

一种改进的双 U 型网络的结肠息肉分割方法

刘佳伟¹, 刘巧红^{2*}, 李晓欧², 凌晨², 刘存珏¹¹上海理工大学医疗器械与食品学院, 上海 200093;²上海健康医学院医疗器械学院, 上海 201318

摘要 针对结肠息肉图像中息肉颜色与背景颜色相似、息肉大小不一等影响分割效果的问题,提出了一种改进的双 U 型结构的结肠息肉图像分割网络。所提模型以 DoubleU-Net 为基础架构,首先,在 U 型结构的跳跃连接处融入空间注意力机制以提取空间特征的相关性信息;其次,在跳跃连接处同时并入通道注意力机制以有效地表达有用通道的依赖性,抑制与息肉分割任务无关的特征;最后,在第二个 U 型结构编码器的输入端引入多尺度选择核心通道注意力模块,以自适应地选择不同的感受野、提高目标的分割精度。实验结果表明,所提出的方法无论从客观指标还是视觉效果上均优于现有的方法,研究成果可以为结肠息肉早期检查和手术计划提供新的参考。

关键词 图像处理;多注意力 DoubleU-Net;结肠息肉;图像分割;注意力机制

中图分类号 TP391.4; TP183

文献标志码 A

doi: 10.3788/AOS202141.1810001

Improved Colonic Polyp Segmentation Method Based on Double U-Shaped Network

Liu Jiawei¹, Liu Qiaohong^{2*}, Li Xiaoou², Ling Chen², Liu Cunjue¹

¹ School of Medical Instrument and Food Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;

² College of Medical Instruments, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China

Abstract To solve the problems that the color of colonic polyps is similar to the background color, and the different sizes of polyps affect the segmentation effect, this paper proposes an improved image segmentation network for colonic polyps with a double U-shaped structure. The proposed model is based on the DoubleU-Net architecture. Firstly, the spatial attention block is integrated into the skip connection of the U-shaped structure to extract the correlation information of spatial features. Secondly, the channel attention block is incorporated into the skip connection to clearly express the dependency of useful channels and suppress the features unrelated to the polyp segmentation task. Finally, at the input of the encoder with the second U-shaped structure, the selective kernel block is introduced to adaptively select different receptive fields and improve the target segmentation accuracy. The experimental results show that the proposed method is better than the existing methods in terms of objective indices and visual effects. The research results can provide new reference for the early detection and surgical planning of colonic polyps.

Key words image processing; multi-attention DoubleU-Net; colonic polyps; image segmentation; attention mechanism

OCIS codes 100.3008; 100.4996; 200.4260

收稿日期: 2021-03-03; 修回日期: 2021-03-24; 录用日期: 2021-04-01

基金项目: 国家自然科学基金(61801288)

通信作者: *hqlqh@163.com

1 引言

据 2018 年癌症统计报告^[1]数据显示,全球结直肠癌的发病率为 10.2%,死亡率为 9.2%,在所有恶性肿瘤中,结直肠癌的发病率、死亡率均位于第三位。而在我国,据国家癌症中心的数据资料显示^[2-3],2015 年全国结直肠癌新发病例数约 38.76 万,占全部恶性肿瘤发病的 9.87%,死亡人数约 18.71 万,占全部恶性肿瘤死亡的 8.01%。结直肠癌的产生与息肉的发生发展有着紧密的联系,其中腺瘤性息肉被公认为癌前病变,从癌变发展到结直肠癌一般需要 5~10 年的时间,如果在此期间能早筛查早诊治,可以有效地防止结直肠癌的发生,降低结直肠癌的死亡率。结肠镜检查肠道中的息肉已经成为肠癌早期筛查的黄金标准,其不仅能对病变情况进行清晰的显示,还能对部分的肠道疾病进行治疗。然而结肠镜检查完全依赖于临床医生的手持操作,其不仅对医生的临床技术和经验水平有着较高的要求,同时该工作的长时间重复性易于造成医生的身心疲劳,结肠镜对于息肉检查的漏检率高达 25%^[4]。因此,利用结合人工智能技术的计算机辅助诊断帮助病灶识别,从而降低息肉漏检率变得极为重要。

随着近年来深度学习在图像处理、目标识别、机器翻译、自然语言处理等领域取得一系列突破性的进展,深度学习在医学领域也越来越受到重视,并逐渐应用于如医学图像^[5-7]、电子病历、药物研发和基因组学等各个方面的任务中。利用深度学习实现医学影像中病灶部位的自动识别和分割,其高准确性和可复用性能大大地减轻医生的工作压力,深度学习可辅助医生对疾病的诊治,提高工作效率。深度学习目前也被应用于结肠镜视频的息肉检测中,尤其在息肉图像分割方面取得了显著的效果。

