

边缘增强结合多尺度信息融合的皮肤病变分割^①



齐向明, 张志伟

(辽宁工程技术大学 软件学院, 葫芦岛 125100)

通信作者: 张志伟, E-mail: 1420306290@qq.com

摘要: 针对皮肤病灶大小不一、与周围皮肤对比度低、边界模糊不规则、存在伪影及毛发干扰等问题, 本文提出一种边缘增强结合多尺度信息融合的皮肤病变分割算法, 该算法由编码器、多尺度感知模块、边缘增强模块、轻量化解码器组成. 首先在编码器中构建 Transformer 模块以提取全局信息, 使用卷积操作以提取局部信息; 其次设计多尺度感知模块, 采用密集连接结构的门控空洞卷积金字塔模块来融合多尺度特征; 接着构建边缘增强模块, 利用深层特征促进对边缘特征的探索以更好的保留细节和边缘信息; 最后设计轻量化解码器, 采用 CARAFE 轻量化算子进行上采样, 在使用相对较少参数的情况下保持较高的分割精度. 在公开数据集 ISIC2016 和 ISIC2018 上做对比实验, 结果表明, 本文算法分割精度高于其他流行算法.

关键词: 皮肤病变; 图像分割; 注意力机制; 边缘增强; 多尺度感知模块

引用格式: 齐向明, 张志伟. 边缘增强结合多尺度信息融合的皮肤病变分割. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/9676.html>

Skin Lesion Segmentation Based on Edge Enhancement Combined with Multi-scale Information Fusion

QI Xiang-Ming, ZHANG Zhi-Wei

(Software College, Liaoning Technical University, Huludao 125100, China)

Abstract: To address the problems of skin lesions, such as varied sizes, low contrast with surrounding skin, blurred and irregular boundaries, artifacts, and hair interference, this study proposes a skin lesion segmentation algorithm that combines edge enhancement with multi-scale information fusion. The algorithm consists of an encoder, a multi-scale sensing module, an edge enhancement module, and a lightweight decoder. Firstly, a Transformer module is built in the encoder to extract global information, and convolution operations are used to extract local information. Secondly, a multi-scale sensing module is designed to integrate multi-scale features using a gated atrous convolution pyramid module with a dense connection structure. An edge enhancement module is constructed, utilizing deep features to promote the exploration of edge features to better retain details and edge information. Finally, a lightweight decoder is designed, employing the CARAFE lightweight operator for upsampling, to maintain high segmentation accuracy with fewer parameters. Comparative experiments on open data sets ISIC2016 and ISIC2018 show that the segmentation accuracy of the proposed algorithm is higher than that of other popular algorithms.

Key words: skin lesion; image segmentation; attention mechanism; edge enhancement; multi-scale perception module

近些年来, 皮肤病变发病率一直升高^[1], 如果不能早期治疗可能会恶化为皮肤癌, 死亡率较高. 皮肤癌若

能够早期被发现, 拥有非常高的被治愈率. 人工检查的准确性由医生自身的水平决定, 人工检查是一项既耗

^① 基金项目: 国家自然科学基金面上项目 (62173171)

收稿时间: 2024-03-14; 修改时间: 2024-04-10, 2024-05-20; 采用时间: 2024-06-04; csa 在线出版时间: 2024-09-24

精力又耗时间的工作,为提高医生诊断效率和精确程度,计算机辅助诊断系统被研发出来,计算机辅助系统需要对病变位置进行定位,这一过程对系统的性能起到决定性的作用.因此,国内外大量科研工作者对此开展了很多研究工作,并提出了许多基于深度学习的皮肤病变的分割方法.

当前,皮肤病变分割算法主要分为两类,第1类是传统分割算法,主要包括3种:专家手动分割,耗时且易错;半自动分割,使用形态学方法进行分割,由于噪声点等因素干扰,可能导致图像被分得太细,即过度分割;自动分割,包括基于阈值^[2]、基于边缘^[3]、基于区域合并^[4]和基于图论^[5]分割算法等.基于阈值算法和基于边缘算法计算简单、运算效率较高,但对噪声敏感,图像中遮挡和人工标注容易影响分割精度,基于区域合并分割算法能够较好提取出病变区域特征,但容易造成过分割,基于图论分割算法处理背景不均匀的图像效果良好,但计算复杂度高,通用性有限.

第2类是基于深度学习的皮肤病变分割算法,由于在学习或感知病变图像特征时,传统方法很难将大量历史数据应用于学习过程,病理诊断存在局限性,而深度学习方法^[6-8]分割通过训练端到端框架,提取高级语义特征,性能高无需人工交互.卷积神经网络从原始图像中自动学习多种特征,取得开创性进展,鲁棒性优于传统分割算法.2015年,Long等人^[9]第1次提出全卷积神经网络(fully convolutional network, FCN),2017年,Yuan等人^[10]基于FCN实现端到端的皮肤病变像素级分割,省去传统算法预处理过程,但没充分利用全局上下文信息,无法结合浅层和深层特征,提取特征能力有限,不同皮肤条件下视觉复杂性和模糊性,对现有自动分割方法仍具有挑战性.

编码-解码结构的网络在医学图像分割中应用广泛,2015年,Ronneberger等人^[11]提出具有编解码结构的U-Net网络,采用跳跃连接融合高级和低级语义信息,U-Net对各种噪声和伪影过于敏感,从而影响其分割的准确性;2018年,Zhou等人^[12]提出UNet++,使用更深的网络并引入多级跳跃连接,在对比度低的皮肤病变图像中,分割的精确度仍然不够高;2019年,Ni等人^[13]提出RAU-Net相较于U-Net引入注意力机制并使用残差连接,在数据有限的情况下,很难达到理想的分割效果;2021年,Gu等人^[14]提出CE-Net,在编解码中应用密集空洞卷积模块和多核池化模块,CE-Net在

