

周卉, 武俊珂. 具有深层特征约束的轻量化胃粘膜下肿瘤分割网络[J]. 智能计算机与应用, 2025, 15(11): 191–199. DOI: 10.20169/j. issn. 2095-2163. 24031801

具有深层特征约束的轻量化胃粘膜下肿瘤分割网络

周 卉¹, 武俊珂²

(1 上海理工大学 理学院, 上海 200093; 2 上海理工大学 管理学院, 上海 200093)

摘要: 针对胃黏膜下肿瘤因层次较多且边缘难分辨等导致分割精度低的问题, 该文提出一种具有深层特征约束的轻量化胃黏膜下肿瘤分割网络。通过引入扩增感受野模块, 在保留特征丰富性的同时大幅减少模型参数量; 结合轻量化空间和通道重构卷积和无参数注意力机制对特征提取模块进行改进, 提升模型性能; 通过有效融合深层与浅层特征, 得到更精细化的边缘信息和定位信息。在该数据集上, 该方法的平均 Dice 系数为 0.778, 提高了 8.2%; 平均交并比为 0.679, 提高了 8.3%; 边界 F 测度、结构相似性度量、平均误差、最大误差以及平均绝对误差等多个不同的指标均提高了 3%~8%, 模型参数量减少为原来的 54%。实验结果证明该方法能有效减少参数量, 并提高胃黏膜下肿瘤分割的精准度和细致性。

关键词: 胃黏膜下肿瘤; 肿瘤分割; 轻量化; 深浅层融合

中图分类号: TP391 文献标志码: A 文章编号: 2095-2163(2025)11-0191-09

A lightweight submucosal tumor segmentation network with deep feature constraints

ZHOU Hui¹, WU Junke²

(1 College of Science, University of Shanghai for Science and Technology, Shanghai 200093, China;
2 Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: To solve the problem of low segmentation accuracy of submucosal tumors due to multiple layers and indistinguishable edges, a lightweight submucosal tumor segmentation network with deep feature constraints was proposed in this paper. This method introduced an expanded receptive field module to significantly reduce the number of model parameters while retaining feature richness. It combined lightweight spatial-channel reconstruction convolution and parameter-free attention mechanism to improve the feature extraction module to improve the model performance. Through the effective fusion of deep and shallow features, more refined edge information and positioning information can be obtained. On this dataset, the proposed method achieves a mean Dice coefficient of 0.778, an increase of 8.2%. The mean intersection over union score is 0.679, an increase of 8.3%. Various metrics, including boundary F-measure, structural similarity, average error, maximum error, and mean absolute error have increased by 3% to 8%, and the number of model parameters has been reduced by 54%. Experimental results validate that this method can effectively reduce the number of parameters while improving the accuracy and detail of submucosal tumor segmentation.

Key words: gastric submucosal tumors; tumor segmentation; lightweight; deep and shallow layer fusion

0 引言

随着公众对自身健康认知的提高, 每年胃黏膜下肿瘤^[1] (Submucosal Tumor, SMT) 的检出率不断提高, 由于此类肿瘤类型繁多, 情况复杂, 如何更精准的判断肿瘤的位置以及定性是目前在临床医学上需要解决的紧迫问题。SMT 是源于消化道粘膜层

的一类病变的总称, 大多数源于粘膜肌层、黏膜下层和固有肌层均被正常粘膜覆盖, 常见的类型有平滑肌瘤, 胃肠道间质瘤、脂肪瘤、异位胰腺等, 其分类定性需要基于肿瘤的在粘膜层的位置^[2]。

超声内镜(Endoscopic Ultrasonography, EUS)是诊断 SMT 的首选方法^[3], 本文使用的数据集均通过 EUS 得到, 能清晰显示病变壁及其周围结构, 并得到

基金项目: 国家自然科学基金(62273239)。

作者简介: 武俊珂(1999—), 男, 博士研究生, 主要研究方向: 医学图像处理, 视觉定位。

通信作者: 周 卉(1999—), 女, 硕士研究生, 主要研究方向: 医学图像处理及人工智能。Email:hannazhz99@126.com。

收稿日期: 2024-03-18

哈尔滨工业大学主办 ◆ 科技创见与应用

肿瘤位置的超声图像。EUS 结合了超声波和内窥镜检查的优点。通过内窥镜观察 SMT 的外部形态,用小超声探头^[4]扫描肿瘤。此外,EUS 可以探测 SMT 起源层的信息,帮助评估其病理^[5],其还提供了关于膜完整性和其他预测手术并发症和患者预后^[6]的因素等有价值的信息。然而传统方法对该类肿瘤的定性识别都来源于医生的主观判断,仅仅依靠医生的专业知识能力和影像学经验,但缺乏经验的医生经常会产生难以判断的情况。因此,SMT 的判断结果往往存在着一些错误,这对医生和患者来说都是影响巨大的,会造成后续手术治疗等情况的风险性和不确定性^[7]。

近年来,由于计算机技术的飞速发展,人工智能与深度学习技术已逐渐渗透到各行各业中,并大大推动了多领域的技术进步和革新。在医学图像邻域中深度学习技术也取得了令人瞩目的应用成果^[8~10]。在影像学分割领域,利用深度学习进行图像分割也已经成为了研究的热点,从海量的医学数据中获取相关信息,能帮助医生在对医学图像分析和诊断时提供较为专业的建议^[11]。许多研究表明,利用深度学习技术,医学图像分割的准确性和效率都有了显著的提高。近期国内外研究都通过对模型的深度优化或多模态数据融合等方法^[12],提高了肿瘤分割的准确性,为解决临床问题以及后续的治疗提供了参考信息。

