

文章编号: 2096-1618(2023)03-0291-07

基于 Dual-Hard-Net 网络的结直肠息肉分割算法

高瑜曼¹, 高琳², 何晋²

(1. 成都信息工程大学电子工程学院, 四川 成都 610225; 2. 成都信息工程大学区域链产业学院, 四川 成都 610225)

摘要: 由于结直肠息肉在形状、色泽和质地等方面各异, 且息肉与其周边正常区域的边界分辨模糊等特点, 导致息肉医学图像分割存在较大挑战。为提高结直肠息肉的分割准确率, 提出一种改进的 Dual-Hard-Net 网络分割算法。算法采用经典的编解码结构, 以 Hardnet 为共享编码器提取必要的多尺度特征, 运用两对并行排列的解码器对提取的特征加以充分利用, 并引入改进的注意力机制和残差模块减少计算量, 解决梯度消失等问题。分别在两个不同的数据集上进行比较和实验, dice 及 mIoU 系数取得了 0.895 和 0.859 的准确率。与当前的多种主流算法相比, 算法具有更高的分割准确性和精确度。

关键词: 医学图像分割; 结直肠息肉; 编解码结构; 多尺度特征; 注意力机制

中图分类号: TP391

文献标志码: A

doi: 10.16836/j.cnki.jcuit.2023.03.007

0 引言

在人工智能和医学大数据快速发展趋势下, 基于深度学习的医学图像分割在医学辅助领域得到更普遍和深入的运用, 有重要的应用前景和潜在发展趋势。和自然图像相比, 医学图像存在分辨率高、数据量小、样本大小不平衡等特点, 因此基于不同医学分割任务的多种图像分割算法应运而生, 解决不同分割领域的问题且得到了有效的应用。如 Long J 等^[1]构建 FCN(完全卷积)网络, 通过由端到端、像素到像素训练的卷积网络进行语义分割。任意大小的图像都可输入到该网络中, 最后通过有效的推理和学习, 产生相应大小的分割结果。并采用全卷积化的方式来解决逐像素的预测问题, 使用跨层的跳跃连接结构, 将浅层的纹理信息和深层次的语义信息进行融合, 实现精准的分割任务。Ronneberger 等^[2]提出一个对称的 U 型网络, 可融合多层多尺度特征, 由一个收缩路径获取全局上下文信息以及一个对称的扩张路径进行局部精确定位组成。U-Net 网络采用对称编解码结构, 在编码网络部分对图像进行卷积获取上下文信息, 在解码网络部分采用上采样操作将图像恢复到相应的尺寸大小, 并通过跳跃连接将两部分同样分辨率大小的特征进行融合, 最后恢复到原尺寸大小。Zhou 等^[3]提出了 UNet++, 是一个结合深度监督的编码器网络, 编码器和子网络由嵌套及密集的跳跃长短连接构建联系, 短连接使模型能够得到更有效的训练, 长连接可获得上下文更多信息。它还整合了不同层次的特征

提高精度, 运用剪枝对网络结构进行简化并配合深度监督, 在可接受的精度范围内让深度网络参数量大幅度缩减。Fabian Isensee 等^[4]提出 nnunet, 这是一种根据 2D 和 3D 的 unet 进行改进的强大自我适应的框架。对包括不同的实体、图像模式、图像几何和数据集大小进行自动适应, 而无需手动调整数据集。Jha D^[5]提出了新的 DoubleUNet 框架, 即两个 UNet 网络叠加的结构, 并使用金字塔池化来获取上下文信息。Guan S^[6]将密集连接机制引入到 UNet 用于图像重建, 在降低网络参量的同时提升了图像重构的质量。MultiResUNet^[7]网络使用残差路径替代传统 UNet 中的跳跃连接, 用多分辨率思路替代卷积层。Huang H^[8]结合全尺度的跳跃连接和深度监督, 把不同尺度的高级语义和低级语义的特征图融合, 提高了计算效率和分割精度。

