

基于边缘感知聚合网络的息肉分割模型*

周涛^[1] 张轶哲^[1] 陈耿^[2] 周毅^[3] 吴烨^[1] 范登平^[4]

¹⁾ 南京理工大学计算机科学与工程学院 ²⁾ 西北工业大学计算机学院 ³⁾ 东南大学计算机科学与工程学院 ⁴⁾ 南开大学计算机学院

摘要 精准的息肉分割对于临床实践中的结直肠癌 (Colorectal Cancer, CRC) 的早期诊断和预防有着重要的意义。然而, 由于尺度的变化和模糊的息肉边界, 实现令人满意的分割性能仍然是一项具有挑战性的任务。在这项研究中, 本文提出了一种新颖的边缘感知特征聚集网络 (Edge-aware Feature Aggregation Network, EFA-Net), 该网络可以充分利用跨层级和多尺度特征来增强息肉分割的性能。具体而言, 本文首先提出了一种边缘感知引导模块 (Edge-aware Guidance Module, EGM), 该模块将低层特征与高层特征相结合, 学习出一种边缘增强的特征, 并使用逐层级策略将其融入到每个解码器单元中。此外, 本文提出了一种尺度感知卷积模块 (Scale-aware Convolution Module, SCM), 通过使用具有不同比例的扩张卷积来学习刻度感知特征, 以有效处理尺度变化。此外, 本文提出了一种跨层级融合模块 (Cross-level Fusion Module, CFM), 以有效地集成跨层级特征, 充分利用局部和全局的背景信息。最后, 通过学习得到的边缘感知特征自适应地对 CFM 的输出进行加权, 然后用于生成多个分割结果图。在五个广泛采用的结肠镜数据集上的实验证明, 本文的 EFA-Net 在泛化和有效性方面优于其他最先进的息肉分割方法。本文的实现代码和分割结果图公布在 <https://github.com/taozh2017/EFANet>。

关键词 结直肠癌, 息肉分割, 边缘感知引导模块, 尺度感知卷积模块, 跨层级融合模块

1 简介

结直肠癌 (Colorectal cancer, CRC) 是全球第三大最常见的恶性肿瘤^[2], 通常起源于腺瘤性息肉。如果不及及时得到有效地治疗, 这种亚型的息肉很容易转化为恶性肿瘤。结肠镜是一种金标准的检查技术, 可以提供关于息肉的位置和外观的信息。然而, 结肠镜检查高度依赖操作者, 经常会漏检可能的息肉。此外, 人工定位和分割息肉存在主观性强、耗时长的的问题。因此, 开发自动息肉分割方法可以有效地为临床医生提供准确的息肉位置。

目前, 基于编码器-解码器结构的 U-Net^[3] 已经成为主流框架, 广泛应用于医学图像分割任务。在 U-Net 结构的启发下, 两个扩展模型 (即 U-Net++^[4] 和 ResUNet++^[5]) 已被用于息肉分割任务中, 并取得了较好的结果。增强型的 U-Net^[6] 利用多尺度语义上下文来增强病变分割的特征有效性。为了捕捉更多的上下文信息, ACSNet^[7] 利用自适应上下文选取方案来探索全局和局部上下文, 有助于提升病变分割性能。CCBANet^[8] 利用级联上下文模块来捕捉区域和全局上下文, 其中引入了平衡注意力块来分别对病变、边界和背景进行注意力机制。由于息肉与背景之间存在不清晰的边界, 因此精准定位息肉与背景之间的边界是

*本文为 Machine Intelligence Research 论文^[1]的中译版。

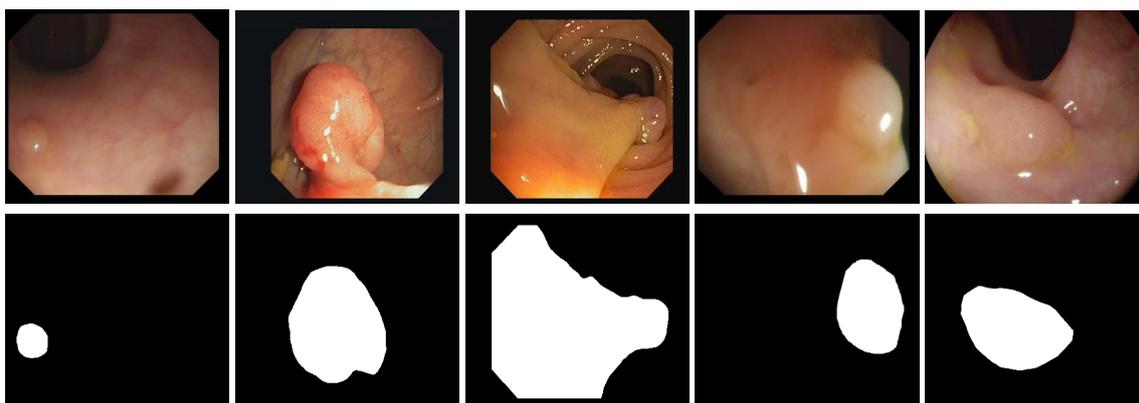


Fig. 1 息肉分割中不同挑战的说明，即，1) 相同类型的息肉具有不同的尺度（见第一个至第三个列），2) 息肉及其周围黏膜之间的边界模糊（见第四和第五个列）。

分割息肉的关键。在这个例子中，有一些方法^[9-11]通过考虑捕捉边界线索或提取边界感知特征来提高分割性能。尽管取得了一定的成果，但由于形状、大小和位置的差异，仍有很大的空间来提高息肉分割的性能。首先，主要挑战之一是息肉分割中的尺度变化。如图 1 所示，前三列显示了不同尺度和形状的息肉。在这种情况下，如何充分整合跨层级特征并捕捉多尺度信息仍然具有挑战性。其次，如图 1 的第四和第五列所示，息肉的外观与背景相似，而息肉与其周围黏膜之间的边界并不清晰。因此，如何有效地利用边界信息或者将边缘增强特征融入到解码器中仍然需要进一步探索。

为此，提出了一种新颖的边缘感知特征聚合网络（EFA-Net）用于结肠镜图像中的息肉分割，该网络能够从跨层级、多尺度的特征中学习上下文知识，并准确地利用边缘语义信息来提升分割性能。具体来说，提出了一种适应边缘引导模块（EGM），用于学习边缘感知特征并对解码器中的特征加权。然后，提出了一种尺度感知卷积模块（SCM），用于从单层特征中提取多尺度信息。此外，提出了一种跨层融合模块（CFM）来整合相邻的跨层特征，这种跨层融合可以增强不同分辨率特征的代表能力。然后，通过学习得到的边缘感知特征自适应加权融合特征，并使用加权特征生成分割结果图。本文的主要贡献有三个方面：

- 本文提出了一个尺度感知的卷积模块来利用多尺度信息，以及一个跨层级的融合模块来整合相邻特征，该模块可以学习尺度相关特征并增强其表达能力，从而更好地应对尺度变化。
- 提出了一种边缘感知引导模块，可以有效地捕捉息肉和背景之间的边缘语义，从而得到更好的分割性能。
- 在五个常用的结肠镜数据集上进行了比较实验，结果表明本文的 EFA-Net 在与十一个最先进的分割方法进行比较时取得了更好的分割结果。

本文的其余部分如下。本文在 2 讨论了与我们的研究相关的两种类型的工作。然后，本文在 3 中提供了所提出的息肉分割框架的详细信息。在 4 进行了对比实验，以验证本文的分割方法的有效性。最后，在 5 中得出了本文研究的结论。

2 相关工作

在这一部分中, 本文简要回顾了某些与息肉分割、基于边界的方法以及多尺度/层级特征表示相关的工作。

2.1 息肉分割

息肉分割的目标是准确地定位和分割给定结肠镜图像中的息肉区域。传统的息肉分割方法主要提取手工设计的特征^[12-16], 例如形状、纹理、轮廓等。然而, 由于手工设计特征的表达能力有限, 这些传统方法往往存在较高的误检率。随着深度学习技术的进步, 各种基于卷积神经网络 (CNNs) 的方法在息肉分割领域得到了发展。早期的全卷积神经网络 (FCN) 已经被用于结肠镜图像的息肉分割^[17]。基于 FCN 的方法由于特征分辨率较低, 分割结果往往比较粗糙, 容易出现边界误差。随后, Lonneberger^[3] 等人提出了一个 U 型网络 (U-Net) 框架, 该框架已成为生物医学图像分割任务的流行 CNN 架构。在 U-Net 结构的基础上, 研究人员开发了几个扩展模型^[2,18-19], 用于息肉分割。然而, 现有的 U-Net 结构, 如 ResUNet^[20], H-DenseUNet^[21] 和 Attention U-Net^[22], 直接采用跳跃连接将编码器和解码器的特征表示融合在一起。然而, 简单的跳跃连接可能会降低分割性能, 因为编码器和解码器之间存在特征差异。为了解决这个问题, Zhou^[19] 等人通过设计一种密集跳跃连接策略, 提出了 UNet++ 来改善分割性能。此外, 一些方法^[9-11,23-24] 提出学习边界感知特征来提升分割性能。例如, Psi-Net^[9] 同时利用边界和区域信息, 但不能有效地探索边界和区域之间的关系。SFA^[10] 利用区域和边界约束为息肉分割设计了一个选择性特征聚合模块。PraNet^[11] 通过有效地聚合高级特征来利用上下文信息, 并采用多个逆向注意力模块来捕捉区域和边界之间的相关性, 以校准一些不准确的预测。Zhou^[23] 等人设计了一个边界预测网络来学习边界感知特征, 并将其融合到分割网络中以提高分割性能。此外, 一些息肉分割方法^[25-27] 充分利用多尺度特征并进行特征融合来增强模型性能。另外, 也有其它方法^[28-32] 利用注意机制来实现息肉分割。