目前针对息肉图像的分割方法大致可以分为两类:1)基于阈值、区域生长、统计形状、边缘检测等的传统分割方法;2)基于深度学习的方法。传统方法虽然实现简单,但是在实现图像分割的时候具有一定的局限性^[8]。更多研究使用深度学习的方法进行息肉图像的分割,U-Net^[9]是目前生物医学图像分割任务中最主流的基于编码器解码器结构的网络。Wang 等^[10]设计了一个基于 U-Net 的计算机辅助诊断模型以自动检测息肉,该方法在检测速度上明显快于人工操作。U-Net 网络在医学图像语义分割任务上的成功得益于其采用了跳跃连接,实现了高

低层特征的融合,基于编码器与解码器的上采样及下采样操作在原图像尺寸大小的特征图上进行像素级分类。然而,U-Net 的特征融合方式仅实现浅层信息的简单融合,未能充分利用更深层的信息,造成了低级特征与高级特征之间存在语义鸿沟。Zhou 等^[11]在 U-Net 原始的编码器解码器结构的基础上,将直接跳跃连接部分调整为密集跳跃连接,提出了 U-Net++ 网络。密集跳跃连接有效地将解码器中深层的、语义的、粗糙的特征图和编码器中浅层的、低级的、精细的特征图相结合,达到减小编码器与解码器之间的语义鸿沟的作用。Jha 等^[12]在标准的 ResUNet 基础上增加了 Squeeze-and-Excitation (SE)^[13]模块来对特征通道之间的相互依赖关系进行建模,同时增加了空洞空间金字塔池化 (ASPP)^[14]以实现多尺度信息融合并扩大了感受野,进而提出了 ResUNet++ 网络,此方法在两个息肉图像数据库的分割实验中明显改善了结肠息肉的分割效果。受到 ResUNet++ 网络及 U-Net 和其相关改进工作的启发,Jha 等^[15]进一步设计了双 U 型结构的网络模型 DoubleU-Net,该网络依次使用两个 U-Net 网络,即两个编码器和两个解码器。该网络利用预训练的 VGG-19^[16]作为第一个编码器提升模型精度,网络结构中增加了 SE 模块和 ASPP,目标在于提高医学图像分割应用的鲁棒性和泛化能力。该网络通过在结肠息肉等 4 个公开数据集上进行验证,取得优于 U-Net 和 U-Net++ 的效果,可以作为一种医学图像分割的基准网络。

本文提出一种改进的双 U 型结肠息肉图像分割网络,该网络针对结肠息肉图像中息肉颜色与背景色相似,以及分割目标尺寸小的问题,以 DoubleU-Net 为基础架构,在跳跃连接处融入联合空间注意力机制 (SAB) 和通道注意力机制 (CAB)^[17],同时在第二个编码器的输入端引入多尺度选择核心通道注意力模块 (SKB)^[18],形成了一个多重注意力机制的 MADoubleU-Net (Multi-Attention DoubleU-Net)。提出的 MADoubleU-Net 相对于 DoubleU-Net 具有以下几方面的优势:1)利用 SAB 有效地提取空间特征的相关性信息,加大空间维度重要特征的权重比例,克服了 DoubleU-Net 不能对空间维度信息进行联系上文信息相关性计算的问题;2)DoubleU-Net 在通道维度只使用了全局池化的压缩特征算法,而 MADoubleU-Net 采用的 CAB 在使用平均池化的同时加入最大池化,能够更加有效地表达有用通道的依赖性;

3) MADoubleU-Net 在第二编码器的输入端加入多尺度选择核心 CAB, 能够自适应地调节接收域的大小, 自适应地选择不同的感受野, 从而加强对通道维度的特征提取能力, 提高小目标的分割精度。

2 息肉图像基本分割框架

2.1 MADoubleU-Net 网络结构

针对结肠息肉图像中目标息肉区域颜色与背景颜色相近, 分割目标尺寸较小且尺寸大小不一等问题, 本文以双 U 型结构的 DoubleU-Net 作为基准网络, 提出了一种新的网络 MADoubleU-Net, 其主要结构包括双 U 型网络框架、多重注意力模块 (MAB)、空洞空间卷积池化金字塔模块 (ASPP), 以及在 ImageNet^[19] 上预训练的 VGG-19 编码器。

MADoubleU-Net 的网络结构如图 1 所示, 上半部分是第一个 U 型网络, 包含 4 个编码器块和 4 个解码器块, 编码器使用的是 ImageNet 预训练的 VGG-19, 其作为一个轻量级模型可以方便地与 U-Net 结构连接; 4 个解码器块对输入的特征图均采用 2×2 的双线性插值上采样, 以恢复成原特征图的尺寸大小。下半部分中的第二个 U 型网络与第一个基本相似, 同样包含 4 个编码器块和 4 个解码器块, 每个编解码器块会执行 3×3 的卷积操作, 且进行一个批量标准化 (BN) 操作, 在一定程度上避免数据偏移误差, 同时加快模型收敛, 然后利用 ReLU (Rectified Linear Unit) 激活函数加强非线性, 并将三个操作重复进行一次。此外解码器还会多执行一次 2×2 的池化层操作以降低计算复杂度。

