

基于 MultiResUNet-SMIS 的皮肤黑色素瘤图像分割^①



张 潮, 宋亚林, 袁明阳

(河南大学 软件学院, 开封 475004)

通信作者: 宋亚林, E-mail: syl@vip.henu.edu.cn

摘 要: 针对现有的皮肤黑色素瘤病灶分割精度不高的问题, 结合现有卷积神经网络方法提出皮肤黑色素瘤图像分割方法 MultiResUNet-SMIS. 首先, 依据皮肤黑色素瘤成像特点, 引入不同空洞率的空洞卷积替换普通卷积, 在参数量相同的前提下扩大感受野, 使网络模型能够适用于多尺度病灶分割任务; 其次加入空间和通道注意力机制以重新分配特征权重, 扩大感兴趣特征影响, 抑制无关特征; 最后融合 Focal loss 与 Dice loss 提出一种新的 loss 函数 FD loss 用于计算回归损失, 解决前景背景像素不均衡问题, 进一步提高网络模型的分割精度. 实验结果表明, MultiResUNet-SMIS 在 ISIC-2018 数据集上的 *Dice* 指数、*IoU* 指数以及 *Acc* 准确率分别达到了 89.47%、82.67%、96.13%, 与原 MultiResUNet 以及 UNet、UNet++、DeepLab V3+ 等主流方法相比, MultiResUNet-SMIS 在皮肤黑色素瘤图像分割中具有更好的效果.

关键词: 皮肤黑色素瘤; 图像分割; 注意力机制; 空洞卷积; 损失函数

引用格式: 张潮, 宋亚林, 袁明阳. 基于 MultiResUNet-SMIS 的皮肤黑色素瘤图像分割. 计算机系统应用, 2023, 32(6): 221-230. <http://www.c-s-a.org.cn/1003-3254/9113.html>

Skin Melanoma Image Segmentation Based on MultiResUNet-SMIS

ZHANG Chao, SONG Ya-Lin, YUAN Ming-Yang

(School of Software, Henan University, Kaifeng 475004, China)

Abstract: In order to address the problem of low accuracy of skin melanoma lesion segmentation in existing image segmentation methods, a MultiResUNet-SMIS method is proposed based on existing convolution neural network methods. Firstly, according to the imaging characteristics of skin melanoma, the dilation convolution with different dilation rates is introduced to replace the normal convolution, and the receptive field is expanded on the premise of the same parameters so that the model can segment the lesion at multiple scales. Secondly, spatial and channel attention mechanisms are added to the model to redistribute feature weights, expand the influence of features of interest, and suppress irrelevant features. Finally, by combining Focal loss with Dice loss, a new loss function, i.e., FD loss, is proposed to calculate the regression loss and solve the problem of unbalanced foreground and background pixels, so as to further improve the segmentation accuracy of the network model. The experimental results show that *Dice*, *IoU*, and *Acc* of MultiResUNet-SMIS on the ISIC-2018 dataset have reached 89.47%, 82.67%, and 96.13%, respectively, which are better than the original MultiResUNet and mainstream methods such as UNet, UNet++, and DeepLab V3+ in skin melanoma image segmentation.

Key words: skin melanoma; image segmentation; attention mechanism; dilation convolution; loss function

① 基金项目: 河南省科技研发项目 (212102210078); 河南省重点研发与推广专项 (科技攻关)(202102210380)

收稿时间: 2022-11-15; 修改时间: 2022-12-23; 采用时间: 2023-01-06; csa 在线出版时间: 2023-03-17

CNKI 网络首发时间: 2023-03-20

皮肤黑色素瘤是一种高度恶性肿瘤,由黑色素细胞病变形,常见于成人,极少见于儿童。近年来,皮肤黑色素瘤的患病率呈现上升趋势,已经成为严重的公共卫生问题^[1]。皮肤黑色素瘤的致死率较高且常发生转移,因此早期诊断和治疗具有十分重要的意义。

皮肤黑色素瘤诊断方法包括 ABCD 法、七分列表法、三色法等^[2],通过观察颜色、形状、大小等主观规律判断在临床诊断中占主导地位。早期病灶色素沉积较为均匀,边界清晰且大多较小,依靠规律诊断易造成误诊、漏诊。

随着计算机辅助医学图像分割技术的发展,减少了因主观规律与经验造成的漏诊、误诊,医学图像分割成为热门研究内容。Soler 等人用阈值法分割肝脏和肿瘤区域,通过对比图像灰度直方图确定病变强度^[3]; Ben-Dan 等人用直方图和统计分布函数及方差最小化方法分割图像的病变轮廓^[4]; Dunn 提出模糊 C 均值聚类,通过计算像素点到聚类中心的加权灰度方差进行医学图像分割任务^[5]。虽然传统图像分割方法早期被广泛应用于医学图像分割,但计算复杂,需要依赖繁琐、准确的特征提取过程,对专业知识的掌握要求高且难度大,因此传统医学图像分割方法近年来并未有显著的发展。

近年来,卷积神经网络具有良好的泛化能力和特征提取能力,能够准确分割不同大小、形状和位置的目标,因此基于卷积神经网络的医学图像分割方法成为目前主流方法^[6],广泛应用于医学图像分割领域,取得诸多研究成果。如 FCN 网络模型^[7]对图像进行像素级的分类,从而解决了语义级别的图像分割问题; SegNet 网络模型^[8]采用编码解码结构,编码部分在采样池化层中记录池化标记,解码部分通过标记还原特征信息; UNet 网络模型^[9]采用编码解码及跳跃连接的拓扑结构,能够实现在少量训练图像上进行更加精准的分割,对医学图像特征提取具有较大帮助,被广泛用在医学图像分割领域^[10]。但卷积神经网络在不断下采样的过程中会造成空间信息丢失,对于分割任务来说,每一处空间特征都十分重要,且在各类分割任务中,待分割目标大小、形状、位置都不尽相同,仅靠传统的卷积神经网络方法不能够精确分割目标,因此需要对卷积神经网络分割方法进行改进,增强特征提取能力,提高分割精度。为扩大网络模型感受野,提高多尺度提取特征能力,DeepLab V3 网络模型^[11]使用空洞卷积,扩大了

