

# 结合 HarDNet 和反向注意力的息肉分割方法

撒子奇<sup>1</sup>, 刘巧红<sup>2\*</sup>, 凌晨<sup>2</sup>, 刘佳伟<sup>1</sup>, 刘存珏<sup>1</sup>

<sup>1</sup>上海理工大学医疗器械与食品学院, 上海 200093;

<sup>2</sup>上海健康医学院医疗器械学院, 上海 201318

**摘要** 针对结肠息肉的形状、大小、颜色和纹理多样性, 息肉与背景相似及结肠镜图像的对比度低等影响分割效果的问题, 提出了一种结合 HarDNet 和反向注意力的 U 型结构的结肠息肉图像分割网络。所提模型以 U 型的编码器解码器结构为基础架构: 首先, 编码器采用 HarDNet68 为主干网络提取特征, 以提升推理速度和计算效率; 其次, 解码器采用 3 个反向注意力模块进行边界特征的融合和细化; 最后, 在编码器和解码器之间通过感受野模块实现多尺度信息的融合, 为解码器提供更为详细的边缘信息。该模型编码器和解码器之间的迭代交互机制能够有效地校正预测结果中矛盾冲突的区域, 达到提高分割精度的目的。实验结果表明, 与现有的方法相比, 所提方法在提高分割精度的同时, 具有良好的实时性和泛化能力。研究成果可以为结肠息肉的早期筛查提供可靠的依据。

**关键词** 医用光学; 反向注意力模块; 感受野模块; HarDNet; 图像分割; 结肠息肉

中图分类号 R445.9; TP391.4; R735.3+5; TP183

文献标志码 A

DOI: 10.3788/LOP212665

## Polyp Segmentation Method Combining HarDNet and Reverse Attention

Han Ziqi<sup>1</sup>, Liu Qiaohong<sup>2\*</sup>, Ling Chen<sup>2</sup>, Liu Jiawei<sup>1</sup>, Liu Cunjue<sup>1</sup>

<sup>1</sup>College of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;

<sup>2</sup>College of Medical Instruments, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China

**Abstract** A U-shaped colon polyp image segmentation network combined with HarDNet and reserve attention is proposed with the aim of solving the problems in the diversity of shape, size, color, and texture of colon polyps, the similarity between polyps and the background, and the low contrast of colonoscopy images, which affects the segmentation effect. The proposed model is based on the U-shaped encoder-decoder structure. First, the encoder uses HarDNet68 as backbone network to extract features for improving the reasoning speed and computational efficiency. Second, the decoder uses three reverse attention modules for fusing and refining the boundary features. Finally, multi-scale information fusion is realized between encoder and decoder through a receptive field module to provide more detailed edge information for the decoder. The iterative interaction mechanism between the encoder and decoder can effectively correct conflicting regions in the prediction results, improving the segmentation accuracy. The experimental results show that compared with existing methods, the proposed method improves segmentation accuracy and also has good real-time and generalization ability. The research results can provide a reliable basis for the early screening of colonic polyps.

**Key words** medical optics; reverse attention block; receptive field block; HarDNet; image segmentation; colonic polyp

## 1 引言

结直肠癌的发病率多年来一直居世界第 3 位<sup>[1]</sup>, 是死亡率第 2 高的癌症, 占癌症总死亡人数的 9.4%<sup>[2]</sup>。大多数结直肠癌起源于腺瘤性息肉, 最初良性生长于结肠和直肠内壁, 随着时间的推移, 息肉可能向恶性转

变, 并扩散到附近的器官。结直肠癌的 5 年相对生存率从早期诊断患者的 90% 到晚期诊断患者的 14% 不等<sup>[3]</sup>, 因此早期发现和切除息肉对预防结直肠癌具有重要意义。如今, 结肠镜检查是一种有效的结直肠息肉检测技术, 并在不同社会时期的许多参考中得到建议<sup>[4]</sup>。由于息肉的不同性质, 包括其形状、大小、质地

收稿日期: 2021-10-08; 修回日期: 2021-11-10; 录用日期: 2021-11-22; 网络首发日期: 2021-12-01

基金项目: 国家自然科学基金(61801288)

通信作者: hqllqh@163.com

及静脉和照明等不可预测的因素,结肠镜检查的结果可能会产生误导。此外,结肠镜检查的结果在很大程度上取决于医生的经验和能力等各种人为因素。临床研究表明,内镜检查中结肠息肉缺失率可能在 20%~47% 之间<sup>[5]</sup>。因此,关于使用计算机辅助系统提高息肉的查出准确率并减轻医生负担的研究是非常重要的。

