

文章编号 1004-924X(2023)18-2765-16

多尺度 ResNeSt-50 聚合网络与置信度传播的 息肉图像分割

夏 平^{1,2}, 张光一^{1,2}, 雷帮军^{1,2*}, 邹耀斌^{1,2}, 唐庭龙^{1,2}

(1. 三峡大学 水电工程智能视觉监测湖北省重点实验室, 湖北 宜昌 443002;

2. 三峡大学 计算机与信息学院, 湖北 宜昌 443002)

摘要:针对大肠的息肉组织与正常组织间无明显边界,准确定位息肉位置困难的问题,提出了一种多尺度 ResNeSt-50 聚合网络融合顺序树重加权置信度传播(sequential Tree-Rewighted Message Passing, TRW-S)的息肉图像分割方法。为提高网络对息肉信息的表达能力,构建编码-解码结构的多尺度 ResNeSt-50 聚合网络,编码器由卷积模块和 4 级 ResNeSt 模块级联构建 ResNeSt-50 骨干网络,实现跨通道信息间的线性整合与连接;ResNeSt-50 采用拆分注意力机制加强重要通道组的表现能力,增强了残差模块提取息肉图像信息的能力;解码部分下三层构建多层感受野模块(receptive field block, RFB)获取多尺度信息,然后用密集聚合模块整合其输出,并以快速解码方式输出解码信息,保证其分割性能的同时减少参数量;其次,生成预测图时采用测试时图像增强(Test-Time Augmentation, TTA)模块提升预测准确度,并增强网络的泛化能力;最后,构建基于马尔科夫随机场的 TRW-S 算法对输出的预测图进行后处理,以实现分割边缘的连续性和分割区域内部的一致性。对大肠息肉数据集 Kvasir-SEG 的测试结果表明,本文方法相比于 U-Net, U-Net++, ResUnet, SFA, PraNet 等算法, mDice 值达 91.6%, mIoU 达 86.3%, Smeasure 达 0.921, MAE 为 0.023, 优于其他五种息肉分割算法;在未知数据集 ETIS-LaribPolypDB, ColonDB 上测试结果表明,相比于 PraNet 模型,本文模型的 mDice 值分别提升了 14.2%, 7.7%;从本文模型在 ETIS-LaribPolypDB 数据集上的分割表现看,本文算法对微小病变十分敏感;因此,本文算法分割的息肉图像,在分割区域内部的一致性、分割边缘的连续性、轮廓清晰度、捕捉微小病变能力等方面均表现出优良的性能,同时,对未知数据集具有较好的泛化能力。

关键词:息肉图像分割;多尺度密集聚合网络;拆分注意力机制;顺序树重加权置信度传播;多尺度感受野

中图分类号: TP391 **文献标识码:** A **doi:** 10.37188/OPE.20233118.2765

Polyp image segmentation based on multi-scale ResNeSt-50 aggregation network and message passing

XIA Ping^{1,2}, ZHANG Guangyi^{1,2}, LEI Bangjun^{1,2*}, ZOU Yaobing^{1,2}, TANG Tinglong^{1,2}

(1. Hubei Key Laboratory of Intelligent Vision based Monitoring for Hydroelectric Engineering,

Three Gorges University, Yichang 443002, China;

2. College of Computer and Information Technology, Three Gorges University, Yichang 443002, China)

* Corresponding author, E-mail: Bangjun.Lei@ieee.org

Abstract: There boundary between colorectal polyps and normal tissues is not typically evident. Therefore, accurately locating polyp positions is challenging. This study developed a novel polyp image segmen-

收稿日期:2023-03-30;修订日期:2023-05-04.

基金项目:国家自然科学基金资助项目(No. U1401252);湖北省重点实验室开放基金资助项目(No. 2018SDSJ07).

tation method based on a combination of multiscale ResNeSt-50 aggregation network and sequential tree-reweighted message passing (TRW-S). First, a multiscale ResNeSt-50 aggregation network with an encoding - decoding structure was constructed to improve the expressiveness of the network. The encoder of the network is cascaded by convolution module and four-level ResNeSt module to build the ResNeSt-50 backbone network, which realizes linear integration and communication between cross-channel information, ResNeSt-50 uses split attention to strengthen the performance of important channel groups and enhance the ability of the residual module to extract polyp image information. In the bottom three layers of the decoder, a multilayer receptive field block (RFB) was used to obtain multiscale information. Subsequently, the dense aggregation module was used to integrate the output. The decoding information was output by using a fast decoding method, which ensured consistent segmentation performance and reduced the number of parameters. Second, the test-time augmentation (TTA) module was used to improve the prediction accuracy and enhance the generalization ability of the network when generating predictive images. Finally, a sequential tree-reweighted message passing (TRW-S) algorithm based on Markov random fields was constructed to postprocess the predicted image output of the model. This helped achieve continuity of the segmentation edge and consistency within the segmentation region. The experimental results on Kvasir-SEG, an open-access dataset for gastrointestinal polyps images, show that our method achieved an mDice value of 91.6%, mIoU of 86.3%, Smeasure of 92.1%, and MAE of 2.3%, which are higher than those of the polyp segmentation algorithms based on U-NET, U-Net++, ResUNet, SFA, and PraNet. Test results on the unknown datasets ETIS-LaribPolypDB and ColonDB indicate that the proposed model affords improvements in the PraNet and mDice values by 16.4% and 7.7%, respectively. As regards the segmentation performance on the ETIS-LaribPolypDB dataset, the proposed model was found to be highly sensitive to small lesions. Thus, the proposed model exhibits excellent performance in terms of consistency of segmentation area, continuity of segmentation edge, sharpness of contour, and ability to capture small lesions. In addition, it exhibits good generalization ability in the case of unknown datasets.

Key words: polyp image segmentation; multiscale dense aggregation network; split-attention; sequential tree-reweighted message passing; multiscale receptive field

1 引言

结直肠癌是现今发病率排第三的癌症^[1-2],而大肠息肉是结直肠癌的早期病变特征,早期诊断是预防结直肠癌的重要手段。研究表明^[3],定期进行结肠镜检查能减少 30% 的结直肠癌发病率,临床中结肠息肉的判断往往需要经验丰富的医生进行定位,存在一定的误诊概率;因此,对结直肠息肉图像分割是为提高医生充分利用结肠镜影像来发掘影像中所反映的解剖及病例信息,便于医生的诊断及病情评估。由于息肉的形状、颜色、大小往往各不相同;此外,息肉与周围黏膜的边界不清晰,影响医生对息肉的判断,因此,息肉图像的分割和定位存在一定困难。由于大肠息肉图像的

复杂性,传统方法对大肠息肉图像分割效果不理想,随着医学图像处理技术的发展,深度学习技术在医学图像处理中的应用大放异彩^[4-5],基于全卷积网络的医学图像分割算法逐渐取代了传统的基于颜色、纹理、外观及形状等特征组合的息肉分割方法^[6-7],分割的效果大幅提高。

作为 FCN(Fully Convolutional Networks)^[8]网络的变体,U-Net^[9]使用编码-解码模型实现全卷积医学图像分割,使用跳跃连接保留浅层细节信息和深层空间信息,并融合不同深度的特征信息,在医学图像分割中取得了很好效果;但当图像特征较复杂时,其编码结构无法有效地提取其特征信息,因而,基于 U-Net 的息肉图像分割效果不佳。U-Net++^[10],ResUNet++^[11]等网络

用于息肉图像的分割,分割效果有了较大提升,但这些网络与 U-Net 存在同样的问题,无法清晰地分割病变边缘,且对微小病变存在完全漏检的情况。

研究者提出了其他一些大肠息肉图像分割方法,如,采用 PraNet(Parallel reverse attention Network)^[12],DoubleUNet^[13]网络等,网络中均采用了多尺度感受野模块,提升了息肉图像分割的效果,从而说明了采用多尺度感受野对提高息肉分割效果是一种有效的方式。ABCNet(Area-Boundary Constraint Network)^[14],ACSNet(Adaptive Context Selection Network)^[15]网络在提高息肉分割精度的同时关注了网络的泛化能力,随后提出的改进型双 U 型网络^[16]引入多尺度选择核心通道注意力和空间注意力等方式提升网络的分割精度;结合 HarDNet 和反向注意力的息肉分割^[17]一文中提出了用 HarDNet 来提高预测速度;这些都取得了较好的分割效果。

注意力机制在网络构建中被广泛应用。ResNet^[18]网络加入 SE 通道注意力模块形成 SENet^[19],SE-Net 通过自适应重新校准通道特征响应来引入通道注意力机制;SK-Net^[20]则利用两个网络分支引入特征图注意力,自适应调整自身感受野;ResNeXt^[21]则在瓶颈层(ResNet bottle)中采用组卷积,将多通道结构变为统一操作。ResNeSt^[22]借鉴了 ResNeXt^[21]网络思想,在 ResNet 中加入拆分注意力模块^[22-23],跨不同的特征图组实现特征图注意力。