基于以上分割模型, 注意力块可以根据不同的重要性有选择地更改输入或分配不同权重, 得出每个像素在各个域中的特征重要性, 从而重点关注图像的关键信息, 较显著地提升分割的准确率。常见注意力机制有局部空间注意力、通道注意力和混合注意力。Biparva M 等^[9]提出一个网络(ST-NET), 将原始图像的信息从一个空间转换为另一个空间并保留关键信息。Oktay 等^[10]提出了注意 U-NET, 注意力块输出一个门控信号, 对不同空间位置的特征重要性加以控制。Roy A G 等^[11]提出的 SE-NET 使用 3 个步骤在通道上实现注意力加权。首先是挤压操作, 全局平均池在输入功能上执行以获取通道特征映射; 第二个是激发操作, 先对通道的数量进行压缩, 然后将压缩后的通道特征重建回到通道数; 最后, 将 Sigmoid 函数生成 [0, 1] 的特征权重映射乘以原始输入得到最终输出。Li

Xiang 等^[12]提出 SKNet, 设置一组动态选择的卷积, 分 Split、Fuse、Select 3 个步骤进行操作。空间注意力可以看作是一种自适应的空间区域选择机制, 而忽略了通道信息; 而通道注意力直接汇总全局信息, 同时忽略每个通道中的本地信息。因此, 为将两种注意力机制的优势加以结合, Fu Jun 等^[13]提出双注意力网络 (DANet), 在主干网络上附加两种注意力模块, 分别对空间维度和通道维度上的语义依赖关系进行建模, 并对输出进行融合。上面提到的注意力机制都提高了最终分割性能, 但改善图像分割性能最有效的是以渠道为中心的注意力机制。混合注意力机制可以同时利用空间和通道的优势。

由于息肉图像背景复杂且息肉很容易受其他正常区域干扰, 使编码器提取有效信息的能力有限; 部分息肉图像类别分布不均衡, 息肉区域像素在图像中比例较小, 图像数据少, 网络训练较困难, 导致出现漏检, 且分割精度不高等情况。针对以上问题, 受 U-Net 模型、DDANet 模型^[14]、Residual 模型和注意力模型的启发, 本文提出一种改进的 DHN (Dual-Hard-Net) 网络的结肠息肉分割算法。主要包括以下几个方面的内容:

(1) 基于 DDANet 网络架构, 通过编解码的形式, 在编码部分运用 hardnet 网络对特征进行有效的多尺度提取, 并结合一个由两个并行解码器共享的单个编码器结构对多尺度特征加以充分利用和融合, 解决了特征能力不够, 数据量少等问题。

(2) 多支路形式, 用不同尺寸大小卷积核的卷积层和空洞卷积层来提取具有不同感受野的特征, 拓展网络特征提取的宽度, 最后, 用 1×1 的卷积来合并这些多尺度特征。

(3) 在解码网络结构和残差模块中嵌入注意力机制模块, 捕获多尺度特征进行融合的同时, 根据其特征之间的相关性分配相关的系数, 对息肉区域加以重点关注, 解决了息肉像素占比小, 关注度不高, 无区别等问题, 提高了分割精度。

1 基于 DHN 网络的分割算法

1.1 算法框架

本文提出的 DHN (Dual-Hard-Net) 结构如图 1 所示, 网络主要包括 3 个部分: 编码部分、解码部分、注意力机制和残差模块。在编码过程中, 将训练图像输入到模型中, 经过由 hardnet 网络和卷积层以及池化层组成的编码器结构, 由卷积层分别输出 4 个不同尺度的特征图, 再分别输入到解码器中; 在解码过程中, 上

面第一个解码器充当分割网络, 下面第二个解码器充当自动编码器网络。自动编码器网络有助于加强编码器网络中的特征图, 被用作生成注意力图以便和第一个解码器的输出相乘, 然后将结果输入到下一层的解码器中。在最终的输出中, 预测图像与输入图像尺寸相同。在解码网络结构和残差模块中嵌入注意力机制模块, 根据像素特征之间的相似关联性, 关注有效信息, 忽略无关信息。将多分支和残差模块嵌入在编解码网络及各个卷积层之中, 提取多尺度特征信息, 拓展网络宽度。