网络感受野,适用于多尺度分割目标,但大空洞率的并列卷积层会产生特征图栅格化,造成病灶空间位置信息丢失; DeepLab V3+网络模型^[12]在 V3 基础上引入编码解码结构,解码部分采用 4 倍上采样的策略,但大的采样倍数仍然不利于恢复空间特征信息;为区分待分割区域与背景特征, Attention UNet 网络模型^[13]在深层次特征融合之前引入 Attention Gate 对特征响应进行调整,提高了待分割区域特征权重,增强了网络模型对待分割区域特征提取能力,但通道特征与空间特征同等重要,需要突出有效通道特征表达。在医学图像分割中,待分割区域占图像比例通常较小,造成网络模型提取待分割区域特征较弱, V-net 提出 Dice loss^[14],通过计算预测集合与真实集合的相似度对网络模型进行优化,使网络模型对待分割区域的特征提取能力增强,提高分割精度,但病灶分割中,每一个像素点特征都十分重要,因此需要对像素点进行损失计算。

为解决上述分割方法中的不足,本文提出一种改进 MultiResUNet^[15]的皮肤黑色素瘤图像分割方法 MultiResUNet-SMIS,创新点如下。

(1) 引入不同空洞率的空洞卷积替换普通卷积,通过卷积层叠加的方式避免空洞卷积栅格化造成的空间信息丢失,增大网络的感受野以多尺度提取病灶特征,使网络模型能够分割不同大小、形状的病灶,提高网络模型分割精度。

(2) 加入通道与空间双注意力模块 att-Block,在深层次特征融合之前重新分配通道和空间权重,增强相关空间和通道特征响应,抑制无关特征,提高分割精度。

(3) 提出一种新的损失函数 FD loss,结合 Focal loss^[16]像素层面计算回归损失的特点与 Dice loss^[14]在集合层面计算回归损失的特点,从像素点到集合层面计算回归损失,优化网络模型提取病灶特征能力,进一步提高分割精准度。

1 MultiResUNet 结构

UNet 是一种带有跳跃连接的编码解码结构卷积神经网络,在提取和重建图像特征方面具有良好表现,编码结构对输入图像进行卷积与下采样操作,不断提取图像特征,随着网络模型层数不断叠加,图像分辨率越来越小。解码结构对编码结构得到的低分辨率图像进行卷积与上采样操作,将特征还原到与输入图像相同分辨率大小作为结果输出。此外 UNet 通过跳跃连

接,将分辨率大小相同的编解码特征拼接防止不断下采样造成的特征丢失。

MultiResUNet 网络模型以 UNet 为基础. 该网络模型提出 MultiRes Block 模块取代 UNet 中堆叠的 3×3 卷积模块, 如图 1 所示. 该模块由 3 个大小为 3×3 的卷积层堆叠, 每个卷积层输出特征拼接后与残差路径输出特征相加. 卷积神经网络中, 通过卷积层堆叠可扩大感受野^[16], 但随着卷积层数增多会增加更多可学习参数, 占用过多硬件资源. 因此模块中每个卷积层的通道数为该模块输出通道数的 0.167、0.333、0.5, 在扩大感受野同时达到减少学习参数目的, 且该模块将对应层的输出通道数扩充为原通道数的 1.67 倍, 以提取更丰富的图像特征, 扩充后的通道数分别为 [51, 105, 212, 426, 853].

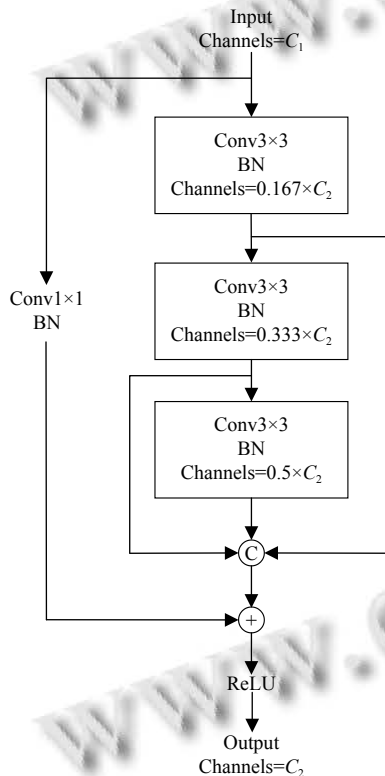


图 1 MultiRes Block 模块

同时, UNet 结构深浅层特征差异较大, 编解码结构相应层特征直接拼接导致空间信息丢失^[17]. 因此 MultiResUNet 网络在跳跃连接处加入若干个残差模块^[18] 组成 Res Path 结构, 如图 2, 用于缩小深浅层特征差异. 由于 U 型网络的特殊结构, 从上至下特征差异不断变小, 因此每层需要的残差模块个数也不同, 分别为 [4, 3, 2, 1, 0].

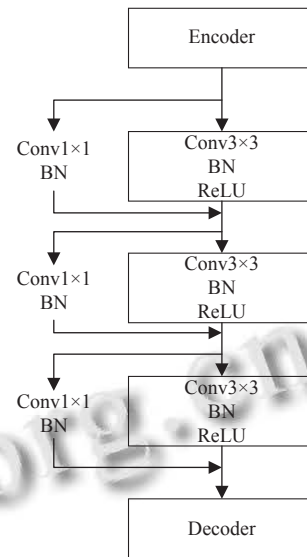


图 2 Res Path 结构

2 MultiResUNet-SMIS 方法设计

皮肤黑色素瘤图像具有病灶大小、形状等各不相同的特点, 为更好地完成分割任务, 需要考虑从多方面提取特征. 本文针对原 MultiResUNet 网络模型中存在的感受野小的问题提出改进的空洞卷积 dilation_MultiRes Block 模块设计; 针对图像中干扰因素多的问题提出双注意力机制 att-Block 模块设计, 针对皮肤黑色素瘤图像前景背景区域大小失衡问题使用新设计的损失函数 FD loss.