在众多的医学图像分割模型中,UNet^[13]是最具代表性的网络结构,也是应用最广泛的。其在全卷积神经网络^[14]的基础上通过其特有的上采样和下采样机制,显著增强了对医学图像中细节的捕获能力。由 Zhou 等^[15]提出的 UNet++则在 UNet 的基础上,加入了多尺度融合和深度链接,增强了模型的表达能力和分割精度。由 Huang 等^[16]提出的 UNet3+采用了全面跳跃连接的架构且增加了注意力机制,但全尺度的特征融合会导致过多的冗余计算,降低模型效率。近几年,随着医学图像数据的逐渐丰富,研究者们已经提出了众多基于这些经典模型的改进策略,并在各种医学图像任务中取得了优异的结果。

在肿瘤分割领域,由 Fan 等^[17]提出 PraNet 是一种聚集高层的语义特征且关注边缘的图像分割网络结构,在结肠息肉分割任务中的表现较好,但对于边界信息较弱的图像处理能力还有待提升,且该网络计算成本较大。由 Wei 等^[18]提出的 SANet 针对息肉分割任务中的颜色分布差异提出一种颜色交换的增广策略,可捕捉到肿瘤细微区别,但不适用于颜色

区别不明显的图像。由 Dong 等^[19]提出 Polyp-PVT 是一种针对结肠镜检查图像的融合跨层特征的网络,能有效抑制特征中的噪声并找到息肉的隐藏特征,但对于边缘细节识别不够准确。

虽然卷积神经网络(Convolutional Neural Networks,CNN)在各种计算机任务中取得显著的性能,包括医学图像分割领域,但通常以大量的计算资源为代价。由 Li 等^[20]提出的利用特征之间空间和通道冗余对 CNN 压缩的卷积模块 SCCConv,通过重构卷积的方式来抑制模型计算量,且是一个即插即用的模块,本文方法中也使用到该模块,作为轻量化策略的一部分。此外,由 Li 等^[21]在 UNet++ 的基础上提出了 attention UNet++,在其卷积块之间引入注意力机制,使得对特征的针对性更强,大大加快了网络的训练速度。同样是减轻注意力机制的计算,由 Yang 等^[22]提出的简单无参数的卷积神经网络注意模块 SAM,不需要在原始网络中添加参数,而是通过评估每个神经元的重要性,提高卷积网络的性能。

尽管上述研究都取得了显著的进展,但依然存在一些不足之处。由于肿瘤中的层次和器官等形状复杂多样,且可能存在遮掩或重叠的情况,致使错误分割的概率很大,影响医学判断。其中最为突出的问题是,这些模型往往参数量庞大,导致计算量增加,实际应用中难以满足快速处理的需求。一些基于 UNet 做出改进的网络结构^[23~25],虽然能提高分割结果的准确性,但模型的冗余性有待提高。此外,对于一些细小的肿瘤结构,模型往往处理不够精准,可能会导致误诊或漏诊。

为了解决上述问题,本文设计了一种具有深层特征约束的轻量化胃粘膜下肿瘤分割网络。针对胃黏膜下肿瘤因种类复杂多样、分层不清晰且边缘难界定等问题,基于 UNet++模型进行改进,以减少模型参数量并提高模型对 SMT 分割的精准度和细致性。首先,本文方法通过引入扩增感受野(Receptive Field Block, RFB)模块,在保留特征丰富性的同时大幅减少模型参数量。其次,本文开发了一种轻量化无参数的重构卷积 (Lightweight Reconstruction Convolution with parameter-free Attention, LRCA) 模块,不仅对特征提取模块进行了有效改进,还提升了模型的整体性能。最后,为了使本文方法能够有效融合深层与浅层特征,设计了全局平均池化融合 (Global Average Pooling Fusion, GF) 模块,从而获取更精细化的边缘信息和定位信息。在针对本数据集的实验中,本文方法展示了显著的性能提升:平均

Dice 系数达到了 0.778, 相比原先提高了 8.2%; 平均交并比提升至 0.679, 增长了 8.3%。此外, 在边界 F 测度、结构相似性度量、平均误差、最大误差以及平均绝对误差等多个不同的性能指标上, 均实现了 3% 至 8% 的提升, 同时模型参数量降低至原来的 54%。这些实验结果不仅证明了该方法在减少参数量方面的有效性, 也展现了其在提高胃黏膜下肿瘤分割的精准度和细致性方面的显著优势。

1 具有深层特征约束的轻量化胃粘膜下肿瘤分割网络

1.1 整体网络结构

本文所提的针对胃黏膜下肿瘤分割的具有深层约束浅层特点的轻量化网络整体结构如图 1 所示。基于 Unet++ 网络的基本思想进行改进, 针对编码器设计了轻量化的编码策略和特征提取模块, 同时针对解码器设计了基于全局平均池化的深浅层特征融合模块。