早期的息肉分割方法基于提取颜色、图案等特征,然后使用分类器将息肉与其周围组织区分开来。然而,手工提取特征不仅速度慢,而且存在较高的漏检率,效果并不理想。由于每个息肉的位置、形状、大小、颜色和纹理等特性各不相同,在很长一段时间实现息肉的自动准确分割是一项具有挑战性的任务。近年来,随着深度学习在图像分割领域取得了一些突破性的进展<sup>[6-7]</sup>,针对息肉图像的分割方法也逐渐受到关注并涌现出了大量基于深度学习技术的息肉分割网络。Brandao 等<sup>[8]</sup>将预训练模型和 fully convolution neural network (FCN) 结合用于识别和分割息肉,实现了像素级的预测分割息肉。Akbari 等<sup>[9]</sup>针对结肠镜镜框中息肉的形状和颜色强度不同这一点,提出了 FCN 的改进版本,结合斑块选择机制来提高息肉分割的准确性。随后编码器-解码器结构被广泛应用在生物医学图像分割中,如 U-Net<sup>[10]</sup>采用跳跃连接实现了高低层特征的融合,编码器与解码器用于捕获上下文和恢复精确细节。U-Net++<sup>[11]</sup>使用嵌套和密集的跳跃连接来减少编码器和解码器的特征映射之间的语义差距。ResUNet++<sup>[12]</sup>结合了一些先进技术,如残差计算、挤压与激发、空洞空间金字塔池化和注意力机制等,进一步改善了息肉分割的性能。然而这些方法只是对息肉区域进行了较好的处理,却导致分割的边界较为模糊。为此,SFANet<sup>[13]</sup>考虑了区域边界约束,提出了在共享编码器和两个相互约束的解码器下的选择性特征聚集结构和边界敏感损失函数,解决了边界模糊问题,但泛化性能较弱。随后提出的 DoubleU-Net<sup>[14]</sup>、ACSNet<sup>[15]</sup>、PraNet<sup>[16]</sup>和 ABCNet<sup>[17]</sup>在恢复息肉与其周围黏膜之间尖锐边界的同时更加关注到模型本身的泛化能力。另外,在结肠息肉分割精度和泛化能力不断提高的同时,模型高效的推理速度也越来越受业内人士的关注<sup>[16]</sup>。

针对息肉形状大小不一、边界不清等影响分割精度及分割网络的模型推理速度等问题,本文提出了一种结合 Harmonic DenseNet (HarDNet)<sup>[18]</sup>和反向注意力 (RA)<sup>[19]</sup>的 U 型结肠息肉图像分割网络,并命名为 HraNet。该网络以编码器-解码器结构为基本架构进行设计,编码器采用 HarDNet68<sup>[18]</sup>作为主干网络提取特征,提升模型推理速度,解码器采用 3 个平行的 RA 模块进行边界特征的融合和细化,同时编码器和解码器之间通过感受野模块 (RFB)<sup>[20]</sup>进行多尺度信息融合。所提 HraNet 具有以下几个方面的优势:1) 编

器利用 HarDNet68 作为主干,相对于以 ResNet、DenseNet 和 VGGNet 为主干的网络,HarDNet 考虑到计算量和内存访问量对模型的影响,减少了各层之间的连接数量以提高推理速度,同时增加了关键层的通道宽度,以弥补精度损失。2) 以 3 个 RA 模块作为解码器进行边界特征的融合和细化,提高了网络对息肉边界的识别能力。3) 编码器中 HarDNet68 产生的 3 个边输出<sup>[21]</sup>跳跃连接到对应 RA 模块作为其中一个输入,进一步保证学习到更多的不同尺度的特征。4) 在最底层的解码器输入端加入 RFB,获取全局特征后为解码器提供详细的边缘信息。该模型编码器和解码器之间的迭代交互机制能够有效校正预测结果中矛盾冲突的区域,从而提高分割精度。5) 相比于 PraNet, HraNet 选取轻量化的 HarDNet68 学习全局特征,在获得好的分割效果和精度的同时降低了计算复杂度,为整个结构的轻量化奠定了基础。随后通过 RFB 对深层特征进行补充,使模型更加专注于结肠息肉图像中的息肉部分,整个模型在仅使用参数为 22.8 MB 的前提下得到更为清晰的息肉边界,与 PraNet (32.5 MB) 相比减少了约 29.8%。实验结果表明,在 5 个公开的息肉图像数据集上,所提 HraNet 不仅在分割精度上优于大多数先进模型,而且具备良好的泛化能力和推理速度。

## 2 息肉分割网络框架

所提 HraNet 以传统的 U 型结构为基本框架,在网络中融合了 HarDNet68、反向注意力模块、感受野模块和边输出形成了息肉图像分割网络,HraNet 的结构如图 1 所示。首先,对输入图像进行  $3 \times 3$  卷积操作,随后经过 5 个连续的 HarDNet 集成块,HarDNet68 将原始的 DenseNet 的全局密集连接改进为稀疏连接,并以卷积 (convolution)、批标准化 (BN) 和 rectified linear unit (ReLU) 激活函数的顺序实现批量标准化的重复,在维持精度的同时,减少参数量以获得更快的推理速度。编码器的输出作为 RFB 的输入,从而捕获不同比例特征的更多视觉信息。接下来通过  $1 \times 1$  的卷积操作和上采样调整通道数和对应的特征图尺寸,并输入解码器。解码器由 3 个 RA 组成,每个 RA 以上一级的输出和 HarDNet68 中对应的边输出同时作为输入,进行取反和点积运算来获得反向注意力特征的输出,从而顺序地融合和细化互补区域和细节。此外,HarDNet68 中 3 个边输出跳跃连接到对应 RA 模块,边输出产生的损失直接反向传导到对应的卷积层,缓解了部分梯度消失,同时捕获了不同尺度的特征。解码器的输出端通过 Sigmoid 函数得到最终的分割结果。HraNet 形成了编码器和解码器之间的迭代交互机制,能够有效校正预测结果中如边界等矛盾冲突的区域从而提高分割精度,同时提升模型的推理速度和计算效率。