2.1 MultiResUNet-SMIS 总体网络结构

MultiResUNet-SMIS 网络结构如图 3, 整体结构包括编码路径和解码路径两个部分, 编码路径由改进的空洞卷积模块 dilation_MultiRes Block 和下采样模块 Maxpool 组成, 通过空洞卷积和下采样获取多尺度图像特征. 解码部分由 dilation_MultiRes Block 与上采样模块组成, 通过上采样恢复特征图的空间特征. 在编码与解码部分中间分别有数量不等的残差模块组成的 Res Path, 旨在减少深浅层特征差异. 在深浅层特征拼接之前引入双注意力机制模块 att-Block, 旨在重新分配特征权重, 突出感兴趣区域特征.

2.2 基于空洞卷积的 dilation_MultiRes Block 模块

皮肤黑色素瘤病灶大小、形状不同, 需要扩大网络模型的感受野以多尺度提取病灶特征. 本文引入空洞率不同的空洞卷积替换原网络中的普通卷积, 通过空洞卷积层叠加的方式避免空间信息丢失, 有效提升

网络模型感受野,多尺度提取图像特征,改进网络表现.

原 MultiRes Block 模块感受野大小为 3×3、5×5、7×7,但皮肤黑色素瘤图像的大小、形状、边缘复杂程度不同,原结构的感受野大小并不能够有效提取全部图像特征,需进一步提升感受野大小.卷积神经网络中多次叠加卷积核能够扩大感受野,但随着卷积核数量的增加,模型需要学习的参数量也会增加,容易出现过拟合问题.Yu 等人提出空洞卷积 (dilation convolution)^[19],在不增加参数量的同时扩大感受野,被广泛应用于密集预测任务^[20].空洞卷积与普通卷积不同,是带有空洞率的卷积,空洞率为 1 时,空洞卷积可以当作普通卷积,

当空洞率大于 1 时,卷积核以空洞率减 1 进行间隔采样.Wang 等人提出,空洞卷积易造成网格化,丢失重要特征信息^[21],因此空洞率的设定需要遵循一定的规律以避免丢失特征,规律表达式为:

$$M_i = \max \{M_{i+1} - 2r_i, M_{i+1} - 2(M_{i+1} - r_i), r_i\} \quad (1)$$

表达式需要满足条件 $M_i = r_i$ 且 $M_2 \leq k$,其中 k 为卷积核大小.例如当卷积核大小为 3×3,3 个堆叠卷积核的空洞率为 [1, 2, 5], $M_2 = 2$ 符合规律,不会丢失信息.文中将 3 个叠加卷积核空洞率分别设为 [1, 2, 5],进而增大网络感受野,改进的 dilation_MultiRes Block 结构如图 4.

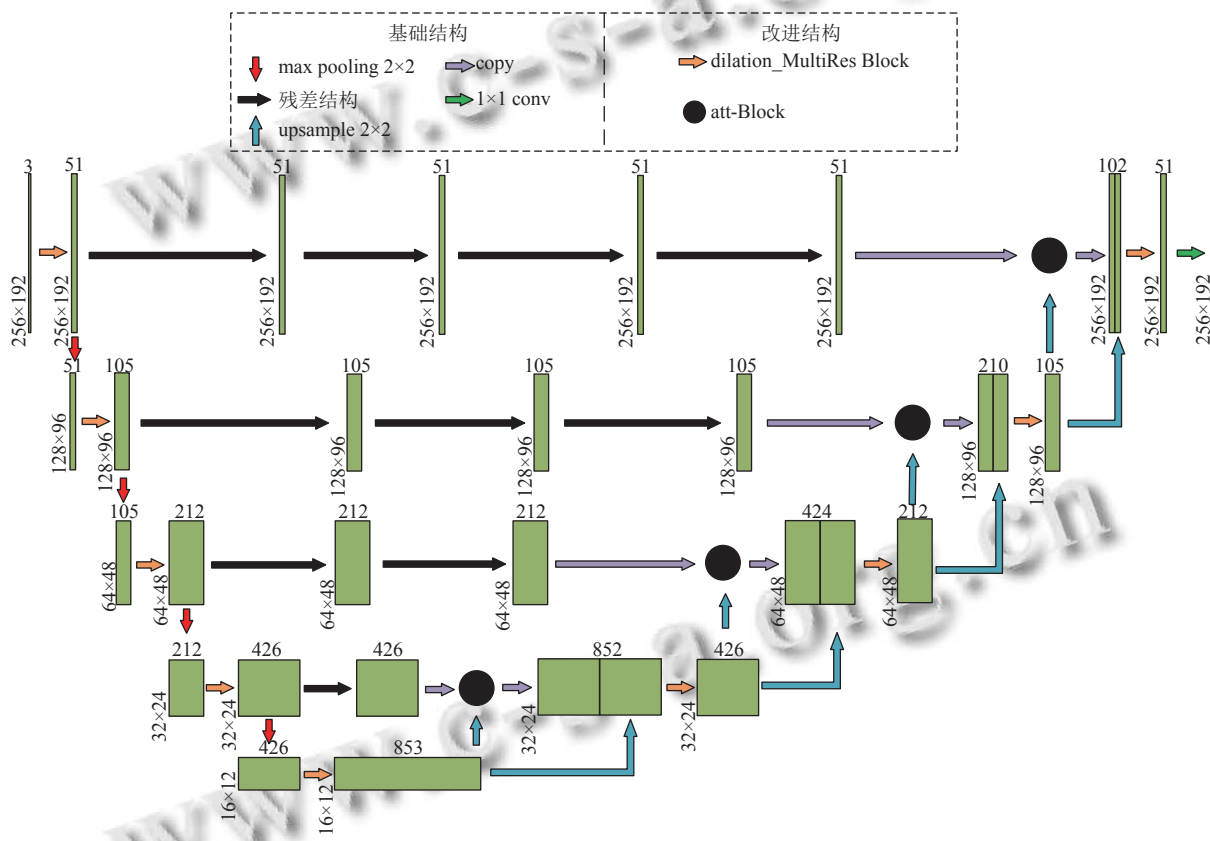


图 3 改进的 MultiResUNet 总体结构

根据感受野计算式 (2),其中 l_{k-1} 为第 $k-1$ 层感受野大小, f_k 为 k 层卷积核大小, s_i 为卷积步距,此时模块感受野大小为 3×3、7×7、17×17,在参数量一定的基础上有效扩增感受野且避免特征丢失,使网络模型多尺度提取病灶特征.

$$l_k = l_{k-1} + \left[(f_k - 1) \times \prod_{i=1}^{k-1} s_i \right] \quad (2)$$

2.3 基于通道注意力与空间注意力的 att-Block 模块

皮肤黑色素瘤图像中病灶边界模糊,对分割任务造成较大干扰,传统的卷积神经网络分割方法不能准确预测模糊边缘病灶图像,容易造成漏诊、误诊.因此需要扩大病灶特征的特征响应,抑制干扰特征表达.文中提出一种双注意力机制 att-Block,通过重新分配空间特征权重与通道特征权重,增强病灶特征响应以区分病灶与背景区域.实验证明,该方法能够减小病灶模

糊边缘、气泡等因素对分割任务的干扰,提高分割精度.