对于输入图像在经过编码器时, 增加了扩增感受野 RFB 模块, 实现对骨干网络每一层输出的多尺度特征提取。在特征提取的过程中, 本文开发了一种轻量化无参数的重构卷积 LRCA 模块。对于每一个节点进行密集卷积操作的同时, 该模块结合了空间和通道重构卷积 (Spatial and Channel Reconstruction Convolution, SCConv) 和无参数注意力机制 (Simple Attention Module, SAM), 以此减少模型冗余计算并提高特征学习的有效性。此外, 本文方法对特征图的融合方式进行了创新性调整, 传统的沿通道拼接方式被替换为基于注意力机制的像素点乘法操作, 以提升模型在肿瘤分割识别中的性能。由于深层特征可以使浅层特征得到更精细化的背景信息和定位信息, 本文方法针对解码器设计了全局平均池化融合 (Global Average Pooling Fusion, GF) 模块, 在原本输出的每层特征图之后将其进行由深层到浅层的进一步融合, 以此获得细化的肿瘤边缘细节。

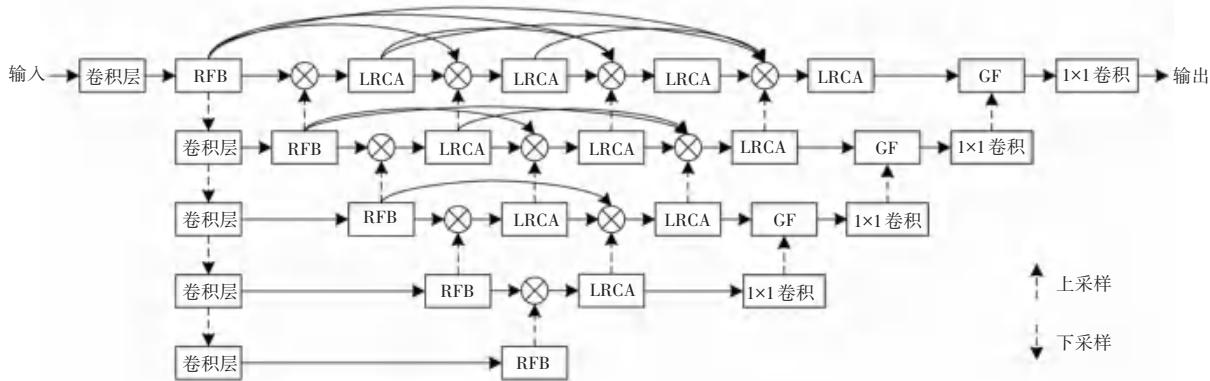


图 1 整体网络结构

Fig. 1 Network architecture

1.2 扩增感受野模块

为了实现精度与网络速度的平衡, 我们使用了 RFB 模块实现感受野扩增, 如图 2 所示。该模块具有多个分支, 每个分支的第一层都由特定大小卷积核的卷积核构成, 例如图上的 1×1 卷积, 3×3 卷积, 5×5 卷积。同时, RFB 模块为了增加感受野引入了空洞卷积, 空洞卷积本就是应用在分割网络^[26]中, 应用于本任务是合适的, 使用其以获得更大的感受野。该模块被应用于骨干网络中, 对每一层的输出进行有效的多尺度特征提取。

RFB 模块的设计结构源自于视觉系统中的感受野机制, 其目的是有效捕捉图像中的多尺度信息。这一特性对于肿瘤分割任务尤为关键, 鉴于肿瘤在形态和尺寸上的多样性, RFB 能够提取这些复杂形

态的关键特征, 进而提高识别的准确性。该方法不仅增强了模型对于不同尺寸输入的处理能力, 同时也确保了特征的丰富性和细粒度信息的捕获。此外, RFB 通过优化特征提取过程, 还有助于提升模型的计算效率, 这对于实时医学影像处理至关重要。

在 RFB 模块中, 最后将不同尺寸和感受野的输出特征图进行 Concat 操作, 以达到融合不同特征的目的。但本文方法中为了进一步降低计算负担, 将该模块的输出通道数设定为 32, 显著减少模型的参数量, 从而提升了模型的推理速度, 同时也减轻了对计算资源的依赖。最后, 通过这种结合了 RFB 和优化通道数的设计, 本文模型在保持高精度检测的同时, 实现了对推理速度的显著提升, 为实时应用提供了可能。

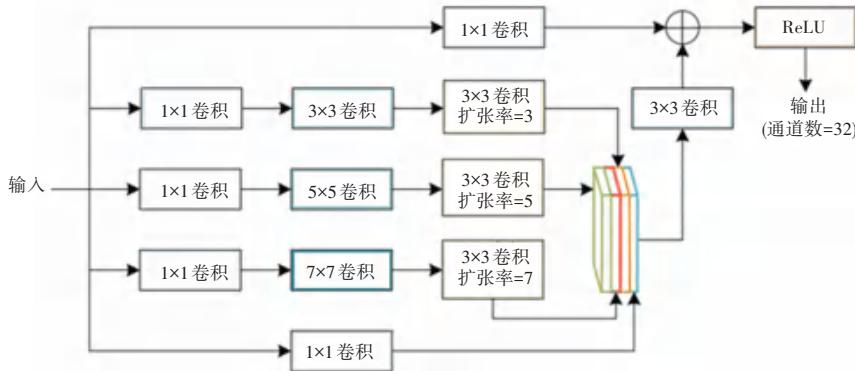


图 2 RFB 网络结构图

Fig. 2 Network architecture of RFB

1.3 轻量化无参数的重构卷积模块

为了有效减少模型参数量并提升特征提取的有效性,本文设计了轻量化无参数的重构卷积 LRCA 模块,如图 3 所示。该模块包涵了空间和通道重构卷积(SCCConv)模块以及无参数注意力机制(SAM)。

