

卷积神经网络在结直肠息肉辅助诊断中的应用综述

考文涛, 李明, 马金刚^{*}

山东中医药大学 智能与信息工程学院, 济南 250355

+ 通信作者 E-mail: ma_jingang@126.com

摘要: 结直肠癌是一种恶性肿瘤, 主要发生在结肠和直肠的组织中, 其早期发现和治疗具有重要意义。结直肠癌的早期检测和预防主要是对病人的肠道进行视觉检查, 从而筛查结直肠息肉, 但人工检查存在漏诊率高等弊端。基于卷积神经网络(CNN)的辅助诊断系统在结直肠息肉的诊断方面表现出最先进的性能, 是目前计算机辅助诊断领域的研究热点。根据近几年发表的相关重要文献, 对卷积神经网络在结直肠息肉辅助诊断中的应用进行系统综述。首先介绍了结直肠息肉诊断领域的常用数据集, 其中包括图片和视频数据集; 其次分别对CNN在结直肠息肉检测、分割以及分类中的应用进行系统阐述, 对各算法的主要改进思路、优缺点以及性能进行深入分析, 旨在为研究人员提供更系统的参考, 并对深度学习模型的可解释性进行总结; 最后对基于CNN的结直肠息肉辅助诊断的各类算法进行总结, 并对未来的研究方向进行展望。

关键词: 结直肠息肉; 卷积神经网络(CNN); 计算机辅助诊断; 可解释性

文献标志码: A **中图分类号:** TP391

Review of Application of Convolutional Neural Network in Auxiliary Diagnosis of Colorectal Polyps

KAO Wentao, LI Ming, MA Jingang^{*}

College of Intelligence and Information Engineering, Shandong University of Traditional Chinese Medicine, Jinan 250355, China

Abstract: Colorectal cancer is a malignant tumor that mainly occurs in the tissues of the colon and rectum, and its early detection and treatment are of great significance. The early detection and prevention of colorectal cancer mainly involves visual examination of the patient's intestines to screen for colorectal polyps, but manual examination has the disadvantage of high misdiagnosis rate. The auxiliary diagnostic system based on convolutional neural networks (CNN) has shown the most advanced performance in the diagnosis of colorectal polyps, and is currently a research hotspot in the field of computer-aided diagnosis. Based on important literature published in recent years, a systematic review of the application of convolutional neural networks in the auxiliary diagnosis of colorectal polyps is conducted. Firstly, the commonly used datasets in the field of colorectal polyp diagnosis are introduced, including image and video datasets. Secondly, the application of CNN in colorectal polyp detection, segmentation, and classification is systematically elaborated. The main improvement ideas, advantages and disadvantages, and performance of each algorithm are analyzed in depth, aiming to provide researchers with a more systematic reference, and summarize the interpretability of deep learning models. Finally, a summary of various algorithms for assisting the diagnosis of

基金项目: 国家自然科学基金面上项目(82074579, 82174528); 2022年山东省研究生优质教育教学资源项目(SDYAL2022041)。

This work was supported by the National Natural Science Foundation of China (82074579, 82174528), and the Graduate High-Quality Education Teaching Resources Project of Shandong Province in 2022 (SDYAL2022041).

收稿日期: 2023-10-18 **修回日期:** 2024-01-03

colorectal polyps based on CNN is provided, and future research directions are prospected.

Key words: colorectal polyps; convolutional neural networks (CNN); computer aided diagnosis; interpretability

结直肠癌(colorectal cancer, CRC)位居全球癌症死亡率第二位,在2020年夺去了近100万人的生命,严重危害人类健康^[1]。结直肠癌的筛查过程是通过内镜检查发现结直肠息肉,进而通过对结直肠息肉的活检明确病变性质^[2]。但其效果受多种因素的影响,如医生的操作技术水平、患者的术前准备以及结直肠息肉的可见性等因素都会对检查结果产生影响,从而导致结直肠息肉和早期结直肠癌的漏诊率居高不下^[3]。因此,越来越多的计算机辅助诊断(computer aided diagnosis, CAD)系统被开发,有效降低了结肠镜检查中的病变漏检率。