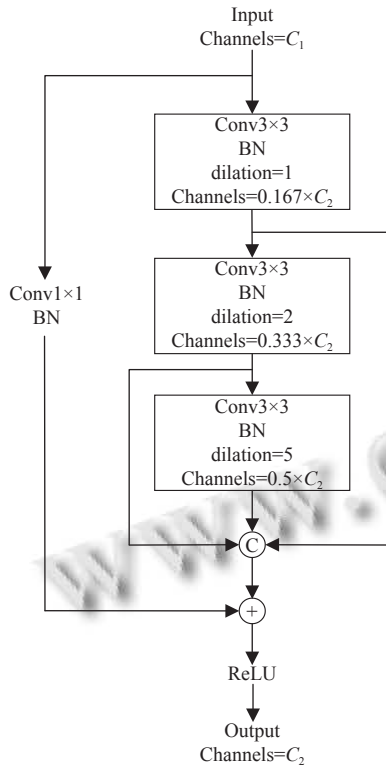


图4 改进的 MultiRes Block

att-Block 模块由空间注意力 Attention Gate 模块^[13]与通道注意力 AAM 模块^[22]组成,在深浅层特征拼接之前重新分配特征权重,增强感兴趣区域特征响应,抑制无关特征表达.

皮肤黑色素瘤图像中存在背景特征,易对待分割区域产生干扰,造成分割准确度不高,因此需要加强待分割区域特征权重,抑制背景特征表达以提高网络模型学习能力. Attention Gate 能够识别显著图像区域,对特征响应进行调整.此结构利用语义较为丰富的深层特征对浅层特征进行引导,通过深浅层特征融合对浅层特征逐像素产生权重系数以加强待分割区域特征响应,且此模块适用于端到端训练,通过反向传播不断更新权重系数,使分割结果更精确.模块如图5所示.

此模块有两个输入,分别为深层特征 y 与浅层特征 x . 特征 y 与特征 x 首先经过 1×1 卷积统一特征分辨率大小与通道数并求和,之后对求和结果进行 ReLU 函数(式(3))激活并再次进行卷积,将特征通道数压缩为1,最后特征经过 Sigmoid 函数(式(4))激活得到注

意力系数 α , 过程表达式为:

$$\sigma_1(x) = \max(0, x) \quad (3)$$

$$\sigma_2(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

$$q(x, y) = \psi(\sigma_1(W_x x + W_y y + b_g)) + b_\psi \quad (5)$$

$$\alpha = \sigma_2(q(x, y)) \quad (6)$$

其中, W_x 与 W_y 为 1×1 卷积, b_g 与 b_ψ 为偏置项, ψ 为卷积操作. 之后将注意力系数 α 与特征 x 逐像素相乘,得到输出 Out_1 , 方法过程描述如算法1.

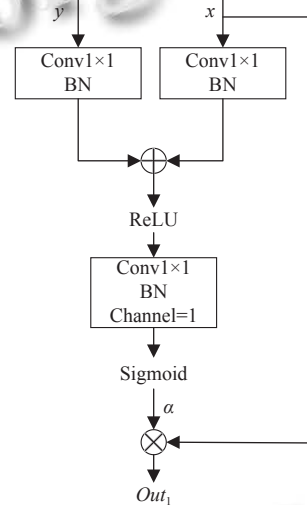


图5 Attention Gate 结构

算法1. Attention Gate 算法描述

输入: 深层特征 y , 浅层特征 x

输出: 加权待分割区域特征 Out_1

- 1) 通过卷积统一 y 、 x 分辨率与通道数得到 $W_x x$ 与 $W_y y$;
- 2) 将 $W_x x$ 与 $W_y y$ 进行求和;
- 3) 对求和结果进行 ReLU 函数激活;
- 4) 通过卷积将激活后的特征通道数压缩为1;
- 5) 将通道为1的特征图进行 Sigmoid 函数激活得到 α ;
- 6) 将浅层特征 x 逐像素与 α 相乘得到输出 Out_1 .

在卷积神经网络中,不同通道包含图像不同的特征信息,对通道进行选择使网络注意对分割有用的信息. AAM 结构对通道进行过滤,用语义信息更丰富的深层特征引导浅层特征,对浅层通道特征重新分配权重以增强有关通道特征响应,训练过程中通过反向传播不断更新权重系数,逐步提升分割精度,此模块也适用于端到端训练.

AAM 模块如图6所示, x 、 y 分别为浅层与深层特

征. x 、 y 首先经过全局平均池化将特征图分辨率调为 1, 得到 $g(x)$ 与 $g(y)$ 后经 ReLU 函数 (式 (3)) 激活并求和, 经 Sigmoid 函数 (式 (4)) 激活得到通道注意力系数 β , 过程表达式为:

$$\beta = \sigma_2(\sigma_1(g(x)) + \sigma_1(g(y))) \quad (7)$$

其中, $g(x)$ 与 $g(y)$ 为全局平均池化, 之后将注意力系数 β 与浅层特征 x 相乘, 结果与深层特征 y 相加得到输出 Out_2 , 方法过程描述如算法 2.