2.2 基于边界的方法

基于边界的方法利用网络内的边缘信息来增强特征表示, 最终提升模型性能^[33-39]。例如, Ding 等人^[33] 创新地引入了一个边界感知的特征传播模块。该模块旨在加强相同分割区域内的局部特征的相似性, 同时保持与不同分割区域相关特征的独特性。在此基础上, Zhou 等人^[36] 提出了一个边界引导模块, 用于学习边界增强的特征表示, 这个模块能够捕捉局部特征和边界细节, 提高伪装目标的检测性能。除了这些应用之外, 还有许多研究调查利用边界/边缘信息来提升医学任务的性能^[40-45]。例如, Zhang 等人^[41] 提出了一个通用的医学分割框架, 利用边缘注意力特征来引导分割过程。类似地, Wang 等人^[45] 引入了一个边界保持模块, 将学习到的边界融入到中间层, 提高了分割的准确性。此外, SFA^[10] 引入了一种创新的边界敏感损失, 其目的在于相互改进息肉区域分割和边界检测。BA-Net^[44] 采用多任务学习方法, 共同分割物体掩膜并识别损伤边界。在这个模型中, 交互式注意力机制被设计用来利用不同任务的互补信息, 从而提升分割的性能。

2.3 多尺度/层级特征表示

多尺度特征表示在许多计算机视觉任务中得到了广泛的研究, 例如目标检测^[46-48]、图像超分辨率^[49-50] 和语义分割^[51-52]。此外, 两种代表性策略 (即金字塔池化^[53] 和空洞空间金字塔池化 (ASPP)^[54]) 已经被广泛应用于多尺度特征提取。目前, 多尺度或跨层级的特征融合^[24,55-58] 在医学图像分割领域受到越来越多的关注, 有助于开发区域上下文信息并处理尺度变化的问题。例如, Yang 等人^[55] 提出了一种跨层级融合策略, 以集成来自不同层级的多尺度特征用于慢性中风损伤分割。Srivastava 等人^[57] 提出了一种专门用于

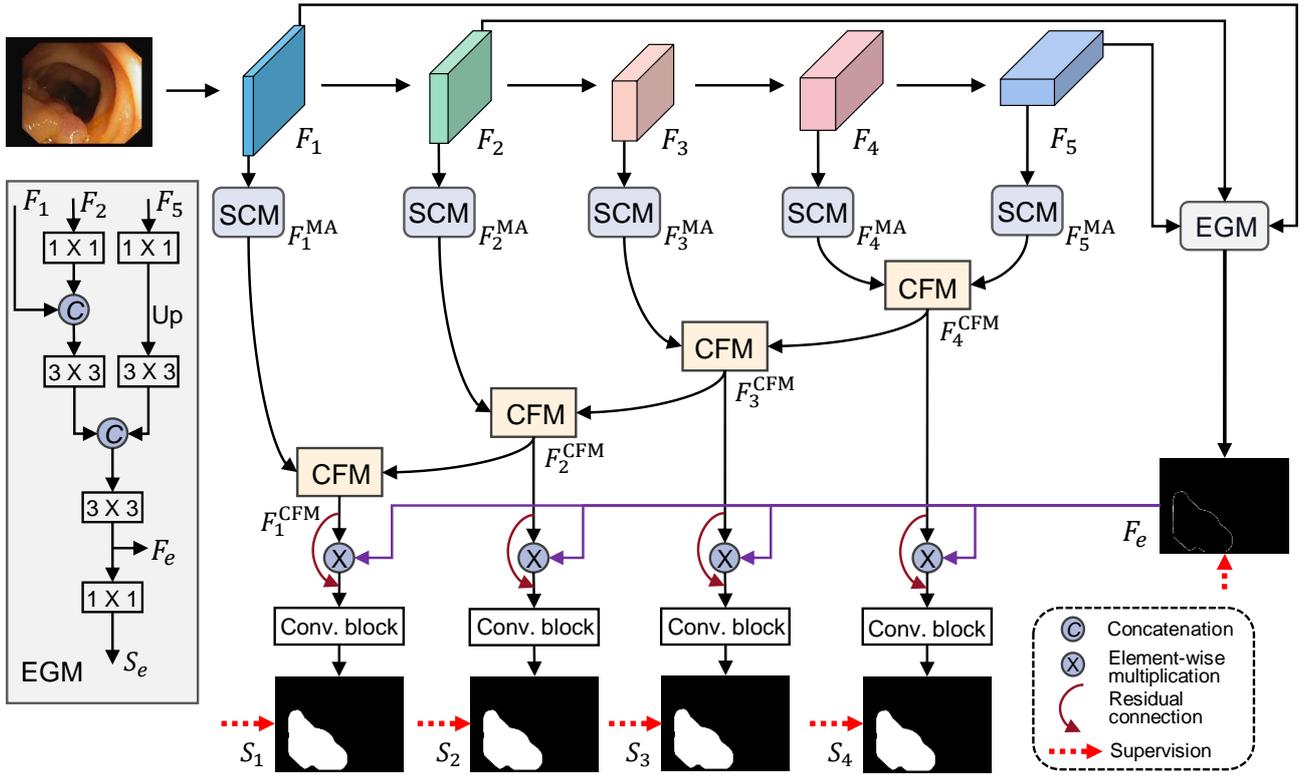


Fig. 2 所提出的 EFA-Net 的架构，包括边缘感知引导模块、尺度感知卷积模块和跨层融合模块。

医学图像分割的多尺度残差融合网络，其中通过可变感受野的双尺度密集融合策略来利用多尺度信息。Yue 等人^[24]充分整合了跨层级特征以捕捉更多的上下文信息，从而改善息肉分割性能。尽管这些方法取得了很大进展，但融合多尺度和跨层级特征表示以捕捉尺度变化并集成区域上下文信息，仍需进一步探索以提高分割性能。

3 方法

在本节中，首先概述了 3.1 节中提出的息肉分割框架。然后介绍了三个关键组件，边缘感知引导模块 (Sec. 3.2)、尺度感知卷积模块 (Sec. 3.3) 和跨层级融合模块 (Sec. 3.4)。此外，本文在 3.5 部分中介绍了学习的细节，包括模型推理和损失函数。

3.1 架构概述

如 Fig. 2 所示，本文介绍了 EFA-Net 的体系架构。具体而言，采用了 Res2Net^[59] 作为骨干网络，用来提取给定图像的多层特征，表示为 $\{F_i, i = 1, 2, \dots, 5\}$ 。然后，提出的 EGM 将低级特征 (*i.e.*, F_1 和 F_2) 和高级特征 (*i.e.*, F_5) 结合起来学习边缘增强特征，并进一步融入解码器中。此外，将每个特征 F_i 输入到所提出的 SCM 中以提取多尺度特征。进一步来说，利用 CFM 融合跨层级特征，以充分利用局部和全局上下文信息。最后，本方法的解码器产生多个侧面分割图 $\{S_i, i = 1, 2, 3, 4\}$ 。

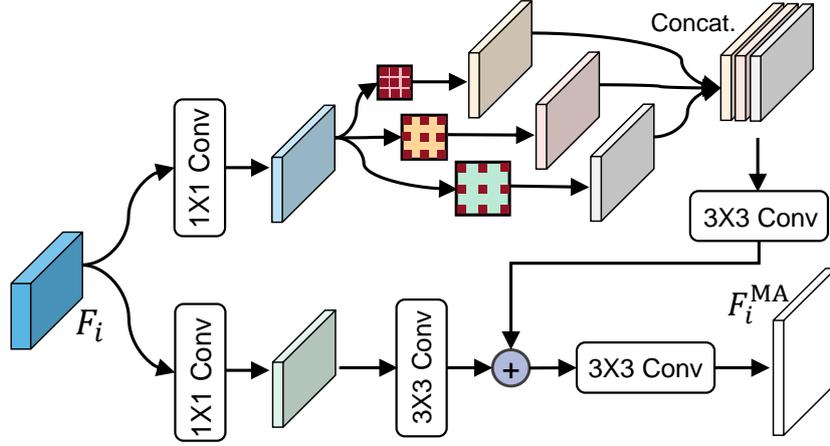


Fig. 3 所提出的尺度适应卷积模块的架构。

3.2 边缘感知引导模块

一些现有的研究指出，边缘信息能够改善分割和检测的性能^[41,60-62]。由于低层级特征保留了足够的边界细节，现有方法^[41,60,62]通常将低层特征整合起来学习增强边缘的表示。尽管低层级特征为学习边缘感知特征提供了更多信息，但它们很可能引入非边缘的粗细节。因此，受^[61]的启发，本文将两个低层级特征 (*i.e.*, F_1 和 F_2) 与高层级特征 (*i.e.*, F_5) 结合起来构建本文的 EGM。具体而言，如图 2 所示，首先将 F_2 和 F_5 分别输入两个 1×1 卷积层，以得到 F'_2 和 F'_5 。然后，将 F_1 和 F'_2 进行级联操作，并通过顺序操作得到 $F_{12} = \mathcal{C}_{3 \times 3}(\text{Concat}(F_1, F'_2))$ ，其中 $\mathcal{C}_{3 \times 3}(\cdot)$ 是一个顺序操作，由一个 3×3 卷积层、批归一化和 $ReLU$ 激活函数组成。此外，对 F_{12} 和 F'_5 进行连接操作，得到增强边缘的特征，可以描述为

$$F_e = \mathcal{C}_{3 \times 3}(\text{Concat}(F_{12}, \text{Up}(F'_5))), \quad (1)$$

