管腔分割网络。将体素的先验知识应用到算法中，并在解码器模块中引入了注意机制，使模型动态地关注显著特征。NvNet在UNet中加入全尺寸连接，在第1层和第2层得到的特征映射中加入最大池化，并利用第4层得到的特征映射进行上采样[35]。CSM-Net在编码器的最后一层提出了级联的多重扩张卷积，用三重空间注意力(Triple spatial attention，TSA)模块作为各解码器层跳跃融合后的自注意，实现了颈动脉超声的精准分割[36]。

2 方法

2.1 架构概述

本研究提出一种超声医学图像分割算法，称为多尺度交叉可学习的SAM(Multi-scale crossing learnable SAM，MSC-LSAM)，网络总体架构如图1所示。MSC-LSAM的编码器部分并行了冻结的SAM图像编码器和UNet编码器，解码器部分采用UNet解码器。

由于超声图像存在大量噪声且边界不清晰，本研究将UNet网络的最大通道数由1 024降为256，在降低运算量的同时减少了过拟合风险。在冻结的SAM图像编码器中加入适配器(Adapter)模块，称为

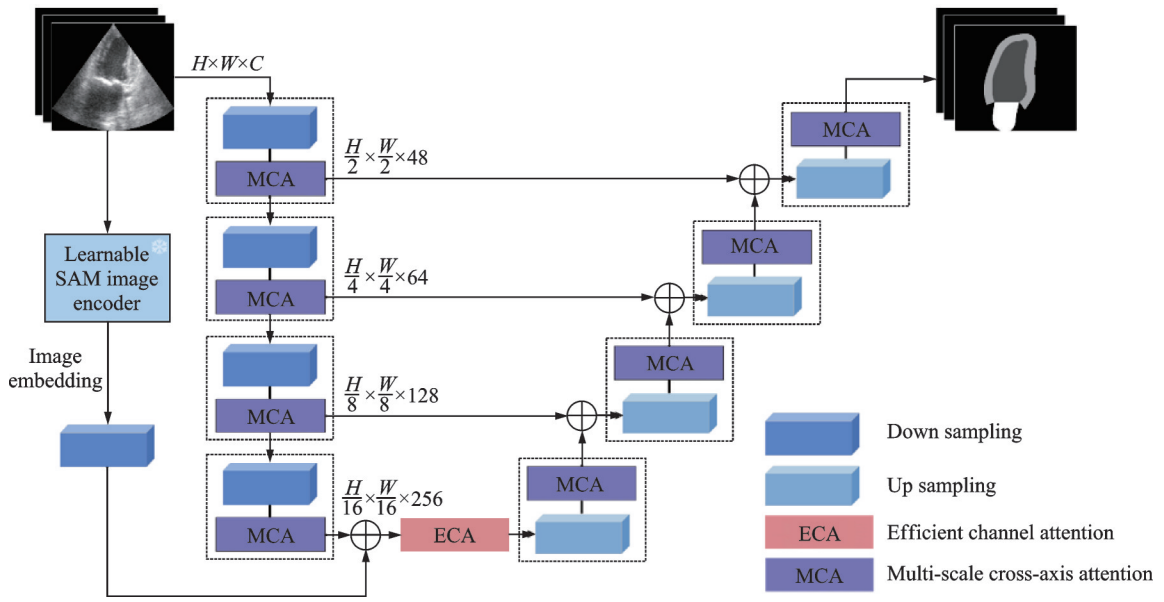

图1 MSC-LSAM的整体架构

Fig.1 Overall architecture of MSC-LSAM

可学习的分割一切模型(Learnable SAM,LSAM),使得SAM视觉编码器在保持较低参数量的同时,保留学习能力。在UNet全局网络引入多尺度交叉注意力(Multi-scale cross-axial attention,MCA),实现多尺度特征的交叉融合,有效提升边缘分割能力,抑制过拟合。在并行编码器特征拼接之后,加入高效通道注意力(Efficient channel attention,ECA)特征融合块,实现并行编码器多尺度特征的高效融合。

2.2 可学习的分割一切模型

SAM的图像编码器具有高度的泛化性能,在零样本(Zero-shot)任务中表现出色,因此本研究使用SAM视觉编码器来增强模型的特征提取能力。为了使SAM视觉编码器输入大小与UNet一致,本研究将SAM视觉编码器输入张量的$H\times W$由1 024×1 024调整为256×256。为了平衡推理速度与分割精度,本研究冻结了SAM视觉编码的Transformer层,并使用轻量级的SAM权重Mobile SAM,使得SAM视觉编码加入到本研究所提出的架构中几乎不会增加运算量。

为了使SAM视觉编码器更好地适应医学图像任务,并保留一些学习能力,本研究在每个Transformer层之间引入简单而高效的Adapter块(图2)。这种设计允许模型对特定任务进行微调,同时保留预训练模型中学习到的通用特征。这种自适应有助于模型更好地适应医学图像分割任务,防止模型出现灾难性遗忘,使模型在医学图像分割任务上更具鲁棒性。

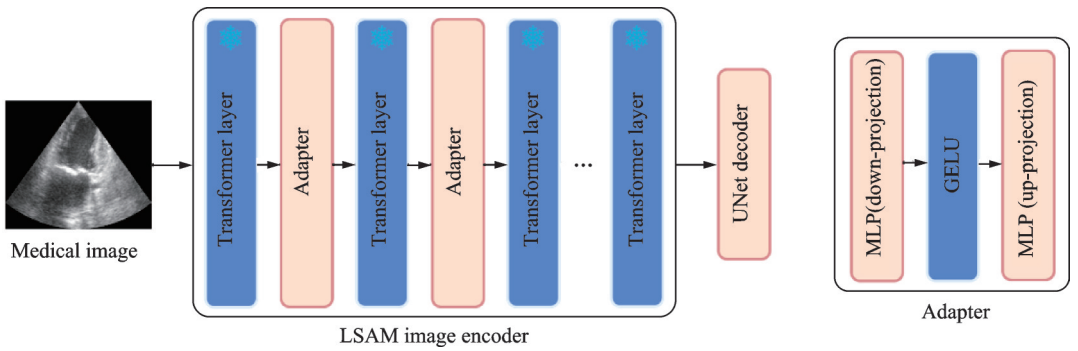

图2 LSAM图像编码器

Fig.2 LSAM image encoder

本研究将Adapter设计为瓶颈模型,包括向下投影的MLP层、GELU激活函数和向上投影的MLP层,可表达为

$$P_i=\mathrm{MLP}_{\mathrm{up}}\left(\mathrm{GELU}\left(\mathrm{MLP}_{\mathrm{down}}\left(F_i\right)\right)\right) \tag{1}$$