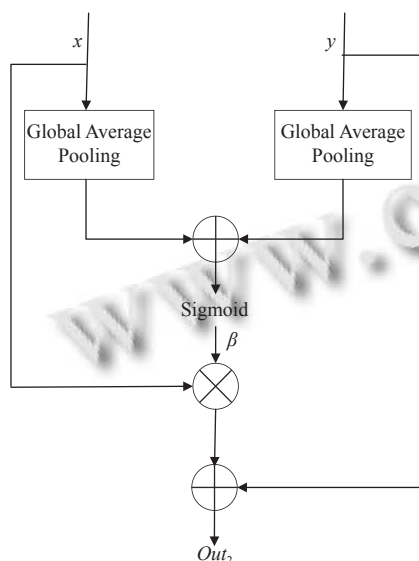


图 6 AAM 结构

算法 2. AAM 算法描述

输入: 深层特征 y , 浅层特征 x
输出: 加权通道特征 Out_2

- 1) x, y 经过全局平均池化得到全局特征 $g(x), g(y)$;
- 2) 将 $g(x), g(y)$ 经过 ReLU 函数激活;
- 3) 对激活后的特征进行求和;
- 4) 将求和后的特征经过 Sigmoid 函数激活得到系数 β ;
- 5) 将 x 与 β 相乘后与 y 相加得到输出 Out_2 .

深浅层特征分别输入 Attention Gate 模块与 AAM 模块得到输出 Out_1 与 Out_2 , 之后将两个输出进行相加, 得到重新分配权重后的特征 Out , att-Block 整体结构如图 7 所示.

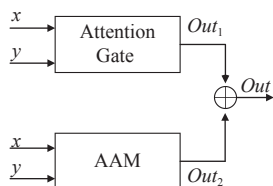


图 7 att-Block 结构

2.4 基于 Focal loss 与 Dice loss 的混合损失函数设计

皮肤黑色素瘤数据集中背景区域较前景区域大, 使得网络模型偏向于预测背景区域. 针对此问题, 本文基于 Focal loss^[16] 与 Dice loss^[14] 提出一种新的损失函数 FD loss. Focal loss 通过像素值接近程度计算损失, Dice loss 通过集合相似度计算损失, 结合两个函数, 从局部和整体优化网络模型的损失计算以提高网络模型分割精度. FD loss 定义为:

$$FDloss = \alpha l_{Focal} + \beta l_{Dice} \quad (8)$$

其中, l_{Focal} 为 Focal loss 函数, l_{Dice} 为 Dice loss 函数, α 、 β 分别为每个损失函数的系数, 且满足条件 $\alpha + \beta = 1$.

Focal loss 是一种在交叉熵损失函数基础上修改的损失函数, 表达式为:

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (9)$$

当像素预测准确时 $p_t \rightarrow 1, (1 - p_t)^\gamma \rightarrow 0$; 当像素预测不准确时 $1 - p_t \rightarrow 1, (1 - p_t)^\gamma \rightarrow 1$, 即该函数对于预测不准确像素损失没有改变, 对于预测准确的像素损失变小, 相当于增大预测不准确的像素损失权重, 使网络模型专注于难分类像素. 文中, γ 取值为 2, α_t 取值 0.25.

Dice 是一种集合相似度度量指标, 通常用于计算两个样本的相似度, Dice 系数的表达式为:

$$Dice = \frac{2 \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2} \quad (10)$$

其中, $p_i g_i$ 为预测分割图与标签之间的点乘相加运算, 在医学图像分割中, 分割图只有 0, 1 两个值, 因此 $p_i g_i$ 将存在于预测分割图而未在标签中的所有像素清零, 以惩罚低置信度的预测. Dice loss 表达式为:

$$Diceloss = 1 - Dice \quad (11)$$

3 实验结果与分析

3.1 数据集

实验所用数据集为公开数据集 ISIC-2018, 该数据集从 ISIC-2017 数据集^[23] 以及 HAM10000 数据集^[24] 中提取, 包含 2 594 张不同种类的皮肤黑色素瘤图像以及图像对应的标注文件. 实验中将数据集以比例 8:2 随机划分为训练集与验证集, 其中训练集有 2 076 张图像, 验证集有 518 张图像. 为方便图像输入图像并进行

训练与验证, 预处理部分将原图像与标注图像尺寸统一为 256×192 , 并将原图像像素值除 255 统一到 $[0, 1]$ 区间. 图 8 所示为皮肤黑色素瘤图像和标注图像, 待分割区域阈值为 1, 对应图 8(b) 白色部分, 背景区域阈值为 0, 对应图 8(b) 黑色部分.

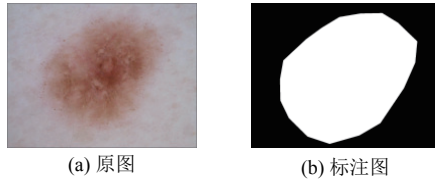


图 8 皮肤黑色素瘤图像及标注

3.2 实验参数与评价指标

实验软硬件环境为: Ubuntu 16.04 LTS; 内存 64 GB; GPU 型号 NVIDIA GeForce GTX 2080TI; CPU 型号 Intel Xeon CPU E5-2650; Python 版本 3.8.5; CUDA 版本 10.1; PyTorch 版本 1.8.0. 在实验中设置的初始学习率为 0.001, 当 *Dice* 评价指标不再上升时学习率衰减为 0.1 倍, 优化器为 Adam; batchsize 为 8, 总训练次数为 100 轮.

本文采用 *Dice* 相似度 (*Dice similarity coefficient, Dice*)、交并比 (*intersection over union, IoU*)、准确率 (*accuracy, Acc*) 作为性能评价指标, 表达式如式 (10)、式 (12)、式 (13); 用 Params 作为网络复杂度指标. 其中 *Dice* 用来衡量预测与真实区域之间的相似度, 在语义分割中常用做训练与评估, 文中作为主要性能评价参考指标.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (12)$$

其中, *A* 为预测、*B* 为标注, $|A \cap B|$ 为预测与标注的交集, $|A \cup B|$ 为预测与标注的并集.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (13)$$

其中, *TP*、*TN*、*FP*、*FN* 分别为真阳性、真阴性、假阳性和假阴性, 分别代表分割正确的病灶像素、分割正确的非病灶像素、分割错误的病灶像素和分割错误的非病灶像素.

3.3 实验结果与分析

为充分探究与分析本文提出的改进算法对皮肤黑色素瘤图像分割的有效性, 对 *dilation_MultiRes Block* 模块、*att-Block* 模块以及损失函数 *FD loss* 进行实验, 具体实验算法如算法 3.