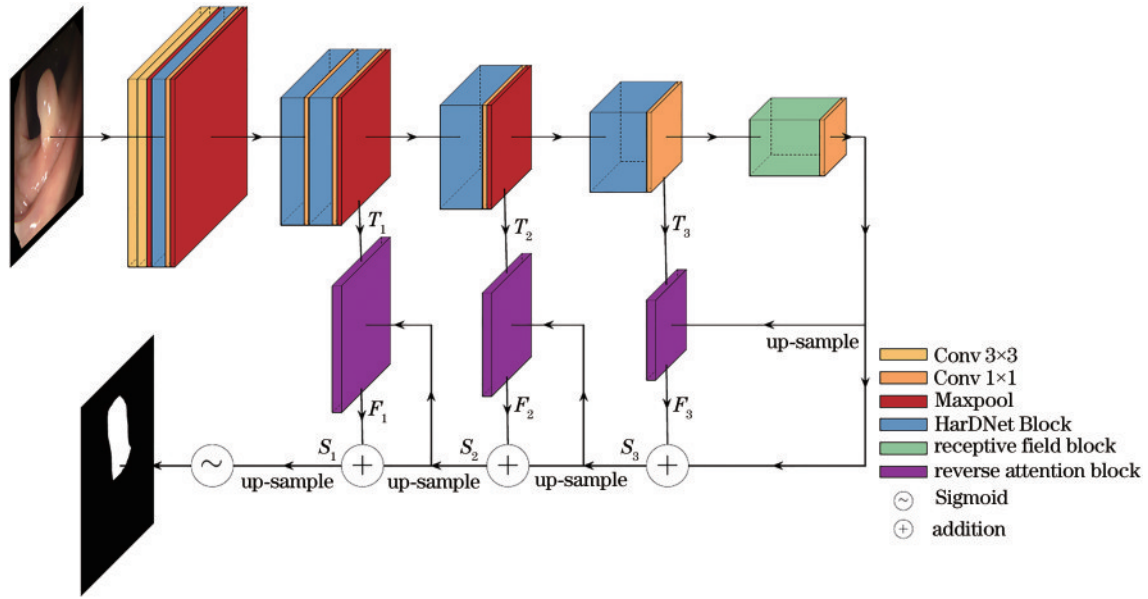


图 1 HraNet 模型结构

Fig. 1 HraNet model structure

### 2.1 HarDNet

HarDNet 也被称作谐波密集连接网络,具有高速、低计算量和内存访问量等优点。HarDNet 相较于原始密集连接的 DenseNet,主要进行了两个方面的改进: 1) 稀疏化连接。假设第  $k$  层连接到第  $k-2^n$  层,  $2^n$  可以整除  $k$ , 其中  $n$  是一个非负整数, 且  $k-2^n \geq 0$ 。在这种

连接下, 如果第  $2^n$  层被处理, 则第 1 层到第  $2^n-1$  层可以从内存中刷新。2) 加权关键层的通道数。通过在部分关键层增加通道数, 来平衡输入和输出之间的通道比率, 从而避免计算量的增加同时降低内存的访问量。HarDNet 和 DenseNet 的结构如图 2 所示。

