

基于多级特征组聚合的皮肤病U型分割网络

廖骏卿¹, 高琳¹, 邹茂扬¹, 黄海莹², 何晋¹

(1. 成都信息工程大学区块链产业学院, 四川成都 610225;

2. 四川大学华西第二医院, 四川成都 610044)

摘要: 皮肤病变区域的精准分割是计算机辅助皮肤病诊断的关键性研究任务, 它为医生提供了重要的诊断依据。尽管基于UNet的改进网络在图像分割领域取得了显著进展, 但是它们往往忽视了掩膜与多层级信息的结合运用, 从而限制了分割精度。引入Transformer有助于分割精度的提升, 但是网络的参数量也会显著增加, 从而可能导致计算复杂度和模型部署成本上升。针对这些问题, 提出一种基于多级特征组聚合的皮肤病U型分割网络(MFGA-UNet), 其能够充分融合掩膜信息与多层级信息, 在保证较低参数数量的同时, 实现高精度且轻量级的皮肤病图像分割。MFGA-UNet中, 首先, 采用改进的反瓶颈卷积模块替代标准UNet中的常规卷积块; 其次, 引入多级特征组聚合模块优化网络的跳跃连接, 有效融合掩膜信息与多层级特征, 丰富了特征层次; 最后, 利用深度监督技术改进损失函数, 通过解码器各层输出求取损失, 优化了模型训练过程。在ISIC2017和ISIC2018皮肤病数据集上的评估结果显示, MFGA-UNet的参数量仅有10.246 M, 且精度超越了现有6种医学图像分割网络, 在两个数据集中, Dice分别达到了94.273%、90.028%。

关键词: 图像分割; 皮肤病; U型网络; 反瓶颈卷积模块; 多级特征组聚合

DOI: 10.11907/rjdk.241439

中图分类号: TP391.41

文献标识码: A

开放科学(资源服务)标识码(OSID):

文章编号: 1672-7800(2025)008-0173-09



Multi-level Feature Group Aggregation UNet for Skin Lesion Segmentation

LIAO Junqing¹, GAO Lin¹, ZOU Maoyang¹, HUANG Haiying², HE Jin¹

(1. College of Blockchain Technology, Chengdu University of Information Technology, Chengdu 610225, China;

2. The West China Second University Hospital of Sichuan University, Chengdu 610044, China)

Abstract: The precise segmentation of dermatological lesions is a crucial research task in computer-aided skin disease diagnosis, providing vital diagnostic evidence for physicians. Although improved networks based on UNet have made significant progress in the field of image segmentation, they often overlook the integration of mask and multi-level information, thus limiting segmentation accuracy. The introduction of Transformer helps improve segmentation accuracy, but it also significantly increases network parameters, potentially leading to increased computational complexity and model deployment costs. To address these issues, we propose a skin lesion segmentation network based on Multi-level Feature Group Aggregation UNet (MFGA-UNet), which can fully integrate mask information and multi-level information while ensuring a low parameter count, achieving high accuracy and lightweight skin lesion segmentation. In MFGA-UNet, we first replace standard UNet convolution blocks with improved bottleneck convolution modules. Secondly, we introduce a multi-level feature group aggregation module to optimize the network's skip connections, effectively integrating mask information and multi-level features, enriching the feature hierarchy. Finally, we utilize deep supervision techniques to improve the loss function, computing losses from decoder layer outputs, optimizing the model training process. Evaluation on the ISIC2017 and ISIC2018 skin disease datasets shows that MFGA-UNet has only 10.246M parameters and outperforms six existing medical image segmentation networks, with Dice coefficients reaching 94.273% and 90.028% on the two datasets, respectively.

Key Words: image segmentation; skin lesion; UNet; inverted bottleneck convolution; multi-level feature group aggregation

收稿日期: 2024-06-18

扫描二维码阅读全文:

基金项目: 四川省科技计划项目(2020YFS0316); 四川省自然科学基金项目(2023NSFSC0482)

作者简介: 廖骏卿(2000-), 男, CCF 学生会会员, 成都信息工程大学区块链产业学院硕士研究生, 研究方向为医学图像分割; 高琳(1976-), 男, 博士, 成都信息工程大学区块链产业学院讲师, 研究方向为计算机视觉。本文通讯作者: 高琳。



0 引言

现有文献表明,皮肤癌是全球最常诊断的癌症类别,其中恶性黑色素瘤约占1/5^[1]。恶性黑色素瘤在历史上是一种罕见疾病,在过去50年中,在诸多国家其发病率持续上升。该疾病的风险通常随着年龄的增长而增加,在老年人群中的发病率更高。该疾病在各国均有分布,并且预计在2020年全球约有325 000例病例,约57 000人(32 000名男性,25 000名女性)死于该疾病^[2-3]。现阶段,国内有很多针对视网膜、肿瘤等的医学图像分割技术正不断产生,这些技术在疾病诊断等领域都取得了不错的效果^[4-7]。因此,利用人工智能,实现一种自动化皮肤病变图像分割技术,辅助医护人员快速识别皮肤病变区域,并为后续治疗过程提供指导具有良好价值。

当前,很多医学图像分割网络主要以UNet^[8]为基础,并在其基础上进行不断创新^[9-11]。这些改进网络针对UNet的卷积块、跳跃连接、损失函数加以改进,在其中使用新的卷积技术替换原始卷积,或者融合Transformer^[12]等前沿技术,实现精度、参数量、效率优化。基于Transformer改进的UNet的网络,由于本身的自注意力技术以及包含有多层的前馈神经网络,故普遍存在参数量大的情况。近年来,多尺度信息利用得到了研究人员的重视,其中较具代表性的ACC-UNet^[13]、EGE-UNet^[14]网络为人们提供了宝贵思路。在ACC-UNet中,通过逐点卷积与拼接的方式,将多层级的信息融合起来,使各层级的特征图能够从其他层级特征图中获取信息以丰富自身,进而提高了一定的分割精度。不足之处在于,该网络在针对多尺度信息改进时未考虑到掩膜信息的利用。在EGE-UNet中,通过将低维信息融合进高维信息中,实现了对高维信息的丰富效果,还引入了研究者们易忽视的掩膜信息,将这些多尺度信息进行合理利用,实现分割精度的提高。虽然该网络提出了对掩膜信息进行利用的方法,但是对于其他层级信息利用不充分。