近年来,卷积神经网络(convolutional neural networks, CNN)在医学图像处理领域取得了显著的进展,已经被广泛用于肿瘤检测、糖尿病眼底病变检测、皮肤病变识别等领域,能够帮助医生更准确、快速地进行诊断,提高了效率和准确性。CNN主要通过卷积层、池化层和全连接层等组件,对医学图像进行多层次的特征提取和抽象。这些特征不仅包括形态、纹理等视觉特征,还包括一些细微的特征^[4]。此外, CNN还具有良好的泛化能力,能够适应不同类型和尺度的医学图像。

在结直肠息肉辅助诊断领域,近年来发表了几篇优秀综述^[5-8],文献[5-6]只对深度学习在结直肠息肉区域提取的应用进行论述,文献[7-8]对人工智能在结直肠息肉辅助诊断中的应用做了简要叙述,但都未对基于CNN的结直肠息肉辅助诊断方法进行系统的概述。针对上述不足,本文主要从数据集介绍、CNN在结直肠息肉区域提取、分类中的应用以及基于CNN的深度学习模型的可解释性四方面进行总结。在结直肠息肉区域提取方面,将其按照检测和分割进行分类,并进一步对不同网络模型进行分别阐述。在结直肠息肉分类方面,将其按照分类任务的不同依次进行阐述。在模型可解释性方面,对常用的可解释性方法进行总结,并进一步分析各个方法的优缺点。最后对CNN在结直肠息肉辅助诊断领域的进展进行全面总结,并对未来研究方向进行展望。

1 数据集

结直肠息肉辅助诊断是医学影像领域的一个重

要任务,旨在帮助医生准确诊断结直肠息肉病变,并为手术治疗提供指导。为了推动相关算法模型的发展和评估,研究人员开发了一些公开数据集,这些数据集包含了丰富的结直肠影像数据,可供科研机构 and 学术界使用。通过利用这些数据集,研究人员可以训练和测试各种算法模型,以提高结直肠息肉诊断的准确性和效率,为医生提供更可靠的诊断支持。本章将介绍一些常见的结直肠息肉诊断数据集。

1.1 CVC-ClinicDB

CVC-ClinicDB^[9]数据集是用于结直肠镜图像分析的一个公开数据集,于2015年提出,该数据集主要用于结直肠息肉检测任务,即识别结肠镜图像中的息肉。该数据集由从29个结肠镜检查视频中提取的612个息肉静态图像组成,所有图像的分辨率固定为384×288像素。因其同一个片段内的各图像内容十分相似,所以对于提高模型的泛化能力较为困难。CVC-ClinicDB数据集常被研究人员用于模型的测试和验证。

1.2 CVC-ColonDB

CVC-ColonDB^[10]数据集中有300个静态样本,每个图像的大小为574×500像素,是基于15个结肠镜检查视频构建的。该数据集特点在于它包括了不同程度病变的结直肠息肉,从轻微的变形到明显的肿瘤,这使得研究人员能够对不同类型和不同程度的息肉进行研究和评估,有利于提高模型的泛化能力。CVC-ColonDB已被广泛用于结直肠息肉辅助诊断算法的评估和比较。

1.3 Kvasir-SEG

Kvasir-SEG^[11]数据集于2020年提出,是息肉分割领域一个常用的数据集。Kvasir-SEG数据集由1000张息肉图像及其相应的掩码组成,样本的分辨率在332×487到1920×1072像素之间。因其每个图像都配有对应的像素级别标注,更有利于模型进行精细化分割,这对于深度学习中的图像分割任务是关键的。并且数据集中包含的多样性图像有助于提高CNN模型的泛化能力。