式中:F_i表示每个Transformer层的输出,$\mathrm{MLP}_{\mathrm{down}}$负责向下投影输出,将输入数据压缩到更低的维度。引入GELU函数作为激活函数,增加了非线性,帮助模型学习更复杂的特征表示。随后,$\mathrm{MLP}_{\mathrm{up}}$层将数据从低维度恢复到原始形状,实现适配器的上投影。这种瓶颈结构设计使适配器具有更多的灵活性,允许模型在保持计算效率的同时学习特定任务的适应性。

2.3 多尺度交叉注意力

MCA是一种新颖的方法,它将多尺度特征嵌入到轴向注意力中,旨在更好地分割具有不同个体大小和形状的区域[37]。

轴向注意力在大规模数据集上捕获全局上下文效率高,但在小数据集上容易学习位置偏差。MCA通过引入双交叉注意力,更好地利用轴向注意力提取的方向信息,从而更适应小数据集的特点,提高了信息的利用效率[21]。MCA的总体架构如图3所示。

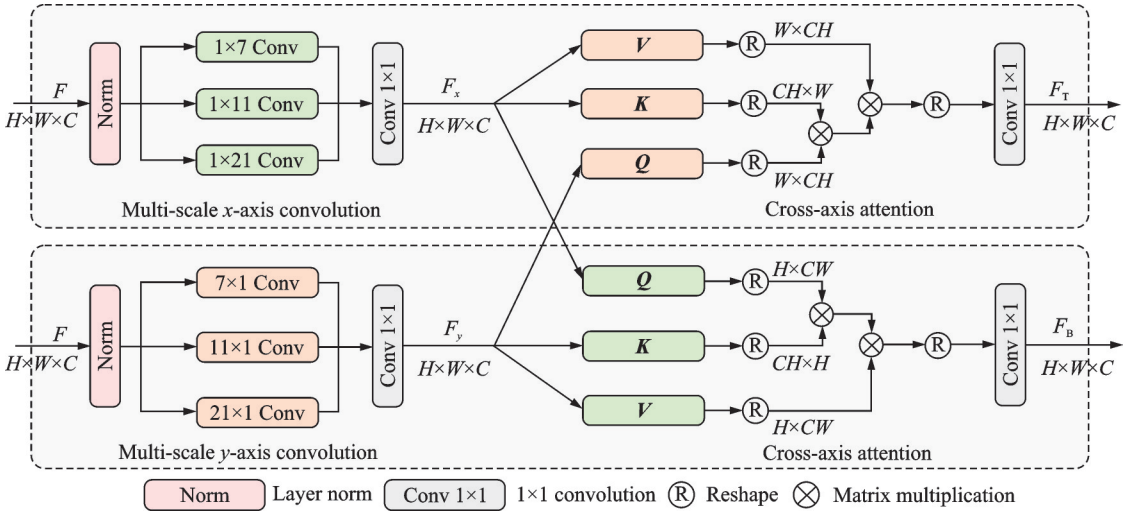

图3 多尺度交叉轴注意力的整体架构

Fig.3 Overall architecture of MCA

MCA分为两个并行分支,分别计算水平和垂直轴向注意力。每个分支,首先经过3个不同内核尺寸(分别为1×7、1×11和1×21)的1D卷积,沿一个空间维度编码多尺度上下文信息,公式为

$$F_x = \mathrm{Conv}_{1\times1}\left(\sum_{i=0}^{2}\mathrm{Conv1D}_i^x(\mathrm{Norm}(F))\right) \tag{2}$$

$$F_y = \mathrm{Conv}_{1\times1}\left(\sum_{i=0}^{2}\mathrm{Conv1D}_i^y(\mathrm{Norm}(F))\right) \tag{3}$$

式中:$\mathrm{Conv1D}_i^x(\cdot)$和$\mathrm{Conv1D}_i^y(\cdot)$分别表示沿$x$轴和$y$轴的一维卷积;$\mathrm{Norm}(\cdot)$表示层归一化;$F_x$和$F_y$分别表示$x$轴和$y$轴的输出。

接着是交叉注意力,输入数据被转换为3个不同的向量集:Key($\boldsymbol{K}$)、Query($\boldsymbol{Q}$)和Value($\boldsymbol{V}$)。但与轴向注意力不同的是,其中一个分支的$\boldsymbol{Q}$矩阵传给另一个分支,从而在两个并行轴之间建立相互作用,更有效地利用多尺度特征和全局信息,即

$$F_{\mathrm{T}} = \mathrm{MHCA}_y(F_{yQ}, F_{xV}, F_{xK}) \tag{4}$$

$$F_{\mathrm{B}} = \mathrm{MHCA}_x(F_{xQ}, F_{yV}, F_{yK}) \tag{5}$$

式中:F_{xK}和F_{xV}分别表示x轴的键矩阵和值矩阵,F_{yQ}表示y轴的查询矩阵;同理F_{yV}和F_{yK}分别表示y轴的键矩阵和值矩阵,F_{xQ}表示x轴的查询矩阵;$\mathrm{MHCA}_y(\cdot,\cdot,\cdot)$表示沿$x$轴的多头交叉注意力,$\mathrm{MHCA}_x(\cdot,\cdot,\cdot)$表示沿$y$轴的多头交叉注意力;$F_{\mathrm{T}}$和$F_{\mathrm{B}}$分别为MCA水平和垂直方向的输出结果。

多尺度交叉轴注意力的输出可以表示为

$$F_{\mathrm{out}} = \mathrm{Conv}_{1\times1}(F_{\mathrm{T}}) + \mathrm{Conv}_{1\times1}(F_{\mathrm{B}}) + F \tag{6}$$

2.4 高效通道注意力

通过多编码器结构能够获取到各模态丰富的语义信息,但是由于不同模态对分割结果的贡献不同,因此如何更好地合并这些特征信息值得思考。

ECA模块是一种高效的特征融合模块,能够在不降低通道维数的前提下,有效地实现跨通道的信息交互(图4)。该模块首先通过全局平均池化(Global average pooling, GAP)操作获取输入特征图的每

个通道的全局空间信息,生成表示全局信息的向量;然后使用一维卷积核来实现通道之间的动态关联,通过调整卷积核的大小,ECA能够灵活地捕获不同范围的通道依赖关系。这种方法使得每个通道不仅考虑自身的信息,还结合了其邻近通道的信息,增强了特征的表达能力[38]。

