

DOI:10.19651/j.cnki.emt.2519021

# LMSCI-Net:轻量级多尺度信道交互皮肤病灶 分割网络\*

王斯豪 张笃振 杨昌昌

(江苏师范大学计算机科学与技术学院 徐州 221116)

**摘要:** 准确的病灶分割结果对于皮肤病早期诊断和后续治疗具有重要意义。现有神经网络倾向于设计更深更复杂的结构达到高分割精度的目的,然而较高的模型参数和计算成本限制了实际应用,针对该问题,提出一种轻量级多尺度信道交互分割网络 LMSCI-Net,对于输入图像,设计基于信道分离和卷积分解的轻量级多尺度编码模块,结合局部-全局通道注意力机制,在保证特征提取能力的同时,实现轻量化网络编码器设计;其次,设计多尺度信道交互增强模块,集成网络多阶段输出信息优化跳跃连接的信息处理过程,为解码器提供丰富且精确的细节信息;最后,在解码器部分设计自适应融合解码模块实现更高效的细节恢复,最终输出精确的分割结果。该网络训练过程采用深度监督机制,在 ISIC2017、ISIC2018、ISIC2016 这三个公开皮肤病灶分割数据集上进行对比实验并在 PH2 数据集上进行泛化实验,实验表明该网络在实现轻量化的同时保持较高的分割精度与泛化能力,与基线网络 U-Net 相比在参数量和计算复杂度上分别降低了 99.38% 和 98.78%,充分验证该网络的有效性与轻量化。

**关键词:** 皮肤病灶分割;轻量化;注意力机制;多尺度特征融合

**中图分类号:** TP391;TN911 **文献标识码:** A **国家标准学科分类代码:** 520.6040

## LMSCI-Net: Lightweight multiscale channel interactive skin lesion segmentation network

Wang Sihao Zhang Duzhen Yang Changchang

(School of Computer Science and Technology, Jiangsu Normal University, Xuzhou 221116, China)

**Abstract:** Accurate lesion segmentation is crucial for early diagnosis and subsequent treatment of dermatological diseases. Existing neural networks often employ increasingly deep and complex architectures to achieve high segmentation accuracy; however, large parameter counts and high computational costs limit practical deployment. To address these challenges, a lightweight multi-scale channel interaction segmentation network (LMSCI-Net) is proposed. For each input image, a lightweight multi-scale encoding module based on channel separation and convolutional decomposition is designed, augmented by a local-global channel attention mechanism to ensure robust feature extraction while maintaining an efficient encoder. A multi-scale channel interaction enhancement module is then introduced to integrate multi-stage outputs and refine skip connections, providing the decoder with rich and precise detail information. Finally, an adaptive fusion decoding module is developed to progressively restore fine-grained details and produce accurate segmentation masks. The network is trained under a deep supervision regime and evaluated on three public skin-lesion segmentation datasets (ISIC2017, ISIC2018 and ISIC2016) as well as the PH2 dermoscopic image database. Experimental results demonstrate that, compared with the U-Net baseline, LMSCI-Net reduces parameter count and computational complexity by 99.38% and 98.78%, respectively, while maintaining high segmentation accuracy and strong generalization, thus validating its effectiveness and lightweight design.

**Keywords:** skin lesion segmentation;lightweighting;attentional mechanisms;multiscale feature fusion

## 0 引言

皮肤癌是全球范围内发病率最高的癌症类型之一,据

统计<sup>[1]</sup>每年新增确诊病例超过 150 万例,且其致死率不容忽视。在临床实践中,早期诊断和及时治疗是提高皮肤癌治愈率的关键因素<sup>[2]</sup>,目前,皮肤镜成像技术因其非介入、

收稿日期:2025-06-05

\* 基金项目:江苏省高等学校自然科学研究面上项目(19KJB520032)、江苏师范大学博士学位教师科研支持项目(20XSRS018)资助

无创的特性,成为早期筛查的主要手段,该技术能够提供皮肤表面的实时图像。然而,这些图像往往存在对比度不足、病灶边界模糊、光照不均匀以及噪声干扰等技术性缺陷<sup>[3]</sup>,加之医生在诊断过程中可能受到主观经验差异和视觉疲劳的影响,导致诊断结果存在不一致性。为了克服这些局限性,基于计算机辅助诊断的自动分割方法应运而生,然而这类技术早期依赖亮度、纹理、边缘和区域连通性这些低级特征,易受噪声和毛发干扰,分割效果难以满足临床需求<sup>[4]</sup>。

近年来,基于深度学习技术的医学图像分割算法展现出了超越传统图像处理技术的性能和潜力,卷积神经网络(convolutional neural network,CNN)<sup>[5]</sup>能够实现图像特征的自动获取从而取代了传统手工标注,2015年Long等<sup>[6]</sup>首次提出了全卷积神经网络(fully convolutional network,FCN)的概念,实现了像素级别的图像分类也就是图像分割,同年,Ronneberger等<sup>[7]</sup>提出了具有跳跃连接的U型编解码器结构网络U-Net,能通过卷积和下采样获取高级特征表示,并将不同层级的信息融合从而实现高水准的医学图像分割效果。为了进一步提高分割性能,一些研究聚焦于U型网络的变体,比如引入注意力机制提高对于重要信息通道和区域的关注度<sup>[8]</sup>,引入金字塔池化捕获多尺度上下文信息<sup>[9]</sup>,与Transformer架构结合以提高网络的全局信息处理能力等。

然而现有研究仍然存在局限,首先,现有网络偏向于设计复杂的编解码器结构以获取更高的特征提取能力<sup>[10]</sup>,比如TransUNet<sup>[11]</sup>、SETR<sup>[12]</sup>、CSWinTransformer<sup>[13]</sup>这类结合Transformer架构的网络,虽然提高了全局特征捕捉能力,但显著增加了模型参数量和计算复杂度。其次,编码器的浅层细节信息对于解码器恢复图像细节至关重要,现有网络缺乏对于多层编码输出的探索,导致模型在应对大小不一、形态各异的病灶目标时表现受限。另外,还有部分网络在处理多尺度信息时未经过充分的交互,导致部分对分割任务至关重要的通道信息可能被弱化,而噪声信息则难以被有效抑制。此外,现有网络还缺乏对编解码器低层与解码器高层间之间存在语义鸿沟问题的考虑,直接将不同语义深度的特征进行简单融合会影响解码器对病灶区域的准确恢复,导致次优的分割结果。

为解决上述问题,本文提出一种用于皮肤病灶分割的轻量级多尺度信道交互分割网络(lightweight multi-scale channel interaction segmentation network,LMSCI-Net),主要贡献如下:

1)设计局部-全局通道注意力机制(local-global feature attention,LGFA),在空间维度切片特征图后单独计算局部信息描述符,最后在全局信息描述符中引入局部信息描述符,以弥补获取通道级统计数据时局部信息的缺失,使得模型能够更好地捕捉通道之间的复杂关系,得到更加精确的通道注意力权重。

2)设计轻量级多尺度编码模块(lightweight multi-

scale encoder,LMSE),采用增加网络宽度代替传统增加深度的设计方式,基于信道分离和空洞卷积实现多尺度信息捕捉,并结合局部-全局通道注意力机制LGFA卷积分解保证特征提取能力的同时,实现轻量化网络编码器设计。

3)设计多尺度信道交互增强模块(multi-scale interaction enhancer,MSIE),基于分组卷积和通道混洗策略实现多尺度信息的高效交互,充分挖掘各尺度特征之间的互补性,并利用局部-全局通道注意力机制LGFA突出关键信道,抑制冗余和噪声信息。