其中 $\text{Up}(\cdot)$ 表示一个上采样操作。最后， F_e 经过一个 1×1 卷积层生成边缘图 (S_e)，并且边缘图被上采样以确保与原始图像具有相同的分辨率。在这里，我们采用二元交叉熵 (BCE) 损失 (称为 \mathcal{L}_{BCE}) 来衡量生成的边缘图和真实的边缘图 (G_e ，由 Sobel 检测算子获得) 之间的差异，相关损失可以被定义为 $\mathcal{L}_e = \mathcal{L}_{BCE}(S_e, G_e)$ 。值得注意的是， F_e 提供了边缘感知特征，以加权解码器中的特征来提升分割性能。

3.3 尺度感知卷积模块

由于同一类型的息肉具有不同的尺度，学习尺度相关的特征表示对模型分割息肉具有重要意义。为此，本文提出一种尺度感知卷积模块 (SCM) 来提取尺度感知特征，利用不同的尺度相关信息来增强特征表示。具体来说， F_i 首先通过两个 1×1 卷积来获得 F_i^1 和 F_i^2 。然后，为了学习尺度感知的特征表示，本文分别使用不同的扩张率将 F_i^1 输入到三个 3×3 的卷积中，从而可以得到

$$E_l = \mathcal{C}_{3 \times 3}^{r_i}(F_i^1), \quad (2)$$

其中 $\mathcal{C}_{3 \times 3}^{r_i}$ 表示扩张率为 r_i 的 3×3 扩张卷积，然后是 $ReLU$ 激活的批量归一化操作。本研究采用三种不同的扩张率 *i.e.*, $r_i = 2, 4, 8$ 。然后，将三个尺度感知特征 (*i.e.*, $E_l, l = 1, 2, 3$) 连接起来，通过 3×3 卷积层进行自适应聚合，得到连接后的特征。最后，将拼接后的特征与原始特征 F_i^2 进行融合，从而将该过程表

示为

$$F_i^{\text{SCM}} = \mathcal{C}_{3 \times 3}(\mathcal{C}_{3 \times 3}(\text{CAT}(E_1, E_2, E_3)) \oplus \mathcal{C}_{3 \times 3}(F_i^2)), \quad (3)$$

其中 \oplus 表示元素加法操作。由此得到尺度感知特征 F_i^{SCM} ，将其作为 i -th SCM 的输出。

值得注意的是，我们的 SCM 可以通过利用不同的扩张卷积来增强多尺度特征表示，其中小尺度特征（带有较小的扩张率）和大尺度特征（带有较大的扩张率）可以聚合起来以处理尺度变化的挑战。

3.4 跨层级融合模块

本文提出 CFM 用于融合相邻层次的特征，可以有效地利用跨层次的上下文语义来提高息肉分割性能。具体而言，以 F_a 和 F_b 为例，首先将它们连接起来得到融合特征 F_{cat} 。然后， F_{cat} 通过三个连续的卷积操作，从而得到 F_{cat}^1 、 F_{cat}^2 和 F_{cat}^3 。因为局部上下文可以避免小息肉被忽略，而全局上下文则有助于强调特征学习来定位一些大息肉。因此，在 CFM 中，提出了捕获局部上下文以及利用全局上下文来增强特征表示。为了实现这一点，在图 3 所示的“局部注意力”分支中，本文进行了两个逐点卷积 (PWC)^[63] 来学习本地注意力权重。此外，如图 3 所示，在“全局注意力”分支中，首先进行全局平均池化 (GAP) 操作，然后利用 PWC 操作来学习全局权重。因此，可以得到基于局部和全局注意力的权重：

$$\begin{cases} W_{local} = \mathcal{S}(\mathcal{P}_{conv2}(\mathcal{R}(\mathcal{P}_{conv1}(F_{cat}^1))))), \\ W_{global} = \mathcal{S}(\mathcal{P}_{conv2}(\mathcal{R}(\mathcal{P}_{conv1}(\text{GAP}(F_{cat}^2))))), \end{cases} \quad (4)$$

其中， \mathcal{P} 表示逐点卷积， \mathcal{P}_{conv1} 和 \mathcal{P}_{conv2} 的卷积核大小分别为 $\frac{K}{t} \times K \times 1 \times 1$ 和 $K \times \frac{K}{t} \times 1 \times 1$ 。这里， t 是通道缩减比例， K 是通道大小。此外， $\mathcal{S}(\cdot)$ 和 $\mathcal{R}(\cdot)$ 表示 *Sigmoid* 和 *ReLU* 激活函数。接下来，通过残差连接对特征和相应权重进行逐元素乘法运算，从而得到增强的特征。

$$\begin{cases} F_{en}^1 = F_{cat}^1 \otimes W_{local}, \\ F_{en}^2 = F_{cat}^2 \otimes W_{global}, \end{cases} \quad (5)$$

其中 \otimes 表示逐元素相乘。然后，将两个注意力增强特征和原始特征 F_{cat}^3 连接起来，并将其输入到序列卷积操作中，生成最终输出 F_i^{CFM} ，得到以下结果：

$$F_i^{\text{CFM}} = \mathcal{B}_{conv3 \times 3}(\text{Concat}(F_{en}^1, F_{en}^2, F_{cat}^3)). \quad (6)$$

3.5 学习策略细节

模型推理：我们得到了跨层融合的特征 $F_i^{\text{CFM}}(i = 1, 2, 3, 4)$ ，因此集成边缘感知特征以提供更多的边界细节是很重要的。为了实现这个整合，首先将边缘感知特征 F_e 通过 *Sigmoid* 激活函数进行归一化，将值映射到 $[0, 1]$ 区间，并作为边缘特征级别的注意力图。为了保留 F_i^{CFM} 中的原始信息，还采用了残差连接策略将边缘增强特征与 F_i^{CFM} 相结合，从而获得边缘感知特征。

$$F_i^{\text{edge}} = F_i^{\text{CFM}} \otimes \mathcal{S}(F_e) + F_i^{\text{CFM}}. \quad (7)$$

然后， F_i^{edge} 将通过一个“Conv. block”(如图 2 所示) 生成多个边输出分割图，即 $S_i(i = 1, 2, 3, 4)$ 。这里“Conv. block”包含两个连续的卷积层 (*i.e.*, $\mathcal{B}_{conv3 \times 3}$) 和一个 1×1 卷积层。

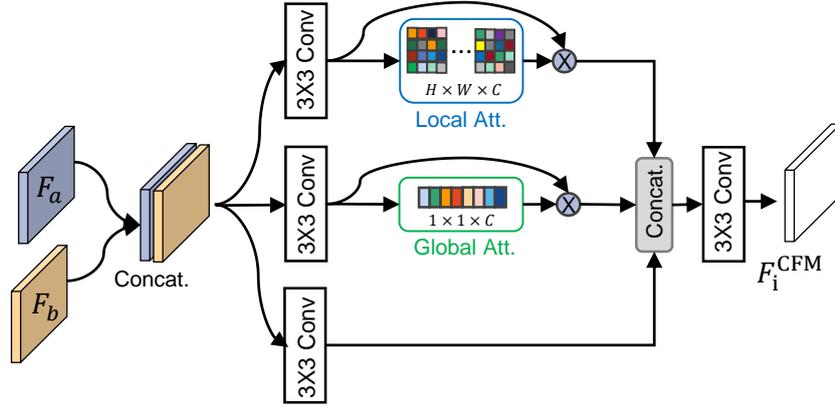


Fig. 4 所提出的跨层融合模块的架构。

Table 1 在 CVC-ClinicDB 和 Kvasir 数据集上的结果比较。最佳结果已突出显示。

Methods	CVC-ClinicDB ^[65]					Kvasir ^[66]				
	mDice	mIoU	S_α	F_β^w	E_ϕ^{mean}	mDice	mIoU	S_α	F_β^w	E_ϕ^{mean}
UNet ^[3]	0.823	0.755	0.889	0.811	0.914	0.818	0.746	0.858	0.794	0.881
UNet++ ^[4]	0.794	0.729	0.873	0.785	0.891	0.821	0.744	0.862	0.808	0.887
SFA ^[10]	0.700	0.607	0.793	0.647	0.840	0.723	0.611	0.782	0.670	0.834
PraNet ^[11]	0.899	0.849	0.936	0.896	0.963	0.898	0.840	0.915	0.885	0.944
ACSNet ^[7]	0.882	0.826	0.927	0.873	0.947	0.898	0.838	0.920	0.882	0.941
MSEG ^[67]	0.909	0.864	0.938	0.907	0.961	0.897	0.839	0.912	0.885	0.942
DCRNet ^[31]	0.896	0.844	0.933	0.890	0.964	0.886	0.825	0.911	0.868	0.933
EU-Net ^[6]	0.902	0.846	0.936	0.891	0.959	0.908	0.854	0.917	0.893	0.951
MSNet ^[56]	0.918	0.869	0.946	0.913	0.973	0.905	0.849	0.923	0.892	0.947
CCBANet ^[8]	0.868	0.812	0.916	0.861	0.935	0.853	0.777	0.887	0.831	0.901
LDNet ^[68]	0.881	0.825	0.924	0.879	0.960	0.887	0.821	0.905	0.869	0.941
EFA-Net	0.919	0.871	0.943	0.916	0.972	0.914	0.861	0.929	0.906	0.955

