Skin Lesion Image Segmentation Method Based on Ultra-Lightweight Real-Time Segmentation Network*

MEI Xiajin[1], WU Liping[1], ZHANG Wenxin[1], MA Yuliang[1,2*]
(1.School of Automation, Hangzhou Dianzi University, Hangzhou Zhejiang 310018, China;
2.International Joint Research Laboratory for Autonomous Robotic Systems, Hangzhou Zhejiang 310018, China)

Abstract: Skin lesion image segmentation is crucial in diagnosing skin lesions. However, deep learning-based segmentation models typically have high computational costs and slow inference speeds, making them difficult to be be deployed on dermatoscopic devices with limited computing power. To address this issue, an ultra-lightweight real-time segmentation network, named ULRTS-Net, is proposed. First, depthwise separable convolutions are adopted to replace standard convolutions, and a relatively lightweight encoder-decoder network architecture is designed to reduce model complexity and computational load. Second, multi-level semantic feature fusion modules are added at skip connections to effectively bridge the semantic gap between shallow and deep features. Additionally, a multi-scale feature fusion module is proposed to enhance the model's ability to learn contextual information. Finally, spatial and channel attention modules are introduced to focus on important features. Experiments show that ULRTS-Net achieves JI scores of 85.78% and 89.95% on the ISIC2016 and PH2 datasets, respectively, with only 0.407M model parameters and 1.51 GFLOPs. Compared to other methods, ULRTS-Net achieves fast and accurate segmentation with low computational costs, demonstrating its effectiveness.

Key words: skin lesion segmentation; ultra-lightweight real-time segmentation; feature fusion; attention mechanism
EEACC: 7510 **doi**: 10.3969/j.issn.1004-1699.2025.10.008

基于超轻量实时分割网络的皮肤病变图像分割方法*

梅厦锦[1], 巫笠平[1], 张文新[1], 马玉良[1,2*]
(1.杭州电子科技大学自动化学院, 浙江 杭州 310018; 2.自主机器人系统联合实验室, 浙江 杭州 310018)

摘 要: 皮肤病变图像分割在皮肤病变诊断中至关重要。然而, 基于深度学习的皮肤病变分割模型计算成本高和推理速度慢, 难以部署在计算能力有限的皮肤镜设备上。针对上述问题, 本研究提出了一种超轻量实时分割网络。首先, 采用深度可分离卷积代替标准卷积, 并设计了相对轻量的编解码网络架构, 以降低模型复杂度和计算量。其次, 在跳跃连接处加入了多级语义特征融合模块, 有效缩小了浅层特征与深层特征的语义差距。此外, 提出了多尺度特征融合模块以增强模型学习上下文信息的能力。最后, 引入空间和通道注意力模块以关注重要特征。实验表明, ULRTS-Net 在 ISIC2016 和 PH2 数据集上的 JI 指标分别达到 85.78%和 89.95%, 且仅有 0.407M 的模型参数量和 1.51GFLOPs。相比其他方法, ULRTS-Net 在低计算成本下实现了快速精准的分割, 验证了其有效性。

关键词: 皮肤病变分割; 超轻量实时分割; 特征融合; 注意力机制
中图分类号: TP183; TP391.4 **文献标识码**: A **文章编号**: 1004-1699(2025)10-1784-09

黑色素瘤是世界上最常见的恶性肿瘤之一, 具有极高的致死率[1]。研究数据显示, 如果患者能够在早期发现皮肤黑色素瘤并进行干预治疗, 他们的存活率将显著提高[2]。皮肤镜是一种用于观察色素性皮肤病的工具, 有助于医生在早期识别黑色素瘤。在临床诊断中, 皮肤病变的检测方法通常依赖于皮肤科医生对皮肤镜图像的人工观察, 并手动分割出病变区域。然而, 这种人工分割的过程既繁琐又耗时, 不仅需要大量人力投入, 而且容易受到主观因素的影响。

目前, 用于皮肤病变图像分割的计算机辅助诊断方法主要分为基于传统数字图像处理方法和基于深度学习的图像分割方法[3-4]。在皮肤病变图像分割方法的早期研究中, 传统数字图像处理方法, 如边缘检测[5]、区域[6]和阈值[7]等, 专注于从皮肤病变图像中提取浅层特征, 如形状[8]、颜色[9]和纹理[10]。然而, 这些传统方法容易受到各种因素的干扰, 其过程繁琐且缺乏鲁棒性。

近年来, 深度学习方法取得了广泛关注, 因而也被引入到皮肤镜图像病变区域分割领域。随着全卷积神经网络(Fully Convolutional Neural Network, FCN)[11]的提出, 它在图像分割领域得到了广泛应用。Ronneberger 等[12]在 FCN 的基础上提出了

项目来源: 国家自然科学基金项目(62071161); 浙江省教育厅科研项目(Y202351775)
收稿日期: 2024-04-11 **修改日期**: 2024-06-12