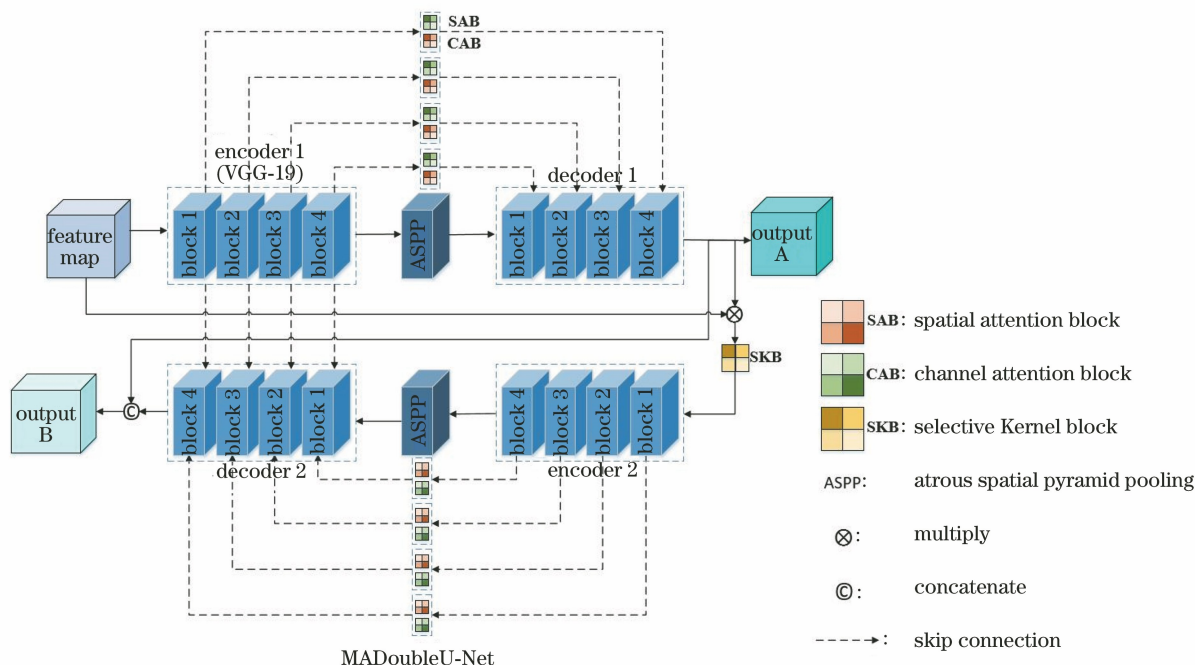


图 1 MADoubleU-Net 网络模型结构

Fig. 1 MADoubleU-Net network model structure

第一个 U 型网络的输出经过一个 1×1 卷积运算后, 通过 Sigmoid 函数形成 Output A。第二个 U 型网络解码器的输入包含了第一个 U 型网络编码器的跳跃连接、第二个 U 型网络编码器的输出, 以及经过多重注意力模块的跳跃连接, 该结构保持了高空间分辨率的同时, 提升了输出特征图的质量。最终的输出结果 Output B 综合了当前网络解码器的输出与 Output A。在两个 U 型网络的编码器和解码器之间利用 ASPP 来捕捉不同尺寸大小的结肠息肉的特征信息, 经实验证明该方法能够有效解决结肠息肉分割任务中息肉大小不一的问题。

本文的 MADoubleU-Net 网络相比于原始的 DoubleU-Net 网络在网络结构上进行了以下改进:

1) 空间注意力模块和通道注意力模块的引入。为了更深层次地提取息肉特征, MADoubleU-Net 在 2 个 U 型网络的上采样和下采样之间且具有相同分辨率的编码器块和解码器块的跳跃连接处引入空间注意力模块和通道注意力模块, 该结构在利用空间上下文信息获取空间特征相关性的同时, 赋予通道不同的权重, 提高了重要通道的敏感度, 使得分割结果更加专注于结肠息肉图像中的息肉部分, 从而减小不重要的背景特征的影响。

2) SKB 的引入。第二个 U 型网络编码器的输入特征图为第一个 U 型网络解码器的输出 Output A 与原始输入特征图的相乘后,再经 SKB 处理的结果,针对不同尺寸大小的卷积核,该模块可以自适应地为不同通道选择合适的感受野,以加强不同尺寸大小目标的特征表征能力,提高分割精度。经实验证明,改进的多重注意力模块能够有效解决结肠息肉分割任务中息肉与无用背景颜色相似的问题。

2.2 多重注意力模块

本文在原始的 DoubleU-Net 网络基础上融入了多重注意力模块,包括了空间注意力模块 SAB、通道注意力模块 CAB 以及多尺度选择核心通道注意力模块(SKB)。SAB 能够通过联系空间上下文特征信息,探究空间维度的相关性;CAB 通过计算通道维度的依赖性,来推理整个通道维度的重要性;