针对上述参数量和多尺度信息利用不充分等问题,本文以UNet为基础,吸收ACC-UNet、EGE-UNet等创新思路,提出了一种新型的多级特征组聚合的皮肤病U型分割网络(Multi-level Feature Group Aggregation UNet, MFGA-UNet),该网络通过对UNet的卷积块、跳跃连接、损失函数进行了修改,合理优化了参数量,并且提升了精度。本文贡献如下:①提出了多尺度特征组聚合融合方法,针对跳跃连接进行修改,使跳跃连接中每个层级的特征图都能通过其他层级的信息与每层解码器中的掩膜信息以丰富自身信息,提升多尺度信息利用率;②采用由深度可分离卷积实现的反向瓶颈卷积块,在反向瓶颈卷积块上补充逐点卷积进行优化,使其能够直接替换UNet网络中的原始卷积块,从而提升模型对图像特征的提取能力;③利用深度

监督方法修改了损失函数,该方法将解码器中的每一层都进行输出,并计算不同阶段的损失,优化模型训练过程。

1 相关研究

深度学习技术已普遍用于医学图像分割任务,并逐渐朝精度、效率方面优化,研究人员通过不断融合新的技术以提高深度学习模型分割能力。其中,UNet网络在医学图像分割领域表现良好,它包含对称的编码器—解码器架构,能提取图像中的特征信息,并采用跳跃连接为解码器提供了在编码器的池化操作期间可能丢失的信息。目前,诸多医学图像分割网络都是在其基础上,针对性地在损失函数、卷积块、跳跃连接上进行改进。

1.1 基于卷积技术的分割网络

这类网络利用卷积构建一种新结构或运用新的卷积技术进行修改,以达到优化模型结构的目的。在损失函数改进上,UNet++提供了一种名为深度监督的损失函数,将每层的特征图单独使用一个1×1卷积进行输出,使网络能够求取不同层级的损失函数,进而实现精度提升^[15]。在跳跃连接的改进上,MultiResUNet通过在跳跃连接过程中添加额外的卷积块,并将其作为语义间隙的补救措施,合理地减少了语义间隙对训练和预测造成的影响^[16]。在卷积块的改进上,Unext与HFRUNet网络分别使用MLP与多孔空间金字塔池模块等方法替换高维卷积块,实现增强对象的上下文信息等效果^[17-18]。脑肿瘤轮廓分割网络将残差与增强上下文信息的模块相结合,提高了脑肿瘤分割精度^[19]。还有很多网络对UNet的多个部分进行混合改进,PRA-UNet3+网络使用残差注意力结构替换原有卷积,提升特征捕获能力,并且在跳跃连接中通过对多层级信息进行融合,丰富了全局信息^[20]。EGE-UNet提出组多轴Hadamard产品注意力模块,从不同角度提取病理信息,并且运用组聚合桥模块对低级特征、高级特征以及解码器在每个阶段生成的掩码进行分组聚合,有效地融合了多尺度信息,并使用深度监督技术提升模型训练能力。ACC-UNet提出了邻域上下文的分层聚合块替换了经典的卷积块,该模块模拟了自注意力机制,并且提出一个多层次特征编译块以实现跳跃连接,获取来自不同层级的编码器信息。上述网络中针对卷积块结构和损失函数的改进表现出了较好效果,但是这些网络对于多尺度信息利用并不充分,未考虑到多层级信息和掩膜信息的组合利用。本文针对此问题,提出了一种新的多级特征组聚合方法对其加以改进。

1.2 基于Transformer技术的分割网络

Transformer最初是自然语言处理任务的一种方法,其利用了自注意力机制使得模型能够捕捉长距离的依赖关系并且可以并行计算。Vision Transformer的出现引起了广泛关注,它能让模型在像素级别上对图像进行处理和分

析,并且通过自注意力机制使模型能够对图像中的不同像素点进行全局的上下文关联^[21]。因此,众多研究员将其添加到了UNet中,TransUNet使用Transformer对来自卷积神经网络特征图的标记化图像块进行编码,作为用于提取全局上下文的输入序列^[22]。Swin-UNet提出了具有移位窗口的分层Swin Transformer,并将其作为编码器以提取上下文特征,实现了从局部到全局的自注意力机制^[23]。UC-TransNet和MCTrans在跳跃连接中分别添加通道Transformer模块和多复合Transformer模块以解决原始跳跃连接的语义间隙问题^[24-25]。SMESwin-UNet在卷积块改动上沿用Swin Transformer,在跳跃连接上提出了CNN和Vision Transformer复合结构以融合多尺度语义特征和注意力,通过同时改进卷积块和跳跃连接以提高网络性能^[26]。这些基于Transformer改进的UNet网络,在跳跃连接的修改上能够很好地融合多层次信息,但是普遍存在参数量大的问题,一定程度上增加了模型部署成本。本文针对此问题,使用改进的反向瓶颈卷积块对其进行优化。

2 MFGA-UNet分割网络设计

2.1 网络总体结构设计

本文提出的多级特征组聚合的皮肤病分割网络如图1所示,该网络由4部分组成,分别为编码器、解码器、跳跃

连接、损失模块。在编码器、解码器中,均使用反瓶颈卷积块替换UNet原有的卷积块。在编码器中,每层反瓶颈卷积块执行两次,与UNet类似,第一次执行卷积运算并实现通道数修改,第二次仅执行卷积运算,之后使用最大池化(Max Pooling)实现特征图空间降维。在解码器中,所有反瓶颈卷积块仅执行卷积运算,在每层反瓶颈卷积运算执行后使用转置卷积(Transposed Convolution)实现上采样与通道数修改。在跳跃连接中使用多级特征组聚合模块,将编码器中完成计算的特征图与解码器对应的上一层级的掩膜传入其中,实现多层次信息分组特征聚合。在损失模块中,将解码器每个层级的信息都单独使用1×1卷积进行输出,得到每层结果,将这些结果使用双线性插值扩充到原有的掩膜大小,然后运用深度监督的方式进行损失函数计算。