针对息肉图像特点,本文提出了多尺度 ResNeSt-50 聚合网络与顺序树重加权置信度传播(Tree-reweighted message passing, TRW-S)的息肉图像分割方法。为定位和分割轮廓模糊的息肉图像,通过构建的多尺度 ResNeSt-50 聚合网络获取图像的弱特征信息,该网络架构由编码、解码两部分组成,编码部分用卷积模块和 ResNeSt 模块构建 ResNeSt-50 骨干网络,解码部分由多层分支感受野(Receptive Field Block, RFB)模块和密集聚合(Dense Aggregation)模块构成,以提高网络的表达能力与微小病变检测能力。网络输出的息肉分割图像经测试时数据增强(Test Time Augmentation, TTA)模块及构建的顺序树重加权置信度传播(TRW-S)算法强化

分割边缘,增强息肉图像内部分割的连贯性,提升模型的泛化能力,实现了息肉图像的有效分割。

2 本文方法

本文构建的息肉图像分割模型由编码-解码器、TTA、及 TRW-S 等三部分组成。为提升模型的表达能力与泛化能力,编码网络采用一级卷积模块和四级 ResNeSt 模块^[22]级联构建 ResNeSt-50 骨干网络;解码部分采用多层分支感受野(RFB)^[24]模块和密集聚合(Dense Aggregation)模块^[25]构建而成,以融合上下文信息,并采用 8 倍上采样跳过最上两层信息实现快速解码输出。

采用测试时数据增强(TTA)模块^[26]对训练后的模型采用水平和竖直翻转两种方式实现数据增强,并将增强的结果与原图预测结果取均值得到最终预测图,以此提升本文模型的预测及泛化能力。

模型最后在 TTA 输出的预测图中构建马尔科夫随机场的顺序树加权置信度传播(TRW-S)算法,结合全局信息对分割图进行平滑,以提高分割图像边缘的连续性和分割区域内部的一致性。TRW-S 采用动态规划方法,正反向遍历节点来迭代传播树的信息并更新结点权重,以此实现其全局收敛;相对于树加权信息传递算法,本文 TRW-S 算法收敛性更好^[27-28]。如图 1 所示。

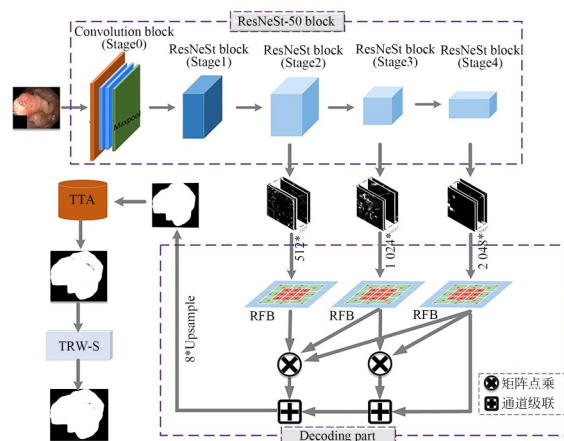


图1 本文模型

Fig. 1 Model of this paper

2.1 ResNeSt-50 编码模块

ResNeSt 网络将输入分成 k 个基数组 (Cardinal), 每个记为 Cardinal i , 每个基数组拆分为 r 个部分组成, 每部分内部通过卷积核分别为 1×1 , 3×3 两级卷积运算, 提取特征信息后, 将拆分的 r 个部分模块信息输入拆分注意力模块中, 实现基数组信息输出。将 k 个基数组信息全连接, 经 1×1 模块卷积运算输出, 特征图组数量取决于基数超参数 k , 引入基数超参数 r 指示基数组内的拆分数, 共有 $G = k \times r$ 组。拆分注意力思想借鉴了 SE-Net^[19] 与 SK-Net^[20] 中通道选择模块和感受野选择模块的思路, 由特征图组和分割注意力操作组成; 如图 2 所示。基数组通过求权重再融合的方式实现了特征通道间权重分配; 在其上进行全局平均池化, 得到 $c(C/K)$ 维特征向量, 表示各通道权重。第 k 个基数群表示^[22]:

$$\hat{U}^k = \sum_{j=r(k-1)+1}^{rk} U_j, \quad (1)$$

其中: $\hat{U}^k \in R^{h \times w \times c/k}$, h, w 与 c 表示块输出特征图大小。

全局上下文信息^[22]:

$$s_c^k = \frac{1}{h \times w} \sum_{i=1}^h \sum_{j=1}^w \hat{U}_c^k(i, j). \quad (2)$$

基数群加权融合实现通道注意力聚合, 经两个全连接层权重重新分配; 由于全连接或卷积层的输出为对称、非稀疏分布, 对其使用 BN 与 ReLU 能产生更加稳定的分布, 因此, BN 层和 ReLU 位于线性单元之后, 并用 softmax 操作保证基数组间特征层拆分权重独立分布, 有^[22]:

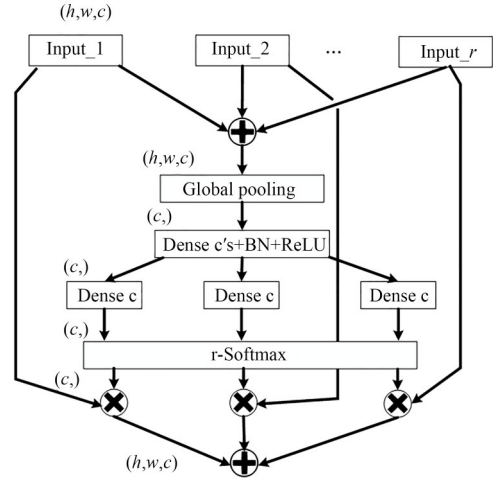


图 2 ResNeSt 中拆分注意力模块

Fig. 2 Split-attention in ResNeSt Block

$$a_i^k(c) = \begin{cases} \frac{\exp(G_i^c(s^k))}{\sum_{j=0}^R \exp(G_j^c(s^k))} & \text{if } R > 1 \\ \frac{1}{1 + \exp(-G_i^c(s^k))} & \text{if } R = 1 \end{cases}, \quad (3)$$

其中: $a_i^k(c)$ 表示通过 softmax 函数分配的权重, 映射 G_i 根据全局上下文表示 s^k 确定第 c 个通道的权重。

最后与原始特征组相乘后按式(4)得基数组输出。

$$V_c^k = \sum_{i=1}^R a_i^k(c) U_{r(k-1)+i}. \quad (4)$$

本文采用一级卷积模块和四级 ResNeSt 模块^[22]级联构建 ResNeSt-50 的骨干网络如图 3 所

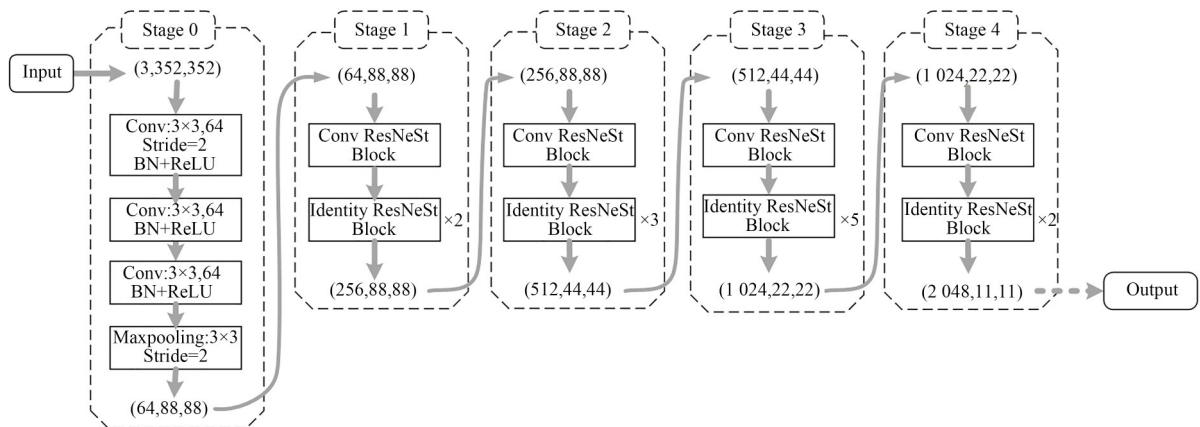


图 3 本文 ResNeSt-50 编码结构

Fig. 3 ResNeSt-50 encoding structure of this paper

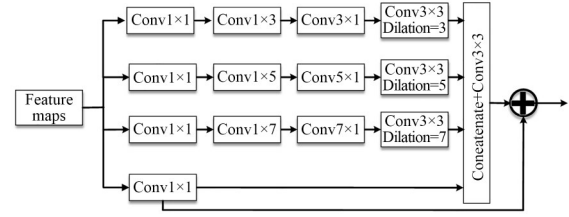
示。文后的消融实验验证了本文构建的 ResNeSt-50 骨干网络对模型的泛化能力有较明显地提升。