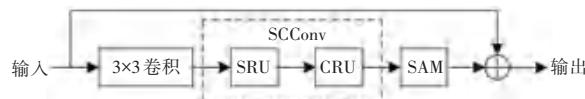


图 3 LRCA 模块网络结构图

Fig. 3 Network architecture of LRCA block

LRCA 模块由两个卷积层构成,第一个卷积层(Convol1)为 3×3 卷积,第二个卷积层(Convol2)为带有

SCConv 和 SAM 的残差卷积层,具体的计算步骤如下:

$$\text{Conv2} = \text{SAM}(\text{SCConv}(x)) + x \quad (1)$$

其中,SCConv 模块包括两个关键部分,空间重构单元(Spatial Reconstruction Unit, SRU)和通道重构单元(Channel Reconstruction Unit, CRU),整个模块通过处理 CNN 中的空间和通道冗余来实现压缩,并提高 CNN 的效率。除此之外,该卷积模块可以减少冗余计算并促进更有效的特征学习,属于本文方法轻量化策略中的一部分。如图 4 所示,SRU 通过分离和重构的方法来减少空间冗余,如图 5 所示,CRU 则通过通道分割、通道转换、通道融合以减少冗余。

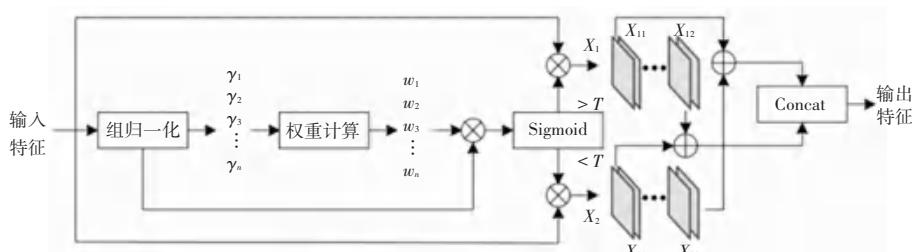


图 4 空间重构单元

Fig. 4 Spatial reconstruction unit

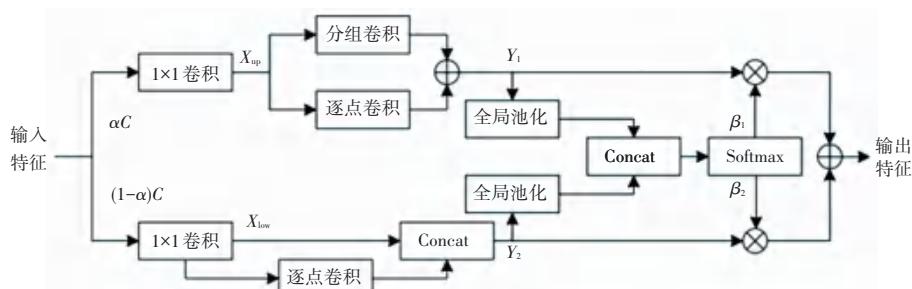


图 5 通道重构单元

Fig. 5 Channel reconstruction unit

SRU 是用于减少 CNN 中的空间冗余, 也就是指神经网络中相邻特征之间的重复或过于相似的信息。首先, SRU 将输入特征图分离成较小的区块, 并计算每个小块的权重, 基于该权重分离开那些具有高度相关性或重复性的特征区域。然后通过学习重要特征的表示并重新对原始特征图进行合并来压缩信息。这个过程涉及提取关键信息并抑制冗余信息, 该方法能有效减少模型复杂度和计算成本, 同时保持或提高网络性能。

CRU 用于减少 CNN 中的通道冗余, 也就是指在不同通道中出现的重复或相似的特征信息。其大致流程分为 3 个步骤: 通道分割、通道转换、通道融合。CRU 首先将输入特征图的通道分割成多个子集, 以专注于独立通道的特征。在分割过程之后, CRU 对每个通道子集应用特定的转换, 并提取每个通道中的有效信息。最后, CRU 将转换后的通道进行重新融合以得到输出特征。该流程能有效提取每个通道的关键信息, 使得 CRU 生成一个综合的、信息丰富的输出, 并在融合过程中减少原始通道的冗余信息。

LRCA 模块中还包括了一种简单无参数的卷积神经网络注意力机制 SAM, 与传统侧重于通道和空间的注意力机制不同, SAM 无需额外的子网络或额外参数即可计算注意力权重。为了估计每个神经元的重要性, 通过使用式(2)中的能量函数, 评估目标神经元和其他神经元之间的线性可分性。

$$e_t(w_t, b_t, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_i - \hat{x}_i)^2 \quad (2)$$

其中, $\hat{t} = w_t t + b_t$, $\hat{x}_i = w_i x_i + b_i$ 是输入特征 $X \in \mathbb{R}^{C \times H \times W}$ 的单个通道中的目标神经元和其他神经元。 i 是空间维度上的指数, $M = H \times W$ 是该通道上的神经元数量。 w_i 和 b_i 都是权重和偏差的变换。通过迭代求解器来求解该方程, 可得到一个快速的封闭形式的解公式如下:

$$w_t = -\frac{2(t - \mu_t)}{(t - \mu_t)^2 + 2\sigma_t^2 + 2\lambda} \quad (3)$$

$$b_t = -\frac{1}{2}(t + \mu_t)w_t \quad (4)$$

那么最小能量值可以通过下式得到,

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (5)$$

其中, $e_t^* = \frac{1}{M} \sum_{i=1}^M x_i$ 且 $\hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{\mu})^2$, M