损失函数: 受到^[11,34,64]的启发, 本文的分割损失由 $\mathcal{L}_{seg} = \mathcal{L}_{BCE}^w + \mathcal{L}_{IoU}^w$ 定义, 其中 \mathcal{L}_{BCE}^w 和 \mathcal{L}_{IoU}^w 分别表示加权二进制交叉熵损失和加权交叉比例 (IoU) 损失。最后, 总体损失函数为:

$$\mathcal{L}_{total} = \sum_{i=1}^4 \mathcal{L}_{seg}(S_i, G) + \beta \mathcal{L}_e, \quad (8)$$

其中超参数 β 被设置为 5, G 代表真实标签。

4 实验与结果

在本节中, 首先展示实验设置的细节, 包括数据集、评估指标和实施细则。然后给出该模型和其他对比方法的分割结果。此外, 进行了消融研究, 以分析所提出的分割模型中每个关键组件的重要性。

4.1 实验设置

4.1.1 数据设置

本文在五个基准数据集上进行了息肉分割任务, 下面介绍每个数据集中的细节。

Table 2 在两个未见数据集 (*i.e.*, CVC-300 和 ColonDB) 的结果比较。

Methods	CVC-300 ^[69]					ColonDB ^[16]				
	mDice	mIou	S_α	F_β^w	E_ϕ^{mean}	mDice	mIou	S_α	F_β^w	E_ϕ^{mean}
UNet ^[3]	0.710	0.627	0.843	0.684	0.848	0.504	0.436	0.710	0.491	0.692
UNet++ ^[4]	0.707	0.624	0.839	0.687	0.834	0.482	0.408	0.693	0.467	0.680
SFA ^[10]	0.467	0.329	0.640	0.341	0.644	0.456	0.337	0.629	0.366	0.661
PraNet ^[11]	0.871	0.797	0.925	0.843	0.950	0.712	0.640	0.820	0.699	0.847
ACSNet ^[7]	0.863	0.787	0.923	0.825	0.939	0.716	0.649	0.829	0.697	0.839
MSEG ^[67]	0.874	0.804	0.924	0.852	0.948	0.735	0.666	0.834	0.724	0.859
DCRNet ^[31]	0.856	0.788	0.921	0.830	0.943	0.704	0.631	0.821	0.684	0.840
EU-Net ^[6]	0.837	0.765	0.904	0.805	0.919	0.756	0.681	0.831	0.730	0.863
MSNet ^[56]	0.865	0.799	0.926	0.848	0.945	0.751	0.671	0.838	0.736	0.872
CCBANet ^[8]	0.888	0.815	0.935	0.862	0.964	0.706	0.626	0.812	0.676	0.852
LDNet ^[68]	0.869	0.793	0.923	0.841	0.948	0.740	0.652	0.830	0.717	0.876
EFA-Net	0.894	0.830	0.941	0.878	0.961	0.774	0.696	0.855	0.753	0.884

- **Kvasir**^[66]: 这个数据集是由挪威的健康信托基金会从胃肠道内部收集来的, 包含了 1,000 张结肠镜图像。
- **CVC-ClinicDB**^[65]: 这个数据集是由 29 个结肠镜视频序列收集而成的 612 张图像, 分辨率为 288×384 。
- **CVC-ColonDB**^[16]: 这个数据集由 380 张分辨率为 500×570 的图像组成。
- **ETIS**^[70]: 这个数据集包含 196 个大小为 966×1225 的结肠镜图像。
- **CVC-300**^[69]: 这个数据集包括 60 张结肠镜图像, 分辨率大小为 500×574 。

根据^[11]中的数据分割方法, 本文随机选择了 1,450 张图像来构建训练集, 其中 900 张来自 vasir 数据集, 550 张来自 CVC-ClinicDB 数据集。剩余的图像是来自两个数据集 (*i.e.*, Kvasir 和 CVC-ClinicDB), 以及其他三个数据集 (*i.e.*, CVC-ColonDB、ETIS 和 CVC-300) 用于测试。

4.1.2 评估指标

为了评估所提出模型的有效性, 本文首先使用了分割任务中常用的两个指标^[71], 即平均 Dice 分数 (Dice) 和平均交并比 (IoU)。此外, 本文还采用了目标检测领域中广泛使用的四个指标^[72-73], 即 S-measure (S_α)^[74], F-measure (F_β^w)^[75], E_ϕ^{mean} ^[76], 以及精确率-召回率曲线。

4.1.3 实现细节

所提出的框架采用 PyTorch 实现, 并在一块 24GB 内存的 NVIDIA Tesla P40 GPU 上进行训练。本文采用 Adam 算法来优化我们的框架, 学习率为 $1e-4$ 。为了增加输入图像的多样性, 采用了不同的数据增强策略, 即随机翻转、裁剪和旋转。此外, 使用三种缩放比例 $\{0.75, 1, 1.25\}$ 训练了提出的模型, 所有输入图像的分辨率均统一缩放到 352×352 。此外, 本文的框架在训练阶段进行了 200 个 epochs 的训练, 批量大小设为 16。在模型推理阶段, 测试图像首先被调整为 352×352 的大小, 然后输入网络获得预测的地图, 在最终评估中将其还原为原始大小。

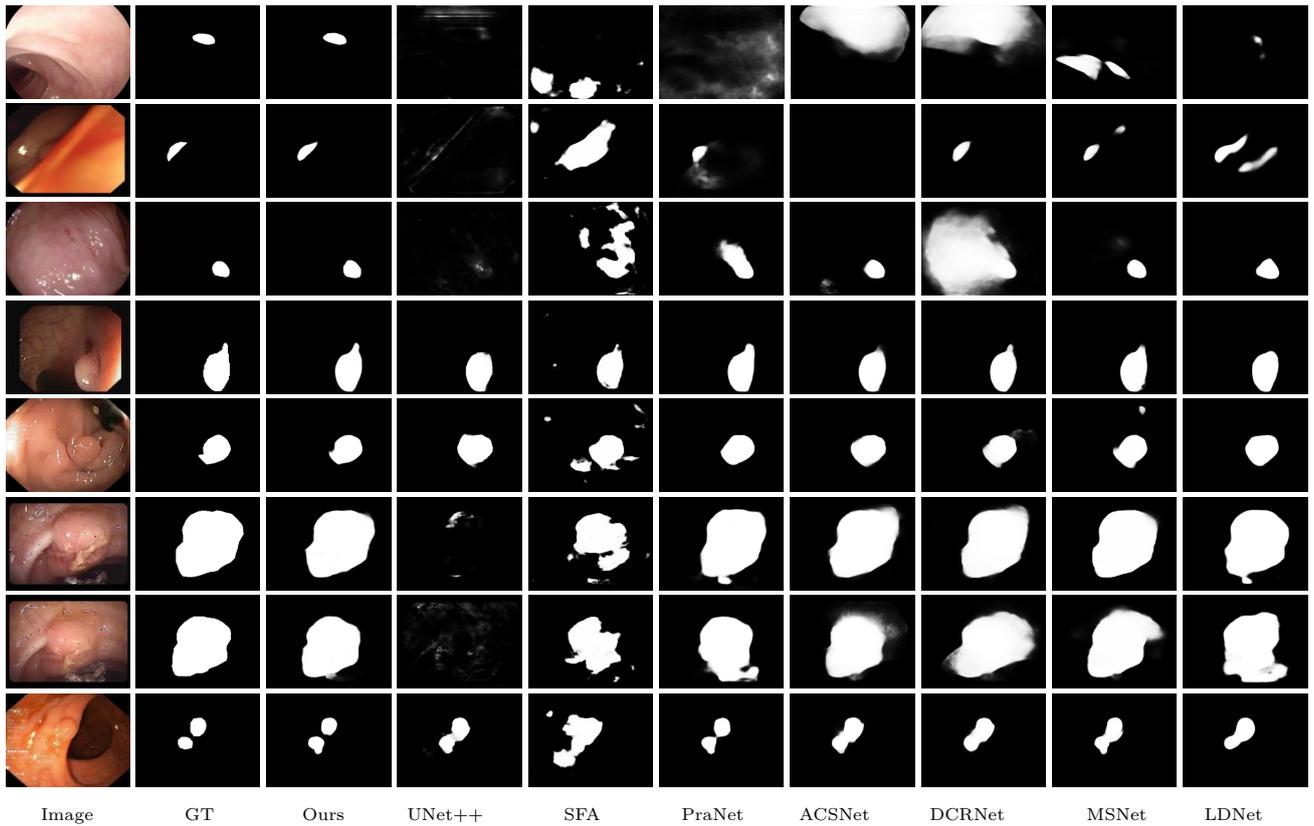


Fig. 5 定性分割结果对比。

Table 3 ETIS 数据集上的结果比较。

Methods	mDice	mIou	S_α	F_β^w	E_ϕ^{mean}
UNet ^[3]	0.398	0.335	0.684	0.366	0.643
UNet++ ^[4]	0.401	0.344	0.683	0.390	0.629
SFA ^[10]	0.297	0.217	0.557	0.231	0.531
PraNet ^[11]	0.628	0.567	0.794	0.600	0.808
ACSNet ^[7]	0.578	0.509	0.754	0.530	0.737
MSEG ^[67]	0.700	0.630	0.828	0.671	0.855
DCRNet ^[31]	0.556	0.496	0.736	0.506	0.742
EU-Net ^[6]	0.687	0.609	0.793	0.636	0.807
MSNet ^[56]	0.723	0.652	0.845	0.677	0.875
CCBANet ^[8]	0.559	0.483	0.751	0.513	0.783
LDNet ^[68]	0.645	0.551	0.788	0.600	0.841
EFA-Net	0.749	0.670	0.858	0.698	0.872