面对少见或异常的病变类型时,难以准确分割;2020年,Jha等人^[15]提出DoubleU-Net模型,将两个U-Net集成起来捕获更多语义切分以提高分割性能,显著增加模型的参数数量,导致了训练时间过长;2021年,Tong等人^[16]通过使用注意力门、空间注意力和通道注意力改进U-Net,捕获更多上下文信息和特征之间空间相关性,提高了网络的复杂性和参数数量,导致了模型训练过程中速度变慢;2021年,Zuo等人^[17]提出一种高效深入R2AU-Net,利用循环残余卷积和注意力机制得到一个鲁棒跳跃连接,但空间信息利用不充分,分割精度仍有提升空间;2021年,Chen等人^[18]提出TransUNet,Transformer将卷积神经网络特征图编码为上下文序列,解码器对编码的特征进行上采样并与编码路径的相应层的不同特征图组合,实现精准分割,引入自注意力机制和过多参数数量,容易带来梯度消失或梯度爆炸问题;2022年,邱拴虎等人^[19]在RAU-Net深度U型网络中融合带有反卷积操作的残差模块和注意力机制,增强图像特征并抑制无关信息,在反卷积层中若对权重和步长设置不当,会引起棋盘格效应,导致边界处理不精确;2023年,崔少国等人^[20]将深度可分离卷积和卷积块注意力机制嵌入到U-Net,深度可分离卷积中的深度卷积和点卷积的分离不能同时捕捉空间和通道间的特征;2023年,梁礼明等人^[21]在U-Net中融合可伸缩级联模块、自校准注意力模块和Transformer特征提取模块提高算法性能,多模块的融合显著增加训练阶段的计算需求,无法满足需要快速反馈的临床实施需求;2023年,Le等人^[22]提出移动抗锯齿注意力U-Net模型,采用反锯齿层以减小图像的空间分辨率,模型需要精细调参,增加训练复杂度;2023年,Ruan等人^[23]整合组向量乘注意力模块和组桥接模块以多视角提取病理信息,但其参数仍可以继续优化来达到更优效果;2024年,Wu等人^[24]提出MHorU-Net模型,采用高阶空间交互模块和多级多尺度信息融合机制显著增加了模型的深度和参数数量,导致了训练时间过长;2024年,Zhi等人^[25]整合了教师-学生架构,利用自我蒸馏分层机制以提高分割精度,但很难找到教师损失函数权重的最优比例,训练难度较大;2024年,Zhu等人^[26]提出MSS-UNet模型,结合多层感知机和多层移位感知模块,对低对比度的皮肤病变图像分割精度仍有提升空间;2024年,Xu等人^[27]将空洞卷积与深度可分离卷积相结合以构建HC-MAMBA模型,通过在卷积核中

插入 0 扩展感受野会导致捕捉到的特征缺乏连贯性; 2024 年, Nguyen 等人^[28]通过结合卷积神经网络与视觉 Mamba 框架, 构建混合特征提取器, 但模型架构复杂和使用多种高级组件会影响其在不同种类的皮肤病变图像的泛化能力. 2024 年, 陶惜婷等人^[29]通过构建特征融合模块来整合卷积神经网络和 Transformer 分支中不同分辨率的特征图, 但其计算复杂度较高, 对硬件设备要求较高.

在皮肤病变分割中, 卷积结构仍占主流, 卷积着重于对局部信息特征的提取, 无法对全局上下文特征建立有效的长程联系, 正常皮肤背景和病变皮肤前景外观相似, 单一尺度信息受限于感受野大小, 无法提取到有区别的上下文特征, 同时没有着重于边界轮廓信息的提取, 导致分割效果不够好.

基于上述问题, 本文提出一种边缘增强结合多尺度信息融合的网络, 采用对称编码-解码结构, 在编码器中构建一种新颖的变形器 (Transformer) 模块和卷积模块以并行提取全局特征和局部特征; 为融合不同

尺度的特征信息, 设计多尺度感知模块; 为增强边缘的细节特征, 设计一种边缘增强模块, 使用通道协调注意力 (coordinate attention, CA) 和全局上下文注意力 (global context, GC) 并行提取通道和空间特征信息; 构建轻量化解码器, 在保持高精度的同时使用较少的参数量; 使用 Focal Loss 和 IoU Loss 构建混合损失函数.

1 本文分割算法

本文网络采用对称编码-解码结构, 如图 1 所示, 整个网络由 5 部分组成, 包括编码器、多尺度感知模块、边缘增强模块、解码器、混合损失函数. 每层编码器由卷积模块和 Transformer 模块构成, 以提高网络对局部特征和全局特征的提取能力, 每层多尺度感知模块中使用多种扩张率的门控空洞卷积模块并行提取特征, 以获取不同感受野的多尺度信息, 边缘增强模块将低阶特征图与高阶特征图进行卷积聚合, 以提高对病灶边缘部分的分割精度, 构建轻量化解码器, 在使用较少参数量的同时维持较高的分割精确度.

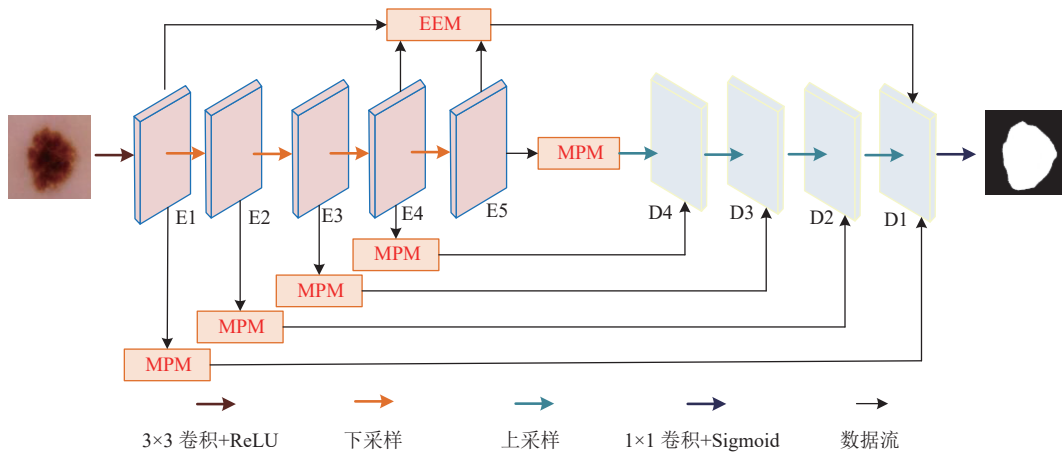


图 1 本文分割算法

1.1 编码器

由 5 层特征提取模块组成, 每 1 层结构由图 2 所示, 卷积块采用可变形卷积 (deformable convolution, DConv)^[30]进行提取信息, 可变形卷积引入可学习偏移量, 允许在卷积过程中对采样点位置进行微小变化, 可变形卷积公式如下:

$$F(x, y) = \sum_{i=1}^K \sum_{j=1}^K I(x + f_x(i, j), y + f_y(i, j)) \times K_d(i, j) \quad (1)$$

其中, (x, y) 是输出特征图 F 的像素坐标, $I(x + f_x(i, j), y +$

$f_y(i, j))$ 是输入图像 I 中, 根据核位移调整后的像素值, $K_d(i, j)$ 是权重, $f_x(i, j)$ 和 $f_y(i, j)$ 是可学习位移偏置, 它们用于调整卷积核在输入图像上的位置. 因此可变形卷积可以在感受野内自适应调整采样点, 更好捕捉局部细节信息.