2.2 反向瓶颈卷积块优化

本文使用反向瓶颈卷积对原始卷积块进行改进,反向瓶颈卷积是由Sandler等^[27]提出,Liu等^[28]对其进行了深入探讨。该卷积通过深度可分离卷积实现了类似Transformer的反向瓶颈结构,该方法有助于减少参数量并提升精度^[29]。原始卷积的参数量计算公式如式(1)所示。

$$Params_{conv} = C_{in} \times K \times K \times C_{out} \quad (1)$$

其中, C_{in} 为输入通道数, K 为卷积核大小, C_{out} 为输出通道数。

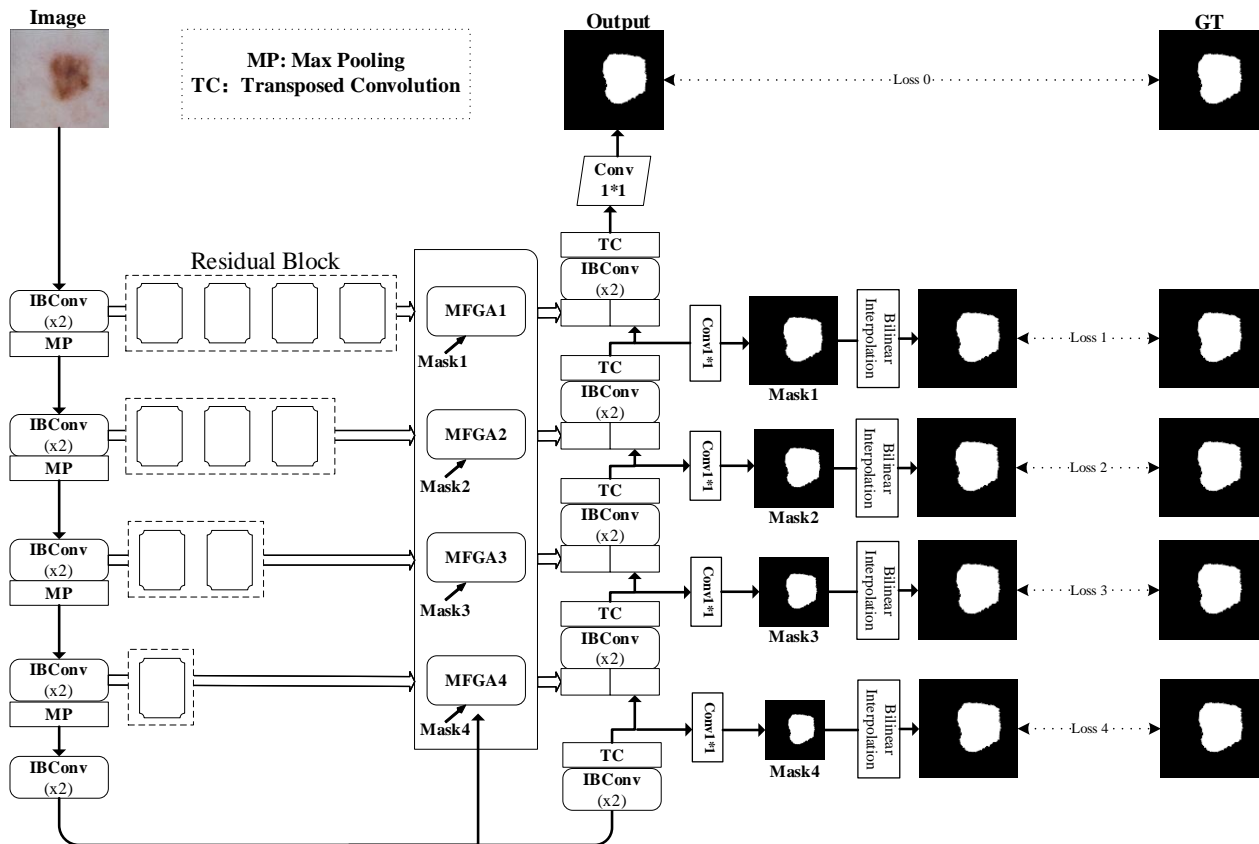


Fig. 1 Multi-level feature group aggregation UNet

图1 多级特征组聚合U型图像分割网络

而深度可分离卷积的参数量计算公式如式(2)所示。

$$Params_{dwc} = C_{out} \times (C_{in} + K \times K) \quad (2)$$

因此,普通卷积和深度可分离卷积之比为 $(C_{in} \times K \times K):(C_{in} + K \times K)$ 。当输出通道数为输入通道数的2倍时,反向瓶颈卷积块参数量可以写为 $4C_{in}(\frac{5}{2}C_{in} + K \times K)$,卷积核大小 K 为3时,原始卷积与反向瓶颈卷积块的参数量差值可以写为 $4C_{in} \times (2C_{in} - 9)$,故在输入通道数大于5时即可保证实现参数量优化。但是,如果不进行通道数修正,那么反向瓶颈卷积块参数量相较于原始卷积会增加 $36C_{in}$ 。

为保证模型的参数量与现有各种模型持平或更优,并且提升网络对通道信息的学习能力,本文使用了由深度可分离卷积实现的反向瓶颈卷积块以替换所有的原始卷积。

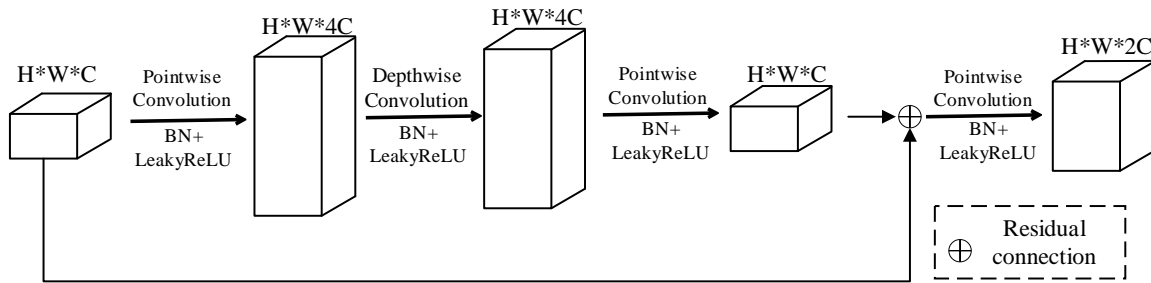


Fig. 2 Inverted bottleneck convolution block

图2 反向瓶颈卷积块

2.3 多级特征组聚合模块设计

本文使用多级特征组聚合模块对跳跃连接进行改进,多级特征组聚合模块由两部分构成:一是残差块(见图3),二是多级特征组聚合块。MultiResUNet中提到,在跳跃连接中添加卷积可以减少特征之间存在的语义间隙,这些语义间隙是输入特征与高层语义之间存在的差异,这些差异会对整个模型训练过程造成影响,进而对预测过程产生不利影响。鉴于此,本文在对多级特征进行融合之前添加残差块,以减少语义间隙带来的影响,实现对每层跳跃连接的特征图进行优化。残差连接虽然能在一定程度上减少语义间隙,改善梯度流动,解决梯度消失问题,但是过多的残差块堆叠可能导致计算开销增加等情况出现,并且假如在高层级的跳跃连接中过多增加残差块,会使网络过多学习训练集样本的边界等细节所包含的高维信息,从而导致网络在未见过的数据上表现不佳,出现过拟合现象,使模型在实际应用中的泛化性能下降。因此,跳跃连接每层残差块的数量随网络层数递增而逐渐减少,在第4层时仅设置一个残差块实现语义间隙消除。