U-Net,通过跳跃连接将浅层信息与深层信息拼接,以解决 FCN 在分割细节上较差的问题。为了提高皮肤病变图像的分割效果,许多研究对模型进行了改进,增加了模型复杂度和计算量,但忽略了模型的轻量化设计。Xie 等[13]提出了 HRC-Net,通过引入一个高分辨率特征提取模块,能够保留皮肤病变区域的边界信息,从而提高分割精度,但导致了计算成本的增加。Wu 等[14]提出了 FAT-Net,通过将 Transformer 模块加到编解码器架构中,以学习特征的长距离依赖关系,但也带来了较高的计算成本。Lei 等[15]提出了 DAGAN,在密集卷积 U-Net 的基础上,增加了一个双判别模块来优化分割结果的边界,同时也增加了模型的复杂度。同时,Goyal 等[16]提出了一种全自动的深度学习集成方法,同时使用两个模型以实现在皮肤病变图像分割中的高灵敏度和高特异性,但相应地增加了计算资源的消耗。Xie 等[17]提出了一种同时进行皮肤病变图像分割和分类的方法,这种方法虽然提高了分割和分类的性能,但增加了对计算资源的需求。Lee 等[18]提出了一种结构边界保持分割框架,通过对边界特征的学习以提高分割性能,但也增加了模型复杂度。Feng 等[19]提出了 CPFNet,在 U-Net 的基础上引入了金字塔模块用于学习图像中的上下文信息,提高了分割精度,但以增加模型的复杂度为代价。Yu 等[20]提出了 EIU-Net,通过学习局部和全局上下文信息,提升了模型的分割性能,但也增加了模型的计算成本与复杂度。Naqvi 等[21]提出了一种用于眼部区域语义分割的残差编解码网络 ORED-Net,通过使用非恒等残差连接,减少了从编码器到解码器的信息丢失。Yamanakkanavar 等[22]提出了 M-SegNet,通过全局注意力来提高对上下文信息的理解,从而提升分割效果。Xu 等[23]提出 DCSAU-Net,该网络通过整合低级和高级语义信息,有效提升了网络的分割性能。然而,这些改进都提升了网络的分割性能,但造成了模型复杂度和计算成本的增加。而 DSNet[24]和 Separable-UNet[25]都使用深度可分离卷积代替标准卷积,MF2-Net[26]使用不同大小的非对称卷积核替换了相应的标准卷积。尽管这些方法降低了模型复杂度和浮点计算量,但却牺牲了分割精度。现有的经典轻量级图像分割模型 ENet[27],已被应用于皮肤病变分割,在推理速度上表现出显著的优势,但 ENet 的分割性能远远落后于目前其他先进的皮肤病变图像分割方法。巫笠平等[28]提出了一种结合八度卷积和分层残差跳连模块的 OctResU-Net,采用八度卷积代替了标准卷积,有效降低了模型复杂度,但在推理速度上表现较慢。Ma 等[29]提出了一种针对皮肤病变分割的超轻量级全非对称卷积网络,以较低的模型复杂度取得了出色的分割效果,但该模型存在实时性不足的问题。

本研究提出了一种超轻量实时分割网络(Ultra-Lightweight Real-Time Segmentation Network, ULRTS-Net),旨在快速地实现皮肤病变图像的精准分割。首先,采用深度可分离卷积代替标准卷积,并设计了一个相对轻量的编解码网络架构,有效降低了模型复杂度和计算量。其次,为了缩小浅层特征与深层特征之间的语义差距,提出了多级语义特征融合模块。同时,为了增强模型对上下文信息的理解,提出了多尺度特征融合模块。最后,为了增强模型的表征能力,引入了空间注意力模块和通道注意力模块。在两个公开皮肤病变分割数据集(PH2 和 ISIC2016)上进行实验评估,以证明该方法的有效性。

1　材料和方法

1.1　数据集

ISIC2016 数据集[30]和 PH2 数据集[31]均为皮肤病变分割任务的公开挑战赛数据集,其中,ISIC2016 数据集包含 1 279 张皮肤病变图像,训练集为 900 张图像,测试集为 379 张图像。而 PH2 数据集包含 200 张皮肤病变图像,本文按照 4∶1 的比例将其随机划分为训练集和测试集。先将皮肤病变图像的分辨率大小统一调整为 224×224,并对其进行灰度化。在训练时,使用了随机翻转、随机旋转、随机放缩和随机亮度变化等实时数据增强方法来增加样本多样性。

1.2　分割网络

所提出的 ULRTS-Net 由编码器和快速解码路径连接路径组成,整体结构如图 1(e)所示。编码器主要由卷积层和池化层构成,逐步提取皮肤病变图像中的低级到高级语义特征。编码器的输出特征将作为解码器的输入,为解码器提供丰富的特征表示。现有基于编解码架构的皮肤病变分割模型通常采用连续多次下采样来降低特征图的分辨率,以获取足够大的感受野并降低模型的计算量,但同时也会丢失皮肤图像中的空间细节信息。ULRTS-Net 的编码器仅通过两次下采样,有效地平衡了空间细节的保留和模型计算量的优化。在 ULRTS-Net 中,设计了一条极具性价比的快速解码路径。该路径没有使用复杂的子模块,从而显著降低了模型的复杂度。为了进一步提高效率,在快速解码路径中,特征通道数在两次上采样过程中被快速压缩到较小的值。同时,快速解码路径通过跳跃连接融合编码器的输出特征。