卷积核为 3×3 , 步长为 1, 填充大小为 1, 针对皮肤病变形状不规则, 可变形卷积能够自适应的调整卷积核的采样位置, 以更好地适应不规则形状的特征, 更好的捕捉局部细节信息. 同时在不增加卷积核尺寸的情况下增大感受野, 捕捉较大范围的上下文信息, 有助于

提升皮肤病分割的准确性。

为解决卷积神经网络对特征图重复提取相同特征的问题,本文构建了多尺度池化块 (multi scale pooling block, MSPB),由图3所示,使用3个平均池化层 (avg

pool),来保留重要信息和聚焦重要信息,池化核大小分别为 (3, 5, 9)。首先采用 Conv1×1 聚合局部信息,然后通过3个并行池化层捕获多尺度特征,随后将不同尺度的特征图进行拼接,使用 Conv1×1 降低维度。

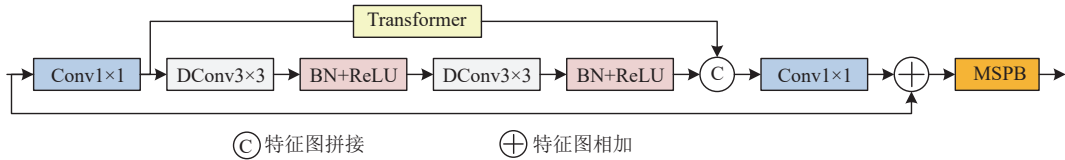


图2 每一层编码器结构

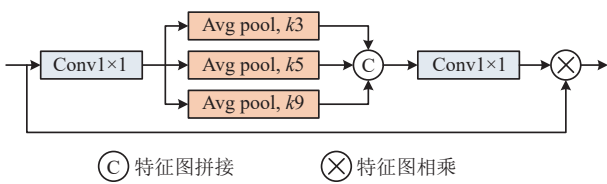


图3 多尺度池化块

为获取丰富的全局特征,针对Transformer原生多头注意力机制的计算成本高、效率低的问题,本文基于轻量化的无参数 SimAM (a simple parameter free attention module for convolutional neural network) 注意力^[31]构建了一个新颖的Transformer架构,如图4所示,以处理图像中不同区域之间的长距离依赖关系,来弥补卷积操作只能捕获局部信息的缺点,包括深度可分离卷积 (depthwise separable convolution, DWConv)^[32], SimAM 注意力机制,全连接前馈神经网络 (feed forward network, FFN), SimAM 无参数注意力捕获上下文信息,深度可分离卷积捕获局部信息,FFN 由两个线性变换层和 ReLU 激活函数组成,FFN 通过引入非线性变换

来加强模型建模能力和提高模型表达能力,使用残差连接将不同分支的特征图进行融合。

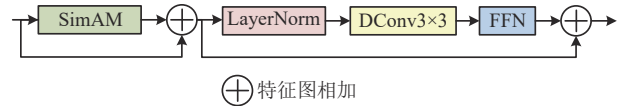


图4 Transformer 模块

1.2 多尺度感知模块

针对单一尺度难以融合上下文信息的问题,为学习不同尺度的特征,本文构建一个多尺度感知模块 (multi scale perception module, MPM),如图5所示,为了融合多尺度信息,该模块采用稠密连接金字塔结构^[33],稠密连接结构能够使得每一层都能够接收到之前所有层的信息,使得网络中的信息流动更加顺畅,相比在传统卷积神经网络中依靠局部感受野及网络卷积层和全连接层中参数通过独立学习得到,稠密连接能够利用全局信息,提高模型的准确性和鲁棒性,同时将每一层输出传递给之后的所有层,使每一层共享之前层的参数,达到减少参数数量的目的。

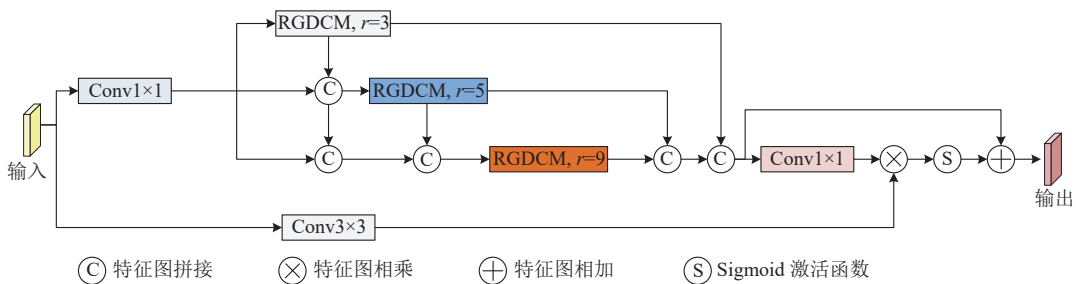


图5 多尺度感知模块

受到谭岩杰等人工作^[34]的启发,本文在原有的空洞卷积基础上增加残差门控机制,形成残差门控空洞卷积结构 (residual gated dilated convolution module, RGDCM),相较于普通的空洞卷积,残差门控空洞卷积

中的门控机制用于控制数据的流通,从而降低梯度消失的风险,残差机制实现信息的多通道传输。残差门控空洞卷积结构如图6所示,其中, ConvD₁ 和 ConvD₂ 均代表空洞卷积 (dilated convolution),二者卷积核个数大

小一致权重不共享, \odot 是 Sigmoid 激活函数, \oplus 为向量的哈达玛乘积, \oplus 为特征图的相加操作。

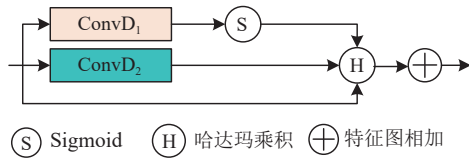


图6 残差门控空洞卷积模块

使用不同大小空洞卷积率扩大感受野, 以有效获取多尺度信息, 为避免因多次叠加相同尺寸卷积核所产生的网格效应, 本文将网络中扩张率 (dilation rate) 设置为 (3, 5, 9), 特征图通过 Conv1×1 升高维度, 同时并行增加一条扩张率为 1 的卷积, 即普通卷积, 以加强细节信息的权重, 同时引入局部残差连接, 防止出现梯度消失问题, 使用 Sigmoid 激活函数进行概率映射。