为每个通道能量函数的个数。该式表明, 能量值越低, 目标神经元与周围神经元越不同, 也表示其越重要。

这种设计选择简化了注意力机制的实现, 并减少了与额外参数相关的计算量。该模块解决了卷积过程中由于全局上下文信息丢失而导致的特征丢失问题, 在图像模糊的情况下, 可以有效提高模型的检测准确性。

最后, 为了进一步提升模型的精度和运行速度, 本文方法将原有的多特征图链接方式进行了替换, 由原始的沿通道拼接变成基于注意力思路的对应像素点相乘的操作。以解码器第 4 层为例, 由 4 个特征图进行融合, 具体的计算步骤如下式所示,

$$x^{0,4} = \prod_{i=0}^3 x^{0,i} \quad (6)$$

其中, $x^{0,i}$ 是第 i 层卷积后输出的特征图。

1.4 全局平均池化融合模块

在通过解码器处理后, 会获得 4 个不同层级的特征图, 以用于进行初步的结果预测。为了更有效地利用这些特征图中蕴含的丰富上下文信息, 并实现从深层到浅层的特征融合, 本文方法设计了一个全局平均池化融合模块 GF 模块, 其结构如图 6 所示。该模块的核心思想是利用深层特征的定位信息来指导浅层特征的融合过程, 以增强模型对目标区域的识别能力, 同时抑制冗余信息的干扰。首先, 对深层特征 f_1 和浅层特征 f_2 执行注意力计算。这一过程可以突出与目标区域相关的特征, 同时将那些与背景相关的特征分离开来。具体计算步骤如下式:

$$f_2 = \text{Conv}(f_2 \cdot f_1 + f_2) \quad (7)$$

随后, 引入全局平均池化 (Global Average Pooling, GAP), 将每个通道的特征图缩减为单个数值, 从而提取全局信息, 以进一步筛选有效区域。该步骤类似人类识别物体的过程, 先定位目标物体的大致区域, 随后关注其边缘和细节, 最终定位物体的准确位置。计算步骤如下式:

$$f_2 = f_2 \cdot \text{Sigmoid}(\text{GAP}(f_2) \cdot \text{Mean}(f_2)) \quad (8)$$

其中, Sigmoid 代表激活函数, Mean 用于计算所有通道的平均值, 以得到最终的特征向量。

通过上述步骤, 有效地结合了深层和浅层特征, 同时保留了对目标区域的精确定位和细节描绘能力。这一特征融合策略预计将显著提升模型在复杂医学图像分割任务中的表现。

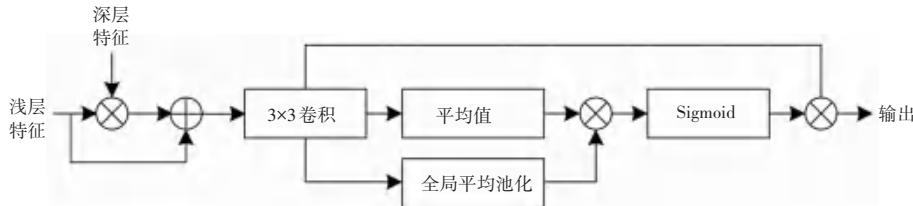


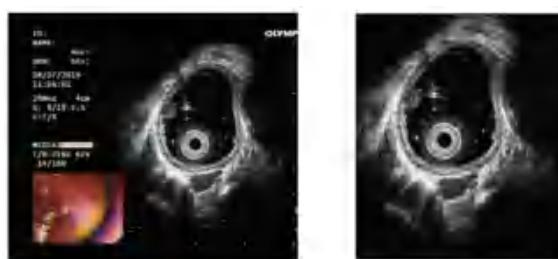
图 6 全局平均池化融合模块

Fig. 6 Global average pooling fusion block

2 实验及结果分析

2.1 数据集

本数据来源于上海交通大学附属第六人民医院消化内镜中心,从 2019 年 9 月到 2020 年 10 月总共收集了 144 名患者的 2 042 张图片,数据包括不同年龄段、不同性别,具有良好的普适性。共包括 5 种肿瘤病例,其中平滑肌瘤 57 例、胃肠道间质瘤 39 例、神经内分泌肿瘤 10 例、异位胰腺 10 例和脂肪瘤 15 例,还有 13 例特殊病例不具有代表性,病例太少或者其他成病原因,所以在进行分析时将此 13 个病例剔除。在收集数据时,均征得所有患者知情同意,在术前对患者均进行多模态超声检查,包括普通白光内镜和超声内镜。所有 144 例患者的最终诊断均基于手术中活检的组织病理学结果。原始数据集像素大小为 764×572,其中包含日期、机器参数以及对应白光图像等无关信息。对原始图像进行裁剪处理,得到最终图片尺寸为 388×457,如图 7 所示。其中训练集和测试集的比例为 7 : 3。



(a) 预处理前 (b) 预处理后

图 7 图片预处理前(a)后(b)

Fig. 7 Comparison chart before (a) and after (b) image preprocessing

2.2 参数设置

实验的计算机操作系统为 Ubuntu20.04, 使用 Pytorch 架构, 算法实现编程语言为 Python, 使用 GPU 加速网络模型的训练和测试。在训练过程中选用 Adam 优化器进行训练, 这是因为其能够动态地调整学习率, 并带来更快的收敛效果。由于计算资源的限制, 本文方法将训练的批量大小定为 2, 学

习率初始设置在 0.01, 每经过 50 个 Epoch 学习率变为 0.1 倍, 总共进行 500 个 Epoch 的训练。无论是训练集还是测试集, 输入的图片尺寸都是 352×352×3。