多级特征组聚合块是本文提出的一个新颖块,现阶段部分网络考虑到了对多层次信息的融合,但是忽视了对掩膜信息的利用,掩膜信息属于多尺度信息的一种,而多尺度信息的获取对于医学图像分割任务而言十分重要,故本文将每个层级的信息与掩膜信息加以分组融合,实现更好

原始的反向瓶颈卷积结构不能实现通道修改的效果,故在其尾部补充一个逐点卷积进行优化,使卷积块能够提高对特征的理解和表达能力,并且能实现通道数修改。在激活函数选择上,使用了LeakyReLU替代原有的ReLU,以解决训练中可能遇到的“死亡神经元”问题,并提升一定的特征提取和表示能力。改进后的反向瓶颈卷积块能够直接替换原有卷积块,如图2所示。首先,使用逐点卷积将特征图通道扩充为原来的4倍,增加网络的表示能力,使模型能够捕捉更多特征信息;其次,用 3×3 的深度卷积进行运算,实现降低参数量的效果;再次,用逐点卷积将特征图通道还原,而后使用残差连接防止梯度消失和网络退化问题;最后,使用逐点卷积实现对卷积块特征图通道数修改的操作。在每次卷积运算后使用归一化及LeakyReLU激活函数以增强模型的非线性处理能力和提高训练稳定性。

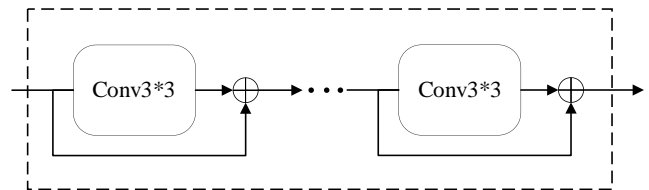


Fig. 3 Residual block

图3 残差块

的信息利用,达到优化网络精度的目的。多级特征组聚合块如图4所示,以第三层级特征图为例,其宽高分别为 $\frac{H}{4}$ 和 $\frac{W}{4}$,通道数为 $4C$ 。首先,通过双线性插值和逐点卷积,将网络中第一二层级的低维特征图和第四五层级的高维特征图的通道数和宽高,根据第三层级特征图的大小进行修改;其次,对特征图进行分块操作,将每个特征图按照通道维度平均分为4块,每块大小为 $\frac{H}{4} \times \frac{W}{4} \times C$;接着,为保证掩膜信息能够充分利用,将每个特征图与掩膜进行分组拼接,得到4个大小为 $\frac{H}{4} \times \frac{W}{4} \times 5C$ 的新子特征图,这4个子特征图中均包含有mask信息及各层级信息,并使用扩张率为 $\{1, 3, 5, 7\}$ 的空洞卷积分别对新特征图进行处理,以获得不同感受野的信息,增加多尺度信息获取能力^[30];而后,将所有特征图使用逐点卷积修改为分块的特征图大

小,并对其进行拼接,让特征图复原为第三层级特征图的大小;最后,由于多级特征组聚合模块是为了实现利用其他层级特征图丰富自身功能,为减少网络中每个层级本身

的信息损失,平衡网络对各维度信息的提取能力,使用残差连接的方式将原特征图与新特征图进行融合,从而获得最终结果。

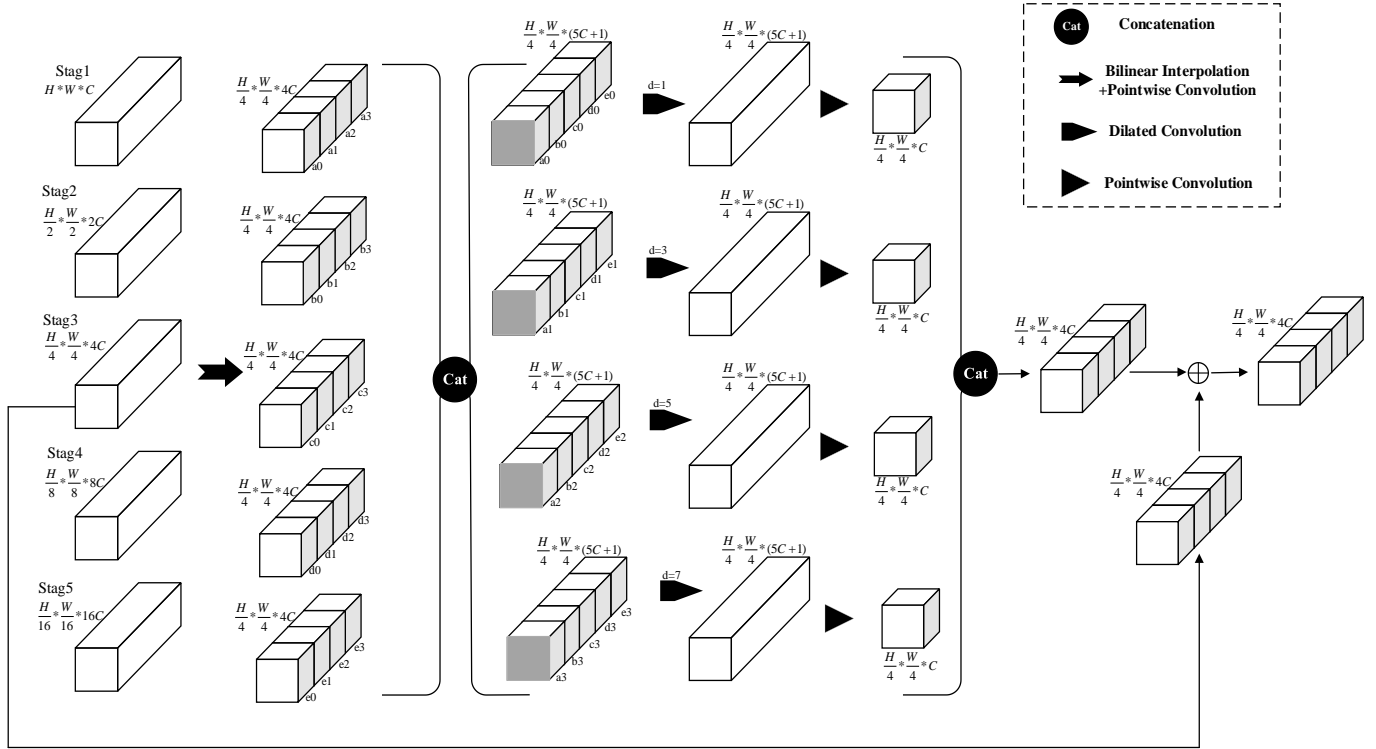


Fig. 4 Multi-level feature group aggregation block

图4 多级特征组聚合块

2.4 损失函数设计

本文多级特征组聚合块需要每层解码器输出的掩膜信息作为输入,故使用深度监督的方式以实现损失函数计算,深度监督方法是在网络的不同层级添加额外的损失函数,在高层级将部分分割结果和特征图信息进行输出,使网络能够在更早的阶段提供反馈信息,这种方法能让网络更快地学习有效的特征表示,加速收敛,提高泛化能力,从而改善分割精度。在解码器每层的损失函数计算中,本文使用由骰子(Dice)损失函数和二元交叉熵(BCE)损失函数的混合损失函数。