1.4 CVC-PolypHD

CVC-PolypHD^[12]数据集于2017年提出,包含164个高分辨率静态图像,每个图像的大小为1920×1080

像素。该数据集包含超光谱图像,这意味着每个像素都有多个波段的光谱信息,提供了更为丰富的数据,有助于更准确地分析结直肠息肉的特征。该数据集主要用于超光谱图像中结直肠息肉检测的研究,由于样本数量有限,数据集主要用于测试。

1.5 CVC-ClinicalVideoDB

CVC-ClinicalVideoDB^[13]于2017年首次提出,它是一个由36个视频序列组成的视频数据集,分辨率为768×576。与静态图像数据集不同,该数据集提供了更丰富的时空信息,使模型可以更全面地学习结直肠息肉在不同时间的变化和特征。该数据集包含各种特性的息肉,被用于训练实时检测模型。

1.6 ASU-Mayo

ASU-Mayo^[14]数据集是一个在2017年推出的视频数据集,包括三种不同的尺寸,大小为712×480、856×480、1 920×1 080,通常用于训练和测试实时息肉检测系统。因其中包含丰富的息肉细节信息,更有利于模型学习深层特征以及提升泛化性。数据集包含38个视频序列,进一步被划分为20个训练视频和18个测试视频。表1总结了结直肠息肉辅助诊断常用数据集。

2 CNN在结直肠息肉区域提取中的应用

结直肠息肉区域提取是结直肠息肉诊断与分析中的关键步骤,它旨在从结直肠影像中准确地定位和提取出结直肠息肉的区域,以便进行后续的特征分析和分类,这一步骤的准确性和效果直接影响到结直肠息肉的诊断和治疗方案的确定。为了实现结直肠息肉区域的提取,研究人员已经开发出多种检测和分割算法。其中,基于卷积神经网络(CNN)的方法,已经取得了显著的进展。下面将分别对结直肠息肉的检测算法和分割算法进行讨论。

2.1 基于CNN的结直肠息肉检测算法

结直肠息肉检测是指利用医学影像技术对结直

肠内的息肉进行识别和定位的过程。传统的结直肠息肉检测方法主要依赖于人工检查,即通过将内窥镜引入结直肠腔内,通过直接观察结直肠黏膜的方式来检测结直肠息肉。然而,这种方法存在一些局限性,如操作依赖医生的经验和技能,易受到视觉疲劳和主观因素的影响,导致漏诊和误诊。近年来,随着深度学习技术的发展,基于CNN的结直肠息肉检测方法受到了广泛关注。基于CNN的目标检测方法可以分为两阶段目标检测和单阶段目标检测。两阶段目标检测首先生成候选目标区域,然后对每个候选区域进行特征提取和分类。代表性的方法包括R-CNN^[15]、Fast R-CNN^[16]和Faster R-CNN^[17]等。单阶段目标检测直接在输入图像中预测目标的位置和类别,无需生成候选区域,它们以密集的采样方式进行预测,通常具有较高的检测速度。代表性的方法包括YOLO(you only look once)^[18]和单发多框检测器(single shot multibox detector, SSD)^[19]等。

在深度学习方法普及之前,研究人员普遍使用基于机器学习的传统方法进行息肉检测。后来,Zhu等人^[20]使用机器学习加深度学习的方法进行内窥镜图像病变检测。他们收集了180张内窥镜图像,使用CNN从这些图像中提取特征,并将特征输入支持向量机(support vector machine, SVM)以增强泛化能力。所提出的方法比以前的基于颜色和纹理特性的传统模型提供了更好的性能,但是该方法不能实现端到端的检测。后来,R-CNN系列算法被提出。R-CNN是于2014年引入的一种目标检测算法,使用选择性搜索来定位感兴趣的对象,并从每个感兴趣的区域独立提取特征进行处理。Sornapudi等人^[21]提出了一种基于R-CNN的改进算法,用于检测图像中的息肉。该方法分别使用ResNet-50^[22]和ResNet-101^[22]主干进行特征提取,通过迁移学习提高模型性能,并且使用了一部分类似息肉的自然图像进行训练提高了息肉检测率。与之前发表的研究相比,此模型提