2.2 并行解码模块

本文并行解码模块由多层分支感受野 (RFB)^[24] 模块和密集聚合模块^[25]融合上下文信息构成。解码器的三层 RFB 模块与编码模块中第三、第四、第五级级联,编码输出的特性信息分别输入解码器的各层 RFB 中。RFB^[24] 模块采用并行的多尺度感受野,前三个分支先用大小分别为 $1 \times 3, 1 \times 5, 1 \times 7$ 卷积核进行卷积运算,再采用大小分别为 $3 \times 1, 5 \times 1, 7 \times 1$ 的卷积核进行卷积运算,最后用 3×3 的空洞卷积进行运算,膨胀系数分别为 3, 5, 7;第四个分支采用 1×1 的卷积运算,实现输入、输出特征图长宽比例的同时获得不同尺度的感受野,因此,本文使用多尺度感受野模块提升小病变区域的分割能力。

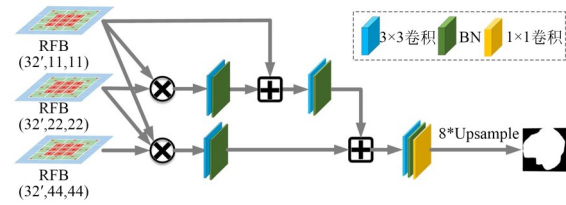
RFB 模块每一层包含不同尺度的膨胀卷积,有利于增大分割网络的感受野,且不会引入过多的计算量和增加网络深度,本文 RFB 模块输出通道数均为 32,这也减少了网络的参数量。文后的消融实验说明,加入 RFB 模块有利于提升模型的泛化能力。文献[12]亦证明,该模块有利于改善大肠息肉分割的效果。如图 4(a) 所示。

密集聚合模块采用逐元素相乘与上采样的方式连接不同尺度的特征图信息,经通道级联,将特征图信息合并输出。本文模型改变了 U-Net 中跳跃连接的线性信息传播方式,使用聚合方式实现不同阶段的信息融合,通过聚合与短连接形成的密集网络加强了新信息的传递,对特征图信息的反复利用强化了网络的理解能力,减少了网络参数量。RFB 模块之后通道数均降至 32,密集聚合后的网络达到了较理想的分割效果。若采用 U-Net 的对称解码结构,网络参数量会大大增加。文献[29]已证明,最上两层的跳跃连接对模型整体准确度贡献小,但参数数目大量增加,因此,本文采用 8 倍上采样跳过最上两层信息实现快速解码输出。如图 4(b) 所示。



(a) 多尺度感受野模块

(a) Multiscale receptive field block



(b) 密集聚合模块

(b) Dense aggregation module

图 4 RFB 与密集聚合模块

Fig. 4 RFB and dense aggregation module

2.3 置信度传播后处理

本文构建的 TRW-S 算法中,节点概率分布状态以消息传播方式传递给相邻节点,经多次迭代后,每个节点的置信度收敛到稳态。如图 5(a) 中,节点 i 对应状态值 x_i ,观测值 y_i 为其像素值,二者统计依赖性用似然函数 $\phi(x_i, y_i) = P(y_i | x_i)$ 描述;相邻节点 i, j 间相互作用以势能量 $\psi(x_i, x_j)$ 表示,其联合概率:

$$p(x_i, x_j) = \frac{1}{Z} \prod_i \phi(x_i, y_i) \prod_{i, j \in N(i)} \psi(x_i, x_j), \quad (5)$$

其中, Z 为归一化常数。

由图 5(b),所有传入节点 i 消息:

$$b_i(x_i) = k \phi(x_i, y_i) \prod_{i, j \in N} m_{ji}(x_i), \quad (6)$$

其中: k 为归一化常数,消息 m_{ji} 表示节点 j 对节点 i 当前标记的影响。

相邻节点 i, j 消息更新时,由节点 j 向节点 i 发送消息为 m_{ji} ,其满足:

$$m_{ji}(x_i) = \sum_{x_j} \phi(x_j, y_j) \cdot \psi(x_i, x_j) \prod_{k \in N(j) \setminus i} m_{kj}(x_j), \quad (7)$$

其中: $N(j) \setminus i$ 表示节点 j 四邻域系统中不含节点 i ;向量 $m_{ji}(x_i)$ 每一维元素描述节点 j 在相应标记

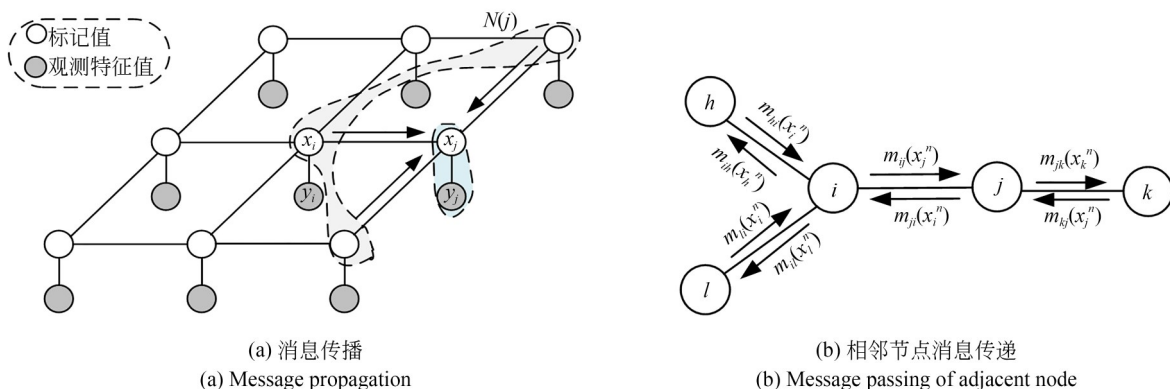


图 5 置信度的消息传播

Fig. 5 Message propagation of confidence level

$x_j (x_j \in L)$ 对节点 i 状态的影响。 $m_{ji}(x_i)$ 为:

$$m_{ji}(x_i) = \begin{bmatrix} m_{ji}(1) \\ m_{ji}(2) \\ \vdots \\ m_{ji}(L) \end{bmatrix}. \quad (8)$$

由式(6)和式(7)可知,节点 i 、节点 j 的联合概率:

$$b_{ij}(x_i, x_j) = k\phi(x_i, y_i)\phi(x_j, y_j)\phi(x_i, x_j) \cdot \prod_{i,j \in N(i)} m_{ij}(x_j) \prod_{i,k \in N, k \neq i} m_{ki}(x_i), \quad (9)$$

结合式(6)和式(9),节点 i 的边缘概率:

$$b_i(x_i) = \sum_{x_j} b_{ij}(x_i, x_j). \quad (10)$$

TRW-S算法利用节点间相互传递信息来更新当前MRF标记的状态,并依据传递后的信息更新节点的置信度。

定义标记场能量^[30-31]:

$$E(x_{ji}) = -\ln \phi(x_i, x_j) = \sum_{c \in C} V_c(x_i), \quad (11)$$

其中:二元势函数 $V_c(x_i, x_j) = \begin{cases} +\beta, & x_i = x_j \\ -\beta, & x_i \neq x_j \end{cases}$;

邻域系统基团势函数参数与标记值无关^[32], $V_c(\cdot)$ 由其邻域系统决定,选 $\beta_i = \beta (i = 1, 2, 3, 4)$ 为对应势团参数。

定义特征场能量:

$$E(y_i|x_i) = -\ln \phi(x_i, y_i) = -\ln P(y_i|x_i) = \frac{1}{2} (y_i - \mu_m)^T (\sum_m)^{-1} (y_i - \mu_m) + \ln(\sqrt{2\pi} |\sum_m|^{1/2}). \quad (12)$$

后验概率 $P(x_i|y_i, x_{N(i)})$ 最大对应的最优

标记 $\hat{x}_{iMAP}$:

$$\begin{aligned} \hat{x}_{iMAP} &= \arg \max_{x_i} P(x_i|y_i, x_{N(i)}) = \\ &= \arg \max_{x_i} (P(y_i|x_i)P(x_i|x_{N(i)})) = \\ &= \arg \max_{x_i} (\phi(x_i, y_i) \cdot \phi_{ji}(x_j, x_i)), \end{aligned} \quad (13)$$

其中: $P(x_i|x_{N(i)})$ 为标记局部概率, $N(i)$ 为节点 i 邻域。

式(13)等价于能量最小,即:

$$\hat{x}_{iMAP} = \arg \min_{x_i} \{ \sum_{(i,j) \in S} [E(y_i|x_i) + E(x_j, x_i)] \}. \quad (14)$$