4)设计自适应加权解码模块AdaptiveFusion。基于局部-全局通道注意力机制LGFA计算融合权重,使网络能动态学习和调整浅层和深层特征图的权重实现有效的信息融合完成细节恢复。采用深度监督策略监督模型训练,提高模型训练效率并促进深层网络的优化。

## 1 相关工作

具有跳跃连接的对称编解码器结构U-Net在皮肤病灶分割任务上效果突出,但容易受到噪声和伪影信息影响而降低分割精度。受U-Net的启发,各种基于U-Net的变体被提出。例如:2018年,Zhou等<sup>[14]</sup>提出的Unet++,设计深层网络,并引入多层跳跃连接设计改进了特征融合,但在处理模糊边界和复杂病灶形状时会出现精度不足的问题。同年Oktay等<sup>[15]</sup>提出Attention U-Net引入注意力机制显著提升了对复杂区域的分割能力,但模型可能会忽略微小的病灶导致精度降低,并且在数据不足的情况下会过度依赖少量样本中强调的特征导致泛化性减弱。2021年,Liu等<sup>[16]</sup>提出的Swin-Transformer可作为各类视觉任务的通用骨干网络,采用移动窗口多头注意力实现了相邻局域信息交互,在多尺度特征的提取和区域分割精度上表现突出,但其复杂度较高使得计算和内存开销限制了其实际应用。同年,TransUNet在原本的U-Net瓶颈位置引入了Transformer块,简单集成两种架构的优势改善了分割效果,但对于一些复杂的边界细节和精确定位任务,可能仍然不如纯粹的卷积网络表现得那样精确。梁礼明等<sup>[17]</sup>提出基于粗分和细分U型网络的双U型皮肤病灶分割网络,分别捕捉细粒度信息和粗粒度信息用以改善最终病灶分割精度,但存在模型复杂性较高且训练时间和资源需求较大的问题。2022年,Ruan等<sup>[18]</sup>提出了MALUNet模型,通过集成多种注意力机制以及轻量级设计有效平衡了参数量和分割性能。2023年,Ruan等<sup>[19]</sup>提出了轻量级的皮肤病灶分割模型EGE-UNet,通过引入组多轴Hadamard积注意力模块GHPA和组聚合桥接模块GAB,显著降低了模型的参数量和计算成本。同年,Tang等<sup>[20]</sup>提出了轻量化的医学图像分割网络CMUNeXt,旨在通过大核卷积和倒置瓶颈设计有效提取全局上下文信息,并引入跳跃融合块以增强特征融合。2024年,Wu等<sup>[21]</sup>提出MHorU-Net模型,



通道的空间特征信息,得到大小为  $S \in R^{1 \times 1 \times C}$  的特征图,接着传入由两个全连接层组成的瓶颈结构中,该瓶颈结构在保持了较高的信息表达能力的同时降低了参数量和计算量,符合轻量化模块的设计,第一个全连接层沿着通道压缩实现降维并经过 *ReLU* 非线性激活层,接着由第二个全连接层将通道数恢复至原通道数并通过 *Sigmoid* 激活,最终得到精炼的局部区域描述符,随后将各分支得到的局部区域描述符沿通道维度拼接得到大小为  $S \in R^{1 \times 1 \times C}$  的局部总信息描述符。在底层路径,对输入图像  $S$  进行全局平均池化得到大小为  $S \in R^{1 \times 1 \times C}$  的全局信息描述符,并通过通道维度拼接的方式引入中间路径得到的局部总描述符生成更丰富的通道级统计数据,以弥补全局平均池化所造成的局部信息损失,接着采用与上文完全一致的瓶颈结构,通过降维、

激活和升维操作生成精确的通道注意力权重,最终利用得到的通道注意力权重实现对原始输入特征图的加权。该局部-全局通道注意力计算过程可用式(1)~(4)表示:

$$F_i = \text{BM}(\text{GAP}(S_i)), i = 1, 2, 3, 4 \quad (1)$$

$$F_{\text{local}} = \text{Concat}(F_1, F_2, F_3, F_4) \quad (2)$$

$$F_{\text{global}} = \text{GAP}(S) \quad (3)$$

$$F_{\text{att}} = \text{BM}(\text{Concat}(F_{\text{global}}, F_{\text{local}})) \quad (4)$$

其中,  $S_i$  表示输入图像  $S$  经过空间等分后得到的子区域特征图,  $\text{GAP}$  表示全局平均池化,  $\text{BM}$  表示瓶颈模块处理,  $\text{Concat}$  表示张量拼接,这里均是在通道维度进行拼接。该局部-全局通道注意力弥补了获取通道级统计数据时局部信息的缺失,使得模型能够更好地捕捉通道之间的复杂关系,进一步缓解通道依赖性,得到更加精确的通道注意力权重。

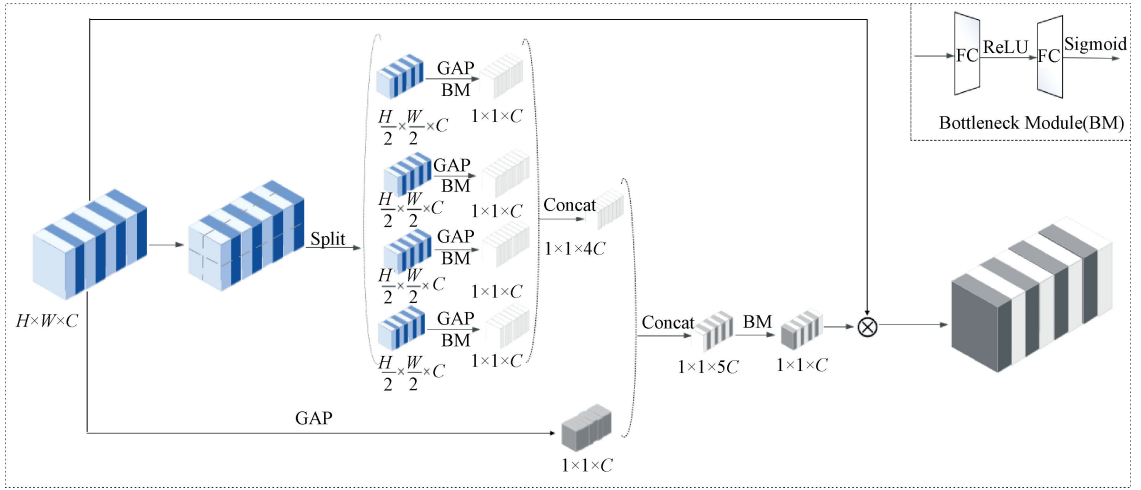


图2 局部-全局通道注意力机制

Fig. 2 Local-global channel attention mechanism

### 2.3 基于卷积分解的轻量级多尺度编码模块 LMSE

现有网络偏向于采用更深更复杂的网络结构来达到更高性能的目的,然而难以应用于存储和计算资源有限的临床环境,更深的网络结构往往伴随着更多的下采样,在连续下采样的过程中会导致病灶图像细节的大量丢失。为了应对上述问题,在编码器部分,本文设计轻量化的多尺度编码模块 LMSE 完成病灶特征提取过程,采用增加网络宽度代替传统增加深度的设计方式,在保证特征提取能力的同时,实现轻量化网络编码器的设计。