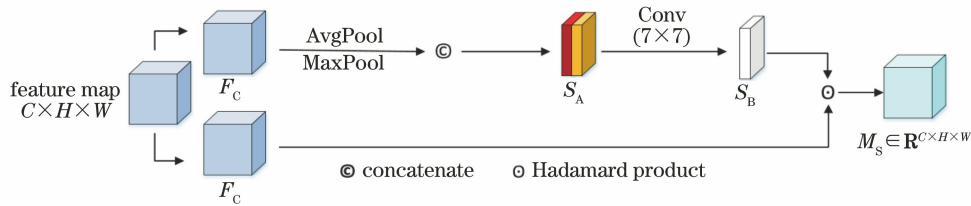


图 2 SAB 示意图

Fig. 2 Schematic of SAB

与传统 SE 模块的平均池化处理特征的方式不同,本文同时采用最大池化和平均池化来进行特征描述,二者结合使得提取的空间特征更加丰富且准确。空间特征描述 S_A 的计算公式为

$$S_A = \text{Con}(F_C^{\text{AvgP}}; F_C^{\text{MaxP}}), \quad (1)$$

式中: F_C^{AvgP} 和 F_C^{MaxP} 分别表示对特征图进行平均池化和最大池化后的特征图, $F_C^{\text{AvgP}}, F_C^{\text{MaxP}} \in \mathbf{R}^{1 \times H \times W}$ 。

相关性计算是对相连后的空间特征描述 S_A 进行了 7×7 的卷积操作后,进一步使用 Sigmoid 函数获得最终的空间相关性特征图 S_B , 其计算公式为

$$S_B = \sigma[\text{Conv}(S_A)], \quad (2)$$

式中: σ 表示 Sigmoid 激活函数; $S_B \in \mathbf{R}^{1 \times H \times W}$ 。

最后将原始特征图 F_C 与空间相关性特征图 S_B 的元素相乘,获得最终的空间注意力特征图 M_S :

$$M_S = S_B \odot F_C, \quad (3)$$

式中: \odot 表示哈达玛积; $M_S \in \mathbf{R}^{C \times H \times W}$ 。

2.2.2 通道注意力模块

一般情况下的卷积运算是默认对所有的通道进行计算融合,不同通道对特征提取的重要性和相关性是不同的。本文利用不同通道之间的依赖性来改

SKB 针对不同尺寸大小的卷积核,选择不同通道适合的不同感受野的信息,来让网络能够抑制结肠息肉图像的不重要的背景信息对分割产生的影响,从而让息肉特征信息能够发挥更大的作用,提升分割精度。SAB、CAB 和 SKB 在 MADoubleU-Net 网络中的位置如图 1 所示。

2.2.1 空间注意力模块

由于息肉图像中的目标区域的局部特征具有相关性,为了获得图像的空间上下文信息,得到相应的具备空间相关性信息的空间特征注意力图,本文使用空间注意力模块 SAB 来避免因局部不重要特征获得过大的权重而影响分割结果。SAB 主要包含空间特征描述、相关性计算和特征恢复三个部分,其结构图如图 2 所示。图中 C 为特征图的通道数, H 为特征图的高, W 为特征图的宽。

善特征提取时的通道语义信息,从而增强重要信息的提取并抑制非重要信息。通道注意力模块主要包含通道特征描述、相关性计算和特征恢复三个部分,其结构图如图 3 所示。

与空间注意力模块相似,本文的通道注意力模块使用最大池化和平均池化对特征图进行通道维度的特征描述。接下来的相关性计算中,将得到的两个特征描述图送入权值共享的全连接网络。此全连接网络包括一个全连接层的降维操作,ReLU 激活函数(以降低计算复杂度)和一个恢复到原始尺寸的升维操作。将输出的两个通道注意力映射的元素相加,并使用 Sigmoid 激活函数获得最终的通道相关性矩阵。特征描述 C_A 的计算式为

$$C_A = \sigma[\sigma_{\text{ReLU}}(F_C^{\text{AvgP}}) + \sigma_{\text{ReLU}}(F_C^{\text{MaxP}})], \quad (4)$$

式中: σ_{ReLU} 表示全连接层使用的 ReLU 激活函数; $C_A \in \mathbf{R}^{C \times 1 \times 1}$ 。

最后将原始特征图 F_C 与通道相关性特征图 C_A 的元素相乘,获得最终的通道注意力特征图 M_C , 其计算公式如下:

$$M_C = C_B \odot F_C, \quad (5)$$

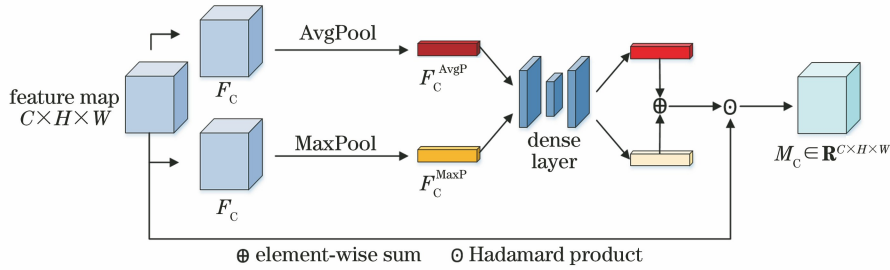


图 3 CAB 原理图

Fig. 3 Schematic of CAB

式中: \odot 表示哈达玛积; $M_C \in \mathbf{R}^{C \times H \times W}$ 。

2.2.3 多尺度选择核心通道注意力模块

上述通道注意力使用单一卷积核,它无法实现对不同大小的卷积核的卷积,无选择不同通道适合的感受野的能力。由于现实生活中人眼能够自适应地调节接收域的大小,因此本文的 MADoubleU-