树加权置信度传播算法: m_{ji}^t 为在 t 次迭代时节点 j 向节点 i 传递信息, c_{ji} 为 ρ_{ji}/ρ_j , 即在 ρ 中任一颗树, 给定包含节点 j 的条件下, 包含边 (j, i) 的可能性; 信息更新规则^[27]:

$$m_{ji}^t(x_i) = \min \left(c_{ji} \left(E(y_j|x_j) + \sum_{k \in N(j)} m_{kj}^{t-1}(x_j) \right) - m_{ij}^{t-1}(x_j) + E(x_j, x_i) \right). \quad (15)$$

经 t 次迭代后, 节点 i 的置信度:

$$b_i(x_i) = E(y_i|x_i) + \sum_{i,j \in N(i)} m_{ji}^t(x_i). \quad (16)$$

由最大置信度计算节点 i 的边缘概率最大时对应的标记:

$$\hat{x}_{iMAP} = \arg \max_{x_i} (b_i(x_i)). \quad (17)$$

TRW-S^[22]算法中消息的迭代从正反向遍历整个随机场内节点, 节点按扫描线遍历, 实现正反向传递消息; 正向遍历时, 每个节点向其右下邻域节点传递消息, 后向则向其左下邻域传递消

息,直至达到最终的收敛状态。

3 息肉图像分割算法

采用第 2 部分构建的模型对息肉图像进行分割。具体步骤:

Step1. 数据集。用公开数据集 CVC-ClinicDB^[33]、Kvasir-SEG^[34]作为训练样本;CVC-ClinicDB 包含 612 幅大小为 384×288 的息肉图像及

其人工标注,数据集 Kvasir-SEG 包含 1000 幅大小不一的息肉图像与人工标注;两数据集均用 9:1 比例分割成训练集与测试集,分别在 Kvasir-SEG^[34]、CVC-ClinicDB^[33]、ColonDB^[35]、ETIS-LaribPolypDB^[1]、CVC-300^[36]等五个数据集上对分割模型进行测试;其中,ColonDB、ETIS-LaribPolypDB、CVC-300 三个数据集对训练样本而言为未知数据集,采用该三个数据集测试和检验本文模型的泛化能力。如表 1~表 2 所示;

表 1 本文使用的五个数据集

Tab. 1 Five datasets used in this paper

数据集	Kvasir-SEG ^[34]	CVC-ClinicDB ^[33]	ColonDB ^[35]	ETIS-LaribPolypDB ^[1]	CVC-300 ^[36]
数量	1 000	612	380	196	60
图像大小	$500^+ \times 500^+$	384×288	574×500	$1\,225 \times 966$	574×500

表 2 实验的训练集与测试集分配

Tab. 2 Allocation of training dataset and test dataset of experiment

	训练集	测试集
实验 A	Kvasir-SEG(900),CVC-ClinicDB(550)	Kvasir-SEG(100)
实验 B	Kvasir-SEG(900),CVC-ClinicDB(550)	CVC-ClinicDB(62)
实验 C	Kvasir-SEG(900),CVC-ClinicDB(550)	ColonDB(380)
实验 D	Kvasir-SEG(900),CVC-ClinicDB(550)	ETIS-LaribPolypDB(196)
实验 E	Kvasir-SEG(900),CVC-ClinicDB(550)	CVC-300(60)

Step2. 模型训练。本文未采用传统数据增强方法,而使用多尺度训练方式,将数据分别以 352×352 大小的 0.75 倍、1 倍、1.25 倍输入模型进行训练,既达到数据增强目的,又缩短了训练时间。优化函数采用 RAdam^[38],学习率为 1×10^{-4} ,训练 30 轮后收敛;

Step3. 息肉图像分割。将数据集 Kvasir-SEG、CVC-ClinicDB、ColonDB、ETIS-LaribPolypDB、CVC-300 的测试图像输入模型得到分割结果,使用测试时图像增强(TTA)提升模型预测及泛化能力;

Step4. 分割图像后处理。采用构建的 TRW-S 算法对 TTA 模型输出结果进行后处理,提高分割图像边缘的连续性和分割区域内部的一致性;

Step5. 模型评价。

4 实验结果与分析

4.1 实验环境与评价指标

实验环境:Ubuntu18.04 LTS,Python3.6.5,pytorch2.1.0,CPU 为 Intel(R) Core(TM) i7-8700 K,频率为 3.70 GHz,GPU 为 GeForce GTX 1080。

将本文模型与 U-Net^[9]、U-Net++^[10]、SFA^[38]、ResUnet++^[11]、PraNet^[12]等模型进行对比实验,从定性和定量分析两方面评价算法的分割效果;并将本文模型与本文编码-解码模型、本文编码-解码+TTA 模型进行对比实验验证模型构建的合理性。

本文引入组合损失函数进行模型训练,该损失函数 L_{seg} 由加权交并比(L_{IoU}^w)^[39]和加权二元交叉熵(L_{BCE}^w)^[40]组成:

$$L_{\text{seg}} = L_{\text{IoU}}^w + \lambda L_{\text{BCE}}^w, \quad (18)$$

其中, λ 表示权重, 本文设为 1。

采用平均交并比 (mIoU)、平均置信度 (mDice)、结构性度量 (Smeasure)^[41] 及平均绝对误差 (MAE) 作为客观评价指标; 用平均交并比 (mIoU) 衡量算法分割的目标区域与人工分割的目标区域交叠率, mIoU 为每一类 IoU 结果的累加平均, 实现基于全局分割效果评价; mIoU $\in [0, 1]$, 值越大则分割效果越好。定义:

$$mIoU = \frac{1}{L} \sum_{i=1}^L \frac{TP}{TP + FP + FN}, \quad (19)$$

其中: L 为图像分类数, TP, FP, TN, FN 分别表示真阳性、假阳性、真阴性、假阴性。

用平均置信度 (mDice) 系数度量算法分割图像与人工分割的相似性; mDice 为每类置信度 (Dice) 累加平均, mDice $\in [0, 1]$, 值越大分割效果越好。定义:

$$mDice = \frac{1}{L} \sum_{i=1}^L \frac{2 * TP}{2 * TP + FP + FN}. \quad (20)$$

Smeasure^[41] 是结合图像结构相似性度量与区域结构相似性度量的分割评价指标, S_0 为图像结构相似性度量, S_r 为面向区域的结构相似性度量, α 取值为 0.5。Smeasure $\in [0, 1]$, 值越大分割效果越好。

$$S = \alpha * S_0 + (1 - \alpha) * S_r. \quad (21)$$

采用平均绝对误差 (MAE) 度量分割的偏差。MAE 为算法分割区域和人工分割区域值间绝对误差的平均值, 范围为 $[0, +\infty)$, 算法分割区域与人工分割区域重合时为 0, MAE 值越接近 0 表示分割效果越好。

$$MAE = \frac{1}{n} \sum_{i=1}^n \left| y_i - \hat{y}_i \right|, \quad (22)$$

式中, $\hat{y}_i$ 、 y_i 分别为算法分割区域和人工分割区域。

4.2 实验结果及对比分析

4.2.1 模型的单元模块对分割效果影响

4.2.1.1 ResNeSt-50 与 RFB 模块的消融实验

为讨论 ResNeSt-50 与 RFB 模块对整体分割性能的影响, 本文在 ColonDB 与 ETIS-Larib-PolypDB 数据集中进行消融实验。其中, 模型 a 为 UNet 模型; 模型 b 是将模型 a 的编码器改为 ResNeSt-50; 模型 c 为在模型 a 基础上加入 RFB 模块; 模型 d 为图 1 中 TTA 前的分割模型。实验结果如表 3 所示。模型 b 分割效果相比于模型 a,

表 3 ResNeSt-50 模块的消融实验 (训练集为 Kvasir-SEG, ClinicDB)

Tab. 3 Ablation experiment of ResNeSt-50 module (training on Kvasir-SEG and ClinicDB dataset)

模型	测试集	mDice	mIoU	Smeasure	MAE
模型 a	ColonDB	0.512	0.444	0.712	0.061
模型 b		0.608	0.523	0.760	0.052
模型 c		0.676	0.606	0.799	0.043
模型 d		0.757	0.675	n/a	n/a
模型 a	ETIS-LaribPolypDB	0.398	0.335	0.684	0.036
模型 b		0.522	0.447	0.738	0.026
模型 c		0.614	0.547	0.781	0.028
模型 d		0.706	0.628	n/a	n/a

mDice 及 mIoU 结果均有所提升; 模型 d 实验结果最好。如此实验效果得益于 ResNeSt-50 中拆分注意力结构能有效地聚焦不同分组, RFB 模块通过不同空洞感受野分支获取多尺度上下文信息, 提升了模型的 mIoU 值。消融实验验证了本文采用模块的有效性和方法的合理性。