2.3 评价指标

本文使用平均 Dice 系数 (mean Dice, mDic)、平均交并比 (mean Intersection over Union, mIoU) 这两个指标来进行定量评估, 其二者都是衡量模型分割准确度的指标, 其值越接近 1 表示分割效果越好。计算步骤如下式:

$$mDic = \frac{2 \times |X \cap Y|}{|X| + |Y|} \quad (9)$$

$$mIoU = \frac{|A \cap B|}{|A \cup B|} \quad (10)$$

在目标检测领域中广泛使用的其他 5 个指标中包括边界 F 测度 (F_β)、结构相似性度量 (S_ω)、平均误差 (mE)、最大误差 (max E) 以及平均绝对误差 (MAE)^[17] 多个不同的指标衡量分割结果的准确性, 更为公正地对分割结果做出评价。

边界 F 测度 (F_β) 用来关注分割结果中边界的准确度, 其是基于预测边界和真实边界之间的精确度和召回率的加权调和平均, 其中的 Precision 是边界预测的准确率, Recall 是边界预测的召回率, β 是调节精确率和召回率相对重要性的参数。在图像分割中, 边缘的准确识别非常关键, 因为其往往是区分不同结构的主要因素。结构相似性度量 (S_ω) 考虑了亮度、对比度和结构 3 个方面的变化, 是对分割结果质量的综合评价。其中, μ_x, μ_y 是预测图像和真实图像的平均亮度, σ_x, σ_y 是其方差, σ_{xy} 是协方差, c_1, c_2 是为了维持稳定性而加入的小常数。计算步骤如下式:

$$F_\beta^\omega = (1 + \beta^2) \frac{\text{Precision}^\omega \cdot \text{Recall}^\omega}{\beta^2 \cdot \text{Precision}^\omega + \text{Recall}^\omega} \quad (11)$$

$$S_\omega = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (12)$$

平均误差(mE)反映了误差的平均水平,而最大误差(maxE)则指出了所有误差中的最大值,其对于理解模型可能出现的最坏情况很有帮助。平均绝对误差(MAE)是一个衡量预测像素值和真实像素值之间差异的指标,是所有单个像素误差绝对值的平均。

这些指标合在一起为我们提供了一个全面的视角来评价医学图像分割模型的性能,不仅仅是在分割精度上,还包括了边界的准确识别、结构的保存以及

误差的大小。在图中,这些指标用于比较不同模型的性能,以确定哪一个模型在特定任务上表现最好。

2.4 实验结果与分析

1) 对比实验。

为了验证模型的性能,本文对比了 6 种医学图像分割邻域拥有最佳性能的分割模型,包括 UNet、UNet++、UNet3+ 和 PraNet、SANet、PolyNet。具体的实验结果见表 1。

表 1 对比实验结果

Table 1 Comparative experimental results

模型	mDic	mIoU	F_{β}^{ω}	S_{ω}	mE	maxE	MAE	Param/M
UNet	0.662	0.561	0.655	0.804	0.842	0.871	0.010	7.79
UNet++	0.719	0.627	0.712	0.835	0.880	0.906	0.008	36.63
UNet3+	0.657	0.562	0.650	0.801	0.839	0.885	0.009	26.97
PraNet	0.673	0.560	0.633	0.808	0.884	0.902	0.011	32.39
SANet	0.668	0.566	0.660	0.808	0.849	0.884	0.009	25.79
PolyNet	0.695	0.600	0.689	0.822	0.868	0.907	0.008	25.91
本文网络	0.778	0.679	0.767	0.862	0.943	0.949	0.007	19.79

如表 1 所示,在各项指标上,本文方法在数据集上表现出了最佳性能。特别是在平均 Dice 系数(maxE)上达到了 0.778,相较于改进前的基线模型 UNet++ 的 0.719,提高了 8.2%。此外,在平均交并比(mIoU)上,本文方法也有所提升,得分为 0.679,提高了 8.3%。

综合考量各项其他指标,本文方法在边界 F 测度(F_{β}^{ω})上也领先,达到了 0.767,提高了 7.7%。在结构相似性度量(S_{ω})上得分为 0.862,提高了 3.2%。此外,在平均误差(mE)和最大误差(maxE)上也分别实现了 0.943 和 0.949 的高分,均高于 UNet++ 的 0.880 和 0.906。在平均绝对误差(MAE)这一指标上,本文模型与 UNet++ 相比,也有细微的提升,从 0.008 降低到了 0.007。

其中值得注意的是,本文方法的参数量达到了 UNet++ 参数量的一半,约为 54%,这证明本文方法

在轻量化策略上有明显效果。

综上所述,本文方法在多个关键性能指标上都实现了显著提升,不仅有着更少的参数量,且在各项指标上有着最好的分割精度,证明了所提出方法的有效性和优越性。

2) 消融实验

为了证明本文添加的 RFB 模块、LRCA 模块和 GF 模块对本文胃粘膜下肿瘤分割网络的有效性和参数量情况,本文对上述模块进行消融实验得分验证。所有的实验训练配置均相同。依次添加模块,并对平均 Dice 系数(mDic)、平均交并比(mIoU)、边界 F 测度(F_{β}^{ω})、结构相似性度量(S_{ω})、平均误差(mE)和最大误差(maxE)以及平均绝对误差(MAE)等多个不同的指标进行测试结果对比,还包括模型参数量,以此来衡量分割结果的准确性。实验结果见表 2。