如图3所示,对于输入图像  $X \in R^{H \times W \times C}$ ,首先采用顺序堆叠的局部-全局通道注意力模块 LGFA 和空间注意力模块 SA, LGFA 模块通过结合特征图局部和全局信息自适应调整各通道的权重,抑制冗余和噪声信息,加强病灶边界、纹理或形态等关键信息的通道,提升模型对于不同病灶的识别能力以及分割性能。SA 模块结构参照文献[8]提出的 CBAM 中的空间注意力部分,将特征图沿着通道轴进行最大池化和平均池化操作,产生两个大小为  $H \times W \times 1$  的注意力图后将它们在通道维度拼接,然后经

过  $7 \times 7$  的卷积操作与 *Sigmoid* 激活函数获取特征间的空间关系,学习与病灶区域相关的关键区域信息。随后将得到的病灶特征图沿通道维度分割成4个部分并进入并行多分支卷积结构中,每个分支内采用不同空洞率的卷积,在不增加额外参数量的同时捕获多尺度信息。同时,分支内将传统的  $3 \times 3$  卷积分解为两个连续的条状卷积操作来实现卷积分解,其中先使用一个  $3 \times 1$  的卷积提取垂直方向上的特征,再使用  $1 \times 3$  卷积提取水平方向上的特征,从而在保证相似感受野的前提下进一步减少参数量和计算量,在最后引入 LGFA 注意力模块和 SA 模块进一步自适应地调整特征的重要性,使得模型在计算效率和性能之间取得更好的平衡。将多分支的结果在通道维度拼接后恢复原通道维度,由于在信道分离的过程中,不同组内的卷积操作往往会使得信息局限于组内,为实现跨组的信息交换,采用通道混合策略<sup>[26]</sup>打乱通道顺序,重新分配和混合通道信息,增强不同特征之间的互补性,使得模型对复杂特征的表达更加充分,最后将结果与原始输入做残差相加,保持原始信息的流动,提高模型的鲁棒性。上述过程

可用式(5)~(7)表示:

$$\mathbf{F} = \text{SA}(\text{LGFA}(\mathbf{X})) \quad (5)$$

$$\mathbf{F}' = \text{ChannelShuffle}(\text{Concat}_{i=1}^4$$

$$(\text{SA}(\text{LGFA}(\text{Conv}_{1 \times 3}^{d_i}(\text{Conv}_{3 \times 1}^{d_i}(\mathbf{F})^{(d)})))) \quad (6)$$

$$\mathbf{F}'' = \mathbf{F}' \oplus \text{Conv}_{1 \times 1}(\mathbf{X}) \quad (7)$$

其中,  $\mathbf{X}$  表示输入图像,  $\text{Concat}$  表示在通道维度进行张量拼接,  $\mathbf{X}_1 \in R^{H \times W \times 8}$  表示分成的 4 个分支,  $\mathbf{X}_i \in R^{H \times W \times 8}$  表示每个分支的空洞率,  $\text{ChannelShuffle}$  表示通道混合操作。

在模型编码器部分包含 5 个阶段, 对于给定编码器输入  $\mathbf{X}_1 \in R^{H \times W \times 8}$ , 经过第 1 阶段的浅层普通卷积块和 BatchNorm 后输出  $\mathbf{X}_1 \in R^{H \times W \times 8}$ , 然后经过堆叠的 4 块 LMSE 模块, 每阶段间通过 DownSample 模块进行下采样, 中间输出分别用  $X_2, X_3, X_4, X_5$  表示, 大小分别为  $\frac{H}{2} \times \frac{W}{2} \times 16, \frac{H}{4} \times \frac{W}{4} \times 32, \frac{H}{8} \times \frac{W}{8} \times 48, \frac{H}{16} \times \frac{W}{16} \times 64$ , 将各阶段中间输出传入跳跃连接, 并将  $X_5$  作为编码器的输出传递到解码器部分。

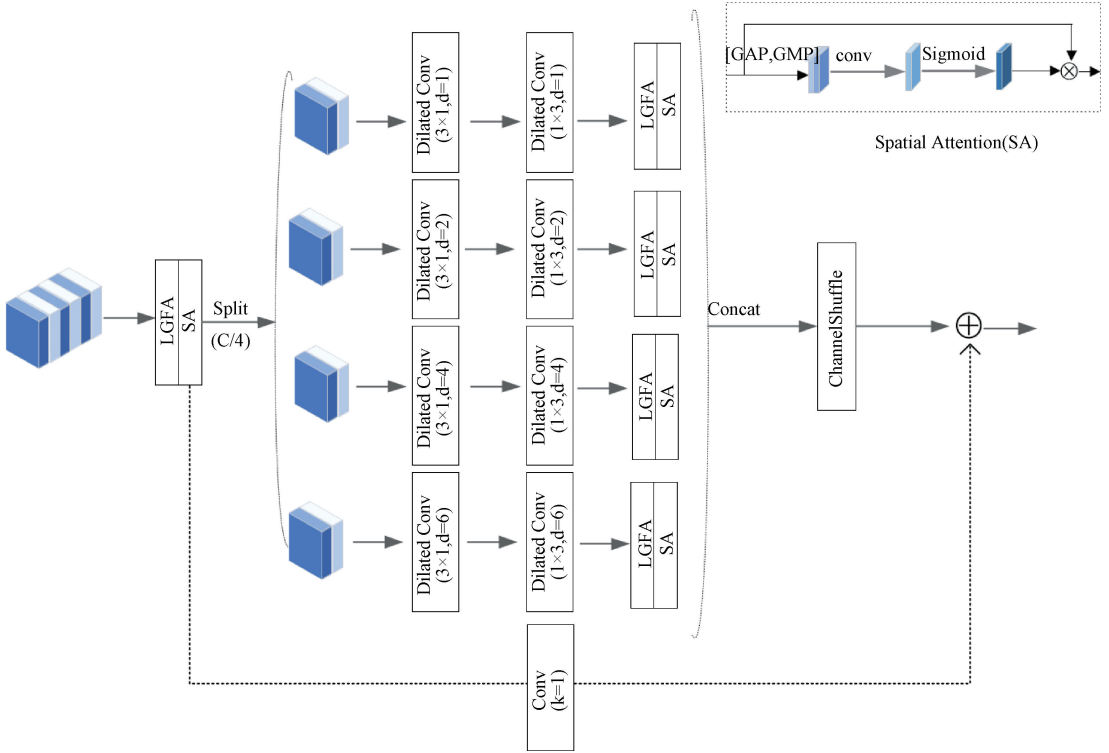


图 3 轻量级多尺度编码模块

Fig. 3 Lightweight multi-scale coding module

### 2.4 轻量级多尺度信道交互增强模块 MSIE

在编码器下采样过程中, 图像分辨率逐阶段降低会丢失部分病灶空间细节信息, 解码部分单纯依靠同阶段的跳跃连接传递无法集成更丰富的细节信息, 导致难以恢复高质量的病灶图像细节, 如果采用多个阶段编码层信息融合的方式能缓解这一问题, 但又会导致计算复杂度和参数数量的增加, 并且在不同阶段的特征图间还存在语义差异以及噪声干扰的问题。因此, 本文提出了一种高效轻量的多尺度特征交互增强模块 MSIE, 用于在当前阶段跳跃连接传递的特征图中引入其余阶段的信息, 基于分组卷积、通道混合机制以及 LGFA 通道注意力机制实现高效的特征交互与融合, 在集成多阶段的信息增强当前阶段的输出, 为当前阶段解码器提供所需要的丰富且精确的细节信息, 并且不同阶段特征中的信息互为补充, 使模型在处理复杂背