本研究引入ECA模块以适应不同尺度的特征表达,有效地整合来自不同模态和不同尺度的信息,降低无用特征的通道权重。

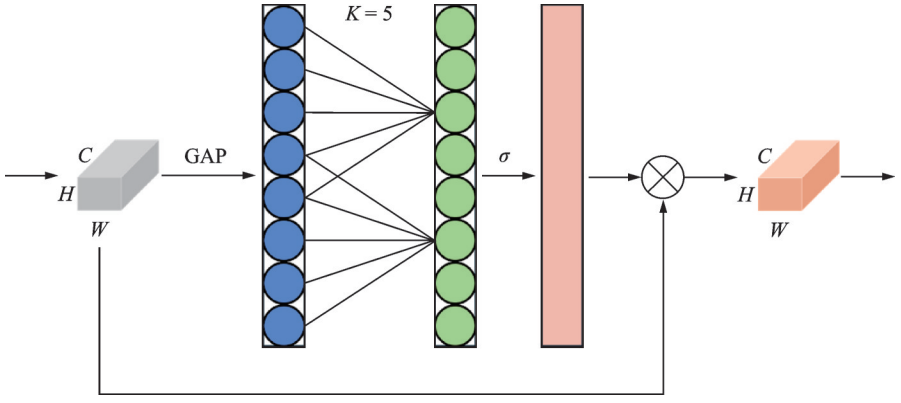

图4 高效通道注意力的整体架构

Fig.4 Overall architecture of efficient channel attention

3 实验结果

3.1 数据集

本研究在公开心脏超声数据集CAMUS和自建颈动脉超声数据集CAUS上对所提模型进行评估。

CAMUS数据集包含二腔切面(2CH)和四腔切面(4CH)各500个2D超声序列,以及每个超声序列在心脏舒张末期(ED)和心脏收缩末期(ES)帧手动注释的标签,其中900个用于训练,100个用于测试。手工标注的心脏结构包含左心室心内膜、心肌和左心房心内膜[39]。

CAUS是来自南通大学附属医院的私有数据集,包含来自于91名患者的2 439个2D颈动脉超声图像,其中1 952张用于训练,487张用于测试。该数据集由5位具有5年以上临床经验的超声科医生标注,手工标注的颈动脉结构包含颈内动脉和颈外动脉。

此项研究在南通大学伦理委员会的批准下进行,涉及对获取的非侵入性数据的智能评估和研究,包括人体医学影像和血管成像。

3.2 评价指标

本研究采用DSC和95%豪斯多夫距离(95% Hausdorff distance,HD95)来评估模型的分割精确度。

DSC通过计算分割结果和分割标签之间的重叠度来度量模型的性能[40],其计算公式为

$$\mathrm{DSC}(T,P)=2\times\frac{|T\cap P|}{|T|\cup|P|}=2\times\frac{T\times P}{T^2+P^2} \tag{7}$$

式中T和P分别代表所有体素的真实值和预测值。

HD95通过计算分割结果与分割标签95%的像素之间的最大距离来度量模型的性能[41],其计算公式为

$$d_{\mathrm{H}}(T',P')=\max(d_{T'P'},d_{P'T'}) \tag{8}$$

式中:$d_{T'P'}$表示真实体素集合T'和预测体素集合P'之间的最大豪斯多夫距离,$d_{P'T'}$表示从预测体素集合

P'到真实体素集合T'之间的最大豪斯多夫距离。

此外，本研究采用准确率(Accuracy，ACC)、精确率(Precision，PRE)、灵敏度(Sensitivity，SE)和F_1分数(F_1-score)来评估模型的异常情况处理能力[42]，其计算公式分别为

$$\mathrm{ACC}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \tag{9}$$

$$\mathrm{SE}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{10}$$

$$\mathrm{PRE}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{11}$$

$$F_1\text{-score}=\frac{2\times\mathrm{PRE}\times\mathrm{SE}}{\mathrm{PRE}+\mathrm{SE}} \tag{12}$$

式中：TP表示被正确分割为目标的像素数；FP表示被错误分割为目标的背景像素数；TN表示被正确分割为背景的像素数；FN表示被错误分割为背景的目标像素数。

3.3 对比实验

MSC-LSAM与各种主流医学分割算法性能的对比实验结果及可视化对比结果分别如表1～3和图5～7所示，其中在表1～3中，R50表示ResNet50，R101表示ResNet101；在图5～7中，GT表示参考标准，DeepLabv3＋表示DeepLabv3＋(ResNet50)。

表1 在CAMUS数据集2CH上的对比实验结果

Table 1 Comparative experiment results on 2CH of CAMUS dataset

算法	DSC	HD95	ACC	PRE	SE	F_1-score
DeepLabv3(R50)	0.923	5.926	0.955	0.868	0.908	0.887
DeepLabv3＋(R50)	0.902	8.796	0.955	0.837	0.887	0.861
DeepLabv3＋(R101)	0.891	8.606	0.956	0.806	0.867	0.835
UNet(Baseline)	0.917	6.696	0.954	0.858	0.892	0.874
UNet＋＋	0.915	7.044	0.955	0.834	0.898	0.865
Trans-Unet	0.919	6.586	0.965	0.849	0.902	0.875
Swin-Unet	0.922	13.997	0.985	0.872	0.911	0.891
nn-UNet	**0.927**	3.952	**0.986**	0.876	0.911	0.893
U-Mamba	0.925	5.879	0.981	**0.888**	**0.917**	**0.902**
MSC-LSAM(Ours)	**0.927**	**3.741**	0.983	0.887	0.911	0.899

表2 在CAMUS数据集4CH上的对比实验结果

Table 2 Comparative experiment results on 4CH of CAMUS dataset

算法	DSC	HD95	ACC	PRE	SE	F_1-score
DeepLabv3(R50)	0.931	5.392	0.996	**0.969**	0.970	**0.970**
DeepLabv3＋(R50)	0.916	7.126	0.996	0.932	0.972	0.952
DeepLabv3＋(R101)	0.905	8.170	0.996	0.925	0.974	0.949
UNet(Baseline)	0.927	5.918	0.995	0.961	0.965	0.963
UNet＋＋	0.924	6.284	0.996	0.951	0.974	0.962
Trans-Unet	0.928	5.852	0.995	0.959	**0.979**	0.969
Swin-Unet	0.913	14.256	0.994	0.934	0.973	0.953
nn-UNet	0.932	**4.951**	**0.997**	0.951	0.973	0.962
U-Mamba	**0.934**	6.021	0.995	0.965	0.974	0.969
MSC-LSAM(Ours)	**0.934**	5.081	0.996	0.967	0.971	0.969

表3 在CAUS数据集上的对比实验结果

Table 3 Comparative experiment results on CAUS dataset

算法	DSC	HD95	ACC	PRE	SE	F_1-score
DeepLabv3(R50)	0.897	3.719	0.997	0.892	**0.949**	0.920
DeepLabv3+(R50)	0.879	5.021	0.997	0.858	0.939	0.897
DeepLabv3+(R101)	0.870	5.448	0.996	0.823	0.930	0.873
UNet(Baseline)	0.892	4.285	0.996	0.879	0.931	0.898
UNet++	0.874	4.070	0.996	0.840	0.934	0.885
Trans-Unet	0.905	5.365	**0.998**	0.901	0.929	0.915
Swin-Unet	0.885	9.007	0.996	0.860	0.933	0.895
nn-UNet	**0.918**	3.745	0.996	**0.918**	0.929	0.923
U-Mamba	0.916	5.945	0.996	0.915	0.936	**0.925**
MSC-LSAM(Ours)	0.917	**2.246**	0.997	0.914	0.933	0.920