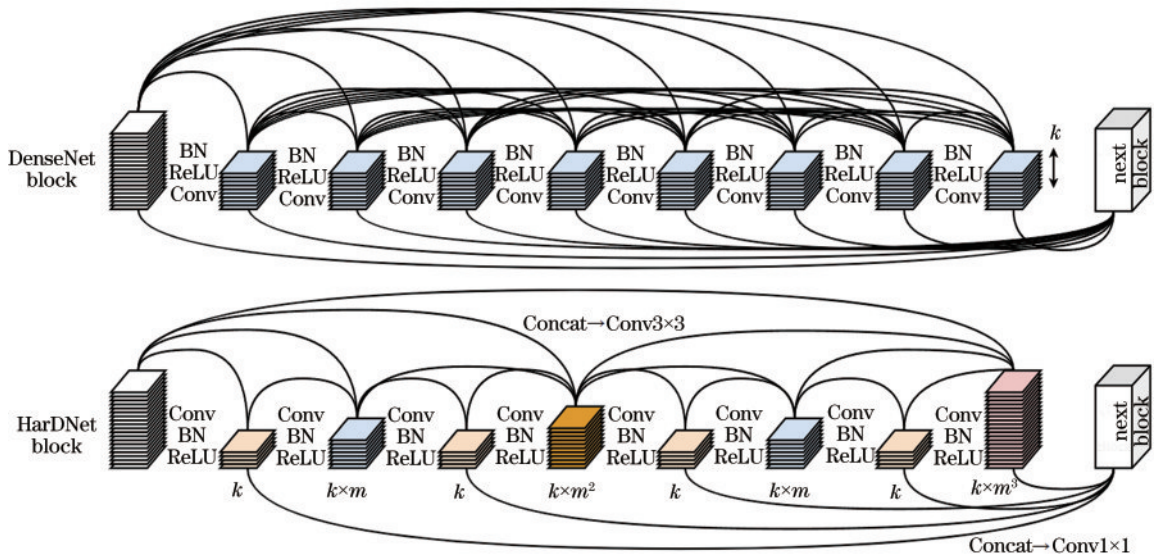


图 2 DenseNet 和 HarDNet 的结构对比图

Fig. 2 Structure comparison diagram of DenseNet and HarDNet

基于以上思想, 6 种不同参数设置的 HarDNet 被提出, 其中 HarDNet68 同样将全局密集连接更改为稀疏连接, 在下采样时采用最大池化 (MaxPool), 并将 DenseNet 的标准处理顺序 BN-ReLU-Conv 顺序转换为 Conv-BN-ReLU 的顺序, 以实现批量标准化的重复。HarDNet68 在第 8 步增加更多的层数, 以增强特

征学习能力达到有助于小尺寸目标检测的目的。事实上, 分割任务更加依赖于全局特征的学习, 因此专注于低分辨率的特征图可以获得更高的分割精度和更低的计算复杂度。所采用的 HarDNet68 的参数实施细节如表 1 所示, 其中乘法器  $m$  为低维压缩因子, “ $3 \times 3, 32$ ” 代表具有 32 个输出通道的  $3 \times 3$  卷积层, HDB 为

表 1 HarDNet68 的详细实施参数  
Table 1 Detailed implementation parameters of HarDNet68

Parameter	$m$	Stride 2	Stride 4	Stride 8	Stride 16	Stride 32
Value	1.7	$3 \times 3, 32,$ stride is 2 $3 \times 3, 64$	8(HDB), $k=14,$ $t=128$	16(HDB), $k=16, t=256$ 16(HDB), $k=20, t=320$	16(HDB), $k=40, t=640$	4(HDB), $k=160,$ $t=1024$

HarDNet 模块的个数,  $k$  为增长率,  $t$  为  $1 \times 1$  卷积过渡的输出通道个数。

### 2.2 反向注意力模块

为了进一步改善推理时间和平衡性能, 在解码器部分利用了 RA 模块, 通过 RA 对 HarDNet68 边输出所提取的高层特征进行互补操作, 增强息肉当前预测中

的互补区域和细节。整个过程采用从粗到精的思想, 当前预测是从其相邻的更深层向上采样获取的。由 RA 模块驱动的由高级特征到低级特征的补全机制最终可以将不精确和粗略的估计细化为准确而完整的边缘预测图, 图 3 为 RA 模块的实现过程。

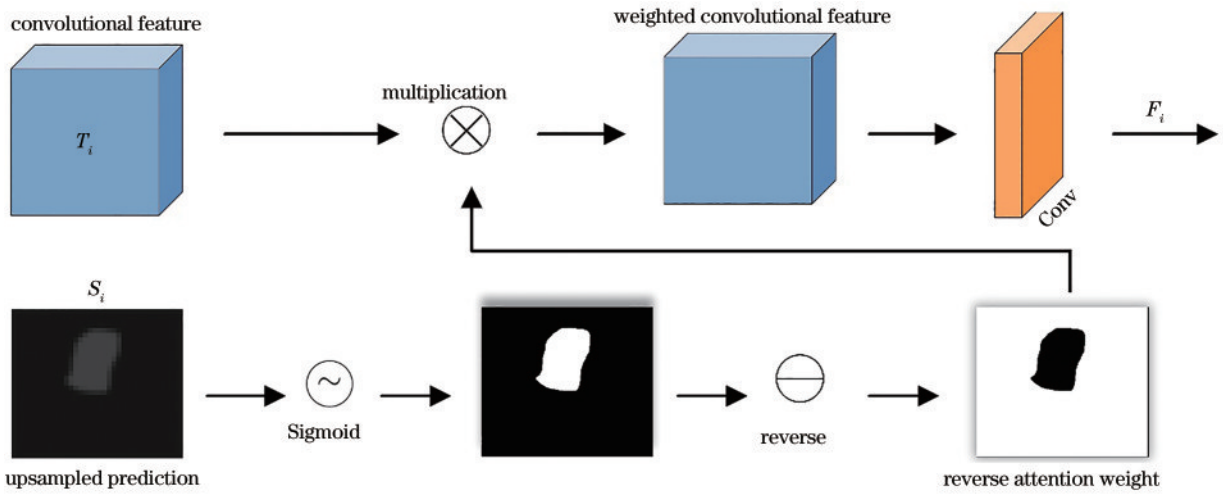


图 3 反向注意力模块

Fig. 3 Reverse attention block

解码器中的反向注意力特征图  $F_i$  通过编码器对应 HarDNet68 模块的边输出特征  $T_i$  和反向注意力权重  $A_i$  进行逐元素相乘计算得到:

$$F_i = A_i \cdot T_i, \quad (1)$$

式中: 反向注意力权重  $A_i$  由前一个 RA 模块的输出  $S_{i+1}$  进行上采样操作和 Sigmoid 计算后, 再进行取反运算获得:

$$A_i = 1 - \text{Sigmoid}(S_{i+1}^{\text{up}}). \quad (2)$$

### 2.3 感受野模块

为了进一步捕获不同比例特征的视觉信息, 所提 HraNet 将编码器的输出送入 RFB 中, 利用 RFB 模块增大感受野的特性, 获取更多的底层特征。RFB 引入了空洞卷积, 利用不同大小的卷积核模拟感受野的多尺度视角, 在不增加参数量的同时提高感受野。最后, 将不同尺寸的感受野输出特征图进行连接和  $3 \times 3$  卷积操作, 在达到融合不同特征目的的同时也获得了较好的全局特征。RFB 的输出作为解码器的输入, 提供了更加详细的边缘特征信息。所采用的 RFB 如图 4 所示。

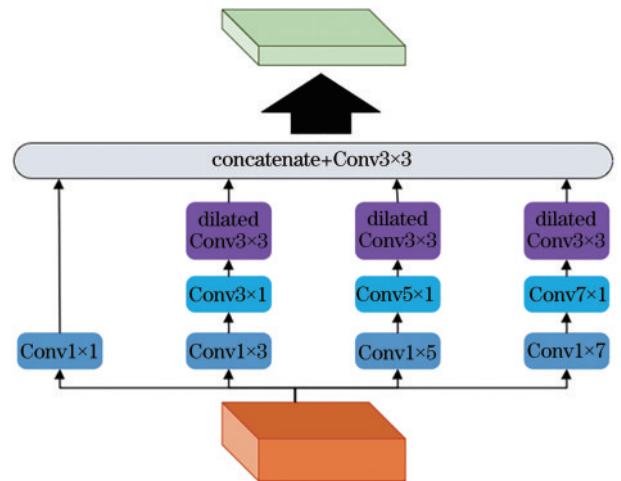


图 4 感受野模块

Fig. 4 Receptive field block

## 3 实验结果分析与讨论

### 3.1 实验环境及参数

实验过程中为了训练 HraNet, 使用了配置为

AMD EPYC 7302 3.0 GHz 16 核处理器、单张 24 GB 显存 NVIDIA GeForce RTX 3090 的计算设备,软件开发环境为 Python 3.8、Pytorch 1.9.0 和 CUDA 11.1。HraNet 的训练中,输入图像分辨率统一为  $352 \times 352$ ,损失函数使用二进制交叉熵和 Dice 损失函数的组合。在 HraNet 中,对 3 个输出 ( $S_1$ 、 $S_2$  和  $S_3$ ) 和 HarDNet68 中对应的边输出 ( $T_1$ 、 $T_2$  和  $T_3$ ) 采用深监督策略,其模型的损失函数可表示为

$$L_{\text{total}} = \sum_{i=1}^3 L(G, S_i^{\text{up}}) + \sum_{i=1}^3 L(G, T_i^{\text{up}}), \quad (3)$$

式中:  $G$  为真实标签;  $S_i^{\text{up}}$  为对 RA 模块的输出  $S_i$  进行上采样操作的结果;  $T_i^{\text{up}}$  为对 HarDNet68 模块的边输出进行上采样操作的结果。实验采用自适应矩估计 (Adam) 优化器,学习率设置为  $1 \times 10^{-4}$ ,批处理量 (batch size) 设置为 8,模型迭代次数设置为 50。

### 3.2 数据集

实验包括分割精度对比、泛化能力对比和时间性能对比,因此采用了 5 个公开的息肉图像数据集,分别为 CVC ColonDB、Endosece、ETISLarib Polyp DB、CVC Clinic DB 和 Kvasir SEG。实验中训练集由 1450 张未经任何增强的训练图像组成,包括 Kvasir SEG 中的 900 张图像和 CVC ClinicDB 中的 550 张图像。测试集由未经过训练的 100 张 Kvasir SEG、62 张 CVC ClinicDB、380 张 CVC ColonDB、60 张 Endosece

和 196 张 ETISLarib Polyp DB 数据集图像组成,其中后 3 个数据集用来检验模型的泛化能力。

### 3.3 评价指标

为了验证所提网络的可行性和有效性,采用了 6 个性能指标对模型分割结果进行评估:平均相似系数 (mDice)、平均交并比 (mIoU)、加权 Dice 度量 ( $F_{\beta}^w$ )、结构相似性度量 ( $S_a$ )、增强对齐度量 ( $E_{\phi}^{\text{max}}$ ) 和平均绝对误差 (MAE)。mDice 和 mIoU 都是衡量两个集合之间相似性的度量,在图像分割领域用来衡量网络分割结果与金标准之间的相似性,是分割网络中最常用的评价指标。 $F_{\beta}^w$  为不同位置的错误分配不同的权重,可以修正 mDice 中的“同等重要的缺陷”问题,MAE 度量用来评估像素级的精度,但是它们都是基于像素级别的评价度量体系,忽略了目标结构的相似性,所以进一步采用  $S_a$  来评估预测与人工标注真实图之间的相似性。 $E_{\phi}^{\text{max}}$  对场景中的全局信息和局部细节高度敏感,它既考虑到像素级的错误,也考虑到了图像级别的错误,是图像分割领域非常有效和高效的评价标准。