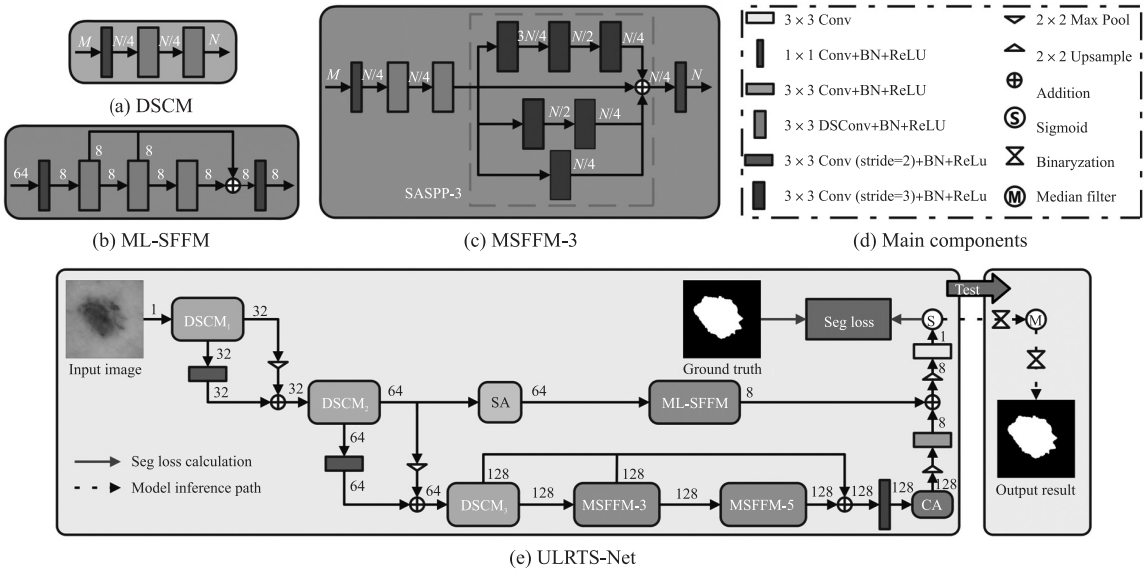

图 1 ULRTS-Net 整体结构和各个模块组成

1.2.1 编码器

ULRTS-Net 的编码器由五个子模块组成,分别为三个深度可分离卷积模块(Depth Separable Convolution Module,DSCM)和两个多尺度特征融合模块(Multi-Scale Feature Fusion Module, MSFFM)。其中,三个 DSCM 所处的深度不同,本文将它们按照从浅到深的顺序命名为 $DSCM_1$,$DSCM_2$ 和 $DSCM_3$。输入图像通过 $DSCM_1$ 进行特征提取,所提取的特征分别通过一个最大池化层和一个卷积核大小为 3×3 的卷积组合(3×3 Conv-BN-ReLU,stride = 2)进行处理,然后将处理后的特征进行相加融合后输入到 $DSCM_2$。$DSCM_2$ 的输出特征同样经过最大池化层和卷积核大小为 3×3 的卷积组合(3×3 Conv-BN-ReLU,stride = 2)进行处理,并通过特征相加融合后输入到 $DSCM_3$。$DSCM_3$ 的输出特征依次经过 MSFFM-3 和 MSFFM-5 进行特征提取,再将 $DSCM_3$、MSFFM-3 和 MSFFM-5 的输出特征相加融合,随后通过一个卷积核大小为 1×1 的卷积组合(1×1 Conv-BN-ReLU)进行跨通道特征融合。

1.2.2 快速解码路径

快速解码路径由三个子模块组成,分别为一个通道注意力、一个空间注意力和一个多级语义特征融合模块。其中,通道注意力模块被添加在编码器之后,使模型更加关注重要的特征通道并抑制冗余特征。空间注意力和多级语义特征融合模块被添加在编码器子模块($DSCM_2$)的跳跃连接路径上,空间注意力用于过滤大量冗余的背景噪声,多级语义特征融合模块则用于增强空间注意力输出特征的语义级别。在快速解码路径中,通过将多级语义特征融合模块的输出特征和第一次上采样后的特征进行特征相加融合,接着通过上采样操作将特征图的分辨率恢复到输入分辨率,并使用一个卷积核大小为 3×3 的卷积将特征通道数压缩到 1,最后经过 Sigmoid 函数处理得到快速解码路径的输出。然而,由于快速解码路径输出的初步分割结果比较粗糙,导致皮肤病变区域的边缘不够精细。本文使用中值滤波来平滑初步分割结果的边缘,从而得到最终的分割结果。

1.2.3 深度可分离卷积模块

相比于标准卷积,深度可分离卷积(DSConv)可以有效地降低模型的参数量和浮点计算量,因此常常作为轻量化模型的基本组件之一。考虑到深度可分离卷积的优势,本文将其引入 ULRTS-Net 中代替标准卷积,在降低模型复杂度的同时也保持了特征提取能力。其中,DSCM 内部包含三个基本组件,第一个组件是一个卷积核大小为 1×1 的卷积组合(1×1 Conv-BN-ReLU),将 DSCM 输入特征的通道数(M)压缩到 DSCM 输出特征通道数(N)的四分之一,以节省运算成本。第二个和第三个组件都是卷积核大小为 3×3 的深度可分离卷积组合(3×3 DSConv-BN-ReLU),用于降低计算成本并提取非线性特征。通过第三个组件,特征通道数被重新恢复到 N。模块结构如图 1(a)所示。

1.2.4 多级语义特征融合模块

在编解码器结构中,跳跃连接通过将浅层特征与深层特征连接起来,有效解决了特征信息丢失的问题。然而,由于浅层特征与深层特征之间存在语义差距,本文提出了一种多级语义特征融合模块(Multi-Level Semantic Feature Fusion Module,ML-SFFM),旨在缩小浅层与深层特征之间的语义差距。ML-SFFM 内部包含五个组件和一个特征相加融合操作,第一个组件是一个

卷积核大小为 1×1 的卷积组合(1×1 Conv-BN-ReLU),用于将 ML-SFFM 输入特征的通道数(64)压缩到 8,显著降低了计算复杂度。第 2~4 组件是三个卷积核大小为 3×3 的深度可分离卷积组合(3×3 DSConv-BN-ReLU),用于提取特征和强化输入特征的语义级别。通过特征相加操作,将第 2~4 个组件的输出特征进行融合,保留了前三层的重要特征信息。最后,通过一个卷积核大小为 1×1 的卷积组合(1×1 Conv-BN-ReLU),对融合后的特征进行跨通道融合,增强特征的表征能力。模块结构如图 1(b)所示。