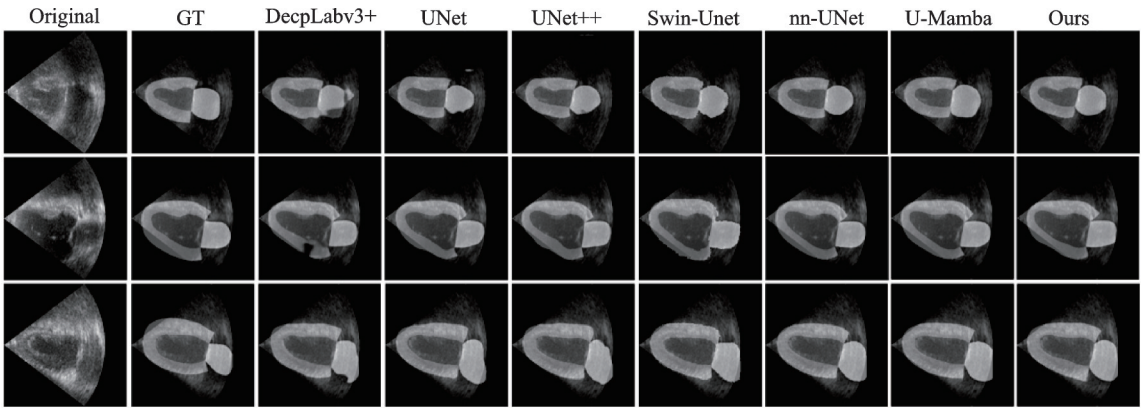

图5 MSC-LSAM与其他算法在CAMUS数据集2CH上的可视化比较

Fig.5 Visualization comparison of MSC-LSAM and other algorithms on 2CH of CAMUS dataset

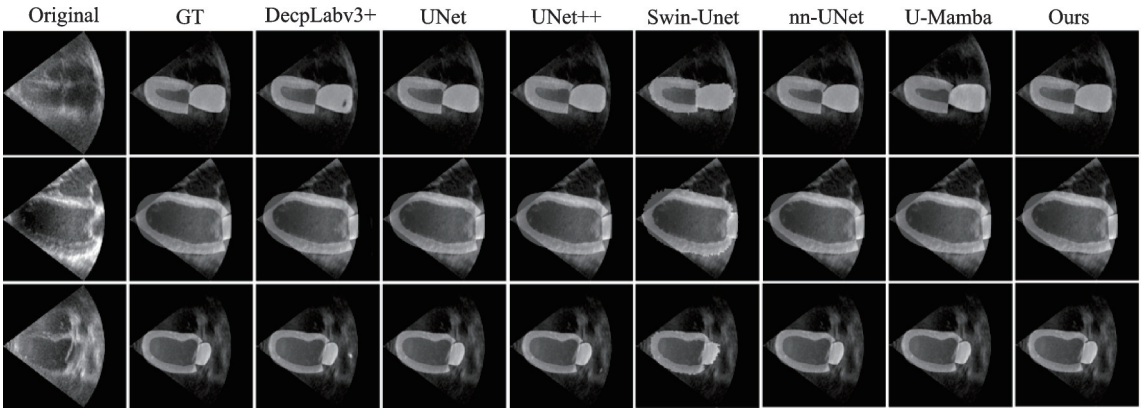

图6 MSC-LSAM与其他算法在CAMUS数据集4CH上的可视化比较

Fig.6 Visualization comparison of MSC-LSAM and other algorithms on 4CH of CAMUS dataset

对比结果表明,MSC-LSAM在超声医学图像分割任务中,DSC、HD95、ACC、PRE、SE和F_1-score等指标均取得了较高的水平,表现出优秀的分割性能和异常情况处理能力。

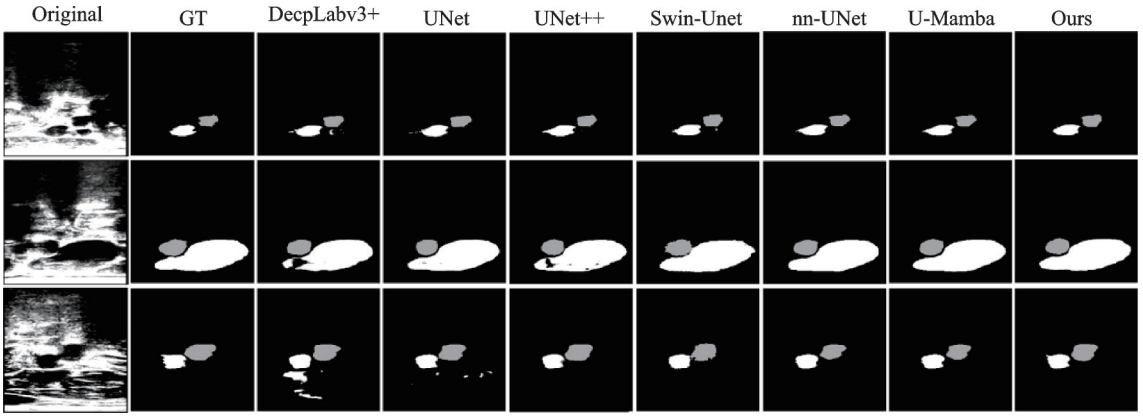

图7 MSC-LSAM与其他算法在CAUS数据集的可视化比较

Fig.7 Visualization comparison of MSC-LSAM and other algorithms on CAUS dataset

DeepLabv3＋、UNet和UNet＋＋作为传统的卷积神经网络架构，虽然在许多任务中表现出色，但在面对心脏腔体和颈动脉管腔这类与背景阈值接近的区域时，特征提取能力受到限制，进而导致了不同程度的误分割。这种现象尤其体现在这些模型对复杂解剖结构边界的识别上，泛化能力的不足使得它们在应对不同样本时易发生边界不清晰或误识别的情况。相比之下，Swin-Unet和Trans-Unet引入了Transformer架构，打破了传统卷积神经网络在局部特征提取方面的局限性，其自注意力机制可以更有效地捕捉全局信息，使得特征提取能力和泛化能力得到了显著提升。然而，由于超声图像中边缘区域通常伴随大量噪声，Swin-Unet在这些区域出现了过拟合的问题，没有准确分割出实际的结构边界。这表明虽然Transformer块在全局信息提取方面具有优势，但在面对噪声较多的超声图像时，其边缘检测能力可能受到影响。

nn-UNet和U-Mamba是目前公认的两个先进的模型。nn-UNet结合了自动化的超参数调整策略，能够根据不同的任务自适应优化模型结构和参数，并通过数据增强技术生成多样性更高的训练样本，以提高模型的泛化能力。而U-Mamba在nn-UNet的基础上引入了先进的Mamba架构，在保持局部空间信息的同时，充分利用全局特征，提升了特征提取的深度和局部特征提取能力。这两个模型在超声图像分割中都显示出了较好的鲁棒性，尤其在边界模糊和噪声干扰较大的区域，误分割率明显低于其他模型。