算法 3. 实验过程

输入: 待分割图像及标注
输出: 分割图像及各项指标

- 1) $e =$ 训练轮次; $m =$ 训练图像数量; $n =$ 验证图像数量; $b =$ 批处理量
- 2) for $i=1$ to e do
- 3) for $j=1$ to m/b do
- 4) 训练网络:
- 5) 将训练图像批次输入网络训练获得预测分割图
- 6) 将预测分割图与标注进行对比计算网络损失
- 7) 根据损失计算梯度, 并更新网络参数
- 8) 验证网络:
- 9) 将 n/b 张图像输入训练好的网络得到预测分割图
- 10) 根据评价指标公式计算各项指标
- 11) end
- 12) end

3.3.1 消融实验

为更好地拟合 ISIC-2018 数据集, 文中设置了一系列实验来确定不同 α 、 β 取值对网络性能的影响. 实验结果如表 1 所示, 结果表明, 当 α 取值为 0.2、 β 取值为 0.8 时各项指标均高于其他实验. 皮肤黑色素瘤图像分割是一种逐像素二值分类任务, 图像中待分割区与背景比例失衡, 二分类交叉熵损失函数无法调整权重, 造成网络模型偏向预测背景区, 降低分割精度. *Focal loss* 通过逐像素加权计算损失, 在像素层面惩罚容易分类类别, 扩大待分割区域在损失计算中的比重, 而 *Dice loss* 通过集合计算损失, 融合两个函数, 能够在像素与集合层面计算损失, 提高分割准确度.

表 1 不同 α 、 β 取值实验结果对比

| α | β | <i>Dice</i> (%) | <i>IoU</i> (%) | <i>Acc</i> (%) |
|------------|------------|-----------------|----------------|----------------|
| 1 | 0 | 88.69 | 81.61 | 95.62 |
| 0.9 | 0.1 | 89.23 | 82.31 | 96.02 |
| 0.8 | 0.2 | 88.96 | 82.01 | 95.81 |
| 0.7 | 0.3 | 88.14 | 81.13 | 95.34 |
| 0.6 | 0.4 | 89.06 | 81.98 | 95.82 |
| 0.5 | 0.5 | 89.41 | 82.42 | 95.87 |
| 0.4 | 0.6 | 88.44 | 81.22 | 95.57 |
| 0.3 | 0.7 | 88.79 | 81.78 | 95.69 |
| 0.2 | 0.8 | 89.47 | 82.67 | 96.13 |
| 0.1 | 0.9 | 89.15 | 82.17 | 95.74 |
| 0 | 1 | 87.73 | 80.93 | 95.31 |

为了验证不同改进方法对实验结果的影响, 针对在网络中使用 *dilation_MultiRes Block*、*att-Block* 以及 *FD loss* 对网络的影响进行了实验, 各项指标如表 2, 分割图如图 9. 最终文中所提改进方法 *Dice* 指数、*IoU* 指数、*Acc* 指数结果分别为 89.47%、82.67%、96.13%, 较改进前方法均有提升.

根据表 2, 得出以下结论: 1) 网络中使用空洞卷积模块 dilation_MultiRes Block, 在参数量没有改变的情况下, Dice 指数有小幅度提升, 表明多尺度特征提取能够有效提升分割性能; 2) 在网络中使用注意力模块 att-Block, Dice 指数较未使用之前提高近 1.5 个百分点, 虽然参数量略微有所上升, 但可以显著提高网络性能, 表

明注意力机制能有效提高分割效果; 3) 在网络中同时使用两个改进模块 Dice 指数较使用单个模块有小幅提升, 表明两个模块共同作用下效果好于单个模块使用; 4) 在网络训练中使用 FD loss 损失函数 Dice 指数较使用 BCE loss 损失函数有小幅提升, 表明该损失函数能够更好地处理数据集中存在的前景背景不均衡问题。

表 2 不同模块实验结果的比较

| 方法 | MultiResUNet | dilation_MultiRes Block | att-Block | BCE loss | FD loss ($\alpha=0.2, \beta=0.8$) | Dice (%) | IoU (%) | Acc (%) | Params (M) |
|----|--------------|-------------------------|-----------|----------|-------------------------------------|--------------|--------------|--------------|--------------|
| 1 | √ | | | √ | | 86.94 | 79.37 | 94.86 | 12.33 |
| 2 | √ | √ | | √ | | 87.43 | 80.23 | 95.17 | 12.33 |
| 3 | √ | | √ | √ | | 88.46 | 81.42 | 95.51 | 12.81 |
| 4 | √ | √ | √ | √ | | 88.92 | 81.84 | 95.68 | 12.81 |
| 5 | √ | √ | √ | | √ | 89.47 | 82.67 | 96.13 | 12.81 |