此网络结构的特征为: 采用并结合由两个并行解码器共享的单个编码器和多个并行解码器组成的编解码体系结构, 编码器中采用的 har-net 网络是对 Densenet 的改进, 在网络中减少 shortcut, 增加计算速度, 弥补准确性的丢失, 且能提取多尺度特征从而提高准确率; 其中第一个解码器充当分割网络, 第二个解码器充当自动编码器网络, 第一个解码器将编码器输出的 3 个不同尺度的特征图加以充分利用, 且使用注意力图来改善特征图的语义表示, 关注特征的重要部分, 再与融合的特征图相乘, 输入到下一个解码器。在下一个解码器中对编码器提取的浅层特征和上个解码器的输出重复进行同样的操作, 产生新的注意力图和融合新的特征图, 能够对重点区域加以再次筛选, 最终和高分辨率特征图相乘, 输出预测的分割 mask 和重建的灰度图像。

实验表明, 这种结构对多尺度特征的提取和利用及多次注意力机制的运用, 有助于提高整个网络的分割性能。

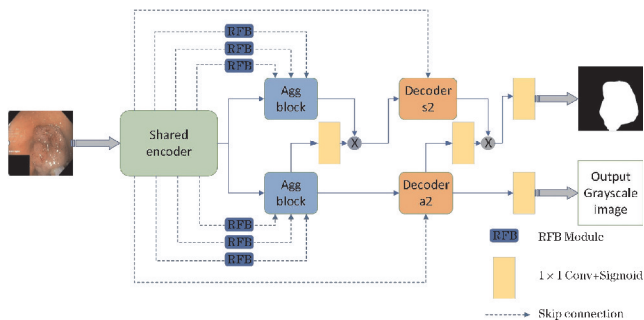


图1 DHN 算法网络框架

1.2 编码网络结构

编码部分由主干网络 HardNet^[15]、卷积层和池化层组成, 结构如图 2 所示。主干网络是一个低内存占用率的 CNN 网络, 其中 hard block 是对 Densenet 网络中的 dense block 进行改进, 减少短连接 shortcuts 的连接数, 增加了推理速度, 还增加了主层的通道宽度^[16]。

因为增加通道的宽度可以让每一层从不同方向、不同频率的纹理特征获取到更加丰富的特征,故可以恢复因为连接在修剪过程中下降的准确性,弥补准确率的丢失。用少量的 $1\times1\text{conv}$ 层来增加计算密度,从而增加运行速度,且达到了比较高的准确率。最后由 $\text{conv}1\times1$ 分别输出不同尺度的特征图,输入解码部分。

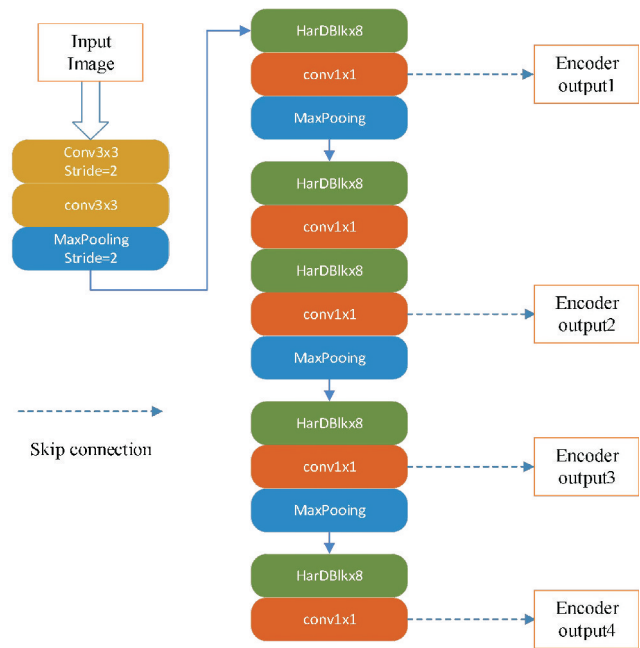


图 2 改进的 Hard-Net 编码器模块

1.3 解码网络结构

在解码器部分,借鉴双解码器注意网络 (DDA-NET),它是一种可以进行自动息肉细分的深度学习体系结构。具体结构为:有 4 个 decoder 块,RGB 图像输入到 encoder 块,将其编码成抽象的特征表示,然后下采样,编码的输出分别对称连到两个解码器中。第一个 decoder s1 块(分割支路)(图 3),有 3 个 RFB module 和一个聚合模块,RFB 模块是先对不同尺度的特征进行提取,然后再聚合为一个特征输出。