### 3.4 实验结果与分析

#### 3.4.1 分割效果

为了验证 HraNet 的性能和分割效果,将 HraNet 分别与 UNet、UNet++、SFA 和 PraNet 等网络在 Kvasir SEG 和 CVC ClinicDB 息肉图像数据集上进行了对比,最终各方法的分割性能指标如表 2 所示。

表 2 不同方法在 Kvasir SEG 和 CVC ClinicDB 数据集上分割效果对比

Table 2 Comparison of segmentation effects of different methods on Kvasir SEG and CVC ClinicDB datasets

Dataset	Method	mDice	mIoU	$F_{\beta}^w$	$S_a$	$E_{\phi}^{\text{max}}$	MAE
Kvasir SEG	UNet	0.818	0.746	0.794	0.858	0.893	0.055
	UNet++	0.821	0.743	0.808	0.862	0.910	0.048
	SFA	0.723	0.611	0.670	0.782	0.849	0.075
	PraNet	0.898	0.840	0.885	0.915	<b>0.948</b>	0.030
	HraNet	<b>0.901</b>	<b>0.845</b>	<b>0.890</b>	<b>0.917</b>	0.944	<b>0.027</b>
CVC ClinicDB	UNet	0.823	0.755	0.811	0.889	0.954	0.019
	UNet++	0.794	0.729	0.785	0.873	0.931	0.022
	SFA	0.700	0.607	0.647	0.793	0.885	0.042
	PraNet	0.899	0.849	0.896	0.936	0.979	0.009
	HraNet	<b>0.930</b>	<b>0.879</b>	<b>0.928</b>	<b>0.949</b>	<b>0.980</b>	<b>0.008</b>

由表 2 可知,除了 Kvasir SEG 数据集的  $E_{\phi}^{\text{max}}$  指标,HraNet 的其余指标均达到了最优效果,同时所有指标也达到了这 5 种方法的平均水准,表明该模型具有很强的学习能力,能够有效提升息肉图像的分割精度。由于 HraNet 结构的设计初衷是将更多的注意力集中在区域和边界,在保证模型轻量化的同时提高边缘切割精确而忽略了对全局统计信息的优化处理,在 Kvasir SEG 数据集上未能展示出最优的性能结果。

图 5 展示了不同方法在 5 个息肉数据集上的分割视觉效果对比,从图 5 可以看出,HraNet 的分割结果相对于其他几个方法更加接近于真实标签,分割效果更

为精确。HraNet 能够很好地克服息肉和背景颜色相似的问题,检测出不同形状大小和颜色的息肉组织,并且对区域和边界的划分更加清晰。

#### 3.4.2 泛化性能

以 CVC ColonDB、Endosece 和 ETISLarib Polyp DB 等 3 个息肉图像数据集作为测试集,验证模型的泛化能力,实验结果如表 3 所示。从表 3 可以看出,UNet、UNet++ 和 SFA 的泛化能力较差,尤其是 SFA 的评估指标数值急剧下降,而 HraNet 在 3 个测试集上各项指标均表现出优越的水平,其指标均达到了这 5 种方法的平均水准。