4.2 分割结果

4.2.1 对比实验

为了验证所提出模型的有效性，本文将 EFA-Net 与十一个最先进的分割方法进行比较，即 UNet^[3], UNet++^[4], SFA^[10], PraNet^[10], ACSNet^[7], MSEG^[67], DCRNet^[31], EU-Net^[6], MSNet^[56], CCBANet^[8], 和 LDNet^[68]。为了公平比较，从原论文中收集了分割图，或使用默认设置运行发布的代码，以获得预测图。

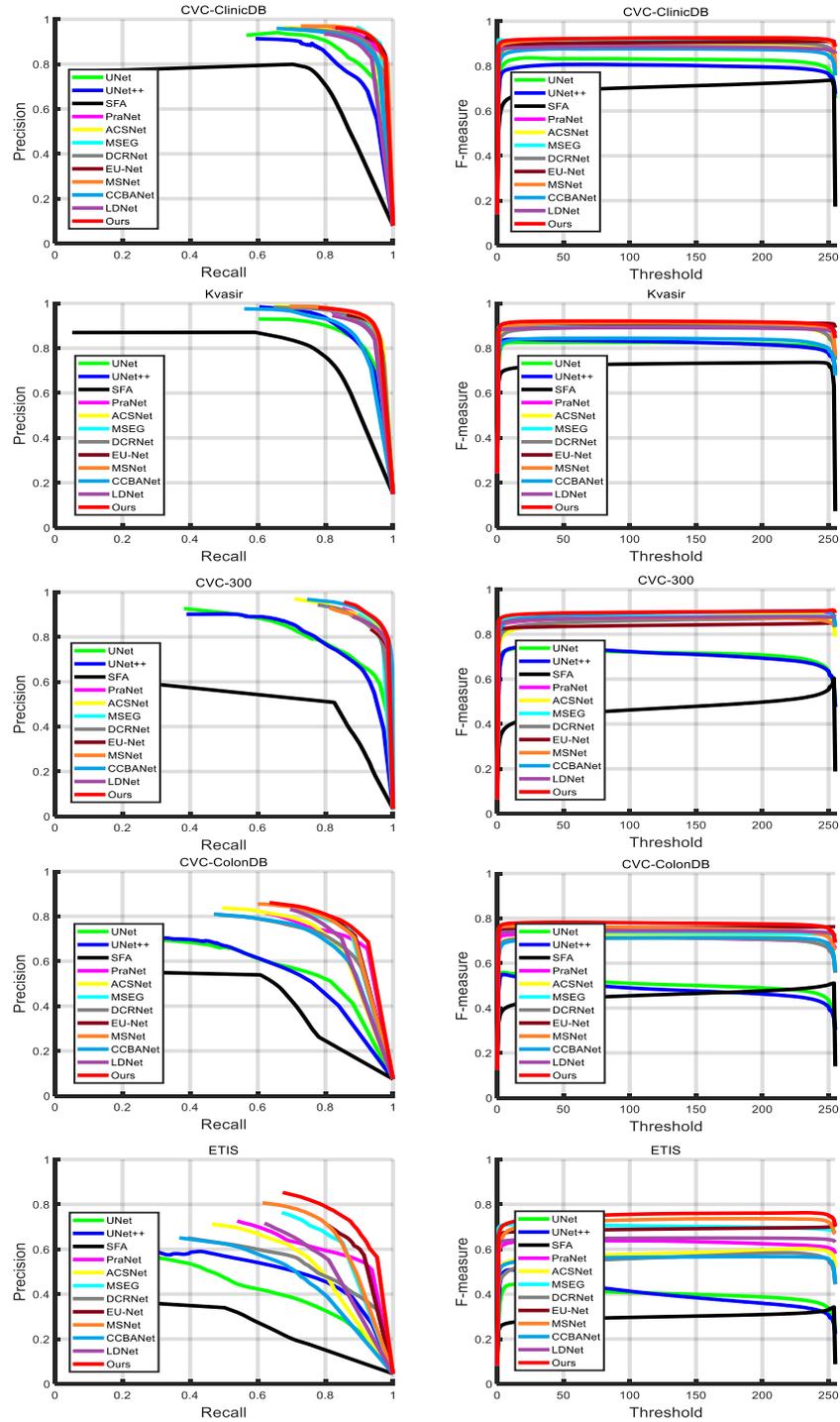


Fig. 6 精确度-召回率和 F-measure 曲线。

4.2.2 定量对比

表格 1 呈现了在 CVC-ClinicDB^[65] 和 Kvasir^[66] 数据集上不同方法的定量结果。如 Table 1 所示, 本文的 EFA-Net 在大多数指标上优于所有对比方法。例如, EFA-Net 在 CVC-ClinicDB 数据集上的 mDice、 F_{β}^w 和 S_{α} 方面相对于 LDNet 实现了 4.3%、4.2% 和 2.1% 的提升。在 Kvasir 数据集上, 本文的模型在 mDice、