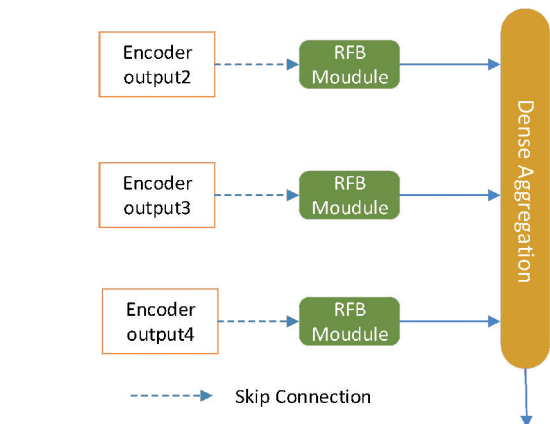


图 3 分割支路的解码模块

第二个 decodera2 块(自动编码器支路)(图 4),有 4×4 的反卷积,将其空间扩大两倍,然后用跳跃连接将编码网络的特征图堆叠到解码器中,跳跃连接增加特征代表性,能加快网络收敛。两个残差块在反向传播中学习必要的特征。紧接着一个 1×1 卷积和 sig 激活函数来产生一个注意力图谱。然后这个图谱乘以第一个解码器(分割支路)的输出,整体的输出作为下个分割支路的解码器的输入,最后 decoder 块输出通过 1×1 卷积和 sig 函数,第一个块输出分割掩码,第二个输出重建的灰度图。采用的改进解码部分的特点主要有两点:

- (1)由多个并列的对称结构的解码器组成,其中第一个解码器充当分割网络,将编码器输出的 3 个不同尺度的特征图加以充分利用,且使用注意图来改善特征图的语义表示,重点关注特征的重要部分。第二个解码器充当自动编码器网络。
- (2)注意力机制的引入使多尺度特征图在不同的区域有不同的权重系数,能够更精确地关注目标区域。

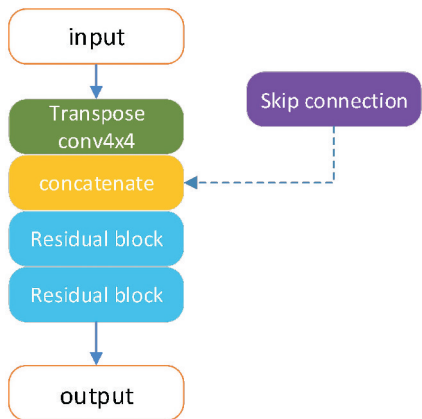


图 4 自动编码器支路的解码器模块

1.4 相关模块组成

多感受野特征提取模块(receptive field block, RFB)(图 5),采用多支路形式,用不同尺寸大小卷积核的卷积层和空洞卷积层^[17]来提取不同感受野的特征,最后用 1×1 的卷积合并这些多尺度特征,得到最后的特征表达,能加强从主干中学习到的深层特征表达。

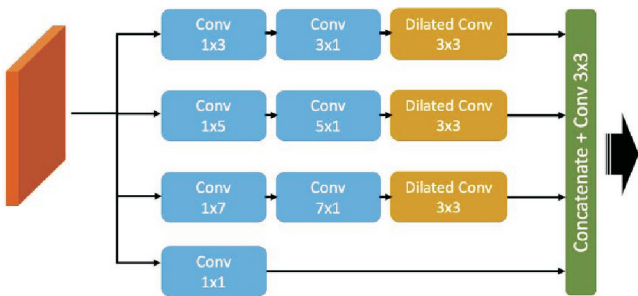


图 5 RFB 模块

Residual 模块(图6)对常见的残差模块进行了改进,引入注意力 SE(squeeze&excitation)模块,该注意力模块结合通道注意力,也可以引入其他的空间注意力和混合注意力模块,其作用是关注不同特征和目标之间的关联相似性,对不同特征进行区别对待,即分配不同的注意力系数,对有关联的特征加以重点关注而忽略不重要的特征,可以节省资源。