景或病灶区域的分割任务中表现得更加鲁棒。

具体如图 4 所示, 当前阶段传递的特征图大小为  $\frac{H}{2} \times \frac{W}{2} \times 16$ , 将其余阶段大小为  $H \times W \times 8, \frac{H}{4} \times \frac{W}{4} \times 32, \frac{H}{8} \times \frac{W}{8} \times 48, \frac{H}{16} \times \frac{W}{16} \times 64$  的特征图与当前阶段特征图一起传递进该模块, 除当前阶段, 其余特征图分别通过和 BI 层将特征图通道数和尺寸调整为  $\frac{H}{2} \times \frac{W}{2} \times 16$ , 随后为减少参数量, 将所有特征图通道数压缩为原来的一半, 然后在通道维度拼接, 随后采用通道混合操作, 初步整合来自各层的信息, 为后续更深层次的特征交互提供融合后的特征表达, 有助于平衡不同语义层次的信息, 然后采用分组卷积和 LGFA 通道注意力机制, 在初步融合的特征上采用分组

卷积,利用分组策略降低计算量的同时局部提取特征;再结合 LGFA 通道注意力机制,对通道上的信息进行动态加权,突出重要特征,抑制噪声,然后再通过一层通道合并并采用  $\text{Conv}_{1 \times 1}$  将通道调整为 16,再通过 BN 和 ReLU 层,帮助缓解梯度消失,提高训练稳定性,并为后续的特征恢复提供非线性增强,最后通过残差加入当前阶段的原始特征图,补充融合过程中的部分信息损失,进一步提升输出特征的准确性和鲁棒性。以上流程可用式(8)~(12)表示:

$$F_i = BI(\text{Conv}_{1 \times 1}(F_i; C)), i \in [1, 3, 4, 5] \quad (8)$$

$$F_i = \text{Conv}_{1 \times 1}(F_i; \frac{C}{\alpha}), i \in [1, 2, 3, 4, 5] \quad (9)$$

$$F_{cat} = \text{ChannelShuffle}(\text{Concat}[F_1, F_2, F_3, F_4, F_5]) \quad (10)$$

$$F' = \text{ChannelShuffle}(\text{LGFA}(\text{GConv}(F_{cat}))) \quad (11)$$

$$F_{res} = \text{ReLU}(\text{BN}(\text{Conv}_{1 \times 1}(F', 16))) + \text{Conv}_{1 \times 1}(F_2) \quad (12)$$

其中,  $\text{Conv}_{1 \times 1}(F_i; C)$  表示通过  $1 \times 1$  卷积调整通道至  $C$ ,  $\alpha$  表示缩放系数,这里  $\alpha = 2$ , GConv 表示分组卷积操作。

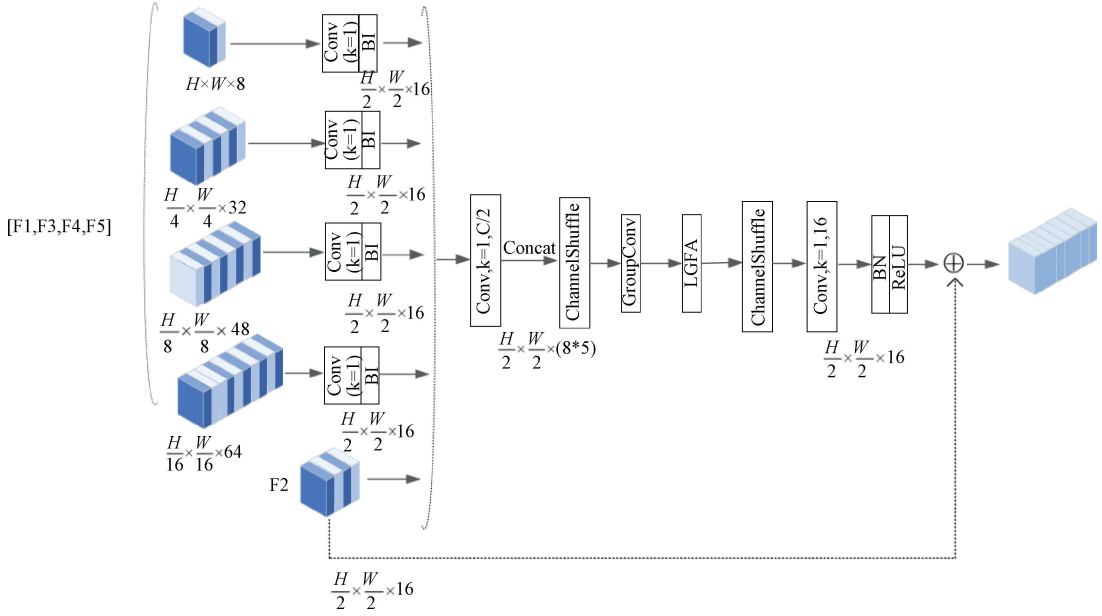


图 4 多尺度信道交互增强模块

Fig. 4 Structure of the multiscale channel interaction enhancement module

### 2.5 轻量级自适应加权解码模块 AdaptiveFusion

在解码器上采样过程中,可以让编码器输出通过逐阶段融合跳跃连接传递的细节信息补偿编码过程中的信息损失,传统 U 型网络通过采取沿通道拼接的方式融合这两部分的特征图,然而简单的通道拼接无法充分建模两部分特征之间的复杂关系,导致在信息传递时可能丢失细微的上下文关联,并且难以根据不同语义层级自适应调整各部分特征的重要性,导致最终分割精度的缺失。为了应对上述问题,本文基于 LGFA 通道注意力机制设计了轻量级自适应加权解码模块,让模型自适应调整两部分的权重,实现更加高效的特征融合。如图 5 所示,将编码器输出进行上采样后与跳跃连接传递的特征图相加,采用 LGFA 模块和 Softmax 层推导出两部分的权重系数,随后各自对原输入进行加权后相加,在通过 BN 和 ReLU 层处理。

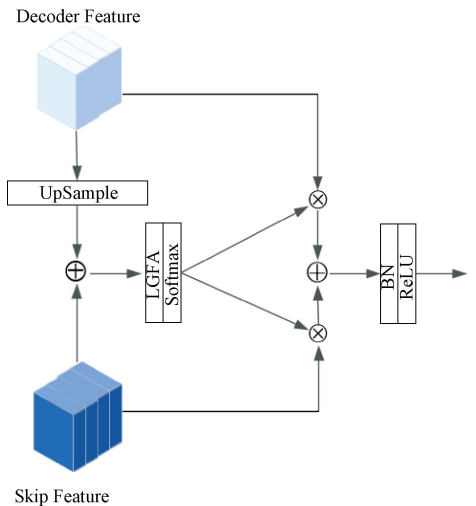


图 5 自适应加权解码模块

Fig. 5 Adaptive weighted decoding module

以上流程可用式(13)~(15)表示,其中  $\otimes$  表示逐通道相乘,  $E$  和  $S$  分别表示编码器和跳跃连接传递的特征图:

$$X = \text{UpSample}(E) + S \quad (13)$$

$$\alpha, \beta = \text{Softmax}(\text{LGFA}(X)) \quad (14)$$

$$Y = \text{ReLU}(\text{BN}(\alpha \otimes X + \beta \otimes S)) \quad (15)$$

### 3 实验结果和分析

#### 3.1 数据集和实验设置

为了验证网络的有效性,本文在 4 个公开皮肤病灶分割数据集 ISIC2017 (international skin imaging collaboration 2017)、ISIC2018、ISIC2016、PH2 (PH2 dermoscopic image database) 上进行综合实验,以训练和评估提出的模型。ISIC2017、ISIC2018、ISIC2016、PH2 数据集分别包含 2 150、2 694、900 张和 200 张带有分割掩码标签的皮肤镜图像,在前 3 个数据集上进行对比实验,将数据集按照 7 : 1.5 : 1.5 的比例随机划分为训练集、验证集和测试集。本文还在 PH2 数据集上进行泛化实验以验证网络的泛化能力,上述数据集中的图像分辨率都在 450 pixel × 600 pixel 以上。为了实验需要,本文将所有图像重塑为 256 pixel × 256 pixel 大小,并采用了包括垂直翻转、水平翻转和随机旋转在内的数据增强操作增加数据多样性。