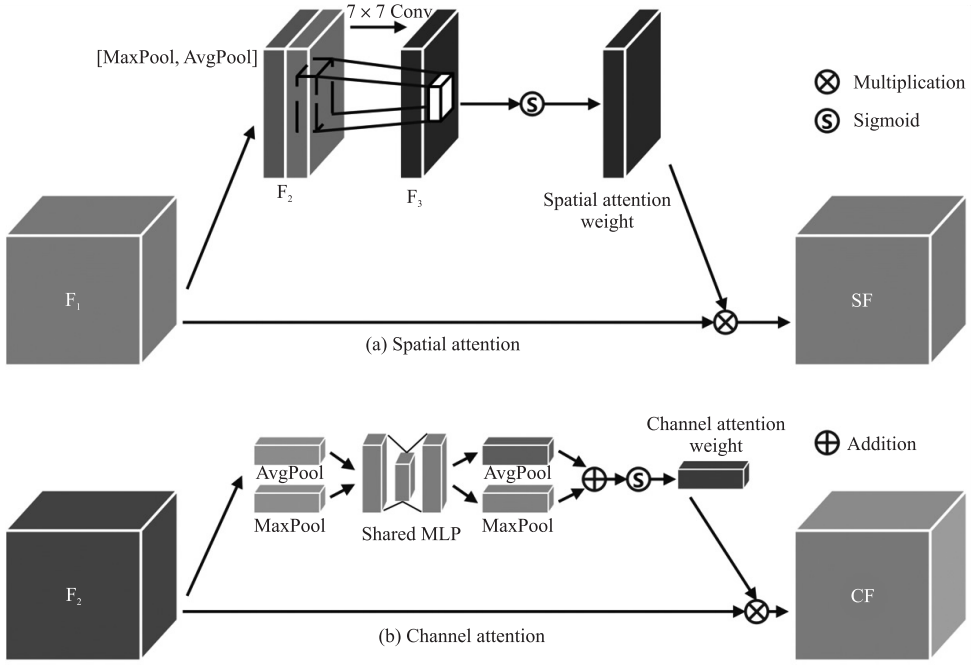

图 2　空间注意力和通道注意力模块结构图

1.2.5　多尺度特征融合模块

空洞空间金字塔池化(Atrous Spatial Pyramid Pooling,ASPP)模块[32]已经广泛用于提取多尺度特征信息。为了解决 ASPP 模块中膨胀率较大导致的卷积层分支稀疏的问题,本文提出了 SASPP-x 模块,将多个具有相同膨胀率(x)的膨胀卷积串联起来,以代替单一膨胀率较大的膨胀卷积,从而获得更大的感受野。SASPP-3 有四个并行分支,其中一个分支直接输出特征,另外三个分支分别串联 1、2 和 3 个膨胀率为 3 的卷积组合(3×3 Conv-BN-ReLU,dilation=3)。模块结构见图 1(c)中所框区域。

MSFFM-3 由 DSCM、SASPP-3 模块和卷积核大小为 1×1 的卷积组合(1×1 Conv-BN-ReLU)组成。首先,DSCM 用于提取和强化语义特征,随后通过 SASPP-3 模块进行多尺度特征的提取,以获得更大的感受野。最后,通过卷积核大小为 1×1 的卷积组合(1×1 Conv-BN-ReLU)进行多尺度特征信息的整合,并将内部特征通道数进行升维。模块结构如图 1(c)所示,其中,MSFFM-3 和 MSFFM-5 的区别在于它们的膨胀率分别被设置为 3 和 5。

1.2.6　空间注意力和通道注意力模块

本文引入了两种经典的注意力模块,分别是空间注意力(Spatial Attention,SA)和通道注意力(Channel Attention,CA)。在皮肤镜临床图像中,通常存在多种干扰噪声和大量冗余的浅层特征,这些噪声可能会严重影响模型对皮肤病变区域进行准确分割。而这两种注意力模块将帮助 ULRTS-Net 分别从空间和通道两个维度来关注重要的特征。

SA 被添加在编码器浅层子模块($DSCM_2$)的特征跳跃连接路径上,由于该路径含有大量冗余的背景噪声,因此采用 SA 来过滤掉这些背景噪声。假设给定 SA 的输入特征图 $F_1 \in \boldsymbol{R}^{H\times W\times C}$,分别在通道维度上进行一次全局最大池化操作和一次全局平均池化操作,并将结果进行特征拼接,得到特征图 $F_2 \in \boldsymbol{R}^{H\times W\times 2}$。然后使用一个卷积核大小为 7×7 的卷积对 F_2 进行特征提取,并将通道数减少到 1,得到特征图 $F_3 \in \boldsymbol{R}^{H\times W\times 1}$。随后,通过 Sigmoid 函数进行激活,得到空间注意力权重。最终,将空间注意力权重和输入特征 F_1 相乘,得到经过空间重标定后的特征图 $\mathrm{SF} \in \boldsymbol{R}^{H\times W\times C}$。模块结构如图 2(a)所示。

CA 被添加在三个深层子模块($DSCM_3$、MSFFM-3 和 MSFFM-5)输出特征的跨通道特征融合步骤之后,旨在减少因通道数压缩而可能造成的重要语义特征丢失。在假设给定 CA 的输入特征图 $F_2 \in$

$\boldsymbol{R}^{H\times W\times C}$的基础上,分别进行一次全局最大池化和全局平均池化,将 F_2 的空间维度压缩为 1,得到两个大小为 $\boldsymbol{R}^{1\times 1\times C}$的通道特征。随后,将两个通道特征分别输入到一个共享的多层感知机中进行特征提取,然后将提取后的特征相加,并通过 Sigmoid 函数进行激活,从而得到通道注意力权重。最终将通道注意力权重和 F_2 进行相乘,得到经过通道重标定的特征图 $\mathrm{CF}\in\boldsymbol{R}^{H\times W\times C}$。模块结构见图 2(b)。

2 实验与结果

2.1 实验环境

本文的实验环境是 PyTorch 框架 1.7.1 版,IDE 是 PyCharm,测试平台是一台装有 Intel® Xeon® Platinum 8358P CPU 和 NVIDIA RTX A5000 GPU 的计算机。使用这两个皮肤病数据集进行训练时,均使用 Adam 优化器,超参数设置情况相同:betas 为(0.9,0.99),初始学习率为 0.001,学习率三段式均匀下降,衰减率为 0.1,训练的 Epoch 为 60。ISIC2016 和 PH2 的 Batch size 分别设置为 5 和 4,所使用的损失函数为 Lovasz hinge loss[33]。