Dice 损失函数是基于 Dice 系数,主要用于评估两个样本的相似度,其公式如式(3)所示。

$$Dice_{loss} = 1 - \frac{2 \sum_i^N p_i y_i}{\sum_i^N p_i^2 + \sum_i^N y_i^2} \quad (3)$$

其中, $p_i \in (0, 1)$ 是模型对第 i 个样本的预测像素概率值, $y_i \in \{0, 1\}$ 是第 i 个样本的真实标签值,当预测像素在目标区域时才有效, N 为图片中像素总个数。

BCE 损失函数是用于二分类问题的常见损失函数,它衡量的是预测值和实际值之间的差异,其公式如式(4)所示。

$$BCE_{loss} = \frac{1}{N} \sum_{i=1}^N [y_i \log p_i + (1 - y_i) \log (1 - p_i)] \quad (4)$$

每层的损失函数计算是 Dice 损失函数值和 BCE 损失函数值分别乘以一个权重之和,其公式如式(5)所示。

$$L_i = \lambda_{dice} \times Dice_{loss} + \lambda_{BCE} \times BCE_{loss} \quad (5)$$

其中, L_i 是指第 i 层的混合损失函数值, λ_{dice} 和 λ_{BCE} 分别为 Dice 损失函数权重和 BCE 损失函数权重,实验内均设置为 0.5, i 为层数。

在损失模块中,将解码器中每一层都使用一个 1×1 的卷积修改通道数,生成每层的掩膜,然后参考 EGE-UNet 使用的双线性插值方式将每层掩膜扩充为原始掩膜大小,对扩充后的掩膜与真实掩膜求取损失。双线性插值是从像素点的水平和垂直方向上进行线性插值,从而得到一个像素点周围 4 个相邻像素点的加权平均,该插值方式能够减少图像中出现的锯齿状或块状伪影,并且能较好地保持图像的细节和质量,减少信息丢失。总损失函数的计算公式如式(6)所示。

$$Loss_{\text{总}} = \sum_{i=0}^n \lambda_i \times L_i \quad (6)$$

其中, λ_i 为第 i 层级的权重, L_i 为第 i 层级的混合损失函数,在实验中将 λ_i 设置为 $\{1, 0.5, 0.4, 0.3, 0.2\}$, i 的范围为 $\{0, 1, 2, 3, 4\}$ 。

3 实验与结果分析

3.1 数据集

实验中使用的数据集为 ISIC2017、ISIC2018 皮肤病数据集和 CVC-ClinicDB 息肉数据集^[31-32]。ISIC2017 数据集由国际皮肤成像协作组织在 2017 年的挑战赛中发布,专门用于皮肤病变的自动检测与分类研究,包含黑色素瘤、非黑色素瘤皮肤癌和良性病变等病例,共计 2 750 张图像及其对应的标注掩膜。ISIC2018 数据集由国际皮肤成像协作组织在 2018 年的挑战赛中发布,包含黑色素瘤、基底细胞癌、鳞状细胞癌等疾病,共计 3 694 张图像及其对应的标注掩膜。CVC-ClinicDB 数据集是 MICCAI 2015 结肠镜视频自动息肉检测子挑战赛训练阶段提供的数据集,该数据库由从结肠镜检查视频中提取的 612 张静态图像组成,每张图像都附带有一个掩膜,用于标识图像中息肉所覆盖的区域,通过对息肉进行分割,可以为肠镜检查 and 消化道肿瘤筛查提供重要的形态、位置信息,辅助医生进行检查治疗。在实验中,将 ISIC2017、ISIC2018 和 CVC-ClinicDB 数据集均按照 7:1.5:1.5 的比例分割成训练集、验证集和测试集。

3.2 实验环境

实验平台选用 Linux 系统,Python 为 3.8,PyTorch 深度学习框架版本为 1.10.0,CUDA 版本为 11.3,CPU 为 15 vCPU AMD EPYC 7543 32-Core Processor,系统内存大小为 80 GB,GPU 为 NVIDIA A40 图形处理器(显存 48 GB)。

3.3 训练设置

在模型训练配置方面,本文采用的 batch size 为 8,迭代次数设置为 1 000 次,为防止模型过拟合,引入了提前停止机制,当模型 dice 指标在 100 次迭代后仍未改进时停止训练。在优化算法上,选取 Adam 优化器,并采用余弦退火策略优化训练过程,初始学习率设置为 10^{-3} ,最小学习率设置为 10^{-4} 。

在训练中,统一将图像尺寸调整为 224×224 像素,为了提高模型泛化能力,还采用了随机翻转和随机旋转的数据增强方法。

3.4 评价指标

为评估模型在 ISIC2018 数据集上的分割效果,本文选择了两个核心的评价指标: Dice 系数和交并比(IoU)。Dice 系数是一种用于量化分割质量的图像分割指标,其主要衡量两个集合之间的相似度,在图像分割任务中用于比较分割结果与真实值之间的重合程度,其公式如式(7)所示。

$$Dice = \frac{2 \times |A \cap B|}{|A| + |B|} \quad (7)$$

其中,A代表模型预测区域,而B代表真实标签区域。Dice 系数值的范围为 0~1,值越接近 1,表示预测与真实之间的一致性越高。

交并比(IoU)是另一种图像分割评价指标,其考虑了面积因素,主要通过计算分割结果与真实值之间的重叠程度以评估模型准确性,其计算公式如式(8)所示。

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (8)$$

其中,A代表模型预测区域,而B代表真实标签区域,IoU值的范围为 0~1,当值为 1 时,表示与真实值相同。

3.5 皮肤病图像分割结果及分析

表 1 显示了本文方法与现有 6 种方法在 ISIC2018 数据集上的实验效果比较情况。其中,UNet 使用的是 ACC-UNet 中提供的低参数量 UNet。其他方法中,Swin-UNet、UCTransNet、SMESwin-UNet 为基于 Transformer 改进的 U 型网络,MultiResUNet、ACC-UNet 以及本文提出的 MFGA-UNet 为基于卷积改进的 U 型网络。在参数量上,MFGA-UNet 仅为 10.246 M,优于基于 Transformer 改进的网络,并且相较于大部分基于卷积的网络仍具有优势。在精度上,本文网络的 Dice 和 IoU 指标优于其他网络,分别达到了 90.028%、83.595%。