相较于nn-UNet和U-Mamba，MSC-LSAM在保持更好的鲁棒性和更低的误分割率的同时，对于噪声的抗干扰能力更强，在细节捕捉和边缘处理方面表现出更强适应性，能够准确分割出更小的结构和细节。这使得MSC-LSAM在需要对噪声环境中的微细结构进行高精度分割的应用场景中更具优势，能够更好地应对高噪声和复杂背景下的分割任务。

3.4 消融实验

本研究对算法网络进行了全面的消融实验，使用LSAM、ECA和MCA作为变量，通过模块的逐步增加来分析每个模块对于神经网络性能和分割效果的提升。消融实验结果及其可视化结果分别如表4～6和图8～10所示。其中在表4～5中，Label 1表示左心室心内膜，Label 2表示心肌，Label 3表示左心房心内膜；在表6中，Label 1表示颈内动脉，Label 2表示颈外动脉；在图8～10中，GT表示参考标准。

在本研究提出的MSC-LSAM中，LSAM的引入显著增强了编码器部分的特征提取能力，使模型聚焦于关键特征，提升了模型对复杂数据模式的理解和捕捉能力。然而，由于LSAM生成的特征与UNet

表4 在CAMUS数据集2CH上的消融实验结果

Table 4 Ablation experiment results on 2CH of CAMUS dataset

LSAM	ECA	MCA	DSC				HD95			
			Label 1	Label 2	Label 3	Average	Label 1	Label 2	Label 3	Average
			0.913	0.908	0.930	0.917	7.560	4.443	8.084	6.696
√			0.920	0.911	0.927	0.919	5.232	3.656	4.200	4.362
√	√		0.920	0.911	0.928	0.920	3.883	3.637	4.444	3.988
√	√	√	0.923	0.919	0.937	0.927	3.663	3.371	4.190	3.741
LSAM	ECA	MCA	ACC				PRE			
			Label 1	Label 2	Label 3	Average	Label1	Label 2	Label 3	Average
			0.945	0.923	0.992	0.954	0.844	0.883	0.847	0.858
√			0.981	0.975	0.992	0.982	0.856	0.880	0.847	0.861
√	√		0.981	0.974	0.992	0.982	0.858	0.882	0.855	0.865
√	√	√	0.980	0.976	0.992	0.983	0.900	0.873	0.889	0.887
LSAM	ECA	MCA	SE				F_1-score			
			Label 1	Label 2	Label 3	Average	Label 1	Label 2	Label 3	Average
			0.896	0.864	0.914	0.892	0.869	0.874	0.879	0.874
√			0.911	0.860	0.944	0.905	0.882	0.870	0.893	0.882
√	√		0.908	0.863	0.943	0.905	0.882	0.872	0.897	0.884
√	√	√	0.909	0.873	0.950	0.911	0.904	0.873	0.919	0.899

表5 在CAMUS数据集4CH上的消融实验结果

Table 5 Ablation experiment results on 4CH of CAMUS dataset

LSAM	ECA	MCA	DSC				HD95			
			Label 1	Label 2	Label 3	Average	Label 1	Label 2	Label 3	Average
			0.915	0.914	0.951	0.927	6.652	3.927	7.176	5.918
√			0.915	0.912	0.951	0.926	6.647	3.923	6.915	5.828
√	√		0.914	0.926	0.953	0.931	5.630	3.587	6.381	5.199
√	√	√	0.919	0.929	0.954	0.934	5.692	3.253	6.297	5.081
LSAM	ECA	MCA	ACC				PRE			
			Label 1	Label 2	Label 3	Average	Label 1	Label 2	Label 3	Average
			0.992	0.993	0.999	0.995	0.925	0.968	0.991	0.961
√			0.992	0.993	0.999	0.995	0.932	0.962	0.992	0.962
√	√		0.993	0.995	0.999	0.996	0.949	0.962	0.990	0.967
√	√	√	0.993	0.994	0.999	0.996	0.943	0.969	0.989	0.967
LSAM	ECA	MCA	SE				F_1-score			
			Label 1	Label 2	Label 3	Average	Label 1	Label 2	Label 3	Average
			0.966	0.948	0.981	0.965	0.945	0.958	0.986	0.963
√			0.959	0.950	0.981	0.963	0.944	0.956	0.986	0.962
√	√		0.958	0.969	0.990	0.973	0.953	0.965	0.990	0.969
√	√	√	0.965	0.960	0.988	0.971	0.953	0.965	0.989	0.969

表6 在CAUS数据集上的消融实验结果

Table 6 Ablation experiment results on CAUS dataset

LSAM	ECA	MCA	DSC			HD95		
			Label 1	Label 2	Average	Label 1	Label 2	Average
			0.859	0.924	0.892	4.883	4.213	4.548
√			0.863	0.919	0.891	5.362	5.581	5.472
√	√		0.895	0.936	0.915	4.070	3.290	3.680
√	√	√	0.900	0.934	0.917	2.559	1.934	2.246
LSAM	ECA	MCA	ACC			PRE		
			Label 1	Label 2	Average	Label 1	Label 2	Average
			0.995	0.997	0.996	0.810	0.949	0.879
√			0.995	0.997	0.996	0.805	0.948	0.877
√	√		0.996	0.997	0.997	0.863	0.958	0.911
√	√	√	0.997	0.997	0.997	0.884	0.945	0.914
LSAM	ECA	MCA	SE			F_1-score		
			Label 1	Label 2	Average	Label 1	Label 2	Average
			0.942	0.921	0.931	0.867	0.929	0.898
√			0.941	0.902	0.921	0.863	0.919	0.891
√	√		0.946	0.919	0.933	0.900	0.933	0.917
√	√	√	0.935	0.931	0.933	0.906	0.934	0.920