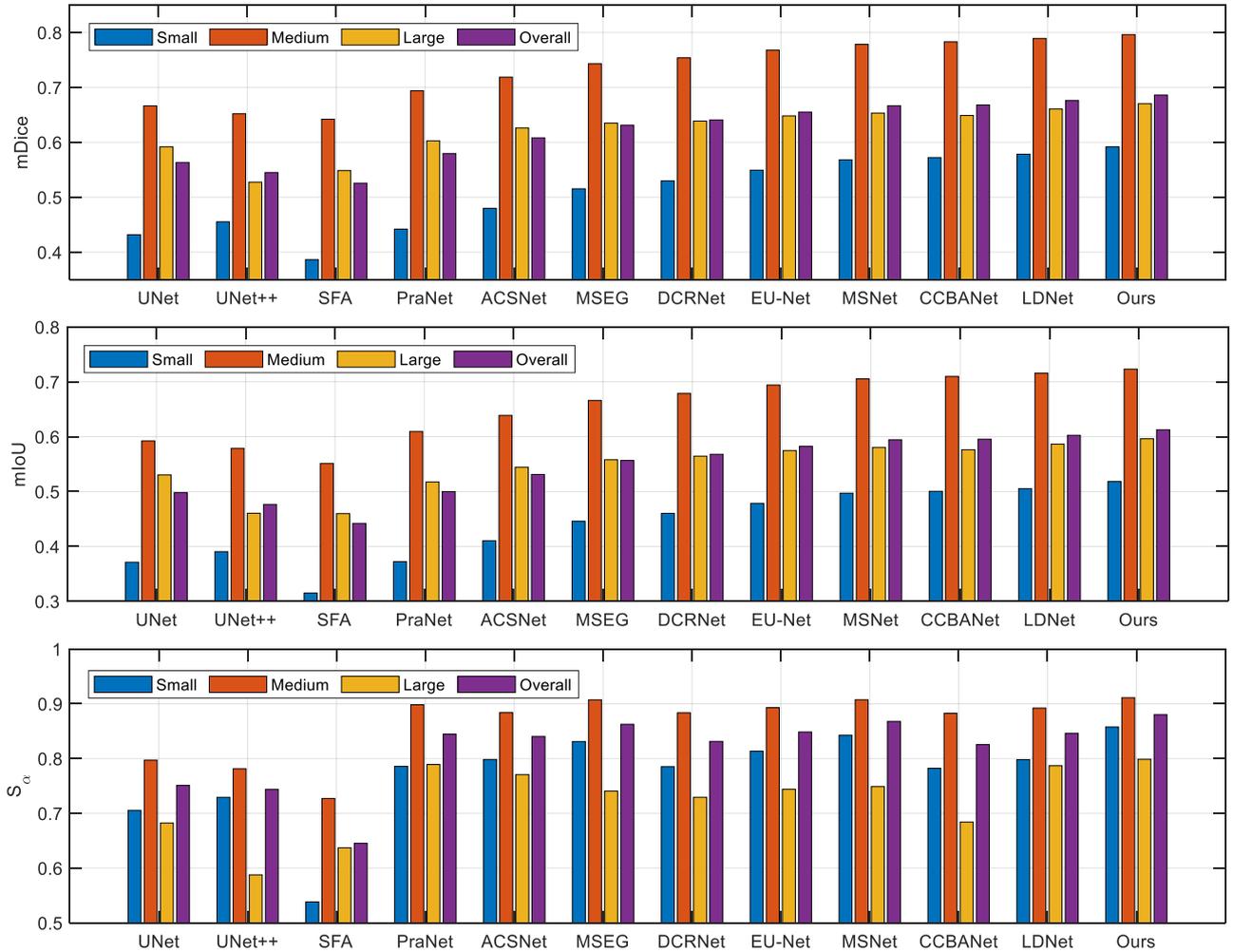


Fig. 7 基于尺度的多个大小范围的息肉评估。

mIoU、 S_α 和 F_β^w 方面相对于 LDNet 实现了 3.0%、4.9%、2.7% 和 4.3% 的提升。此外，本文的方法在 mDice、mIoU、 S_α 、和 F_β^w 方面相对于 CCBANet 分别实现了 7.2%、10.8%、4.7%、和 9.0% 的改进。这些结果表明，本文的 EFA-Net 模型在不同挑战下可以准确地分割息肉。这主要是因为所提出模型能够有效地捕获多尺度信息，并整合跨层级和多尺度特征以捕获语境语义信息，从而获得更好的分割性能。对于三个未公开的数据集上，本文的 EFA-Net 仍然表现出更好的性能，实现了准确的分割，如 Table 2和 Table 3 所示。这进一步验证了 EFA-Net 比最先进的分割方法具有更好的泛化能力。还可以看到，在 CVC-ClinicDB、CVC-300 和 ETIS 数据集上，本文方法在 E_ϕ^{mean} 方面略逊于某些方法，但差距是相对较小的。这可以归因于其在分割较大的息肉区域时可能忽视了一些细节，而这些细节在上述数据集中存在。因此，专注于增强我们的模型捕获全局和局部上下文信息的能力。通过整合这些改进，进一步的目标是在未来的研究中特别提高对于较大息肉区域的分割性能。

此外，在图 6中呈现了精确率-召回率曲线和 F-measure 曲线。如图 6所示，本文方法相对于其他最先进的息肉分割方法实现了更好的效果。因此，这些结果进一步验证了所提出的息肉分割方法的有效性。

Table 4 模型参数和运算量 (FLOPs) 的比较。

Models	UNet	UNet++	PraNet	ACSNet	DCRNet	EU-Net	MSNet	CCBANet	LDNet	EFA-Net
FLOPs	123.87	262.16	13.15	21.75	17.27	23.15	13.16	26.40	113.96	33.20
Param (M)	34.52	36.63	30.50	29.45	28.73	31.36	26.36	31.58	40.32	27.40

Table 5 CVC-ClinicDB 和 Kvasir 数据集上的消融研究结果。

Models	CVC-ClinicDB ^[65]					Kvasir ^[66]				
	mDice	mIoU	S_α	F_β^w	E_ϕ^{mean}	mDice	mIoU	S_α	F_β^w	E_ϕ^{mean}
<i>w/o</i> SCM	0.910	0.865	0.939	0.907	0.966	0.891	0.832	0.906	0.871	0.936
<i>w/o</i> CFM	0.901	0.850	0.930	0.896	0.970	0.897	0.840	0.914	0.880	0.944
<i>w/o</i> EGM	0.892	0.823	0.937	0.869	0.962	0.898	0.842	0.917	0.884	0.940
EFA-Net	0.919	0.871	0.943	0.916	0.972	0.914	0.861	0.929	0.906	0.955

4.2.3 定性比较

图 5展示了不同方法的定性分割结果。可以观察到，EFA-Net 获得的结果与真实情况更相似，而且我们的模型在处理不同挑战因素方面优于其他的对比方法。例如，在前两行中，息肉的尺寸区域非常小，本文方法仍然能够准确地定位和分割它们，而一些方法（如 UNet++ 和 SFA）完全无法定位到息肉区域。在第三行，息肉在其周围的黏膜中，而息肉与背景之间的边界非常模糊。在这种情况下，本文方法比其他对比方法能更准确地分割出息肉。还可以注意到，SFA、PraNet 和 DCRNet 产生了几处过度分割的碎片区域。在第四行和第五行中，可以观察到所有方法都可以定位到息肉的主要区域，但本文的模型获得了更准确的分割结果。在第六行和第七行中，息肉相对较大，这使得准确地分割完整的息肉区域具有挑战性。在这个挑战下，本文的模型仍然比其他方法表现更好，而 PraNet 和 LDNet 则产生了一些非息肉区域的结果。这主要是因为本文的尺度感知卷积模块可以通过利用不同的扩张卷积学习尺度增强特征，进一步聚合不同尺度的特征以处理尺度变化。因此，本文的模型能够有效地定位和分割不同尺度的息肉，并捕捉到完整的息肉区域。此外，边缘感知指导模块的融合提供了至关重要的边界线索，用于区分息肉和背景之间的语义边缘，从而显著提高了分割性能。此外，本文的方法还能有效地分割多个息肉（如第八行）。总的来说，定性比较结果进一步验证了本文的方法在处理息肉分割的挑战因素方面具有良好的性能。

4.2.4 基于尺度的评估

尺度变化是息肉分割任务中最具挑战性的因素之一。本研究提出了完全学习尺度感知特征并有效融合交叉层特征以增强学习特征的区分能力，有助于处理尺度变化。因此，对不同方法进行了基于尺度的评估。本文计算一个涉及给定图像中息肉区域尺度的比例 (r) 来描述息肉区域的尺度。为此，定义了三种类型的息肉尺度，*i.e.*, 1) 当 $r < 0.025$ 时，将其称为“小”型；2) 当 $r > 0.2$ 时，将其称为“大”型；3) 当比例在 $[0.025, 0.2]$ 范围内时，称之为“中等”型。此外，本文从 5 个使用的结肠镜数据集（测试图像）中收集了一个混合数据集，其中 53.1%，36.1%，和 10.8% 的图像分别属于这三种类型。图 7以三个度量指标 (*i.e.*, mDice, mIoU, 和 S_α) 的形式呈现了基于息肉尺度进行属性研究的比较结果。基于结果，可以得出以下观察结果：1) 在不同息肉尺度条件下，本文方法在性能上优于其他对比方法；2) 本文模型和所有对比方法在分割“中等”型息肉时的性能较好，而在分割其他类型时性能相对较低。这表明尺度变化仍然是息肉分割的一个挑战，并且相对容易分割“中等”型息肉。

Table 6 不同数据增强策略在 Kvasir 数据集的影响。

Strategies	mDice	mIou	S_α	F_β^w	E_ϕ^{mean}
Baseline	0.899	0.841	0.914	0.889	0.946
+ Crop	0.905	0.848	0.919	0.894	0.950
+ Random flipping	0.910	0.858	0.924	0.901	0.955
+ Rotation	0.900	0.841	0.915	0.888	0.949
All strategies	0.914	0.861	0.929	0.906	0.955

4.2.5 模型复杂度比较

为了研究所提出模型的复杂性，表 4 中展示模型参数和 FLOPs 的对比结果。在表 4 中，# 参数的测量单位是百万 (M)，浮点运算 (FLOPs) 的测量单位是千兆 (G)。可以观察到，与其它方法相比，本文的模型参数相对较少，表明其在定位和分割息肉方面的效率更高。然而，应该注意的是，与一些比较方法相比，本文的模型需要更多的 FLOPs。这主要是由于使用了尺度感知的卷积模块来学习多尺度特征表示，以及使用了跨层融合模块来获取局部和全局上下文信息。这两个模块都涉及各种卷积操作。因此，在未来的工作中，将开发轻量级网络来提高所提模型的效率，并实现实时的息肉分割。

4.3 消融实验

本节研究了提出的 EFA-Net 中每个模块的重要性，消融结果在 Table 5 显示。此外，还研究了不同数据增强策略的效果。

SCM 的有效性: 为了验证 SCM 的重要性，本文使用了一个 1×1 的卷积层代替它，标记为 “w/o SCM”。如 Table 5 所报告的，观察到 “w/o SCM” 在 Kvasir 数据集上的性能下降较大。结果验证了 SCM 的有效性，它可以学习与尺度相关的特征以处理尺度变化。

CFM 的有效性: 为了研究 CFM 的贡献，本文直接拼接了跨层级特征，记为 “w/o CFM”。可以看到，在没有使用提出的 CFM 的情况下，性能下降，这表明 CFM 可以融合跨层级特征，以利用局部和全局上下文信息来促进息肉分割的贡献。

EGM 的有效性: 如 Table 5 所示，结果表明，通过 EGM 将边缘感知特征引入解码器，本文的 EFA-Net 实现了很好的分割性能。

数据增强的有效性: 为了研究不同的数据增强策略的影响，本文进行了一系列实验，将它们与不进行任何数据增强的“基准”方法进行了比较。不同数据增强策略的比较结果在表 6 中呈现。从结果可以明显看出，当三种数据增强策略都被使用时，本文的模型实现了最高的性能。相反，当没有应用数据增强时，模型的性能明显较差。

4.4 失败案例和分析

以上实验结果证明了所提出的 EFA-Net 的优越性和有效性。然而，在涉及较大的息肉或模糊边界等复杂场景中，本文的模型可能难以准确定位和分割息肉区域。这些失败案例的视觉表现如图 8 所示。在前两个例子中，息肉的庞大尺寸使得完全分割变得复杂。正如这些案例所示，我们的模型未能完全分割整个息肉区域，导致部分段丢失。在第 3 行和第 4 行，息肉与周围黏膜之间的边界不清晰。在这种情况下，本文的 EFA-Net 忽视了边界细节的复杂性或生成了不精确的分割图。本文的未来工作目标是改进 EFA-Net，以