2.2 评估指标

本文所使用的图像分割评价指标包括:准确性(Acc)、Dice 系数(DC)和 Jaccard 相似系数(JI),它们具体计算公式如下:

$$\mathrm{ACC}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FN}+\mathrm{FP}} \tag{1}$$

$$\mathrm{DC}=\frac{2\times\mathrm{TP}}{2\times\mathrm{TP}+\mathrm{FN}+\mathrm{FP}} \tag{2}$$

$$\mathrm{JI}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}+\mathrm{FP}} \tag{3}$$

式中:真阳性(True Positive,TP)是病变区域被正确识别为病变区域的像素数;真阴性(True Negative,TN)是非病变区域被正确识别为非病变区域的像素数;假阳性(False Positive,FP)是非病变区域被错误地识别为病变区域的像素数;假阴性(False Negative,FN)是病变区域被错误地识别为非病变区域的像素数。

为了更全面的评估 ULRTS-Net 的推理速度,本文在三个不同算力平台中对其进行测试,分别是 NVIDIA RTX 2080Ti GPU、NVIDIA RTX A5000 GPU 和 Xeon(R) Platinum 8358P CPU,并以每秒帧数(FPS)为单位。

2.3 结果与分析

2.3.1 与先进的方法比较

表 1 展示了 ULRTS-Net 与其他皮肤病变分割方法在两个皮肤病变公开数据集(PH2 和 ISIC2016)的综合比较结果,包括模型分割性能、模型复杂度以及模型推理速度。本文方法的分割性能指标是在 RTX A5000 GPU 上评估的,而其他皮肤病变分割方法的 FPS 指标均在 RTX 2080Ti GPU 上进行推理速度的测试。为了保持一致,本文将 ULRTS-Net 部署到 RTX 2080Ti GPU 上进行推理速度的测试。

表 1 ULRTS-Net 与其他皮肤病变分割方法的综合比较

Dataset	Methods	Segmentation Performance			Model Complexity		Inference Speed/FPS
		JI/%	DC/%	Acc/%	Params/M	GFLOPs	
PH2	HRC-Net[13]	87.06	93.05	95.20	25	50	—
	EDLM[16]	85.34	91.81	94.52	45	51	40
	DSNet[24]	87.15	91.97	94.82	10	35	57
	MB-DCNN[17]	87.12	93.25	95.87	53	55	36
	MF2-Net[26]	89.00	**94.00**	**96.00**	1.17	—	—
	S-Unet[25]	88.81	93.02	95.92	20	37	46
	ORED-Net[21]	84.00	91.00	94.00	—	—	—
	M-SegNet[22]	87.00	93.00	**96.00**	—	—	—
	ULRTS-Net	**89.95**	92.95	95.67	**0.41**	**1.51**	**133**
ISIC 2016	SBPS[18]	84.34	90.42	94.96	42	41	41
	EIU-Net[20]	85.50	**91.90**	95.90	—	—	—
	S-Unet[25]	84.27	89.95	95.67	20	37	42
	CPFNet[19]	83.81	90.23	95.09	43	16	58
	FAT-Net[14]	85.30	91.59	**96.04**	30	23	44
	DAGAN[15]	84.42	90.85	95.82	56	62	32
	DCSAU-Net[23]	85.30	91.40	95.70	—	—	—
	ULRTS-Net	**85.78**	91.36	95.58	**0.41**	**1.51**	**133**

如表 1 所示,ULRTS-Net 分别在 ISIC2016 和 PH2 数据集上取得了最高的 85.78%JI 和 89.95%JI,相较于其他皮肤病变分割方法表现出了强大的竞争力。此外,在 RTX 2080 算力平台上进行测试,ULRTS-Net 以 133 FPS 的速度显著领先于其他皮肤病变分割方法。具体而言,ULRTS-Net 比 DSNet 快大约 2.3 倍,比 DAGAN 更是快了大约 4.1 倍。在模型复杂度方面,ULRTS-Net 仅需 0.41M 的模型参数量和 1.51 GFLOPs,同样远低于其他皮肤病变分割模型。ULRTS-Net 的模型参数量比 MF2-Net(1.17M)少了约 2.8 倍,更是比 MB-DCNN(53M)少了约 129 倍。所提方法以较低的计算成本实现快速且准确的分割,优于其他皮肤病变分割方法。

表 2 展示了 ULRTS-Net 与七个经典图像分割模型在 ISIC2016 数据集上的综合比较结果,包括模型分割性能、模型复杂度以及模型推理速度。为了确保公平比较,所有的实验都是在与 ULRTS-Net 相同的计算环境和数据增强方法下进行。其中,所有模型的分割性能均在 RTX A5000 GPU 上评估,而 FPS 指标分别在 RTX A5000 GPU 和 8358P CPU 上测试。

从表 2 中的实验结果可以看出,ULRTS-Net 在分割精度和 GFLOPs 方面相较于七个经典图像分割模型都表现出明显的优势。其次,在模型参数量方面,ULRTS-Net 虽略多于实时分割模型 ENet,但远少于其余六个分割模型。最后,ULRTS-Net 在 RTX A5000 GPU 上的推理速度明显快于所有经典模型;而在低算力平台 8358P CPU 上,ULRTS-Net 的推理速度虽然低于 ENet,但远快于其余六个经典分割模型。