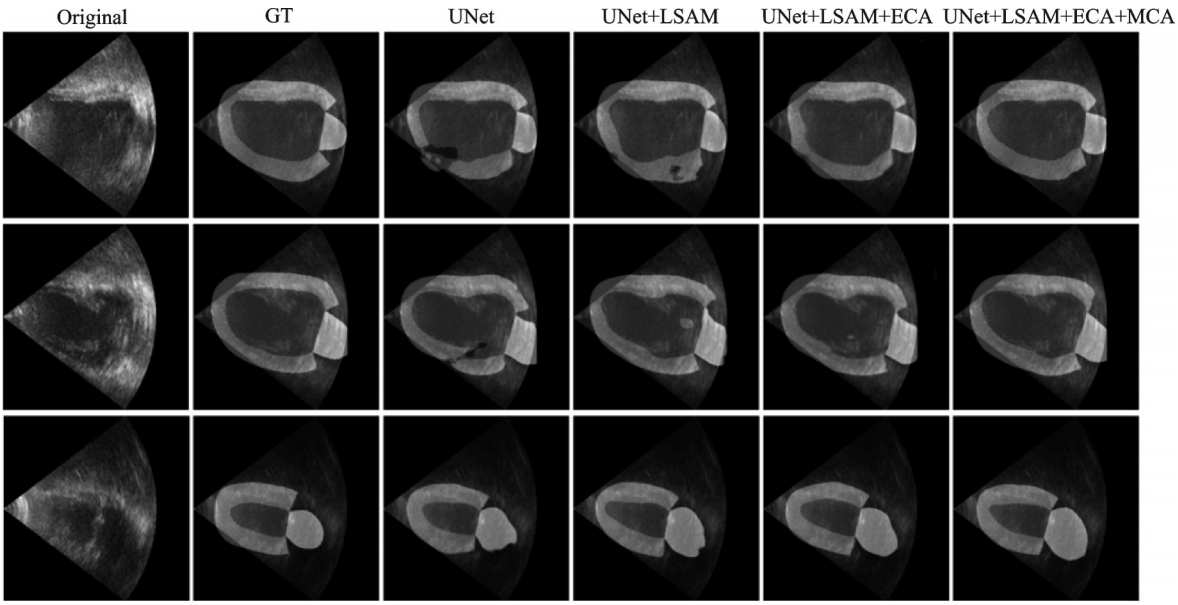

图8 在CAMUS数据集2CH上的消融实验可视化结果

Fig.8 Visualization results of ablation experiment on 2CH of CAMUS dataset

编码器输出的特征在尺度上存在不匹配问题,导致在特征融合过程中出现困难。这种不匹配引发了特征的低效融合,从而导致误分割率增加。

ECA的引入,实现了对双编码器特征的并行融合,增强了多尺度特征的互补性。ECA能够动态调整不同通道的权重,从而有效解决了LSAM与UNet编码器特征融合中的尺度不匹配问题,大幅减少了误分割的发生。

MCA的引入,进一步提升了模型对多尺度特征和全局信息的利用能力。MCA能够从多个轴提取

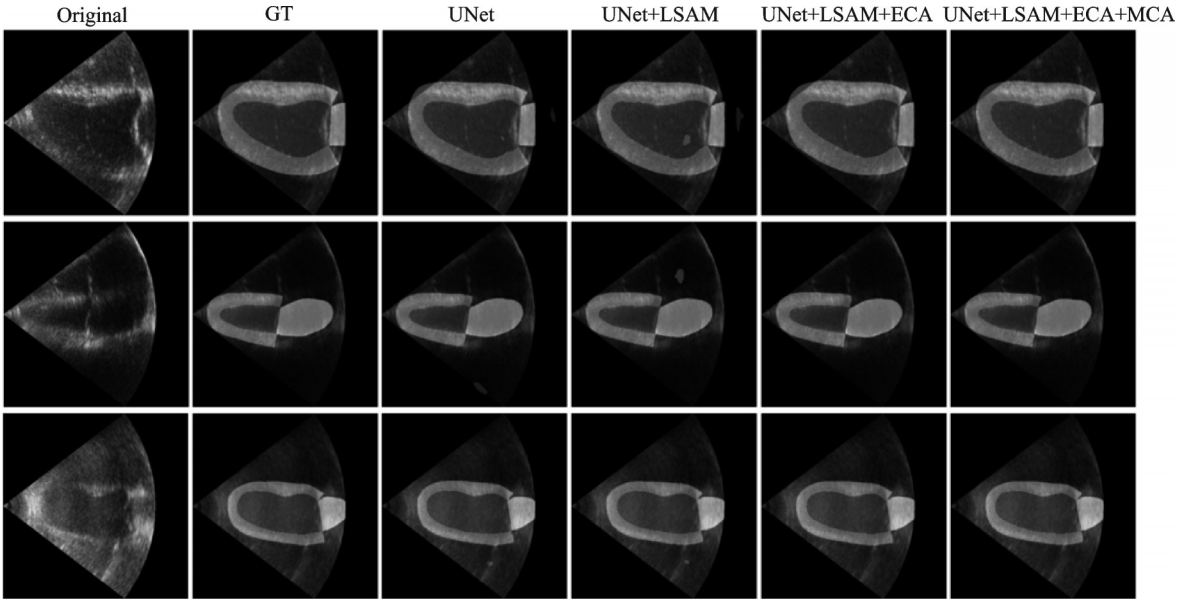

图9 在CAMUS数据集4CH上的消融实验可视化结果

Fig.9 Visualization results of ablation experiment on 4CH of CAMUS dataset

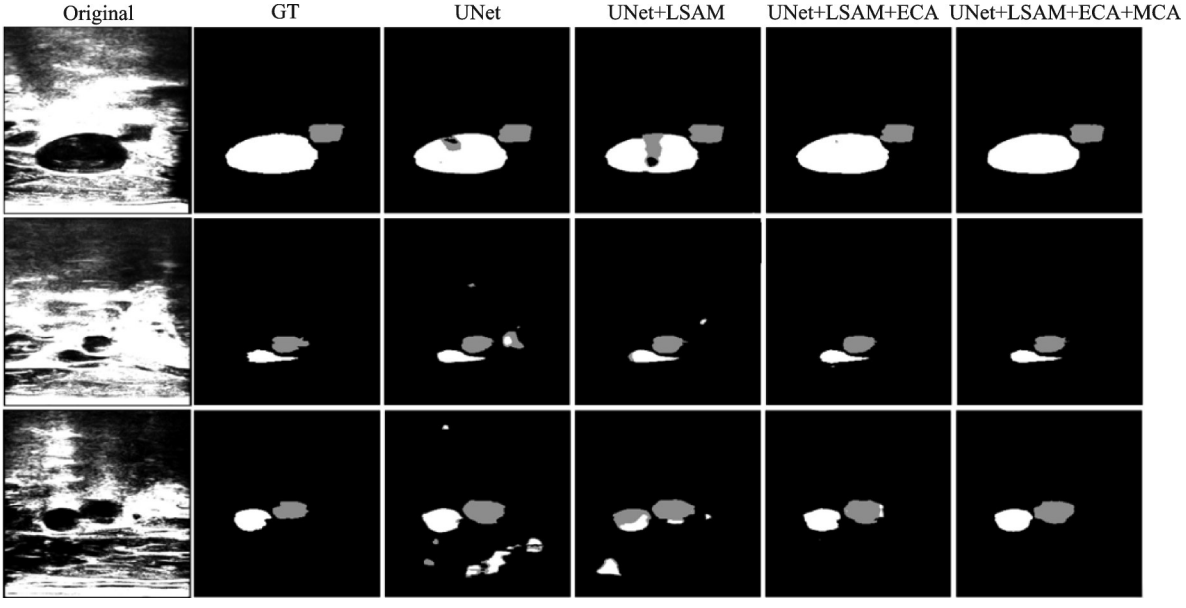

图10 在CAUS数据集上的消融实验可视化结果

Fig.10 Visualization results of ablation experiment on CAUS dataset

上下文信息,并聚焦于特定的区域或边缘特征,使得模型在处理复杂的边缘分割任务时表现更为优异。实验结果表明,MCA的加入不仅改善了边缘分割效果,还进一步减少了误分割,显著提升了模型的鲁棒性和泛化能力。