1.3 边缘增强模块

浅层特征中包括丰富的边缘细节和非边缘细节, 需要利用深层特征信息来促进对边缘特征的探索. 受到 Sun 等人工作^[35]的启发, 本文构建了一个边缘增强模块 (edge enhancement module, EEM), 如图 7 所示,

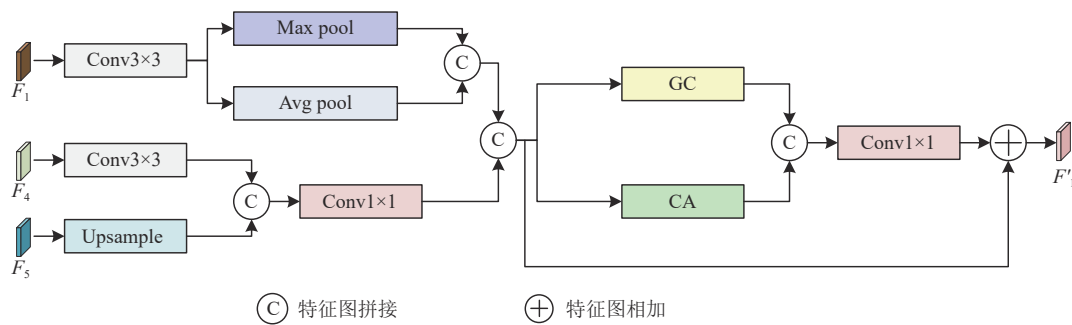


图7 边缘增强模块

1.4 解码器

构建 4 层解码器, 每 1 层如图 9 所示, 使用 Conv3×3 将多种特征进行整合, 有助于全局信息与局部信息相结合, 利用批归一化处理层 (batch normalization, BN) 提高泛化能力, 利用 GELU 激活函数更加平滑的特性来抑制梯度消失, 使用 CARAFE 轻量化算子进行上采样, CARAFE 算子能够在不同位置进行细粒度的特征重组, 这对捕获不同形状和大小的目标非常有帮助, 相对传统的上采样操作如双线性插值, CARAFE 算子需要更少的计算资源。

F_1 进行 Conv3×3 操作, 进入并联池化模块, 并联池化模块由最大池化层 (max pool) 和平均池化层 (avg pool) 并行连接构成, 最大池化层利用最大值以保留重要特征, 它能够捕捉到特征图内最重要的特征, 例如边缘、纹理等; 平均池化层利用计算平均值来提取信息, 它对特征图内信息进行平滑处理, 从而减少由噪声所产生影响, 能够使得网络捕捉到整体特征, 通过并联两种池化层, 可以充分利用两种池化层的优势, 从而更好地进行特征表示和降维, 以提高模型性能. F_4 进行 Conv3×3 操作, 与 F_5 进行拼接, 通过 Conv1×1 降维, 再与 F_1 进行拼接, 进入混合注意力 (mixed attention, MA) 模块. MA 模块包括空间注意力 (global context, GC) 和通道注意力 (coordinate attention, CA), 如图 8 所示, GC 注意力关注像素点之间的位置关系, 强化特征图的边缘和纹理等重要信息, 能够抑制噪声等无关信息的干扰, 减小对干扰信息的误判程度; CA 注意力分支提取通道信息以在通道维度上重新进行加权, 关注不同通道的重要性, 处理不同类型噪声, 更灵活的适应不同噪声场景, 提高网络鲁棒性, 随后通过 Conv1×1 来降低维度去除冗余特征同时加快计算速度, 增加一条局部残差连接, 以解决梯度消失问题。

1.5 损失函数

本文使用 Focal Loss 和 IoU Loss 作为混合损失函数. 使用 Focal Loss 抑制类别不平衡问题, 使用 IoU Loss 使模型更加关注分割结果的边界部分, 通过将 Focal Loss 和 IoU Loss 相加得到总损失 L .

Focal Loss 数学公式:

$$FocalLoss = -\alpha(1 - pt)^\gamma \log(pt) \quad (2)$$

其中, α 是平衡因子来控制正负样本的权重, pt 是模型对真实标签的预测概率, γ 是聚焦参数, 当 $\gamma > 0$ 时, 难以分类的样本会得到更大的权重. 对于本文模型, 经过多

次实验得证,将 α 设置为 0.27, γ 设置为 3 为最佳值.

IoU Loss 数学公式:

$$IoULoss = 1 - \frac{intersection}{union} \quad (3)$$

其中, *interaction* 表示预测值与真实值的交集, *union* 表示预测值与真实值的并集. 数值越小表示预测值与真实值越接近.

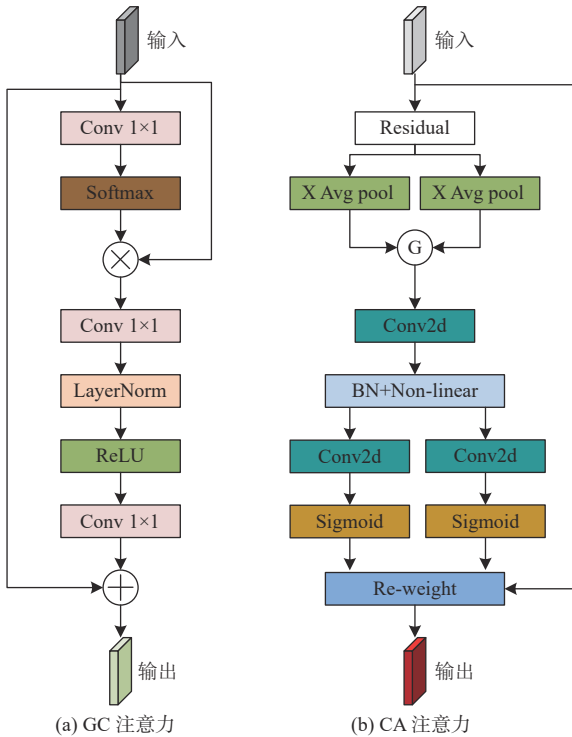


图 8 GC 注意力和 CA 注意力



图 9 每一层解码器结构

2 实验与结果分析

2.1 实验准备

2.1.1 数据集

使用国际皮肤成像合作组织发布的大规模皮肤病变图像数据集,即 ISIC2016 数据集和 ISIC2018 数据集,进行对模型进行训练与评估工作.