表 2　ULRTS-Net 与七个经典图像分割模型的综合比较

Methods	Segmentation Performance			Model Complexity		Inference Speed/FPS	
	JI/%	DC/%	Acc/%	Params/M	GFLOPs	RTX A5000	8358P
U-Net[10]	82.61	89.47	94.86	31	35	68	1.22
ENet[22]	84.02	90.42	95.12	**0.34**	3.78	125	**24.31**
SegNet[34]	82.80	89.81	94.66	7	359	67	1.76
UNet++[35]	83.27	90.02	95.11	37	106	43	0.48
Att-UNet[36]	82.70	89.68	94.86	35	51	66	1.05
UNet3+[37]	83.71	90.46	95.24	27	152	36	0.30
U2Net[38]	84.18	90.75	95.09	44	29	37	1.11
ULRTS-Net	**85.78**	**91.36**	**95.58**	0.41	**1.51**	**169**	14.28

图 3 展示了 ULRTS-Net 和其他七种经典图像分割模型的分割效果,其中,黑色曲线是真实标签中的病灶区域边界,而灰色曲线则是各个模型分割结果的病灶区域边界。图 3 的可视化分割结果直观地展示了 ULRTS-Net 在准确分割皮肤病变区域方面有着显著优势。在对比前 3 张皮肤病变图像的各模型分割结果时,可以清晰地观察到,ULRTS-Net 在克服皮肤图像中的几种常见干扰方面表现更为出色,如黑圈、皮肤色素暗斑等,能够实现较为准确的皮肤病变分割。第 4 张皮肤病变图像的分割结果体现了 ULRTS-Net 的鲁棒性,在处理大小和形状呈现极端变化的皮肤病变区域时,分割结果更接近真实标签。

2.3.2　消融实验

表 3 展示了对 ULRTS-Net 中的两个注意力模块进行消融实验的结果。在实验 1 中,去掉了 ULRTS-Net 中的两个注意力模块,作为基线模型与后续实

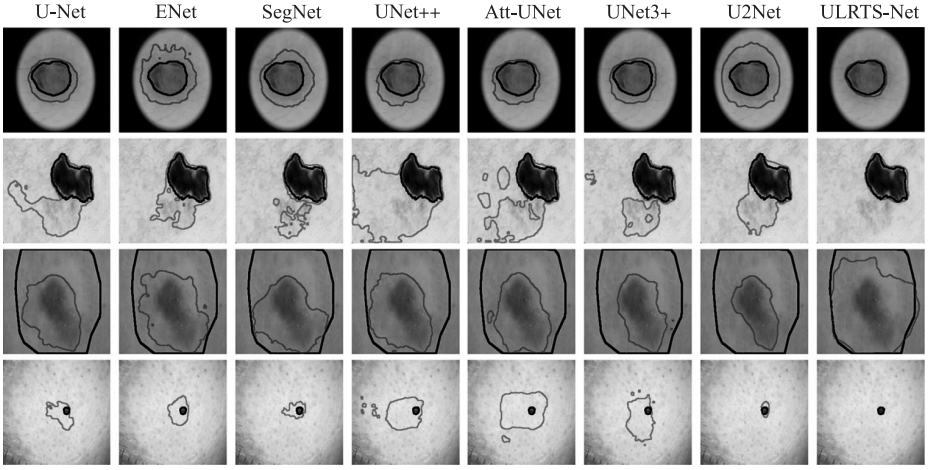

图 3　不同主流分割网络的分割可视化结果比较

验进行对照。接着,实验 2 和实验 3 在实验 1 的基础上分别添加 SA 和 CA。实验 4 交换了 ULRTS-Net 中两个注意力模块的位置,而实验 5 则将两个注意力模块添加在本文设计的位置上。所有消融实验的指标都是在 RTX A5000 GPU 上进行评估和测试。

实验 1~3 的结果表明,与基线模型相比,在 ULRTS-Net 中分别添加 SA 或 CA 后都取得了不同程度的性能提升,并且只增加了相当有限的计算成本和推理时间。其中,$DSCM_2$ 的输出特征中包含着大量的背景噪声,这些噪声可能会干扰后续的特征提取和解码。在 $DSCM_2$ 后添加 SA 有助于过滤这些噪声,并增强模型对重要特征的学习能力。将 CA 放在 $DSCM_3$、MSFFM-3 和 MSFFM-5 之后,三个子模块的输出特征经过跨通道融合,再通过 CA 重新进行通道加权,从而减少因压缩通道数而丢失的重要通道信息。同时,实验 4 和 5 的结果进一步证明了本文选择的注意力安放位置是合理的。对比实验 2、实验 3 和实验 5 的结果,在 ULRTS-Net 中同时使用这两个注意力能够获得最大的性能提升。

表 3 ULRTS-Net 中注意力的消融实验结果

Method	Attention in ULRTS-Net	JI/%	Params/M	GFLOPs	FPS
1	None	84.20	0.405	1.51	192
2	SA	84.89	0.405	1.51	187
3	CA	85.23	0.407	1.51	184
4	SA+CA	84.55	0.406	1.51	178
5	CA+SA	85.78	0.407	1.51	169

表 4 展示了对 ULRTS-Net 中的三个核心模块进行消融实验的结果。在实验 1 中,ULRTS-Net 编码器和跳跃连接处的所有子模块都使用了两个连续的卷积核大小为 3×3 的卷积组合(3×3 Conv-BN-ReLU-3×3 Conv-BN-ReLU),简记为 DC。在实验 2 中,ULRTS-Net 编码器和跳跃连接处的所有子模块均使用 DSCM 来提取特征。实验 3 在实验 2 的基础上,将编码器最后两个子模块替换成 MSFFM-3。而实验 4 在实验 3 的基础上,将编码器最后一个 MSFFM-3 子模块替换成 MSFFM-5。实验 5 则是在实验 4 的基础上,将跳跃连接处的子模块替换为 ML-SFFM。实验 1 和 2 的结果表明,相比于常见的 DC 组合,提出的 DSCM 模块虽然在一定程度上降低了分割性能,但是显著降低 ULRTS-Net 的模型复杂度。