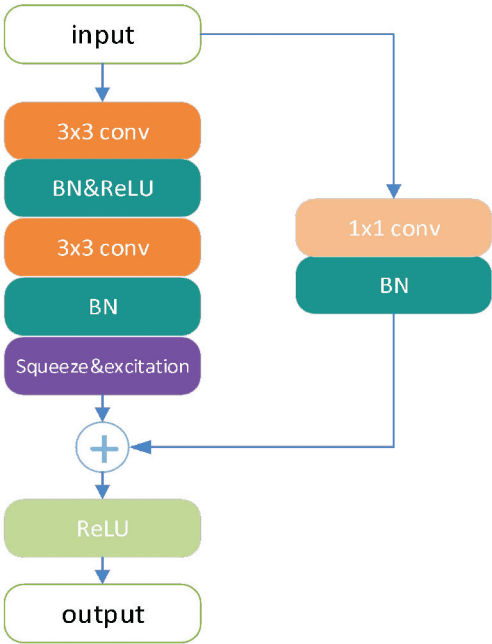


图6 改进的残差模块

而 Aggregation block 部分(图7)将经过 RFB module 处理的多尺度特征图先上采样到同样的尺寸,再与另外相应尺寸的输入特征进行相乘,最终通过 concat 操作叠加所得到的特征图,得到该模块的输出。

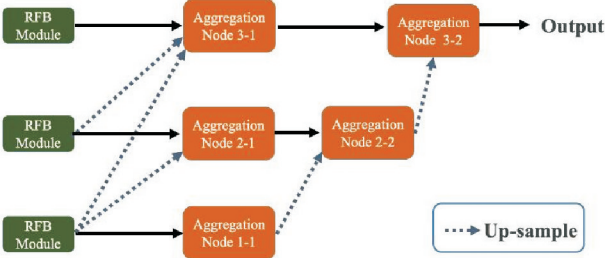


图7 特征聚合模块

1.5 损失函数

交叉熵损失函数不仅可以用在传统的分类问题中,也可用在大多数语义分割场景中使用。本文采用的总体损失为两种函数的混合损失函数,表达式为

$$L=L_{IoU}+L_{BCE}$$

通常 IoU 的计算方式为

$$L_{IoU}=1-\frac{\sum_{r=1}^H\sum_{c=1}^W S(r,c)G(r,c)}{\sum_{r=1}^H\sum_{c=1}^W [S(r,c)+G(r,c)-S(r,c)G(r,c)]}$$

在二分类情况下,使用二元交叉熵^[18]BCE 公式为
$$l_{BCE}=-\sum_{(r,c)}[G(r,c)\lg(S(r,c))+(1-G(r,c))\lg(1-S(r,c))]$$

其中, $G(r,c)\in\{0,1\}$ 是像素点 (r,c) 的 ground truth 标签, $S(r,c)$ 是其为目标显著类别的预测概率。其中 IoU 代表交并比, BCE 代表二元交叉熵,使用加权值的方式来计算损失,可使整个模型重点关注图像中物体的边缘部分。先提取像素值周围 15 个像素值(上下左右各 15 个),形成 31×31 的矩阵,然后计算矩阵中数值的平均值,最后计算平均值与该像素值的差值,这样可以得知该像素点与周围像素点的差异,并根据该差异大小添加相应大小的权值来计算损失,从而使模型更加关注图像的边缘部分。由于只关注图像中每个像素点与周围的区别,且结果只与差异的大小有关,所以只需要取绝对值。

2 实验结果及分析

2.1 数据集及评价指标

采用了 Kvasir-SEG^[19] 息肉数据集,由 1000 张息肉图像和对应的标签图像组成,其中 880 张图像用于训练集,120 张图像用于测试集。息肉数据的病灶区域和背景信息较难区分,边界不易分离,且病灶存在特征差异大、形状大小各异等特性,故该数据集对于医学分割任务具有较大的挑战性。因此采用了随机旋转,水平和垂直翻转等数据增强技术^[20]。

分割性能的评价指标采用的是相似系数(Dice)和平均交并比(mIoU)。相似系数(Dice)通常用于计算两个样本之间的相似度,是一种描述集合相似度的度量函数,其范围在 $[0,1]$,其中 1 表示预测值和真值完全重合,0 表示预测值和实际值毫无重合部分,定义为

$$s=\frac{2|X\cap Y|}{|X|+|Y|}$$

其定义中 $|X\cap Y|$ 表示 X 和 Y 之间的交集大小, $|X|$ 和 $|Y|$ 分别表示 X 和 Y 中元素的个数。在分割任务中, $|X|$ 和 $|Y|$ 则具体表示为分割的 ground truth 和 predict mask。由于分母对 X 和 Y 之间的共同元素存在重复计算,故将分子乘以 2。