4.2.1.2 多尺度训练及 RAdam 优化对模型分割影响

本文网络的特征图被缩小至原图的 1/32, 因此, 识别息肉图像中微小病变的特征存在一定困难; 为此, 通过输入更大、多尺度图像进行训练, 能在一定程度上解决这一问题, 提高模型对息肉图像中微小病变分割的鲁棒性; 表 4 的结果表明, 多尺度训练对提升本文模型分割精度是一种有效的途径。

表 4 多尺度训练对模型分割的影响

Tab. 4 Effect of multi-scale training on model segmentation

	mDice	mIoU	Smeasure	MAE
未采用多尺度训练	0.888	0.827	0.908	0.031
使用多尺度训练	0.913	0.858	0.925	0.023

数据集较小时, 自适应学习率优化器 Adam 在训练开始时学习率方差偏大, 此时其波动较大, 易使启动时因过度跳跃导致模型收敛至局部极小。RAdam^[38] 为防止在迭代早期学习率剧烈变化采用了整流器函数, 动态地开启或关闭自适应动量, 阻止其全速跳跃, 直至数据方差变化稳定。

由表 5,使用 RAdam 优化尽管比 Adam 收敛稍慢,但有利于模型收敛到一个合理值,实验结果的 mDice, mIoU, Smeasure, MAE 等评价指标均有所改善。

表 5 RAdam 优化对模型分割的影响

Tab. 5 Effect of RAdam optimizer on model segmentation

	Epoch	mDice	mIoU	Smeasure	MAE
Adam	20	0.904	0.851	0.920	0.027
RAdam	30	0.913	0.858	0.925	0.023

4.2.1.3 TTA 增强与 TRW-S 后处理对模型分割影响

图 6 给出在 CVC-ClinicDB^[33], Kvasir-SEG^[34]数据集上对本文编码-解码模型、本文编码-解码+TTA 模型及本文模型训练后,在 ETIS-Larib-PolypDB^[1]数据集上分割结果比较。本文模型采用数据增强和 TRW-S 后处理,使得模型对微小病变十分敏感,对微小病变定位准确性更高,分割图像区域一致性更好,在未知数据集中检测微小病变的效果较好;此外,本文模型对大病变组

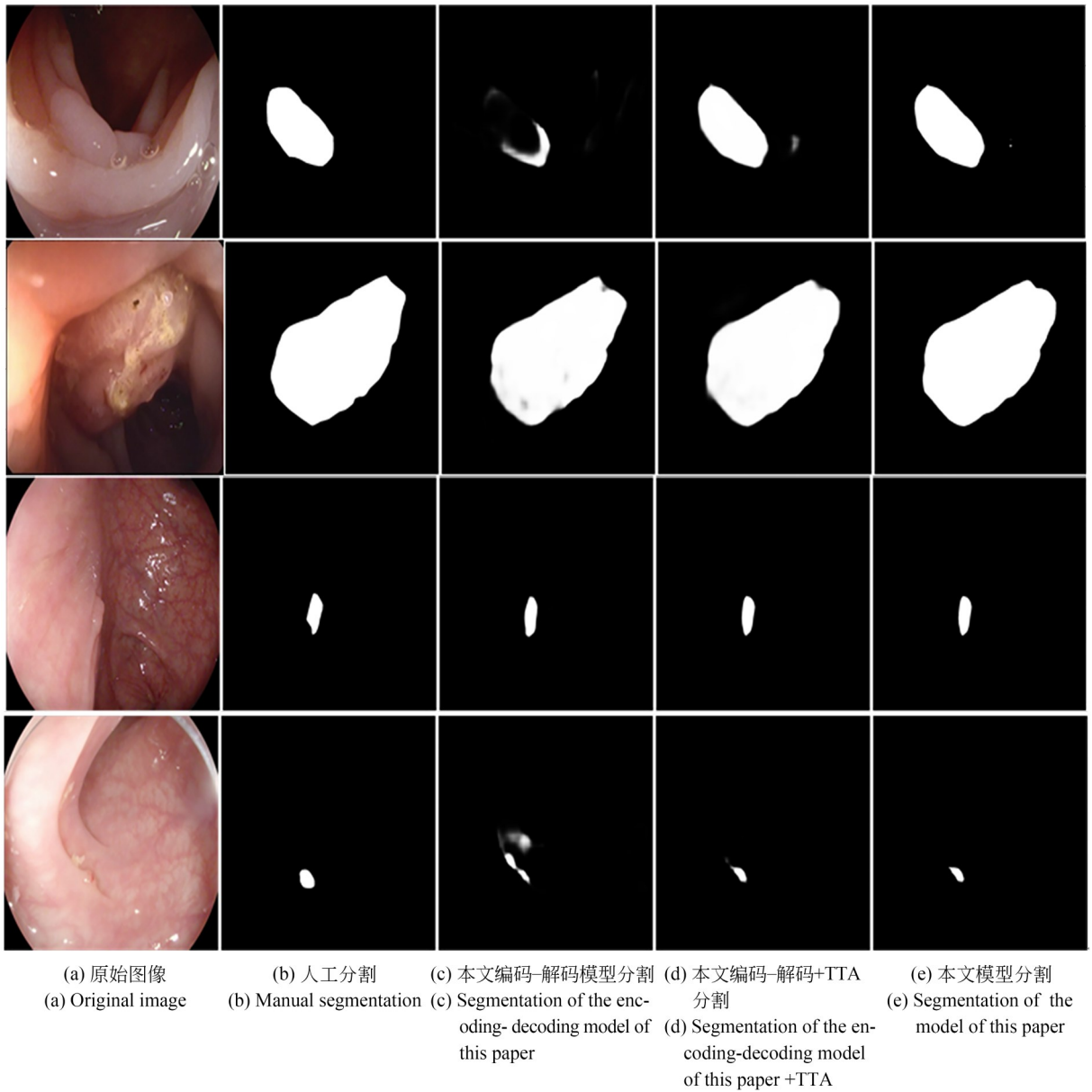


图 6 在 ETIS-LaribPolypDB^[1]数据集下测试的分割结果
Fig. 6 Segmentation result of testing on the ETIS-LaribPolypDB dataset

织的分割边缘亦较清晰;由此说明,模型中加入 TTA 模块及构建 TRW-S 算法,有利于提升模型检测微小病变的能力以及模型泛化能力。

由表 6,模型加入 TTA 模块后,在已知数据集中的 mDice 提高了 2%,mIoU 提升了 3%;因此,TTA 的加入有利于提升模型的预测准确率。相对于本文编码-解码模型和编码-解码+TTA 模型,本文模型的 mDice 提高 1% 以上,mIoU 提高 2% 以上;MAE 值与另两种算法一致,尽管 Smeasure 值有所降低,但仍高达 0.921,反映了本文模型采用 TRW-S 处理后,分割效果整体优于本文编码-解码模型和编码-解码+TTA 模型。

表 7 比较了三种模型在未知数据集上的表现。加入 TTA 模块后,mDice 提高了 1.1% 以上,mIoU 提高了 1.6% 以上,其中,在数据集 ETIS-LaribPolypDB^[1]上提升最多,相比于编码-解码模型,mDice 和 mIoU 分别提升了 5.8%,5.8%;从而说明 TTA 模块对提升模型的泛化能力的有效性。本文模型在三个未知数据集上性

表 6 在 CVC-ClinicDB^[33],Kvasir-SEG^[34]数据集中训练,Kvasir-SEG^[34]数据集中测试效果比较

Tab. 6 Trained on CVC-ClinicDB^[33],Kvasir-SEG^[34]dataset, compare the test results on Kvasir-SEG^[34]dataset

	mDice	mIoU	Smeasure	MAE
本文编码-解码器	0.913	0.858	0.925	0.023
本文编码-解码器+TTA	0.915	0.861	0.925	0.023
本文模型	0.916	0.863	0.921	0.023

能指标均有提升,在 ETIS-LaribPolypDB^[1]数据集上,相比于本文编码-解码模型,mDice 提高了 6.4% 以上,mIoU 提高了 6.3%;相比于本文编码-解码+TTA 模型,本文算法的 mDice 提高了 0.6%,mIoU 提高了 0.7%;在未知数据集 ColonDB 上,本文算法的 mDice 与 mIoU 均提高了 0.3% 以上;在未知数据集 CVC-300 上,mDice 与 mIoU 值亦均有不同程度的提升;结果表明,本文模型加入 TRW-S 后处理有利于提升模型的泛化能力。

表 7 在 CVC-ClinicDB^[33],Kvasir-SEG^[34]数据集下训练,分别在 ColonDB^[35],ETIS-LaribPolypDB^[1],CVC-300^[36]数据集中测试结果比较

Tab. 7 Trained on the CVC-ClinicDB^[33],Kvasir-SEG^[34]dataset, respectively compare the test results on ColonDB^[35], ETIS-LaribPolypDB^[1],CVC-300^[36]