Net 引入了 SKB,使得该网络针对不同尺寸的输入信息自适应地调节感受野的大小,从而加强对通道维度的特征提取能力。本文采用的 SKB 包括多尺度计算、相关性计算和特征选取三个部分,其结构图如图 4 所示。

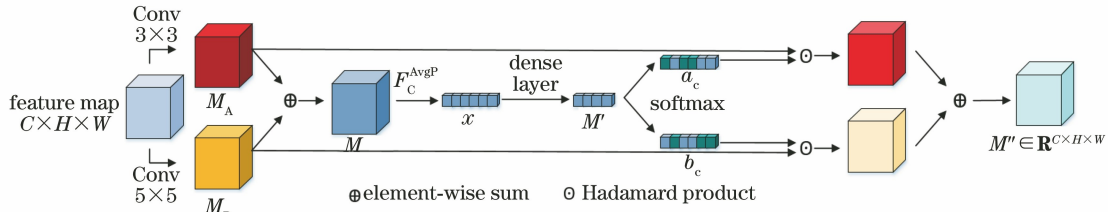


图 4 SKB 原理图

Fig. 4 Schematic of SKB

首先进行多尺度计算,分别利用 3×3 和 5×5 的卷积核输入特征图进行计算,得到特征图 M_A 和 M_B ,将 M_A 和 M_B 相连接,生成特征图 M 。多尺度卷积核在扩大感受野的同时融合了多尺度信息,这有利于后续对不同的通道赋予不同的权重。特征图 M 的计算式为

$$M = M_A + M_B, \quad (6)$$

式中: $M, M_A, M_B \in \mathbf{R}^{C \times H \times W}$ 。

多尺度处理后的特征图与原特征图 M 大小相同,接下来对特征图 M 进行与 SE 模块相似的操作,即平均全局池化,再通过一个全连接层进行降维,在批量标准化后,使用激活函数 ReLU 降低计算复杂度,特征图 M' 的计算式为

$$x = F^{AvgP}(M) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W M(i, j), \quad (7)$$

$$M' = \sigma[B(\omega x)], \quad (8)$$

$$D = \max\left(\frac{c}{r}, L\right), \quad (9)$$

式中: x 表示经过平均池化后的特征图; c 为特征图的通道数; D 为降维后的特征数; F^{AvgP} 表示平均池

化操作; B 表示批量标准化; $\omega \in \mathbf{R}^{D \times C}$; L 设置为 32。

特征图 x 表示通道能够自适应地选择不同尺度的特征信息,针对降维后的特征图 M' ,计算软注意力以选择不同的空间尺度,按信道方向使用 Softmax 函数。将输出的两个通道注意力与多尺度处理后的特征图相乘,经不同尺度选择后的具有通道相关性的特征图 M'' 可表示为

$$M'' = a_c \cdot M_A + b_c \cdot M_B, \quad (10)$$

式中: $M'' \in \mathbf{R}^{C \times H \times W}$; 下标 c 代表矩阵中的行数; a_c 和 b_c 分别表示两个不同卷积核的特征图的软注意力

$$a_c = \frac{e^{A_c x}}{e^{A_c x} + e^{B_c x}}, b_c = \frac{e^{B_c x}}{e^{A_c x} + e^{B_c x}}.$$

在 MADoubleU-Net 中,将 SKB 放在第二个编码器的输入位置,这是因为第一个 U 型结构的计算在获得了特征信息的同时,也存在部分有用特征丢失的现象。因此 SKB 与原始图像的结合有效地聚合了更多的信息,使得通道信息更加丰富。同时 SKB 的应用能够为重要的通道赋予更高的权重以及选择合适的感受野。

3 实验结果分析与讨论

3.1 实验环境与数据集

本文的实验硬件设备包括 Intel © E5-2620V4 2.1GHz 10 核处理器, NVIDIA GTX1080Ti (11 GB)×3。软件环境为 Ubuntu 18.04 LTS 64 位操作系统,开发工具 Python3.6 和 Tensorflow2.0 为后端的 Keras 框架。

实验采用 CVC-ClinicDB^[20] 数据集,该数据集为医学图像计算与计算机辅助干预国际会议于 2015 年发布的公共数据集。数据集包含 612 张 384×288 的原始图像以及 612 张 384×288 的息肉掩模。实验中将训练集、验证集和测试集按照 8:1:1 的比例划分,最后将剩余的 61 张结肠息肉图像用作测试。