4 讨论

颈动脉超声和心脏超声分割可以辅助医生定量地评估患者心脏和颈动脉状况,为评估缺血性脑卒

中风险提供重要参考,但超声图像的分割是一项复杂的工作[7-9]。相较于CT和MRI,超声图像对比度低、噪声多、图像模糊、伪影多,超声医学图像分割更具挑战性[10]。传统深度学习算法在进行心脏腔体和颈动脉管腔超声图像分割时,容易出现误分割,且边缘分割效果较差,因此本研究提出了MSC-LSAM来解决这些难题[12]。

MSC-LSAM在编码器部分并行了SAM编码器和UNet编码器,在解码器部分采用UNet解码器。在参数冻结的预训练的SAM图像编码器中加入Adapter模块,使得SAM图像编码器在拥有较低参数量的同时,保留学习能力。此外,MSC-LSA运用了两组多尺度特征。在UNet全局网络引入MCA,实现了不同轴之间多尺度特征的交叉融合,有效提升边缘分割能力,抑制过拟合。在双编码器特征拼接之后,加入了ECA,实现双编码器多尺度特征的融合。

MSC-LSAM在心脏超声数据集CAMUS和颈动脉超声数据集CAUS上均取得了较好的分割效果。在对比实验部分,MSC-LSAM优于主流分割模型。在消融实验部分,通过LSAM、ECA和MCA这3个模块的逐个加入,确定每个模块对消融效果的贡献和相互作用。实验结果表明,LSAM可以显著提升模型的特征提取能力,提升分割效果;ECA模块可以提升并行编码器的特征融合能力,显著减少误分割[38];MCA使得模型的全局信息提取能力得到了提升,边缘分割效果显著提升,误分割进一步减少,算法的鲁棒性和泛化能力得到了提升[37]。

本研究的创新点在于:(1) 提出了一个颈动脉超声和心脏超声图像的自动分割网络MSC-LSAM,能够辅助医生评估缺血性脑卒中风险;(2) 并行了SAM视觉编码器和UNet编码器,在冻结的预训练的SAM视觉编码器中加入高效的Adapter块,使SAM编码器在保持较低参数量的同时,拥有一定学习能力;(3) 结合两组多尺度,通过ECA块实现并行编码器多尺度特征的高效融合,提升了特征提取能力。在全局网络引入MCA块,实现不同轴之间多尺度特征的交叉融合,提升模型全局信息提取能力。

本研究尚存在以下不足:(1) 缺血性脑卒中评估、预防是一个多因素的复杂过程,有待于与其他风险指标如脑白质高信号相结合;(2) 提出的算法仅在两个数据集上进行实验,模型的泛化能力有待进一步验证。但本文使用的两个数据集共包含4 439张超声图像,具有较好的代表性和数据规模,能够在一定程度上缓解这个问题;(3) 多尺度交叉的双编码器结构增加了模型的复杂度和计算成本,未来的研究可以通过引入轻量级网络或采用模型剪枝技术,来减少计算资源的消耗,在保持或提高分割精度的同时,优化模型的效率和实用性。

总的来说,MSC-LSAM算法在超声医学图像分割中具有显著的应用前景,尤其是在提高心脏和颈动脉超声分割的准确性方面。然而,其在实际临床应用中MSC-LSAM面临计算复杂度、可解释性以及泛化能力等方面的限制和挑战,仍需进一步的研究和优化。未来通过提升模型的计算效率、增强跨设备的适应性,MSC-LSAM有望在未来成为超声医学图像分割的主流工具,为临床诊断提供更强大的技术支持。

5 结束语

本文提出一种颈动脉超声和心脏超声图像的自动分割方法MSC-LSAM,用于实现心脏腔体和颈动脉管腔的精准分割,辅助医生定量地评估患者心脏和颈动脉状况,辅助实现缺血性脑卒中的早期诊断,具有较好的准确度及应用前景。

参考文献:

[1] TU W J, WANG L D. China stroke surveillance report 2021[J]. Military Medical Research, 2023, 10(1): 33.

[2] KLEINDORFER D O, TOWFIGHI A, CHATURVEDI S, et al. 2021 guideline for the prevention of stroke in patients with stroke and transient ischemic attack: A guideline from the American Heart Association/American Stroke Association[J]. Stroke, 2021, 52(7): e364-e467.

[3] FURE B, WYLLER T B, THOMMESSEN B. TOAST criteria applied in acute ischemic stroke[J]. Acta Neurologica