表1 结直肠息肉辅助诊断数据集

Table 1 Common datasets for auxiliary diagnosis of colorectal polyps

数据集	样本数量	分辨率	成像方式	类型	任务
CVC-ClinicDB ^[9]	612	384×288	结肠镜检查	静态图片	检测、定位、分割
CVC-ColonDB ^[10]	300	574×500	结肠镜检查	静态图片	分割
Kvasir-SEG ^[11]	1 000	332×487到1 920×1 072	结肠镜检查	静态图片	分割
CVC-PolypHD ^[12]	164	1 920×1 080	结肠镜检查	静态图片	检测、定位、分割
CVC-ClinicalVideoDB ^[13]	36	768×576	结肠镜检查	视频	检测、定位、分割
ASU-Mayo ^[14]	38	712×480; 856×480; 1 920×1 080	结肠镜检查	视频	检测、定位、分割

供了更好的性能,在训练了1 000个epoch后,平均检测精度达到89.94%,召回率达到91.64%。由此可见,相对于传统方法,使用深度学习方法进行息肉检测,不仅提高了检测准确率,而且省去了很多人工操作。

Ren等人^[17]在2015年提出了Faster R-CNN模型,该算法与原来的R-CNN算法不同之处在于,它使用了与检测网络共享所有特征的区域建议网络(region proposal network, RPN),使其速度更快,可以用于实时检测。Liu等人^[23]使用了带有Inception-ResNet-v2^[24]骨干网络的Faster R-CNN模型来检测息肉,并选择交叉熵和平滑熵作为损失函数。他们使用了一个私人数据集,其中包含2 000多张图像。实验表明,它的平均精度达到了90.645%。Chen等人^[25]基于Faster R-CNN网络提出改进,进一步提升了息肉检测精度。他们使用VGG-16^[26]作为特征提取网络,其中包含13个卷积层以及4个最大池化层,有利于提取不同层次和不同尺度的特征图。并且他们为了进一步提高网络的特征提取能力,在特征提取网络后附加了一个注意力模块,以关注有用的特征通道,削弱无助的特征通道的贡献,网络结构如图1所示。实验结果表明,与现有的算法相比,所提出的息肉检测网络的准确性有了很大的提高,精度达到了94.3%,但是作者并未使用公共数据集进行测试,导致结果缺乏说服力。上述研究表明,基于Faster R-CNN的结肠息肉检测算法不仅大大提升了精度,在运行速度方面,也比基于R-CNN的算法更快。

Mask R-CNN是一种基于深度学习的目标检测和语义分割模型,它是在Faster R-CNN模型的基础上进行扩展和改进的。与传统的目标检测算法相比,Mask R-CNN不仅可以准确地检测出图像中的目标位置,还能够生成每个目标的像素级分割掩码。

而且传统的目标检测算法使用感兴趣区域池化(ROI Pooling)来提取感兴趣区域的特征,这可能导致信息损失和精确度下降。Mask R-CNN引入了感兴趣区域对齐(RoI Align)操作,通过在特征图上进行双线性插值,更精确地对齐像素级别的特征,提高了分割的准确性。Wittenberg等人^[27]提出了一种基于ResNet-101主干的Mask R-CNN模型,用于息肉自动检测。在训练中,他们使用了三个结肠镜数据集,包括两个公共数据集和一个包含2 484张图像的私人数据集,这有利于提高模型的泛化能力,并且使用了迁移学习和图像增强技术提升精度。他们对三个数据集分别进行了测试,实验结果表明该算法的最高检测精度为86%,召回率为93%。虽然该方法在性能上有了明显提升,但与Faster R-CNN模型相比仍然存在一定差距。