实验环境基于 Ubuntu22.04 操作系统,硬件设备是显存为 24 GB 的 NVIDIA RTX 3090 GPU,模型运行在 Python3.8、CUDA 11.8、PyTorch2.0.1 的配置环境,开发工具为 Pycharm 2022.1.2 专业版。训练过程中使用 AdamW<sup>[27]</sup> 作为优化器,初始学习率设置为  $5 \times 10^{-4}$ ,使用余弦退火学习率调度器,最大迭代次数为 100 次,最小学习率为 0.000 01。训练迭代次数设置为 150,批大小为 8,并采用深度监督机制为解码器的每阶段添加额外的监督信号进行模型训练。

#### 3.2 评价指标

本文采用 3 个被广泛使用的指标评价模型在皮肤镜图像上的分割性能,包括 Dice 相似系数 (dice similarity coefficient, Dice)、交并比 (intersection over union, IoU)、准确率 (accuracy, ACC), 计算式分别如式 (16) ~ (18) 所示,另外,使用 Params 表示参数的个数, FLOPS 表示计算复杂度,本文模型的参数量和计算复杂度均是在输入为 256 pixel × 256 pixel 大小的图片时计算得到的。

$$Dice = \frac{2TP}{2TP + FP + FN} \quad (16)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (17)$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (18)$$

其中, TP、FP、FN、TN 分别代表真阳性、假阳性、假阴性、真阴性。

#### 3.3 对比实验

为了验证本文模型的有效性,在相同配置环境下,选取 U-Net<sup>[7]</sup>、Attention U-Net<sup>[15]</sup>、TransUNet<sup>[11]</sup>、Swin-UNet<sup>[28]</sup>、Hiformer<sup>[29]</sup>、MALUNET<sup>[18]</sup>、DCSAUNet<sup>[30]</sup>、EGE-UNet<sup>[19]</sup>、CMUNeXt<sup>[20]</sup>、GA-UNet<sup>[24]</sup> 这些语义分割经典网络、医学图像分割的先进网络和轻量化网络在

ISIC2018、ISIC2017、ISIC2016 三个公开的皮肤病变数据集上进行对比试验,不同网络的分割性能指标如表 1、2 所示,加粗表示最优,下划线表示次优。

表 1 ISIC2017 上各模型指标数据  
Table 1 Segmentation results for each model on the ISIC2017 dataset

模型	Params/M	FLOPS/G	IoU	Dice	Acc
UNet	32.08	50.75	75.53	86.06	94.84
AttUNet	34.88	66.63	84.29	91.47	96.74
TransUNet	54.86	9.376	83.23	90.85	96.56
Swin-Net	41.34	8.69	82.24	90.25	96.25
MALUNET	<u>0.175</u>	<u>0.083</u>	82.88	90.64	96.41
DCSAUNet	2.6	6.92	82.95	90.68	96.45
Hiformer	25.51	51.0	<b>85.38</b>	<b>92.11</b>	<b>96.97</b>
EGE-UNet	<b>0.045</b>	<b>0.072</b>	83.69	91.12	96.56
CMUNeXt	3.15	7.42	84.10	91.36	96.74
GA-UNet	1.92	4.46	84.86	91.81	96.93
LMSCI-Net	0.20	0.62	<u>85.24</u>	<u>92.03</u>	<u>96.95</u>

由表 1 可看出本文网络 LMSCI-Net 在轻量化设计和各项分割性能指标上的优越性,在 ISIC2017 数据集上, Dice、IoU、Acc 指标分别达到了 85.24%、92.03%、96.95%,而模型规模和计算复杂度分别控制在 1 M 和 1 G 内,相较于经典 U-Net 网络在参数量和计算复杂度上分别降低了 99.38% 和 98.78%,与最优网络 Hiformer 相比,分割精度略低但在参数量和计算复杂度上分别降低了 99.22% 和 98.79%,Hiformer 基于 transformer 架构并采取了更加复杂的跨尺度特征融合机制,在分割性能上优于本文的轻量化结构,却大大提升了模型复杂度。

基于 U 型网络结构的 U-Net 和 AttUNet 等复杂模型通常采用每阶段通道翻倍的模式,现有研究<sup>[31]</sup>表明通过简单堆叠特征通道的数量并不能带来性能的提升反而会引引起通道冗余,本文网络 LMSCI-Net 在编码部分只采用较少的通道数,但通过探究多尺度信息以及信道间的交互,实现了轻量高效的结构设计,用较少的资源需求达到了较高的病灶分割结果。

相较于 DCSAUNet、CMUNeXt、GA-UNet 这些轻量化模型,本文模型更加轻量且各个指标均得到提升,而相较于 MALUNET、EGE-UNet 这类超轻量化模型,本文网络略微增加参数和计算复杂度,在 3 个指标上分别得到较高提升,过于追求轻量的网络设计会损失较多的精度,模型资源需求和分割性能是一个权衡的过程,本文网络 LMSCI-Net 在这方面取得了最优。

表 2 展示了各个网络在 ISIC2016 和 ISIC2018 数据集上的测试结果,本文网络 LMSCI-Net 的 IoU 值达到

84.49%和81.05%,Dice值达到91.60%和89.53%,Acc值达到95.82%和95.50%。略低于最优网络Hiformer但大幅度减少了参数量和计算复杂度,相较于超轻量网络EGE-UNet,增加少量参数和计算复杂度,使IoU提高1.27%和1.03%,Dice提高0.76%和0.63%,Acc提高0.36%和0.24%,LMSCI-Net网络在保持较高分割精度的同时大大降低了模型参数量和计算复杂度。

表2 ISIC2016和ISIC2018上各模型指标数据

Table 2 Segmentation results for each model on the ISIC2016 dataset and ISIC2018

模型	ISIC2016			ISIC2018		
	IoU	Dice	Acc	IoU	Dice	Acc
UNet	73.44	84.68	92.49	76.25	86.52	94.29
AttUNet	82.61	90.48	95.27	79.73	88.72	95.09
TransUnet	82.87	90.63	95.27	79.32	88.47	95.06
Swin-Net	83.16	90.81	95.30	78.62	88.03	94.93
MALUNET	83.47	90.99	95.52	79.43	88.54	94.99
DCSAUNet	81.92	90.06	95.01	80.34	89.10	95.32
Hiformer	<b>85.10</b>	<b>91.95</b>	<b>95.92</b>	<b>82.12</b>	<b>90.18</b>	<b>95.75</b>
EGE-UNet	83.22	90.84	95.46	80.02	88.90	95.26
CMUNeXt	83.54	91.03	95.66	79.27	88.43	95.05
GA-UNet	83.49	91.01	95.58	80.40	89.14	95.31
LMSCI-Net	84.49	91.60	95.82	81.05	89.53	95.50

为更直观反映本文模型相较于其他算法的优越性,图6和7可视化展示了原图、Ground Truth图像、各个网络在ISIC2016、ISIC2017和ISIC2018数据集部分病灶图片上的分割结果图像,图中每行表示一例病例,每列表示使用的模型,正确预测区域为白色,漏检区域为浅灰,误检