Scandinavica, 2005, 112(4): 254-258.
[4] LIAO Y, GUAN M, LIANG D, et al. Differences in pathological composition among large artery occlusion cerebral thrombi, valvular heart disease atrial thrombi and carotid endarterectomy plaques[J]. Frontiers in Neurology, 2020, 11: 811.
[5] FLAHERTY M L, KISSELA B, KHOURY J C, et al. Carotid artery stenosis as a cause of stroke[J]. Neuroepidemiology, 2012, 40(1): 36-41.
[6] NEPAL R, CHOUDHARY M K, DHUNGANA S, et al. Prevalence and major cardiac causes of cardio-embolic stroke and in-hospital mortality in Eastern Nepal[J]. Journal of Clinical and Preventive Cardiology, 2020, 9(1): 19-24.
[7] WANG Y, YAO Y. Application of artificial intelligence methods in carotid artery segmentation: A review[J]. IEEE Access, 2023, 11: 13846-13858.
[8] YANG S Y, HUANG M, WANG A L, et al. Atrial fibrillation burden and the risk of stroke: A systematic review and dose-response meta-analysis[J]. World Journal of Clinical Cases, 2022, 10(3): 939.
[9] BOTS M L, HOES A W, KOUDSTAAL P J, et al. Common carotid intima-media thickness and risk of stroke and myocardial infarction: The Rotterdam study[J]. Circulation, 1997, 96(5): 1432-1437.
[10] AZZOPARDI C, CAMILLERI K P, HICKS Y A. Bimodal automated carotid ultrasound segmentation using geometrically constrained deep neural networks[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(4): 1004-1015.
[11] SIDDIQUE N, PAHEDING S, ELKIN C P, et al. U-Net and its variants for medical image segmentation: A review of theory and applications[J]. IEEE Access, 2021, 9: 82031-82057.
[12] GU P, LEE W M, ROUBIDOUX M A, et al. Automated 3D ultrasound image segmentation to aid breast cancer image interpretation[J]. Ultrasonics, 2016, 65: 51-58.
[13] MAHMOUD A, MORSY A, DE GROOT E. A new gradient-based algorithm for edge detection in ultrasonic carotid artery images[C]//Proceedings of 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. [S.l.]: IEEE, 2010: 5165-5168.
[14] XU X, ZHOU Y, CHENG X, et al. Ultrasound intima-media segmentation using Hough transform and dual snake model[J]. Computerized Medical Imaging and Graphics, 2012, 36(3): 248-258.
[15] MANNIESING R, SCHAAP M, ROZIE S, et al. Robust CTA lumen segmentation of the atherosclerotic carotid artery bifurcation in a large patient population[J]. Medical Image Analysis, 2010, 14(6): 759-769.
[16] NAGARAJ Y, TEJA A H S, NARASIMHADHAN A V. Automatic segmentation of intima media complex in carotid ultrasound images using support vector machine[J]. Arabian Journal for Science and Engineering, 2019, 44(4): 3489-3496.
[17] XIAO X, LIAN S, LUO Z, et al. Weighted res-UNet for high-quality retina vessel segmentation[C]//Proceedings of the 9th International Conference on Information Technology in Medicine and Education (ITME). [S.l.]: IEEE, 2018: 327-331.
[18] ZHOU Z, SIDDIQUEE M M R, TAJBAKHSH N, et al. UNet++: Redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 39(6): 1856-1867.
[19] HAN K, XIAO A, WU E, et al. Transformer in transformer[J]. Advances in Neural Information Processing Systems, 2021, 34: 15908-15919.
[20] CAO H, WANG Y, CHEN J, et al. Swin-Unet: Unet-like pure transformer for medical image segmentation[C]//Proceedings of European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022: 205-218.
[21] MA L, ZHANG D, WANG Z, et al. Swin-PSAxialNet: An efficient multi-organ segmentation technique[J]. Journal of Visualized Experiments, 2024, 209: e66459.
[22] LIANG B, TANG C, ZHANG W, et al. N-Net: An UNet architecture with dual encoder for medical image segmentation[J]. Signal, Image and Video Processing, 2023, 17(6): 3073-3081.
[23] WANG H, CAO P, YANG J, et al. MCA-UNet: Multi-scale cross co-attentional U-Net for automatic medical image segmentation[J]. Health Information Science and Systems, 2023, 11(1): 10.
[24] KIRILLOV A, MINTUN E, RAVI N, et al. Segment anything[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2023: 4015-4026.
[25] ZHANG C, HAN D, QIAO Y, et al. Faster segment anything: Towards lightweight SAM for mobile applications[EB/OL]. (2023-06-25). https://doi.org/10.48550/arXiv.2306.14289.
[26] HE K, CHEN X, XIE S, et al. Masked autoencoders are scalable vision learners[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2022: 16000-16009.

[27] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[EB/OL]. (2020-10-05). https://doi.org/10.48550/arXiv.2010.11929.
[28] MA J, HE Y, LI F, et al. Segment anything in medical images[J]. Nature Communications, 2024, 15(1): 654.
[29] WU J, FU R, FANG H, et al. Medical SAM adapter: Adapting segment anything model for medical image segmentation[EB/OL]. (2023-04-25). https://doi.org./10.48550/arViv.2304.12620.
[30] LI Y, JING B, FENG X, et al. nnSAM: Plug-and-play segment anything model improves nnUNet performance[J]. [EB/OL]. (2024-01-10). https://doi.org./10.48550/arViv.2309.16967.
[31] MORADI S, OGHLI M G, ALIZADEHASL A, et al. MFP-Unet: A novel deep learning based approach for left ventricle segmentation in echocardiography[J]. Physica Medica, 2019, 67: 58-69.
[32] LECLERC S, SMISTAD E, ØSTVIK A, et al. LU-Net: A multistage attention network to improve the robustness of segmentation of left ventricular structures in 2-D echocardiography[J]. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 2020, 67(12): 2519-2530.
[33] LIU F, WANG K, LIU D, et al. Deep pyramid local attention neural network for cardiac structure segmentation in two-dimensional echocardiography[J]. Medical Image Analysis, 2021, 67: 101873.
[34] ZHOU R, GUO F, AZARPAZHOOH M R, et al. A voxel-based fully convolution network and continuous max-flow for carotid vessel-wall-volume segmentation from 3D ultrasound images[J]. IEEE Transactions on Medical Imaging, 2020, 39(9): 2844-2855.
[35] ZHANG B, WANG C, LI C. NvNet: An enhanced attention network for segmenting neck vascular from ultrasound images [C]//Proceedings of 2021 International Joint Conference on Neural Networks (IJCNN). [S.l.]: IEEE, 2021: 1-8.
[36] YUAN Y, LI C, XU L, et al. CSM-Net: Automatic joint segmentation of intima-media complex and lumen in carotid artery ultrasound images[J]. Computers in Biology and Medicine, 2022, 150: 106119.
[37] SHAO H, ZHANG Y, HOU Q. Polyper: Boundary sensitive polyp segmentation[C]//Proceedings of the 38th AAAI Conference on Artificial Intelligence and 36th Conference on Innovative Applications of Artificial Intelligence and 14th Symposium on Educational Advances in Artificial Intelligence. [S.l.]: AAAI, 2024: 4731-4739.
[38] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 11534-11542.
[39] LECLERC S, SMISTAD E, PEDROSA J, et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography[J]. IEEE Transactions on Medical Imaging, 2019, 38(9): 2198-2210.
[40] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[EB/OL]. (2019-11-19). https://doi.org./10.48550/arViv.1911.08287.
[41] HUTTENLOCHER D P, KLANDERMAN G A, RUCKLIDGE W J. Comparing images using the Hausdorff distance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(9): 850-863.
[42] SOKOLOVA M, JAPKOWICZ N, SZPAKOWICZ S. Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation[C]//Proceedings of Australasian Joint Conference on Artificial Intelligence. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006: 1015-1021.

作者简介:

王朝欣(2001-),男,硕士研究生,研究方向:人工神经网络和认知神经科学,E-mail:919874984@qq.com。

杨汶汶(1985-),男,教授,研究方向:人工智能和数字信号处理。

戎泽(2004-),男,本科生,研究方向:人工神经网络和医学图像处理。

李铮昱(2003-),男,本科生,研究方向:人工神经网络和医学图像处理。

王行(2002-),女,本科生,研究方向:人工神经网络和医学图像处理。

马磊(1983-),**通信作者**,男,副教授,研究方向:人工智能和生物信号采集与处理,E-mail:malei@ntu.edu.cn。

(编辑:王静)