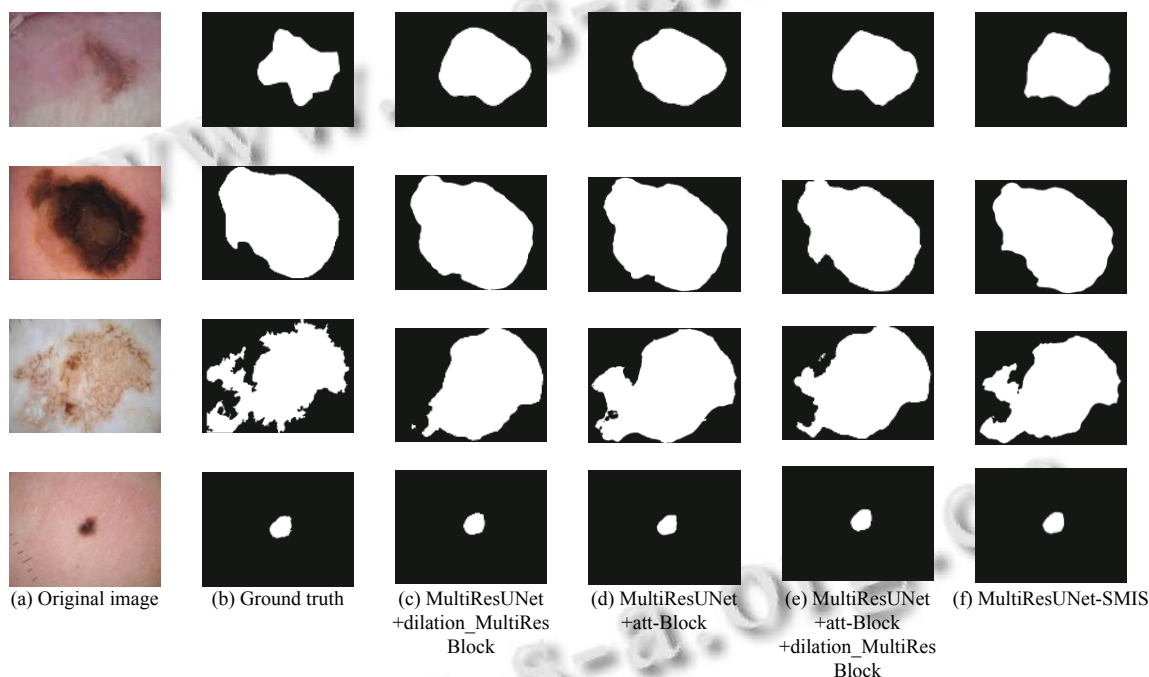


图 9 消融实验分割结果

图 9 为引入空洞卷积、att-Block 模块以及在回归损失计算中采用 FD loss 的分割效果图。图 9(c) 与图 9(d) 为分别在网络模型中引入空洞卷积与 att-Block 的分割结果图, 可直观观察到空洞卷积可一定程度上避免网络模型将背景区域误分为病灶区域, 但同时并不能完全分割整个病灶区域, att-Block 模块可帮助网络模型分割全部病灶区域, 但在边界模糊区域分割结果并不精确, 仍存在背景区域误分为病灶的情况; 图 9(e) 为在网络模型中同时引入空洞卷积与 att-Block 模块, 此时分割结果图较图 9(b) GT 相似度较高, 且对于边界模糊区域分割更为精准; 图 9(f) 为 MultiResUNet-SMIS 分

割方法, 能够直观看出分割图较 GT 最相似, 对于背景区域的模糊边界、气泡等干扰因素的过滤能力较强, 能够较为准确区分待分割区域与背景区域。

3.3.2 对比实验

在 ISIC-2018 数据集上将 UNet^[9]、DeepLab V3+^[12]、UNet++^[25]、Attention UNet^[13]、MultiResUNet 作为对比方法进行实验, 结果如表 3。MultiResUNet-SMIS 取得了最优的分割性能, Dice、IoU、Acc 指标分别达到 89.47%、82.67%、96.13%, 虽然增加了一定参数量 (Params), 但 Dice 指标较其他方法均高出 2-3 个百分点。

表3 不同方法实验结果对比

| 方法 | Dice (%) | IoU (%) | Acc (%) | Params (M) |
|--------------------------|--------------|--------------|--------------|--------------|
| UNet | 86.21 | 78.18 | 94.82 | 8.64 |
| DeepLab V3+ | 87.82 | 80.29 | 95.18 | 59.34 |
| UNet++ | 86.83 | 79.18 | 94.95 | 9.16 |
| Attention UNet | 87.85 | 80.32 | 95.23 | 8.73 |
| MultiResUNet | 86.94 | 79.37 | 94.86 | 12.33 |
| MultiResUNet-SMIS | 89.47 | 82.67 | 96.13 | 12.81 |

从表3可以看出 MultiResUNet-SMIS 分割方法较其他方法在 Dice 指数、IoU 指数、Acc 指数均为最佳,与 U 型结构网络模型 UNet、Attention UNet、UNet++ 相比,虽参数量有所增加,但 Dice 评价指标均有明显提高,与网络模型 DeepLab V3+ 相比,在参数量大幅度下降的情况下 Dice 指数仍有明显提升。

实验结果表明,对比于其他方法,MultiResUNet-SMIS 能够多尺度提取特征,保留边界信息,同时在网络中引入双注意力机制,通过更新模型参数,相关特征权重不断变大,网络模型偏向于提取病灶特征,提高分

割精度.同时通过 FD loss 损失函数,综合像素以及集合计算损失,提升网络模型分割效果。

图10为不同方法分割结果展示,图10(c) UNet、图10(e) UNet++没有对网络进行改进,不能够提取图像模糊边界特征,存在对病灶的过度分割,将部分背景区域误判为病灶区域;图10(d) DeepLab V3+对图像进行多尺度特征提取,对模糊边界较为敏感,能够提取到较多病灶边缘信息,但存在分割不充分问题,如将部分病灶区域误判为背景区域;图10(f) Attention UNet 通过注意力机制加强病灶特征表达,能够提取到较为充足的病灶特征,但对于模糊边界病灶分割效果仍然有待提高;图10(g) MultiResUNet 分割方法在病灶分割任务中表现欠佳,不能准确地分割病灶形状;图10(h) MultiResUNet-SMIS 方法能够过滤掉模糊边缘病灶等干扰因素,在大病灶分割中效果好于其他方法,分割精度最高,效果最好。