	ColonDB ^[35]		ETIS-LaribPolypDB ^[1]		CVC-300 ^[36]	
	mDice	mIoU	mDice	mIoU	mDice	mIoU
本文编码-解码器	0.757	0.675	0.706	0.628	0.858	0.781
本文编码-解码器+TTA	0.783	0.700	0.764	0.686	0.869	0.797
本文模型	0.786	0.703	0.770	0.691	0.871	0.798

4.2.2 已知数据集中模型分割效果比较

对比实验的训练集采用 CVC-ClinicDB^[33],Kvasir-SEG^[34]数据集,测试集采用 Kvasir-SEG 数据集,分割效果如图 7 所示。

图 7 给出采用 U-Net^[9],U-Net++^[10],PraNet^[12],ResUnet++^[11],SFA^[38]以及本文模型的分割结果。从分割效果看,息肉图像经编码-解码模块分割、TTA 处理、TRW-S 处理后,输出较理想的分割图像;由图 7 知,本文算法不仅对大范围息肉病变有很好的分割效果,对小面积息肉病变分割效果亦较理想。本文算法分割图像的边缘清晰,且分割区域内部连贯、区域

内部一致性较好,分割图像边缘与轮廓清晰度、细节信息捕捉能力等方面均好于其他五种模型。

表 8 结果表明,本文算法相比于基准 PraNet^[12]算法,mDice 提高了 1.8%,mIoU 提高了 2.3%;Smeasure 提高了 0.6%,MAE 降低了 0.7%,因此,本文算法的四种评价指标均有不同程度改善。相比于 U-Net^[9],U-Net++^[10],ResUnet++^[11],SFA^[38]等算法,本文模型分割效果的评价指标提高更多;从而说明了本文算法对息肉图像分割的可行性和优势。

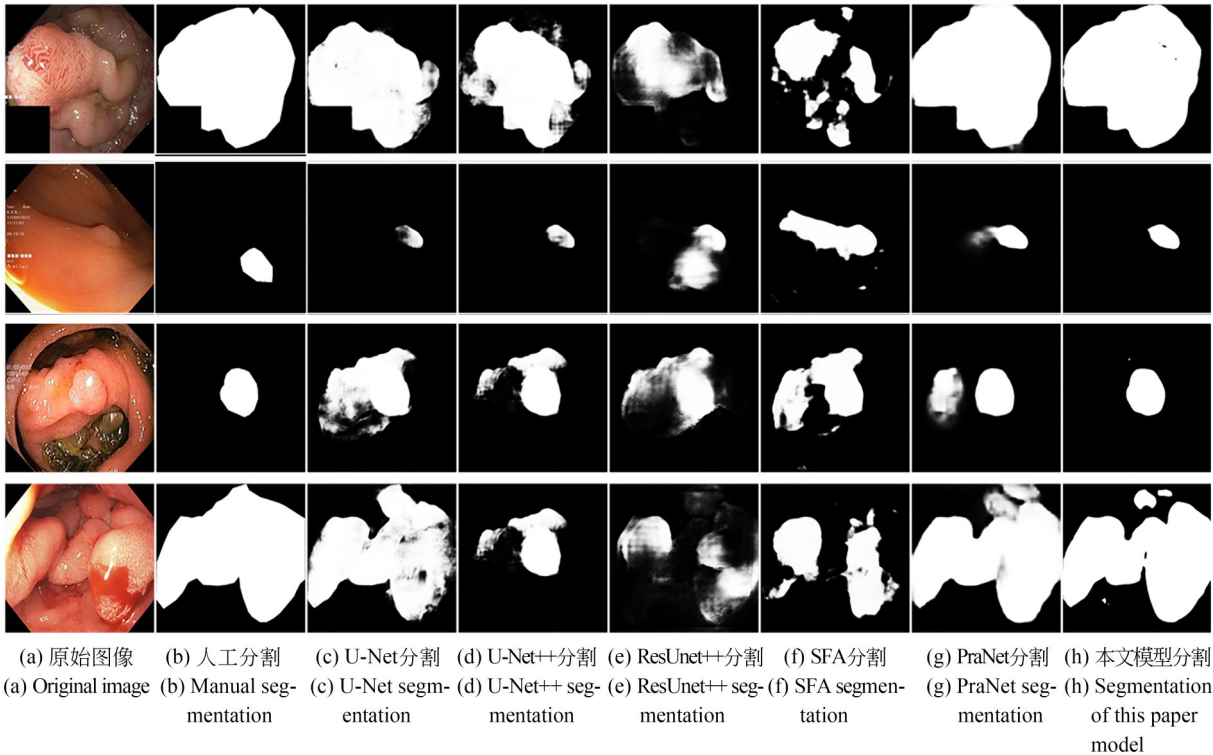


图 7 在 Kvasir-SEG^[34]数据集下测试的分割结果
Fig. 7 Segmentation result of testing on the Kvasir-SEG^[34] dataset

表 8 在 CVC-ClinicDB ^[33] , Kvasir-SEG ^[34] 数据集中训练, Kvasir-SEG ^[34] 数据集中测试效果比较				
Tab. 8 Trained on CVC-ClinicDB ^[33] , Kvasir-SEG ^[34] dataset, compare the test results on Kvasir-SEG ^[34] dataset				
	mDice	mIoU	Smeasure	MAE
U-Net ^[9]	0. 818	0. 746	0. 858	0. 055
U-Net++ ^[10]	0. 821	0. 743	0. 862	0. 048
ResUnet++ ^[11]	0. 734	0. 639	n/a	n/a
SFA ^[38]	0. 723	0. 611	0. 782	0. 075
PraNet ^[12]	0. 898	0. 840	0. 915	0. 030
本文方法	0. 916	0. 863	0. 921	0. 023

4. 2. 3 模型检测微小病变能力比较

为验证本文模型检测微小病变的能力,在 550 幅 CVC-ClinicDB^[33] 图像、900 幅 Kvasir-SEG^[34] 图像中对 U-Net^[9], U-Net++^[10], ResUnet++^[11], SFA^[38], PraNet^[12] 及本文模型进行训练后,在数据集 ETIS-LaribPolypDB^[1] 中进行测试;由图 8 知,本文算法对于微小病变具有较高的检出能力。

由图 8 和表 9,对微小病变而言,本文模型定

位准确,而 PraNet^[12] 等模型几乎无法找到微小病变位置,分割效果也差强人意;此外,采用本文模型分割的各项指标均优于其他 5 种模型,病变边缘的分割亦十分清晰。

图 9 采用箱线图比较 8 组数据中,本文模型与 PraNet 模型分割微小病变的置信度(Dice)性能;8 组数据按照病变由小到大进行分组,由图可知,本文模型的分割效果整体优于 PraNet 模型,且病变组织越大检测效果越好。

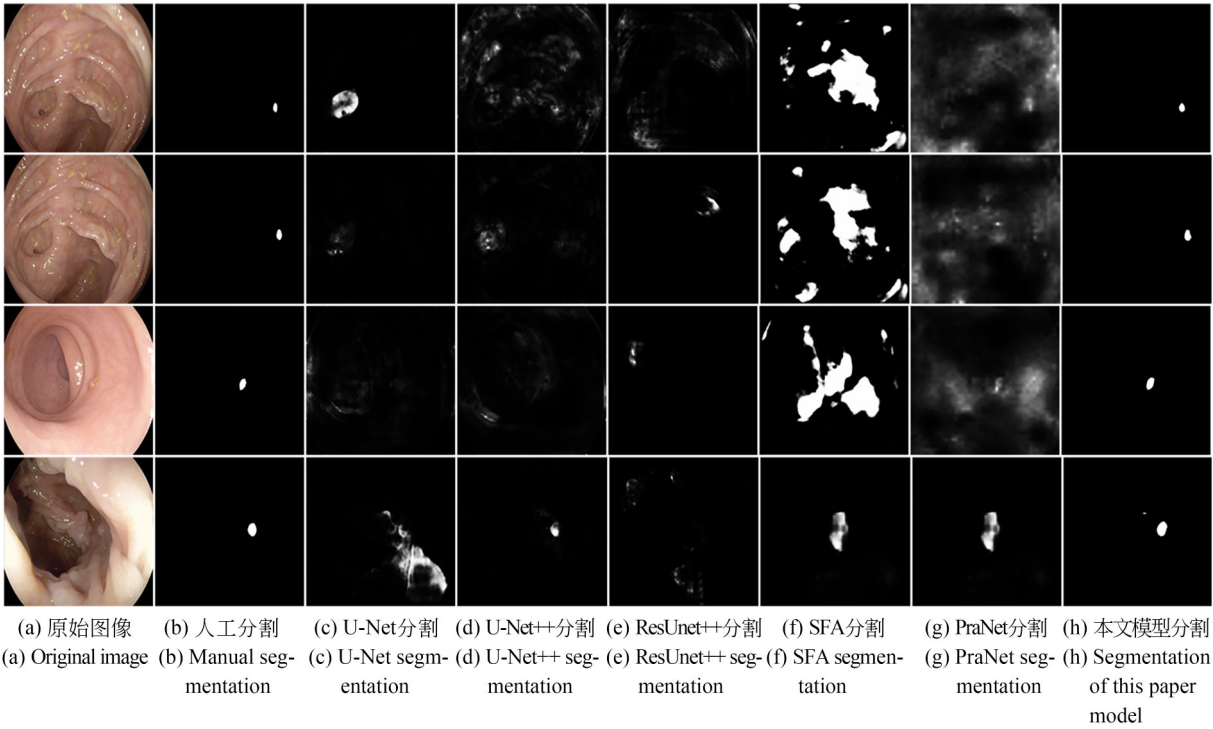


图 8 ETIS-LaribPolypDB^[1]数据集中测试微小病变的分割结果
Fig. 8 Segmentation result of testing small lesion on ETIS-LaribPolypDB^[1]dataset