YOLO是2016年提出的第一个单阶段检测方法,也是第一个实现实时目标检测的方法,但其对微小物体的检测精度非常低。同年,Liu等人提出了单发多框检测器(SSD),它结合了Faster R-CNN中区域提议网络(RPN)和YOLO中单阶段检测网络的优点。2016年底,YOLO推出了升级版的YOLOv2^[28]。在YOLOv2中,每个卷积层之后增加了批量归一化层,以减少样本差异,并从网络中移除全连接层,以消除固定图像大小的限制。Liu等人^[29]实现了基于SSD的息肉检测方法,他们评估了三种不同的特征提取器,包括ResNet50、VGG-16和Inception-V3^[30],并在基础网络之上额外增加了特征层。在所有的特征提取器中,Inception-V3获得了最好的精度和召回率,并且检测速度达到32 FPS,超过了临床应用的最低要求。该研究的局限性在于难以识别肠内容物、浆膜和出血灶以及会漏诊小的、扁平的息肉。Zheng等人^[31]

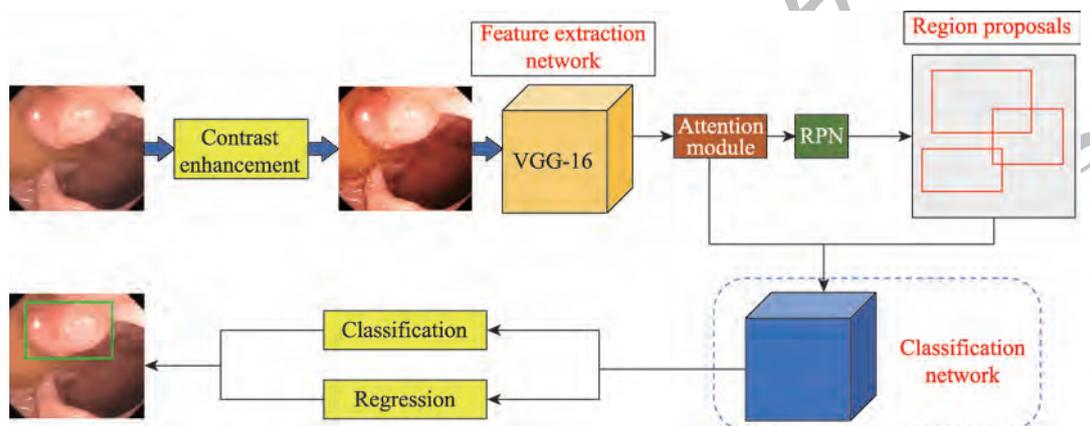


图1 基于改进Faster R-CNN的网络结构图

Fig.1 Network architecture diagram based on improved Faster R-CNN

介绍了一种基于YOLO的实时息肉检测模型。该模型在非医学图像中进行预训练,然后使用三个公共结肠镜检查数据集进行训练,最后在一个包含196张图像的公共数据集上进行了测试。实验结果表明,YOLO在定位任务中达到了79.3%的精度,时间效率为0.06秒/帧,并且使用多个数据集进行训练,提高了模型的泛化性。虽然基于SSD和YOLO的模型在精度上下降了,但是它们在速度上得到了很大的提升,更有利于协助医生在临床检查中定位结直肠息肉。后来,研究人员提出了一系列提高模型检测精度的方法。Wan等人^[32]和司丙奇等人^[33]分别基于YOLOv5^[34]模型提出了自己的改进算法,用于息肉目标检测。Wan等人在每一级特征提取网络的顶层集成了一个自注意力机制模块,该模块包括一个全局平均池化层和两个全连接层。司丙奇等人在标准YOLOv5模型的基础上,在特征提取网络的C3模块中加入了SE注意力机制模块。他们都采用在模型中加入注意力机制的方法,以提高富含信息的特征通道的贡献,削弱无用通道的干扰。实验结果表明,他们的平均精度分别为91.5%和94.7%。该方法能准确识别息肉图像,特别是对小息肉和对比度不明显的息肉,但局限性在于网络参数量大,对硬件配置要求高。Ma等人^[35]提出了一种基于SSD的改进模型,采用Inception-V2^[36]作为特征提取模块,并且使用Momentum梯度下降算法以及Softmax损失算法替换掉原有算法。该模型可以从多个维度提取特征层信息,并融合六个特征层进行预测,在一定程度上提高了对小目标的检测能力。测试结果表明,该方法的准确率为96.04%,使