平均交并比(mIoU)可解释为交集和并集的平均比值,利用分割区域重叠面积的比值测量真值和预测值之间的相关度。mIoU 值可用作衡量图像分割精度的重要指标,可以看作是在每个类别上分别计算的 IoU 值(即真的正样本数/(真的正样本数+假的正样本数+假的负样本数)),定义为

$$mIoU = \frac{TP}{TP+FP+FN}$$

其中, TP (true positive), FP (false positive), TN (true negative), FN (false negative) 分别表示真的正样本,假的正样本,真的负样本,假的负样本。整体实验结果以 mIoU 和 Dice 作为主要评价指标,且其数值取值越大表明分割精度越高,实验效果越好。

2.2 对比实验

为检验本文算法在 Kvasir-SEG、CVC-ClinicDB 数据集上的分割性能,在两个数据集上分别和 8 个经典算法进行了对比实验,具体为 UNet, ResUNet, ResUNet++, FCN8, DoubleUNet, PSPNet, Deeplabv3 + [21], DDANet 算法。在 Kvasir-SEG 息肉数据集上,采用 Adam 优化器,将网络模型与一些经典的医学分割网络以及基准网络 DDANet 做对比,实验结果分别如表 1 所示。

从表 1 中可以看出,本文算法在 Kvasir SEG 数据集的测试中,相对于其他 8 个经典的分割网络来说,在指标 mIoU 和 mDice 上均得到了提升,分别为 0.895 和 0.859。综合所有指标数据,实验结果表明本文算法的分割效果较传统分割网络来说准确率更高,在医学图像分割任务中,将 Dual-HardNet 网络应用于肠道息肉图像分割有一定的实用价值。

表 1 各算法分割结果对比

Method	mIoU	mDice
U-Net	0.471	0.597
ResUNet	0.572	0.69
ResUNet++	0.613	0.714
FCN8	0.737	0.831
DoubleUNet	0.733	0.813
PSPNet	0.744	0.841
Deeplabv3+	0.776	0.857
DDANet	0.78	0.857
本文方法 (DHN)	0.895	0.859

2.3 消融实验

为验证 Decoder 结构对提升分割精度的有效性,本文构建只包含 Decoders1 部分网络和包含 Decoders1 +s2 的完整网络,并分别在数据集上进行实验,结果如表 2 所示。从表中可以看出 Decoder 结构对结果的影响,完整的 Decoder 结构能够提升 mIoU 指标,从而验证了提出的完整模型能够对多尺度图像的特征信息更加充分地学习和使用,使整体的分割精度得到提升。

表 2 消融实验结果

Decoders1 (a1)	Decoders2 (a2)	注意力模块	mIoU
✓			0.87
✓		✓	0.876
✓	✓		0.88
✓	✓	✓ (cbam)	0.892
✓	✓	✓ (se)	0.895

为验证不同注意力模块以及添加注意力模块是否对分割准确率产生影响,且减少其他因素的影响,构建分别包含 SE 或 CBAM 注意力模块和不包含注意力模块的网络,两组实验结果如表所示,第 1 行和第 2 行以及第 3 行与最后一行分别表明引入注意力模块有助于对准确率的提升,第 4 行和第 5 行分别表示引入不同的注意力模块对网络的影响,其中 se 模块效果最好。两组实验表明,在基础网络和多分支结构中添加注意力模块,有益于增强模块提取全局特征的能力,有区别对待不同的特征区域使分割精确度得到了提升,验证了注意力机制的有效性。

2.4 分割结果

分割结果如图 8 所示,可以看出, DHN 算法所得到的分割预测结果相比于实际的标记图,在整体的轮廓、位置和形状上都十分接近和相似,从视觉上来看,输入图像的病变区域分割边界更为平滑,形状和位置更加吻合,对于局部的图像信息也展现的较为详细和清晰。和其他算法的分割结果图比较可以看出, DHN 算法分割效果较好,边缘更清晰,且更加接近 GT 图,对医疗诊断起到更精确的辅助作用。