ISIC2016 数据集包括 1279 张皮肤病变图像,其中训练集 900 张,测试集 379 张,本文通过 Mosaic 和 Mix up 方法进行图像扩充,将训练集扩充到 2250 张图像,将测试集扩充到 947 张图像,将 947 张测试集随机拆分为 319 张验证集和 628 张测试集,以满足训练集、验证集以及测试集的比例约为 7:1:2.

ISIC2018 数据集包括分割数据集和分类数据集,本文使用其中的分割数据集进行训练评估,其中分割数据集包括 3 694 张皮肤病变图像,其中训练集为 2594 张,验证集为 100 张,测试集为 1 000 张,本文通过 Mosaic 和 Mix up 方法进行图像扩充,将训练集扩充到 3 891 张图像,将验证集扩充到 250 张,将测试集扩充到 2 500 张图像,在测试集中随机取出 757 张图像放入训练集中,随机取出 414 张图像放入验证集中,以满足训练集、验证集以及测试集的比例约为 7:1:2.

2.1.2 数据扩充

为适应模型训练,本文使用 Mosaic 和 Mix up 方法进行图片扩充,可视化如图 10 所示, Mosaic 每次处理 4 张图片,首先依据中心位置进行裁剪原始图片,再随机排列拼接; Mix up 每次处理 2 张图片,先将其翻转、缩放以及变化色域,再按照一定比例混合,通过这两种方法进行数据增强,以防止训练模型时发生过拟合.

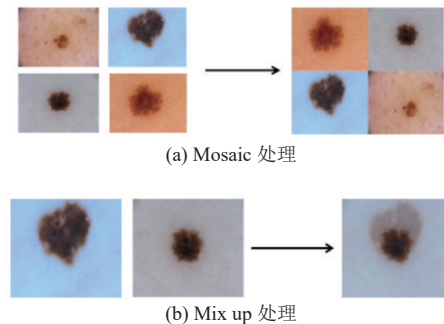


图 10 数据扩充操作

2.1.3 实验环境与参数设置

1) 硬件环境

CPU: 14 vCPU Intel Xeon(R) Gold 6330 CPU @ 2.00 GHz; GPU: NVIDIA RTX 3090 (24 GB); RAM: 16 GB.

2) 软件环境

操作系统: Ubuntu 18.04; 部署环境: Python 3.8; 深度学习框架: PyTorch 1.1.8.1; 加速计算架构: CUDA 11.6.

3) 参数设置

输入图像尺寸: 256×256; 参数学习率: 0.001; 动量: 0.9; 权重衰减系数: 0.0005; 批样本大小 (batch size): 8; epoch: 200.

为加快模型训练的收敛速度,将图像大小调整为 256×256×3,批样本大小 (batch size) 是指一次训练所选择的样本数量,考虑现有 GPU 处理图像的能力,将其

设为 8; 随着训练轮数的增加, 神经网络中权重在不断更新, 训练集误差不断减小, 曲线从欠拟合变的过拟合, 经多次训练可知, 200 为最佳训练轮数, 发现误差接近 0 时提前中止训练能够避免网络发生过拟合现象。

2.1.4 评价指标

本文使用特异度 (specificity, SP)、杰卡尔德系数 (Jaccard index, JA)、骰子系数 (Dice similarity coefficient, DC)、敏感度 (sensitivity, SE)、准确度 (accuracy, AC) 共 5 个性能指标作为评价标准。各指标值分别通过式 (4)–式 (8) 得到。各公式中, TP 为被模型预测为正类的正样本, TN 为被模型预测为负类的负样本, FN 为被模型预测为负类的正样本, FP 为被模型预测为正类的负样本。指标值越大表示分割性能越优, 含义如下。

1) JA : 杰卡尔德系数, 预测分割与真实分割的重叠区域和全部区域的比例。

$$JA = \frac{TP}{TP + FN + FP} \quad (4)$$

2) DC : 骰子系数, 预测分割与真实分割区域空间重叠程度。

$$DC = \frac{2TP}{2TP + FN + FP} \quad (5)$$

3) SE : 敏感度, 所有实际病变区域中被正确识别和分割比例。

$$SE = \frac{TP}{TP + FN} \quad (6)$$

4) AC : 准确度, 整个图像预测准确性, 包括病变区域和正常区域。

$$AC = \frac{TP + TN}{TP + TN + FN + FP} \quad (7)$$

5) SP : 特异度, 正常皮肤区域预测准确度。

$$SP = \frac{TN}{TN + FP} \quad (8)$$

2.2 消融实验

为验证本文所设计算法中各个模块有效性, 在数据集 ISIC2018 上, 本文对编码器模块、多尺度感知模块、边缘增强模块、轻量化解码器进行消融实验, 实验均在相同的环境下进行, 并使用相同的数据增强, 以确保公平性, 结果见表 1, 表中加粗字体为最优值, 分析如下。

表 1 消融实验结果 (%)

序号	Transformer模块	CNN	MPM	EEM	轻量解码器	JA	DC	SE	AC	SP
1	—	√	√	√	√	82.54	89.68	88.74	94.03	93.35
2	√	—	√	√	√	83.73	90.34	88.96	93.14	93.52
3	√	√	—	√	√	83.03	89.85	88.21	94.03	93.41
4	√	√	√	—	√	84.29	90.13	89.03	93.24	94.01
5	√	√	√	√	—	83.87	90.52	89.29	93.78	94.05
6	√	√	√	√	√	84.56	90.78	89.36	94.26	94.12

注: CNN表示编码器中卷积模块

1) 第 1 组, 去除编码器中 Transformer 模块分支, 保留卷积块, 编码器层数不变, 5 项指标均下降。

2) 第 2 组, 去除编码器中卷积块, 保留 Transformer 模块分支, 编码器层数不变, 5 项指标均下降。

3) 第 3 组, 去除多尺度感知模块, 5 项指标均下降。

4) 第 4 组, 去除边缘增强模块, 5 项指标均下降。

5) 第 5 组, 将轻量化解码器替换为传统卷积解码器, 解码器层数不变, 5 项指标均下降。

从表 1 可以看出, 缺少任何部分都会使得网络性能出现不同程度的降低。第 1 组实验证明 Transformer 模块能够有效提取全局上下文信息, 第 2 组实验证明卷积模块能够有效提取局部信息, 第 3 组实验证明多尺度感知模块能够有效融合多尺度信息并有效克服提取上下文受限问题, 第 4 组实验证明深层信息能够有效促进边缘信息的提取精细度及更好的保留细节特征。