用这种方法,可以高精度地定位和检测结肠镜图像中的息肉,但速度慢,网络复杂度高。上述研究表明,基于SSD和YOLO的算法在速度上得到了提升,但检测精度却下降了。在后续的研究中,通过对模型进行改进,例如加入注意力机制或者使用更加复杂的特征提取网络等方法,虽然提升了精度,但是模型复杂度高,速度慢,因此始终无法找到检测精度与检测速度间的平衡。但是在最新的研究中,人们提出了一些方法解决上述问题,可以作为参考,例如薛钦原等人^[37]基于YOLOv7^[38]模型,提出了一种改进的结直肠息肉检测算法,该模型采用了含隐式知识的高效解耦头以及全局注意力机制,增强了模型对浅层特征的提取能力,充分利用隐含信息,并且对SPPCSPC模块进行优化,用串行计算替换原本的并行计算,减少了模型参数。模型的准确率达到95.8%,速度为126 FPS,兼顾了检测精度和速度。但是在后续测试中,模型的泛化性能表现一般,还有待提升。

在SSD和YOLO算法提出之前,人们主要使用R-CNN系列算法检测视频中的息肉。Mo等人^[39]使用基于VGG主干的Faster R-CNN算法来检测内镜视频中的息肉。虽然该方法可以降低结肠镜检查中息肉漏检率,但在测试中息肉的假阳性率很高,而且检测速度较慢,无法满足临床使用的要求。SSD和YOLO算法的提出,为结直肠息肉的实时检测提供了有效的方法。Ma等人^[40]提出YOLOv5m与Swin Transformer相结合的方法,从实时结肠镜检查视频中检测息肉,网络结构如图2所示。作者使用特征提取模块从单个帧中提取特征,并将提取的信息与Swin Transformer