区域为深灰。图6第1张病灶图片边缘非常不规则,图6第2张图片和图7第4张分别存在血管干扰和毛发遮挡的问题,图6第3、4张和图7第2张存在边缘不清晰的问题,图7第3张存在伪影问题,可以清晰地观察到,相较于其他网络存在较多的漏检误检现象,本文网络漏判误判现象较为轻微,能够较为精准地区分病变区域像素和正常区域像素,分割结果更贴合真实掩码图像。

### 3.4 泛化实验

为评估本文网络LMSCI-Net在不同数据集间的泛化性能,进行跨数据集泛化实验,具体而言,将分别在ISIC2017、ISIC2018和ISIC2016数据集上预训练的模型,直接应用于PH2数据集测试分割性能,详细量化结果如表3所示,本文提出的网络LMSCI-Net在跨数据集测试中展现出优越的泛化能力。当使用在ISIC2018数据集上训练后的模型进行测试时,LMSCI-Net在PH2数据集上取得了最优的分割效果,其IoU、Dice系数和准确率(Acc)分别达到了84.50%、91.60%和94.46%,显著优于所有对比的基线网络。同样,使用ISIC2016数据集上训练的结果进行测试,LMSCI-Net在PH2上的性能也达到了最优,IoU、Dice和Acc分别为82.80%、90.59%和93.61%。而使用在ISIC17数据集上训练的结果进行测试,LMSCI-Net的量化结果略逊于Hiformer,优于其余网络。一般来说,由于Hiformer基于更复杂的网络结构和跨尺度语义融合机制,在泛化实验中会比较轻量化网络具有更好的表现,然而本文泛化实验中采用的PH2数据集中病灶较为清晰,伪影较少,在该测试集上轻量化且对局部细节敏感的网络可能会表现更佳。总之,综合数据来看,无论训练源数据集如何变化,本文网络在目标PH2数据集上的表现优越,这足以说明其学习到的特征具有良好的迁移性和泛化性,不易受特定训练数据分布的影响,充分验证了LMSCI-Net的有效性和鲁棒性。

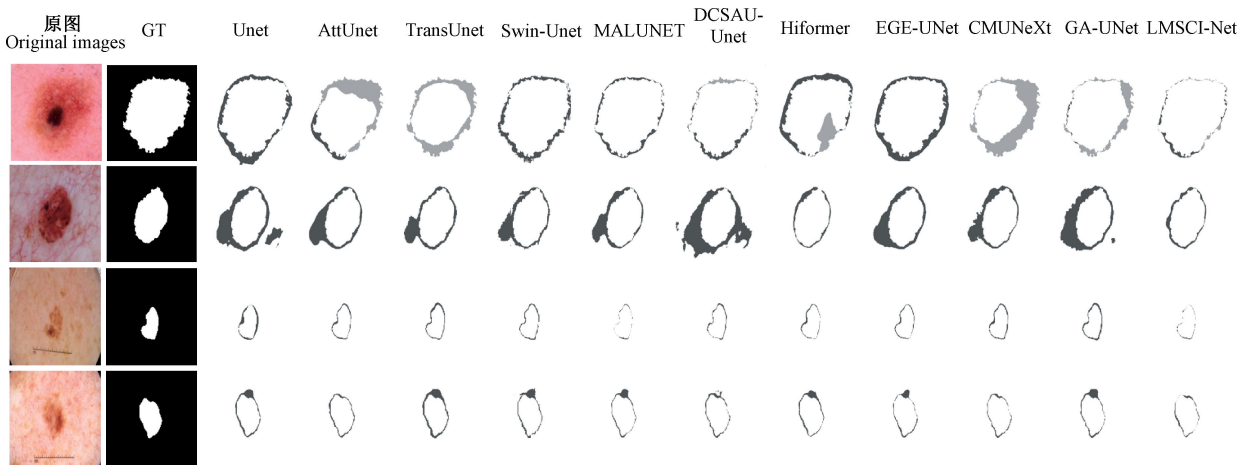


图6 ISIC2016和ISIC2018分割结果可视化

Fig. 6 Visualisation of ISIC2016 and ISIC2018 segmentation results

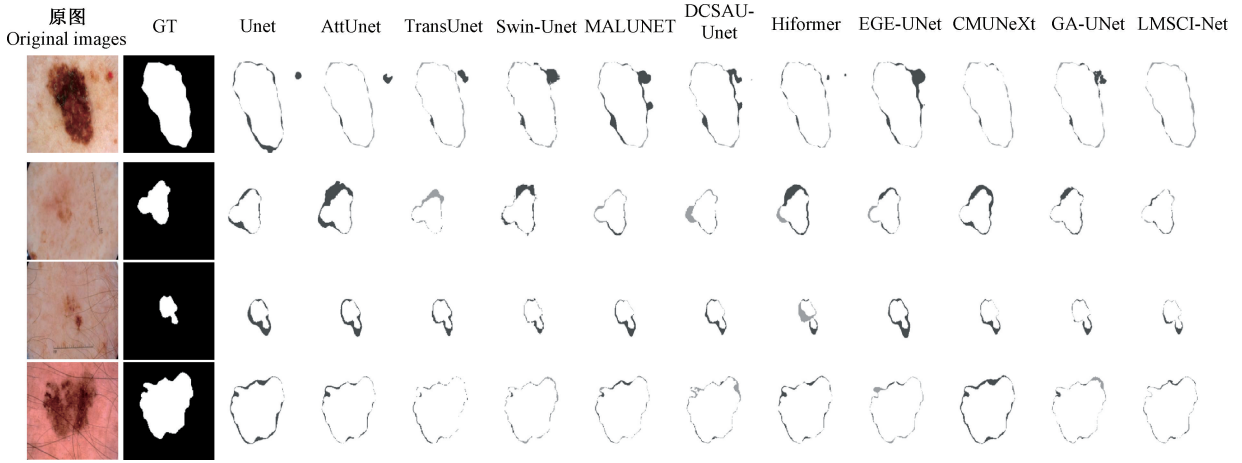


图 7 ISIC2017 分割结果可视化

Fig. 7 Visualisation of ISIC2017 segmentation results

表 3 在 PH2 上进行泛化实验的结果

Table 3 Segmentation results of generalisation experiments on PH2

模型	ISIC2017			ISIC2018			ISIC2016		
	IoU	Dice	Acc	IoU	Dice	Acc	IoU	Dice	Acc
UNet	78.52	87.97	92.42	81.25	89.66	93.35	76.79	86.87	91.32
AttUNet	79.38	88.50	92.81	84.35	91.51	94.35	80.09	88.95	92.43
TransUnet	81.32	89.70	93.51	81.24	89.65	93.14	81.15	89.59	93.02
Swin-Net	81.51	89.81	93.32	79.52	88.59	92.50	77.22	87.14	90.89
MALUNET	82.94	90.67	93.92	82.92	90.66	93.99	77.81	87.51	92.36
DCSAUNet	77.32	87.21	91.89	83.07	90.75	93.90	79.21	88.40	92.16
Hiformer	<b>83.72</b>	<b>91.14</b>	<b>94.10</b>	83.86	91.22	94.08	81.24	89.65	92.70
EGE-UNet	81.84	90.02	93.27	82.50	90.40	93.61	81.42	89.75	92.88
CMUNeXt	80.24	89.03	93.03	85.33	92.09	94.83	78.98	88.26	92.27
GA-UNet	82.26	90.27	93.85	83.11	90.77	93.85	80.57	89.24	92.63
LMSCI-Net	83.11	90.77	93.96	<b>84.50</b>	<b>91.60</b>	<b>94.46</b>	<b>82.80</b>	<b>90.59</b>	<b>93.61</b>