Table 1 ISIC2018 dataset results

表 1 ISIC2018 数据集结果

模型	Params/M	Flops/G	Dice/%	IoU/%
UNet	14.752	25.185	88.862	81.725
MultiResUNet	7.251	17.359	89.067	82.153
Swin-UNet	27.146	5.911	88.724	81.664
SMESwin-UNet	169.768	6.240	88.606	81.625
UCTransNet	66.242	32.984	89.856	83.329
ACC-UNet	16.771	45.337	89.778	83.298
MFGA-UNet	10.246	20.807	90.028	83.641

图 5 展示了各模型在 ISIC2018 数据集上的部分可视化分割结果,共包含 3 个,从上至下分割难度逐渐递增。每张图像底部标有图片(Image)、真实值(GT)或者模型名,每个模型名下部括号中的数字代表其分割的 Dice 指标值。可以看到,相较于其他网络,本文 MFGA-UNet 对于边界区域的分割更加平滑,并且在部分浅颜色病变样本中表现出了更好的效果

为了证明模型在皮肤病分割任务中的有效性和泛化效果,在 ISIC2017 数据集上采用相同的环境、代码与训练方式进行补充实验,实验结果如表 2 所示。可以看出,该模型在 ISIC2017 数据集上的精度指标仍优于其他模型,Dice 和 IoU 分别达到了 94.308%、90.262%。

在 ISIC2017 数据集中,所有网络都可取得较为有效的病变结果,在部分数据上,本文方法对边界分割得更加精确,可视化结果如图 6 所示。

为证明本文方法在其他数据集上的泛化性,在 CVC-ClinicDB 数据集上进行了实验,实验结果如表 3 所示。实验结果表明,在息肉分割数据集中,该模型的精度指标优于除 UCTransNet 以外的 5 种医学图像分割网络,Dice 和 IoU 分别达到了 91.436%、85.603%,并且在参数量上远低于基于 Transformer 改进的图像分割模型。

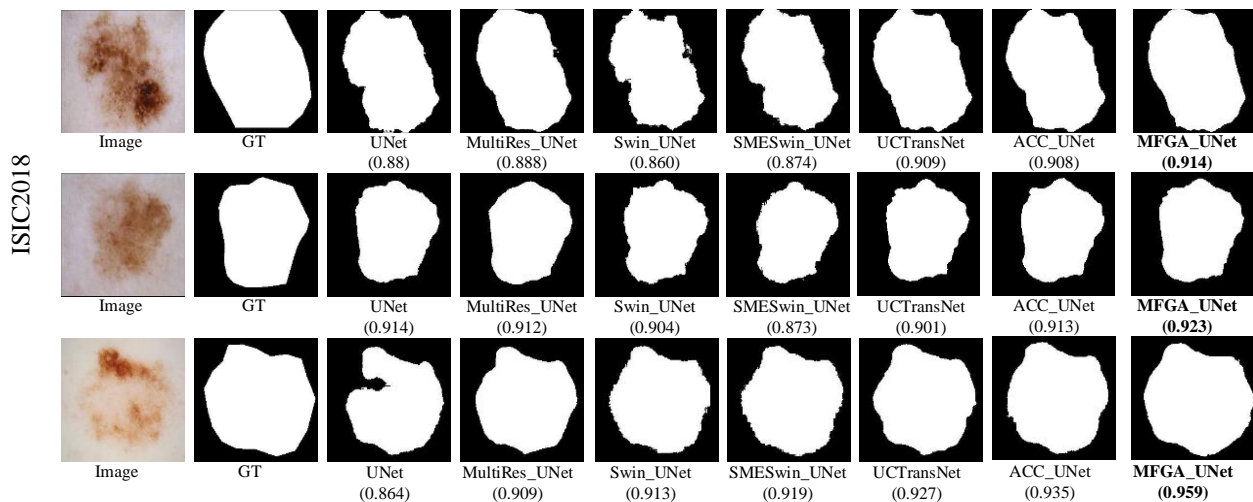


Fig. 5 Experimental result

图 5 实验结果

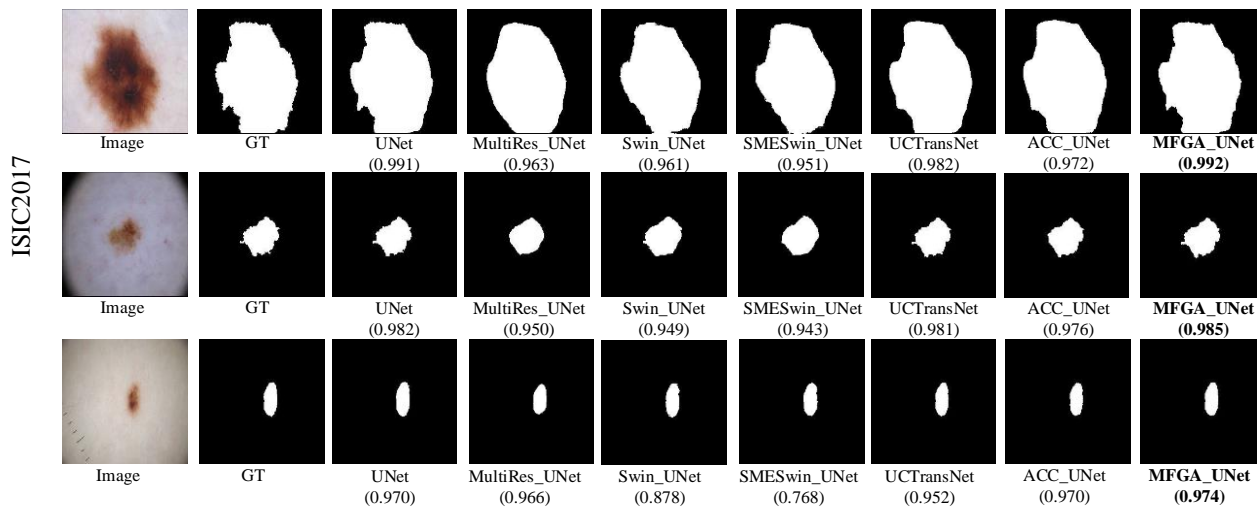


Fig. 6 Experimental result

图 6 实验结果

Table 2 ISIC2017 dataset results

表 2 ISIC2017数据集结果

模型	Params/M	Flops/G	Dice/%	IoU/%
UNet	14.752	25.185	93.654	89.238
MultiResUNet	7.251	17.359	90.885	84.241
Swin-UNet	27.146	5.911	91.675	85.363
SMESwin-UNet	169.768	6.240	91.291	84.891
UCTransNet	66.242	32.984	94.220	89.877
ACC-UNet	16.771	45.337	93.195	88.290
MFGA-UNet	10.246	20.807	94.308	90.262