表 2 消融实验

Table 2 Ablation studies

模型	mDic	mIoU	F_{β}^{ω}	S_{ω}	mE	max E	MAE	Param/M
UNet++	0.719	0.627	0.712	0.835	0.880	0.906	0.008	36.63
RFB	0.726	0.631	0.725	0.836	0.884	0.901	0.008	20.02
RFB+LRCA	0.754	0.649	0.740	0.851	0.913	0.923	0.007	19.95
RFB+LRCA+GF(本文)	0.778	0.679	0.767	0.862	0.943	0.949	0.007	19.79

从表 2 中可知,添加 RFB 模块后各项指标均有增长,且参数量大幅下降了 44%,说明该模块的添

加对于模型轻量化有着不可或缺的重要作用。其中平均 Dice 系数、平均交并比、边界 F 测度和结构相

似性度量都增长了 1% 左右。在此基础之上,添加 LRCA 模块后,平均 Dice 系数增长了 4%,平均交并比增长了 3%,边界 F 测度和结构相似性度量都增长了 2% 左右,平均误差提高了 3.3%,同时最大误差提高了 2.4%。由此可知该模块对于肿瘤分割的细节方面起到重要作用,使模型准确性有效提高。

表 2 中第 4 行是本文方法的测试结果,在 RFB 模块和 LRCA 模块的基础上添加了 GF 模块,进行深层浅层特征的有效融合,进一步提升对肿瘤的精确定位和细节描绘能力。其中平均 Dice 系数增长了 3.2%,平均交并比增长了 4.6%,边界 F 测度增长了 3.6%,对肿瘤细节的识别有明显提升。平均误差提高了 3.3%,最大误差提高了 2.8%。同时模型的参数量也在不断减小。

(3) 分割结果对比

为了直观地显示本文提出模型在本数据集上表现的优越效果,对分割结果进行可视化展示。如图 8 所示,分别给出了 9 种不同的医学分割模型在本文数据集上的分割效果,本文所提出的方法在分割结果中明显更接近原图的真实分割图像,尤其是定位上的准确性和边界分割上的细致性。通过对比可得出其他分割模型只能大概定位出肿瘤的位置和大小且有残缺的情况发生,这是由于胃黏膜的层数较多,较难清晰地分割出边界。对比 Unet++ 的分割结果可以看出,本文模型在其基础之上的改进使得分割边缘更加精准、平滑且分界线更流畅,证明本文方法处理得到的分割图更完整准确。

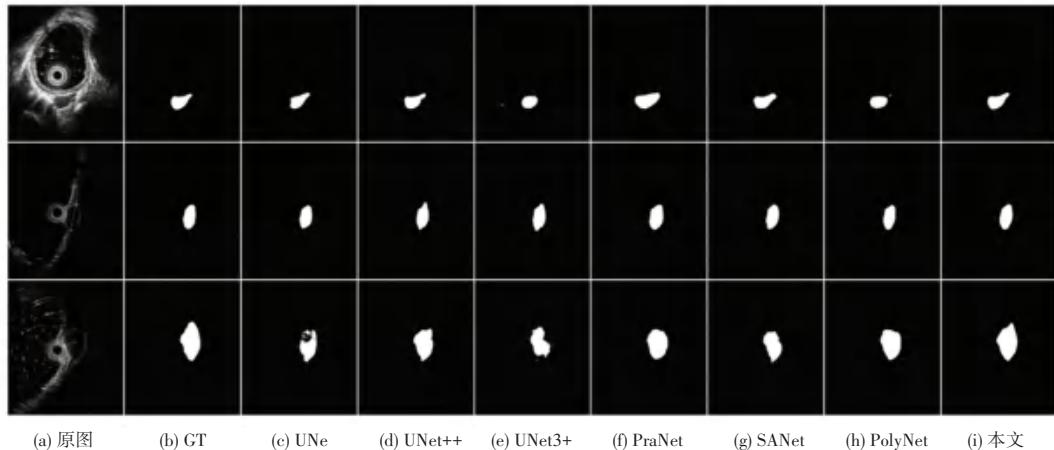


图 8 不同模型在本文数据集上的分割结果对比

Fig. 8 Comparison of segmentation results of different models on our dataset

3 结束语

本文针对胃黏膜下肿瘤层次较多、类型复杂且边界分割困难等问题提出一种具有深层特征约束的轻量化胃黏膜下肿瘤分割网络。该网络通过结合扩增感受野模块和轻量化无参数的重构卷积模块,有效地提高了胃黏膜下肿瘤分割的精度和细节识别能力。此外,还设计了一个全局平均池化融合模块,在深层特征约束的基础上进行了浅层特征融合,进一步优化了肿瘤边缘和定位信息的精确度。实验结果显示,本文方法在多个关键性能指标上都取得了显著提升,在本数据集上得到的平均 Dice 系数和平均交并比得分分别为 0.778 和 0.679,相比 UNet++ 均提高超过 8%;边界 F 测度、结构相似性度量、平均误差、最大误差以及平均绝对误差等多个不同的指标都有显著提升,提高了 3%~8%,整体网络模型参

数量减少到原来的 54%。实验结果证明该方法不仅能提高胃黏膜下肿瘤分割的精准度和细致性,还同时实现了模型参数的大幅减少。这一成果不仅在技术层面上实现了突破,更在临床应用上展现了巨大的潜力。

未来,我们计划进一步改进和优化该方法,以便更好地适应不同的临床需求和医学图像分析场景。此外,探索将该技术应用于其他类型的医学图像分割任务,如其他器官的肿瘤分割等,也将是未来研究的重要方向,旨在为医学诊断提供更广泛、更精准的辅助工具。

参考文献

- [1] GUO Jintao, LIU Zhiyun, SUN Siyu, et al. Endosonography-assisted diagnosis and therapy of gastrointestinal submucosal tumors[J]. Endoscopic Ultrasound, 2013, 2(3): 125.