3.2 评价标准

为了验证所提网络的有效性,采用了 4 个性能指标对模型分割结果的准确性进行评估,即相似系数(R_{Dice}),平均交并比(R_{mIoU}),精确率(P)和召回率(R),分别可表示为

$$R_{Dice} = \frac{2 \times T_P}{2 \times T_P + F_P + F_N}, \quad (11)$$

$$R_{mIoU} = \frac{T_P}{T_P + F_P + F_N}, \quad (12)$$

$$P = \frac{T_P}{T_P + F_P}, \quad (13)$$

$$R = \frac{T_P}{T_P + F_N}, \quad (14)$$

式中: T_P 表示正例被正确判断成正例的样本数; F_N 表示正例被错误判断成负例的样本数; T_N 表示负例被正确判断成负例的样本数; F_P 表示负例被错误判断成正例的样本数。相似系数(R_{Dice})和平均交并比(R_{mIoU})是用于评价医学图像分割的标准指标,金标准与分割结果部分的重合部分面积越大,则数值较大, R_{Dice} 和 R_{mIoU} 能够相互转换。精确率 P 是在预测为正的样本中正确属于正样本的比例,召回率 R 是正样本被正确预测的比例。

3.3 参数设置

本文实验在网络的训练过程中使用 Nadam 优化器,学习率的初始值设置为 10^{-5} ,并且使用二进制交叉熵作为网络的损失函数。为了充分利用实验硬件资源,将批处理量(batch size)设置为 8,迭代次数设置为 300。在训练过程中,学习率会随着训练次数的增加而降低,并且训练过程中,利用相应的回调函数自动停止迭代过程,本文所提网络结构在迭代 56 次时停止迭代。

3.4 数据扩充

由于训练集样本量较少,在训练的过程中易出现特征提取不足导致的过拟合问题,因此对数据集进行数据增强,通过水平翻转、垂直旋转、中心裁剪、随机亮度对比、弹性变换、高斯噪声和通道转置等方式对图像进行扩充。对每张图像进行 25 次变换,最终扩充到 12740 张图像,进行训练。部分结肠息肉扩充后的图像如图 5 所示。

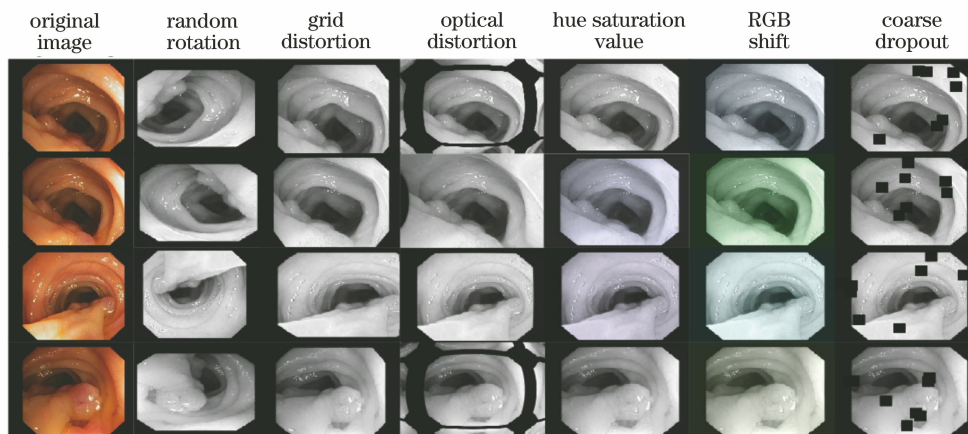


图 5 图像扩充

Fig. 5 Image expansion

3.5 结果与分析

为了验证本文提出的 MADoubleU-Net 的性能和分割结果,将 MADoubleU-Net 对息肉图像的分割结果分别与 FCN^[21], U-Net^[9], U-Net++^[11],

LinkNet^[22], DoubleU-Net(不使用 ASPP 模块)以及 DoubleU-Net^[10]等网络对息肉图像的分割结果进行对比,最终得到的各方法的分割性能指标如表 1 所示。从表 1 的结果可知,本文提出的

MADoubleU-Net 网络在 R_{Dice} 、 R_{mIoU} 和 P 三项指标上获得了最优的效果,分别达到了 0.931, 0.877, 0.964。同时, R 指标为 0.818, 达到了 6 种方法的平均水平。实验结果表明本文方法能够有效提升息肉图像的分割精度。

表 1 不同方法在 CVC ClinicDB 数据集上的分割效果对比

Table 1 Segmentation results of different methods on CVC ClinicDB database				
Method	R_{Dice}	R_{mIoU}	R	P
FCN	0.809	0.753	0.794	0.814
U-Net	0.881	0.835	0.893	0.952
U-Net++	0.874	0.773	0.822	0.921
LinkNet	0.862	0.795	0.858	0.922
DoubleU-Net	0.923	0.861	0.835	0.959
DoubleU-Net (no ASPP)	0.913	0.841	0.809	0.944
MADoubleU-Net	0.931	0.877	0.818	0.964