### 3.5 消融实验

#### 1) 模块消融实验

如表 4 所示, 本文在 ISIC2017 数据集上进行消融实验, 基线网络为经典 U-Net 网络, 将原先采用每阶段通道翻倍的模式改为与本文网络相同的设计, 五层编码器通道数分别为 8、16、32、48、64, 可以看出轻量化后的 U-Net 网络虽然通过减少通道数大幅降低了参数量和计算复杂度, 但也导致网络缺乏足够的特征提取能力, 各项指标大幅度降低, 本文在基线网络 U-Net 的基础上, 通过探究多尺度信息、信道关联与交互以及浅层深层语义鸿沟等问题, 提出 3 个改进模块: 轻量级多尺度编码模块 LMSE、多尺度信道交互增强模块 MSIE 和自适应融合解码模块 AdaptiveFusion。在做单个模块消融时可以看出各个模块均有较大提升, 其中多尺度编码模块 LMSE 大幅度提高模型的特征提取能力, 在基线网络的基础上将 IoU、Dice 和 Acc 分别提高 13.97%、9.03% 和 3.14%。模块组合消融

时也可以看出相较只启用单个模块, 在各项指标上也有进一步提升, 采用深度监督机制后, 为解码器的每阶段添加额外的监督信号进行模型训练, 能迫使网络学习更加鲁棒和更有效的特征, 最终使本文网络达到最优结果。消融实验的实验结果表明各个模块的有效性, 在基线网络的基础上逐步添加各个模块能有效提高网络的各项分割指标。综合各项指标以及网络参数量, 相较于各个对比网络, 本文提出的网络在皮肤病变分割任务上达到了最优效果。

#### 2) 损失函数消融实验

像素级分类任务通常采用二元交叉熵损失函数<sup>[32]</sup>作为网络训练时的损失函数, 而交叉熵损失函数是基于每个类别的预测概率计算的, 在类别不平衡的情况下, 会使网络难以关注比例较少的区域, 在处理微小病灶上效果不佳, 导致最终分割效果并不理想。于是本文引入另一种损失函数 DiceLoss<sup>[33]</sup>, DiceLoss 本质上是对重叠区域与总体区域的比例进行优化, 与类别像素数量无关, 因此能够更

表4 在ISIC2017上进行模块消融实验

Table 4 Modular ablation experiments at ISIC2017

Baseline	LMSE	MSIE	AdaptiveFusion	深度监督	Params/M	FLOPS/G	IoU/%	Dice/%	Acc/%
√	×	×	×	×	0.25	0.68	68.99	81.66	93.25
√	√	×	×	×	0.15	0.57	82.96	90.69	96.39
√	×	√	×	×	0.33	1.15	79.50	88.58	95.73
√	×	×	√	×	0.17	0.31	79.37	88.50	95.75
√	√	×	√	×	0.08	0.18	83.76	91.17	96.65
√	√	√	×	×	0.24	1.02	84.04	91.33	96.65
√	×	√	√	×	0.29	0.76	83.11	90.78	96.53
√	√	√	√	×	0.20	0.62	<u>84.44</u>	<u>91.56</u>	<u>96.78</u>
√	√	√	√	√	0.20	0.62	<b>85.24</b>	<b>92.03</b>	<b>96.95</b>

有效地处理类别不平衡问题,能引导模型关注一些较小的病灶区域。综上,本文采用基于Dice损失函数和二元交叉熵损失函数的融合损失函数 $L_{BceDice}$ 进行模型训练,最终的损失函数公式如式(19)~(21)所示。

$$LBce = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (19)$$

$$LDice = 1 - \frac{2 |X \cap Y|}{|X| + |Y|} \quad (20)$$

$$LBceDice = \lambda_1 LBce + \lambda_2 LDice \quad (21)$$

其中, $N$ 为样本总数, $y_i$ 为真实标签, $p_i$ 为预测值。 $|X|$ 和 $|Y|$ 表示真值和预测值。 $\lambda_1$ 和 $\lambda_2$ 分别表示两个损失函数的权重,本文 $\lambda_1$ 和 $\lambda_2$ 均取1。为了验证该融合损失函数 $BceDiceLoss$ 的有效性,分别采用损失函数 $BCELoss$ 、 $DiceLoss$ 以及本文使用的融合损失函数在ISIC2017数据集上进行消融实验,结果如表5所示,相比只使用 $BCELoss$ 损失函数进行训练在IoU和Dice指标上分别高出1.27%和0.74%,比只使用 $DiceLoss$ 分别提高0.66%和0.38%,表明相较于只启用单一损失函数,使用融合损失函数能得到最高的分割精度。

### 3.6 注意力对比试验

本文设计局部-全局通道注意力模块,在全局信息描述符中引入局部信息描述符,以弥补获取通道级统计数据时局部信息的缺失,使得模型能够更好地捕捉通道之间的复杂关系,得到更加精确的通道注意力权重。SE<sup>[24]</sup>和

CBAM<sup>[8]</sup>注意力机制在计算通道注意力时主要依赖全局池化操作捕获的全局池化信息,忽视了局部细节的重要性,高效通道注意力(efficient channel attention, ECA)<sup>[34]</sup>注意力机制是对SE模块的改进,在计算通道注意力时仍然主要依赖全局平均池化获取的全局信息,而坐标注意力(coordinate attention, CA)<sup>[35]</sup>通过分别沿水平和垂直方向进行全局平均池化,并在后续处理中融入位置信息,在一定程度上考虑了局部信息,缓解了全局信息的不足。

表5 在ISIC2017上进行损失函数消融实验

Table 5 Loss function ablation experiments at ISIC2017

					%
BCE	DICE	IoU	Dice	Acc	
√	×	83.97	91.29	96.75	
×	√	84.58	91.65	96.83	
√	√	<b>85.24</b>	<b>92.03</b>	<b>96.95</b>	

本节实验基于ISIC2017数据集对网络中涉及注意力机制的部分做对比试验,测试不同注意力机制下的网络性能,如表6所示,采用LGFA注意力机制后,相较于不使用任何注意力在IoU、Dice和Acc上分别提高1.79%、1.05%和0.30%,相较于次优注意力机制ECA分别提高1.12%、0.65%和0.23%,足以说明本文LGFA通道注意力机制的优越性,能得到更加精确的通道注意力权重。

表6 在ISIC2017上进行注意力对比实验

Table 6 Comparative attention experiments at ISIC2017

CA(CBAM)	SE	CA	ECA	LGFA	Param/M	FLOPS/G	IoU/%	Dice/%	Acc/%
×	×	×	×	×	0.137	0.603	83.45	90.98	96.65
√	×	×	×	×	0.168	0.608	84.02	91.32	96.68
×	√	×	×	×	0.141	0.604	83.68	91.12	96.69
×	×	√	×	×	0.152	0.613	<u>84.68</u>	<u>91.70</u>	<u>96.82</u>
×	×	×	√	×	0.137	0.608	84.12	91.38	96.72
×	×	×	×	√	0.198	0.625	<b>85.24</b>	<b>92.03</b>	<b>96.95</b>

## 4 结 论

现有神经网络倾向于设计更深更复杂的结构达到高分割精度的目的,然而较高的模型参数和计算成本限制了实际应用,同时还存在缺乏对于网络编码层多尺度输出间关联的探索、信道冗余和信道交互的考虑以及忽视浅层深层特征的语义鸿沟的问题。针对以上问题本文分别提出了局部-全局通道注意力模块 LGFA、轻量级多尺度编码模块 LMSE、多尺度信道交互增强模块 MSIE 和自适应融合解码模块 AdaptiveFusion 用于优化基线网络,融合 *Dice* 损失函数和二元交叉熵损失函数来构建训练的损失函数,并结合深度监督机制,通过在多个数据集上进行综合实验,验证了本文网络的优越性,在实现轻量化的同时保持较高的病灶分割精度,这在实际医疗场景中有较高的参考价值。在未来的研究中,将探索网络在其他医学图像分割任务中的应用,继续优化网络使其在临床环境中能够带来更多的价值。