表 9 在 CVC-ClinicDB^[33],Kvasir-SEG^[34]数据集下训练,在 ETIS-LaribPolypDB^[1]数据集中测试效果比较
Tab. 9 Trained on CVC-ClinicDB, Kvasir-SEG dataset, compare the test results on ETIS-LaribPolypDB dataset

	mDice	mIoU	Smeasure	MAE
U-Net ^[9]	0.055	0.046	0.517	0.021
U-Net++ ^[10]	0.093	0.078	0.545	0.018
SFA ^[38]	0.012	0.006	0.425	0.130
ResUnet++ ^[11]	0.016	0.009	0.500	0.011
PraNet ^[12]	0.099	0.073	0.515	0.074
本文方法	0.487	0.401	0.734	0.003

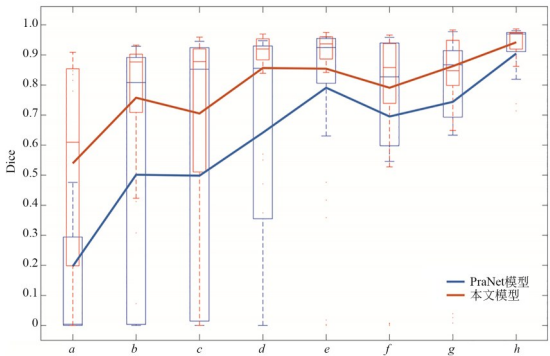


图 9 本文模型与 PraNet 模型分割微小病变的效果比较
Fig. 9 Comparison of the effect of the model used in this paper and PraNet model in segmenting minimal lesions

4. 2. 4 模型的泛化能力比较

在 U-Net^[9], SFA^[38], U-Net++^[10], ResUnet++^[11], PraNet^[12] 以及本文模型中采用在 1 450 幅 Kvasir-SEG^[34] 和 CVC-ClinicDB^[33] 数据集组合而成的训练集上进行训练,训练后模型分别用 380 幅 ColonDB^[35] 图像、196 幅 ETIS-LaribPolypDB^[1] 图像、以及 60 幅 CVC-300^[36] 图像等未知数据集进行测试,检验本文模型的泛化能力。如图 10 和表 10 所示。

测试结果的定性和定量评价表明,本文算法在这三个未知数据集上分割效果均有提升,特别是在 ETIS-LaribPolypDB^[1] 数据集上,相比于

PraNet^[12] 基准算法,本文模型的 mDice 提高了 16.4%,mIoU 提高了 14.5%;说明,本文模型具

有很好的泛化能力,并具有较强的分割微小病变的能力。

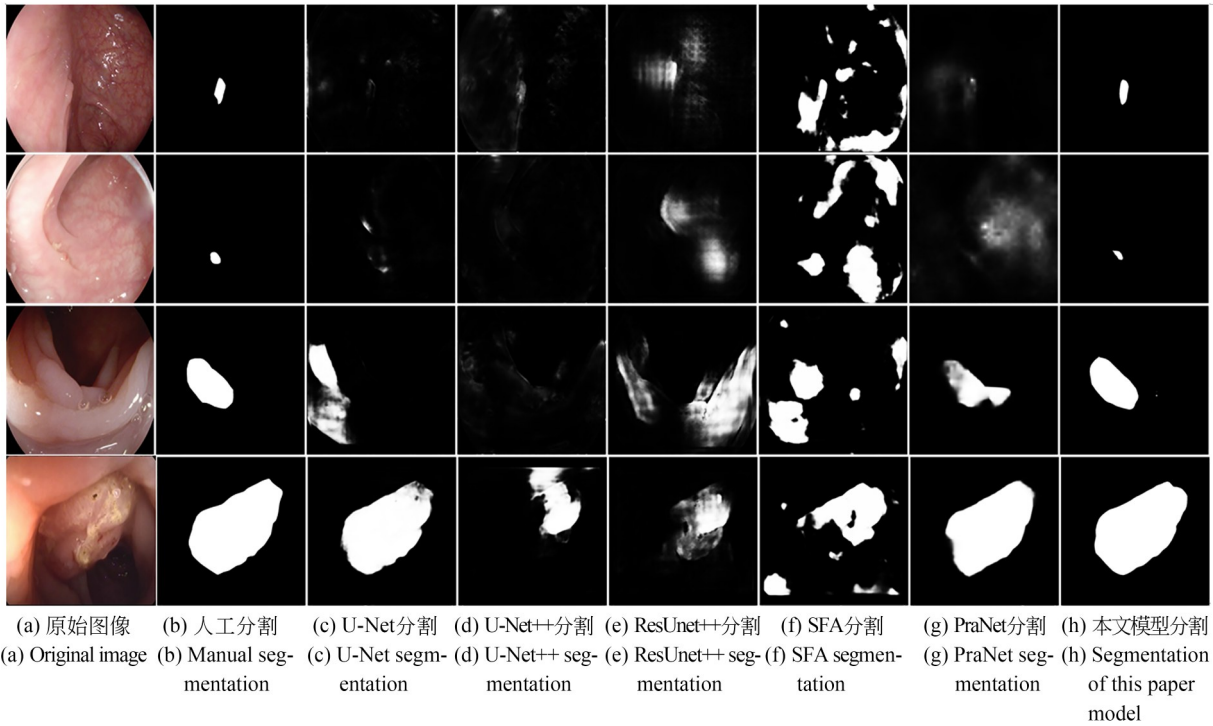


图 10 ETIS-LaribPolypDB^[1]数据集中测试的分割结果

Fig. 10 Segmentation result of testing on the ETIS-LaribPolypDB^[1]dataset

表 10 在 CVC-ClinicDB^[33],Kvasir-SEG^[34]数据集下训练,分别在 ColonDB^[35]、ETIS-LaribPolypDB^[1],CVC-300^[36]数据集下测试结果比较

Tab. 10 Trained on the CVC-ClinicDB^[33], Kvasir-SEG^[34]dataset, respectively compare the test results on ColonDB^[35], ETIS-LaribPolypDB^[1], CVC-300^[36]dataset

	ColonDB ^[35]		ETIS-LaribPolypDB ^[1]		CVC-300 ^[36]	
	mDice	MIoU	mDice	MIoU	mDice	MIoU
U-Net ^[9]	0. 512	0. 444	0. 398	0. 335	0. 71	0. 627
U-Net++ ^[10]	0. 483	0. 410	0. 401	0. 344	0. 707	0. 624
ResUnet++ ^[11]	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
SFA ^[38]	0. 469	0. 347	0. 297	0. 217	0. 467	0. 329
PraNet ^[12]	0. 709	0. 640	0. 628	0. 567	0. 871	0. 797
本文方法	0. 786	0. 703	0. 770	0. 691	0. 871	0. 798

5 结 论

本文提出了多尺度 ResNeSt-50 聚合网络进行粗分割与马尔科夫随机场的 TRW-S 算法进行再分割的两阶段息肉图像分割方法。针对病变区域与正常组织间边界区分度不高的问题,构建

编、解码多尺度聚合网络提升骨干网络的表达能力,捕捉和提取息肉图像的弱特征信息,网络中密集聚合模块融合不同尺度的特征图信息,实现上下文信息的关联,采用 TTA 模块提升模型预测及泛化能力,最后,构建马尔科夫随机场的 TRW-S 算法强化分割边缘,增强息肉内部分割

区域内部的一致性和边缘的连续性,实现息肉图像的有效分割。

理论推导和对比实验验证了本文模型在息肉图像分割的有效性,本文模型能够有效地和精确地实现息肉病灶区域的分割,且具有良好的泛

化能力和检测微小病变的能力,有助于提升医生临床诊断中诊断的准确度和精确率。后续研究将在降低模型的参数量、提高模型的推理速度等方面进一步优化,并进一步关注模型的泛化能力,提升模型的分割性能。

参考文献:

- [1] SILVA J, HISTACE A, ROMAIN O, *et al.* Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer[J]. *International Journal of Computer Assisted Radiology and Surgery*, 2014, 9(2): 283-293.
- [2] 郑荣寿, 孙可欣, 张思维, 等. 2015年中国恶性肿瘤流行情况分析[J]. *中华肿瘤杂志*, 2019, 41(1): 19-28.
ZHENG R S, SUN K X, ZHANG S W, *et al.* Report of cancer epidemiology in China, 2015[J]. *Chinese Journal of Oncology*, 2019, 41(1): 19-28. (in Chinese)
- [3] ARAN V, VICTORINO AP, THULER LC, *et al.* Colorectal cancer: epidemiology, disease mechanisms and interventions to reduce onset and mortality[J]. *Clinical Colorectal Cancer*, 2016, 15(3): 195-203.
- [4] 江贵平, 秦文健, 周寿军, 等. 医学图像分割及其发展现状[J]. *计算机学报*, 2015, 38(6): 1222-1242
JIANG G P, QIN W J, ZHOU S J, *et al.* State-of-the-art in medical image segmentation[J]. *Chinese Journal of Computers*, 2015, 38(6): 1222-1242 (in Chinese)
- [5] 田娟秀, 刘国才, 谷珊珊, 等. 医学图像分析深度学习研究方法研究与挑战[J]. *自动化学报*, 2018, 44(3): 401-424
TIAN J X, LIU G C, GU S S, *et al.* Deep learning in medical image analysis and its challenges[J]. *Acta Automatica Sinica*, 2018, 44(3): 401-424 (in Chinese)
- [6] 李少东, 洪森, 李世权等. 人工智能利用图像深度学习诊疗消化道肿瘤的现状与展望[J]. *中国普外基础与临床杂志*, 2021, 28(11): 1524-1529.
LI S H D, HONG S, LI S H Q, *et al.* Current situation and prospect of artificial intelligence in the diagnosis and treatment of gastrointestinal tumors using image deep learning[J]. *Chinese Journal of Bases and Clinics in General Surgery*, 2021, 28(11): 1524-1529.
- [7] 吴锦珍, 谢玥, 王新颖. 人工智能在结直肠息肉性质鉴别中的应用进展[J]. *中华消化内镜杂志*, 2020, 37(11): 849-852.
WU J Z, XIE Y, WANG X Y. Application progress of artificial intelligence in identification of colorectal polyps[J]. *Chinese Journal of Digestive Endoscopy*, 2020, 37(11): 849-852. (in Chinese)
- [8] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640-651.
- [9] RONNEBERGER O, FISCHER P, BROX T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*[M]. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234-241.
- [10] ZHOU Z, SIDDIQUEE M M R, TAJBAKHSH N, *et al.* UNet++: a Nested U-net Architecture for Medical Image Segmentation[EB/OL]. 2018: *arXiv*: 1807.10165. <https://arxiv.org/abs/1807.10165>
- [11] 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)*. 9-11, 2019, San Diego, CA, USA. IEEE, 2020: 225-2255.
- [12] FAN D P, JI G P, ZHOU T, *et al.* *PraNet: Parallel Reverse Attention Network for Polyp Segmentation*[M]. Medical Image Computing and Computer Assisted Intervention-MICCAI 2020. Cham: Springer International Publishing, 2020: 263-273.
- [13] 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)*. 28-30, 2020, Rochester, MN, USA. IEEE, 2020: 558-564.
- [14] 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.
- [15] ZHANG R F, LI G B, LI Z, *et al.* *Adaptive Context Selection for Polyp Segmentation*[M]. Medical Image Computing and Computer Assisted Intervention - MICCAI 2020. Cham: Springer International Publishing, 2020: 253-262.
- [16] 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.
- 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. (in Chinese)
- [17] 撒子奇, 刘巧红, 凌晨, 等. 结合 HarDNet 和反向注意力的息肉分割方法[J]. *激光与光电子学进展*, 2023, 60(2): 3788/LOP212665.
- HAN Z Q, LIU Q H, LING C, *et al.* Polyp segmentation method combining HarDNet and reverse attention[J]. *Laser & Optoelectronics Progress*, 2023, 60(2): 3788/LOP212665. (in Chinese)
- [18] ZHANG H, WU C R, ZHANG Z Y, *et al.* ResNeSt: split-attention networks [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 19-20, 2022, New Orleans, LA, USA. IEEE, 2022: 2735-2745.
- [19] LIU S T, HUANG D, WANG Y H. *Receptive Field Block Net for Accurate and Fast Object Detection* [M]. Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 404-419.
- [20] ZHU Z, BIAN Z, HOU J, *et al.* When Residual Learning Meets Dense Aggregation: Rethinking the Aggregation of Deep Neural Networks [EB/OL]. 2020: *arXiv*: 2004.08796. <https://arxiv.org/abs/2004.08796>
- [21] YAMADA M, SAITO Y, IMAOKA H, *et al.* Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy[J]. *Scientific Reports*, 2019, 9: 14465.
- [22] WAINWRIGHT M J, JAAKKOLA T S, WILLISKY A S. MAP estimation via agreement on trees: message-passing and linear programming [J]. *IEEE Transactions on Information Theory*, 2005, 51(11): 3697-3717.
- [23] SZELISKI R, ZABIH R, SCHARSTEIN D, *et al.* A comparative study of energy minimization methods for Markov random fields with smoothness-based priors[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(6): 1068-1080.
- [24] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [25] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [26] 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)*. 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 510-519.
- [27] XIE S N, GIRSHICK R, DOLLÁR P, *et al.* Aggregated residual transformations for deep neural networks[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 5987-5995.
- [28] ZHANG H, ZU K K, LU J, *et al.* *EPSANet: an Efficient Pyramid Squeeze Attention Block on Convolutional Neural Network*[M]. Computer Vision-ACCV 2022. Cham: Springer Nature Switzerland, 2023: 541-557.
- [29] WU Z, SU L, HUANG Q M. Cascaded partial decoder for fast and accurate salient object detection [C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 3902-3911.
- [30] KOLMOGOROV V. Convergent tree-reweighted message passing for energy minimization [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(10): 1568-1583.
- [31] YANG X, GAO X B, TAO D C, *et al.* An efficient MRF embedded level set method for image segmentation [J]. *IEEE Transactions on Image*

- Processing*, 2015, 24(1): 9-21.
- [32] 夏平, 施宇, 雷帮军, 等. 复小波域混合概率图模型的超声医学图像分割[J]. 自动化学报, 2021, 47(1): 185-196.
- XIA P, SHI Y, LEI B J, *et al.* Ultrasound medical image segmentation based on hybrid probabilistic graphical model in complex-wavelet domain[J]. *Acta Automatica Sinica*, 2021, 47(1): 185-196. (in Chinese)
- [33] BERNAL J, SÁNCHEZ FJ, 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.
- [34] JHA D, SMEDSRUD P H, RIEGLER M A, *et al.* *Kvasir-SEG: a Segmented Polyp Dataset*[M]. MultiMedia Modeling. Cham: Springer International Publishing, 2019: 451-462.
- [35] TAJBAKSH N, GURUDU S R, LIANG J M. Automated polyp detection in colonoscopy videos using shape and context information[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(2): 630-644.
- [36] VÁZQUEZ D, BERNAL J, *et al.* A benchmark for endoluminal scene segmentation of colonoscopy images[J]. *Journal of Healthcare Engineering*, 2017, 2017: 4037190.
- [37] LIU L, JIANG H, HE P, *et al.* On the variance of the adaptive learning rate and beyond[J]. *arXiv e-prints*, 2019: arXiv: 1908.03265.
- [38] FANG Y Q, CHEN C, YUAN Y X, *et al.* *Selective Feature Aggregation Network with Area-Boundary Constraints for Polyp Segmentation* [M]. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019: 302-310.
- [39] QIN X B, ZHANG Z C, HUANG C Y, *et al.* BASNet: boundary-aware salient object detection [C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 7471-7481.
- [40] WEI J, WANG S H, HUANG Q M. F³Net: fusion, feedback and focus for salient object detection [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 12321-12328.
- [41] FAN D P, CHENG M M, LIU Y, *et al.* Structure-measure: a new way to evaluate foreground maps[C]. 2017 *IEEE International Conference on Computer Vision (ICCV)*. 22-29, 2017, Venice, Italy. IEEE, 2017: 4558-4567.

作者简介:



夏平(1967—),男,教授,1998年于华中理工大学获工学硕士学位,现为三峡大学计算机与信息学院教授。研究方向:计算机视觉、智能信息处理、概率图模型及其应用。E-mail: pxia@ctgu.edu.cn

通讯作者:



雷帮军(1973—),男,博士,2003年获荷兰德尔夫特理工大学博士学位,现为三峡大学计算机与信息学院教授,德尔夫特理工大学(荷兰)客座高级研究员,湖北省“楚天学者”、“百人计划”人选,研究方向:计算机视觉、图像处理、模式识别。E-mail: Bangjun.Lei@ieee.org