将本文提出的网络 MADoubleU-Net 与基准网

络 DoubleU-Net 的息肉图像分割结果进行对比,分割视觉效果如图 6 所示。图 6 每行分别显示每张息肉图像、对应的真实标签、DoubleU-Net 的两个 U 型结构输出以及 MADoubleU-Net 的两个 U 型结构输出。从图中可以看出, MADoubleU-Net 的分割结果相比 DoubleU-Net 更加接近于真实标签。由示例 1 得出 MADoubleU-Net 解决了 DoubleU-Net 对空间维度信息处理不足以及通道维度特征表征较粗糙的问题;示例 2 和示例 3 的单息肉图像的分割结果表明, MADoubleU-Net 能够有效解决目标息肉与背景颜色相似的问题,其多重注意力模块避免了颜色相似的息肉以及背景的干扰,比 DoubleU-Net 的分割结果更加准确;而示例 4 中的多息肉的分割结果表明, MADoubleU-Net 能够有效分割目标息肉大小形状各异的结肠息肉,通过 ASPP 尺度方法将不同尺寸大小的结肠息肉特征准确捕获,使得分割效果更为精准。

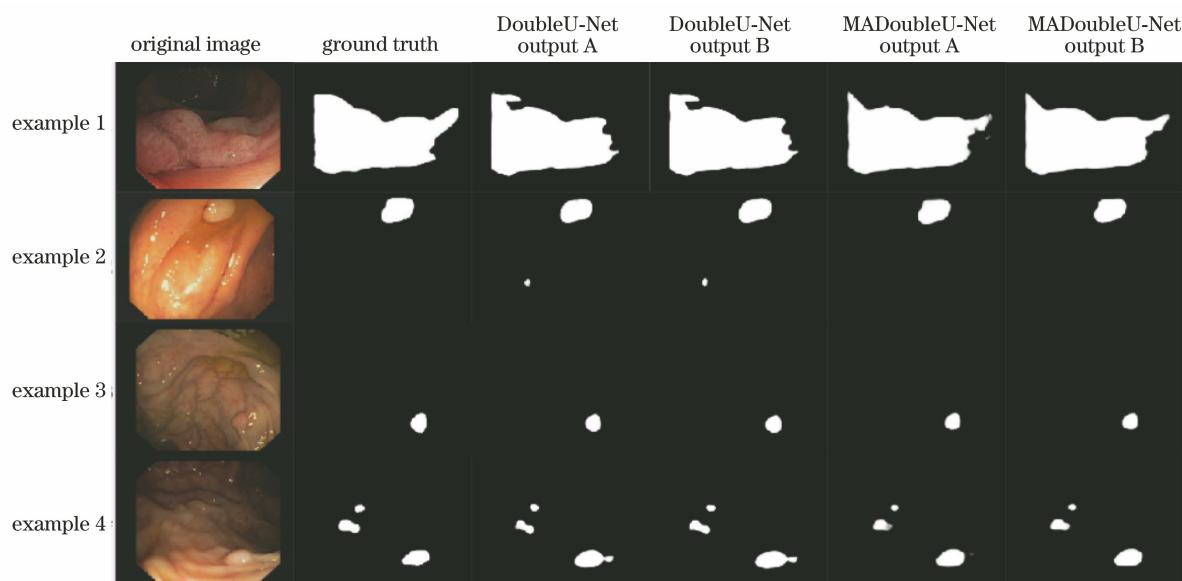


图 6 结肠息肉分割结果对比图

Fig. 6 Comparison of colon polyp segmentation results

综合客观评价指标和视觉效果可知,本文提出的 MADoubleU-Net 网络在 R_{Dice} 、 R_{mIoU} 和 P 指标上获得了最优的表现,同时在视觉效果上优于其他方法。本文提出的 MADoubleU-Net 网络利用多重注意力模块联系图像上下文特征信息,在空间维度和通道维度获取特征相关性和依赖性,可在提高分割精度的同时扩大感受野,让网络专注于目标分割区域,从而减小颜色相似的背景影响,有效地解决了结肠息肉图像中息肉颜色与景色相似的问题。

4 结 论

针对结肠息肉的图像分割问题,在原始 DoubleU-Net 网络的基础上,提出了一种新的融合多重注意力机制的网络模型 MADoubleU-Net。该网络引入 SAB 和 CAB,分别计算空间维度相关性和通道维度依赖性,使得网络更加准确地利用一些重要的通道表征信息,从而提高分割精度;通过引入 SKB 并使用不同尺寸大小的卷积核,让网络能够在不同的通道维度自适应地选择感受野,从而解决了