- [2] EADIE L H, TAYLOR P, GIBSON A P. A systematic review of computer-assisted diagnosis in diagnostic cancer imaging [J]. European Journal of Radiology, 2012, 81(1): e70–e76.
- [3] AL-SHARE B, ALLOGHBI A, AI HALLAK M N, et al. Gastrointestinal stromal tumor: A review of current and emerging therapies [J]. Cancer and Metastasis Reviews, 2021, 40 (2): 625–641.
- [4] KE Chongwei, CAI Jingli, CHEN Danlei, et al. Extraluminal laparoscopic wedge resection of gastric submucosal tumors: A retrospective review of 84 cases [J]. Surgical Endoscopy, 2010, 24 (8): 1962–1968.
- [5] KIDA M, KAWAGUCHI Y, MIYATA E, et al. Endoscopic ultrasonography diagnosis of subepithelial lesions [J]. Digestive Endoscopy, 2017, 29(4): 431–443.
- [6] ZHANG Yu, WANG Zhen, JIN Ting, et al. Hyperechoic demarcation line between a tumor and the muscularis propria layer as a marker for deciding the endoscopic treatment of gastric submucosal tumor [J]. Journal of Turbulence, 2017, 18(8): 707–716.
- [7] LU Yi, ZHUO Xianhua, ZHONG Qinghua, et al. Endoscopic ultrasonography is useful for predicting perforation in the endoscopic resection of gastric submucosal tumors originating from the muscularis propria: A retrospective case-control study [J]. Ultrasonography, 2023, 42(1): 78–88.
- [8] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]// Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 770–778.
- [9] SHI Baoguang, BAI Xiang, YAO Cong. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(11): 2298–2304.
- [10] GAO Y Q, LI K, MOSALAM K M, et al. Deep transfer learning for image-based structural damage recognition [J]. Computer-Aided Civil and Infrastructure Engineering, 2018, 33(9): 748–768.
- [11] SAJDA P. Machine learning for detection and diagnosis of disease [J]. Annual Review of Biomedical Engineering, 2006, 8(1): 537.
- [12] SCHLEMPER J, OKTAY O, SCHAAP M, et al. Attention gated networks: Learning to leverage salient regions in medical images [J]. Medical Image Analysis, 2019, 53: 197–207.
- [13] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]// Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. 2015: 234–241.
- [14] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39 (4): 640 – 651.
- [15] ZHOU Zongwei, SIDDIQUEE M M R, TAJBAKHSH N, et al. U-Net++: A Nested U-Net Architecture for Medical Image Segmentation [M]// Stoyanov D. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 2018: 3–11 .
- [16] HUANG Huimin, LIN Lanfen, TONG Ruofeng, et al. Unet 3+: A full - scale connected U - Net for medical image segmentation [C]// Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing. Piscataway, NJ: IEEE, 2020: 1055–1059.
- [17] FAN Dengping, JI Gepeng, ZHOU Tao, et al. Planet: Parallel reverse attention network for polyp segmentation [C]// Proceedings of International Conference on Medical Image Computing and Computer Assisted Intervention. Cham: Springer, 2020: 263–273.
- [18] WEI Jun, HU Yiwen, ZHANG Ruimao, et al. Shallow attention network for Polyp segmentation [C]// Proceedings of International Conference on Medical Image Computing and Computer Assisted Intervention. Cham: Springer, 2021: 699–708.
- [19] DONG Bo, WANG Wenhui, FAN Dengping, et al. Polyp-PVT: Polyp segmentation with pyramid vision transformers [J]. arXiv, 2108.06932, 2021. DOI:10.48550/arXiv. 2108.06932.
- [20] LI Jiafeng, WEN Ying, HE Lianghua. SCConv: Spatial and channel reconstruction convolution for feature redundancy [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2023: 6153–6162.
- [21] LI Chen, TAN Yusong, CHEN Wei, et al. Attention Unet++: A nested attention - aware U - Net for liver CT image segmentation [C]// Proceedings of IEEE International Conference on Image Processing. Piscataway, NJ: IEEE, 2020: 345–349.
- [22] YANG Lingxiao, ZHANG Ruyuan, LI Lida, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks [C]// Proceedings of International Conference on Machine Learning. PMLR, 2021:11863–11874.
- [23] TREBING K, STANCZYK T, MEHRKANOON S. SmaAT-UNet: Precipitation nowcasting using a small attention-UNet architecture [J]. Pattern Recognition Letters, 2021, 145: 178–186.
- [24] QAMAR S, JIN Hai, ZHENG Ran, et al. A variant form of 3D-UNet for infant brain segmentation [J]. Future Generation Computer Systems, 2020, 108: 613–623.
- [25] GADOSEY P K, LI Yujian, AGYEKUM E A, et al. SD UNet: Stripping down U-Net for segmentation of biomedical images on platforms with low computational budgets [J]. Diagnostics, 2020, 10(2): 110.
- [26] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs [J]. Computer Science, 2014, 4: 357–361.