第 5 组实验证明本文构建的轻量化解码器维持了较高的分割精度。第 6 组实验证明本文网络结构有较强的抑制噪声能力, 达到较高分割精度。

2.3 对比实验

本文选取医学图像分割的代表性算法 U-Net、UNet++、Ms RED、FAT-Net、TransFuse、TransU-Net、CE-Net 进行对比实验。在公开数据集 ISIC2016 和 ISIC2018 上, 将本文算法与上述算法进行对比实验, JA 、 DC 、 SE 、 SP 和 AC 指标实验结果数据 (见表 2), 分析如下。

1) 数值分析

在 ISIC2016 上, JA 、 DC 、 SE 、 SP 和 AC 指标分别达到 84.33%、92.61%、91.87%、93.01%、92.91%, 在 ISIC2018 上, JA 、 DC 、 SE 、 AC 和 SP 指标分别达到 84.56%、90.78%、89.36%、94.26%、94.12%, 5 项

指标均获得最优值,说明针对边缘模糊、形状不规则且存在伪影的问题,本文模型中多尺度感知模块考虑到了多尺度信息,边缘增强模块考虑到了边缘特征信息,对病灶细节部分和边缘分割精确度更高,同时本文网络参数量(Params)约为 TransUNet 的 71.86%,虽然在网络参数和推理速度不如 Ms RED 和 U-Net,但在各项评价指标均高于 Ms RED 和 U-Net,表明本文所提出模型相对于其他 5 种先进性的分割模型来说兼顾了模型分割精度和模型复杂度,充分说明本文所提算法在综合性能优于其他算法。

2) 可视化分析

图 11 和图 12 选取边缘模糊、对比度低、形状不规则、伪影、毛发干扰等具有挑战性图像,其中第 2 列是样本真实标签,第 3-10 列分别为 U-Net、UNet++、Ms RED、TransUNet、FAT-Net、CE-Net、TransFuse 和本文模型的分割结果,比较不同网络的分割结果,分析如下。

表 2 对比实验结果

数据集	方法	JA (%)	DC (%)	SE (%)	AC (%)	SP (%)	Params (M)	Time (s)
ISIC2016	U-Net ^[11]	80.31	81.67	83.32	85.13	87.36	15.21	266
	UNet++ ^[12]	82.87	85.36	85.42	87.64	89.28	16.83	302
	Ms RED ^[6]	82.92	86.73	86.58	88.64	89.36	11.34	253
	TransUNet ^[18]	83.02	86.67	87.39	89.74	91.28	36.71	672
	FAT-Net ^[7]	83.47	89.38	90.41	89.65	91.32	28.38	501
	CE-Net ^[14]	83.36	88.58	90.36	89.16	91.83	29.60	523
	TransFuse ^[8]	83.85	90.32	90.02	91.33	91.76	27.56	458
	Ours	84.33	92.61	91.87	93.01	92.91	26.38	356
ISIC2018	U-Net ^[11]	80.51	85.34	84.31	88.62	89.41	15.21	286
	UNet++ ^[12]	82.12	87.13	86.51	89.65	89.19	16.83	335
	Ms RED ^[6]	82.01	87.36	86.03	89.31	90.05	11.34	268
	TransUNet ^[18]	82.36	88.31	86.61	92.71	93.13	36.71	778
	FAT-Net ^[7]	83.58	88.59	88.36	93.47	92.69	28.38	601
	CE-Net ^[14]	83.26	88.59	88.36	93.47	92.69	29.60	618
	TransFuse ^[8]	83.01	87.96	88.06	92.87	91.43	27.56	578
	Ours	84.56	90.78	89.36	94.26	94.12	26.38	396

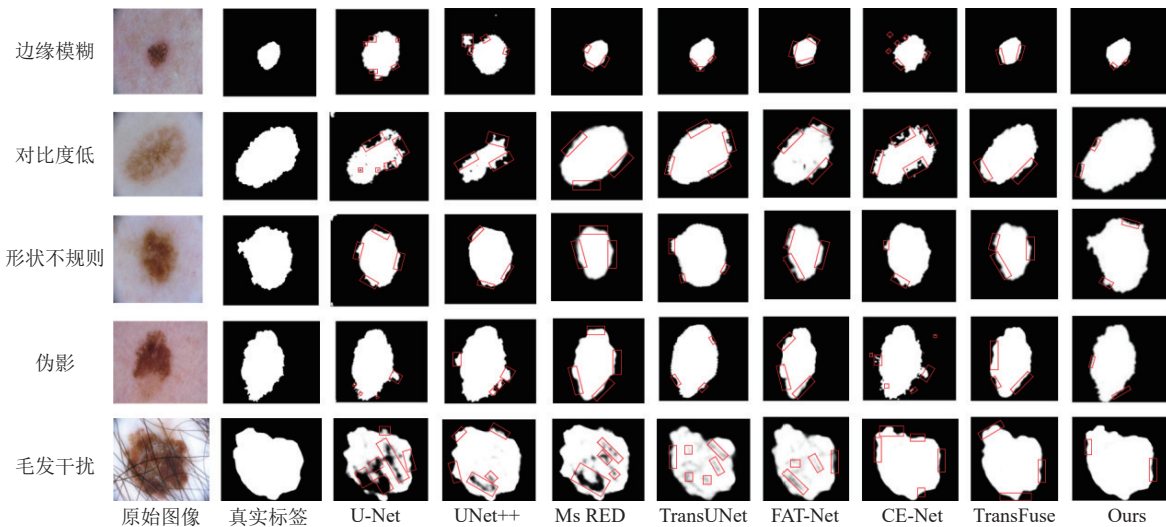


图 11 ISIC2016 对比实验可视化结果

图 11 和图 12 中,使用红色框标出图像分割主要差异点,由图 11 和图 12 中可以看出,本文所提算法整体分割精确度更高,U-Net 和 UNet++对非病变区域存在一些漏检和误检,对于毛发干扰这类问题漏检尤其严重,这是由于其对全局特征利用不充分,提取上下文信息受限,对目标特征重构能力不足,容易丢失边缘特征,在建模远程依赖关系方面显示出局限性,对于形状不规则、对比度较低以及存在毛发干扰的样本尤为明显;Ms RED 对于病变图像边缘分割精度不够高以及存在欠分割问题,这是由于其仅依靠卷积操作,缺乏捕获