结肠息肉与背景颜色相似的问题;同时多尺度的特征信息融合能够解决目标结肠息肉尺寸不一的问题。在 CVC-ClinicDB 数据集上的实验结果表明,本文方法的 R_{Dice} , R_{mIoU} 和 P 分别达到了 0.931, 0.877, 0.964, 均高于对比方法。不过本文方法仍存在一些问題,如网络较深较复杂、网络参数过多、网络的泛化能力不足等。未来研究需进一步优化网络模型,提高结肠镜下的息肉图像分割效果。

参 考 文 献

- [1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. CA: A Cancer Journal for Clinicians, 2018, 68(6): 394-424.
- [2] Chen W Q, Sun K X, Zheng R S, et al. Cancer incidence and mortality in China, 2014 [J]. Chinese Journal of Cancer Research, 2018, 30(1): 1-12.
- [3] Wu C X, Gu K, Gong Y M, et al. Analysis of incidence and mortality of colorectal cancer in China, 2015 [J]. China Oncology, 2020, 30(4): 241-245.
吴春晓, 顾凯, 龚杨明, 等. 2015 年中国结直肠癌发病和死亡情况分析 [J]. 中国癌症杂志, 2020, 30(4): 241-245.
- [4] Leufkens A M, van Oijen M G H, Vleggaar F P, et al. Factors influencing the miss rate of polyps in a back-to-back colonoscopy study [J]. Endoscopy, 2012, 44(5): 470-475.
- [5] Zhang H L, Li Q, Guan X. An improved three-dimensional dual-path brain tumor image segmentation network [J]. Acta Optica Sinica, 2021, 41(3): 0310002.
张恒良, 李镛, 关欣. 一种改进的三维双路径脑肿瘤图像分割网络 [J]. 光学学报, 2021, 41(3): 0310002.
- [6] Huang H, Peng C, Wu R Y, et al. Self-supervised transfer learning of pulmonary nodule classification based on partially annotated CT images [J]. Acta Optica Sinica, 2020, 40(18): 1810003.
黄鸿, 彭超, 吴若愚, 等. 基于部分注释 CT 图像的自监督迁移学习肺结节分类 [J]. 光学学报, 2020, 40(18): 1810003.
- [7] Li D X, Zhang Z. Improved U-net segmentation algorithm for the retinal blood vessel images [J]. Acta Optica Sinica, 2020, 40(10): 1010001.
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] Wang Y G, Xi Y Y, Pan X Y. Method for intestinal polyp segmentation by improving DeepLabv3 + network [J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(7): 1243-1250.
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.
- [9] 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.
- [10] Wang L S, Chen R Z, Hu Y X. IDDF2018-ABS-0261 Polyp detection using an unet based model [C] // International Digestive Disease Forum (IDDF) 2018, June 9-10, 2018, Hong Kong, China. London: BMJ Publishing Group Limited, 2018, 67: A85.
- [11] Zhou Z W, Siddiquee M M R, 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: 225-2255.
- [13] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [14] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation [EB/OL]. (2017-12-05) [2021-02-25]. <http://arxiv.org/abs/1706.05587>.
- [15] Jha D, Riegler M A, Johansen D, et al. DoubleU-net: a deep convolutional neural network for medical image segmentation [C] // 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), July 28-30, 2020, Rochester, MN, USA. New York: IEEE Press, 2020: 558-564.
- [16] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2014-09-04) [2021-02-25]. <https://arxiv.org/abs/1409.1556>.
- [17] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module [M] // Ferrari V,

- Hebert M, Sminchisescu C, et al. Computer vision-
ECCV 2018. Lecture notes in computer science.
Cham: Springer, 2018, 11211: 3-19.
- [18] Li X, Wang W H, Hu X L, et al. Selective kernel
networks [C] // 2019 IEEE/CVF Conference on
Computer Vision and Pattern Recognition (CVPR),
June 15-20, 2019, Long Beach, CA, USA. New
York: IEEE Press, 2019: 510-519.
- [19] Deng J, Dong W, Socher R, et al. ImageNet: a
large-scale hierarchical image database [C] // 2009
IEEE Conference on Computer Vision and Pattern
Recognition, June 20-25, 2009, Miami, FL, USA.
New York: IEEE Press, 2009: 248-255.
- [20] Bernal J, Sánchez F J, Fernández-Esparrach G,
et al. WM-DOVA maps for accurate polyp
highlighting in colonoscopy: validation vs. saliency
maps from physicians [J]. Computerized Medical
Imaging and Graphics, 2015, 43: 99-111.
- [21] Long J, Shelhamer E, Darrell T. Fully convolutional
networks for semantic segmentation[C]//2015 IEEE
Conference on Computer Vision and Pattern
Recognition (CVPR), June 7-12, 2015, Boston,
MA, USA. New York: IEEE Press, 2015: 3431-
3440.
- [22] Chaurasia A, Culurciello E. LinkNet: exploiting
encoder representations for efficient semantic
segmentation[C]//2017 IEEE Visual Communications
and Image Processing (VCIP), December 10-13,
2017, St. Petersburg, FL, USA. New York: IEEE
Press, 2017: 1-4.