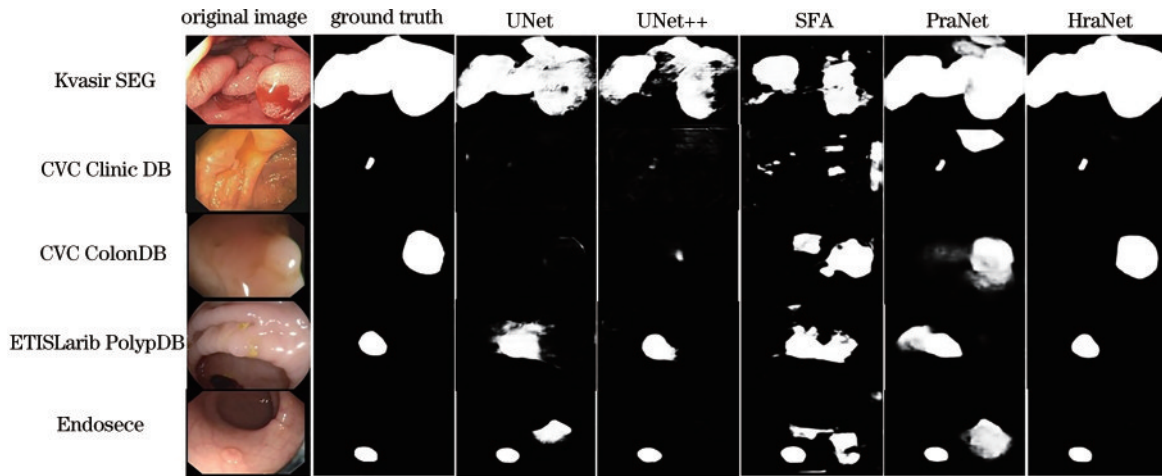


图 5 结肠息肉分割结果对比图

Fig. 5 Comparison of segmentation results of colonic polyps

表 3 泛化能力测试结果

Table 3 Generalization capability test results

Dataset	Method	mDice	mIoU	$F_{\beta}^w$	$S_a$	$E_{\phi}^{\max}$	MAE
CVC ColonDB	UNet	0.512	0.444	0.498	0.712	0.776	0.061
	UNet++	0.483	0.410	0.467	0.691	0.760	0.064
	SFA	0.469	0.347	0.379	0.634	0.765	0.094
	PraNet	0.709	0.640	0.696	0.819	<b>0.869</b>	0.045
	HraNet	<b>0.729</b>	<b>0.660</b>	<b>0.720</b>	<b>0.830</b>	0.858	<b>0.038</b>
ETISLarib Polyp DB	UNet	0.398	0.335	0.366	0.684	0.740	0.036
	UNet++	0.401	0.344	0.390	0.683	0.776	0.035
	SFA	0.297	0.217	0.231	0.557	0.633	0.109
	PraNet	0.628	0.567	0.600	0.794	<b>0.841</b>	<b>0.031</b>
	HraNet	<b>0.650</b>	<b>0.581</b>	<b>0.617</b>	<b>0.797</b>	0.827	0.035
Endosece	UNet	0.710	0.627	0.684	0.843	0.876	0.022
	UNet++	0.707	0.624	0.687	0.839	0.898	0.018
	SFA	0.467	0.329	0.341	0.640	0.817	0.065
	PraNet	0.871	0.797	0.843	0.925	<b>0.972</b>	0.010
	HraNet	<b>0.892</b>	<b>0.824</b>	<b>0.875</b>	<b>0.937</b>	0.964	<b>0.007</b>

### 3.4.3 时间性能

为了验证模型的实时分割效率和轻量化程度,以 CVC ClinicDB 数据集为例,在 AMD EPYC 7302 3.0 GHz 16 核 CPU、单张 24 GB 显存的 NVIDIA GeForce RTX 3090 GPU 上对不同方法进行训练时间和推理速度的比较,测试结果如表 4 所示。由于

HraNet 采用 HarDNet68 为模型主干,在迭代次数为 50 的情况下 33 min 可完成模型的训练,从而实现了模型的轻量化。在同一平台下,HraNet 以 ~148 frame/s 的实时速度进行结肠息肉分割,获得了优于其他模型的分割精度。高效精准的实时分割性能能够将本研究推广至需要实时应用的场合,例如结肠镜下的息肉自动

表 4 不同方法基于同一平台训练时间和推理速度的比较

Table 4 Comparison of training time and reasoning speed of different methods based on same platform

Dataset	Method	Epoch	Training /min	Inference / (frame · s <sup>-1</sup> )	mDice	mIoU
CVC ClinicDB	UNet	50	~70	~126	0.823	0.755
	UNet++	50	~80	~53	0.794	0.729
	SFA	200	>600	~70	0.700	0.607
	PraNet	50	~75	~115	0.899	0.849
	HraNet	50	~33	~148	0.930	0.879

检测、手术机器人导航等领域。

综合以上对比实验可知,所提 HraNet 多项数值指标上取得了最优的结果,同时在视觉效果上优于其他方法的分割结果,HraNet 在分割精度、实时性和泛化能力都具有显著优势。HraNet 通过 HarDNet68 学习全局特征,医学图像中小目标的分割任务更多地依赖于全局特征的学习,因此专注于低分辨率可以获得更好的分割效果、更高的精度和更低的计算复杂度。随后 RFB 增强了从轻量级 HarDNet68 主干中学习到的深层功能,利用不同尺寸大小的卷积核,在不增加参数量的同时提高感受野,加强了对息肉病灶进行特征捕捉的能力。之后通过 3 个 RA 对 HarDNet68 边输出所提取的高层特征进行互补操作,增强息肉当前预测中的互补区域和细节。该结构具备捕捉上下文信息的路径和允许精确定位的迭代交互机制,使得网络将上下文信息向更高层分辨率传播,能够有效地校正预测结果中矛盾冲突的区域,很好地克服了息肉的形状、大小、颜色和纹理的多样性及息肉与其周围黏膜之间的边界不清晰的问题,达到了提高分割精度和边界的目的。

## 4 结 论

提出了一种结合 HarDNet 和反向注意力的息肉分割网络模型 HraNet,用来解决结肠镜图像中息肉的分割问题。该网络由 HarDNet68、RA 和 RFB 组成了一个健壮、统一的 U 型体系结构,其形成的迭代交互机制,能够有效校正预测结果中矛盾冲突的区域。HarDNet68 的 3 个边输出产生的损失直接反向传导到对应的卷积层,缓解了部分梯度消失,很好地解决了息肉的形状、大小、颜色和纹理的多样性、息肉与其背景之间的差异及结肠镜图像的低对比度等问题。在 5 个流行的基准数据集上进行了大量的实验,实验结果表明,所提方法在 Kvasir SEG 数据集上的 mDice 达到了 0.901,在 CVC ClinicDB 数据集上 mDice 达到了 0.930,均高于其他对比方法。HraNet 不需要经过预处理工作,实现了结肠息肉的高精度分割,同时具有良好的泛化能力,表现出了巨大的优势。HraNet 无论是训练速度还是推理速度都优于对比方法,说明 HraNet 在模型的轻量化和实时分割效率之间实现了很好的平衡。未来的工作将通过引入其他的有效注意力模块来进一步提高分割精度和泛化能力。同时,所提网络模型可应用在其他医学图像的分割任务上,如皮肤癌的分割、视网膜血管分割和肺部病变分割等。