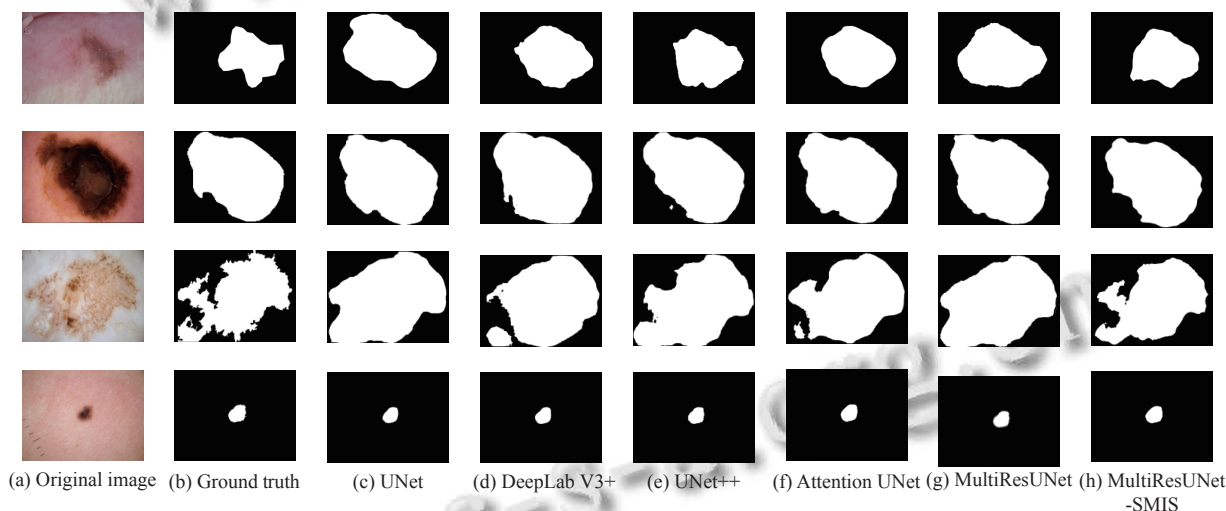


图10 不同方法分割结果展示

4 结语

针对现有图像分割方法存在的皮肤黑色素瘤病灶分割精度不高的问题,文中提出 MultiResUNet-SMIS 分割方法,使用空洞卷积扩大感受野,适用于多尺度分割任务;引入双注意力机制 att-Block 扩大相关特征影响,抑制无关特征表达,提高分割精度;提出适用于前景背景不均衡图像的分割损失函数 FD loss,通过融合 Focal loss 与 Dice loss,在像素层面与集合层面综合考虑损失计算,提高分割模型性能表现.对比实验结果表明,文中提出方法较原方法在 Dice 指数、IoU 指数及准确率评价指标均有较大提高.但实验结果也表明,改

进方法分割精度仍有较大提升空间,需要进一步改进.除此外,改进方法也可适用于彩色图像的二值分割任务,如结肠息肉分割,但应注意在分割任务中包含小目标时需调整空洞卷积率以适应分割任务,也可将 FD loss 函数应用于前景与背景区域比例失衡的分割任务.在未来研究中,本文方法可进一步得到改进与优化,提高复杂边缘图像分割精度。

参考文献

- 林千里, 张文俊, 汪汇, 等. 皮肤黑色素瘤流行病学及防治研究进展. 中国医药导报, 2019, 16(3): 28-32.

- 2 唐志铭, 傅宏阳, 荆梦晴, 等. 皮肤镜在皮肤肿瘤诊断中的应用. *皮肤病与性病*, 2021, 43(3): 352–354, 361. [doi: [10.3969/j.issn.1002-1310.2021.03.013](https://doi.org/10.3969/j.issn.1002-1310.2021.03.013)]
- 3 Soler L, Delingette H, Malandain G, *et al.* Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. *Computer Aided Surgery*, 2001, 6(3): 131–142. [doi: [10.3109/10929080109145999](https://doi.org/10.3109/10929080109145999)]
- 4 Ben-Dan I, Shenhav E. Liver tumor segmentation in CT images using probabilistic methods. *Proceedings of the 2008 Medical Image Computing and Computer Assisted Intervention Workshop*. New York: MICCAI, 2008. 41–43.
- 5 Dunn JC. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 1973, 3(3): 32–57. [doi: [10.1080/01969727308546046](https://doi.org/10.1080/01969727308546046)]
- 6 张鑫, 姚庆安, 赵健, 等. 全卷积神经网络图像语义分割方法综述. *计算机工程与应用*, 2022, 58(8): 45–57. [doi: [10.3778/j.issn.1002-8331.2109-0091](https://doi.org/10.3778/j.issn.1002-8331.2109-0091)]
- 7 Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation. *Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago: IEEE, 2015. 1520–1528.
- 8 Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481–2495. [doi: [10.1109/TPAMI.2016.2644615](https://doi.org/10.1109/TPAMI.2016.2644615)]
- 9 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.
- 10 殷晓航, 王永才, 李德英. 基于 U-Net 结构改进的医学影像分割技术综述. *软件学报*, 2021, 32(2): 519–550. [doi: [10.13328/j.cnki.jos.006104](https://doi.org/10.13328/j.cnki.jos.006104)]
- 11 Chen LC, Papandreou G, Schroff F, *et al.* Rethinking atrous convolution for semantic image segmentation. *arXiv:1706.05587*, 2017. [doi: [10.48550/arXiv.1706.05587](https://doi.org/10.48550/arXiv.1706.05587)]
- 12 Chen LC, Zhu YK, Papandreou G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 833–851.
- 13 Oktay O, Schlemper J, Le Folgoc L, *et al.* Attention U-Net: Learning where to look for the pancreas. *arXiv:1804.03999*, 2018.
- 14 Milletari F, Navab N, Ahmadi SA. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. *Proceedings of the 4th International Conference on 3D Vision*. Stanford: IEEE, 2016. 565–571.
- 15 Ibtihaz N, Rahman MS. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Networks*, 2020, 121: 74–87. [doi: [10.1016/j.neunet.2019.08.025](https://doi.org/10.1016/j.neunet.2019.08.025)]
- 16 Lin TY, Goyal P, Girshick R, *et al.* Focal loss for dense object detection. *Proceedings of 2017 IEEE International Conference on Computer Vision*. Venice: IEEE, 2017. 2999–3007.
- 17 Szegedy C, Vanhoucke V, Ioffe S, *et al.* Rethinking the inception architecture for computer vision. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 2818–2826.
- 18 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 770–778.
- 19 Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. *arXiv:1511.07122*, 2016.
- 20 Chen LC, Papandreou G, Kokkinos I, *et al.* DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834–848. [doi: [10.1109/TPAMI.2017.2699184](https://doi.org/10.1109/TPAMI.2017.2699184)]
- 21 Wang PQ, Chen PF, Yuan Y, *et al.* Understanding convolution for semantic segmentation. *Proceedings of 2018 IEEE Winter Conference on Applications of Computer Vision*. Lake Tahoe: IEEE, 2018. 1451–1460.
- 22 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.
- 23 Codella NCF, Gutman D, Celebi ME, *et al.* Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). *Proceedings of the 15th IEEE International Symposium on Biomedical Imaging*. Washington, DC: IEEE, 2018. 168–172.
- 24 Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 2018, 5: 180161. [doi: [10.1038/sdata.2018.161](https://doi.org/10.1038/sdata.2018.161)]
- 25 Zhou ZW, Siddiquee MR, 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)]

(校对责编: 孙君艳)