对比实验 2 和实验 3 的结果可以看出,MSFFM-3 子模块虽然增加了一定的计算成本,但却在分割精度上取得显著提升。这是因为其内部的 SASPP 模块能够有效扩大模型的感受野,提取到丰富的多尺度特征信息,使模型能够学习到上下文信息,从而提高模型的分割性能。实验 3 和 4 的结果证明了 MSFFM-5 的有效性,其内部的膨胀卷积具有更大的膨胀率,为 ULRTS-Net 提供了更大的感受野,从而带来更好的性能增益。实验 4 和实验 5 的结果证明了 ML-SFFM 的有效性,该模块通过对编码器的浅层特征进行语义强化,缩小了与深层特征之间的语义差距,从而有效地提高了 ULRTS-Net 的分割性能。

表 4 ULRTS-Net 中三个核心模块的消融实验结果

Method	Sub-modules in ULRTS-Net	Skip	JI(%)	Params(M)	GFLOPs	FPS
1	DC×5	DC	84.11	0.959	4.21	245
2	DSCM×5	DSCM	83.59	0.107	0.56	230
3	DSCM×3+MSFFM-3×2	DSCM	84.63	0.406	1.50	188
4	DSCM×3+MSFFM-3+MSFFM-5	DSCM	85.48	0.406	1.50	175
5	DSCM×3+MSFFM-3+MSFFM-5	ML-SFFM	85.78	0.407	1.51	169

3 讨论

经典轻量级图像分割模型 ENet 被应用于皮肤病变图像分割中,但其在皮肤病变图像上的分割性能相较于目前主流的皮肤病变分割方法仍有明显差距。此外,现有基于深度学习的皮肤病变分割方法往往存在实时推理速度较慢的问题。针对以上方法的不足,本文提出了一种新颖的实时皮肤病变分割网络,旨在以较低的模型复杂度实现快速而准确的分割,并在公开数据集上验证了该模型的实时性和准确性。与其他先进的皮肤病变分割方法相比,ULRTS-Net 在模型复杂度和推理速度上具有显著的优势,并保持了有竞争力的皮肤病变分割精度。此外,与常见的经典分割模型相比,ULRTS-Net 在处理皮肤病变图像中常见的几种干扰时,展现出更高的可靠性和鲁棒性,其分割结果更接近真实标签。最

终,消融实验进一步验证了 ULRTS-Net 各个模块的有效性以及改进方法的合理性。

4　总结与展望

为了改善皮肤病变分割模型分割精度差、模型复杂度高和推理速度慢的问题,提出了一种超轻量实时分割网络。该模型采用了相对轻量的编解码网络架构,并使用深度可分离卷积代替了标准卷积,大大降低了模型复杂度以及计算量。其次,为了缩小浅层特征与深层特征之间的语义差距,在跳跃连接处设计了一个多级语义特征融合模块。此外,提出了多尺度特征融合模块,通过增加模型的感受野来增强模型对上下文信息的学习能力。最后,通过引入了空间注意力和通道注意力,进一步增强了模型对重要特征的学习能力。

本研究的局限性在于:尽管 ULRTS-Net 在降低模型复杂度和提升推理速度方面取得了显著进展,实现了精准的皮肤病变图像分割效果,但其在 CPU 上的推理速度仍有提升空间。未来的研究将继续考虑对模型进行更深入的轻量化改进,在不牺牲分割精度的前提下,提高模型在 CPU 上的推理速度。

参考文献:

[1] 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.

[2] Burroni M,Corona R,Dell'Eva G,et al. Melanoma Computer-Aided Diagnosis:Reliability and Feasibility Study[J]. Clinical Cancer Research,2004,10(6):1881-1886.

[3] 张杰,赵惠军,李贤威,等. 基于深度学习方法检测皮肤镜图像中黑色素瘤的研究进展与展望[J]. 医疗卫生装备,2018,39(11):90-95.

[4] 王守印,陈健,万佳泽,等. 基于跨层非局部融合和 DeepLabV3+的 PCB 图像分割算法[J]. 电子器件,2024,47(6):1547-1555.

[5] 张翠萍,杨善超. 叶类中药显微图像的阈值分割和自动分类[J]. 计算机工程与科学,2012,34(11):114-119.

[6] Tang H,Wu E X,Ma Q Y,et al. MRI Brain Image Segmentation by Multi-Resolution Edge Detection and Region Selection[J]. Computerized Medical Imaging and Graphics,2000,24(6):349-357.

[7] Emre Celebi M,Kingravi H A,Iyatomi H,et al. Border detection in Dermoscopy Images Using Statistical Region Merging[J]. Skin Research and Technology,2008,14(3):347-353.

[8] Xie F,Fan H,Li Y,et al. Melanoma Classification on Dermoscopy Images Using a Neural Network Ensemble Model[J]. IEEE Transactions on Medical Imaging,2016,36(3):849-858.

[9] Fan H D,Xie F Y,Li Y,et al. Automatic Segmentation of Dermoscopy Images Using Saliency Combined with Otsu Threshold[J]. Computers in Biology and Medicine,2017,85:75-85.

[10] Glaister J,Wong A,Clausi D A. Segmentation of Skin Lesions from Digital Images Using Joint Statistical Texture Distinctiveness[J]. IEEE Transactions on Biomedical Engineering,2014,61(4):1220-1230.

[11] Long J,Shelhamer E,Darrell T. Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,39(4):640-651.

[12] Ronneberger O,Fischer P,Brox T. U-Net:Convolutional Networks for Biomedical Image Segmentation[C]//18th International Conference on Medical Image Computing and Computer-Assisted Intervention,Munich,Germany,Springer,2015:234-241.

[13] Xie F,Yang J,Liu J,et al. Skin Lesion Segmentation Using High-Resolution Convolutional Neural Network[J]. Computer Methods and Programs in Biomedicine,2020,186:105241.

[14] Wu H,Chen S,Chen G,et al. FAT-Net:Feature Adaptive Transformers for Automated Skin Lesion Segmentation[J]. Medical Image Analysis,2022,76:102327.