## 参 考 文 献

- [1] Shussman N, Wexner S D. Colorectal polyps and polyposis syndromes[J]. *Gastroenterology Report*, 2014, 2(1): 1-15.
- [2] Sung H, Ferlay J, Siegel R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA: A Cancer Journal for Clinicians*, 2021, 71(3): 209-249.
- [3] Siegel R L, Miller K D, Sauer A G, et al. Colorectal cancer statistics, 2020[J]. *CA: A Cancer Journal for Clinicians*, 2020, 70(3): 145-164.
- [4] Issa I A, Noureddine M. Colorectal cancer screening: an updated review of the available options[J]. *World Journal of Gastroenterology*, 2017, 23(28): 5086-5096.
- [5] Ahn S B, Han D S, Bae J H, et al. The miss rate for colorectal adenoma determined by quality-adjusted, back-to-back colonoscopies[J]. *Gut and Liver*, 2012, 6(1): 64-70.
- [6] 刘佳伟, 刘巧红, 李晓欧, 等. 一种改进的双 U 型网络的结肠息肉分割方法[J]. *光学学报*, 2021, 41(18): 1810001.
- [7] Liu J W, Liu Q H, Li X O, et al. Improved colonic polyp segmentation method based on double U-shaped network[J]. *Acta Optica Sinica*, 2021, 41(18): 1810001.
- [7] 李大湘, 张振. 基于改进 U-Net 视网膜血管图像分割算法[J]. *光学学报*, 2020, 40(10): 1010001.
- [8] Li D X, Zhang Z. Improved U-net segmentation algorithm for the retinal blood vessel images[J]. *Acta Optica Sinica*, 2020, 40(10): 1010001.
- [8] Brandao P, Mazomenos E, Ciuti G, et al. Fully convolutional neural networks for polyp segmentation in colonoscopy[J]. *Proceedings of SPIE*, 2017, 10134: 101340F.
- [9] Akbari M, Mohrekesh M, Nasr-Esfahani E, et al. Polyp segmentation in colonoscopy images using fully convolutional network[C]//2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), July 18-21, 2018, Honolulu, HI, USA. New York: IEEE Press, 2018: 69-72.
- [10] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. *Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science*. Cham: Springer, 2015, 9351: 234-241.
- [11] Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N, et al. UNet++: a nested U-net architecture for medical image segmentation[M]//Stoyanov D, Taylor Z, Carneiro G, et al. *Deep learning in medical image analysis and multimodal learning for clinical decision support. Lecture notes in computer science*. Cham: Springer, 2018, 11045: 3-11.
- [12] Jha D, Smedsrud P H, Riegler M A, et al. ResUNet++: an advanced architecture for medical image segmentation[C]//2019 IEEE International Symposium on Multimedia (ISM), December 9-11, 2019, San Diego, CA, USA. New York: IEEE Press, 2019.
- [13] Fang Y Q, Chen C, Yuan Y X, et al. Selective feature aggregation network with area-boundary constraints for polyp segmentation[M]//Shen D G, Liu T M, Peters T M, et al. *Medical image computing and computer*

- assisted intervention-MICCAI 2019. Lecture notes in computer science. Cham: Springer, 2019, 11764: 302-310.
- [14] Wang P, Chung A C S. DoubleU-net: colorectal cancer diagnosis and gland instance segmentation with text-guided feature control[M]//Bartoli A, Fusiello A. Computer vision-ECCV 2020 workshops. Lecture notes in computer science. Cham: Springer, 2020, 12535: 338-354.
- [15] Zhang R F, Li G B, Li Z, et al. Adaptive context selection for polyp segmentation[M]//Martel A L, Abolmaesumi P, Stoyanov D, et al. Medical image computing and computer assisted intervention-MICCAI 2020. Lecture notes in computer science. Cham: Springer, 2020, 12266: 253-262.
- [16] Fan D P, Ji G P, Zhou T, et al. PraNet: parallel reverse attention network for polyp segmentation[M]//Martel A L, Abolmaesumi P, Stoyanov D, et al. Medical image computing and computer assisted intervention-MICCAI 2020. Lecture notes in computer science. Cham: Springer, 2020, 12266: 263-273.
- [17] Fang Y Q, Zhu D L, Yao J H, et al. ABC-net: area-boundary constraint network with dynamical feature selection for colorectal polyp segmentation[J]. IEEE Sensors Journal, 2021, 21(10): 11799-11809.
- [18] Chao P, Kao C Y, Ruan Y S, et al. HarDNet: a low memory traffic network[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 3551-3560.
- [19] Chen S H, Tan X L, Wang B, et al. Reverse attention for salient object detection[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11213: 236-252.
- [20] Liu S T, Huang D, Wang Y H. Receptive field block net for accurate and fast object detection[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11215: 404-419.
- [21] Xie S N, Tu Z W. Holistically-nested edge detection [C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015.