长距离依赖关系的能力,对于毛发遮挡这类挑战性的问题尤为明显;TransUNet 对于多尺度特征的利用操作大多数在上采样过程中进行,对于下采样丢失的细节信息利用不够,对于边缘细节的分割精度仍有提升空间;FAT-Net 和 TransFuse 都是通过结合卷积神经网络和传统 Transformer 架构的优势获得图像分割精度的提高,但这样会显著提高模型的复杂度以及降低推理速度;CE-Net 引入空洞卷积模块和残差多核池化,加强了信息提取能力,但部分存在分割边缘模糊问题。

上述流行算法对全局上下文建模能力不足或缺少

对低级特征的有效利用, 以及存在推理时间较长等问题, 不能有效地保留细节和边缘信息, 对背景噪声不能很好的区分, 本文算法编码器中采用 Transformer 和卷积结构并行处理特征能够使全局特征与局部特征双向融合以提高分割精度, 多尺度感知模块可以实现特征重用及捕获多尺度信息, 能够有效避免单一尺度融合

策略提取上下文受限问题, 边缘增强模块能够有效的结合深层特征和浅层特征, 从而有效地保留细节信息和边缘信息以提高边缘部分的分割精度, 轻量化解码器能够在使用较少参数的同时保持较高的分割精度, 因而可以在背景噪声的影响下实现较高精度的分割, 本文算法对于边缘细节的分割精度有明显提升.

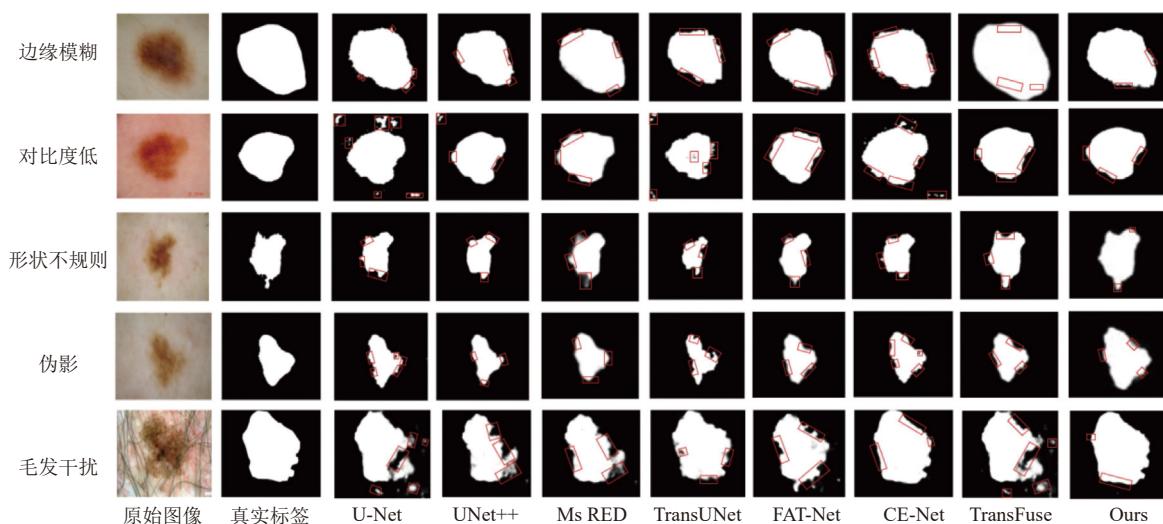


图 12 ISIC2018 对比实验可视化结果

3 结论与展望

针对皮肤病灶大小不一、与周围皮肤的对比度低, 病灶边界不规则、模糊、存在伪影及毛发干扰问题, 本文提出边缘增强结合多尺度信息融合的皮肤病变分割算法, 编码器采用 Transformer 模块和卷积模块并行提取全局信息和局部信息, 构建多尺度感知模块融合上下文多尺度信息, 设计边缘增强模块来使得边缘分割更加清晰, 构建轻量解码器在维持较高分割精度的同时使用较少的参数量. 本文所提出的网络对 ISIC2016 和 ISIC2018 数据集中的皮肤病变像素分割精度数值较高, 对皮肤病的诊断有一定参考价值, 后续将继续研究提高本文网络分割的精细度以及轻量化.

参考文献

- 1 Jin QG, Cui H, Sun CM, *et al.* Cascade knowledge diffusion network for skin lesion diagnosis and segmentation. *Applied Soft Computing*, 2021, 99: 106881. [doi: 10.1016/j.asoc.2020.106881]
- 2 Xia SR, Krishnan SM, Tjoa MP, *et al.* A novel methodology for extracting colon's lumen from colonoscopic images. *Journal of Systemics, Cybernetics and Informatics*, 2003, 1(2): 7–12.

- 3 Celebi ME, Iyatomi H, Schaefer G, *et al.* Lesion border detection in dermoscopy images. *Computerized Medical Imaging and Graphics*, 2009, 33(2): 148–153. [doi: 10.1016/j.compmedimag.2008.11.002]
- 4 Moscheni F, Bhattacharjee S, Kunt M. Spatio-temporal segmentation based on region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(9): 897–915. [doi: 10.1109/34.713358]
- 5 王学松, 周成全, 樊亚春, 等. 彩色图像色度距离权值的图论分割算法. *中国图象图形学报*, 2011, 16(2): 221–226. [doi: 10.11834/jig.20110217]
- 6 Dai DW, Dong CX, Xu SH, *et al.* Ms RED: A novel multi-scale residual encoding and decoding network for skin lesion segmentation. *Medical Image Analysis*, 2022, 75: 102293. [doi: 10.1016/j.media.2021.102293]
- 7 Wu HS, Chen SH, Chen GL, *et al.* FAT-Net: Feature adaptive Transformers for automated skin lesion segmentation. *Medical Image Analysis*, 2022, 76: 102327. [doi: 10.1016/j.media.2021.102327]
- 8 Zhang YD, Liu HY, Hu Q. TransFuse: Fusing Transformers and CNNs for medical image segmentation. *Proceedings of the 24th International Conference on Medical Image Computing and Computer Assisted Intervention*. Strasbourg: Springer, 2021. 14–24.
- 9 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *Proceedings of the 2015*

- IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 3431–3440.
- 10 Yuan YD, Chao M, Lo YC. Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. *IEEE Transactions on Medical Imaging*, 2017, 36(9): 1876–1886. [doi: [10.1109/TMI.2017.2695227](https://doi.org/10.1109/TMI.2017.2695227)]
 - 11 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. *Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention*. Munich: Springer, 2015. 234–241.
 - 12 Zhou ZW, Siddiquee MMR, Tajbakhsh N, *et al.* UNet++: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging*, 2020, 39(6): 1856–1867. [doi: [10.1109/TMI.2019.2959609](https://doi.org/10.1109/TMI.2019.2959609)]
 - 13 Ni ZL, Bian GB, Zhou XH, *et al.* RAUNet: Residual attention U-Net for semantic segmentation of cataract surgical instruments. *Proceedings of the 26th International Conference on Neural Information Processing*. Sydney: Springer, 2019. 139–149.
 - 14 Gu ZW, Cheng J, Fu HZ, *et al.* CE-Net: Context encoder network for 2D medical image segmentation. *IEEE Transactions on Medical Imaging*, 2019, 38(10): 2281–2292. [doi: [10.1109/TMI.2019.2903562](https://doi.org/10.1109/TMI.2019.2903562)]
 - 15 Jha D, Riegler MA, Johansen D, *et al.* DoubleU-Net: A deep convolutional neural network for medical image segmentation. *Proceedings of the 33rd IEEE International Symposium on Computer-based Medical Systems*. Rochester: IEEE, 2020. 558–564.
 - 16 Tong XZ, Wei JY, Sun B, *et al.* ASCU-Net: Attention gate, spatial and channel attention U-Net for skin lesion segmentation. *Diagnostics*, 2021, 11(3): 501. [doi: [10.3390/diagnostics11030501](https://doi.org/10.3390/diagnostics11030501)]
 - 17 Zuo Q, Chen SY, Wang ZF. R2AU-Net: Attention recurrent residual convolutional neural network for multimodal medical image segmentation. *Security and Communication Networks*, 2021, 2021(1): 6625688.
 - 18 Chen JN, Lu YY, Yu QH, *et al.* TransUNet: Transformers make strong encoders for medical image segmentation. *arXiv:2102.04306*, 2021.
 - 19 邱拴虎, 杨文瀚, 廖苗, 等. 基于 RA-Unet 的 CT 图像肝脏肿瘤分割. *仪器仪表学报*, 2022, 43(8): 65–72.
 - 20 崔少国, 文浩, 张宇楠, 等. 带有注意力机制的 OCTA 视网膜血管分割方法. *计算机工程与应用*, 2023, 59(18): 163–171.
 - 21 梁礼明, 陈鑫, 周珑颂, 等. 基于 U-Net 多尺度自校准注意力视网膜分割算法. *计算机应用研究*, 2023, 40(3): 943–948.
 - 22 Le PT, Pham BT, Chang CC, *et al.* Anti-aliasing attention U-Net model for skin lesion segmentation. *Diagnostics*, 2023, 13(8): 1460. [doi: [10.3390/diagnostics13081460](https://doi.org/10.3390/diagnostics13081460)]
 - 23 Ruan JC, Xie MY, Gao JS, *et al.* EGE-UNet: An efficient group enhanced UNet for skin lesion segmentation. *Proceedings of the 26th International Conference on Medical Image Computing and Computer-assisted Intervention*. Vancouver: Springer, 2023. 481–490.
 - 24 Wu RK, Liang PC, Huang X, *et al.* MHorUNet: High-order spatial interaction UNet for skin lesion segmentation. *Biomedical Signal Processing and Control*, 2024, 88: 105517. [doi: [10.1016/j.bspc.2023.105517](https://doi.org/10.1016/j.bspc.2023.105517)]
 - 25 Zhi YC, Bie HX, Wang JL, *et al.* Masked autoencoders with generalizable self-distillation for skin lesion segmentation. *Medical & Biological Engineering & Computing*, 2024: 1–15, doi: [10.1007/s11517-024-03086-z](https://doi.org/10.1007/s11517-024-03086-z).
 - 26 Zhu WH, Tian JY, Chen MZ, *et al.* MSS-UNet: A multi-spatial-shift MLP-based UNet for skin lesion segmentation. *Computers in Biology and Medicine*, 2024, 168: 107719. [doi: [10.1016/j.compbiomed.2023.107719](https://doi.org/10.1016/j.compbiomed.2023.107719)]
 - 27 Xu JS. HC-Mamba: Vision MAMBA with hybrid convolutional techniques for medical image segmentation. *arXiv:2405.05007*, 2024.
 - 28 Nguyen VT, Pham VT, Tran TT. AC-MAMBASEG: An adaptive convolution and Mamba-based architecture for enhanced skin lesion segmentation. *arXiv:2405.03011*, 2024.
 - 29 陶惜婷, 叶青. 融合 CNN 和 Transformer 的并行双分支皮肤病灶图像分割. *计算机应用研究*, 2024, 41(8): 2554–2560. [doi: [10.19734/j.issn.1001-3695.2023.10.0600](https://doi.org/10.19734/j.issn.1001-3695.2023.10.0600)]
 - 30 Dai JF, Qi HZ, Xiong YW, *et al.* Deformable convolutional networks. *Proceedings of the 2017 IEEE International Conference on Computer Vision*. Venice: IEEE, 2017. 764–773.
 - 31 Yang LX, Zhang RY, Li LD, *et al.* SimAM: A simple, parameter-free attention module for convolutional neural networks. *Proceedings of the 38th International Conference on Machine Learning*. PMLR, 2021. 11863–11874.
 - 32 Chollet F. Xception: Deep learning with depthwise separable convolutions. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 1800–1807.
 - 33 Huang G, Liu Z, van der Maaten L, *et al.* Densely connected convolutional networks. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 2261–2269.
 - 34 谭岩杰, 陈玮, 尹钟. 门控空洞卷积与级联网络中文命名实体识别. *小型微型计算机系统*, 2023, 44(6): 1198–1203.
 - 35 Sun YJ, Wang S, Chen CLZ, *et al.* Boundary-guided camouflaged object detection. *Proceedings of the 31st International Joint Conference on Artificial Intelligence*. Vienna: ijcai.org, 2022. 1335–1341.

(校对责编: 孙君艳)