在 CVC-ClinicDB 数据集中,对于一些较为明亮且与周围颜色相近的病变组织上,本文网络略优于其他网络,可视化结果如图 7 所示。

3.6 消融实验结果及分析

为了体现本文网络各结构改进的有效性,对卷积块、跳跃连接、损失函数分别进行了消融实验,如表 4 所示。选用通道数为 32 的 UNet 网络作为基准网络(Baseline),并通过 3 个消融实验探究多级特征组聚合模块、反向瓶颈卷

Table 3 CVC-ClinicDB dataset results

表 3 CVC-ClinicDB数据集结果

模型	Params/M	Flops/G	Dice/%	IoU/%
UNet	14.752	25.185	90.766	84.475
MultiResUNet	7.251	17.359	84.279	77.731
Swin-UNet	27.146	5.911	75.775	66.086
SMESwin-UNet	169.768	6.240	75.810	65.752
UCTransNet	66.242	32.984	91.804	85.710
ACC-UNet	16.771	45.337	91.110	85.260
MFGA-UNet	10.246	20.807	91.436	85.603

积和深度监督损失函数对模型分割精度与参数数量的影响。第 1 个消融实验中,将多级特征组聚合模块添加到跳跃连接中,提高了网络多尺度信息的利用,并且合理地减少了语义间隙,实现了精度优化。第 2 个消融实验中,使用反向瓶颈模块替换原有卷积块,提高了网络对特征的提取能力,实现了细微的精度提升,并保证了参数数量不会明显增加。第 3 个消融实验中,添加了深度监督方式对网络进行优化,通过对每层解码器的特征图求取损失,改善了模型

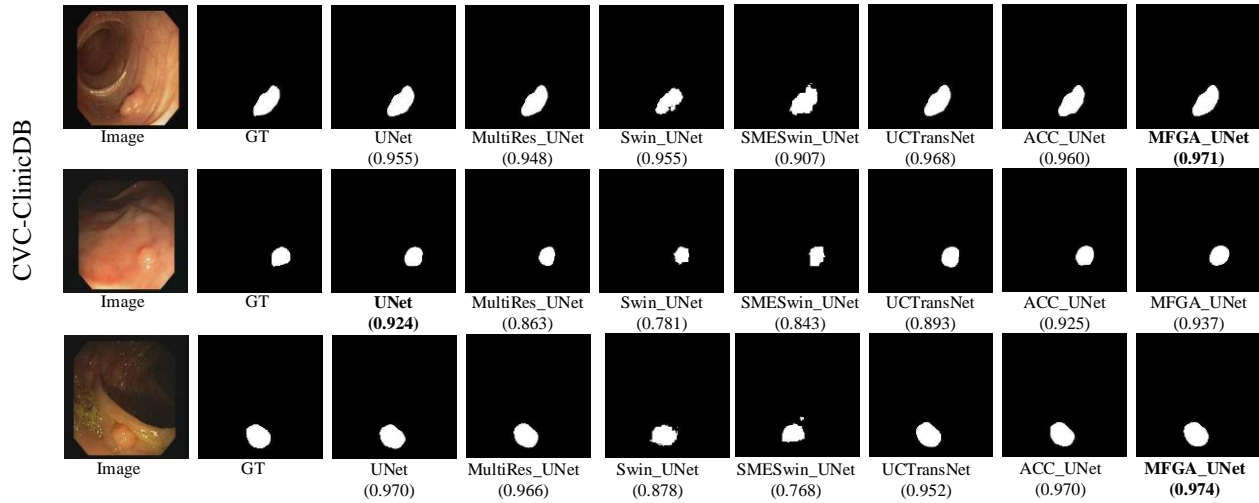


Fig. 7 CVC-ClinicDB experimental result

图7 CVC-ClinicDB实验结果

训练过程,实现了精度提升。实验结果表明,3个部分的模块均可提高网络精度。

Table 4 Ablation experiment

表4 消融实验

模型	Dice/%	IoU/%	Params/M
Baseline	89.175	82.344	7.765
Baseline+MFGA	89.716	83.070	9.361
Baseline+IBConv	89.301	82.568	8.651
Baseline+Loss	89.663	83.004	7.766

4 结语

本文提出了一种基于多级特征组聚合的皮肤病U型分割网络,旨在解决现阶段网络对多级特征与掩膜信息利用不充分的问题,并且减少模型部署成本。相较于其他网络,本文网络考虑到了更多的多尺度信息,在编码器中对每个层级的信息进行融合,丰富了每个层级的特征图信息,并添加了每层解码器的掩膜数据,实现了对多级特征与掩膜信息的组合利用,使网络能更全面地学习多尺度信息。在皮肤病治疗领域,本文提出的分割网络能更精确地实现病变区域提取分割,为后续病变区域面积测量等工作提供支持,以更加有效地辅助医护人员进行疾病治疗,可在一定程度上改善现阶段皮肤病诊疗问题。

在ISIC2017、ISIC2018和CVC-ClinicDB数据集上进行实验,结果表明:在ISIC2017、ISIC2018皮肤病数据集上,Dice指标分别达到了94.308%、90.028%,均优于现有6种分割网络,表明网络在皮肤病数据集上精度较高,能够更有效地实现皮肤病分割;在CVC-ClinicDB数据集上,Dice指标达到了91.436%,精度优于现有的5种网络,表明本文网络在息肉数据上具有一定的泛化性;在ISIC2018数据集上的消融实验结果表明,网络使用的卷积块、跳跃连接、损失函数均能实现精度优化,改进的反向瓶颈卷积块能够改善特征提取能力,多级特征组聚合模块能够提升多尺度信

息学习能力,使用的深度监督损失函数能够优化模型训练过程。虽然基于多级特征组聚合的皮肤病U型分割网络能够提升精度,但是在显存占用上还有优化空间。因此,在保证精度的前提下,可考虑在解码器中使用普通卷积进行特征重建等方式,以合理地减小显存占用,这可作为后续研究方向。