Fig. 8 提出模型得到的一些失败案例的可视化结果。

完全分割更大尺寸的息肉，并更有效地利用边界指标，从而提高整体分割性能。这些发展将对结直肠癌的早期检测和预防措施的有效性做出重要贡献。

4.5 讨论

局限性. 如图 8 所示，本文方法在大规模息肉分割时无法产生令人满意的分割图。为了解决这个问题，本文提出利用一个大型选择性核网络^[77]，它可以动态调整其空间感受野，以有效地捕捉复杂背景下各种物体的上下文。这种方法有助于提高不同大小和变化的息肉的分割性能。此外，值得注意的是，在临床实践中，一个具有高准确性的实时检测系统是至关重要的。这样的系统可以在结肠镜检查过程中协助医生采取必要的措施。然而，当前的模型仍然需要大量的计算资源，这可能限制其在临床环境中的实际应用。为了克服这个限制，计划探索模型压缩的方法，如网络剪枝^[78]和知识蒸馏^[79]。这些技术可以用于开发一个专门针对临床实时息肉分割的轻量级网络。通过降低计算成本，同时保持足够的准确性，使模型在医疗环境中更加可行和易于实际使用。

据我们所知，噪声和图像伪影在医学图像中经常遇到，因此解决这些情况对模型的临床实用性至关重要。首先，在将输入图像输入分割模型之前，应用去噪算法或滤波器对图像进行处理是有益的。这种方法有助于减轻噪声的影响，提高图像质量，并降低噪声对分割过程的影响。其次，通过人工生成带有噪声或伪影的图像来扩充训练数据，可以使模型学习并对现实世界中存在的各种变化更加稳健。这种扩充技术有助于模型更好地概括推断过程中不同类型的噪声和伪影。第三，在获得初始分割结果后，应用后处理技术可以进一步优化输出结果。这些技术可能涉及平滑边界，去除小的伪影或施加几何约束来改善整体分割效果。总

体而言, 本文的息肉分割模型通过采用去噪技术、数据扩充和后处理策略的组合方法, 能够有效处理噪声和图像伪影, 从而增强分割性能和临床适用性。

此外, 在息肉分割中处理非均匀纹理和颜色变化是一个重要挑战, 因为存在不同的组织类型。为了解决这些挑战, 有几个解决方案可以讨论。首先, 非均匀纹理和颜色变化可以在息肉内以不同的尺度表现出来。因此, 引入多尺度分析可以使模型捕捉到局部细节和整体上下文。通过采用包含多个分支或不同大小的感受野的网络架构, 可以实现这一目标, 使模型能够自适应地从各个细粒度级别捕获信息。其次, 通过对训练数据进行各种转换, 如颜色转换, 可以增强模型处理非均匀纹理和颜色变化的能力。第三, 注意力机制可以增强模型对信息区域的关注, 同时抑制不相关或噪声信息。通过将注意力机制融入网络架构, 模型可以关注到呈现非均匀纹理或颜色变化的特定区域, 提高准确分割息肉的能力。最后, 将模型预训练在包含多样化图像(包括具有非均匀纹理和颜色变化的图像)的大规模数据集上, 可以提供强大的初始特征表示。通过在下游息肉分割任务上对模型进行微调, 可以使其在测试过程中有效地推广到纹理和颜色的变化。总的来说, 通过考虑这些方法, 模型对非均匀纹理和颜色变化的处理性能可以得到提高, 从而实现更准确和稳健的息肉分割结果。

5 总结

本文提出了一种新的 EFA-Net 框架, 用于结肠镜图像中的息肉分割。为探索边缘语义信息, 提出一种边缘感知引导模块来学习边缘增强特征, 并将其输入到解码器网络中以提升分割性能。此外, 提出了尺度感知卷积模块和跨层级融合模块, 用于捕获多尺度信息并融合跨层级特征以处理尺度变化问题。在基准数据集上的广泛实验结果验证了所提出的 EFA-Net 在准确分割息肉方面的有效性。

参考文献

- [1] ZHOU T, ZHANG Y, CHEN G, et al. Edge-aware feature aggregation network for polyp segmentation[J]. Machine Intelligence Research, 2024.
- [2] IBTEHAZ N, RAHMAN M S. Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation[J]. Neural Networks, 2020, 121: 74-87.
- [3] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer Assisted Intervention. Munich, Germany: Springer, 2015: 234-241.
- [4] ZHOU Z, SIDDIQUEE M M R, TAJBAKHS N, et al. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 39(6): 1856-1867.
- [5] JHA D, SMEDSRUD P H, RIEGLER M A, et al. Resunet++: An advanced architecture for medical image segmentation[C]//IEEE International Symposium on Multimedia. San Diego, CA, USA: IEEE, 2019: 225-2255.
- [6] PATEL K, BUR A M, WANG G. Enhanced u-net: A feature enhancement network for polyp segmentation[C]//Proc. IEEE Int. Robots and Vision. 2021: 181-188.
- [7] ZHANG R, LI G, LI Z, et al. Adaptive context selection for polyp segmentation[C]//International Conference on Medical Image Computing and Computer Assisted Intervention. Lima, Peru: Springer, 2020: 253-262.

-
- [8] NGUYEN T C, NGUYEN T P, DIEP G H, et al. Ccbanet: Cascading context and balancing attention for polyp segmentation[C]//International Conference on Medical Image Computing and Computer Assisted Intervention. Strasbourg, France: Springer, 2021: 633-643.
- [9] MURUGESAN B, SARVESWARAN K, SHANKARANARAYANA S M, et al. Psi-Net: Shape and boundary aware joint multi-task deep network for medical image segmentation[C]//Engineering in Medicine and Biology Society. Germany, Germany: IEEE, 2019: 7223-7226.
- [10] FANG Y, CHEN C, YUAN Y, et al. Selective feature aggregation network with area-boundary constraints for polyp segmentation[C]//International Conference on Medical Image Computing and Computer Assisted Intervention. Shenzhen, China: Springer, 2019: 302-310.
- [11] FAN D P, JI G P, ZHOU T, et al. Pranet: Parallel reverse attention network for polyp segmentation[C]//Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent. 2020: 263-273.
- [12] HWANG S, OH J, TAVANAPONG W, et al. Polyp detection in colonoscopy video using elliptical shape feature [C]//IEEE International Conference on Image Processing: Vol. 2. IEEE, 2007: II-465.
- [13] BERNAL J, SÁNCHEZ J, VILARINO F. Towards automatic polyp detection with a polyp appearance model[J]. Pattern Recognition, 2012, 45(9): 3166-3182.
- [14] WANG Y, TAVANAPONG W, WONG J, et al. Part-based multiderivative edge cross-sectional profiles for polyp detection in colonoscopy[J]. IEEE Journal of Biomedical and Health Informatics, 2013, 18(4): 1379-1389.
- [15] MAMONOV A V, FIGUEIREDO I N, FIGUEIREDO P N, et al. Automated polyp detection in colon capsule endoscopy[J]. IEEE Transactions on Medical Imaging, 2014, 33(7): 1488-1502.
- [16] Tajbakhsh N, Gurudu S R, Liang J. Automated polyp detection in colonoscopy videos using shape and context information[J]. IEEE Transactions on Medical Imaging, 2016, 35(2): 630-644.
- [17] LI Q, YANG G, CHEN Z, et al. Colorectal polyp segmentation using a fully convolutional neural network[C]//IEEE International Congress on Image and Signal Processing, Biomedical Engineering and Informatics. IEEE, 2017: 1-5.
- [18] JHA D, SMEDSRUD P H, RIEGLER M A, et al. Resunet++: An advanced architecture for medical image segmentation[C]//IEEE International Symposium on Multimedia. 2019: 225-2255.
- [19] ZHOU Z, SIDDIQUEE M M R, TAJBAKHSH N, et al. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 39(6): 1856-1867.
- [20] ZHANG Z, LIU Q, WANG Y. Road extraction by deep residual u-net[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(5): 749-753.
- [21] LI X, CHEN H, QI X, et al. H-DenseUNet: hybrid densely connected unet for liver and tumor segmentation from ct volumes[J]. IEEE Transactions on Medical Imaging, 2018, 37(12): 2663-2674.
- [22] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention u-net: Learning where to look for the pancreas[A]. 2018.
- [23] ZHOU T, ZHOU Y, HE K, et al. Cross-level feature aggregation network for polyp segmentation[J]. Pattern Recognition, 2023, 140: 109555.