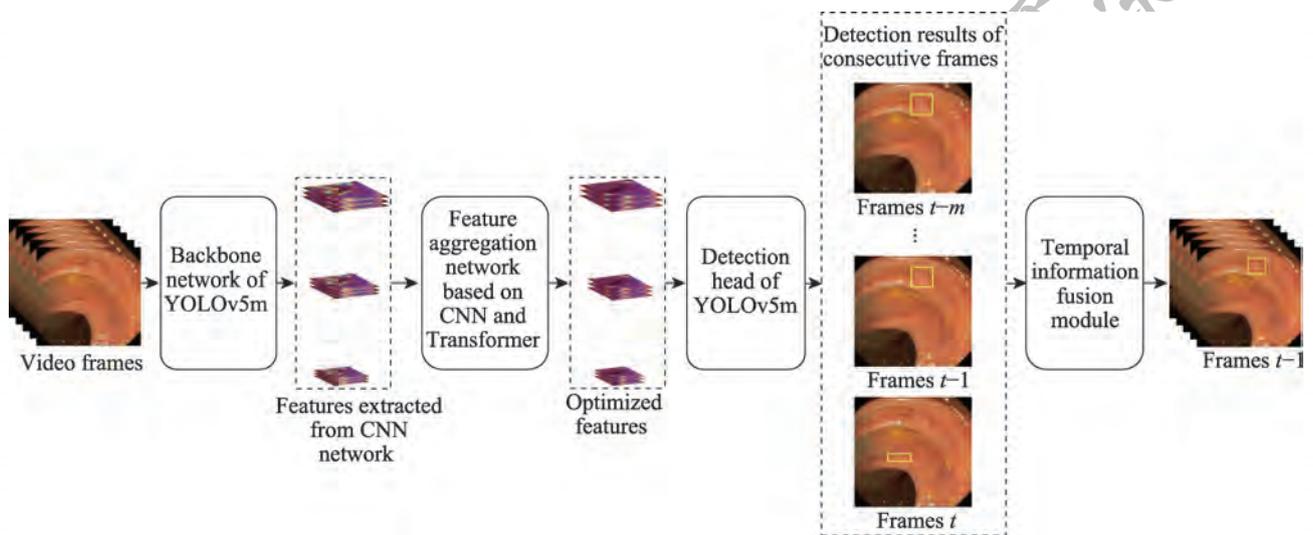


图2 基于改进YOLO的网络结构图

Fig.2 Network architecture diagram based on improved YOLO

提取的特征相结合,以提高模型的整体性能。Transformer 和 YOLOv5m 网络的结合是通过用 Swin Transformer 取代 YOLO 网络的 Neck 模块来实现的,而且引入了一个时间信息融合模块,以最大限度地减少每帧中白光反射的影响。实验结果表明,该模型在 CVC-ClinicalVideoDB 数据集上测试精度为 83.6%,召回率为 73.1%。并且相较于之前基于 Faster R-CNN 的方法,速度也得到了提升。

综上所述,虽然 R-CNN 系列网络在结直肠息肉检测方面取得了显著的进展,并且研究人员通过对网络进行改进,大大提高了检测精度,但在检测速度上还无法满足实时检测的要求,而目前基于 SSD 和 YOLO 的算法虽然精度较低,但速度较快,并且经过研究,通过在特征提取部分加入注意力机制,有利于增强模型特征提取能力,提升检测精度。因此在检测视频中的息肉时往往使用 SSD 和 YOLO 算法来实现。上述研究都是利用结肠镜图像进行检测,而无线胶囊内窥镜也是结直肠检查的一种主要方式,Nadimi 等人^[41]提出了一种基于 Faster R-CNN 目标检测模型,用于在无线胶囊内窥镜图像中自主检测结直肠息肉。他们使用改进的 ZF-Net^[42]网络为主干,并结合迁移学习、预处理和数据增强等方法。通过研究,虽然上述单 CNN 检测算法的推理速度较快,但多 CNN 检测算法则有更高的定位能力和识别准确性。Albuquerque 等人^[43]通过堆叠集成方法将单 CNN 模型结合起来,实验证明了这种新方法的潜力,可以作为结直肠息肉检测算法未来的研究方向。

目前 CNN 在结直肠息肉的检测上尚存在以下问题:结直肠息肉中的一些小型病灶往往难以被准确检测和定位,因为它们图像中的尺寸很小,可能被模糊、噪声或其他结构干扰所掩盖。同时,在临床实践中,结直肠息肉的快速检测对于及时干预和治疗非常重要。然而,一些基于 CNN 的方法在速度和计算效率方面仍存在一定的限制。因此,未来需要进一步研究结直肠息肉检测,以寻求同时提高准确率和速度的方法和技术,这将为结直肠息肉的早期发现和治理提供更可靠和高效的支持。表 2 为不同结直肠息肉检测算法的性能比较。表 3 为不同结直肠息肉检测算法的改进及效果。

2.2 基于 CNN 的结直肠息肉分割算法

结直肠息肉分割是指将结直肠影像中的息肉区域从背景中准确地分离出来的过程。相比于息肉检测,结直肠息肉分割可以更准确地定位和标记出每个息肉的位置和形状,有助于医生进行后续的诊断和治疗计划。最早实现端到端目标分割的网络是 2015 年提出的全卷积网络 (fully convolutional network, FCN)^[44],它首次将卷积神经网络应用于像素级别的语义分割任务。在 FCN 中,通过使用转置卷积操作来进行上采样,将低分辨率的特征图恢复到输入图像的尺寸,从而实现像素级别的目标分割,但是在多层次的下采样过程中,FCN 可能会失去输入图像的某些空间信息。针对该问题,Yu 等人^[45]提出了一种将 FCN 和 DenseNet^[46]相结合的方法,用于分割结肠镜图像中的息肉,有效解决了空间信息缺失问