参考文献:

- [1] MATHUR P, SATHISHKUMAR K, CHATURVEDI M, et al. Cancer statistics, 2020: report from national cancer registry programme, India[J]. JCO Global Oncology, 2020, 6(6): 1063-1075.
- [2] GHAZAWI F M, CYR J, DARWICH R, et al. Cutaneous malignant melanoma incidence and mortality trends in Canada: a comprehensive population-based study[J]. Journal of the American Academy of Dermatology, 2019, 80(2): 448-459.
- [3] ARNOLD M, SINGH D, LAVERSANNE M, et al. Global burden of cutaneous melanoma in 2020 and projections to 2040[J]. JAMA Dermatology, 2022, 158(5): 495-503.
- [4] SHEN X L, WANG K P. A Retinal vessel image segmentation method incorporating dual attention mechanism and U-Net model[J]. Software Guide, 2024, 23(6): 150-156.
沈学利,王昆蓬.融合双注意力机制与U-Net模型的视网膜血管图像分割方法[J].软件导刊,2024,23(6):150-156.
- [5] ZHOU W Q, WANG C L, SUN Z Q, et al. Segmentation of neuroblastoma with PET-CT multimodal images based on tumor shape characteristics and point clouds method[J]. Software Guide, 2024, 23(3): 128-133.
周维钦,王朝立,孙占全,等.基于肿瘤形状特征与点云方法的PET-CT多模态图像神经母细胞瘤分割[J].软件导刊,2024,23(3):128-133.
- [6] QU X B, YU S. Retinal blood vessel segmentation algorithm based on generating adversarial network[J]. Software Guide, 2023, 22(8): 209-215.
曲小波,余粟.基于生成对抗网络的视网膜血管分割算法[J].软件导刊,2023,22(8):209-215.
- [7] XU W, TANG J W, ZHANG C. Image segmentation method of liver cancer based on RA-U-Net++ network[J]. Software Guide, 2023, 22(8): 203-208.
徐微,汤俊伟,张驰.基于RA-U-Net++的肝癌图像分割方法[J].软件导刊,2023,22(8):203-208.

- [8] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]//Proceedings of 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015: 234-241.
- [9] XU G X, FENG C, MA F. Review of medical image segmentation based on UNet [J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(8): 1776-1792.
徐光宪,冯春,马飞. 基于UNet的医学图像分割综述[J]. 计算机科学与探索, 2023, 17(8): 1776-1792.
- [10] FU L Y, YIN M X, YANG F. Transformer based U-shaped medical image segmentation network: a survey [J]. Journal of Computer Applications, 2023, 43(5): 1584-1595.
傅励瑶,尹梦晓,杨锋. 基于Transformer的U型医学图像分割网络综述[J]. 计算机应用, 2023, 43(5): 1584-1595.
- [11] SONG J, LIU C X, LI H T. Review of medical image segmentation based on U-Net network [J]. Computer Technology and Development, 2024, 34(1): 9-16.
宋杰,刘彩霞,李慧婷. 基于U-Net网络的医学图像分割研究综述[J]. 计算机技术与发展, 2024, 34(1): 9-16.
- [12] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 6000-6010.
- [13] IBTEHAZ N, KIHARA D. ACC-UNET: a completely convolutional UNet model for the 2020s [C]//Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention, 2023: 692-702.
- [14] RUAN J C, XIE M Y, GAO J S, et al. EGE-UNet: an efficient group enhanced UNet for skin lesion segmentation [C]//Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention, 2023: 481-490.
- [15] ZHOU Z, SIDDIQUEE M M R, TAJBAKSH 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.
- [16] IBTEHAZ N, SOHEL R M. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation [J]. Neural Networks, 2020, 121: 74-87.
- [17] VALANARASU J M J, PATEL V M. UNeXt: MLP-based rapid medical image segmentation network [C]//Proceedings of 25th International Conference on Medical Image Computing and Computer Assisted Intervention, 2022: 23-33.
- [18] KUSHNUR D T, TALBAR S N. HFRU-Net: high-level feature fusion and recalibration UNet for automatic liver and tumor segmentation in CT images [J]. Computer Methods and Programs in Biomedicine, 2022, 213: 106501.
- [19] ZHAO W Y, DING R, LI Y, et al. A multi-scale fusion segmentation method for MRI brain tumor images [J]. Software Guide, 2022, 21(8): 159-163.
赵惟玉,丁然,李阳,等. 多尺度融合的MRI脑肿瘤图像分割方法[J]. 软件导刊, 2022, 21(8): 159-163.
- [20] ZHONG J W. PRA-UNet3+: full-scale connected CT liver image segmentation model [J]. Software Guide, 2023, 22(2): 15-20.
钟经纬. PRA-UNet3+: 全尺度跳跃连接CT肝脏图像分割模型[J]. 软件导刊, 2023, 22(2): 15-20.
- [21] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale [DB/OL]. <https://arxiv.org/abs/2010.11929>, 2020.
- [22] CHEN J, LU Y, YU Q, et al. Transunet: transformers make strong encoders for medical image segmentation [DB/OL]. <https://arxiv.org/abs/2102.04306>, 2021.
- [23] 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, 2022: 205-218.
- [24] WANG H, CAO P, WANG J, et al. Utransnet: rethinking the skip connections in U-Net from a channel-wise perspective with transformer [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2022: 2441-2449.
- [25] JI Y, ZHANG R, WANG H, et al. Multi-compound transformer for accurate biomedical image segmentation [C]//Medical Image Computing and Computer Assisted Intervention, 2021: 326-336.
- [26] WANG Z, MIN X, SHI F, et al. SMESwin Unet: merging CNN and transformer for medical image segmentation [C]//Proceedings of International Conference on Medical Image Computing and Computer Assisted Intervention, 2022: 517-526.
- [27] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [28] LIU Z, MAO H Z, WU C Y, et al. A convnet for the 2020s [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11966-11976.
- [29] CHOLLET F. Xception: deep learning with depthwise separable convolutions [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1800-1807.
- [30] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [DB/OL]. <https://arxiv.org/abs/1511.07122>, 2015.
- [31] BERSETH M. ISIC 2017 - skin lesion analysis towards melanoma detection [DB/OL]. <https://arxiv.org/abs/1703.00523>, 2017.
- [32] CODELLA N, ROTEMBERG V, TSCHANDL P, et al. Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (ISIC) [DB/OL]. <https://arxiv.org/abs/1902.03368>, 2019.

(责任编辑:孙娟)