[15] Lei B,Xia Z,Jiang F,et al. Skin Lesion Segmentation via Generative Adversarial Networks with Dual Discriminators[J]. Medical Image Analysis,2020,64:101716.

[16] Goyal M,Oakley A,Bansal P,et al. Skin Lesion Segmentation in Dermoscopic Images with Ensemble Deep Learning Methods[J]. IEEE Access,2019,8:4171-4181.

[17] Xie Y,Zhang J,Xia Y,et al. A Mutual Bootstrapping Model for Automated Skin Lesion Segmentation and Classification[J]. IEEE Transactions on Medical Imaging,2020,39(7):2482-2493.

[18] Lee H J,Kim J U,Lee S,et al. Structure Boundary Preserving Segmentation for Medical Image with Ambiguous Boundary[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition,Seattle,WA,USA,IEEE,2020:4816-4825.

[19] Feng S,Zhao H,Shi F,et al. CPFNet:Context Pyramid Fusion Network for Medical Image Segmentation[J]. IEEE Transactions on Medical Imaging,2020,39(10):3008-3018.

[20] Yu Z,Yu L,Zheng W,et al. EIU-Net:Enhanced Feature Extraction and Improved Skip Connections in U-Net for Skin Lesion Segmentation[J]. Computers in Biology and Medicine,2023,162:107081.

[21] Naqvi R A,Hussain D,Loh W K. Artificial Intelligence-Based Semantic Segmentation of Ocular Regions for Biometrics and Healthcare Applications[J]. Computers,Materials & Continua,2021,66(1):715-732.

[22] Yamanakkanavar N, Lee B. A Novel M-SegNet with Global Attention CNN Architecture for Automatic Segmentation of Brain MRI[J]. Computers in Biology and Medicine,2021,136:104761.

[23] Xu Q, Ma Z, 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.

[24] Hasan M K,Dahal L,Samarakoon P N,et al. DSNet:Automatic Dermoscopic Skin Lesion Segmentation[J]. Computers in Biology and Medicine,2020,120:103738.

[25] Tang P,Liang Q,Yan X,et al. Efficient Skin Lesion Segmentation Using Separable-Unet with Stochastic Weight Averaging[J]. Computer Methods and Programs in Biomedicine,2019,178:289-301.

[26] Yamanakkanavar N,Lee B. MF2-Net:A Multipath Feature Fusion Network for Medical Image Segmentation[J]. Engineering Applications of Artificial Intelligence, 2022, 114:105004.

[27] Paszke A,Chaurasia A,Kim S,et al. ENet:A Deep Neural Network Architecture for Real-Time Semantic Segmentation [EB/OL]. (2016-06-07). https://arxiv.org/abs/1606.02147v1

[28] 巫笠平,段晓鹏,马玉良等.基于改进 U-Net 网络的多囊卵巢图像分割方法[J]. 中国生物医学工程学报, 2022,41(6):663-671.

[29] Ma Y,Wu L,Gao Y,et al. ULFAC-NET:Ultra-Lightweight Fully Asymmetric Convolutional Network for Skin Lesion Segmentation[J]. IEEE Journal of Biomedical and Health Informatics,2023,27(6):2886-2897.

[30] Codella N C F,Gutman D,Emre Celebi M,et al. Skin Lesion Analysis Toward Melanoma Detection:A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, Hosted by the International Skin Imaging Collaboration (ISIC) [EB/OL]. (2018-01-08). https://arxiv.org/abs/1710.05006v3.

[31] Mendonça T, Ferreira P M, Marques J S, et al. PH2-A Dermatoscopic Image Database for Research and Benchmarking[C]//2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),Osaka,Japan,IEEE,2013:5437-5440.

[32] Chen L C,Papandreou G,Kokkinos I,et al. Deeplab:Semantic Image Segmentation with Deep Convolutional Nets,Atrous Convolution,and Fully Connected Crfs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,40(4):834-848.

[33] Berman M,Triki A R,Blaschko M B. The Lovász-Softmax Loss:A Tractable Surrogate for the Optimization of the Intersection-Over-Union Measure in Neural Networks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition,Salt Lake City,UT,USA,IEEE,2018: 4413-4421.

[34] Badrinarayanan V,Kendall A,Cipolla R. Segnet:A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(12):2481-2495.

[35] Zhou Z, Rahman Siddiquee M M, Tajbakhsh N, et al. UNet++:A Nested U-Net Architecture for Medical Image Segmentation[C]//4th International Workshop on Deep Learning in Medical Image Analysis and 8th International Workshop on Multimodal Learning for Clinical Decision Support,Held in Conjunction with 21st Medical Imaging and Computer-Assisted Intervention, Granada, Spain, Springer,2018:3-11.

[36] Oktay O, Schlemper J, Le Folgoc L, et al. Attention U-Net:Learning Where to Look for the Pancreas[EB/OL]. (2018-05-20). https://arxiv.org/abs/1804.03999. Preprint arXiv:1804.03999,2018.

[37] Huang H, Lin L, Tong R, et al. Unet 3+:A Full-Scale Connected Unet for Medical Image Segmentation[C]// ICASSP 2020-2020 IEEE International Conference on Acoustics,Speech and Signal Processing(ICASSP),Barcelona,Spain,IEEE,2020:1055-1059.

[38] Qin X,Zhang Z,Huang C,et al. U2-Net:Going Deeper with Nested U-Structure for Salient Object Detection[J]. Pattern Recognition,2020,106:107404.

梅厦锦(2000—)杭州电子科技大学自动化(人工智能)学院硕士研究生,主要研究方向为生物医学图像处理、模式识别;

马玉良(1976—)杭州电子科技大学自动化(人工智能)学院教授,主要研究方向为图像处理、生物信号处理、模式识别等,mayuliang@hdu.edu.cn。