-
- [24] YUE G, HAN W, JIANG B, et al. Boundary constraint network with cross layer feature integration for polyp segmentation[J]. *IEEE Journal of Biomedical and Health Informatics*, 2022, 26(8): 4090-4099.
- [25] LIU F, HUA Z, LI J, et al. Dbmf: Dual branch multiscale feature fusion network for polyp segmentation[J]. *Computers in Biology and Medicine*, 2022, 151: 106304.
- [26] SU Y, CHENG J, ZHONG C, et al. Accurate polyp segmentation through enhancing feature fusion and boosting boundary performance[J]. *Neurocomputing*, 2023, 545: 126233.
- [27] SONG P, LI J, FAN H. Attention based multi-scale parallel network for polyp segmentation[J]. *Computers in Biology and Medicine*, 2022, 146: 105476.
- [28] TOMAR N K, JHA D, BAGCI U, et al. Tganet: text-guided attention for improved polyp segmentation[C]// *International Conference on Medical Image Computing and Computer Assisted Intervention*. Springer, 2022: 151-160.
- [29] WEI J, HU Y, ZHANG R, et al. Shallow attention network for polyp segmentation[C]// *International Conference on Medical Image Computing and Computer Assisted Intervention*. Strasbourg, France: Springer, 2021: 699-708.
- [30] WU H, ZHAO Z, ZHONG J, et al. Polypseg+: A lightweight context-aware network for real-time polyp segmentation [J]. *IEEE Transactions on Cybernetics*, 2022.
- [31] YIN Z, LIANG K, MA Z, et al. Duplex contextual relation network for polyp segmentation[C]// *Proc. Int. Symp. Biomed. Imaging*. 2022: 1-5.
- [32] LIN Y, WU J, XIAO G, et al. Bsca-net: Bit slicing context attention network for polyp segmentation[J]. *Pattern Recognition*, 2022, 132: 108917.
- [33] DING H, JIANG X, LIU A Q, et al. Boundary-aware feature propagation for scene segmentation[C]// *Proceedings of the IEEE International Conference on Computer Vision*. 2019: 6819-6829.
- [34] QIN X, ZHANG Z, HUANG C, et al. Basnet: Boundary-aware salient object detection[C]// *Conference on Computer Vision and Pattern Recognition*. 2019: 7479-7489.
- [35] LIU K, ZHAO Y, NIE Q, et al. Weakly supervised 3d scene segmentation with region-level boundary awareness and instance discrimination[C]// *European Conference on Computer Vision*. Springer, 2022: 37-55.
- [36] ZHOU T, ZHOU Y, GONG C, et al. Feature aggregation and propagation network for camouflaged object detection [J]. *IEEE Trans. Image Process.*, 2022, 31: 7036-7047.
- [37] LI F, DU X, ZHANG L, et al. Image feature fusion method based on edge detection[J]. *Information Technology and Control*, 2023, 52(1): 5-24.
- [38] NAWAZ M, NAZIR T, MASOOD M, et al. Melanoma segmentation: A framework of improved densenet77 and unet convolutional neural network[J]. *International Journal of Imaging Systems and Technology*, 2022, 32(6): 2137-2153.
- [39] MASKELIUNAS R, DAMASEVICIUS R, VITKUTE-ADZGAUSKIENE D, et al. Pareto optimized large mask approach for efficient and background humanoid shape removal[J]. *IEEE access*, 2023.
- [40] CHEN H, QI X, YU L, et al. Dcan: Deep contour-aware networks for object instance segmentation from histology images[J]. *Medical Image Analysis*, 2017, 36: 135-146.

-
- [41] ZHANG Z, FU H, DAI H, et al. Et-net: A generic edge-attention guidance network for medical image segmentation [C]//International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, 2019: 442-450.
- [42] FAN D P, JI G P, ZHOU T, et al. Pranet: Parallel reverse attention network for polyp segmentation[C]//International Conference on Medical Image Computing and Computer Assisted Intervention. Lima, Peru: Springer, 2020: 263-273.
- [43] RAMASAMY L K, PADINJAPPURATHU S G, KADRY S, et al. Detection of diabetic retinopathy using a fusion of textural and ridgelet features of retinal images and sequential minimal optimization classifier[J]. PeerJ computer science, 2021, 7: e456.
- [44] WANG R, CHEN S, JI C, et al. Boundary-aware context neural network for medical image segmentation[J]. Medical Image Analysis, 2022, 78: 102395.
- [45] WANG K, ZHANG X, ZHANG X, et al. EANet: Iterative edge attention network for medical image segmentation [J]. Pattern Recognition, 2022, 127: 108636.
- [46] KIM Y, KANG B N, KIM D. San: Learning relationship between convolutional features for multi-scale object detection[C]//Proceedings of the European Conference on Computer Vision. 2018: 316-331.
- [47] CAO J, PANG Y, ZHAO S, et al. High-level semantic networks for multi-scale object detection[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 30(10): 3372-3386.
- [48] ZHOU T, FAN D P, CHEN G, et al. Specificity-preserving rgb-d saliency detection[J]. Computational Visual Media, 2023, 9: 297-317.
- [49] LI J, FANG F, MEI K, et al. Multi-scale residual network for image super-resolution[C]//Proceedings of the European Conference on Computer Vision. 2018: 517-532.
- [50] LI J, FANG F, LI J, et al. MDCN: Multi-scale dense cross network for image super-resolution[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 31(7): 2547-2561.
- [51] HE J, DENG Z, QIAO Y. Dynamic multi-scale filters for semantic segmentation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 3562-3572.
- [52] GU J, KWON H, WANG D, et al. Multi-scale high-resolution vision transformer for semantic segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 12094-12103.
- [53] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]//Proc. IEEE Conf. Comput. Vision Pattern Recognit. 2017: 2881-2890.
- [54] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs[J]. IEEE Trans. Pattern Anal. Mach. Intell., 2017, 40(4): 834-848.
- [55] YANG H, HUANG W, QI K, et al. CLCI-Net: Cross-level fusion and context inference networks for lesion segmentation of chronic stroke[C]//Medical Image Computing and Computer Assisted Intervention. Springer, 2019: 266-274.
- [56] ZHAO X, ZHANG L, LU H. Automatic polyp segmentation via multi-scale subtraction network[C]//International Conference on Medical Image Computing and Computer Assisted Intervention. Strasbourg, France: Springer, 2021: 120-130.

- [57] SRIVASTAVA A, JHA D, CHANDA S, et al. Msrf-net: A multi-scale residual fusion network for biomedical image segmentation[J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 26(5): 2252-2263.
- [58] YANG H, ZHOU T, ZHOU Y, et al. Flexible fusion network for multi-modal brain tumor segmentation[J]. *IEEE Journal of Biomedical and Health Informatics*, 2023.
- [59] GAO S H, CHENG M M, ZHAO K, et al. Res2net: A new multi-scale backbone architecture[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 43(2): 652-662.
- [60] FAN D P, ZHOU T, JI G P, et al. Inf-net: Automatic covid-19 lung infection segmentation from ct images[J]. *IEEE Transactions on Medical Imaging*, 2020, 39(8): 2626-2637.
- [61] SUN Y, WANG S, CHEN C, et al. Boundary-guided camouflaged object detection[C]//International Joint Conference on Artificial Intelligence. 2022: 1335-1341.
- [62] ZHAO J X, LIU J J, FAN D P, et al. EGNet: Edge guidance network for salient object detection[C]//International Conference on Computer Vision. 2019: 8779-8788.
- [63] DAI Y, GIESEKE F, OEHMCKE S, et al. Attentional feature fusion[C]//IEEE WACV. 2021: 3560-3569.
- [64] WEI J, WANG S, HUANG Q. F3Net: Fusion, Feedback and Focus for Salient Object Detection[C]//AAAI Conference on Artificial Intelligence. 2020: 12321-12328.
- [65] 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.
- [66] JHA D, SMEDSRUD P H, RIEGLER M A, et al. Kvasir-seg: A segmented polyp dataset[C]//International Conference on Multimedia Modeling. Daejeon, Korea: Springer, 2020: 451-462.
- [67] HUANG C H, WU H Y, LIN Y L. Hardnet-mseg: A simple encoder-decoder polyp segmentation neural network that achieves over 0.9 mean dice and 86 fps[A]. 2021.
- [68] ZHANG R, LAI P, WAN X, et al. Lesion-aware dynamic kernel for polyp segmentation[C]//International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, 2022: 99-109.
- [69] VÁZQUEZ D, BERNAL J, SÁNCHEZ F J, et al. A benchmark for endoluminal scene segmentation of colonoscopy images[J]. *Journal of Healthcare Engineering*, 2017, 2017: 4037190.
- [70] 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.
- [71] YANG C, GUO X, ZHU M, et al. Mutual-prototype adaptation for cross-domain polyp segmentation[J]. *IEEE J. Biomed. Health Informat.*, 2021, 25(10): 3886-3897.
- [72] ZHOU T, FAN D P, CHENG M M, et al. RGB-D salient object detection: A survey[J]. *Computational Visual Media*, 2021, 7(1): 37-69.
- [73] FAN D P, JI G P, SUN G, et al. Camouflaged object detection[C]//Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE, 2020: 2777-2787.
- [74] FAN D P, CHENG M M, LIU Y, et al. Structure-measure: A new way to evaluate foreground maps[C]//International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 4548-4557.

-
- [75] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency-tuned salient region detection[C]//Conference on Computer Vision and Pattern Recognition. Miami, FL, USA: IEEE, 2009: 1597-1604.
- [76] FAN D P, GONG C, CAO Y, et al. Enhanced-alignment measure for binary foreground map evaluation[C]//International Joint Conference on Artificial Intelligence. Stockholm, Sweden: IJCAI, 2018: 698-704.
- [77] LI Y, HOU Q, ZHENG Z, et al. Large selective kernel network for remote sensing object detection[A]. 2023.
- [78] MOLCHANOV P, MALLYA A, TYREE S, et al. Importance estimation for neural network pruning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019: 11264-11272.
- [79] GOU J, YU B, MAYBANK S J, et al. Knowledge distillation: A survey[J]. International Journal of Computer Vision, 2021, 129: 1789-1819.