## 参考文献

- [1] BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA: A Cancer Journal for Clinicians, 2024, 74(3): 229-263.
- [2] SIEGEL R L, MILLER K D, FUCHS H E, et al. Cancer statistics, 2021[J]. CA:A Cancer Journal for Clinicians, 2021, 71(1): 7-33.
- [3] 张玮智,于谦,苏金善,等.从 U-Net 到 Transformer: 深度模型在医学图像分割中的应用综述[J].计算机应用,2024,44(S1):204-222.
- ZHANG W ZH, YU Q, SU J SH, et al. From U-Net to Transformer: A review of deep models in medical image segmentation [J]. Journal of Computer Applications, 2024, 44(S1): 204-222.
- [4] SAHA A, PRASAD P, THABIT A. Leveraging adaptive color augmentation in convolutional neural networks for deep skin lesion segmentation[C]. IEEE 17th International Symposium on Biomedical Imaging. Piscataway: IEEE, 2020: 2014-2017.
- [5] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [6] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, Piscataway: IEEE, 2015: 3431-3440.
- [7] RONNEBERGER O, FISCHER P, BROX T. U-net:

Convolutional networks for biomedical image segmentation [C]. Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 234-241.

- [8] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]. European Conference on Computer Vision. Cham: Springer, 2018: 3-19.
- [9] ZHAO SH, WANG Y L, WU X, et al. MAFNet: Dual-branch fusion network with multiscale atrous pyramid pooling aggregate contextual features for real-time semantic segmentation [J]. Complex & Intelligent Systems, 2024: 5107-5126.
- [10] 刘拥民,张毅,欧阳凌轩,等.基于改进 U-Net 的轻量级眼底病变分割算法设计[J].电子测量技术,2024, 47(3):127-134.
- LIU Y M, ZHANG Y, OUYANG L X, et al. Design of lightweight fundus lesion segmentation algorithm based on improved U-Net[J]. Electronic Measurement Technology, 2024, 47(3): 127-134.
- [11] CHEN J N, LU Y Y, YU Q H, et al. Transunet: Transformers make strong encoders for medical image segmentation [EB/OL]. (2021-05-20) [2024-05-08]. <https://arxiv.org/abs/2102.04306v1>.
- [12] ZHENG S X, LU J CH, ZHAO H SH, et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 6881-6890.
- [13] DONG X Y, BAO J M, CHEN D D, et al. CSWin transformer: A general vision transformer backbone with cross-shaped windows [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 12124-12134.
- [14] ZHOU Z W, RAHMAN SIDDIQUEE M M, TAJBAKHS N, et al. Unet++: A nested U-Net architecture for medical image segmentation [C]. International Workshop on Deep Learning in Medical Image Analysis and the International Workshop on Multimodal Learning for Clinical Decision Support. Cham: Springer, 2018: 3-11.
- [15] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-net: Learning where to look for the pancreas [EB/OL]. (2018-05-20) [2024-07-24]. <https://arxiv.org/abs/1804.03999v3>.
- [16] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows [C]. IEEE/CVF International Conference on Computer Vision.

- Piscataway: IEEE, 2021: 10012-10022.
- [17] 梁礼明,彭仁杰,冯骏,等. 基于多尺度特征融合双U型皮肤病变分割算法[J]. 计算机应用研究, 2021, 38(9):2876-2880.
- LIANG L M, PENG R J, FENG J, et al. A skin lesion segmentation algorithm based on multi-scale feature fusion and dual U-Net [J]. Application Research of Computers, 2021, 38(9): 2876-2880.
- [18] RUAN J CH, XIANG S CH, XIE M Y, et al. MALUNet: A multi-attention and light-weight unet for skin lesion segmentation[C]. IEEE Conference on Bioinformatics and Biomedicine. Piscataway: IEEE, 2022: 1150-1156.
- [19] RUAN J CH, XIE M Y, GAO J SH, et al. EGE-UNet: An efficient group enhanced unet for skin lesion segmentation [C]. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer Nature Switzerland, 2023: 481-490.
- [20] TANG F H, DING J R, QUAN Q, et al. Cmunext: An efficient medical image segmentation network based on large kernel and skip fusion[C]. 2024 IEEE International Symposium on Biomedical Imaging (ISBI). IEEE, 2024: 1-5.
- [21] WU R K, LIANG P CH, HUANG X, et al. MHorUNet: High-order spatial interaction UNet for skin lesion segmentation [J]. Biomedical Signal Processing and Control, 2024, 88: 105517.
- [22] 陶惜婷,叶青. 融合 CNN 和 Transformer 的并行双分支皮肤病灶图像分割[J]. 计算机应用研究, 2024, 41(8):2554-2560.
- TAO X T, YE Q. Parallel dual-branch skin lesion image segmentation fusing CNN and Transformer[J]. Application Research of Computers, 2024, 41(8): 2554-2560.
- [23] ALRFOU K, ZHAO T. Gctx-unet: Efficient network for medical image segmentation[J]. ArXiv preprint arXiv:2406.05891, 2024.
- [24] PANG B, CHEN L H, TAO Q CH, et al. GA-UNeT: A lightweight ghost and attention u-net for medical image segmentation[J]. Journal of Imaging Informatics in Medicine, 2024, 37(4): 1874-1888.
- [25] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [26] ZHANG X Y, ZHOU X Y, LIN M X, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6848-6856.
- [27] LOSHCHILOV I. Decoupled weight decay regularization[EB/OL]. (2019-01-04) [2024-07-18]. <https://arxiv.org/abs/1711.05101v3>.
- [28] CAO H, WANG Y Y, CHEN J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation[C]. European Conference on Computer Vision. Cham: Springer, 2022: 205-218.
- [29] HEIDARI M, KAZEROUNI A, SOLTANY M, et al. Hiformer: Hierarchical multi-scale representations using transformers for medical image segmentation[C]. IEEE/CVF Winter Conference on Applications of Computer Vision, 2023: 6202-6212.
- [30] XU Q, MA ZH CH, HE N, et al. DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation[J]. Computers in Biology and Medicine, 2023, 154: 106626.
- [31] VASA V K, ZHU W H, CHEN X W, et al. STA-Unet: Rethink the semantic redundant for Medical Imaging Segmentation [J]. ArXiv preprint arXiv: 2410.11578, 2024.
- [32] RUBY U, YENDAPALLI V. Binary cross entropy with deep learning technique for image classification [J]. International Journal of Advanced Trends in Computer Science and Engineering, 2020, 9(10). DOI:10.30534/ijatcse/2020/175942020.
- [33] MILLETARI F, NAVAB N, AHMADI S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation [C]. 2016 Fourth International Conference on 3D Vision(3DV). IEEE, 2016: 565-571.
- [34] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534-11542.
- [35] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.

## 作者简介

王斯豪, 硕士研究生, 主要研究方向为深度学习、医学影像处理。

张笃振(通信作者), 副教授, 硕士生导师, 博士, 主要研究方向为图像处理、模式识别、机器学习。

杨昌昌, 硕士研究生, 主要研究方向为人工智能、目标检测。

E-mail: zhduzhen@jsnu.edu.cn