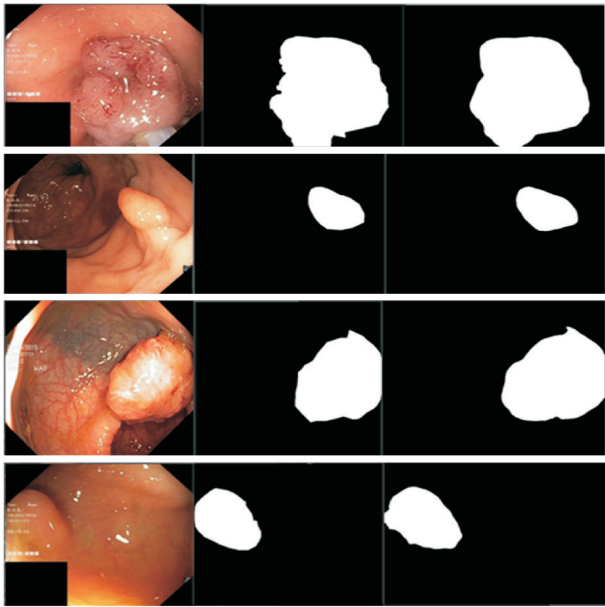


图 8 DHN 分割结果图

对比实验如图9所示,可以看出,DHN算法在Kvasir SEG数据集的测试中,相对于其它8个经典的分割网络来说,指标mIoU和mDice均得到了提升,分别为0.895和0.859。从消融实验中可以看出,在基础网络和多分支结构中添加注意力模块,有益于增强模块提取全局特征的能力,有区别对待不同的特征区域使分割精确度得到了提升,验证了注意力机制的有效性。改进的Decoder结构有助于提升分割精确度,从而验证了提出的完整模型能够对多尺度图像的特征信息更加充分的学习和使用,使整体的分割精度得到提升。综合所有指标数据,实验结果表明DHN算法的分割效果较传统分割网络来说准确率更高,在医学图像分割任务中,将Dual-HardNet网络应用于肠道息肉图像分割有一定的实用价值。

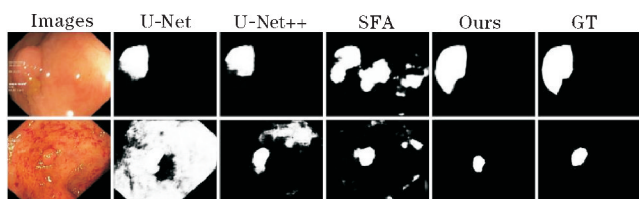


图9 不同算法的分割结果对比图

3 总结

提出一种基于注意力机制的医学图像分割模型,模型在传统的编解码结构上采用共享编码器和并列的解码器结构,结合多分支的结构特点提取并融合多尺度特征,引入注意力机制对特征分配相关权重。在两个公开的医学图像数据集上进行实验验证,并与多个主流分割经典网络进行了对比。实验表明DHN分割算法在分割的准确性和训练的速率等方面都得到了较大的提升,分割的效果更好,且具有一定的通用性,对于以后的临床诊断有一定的辅助应用价值。即使如此,实验中依然存在许多需要改进的地方,如由于肠息肉分割存在背景和目標难以区别的特点;预测的器官边界过于模糊,轮廓不清晰;由专业医学人士标注的医学图像样本量较匮乏,使模型进行训练和验证的样本量较少,导致模型容易出现过拟合等问题。为此,以后将进一步对数据进行图像预处理,使用图像增强技术,最大化增加图像样本量进而使模型得到进一步的优化。

参考文献:

- [1] Long J, Shelhamer E, Darrell T. Fully Convolutional Networks for Semantic Segmentation[J/OL]. <http://arxiv.org/abs/1411.4038>. 2015.
- [2] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation [J/OL]. <http://arxiv.org/abs/1505.04597>. 2015.
- [3] Zhou Zongwei, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, et al. UNet++: A Nested U-Net Architecture for Medical Image Segmentation[J/OL]. <http://arxiv.org/abs/1807.10165>. 2018;3-11.
- [4] ISENSEE F, PETERSEN J, KLEIN A, et al. 2018a. nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation [J/OL]. <http://arxiv.org/abs/1809.10486>. 2018;778-789.
- [5] JHA D, RIEGLER M A, JOHANSEN D, et al. DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation [J/OL]. <http://arxiv.org/abs/2006.04868>. 2020.
- [6] Guan S, Khan A, Chitnis P V, et al. Fully Dense UNet for 2D Sparse Photoacoustic Tomography Reconstruction[J]. *arXiv:1808.10848 IEEE*. 2018;7-23.
- [7] IBTEHAZ N, RAHMAN M S. MultiResUNet: Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation[J/OL]. *Neural Networks*, 2020, 121:74-87.
- [8] HUANG H, LIN L, TONG R, et al. UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation[C/OL]. *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Barcelona, Spain: IEEE:2021;1055-1059.
- [9] BIPARVA M, TSOTSOS J. STNet: Selective Tuning of Convolutional Networks for Object[J/OL]. <http://arxiv.org/abs/1708.06418>. 2022-07-25.
- [10] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. 2018. Attention U-Net: Learning Where to Look for the Pancreas [J/OL]. <http://arxiv.org/abs/1804.03999>. 2018.
- [11] ROY A G, NAVAB N, WACHINGER C. Concurrent Spatial and Channel Squeeze & Excitation in Fully Convolutional Networks [J/OL]. <http://arxiv.org/abs/1803.02579>. 2018.
- [12] LI X, WANG W, HU X, et al. Selective Kernel Networks [M/OL]. <http://arxiv.org/abs/1903.06586>, 2022-07-25.

- [13] FU J, LIU J, TIAN H, et al. Dual Attention Network for Scene Segmentation[J/OL]. <http://arxiv.org/abs/1809.02983>. 2019.
- [14] TOMAR N K, JHA D, ALI S, et al. DDANet: Dual Decoder Attention Network for Automatic Polyp Segmentation [J/OL]. <http://arxiv.org/abs/2012.15245>. 2020.
- [15] CHAO P, KAO C Y, RUAN Y S, et al. HardNet: A Low Memory Traffic Network[M/OL]. <http://arxiv.org/abs/1909.00948>, 2019. 2022-07-25.
- [16] Lu Z, Pu H, Wang F, et al. The expressive power of neural networks: A view from the width[C]. Advances in Neural Information Processing Systems. 2017:6231-6239.
- [17] YU F, KOLTUN V, Multi-Scale Context Aggregation by Dilated Convolutions [M/OL]. <http://arxiv.org/abs/1511.07122>. 2022-07-29.
- [18] Prashant Brahmabhatt, Siddhi Nath Rajan. Skin Lesion Segmentation using SegNet with Binary Cross-Entropy [J]. International Conference on Artificial Intelligence and Speech Technology (AIST2019), 2019:223-245.
- [19] JHA D, SMEDSRUD P H, RIEGLER M A, et al. 2019. ResUNet++: An Advanced Architecture for Medical Image Segmentation[J/OL]. <http://arxiv.org/abs/1911.07067>. 2019.
- [20] PEREZ L, WANG J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning [M/OL]. <http://arxiv.org/abs/1712.04621>. 2023-03-20. 2017.
- [21] 王亚刚, 郝怡媛, 潘晓英. 改进 DeepLabv3+ 网络的肠道息肉分割方法[J]. 计算机科学与探索, 2020, 14(7):1243-1250.

Colorectal Polyp Segmentation Algorithm Using Dual-Hard-Net Network

GAO Yuman¹, GAO Lin², HE Jin²

(1. College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 610225, China; 2. College of Blockchain Technology, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: Colorectal polyps are different in size, color and texture, and the boundaries between the polyps and the surrounding mucosa are not clear, leading to significant challenges in polyp segmentation. In order to improve the segmentation accuracy of colorectal polyps, this paper proposes an improved Dual-Hard-Net network segmentation algorithm which applies the classical encoding/decoding structure. In this structure, we use the Hardnet as the shared encoder to extract multi-scale information, and use two Juxtaposed decoders to further process the extracted features. To enhance the semantic information of the feature map, reduce the amount of computation and settle the Vanishing Gradient Problem, the advanced attention mechanism and residual blocks are introduced for original network. The results of experiments on two medical image data sets show that the Dice coefficient and mIoU are increased, which are 0.895 and 0.859 respectively compared with the original network. The experimental results compared with the main classic algorithms verify the feasibility and high accuracy of this method.

Keywords: medical image segmentation; colorectal polyps; encoding/decoding structure; multi-scale information; attention mechanism