表 2 不同结直肠息肉检测算法的性能比较

Table 2 Performance comparison of different colorectal polyp detection algorithms

参考文献	检测方法	准确率/%	召回率/%	数据集	类型
Zhu 等人 ^[20]	CNN+SVM	80.01		私人	静态图片
Sornapudi 等人 ^[21]	R-CNN	89.94	91.64	CVC-ColonDB	静态图片
Liu 等人 ^[23]	Faster R-CNN	90.65		私人	静态图片
Chen 等人 ^[25]	Faster R-CNN	94.30	92.50	私人	静态图片
Wittenberg 等人 ^[27]	Mask R-CNN	86.00	93.00	CVC-ClinicDB	静态图片
Liu 等人 ^[29]	SSD	73.60	80.30	CVC-ClinicDB	静态图片
Zheng 等人 ^[31]	YOLO	79.30		ETIS-Larib	静态图片
Wan 等人 ^[32]	YOLO	91.50	89.90	Kvasir-SEG	静态图片
司丙奇等人 ^[33]	YOLO	94.70	79.50	Kvasir-SEG	静态图片
Ma 等人 ^[35]	SSD	96.04		私人	静态图片
薛钦原等人 ^[37]	YOLO	95.80	93.90	CVC-ClinicDB	静态图片
Mo 等人 ^[39]	Faster R-CNN	86.20	98.10	CVC-ClinicalVideoDB	视频
Ma 等人 ^[40]	YOLO	83.60	73.10	CVC-ClinicalVideoDB	视频
Nadimi 等人 ^[41]	Faster R-CNN	98.00		私人	静态图片

表3 不同结直肠息肉检测算法的改进及表现

Table 3 Improvement and performance of different colorectal polyp detection algorithms

参考文献	主要改进	效果	局限性
Zhu 等人 ^[20]	使用CNN替换传统提取特征的方法	避免医生主观因素的影响,且能提取到深层次的特征	不能实现端到端的检测
Sornapudi 等人 ^[21]	使用预训练的ResNet-50和ResNet-101网络进行微调	提高模型泛化能力,得到了更好的定位结果	检测速度较慢
Liu 等人 ^[23]	在Faster R-CNN网络中采用Inception-ResNet-v2主干提取特征	检测精度得到提高	未使用不同数据集证明其泛化性
Chen 等人 ^[25]	在Faster R-CNN网络中引入注意力机制	关注有用的特征通道,削弱无用的特征通道的贡献;提高准确率	未使用公共数据集进行测试
Wittenberg 等人 ^[27]	在Mask R-CNN网络中采用ResNet-101主干提取特征,并使用迁移学习	提高模型泛化能力;缓解了数据集过小的问题	检测精度有待进一步提升
Wan 等人 ^[32]	在YOLOv5网络中引入注意力机制	明显提升对小的息肉以及对对比度不明显的息肉检测效果	网络参数量大,对硬件配置要求高
司丙奇等人 ^[33]	在YOLOv5网络中引入SE注意力模块	在不明显增加计算复杂度的情况提升了模型的检测性能	召回率有待进一步提升
Ma 等人 ^[35]	在SSD网络中采用Inception-V2主干提取特征	高精度地定位和检测结肠镜图像中的息肉	速度慢,网络复杂度高
薛钦原等人 ^[37]	在YOLOv7网络中引入含隐式知识的高效解耦头以及全局注意力机制,优化SPPCSPC模块	防止分类和回归任务之间相互干扰;增强模型对浅层特征的提取能力;减少模型参数	泛化性能表现一般,有待提升
Mo 等人 ^[39]	在Faster R-CNN网络中采用VGG主干提取特征	降低结肠镜检查中息肉缺失的风险	测试中息肉的假阳性率很高,而且检测速度较慢
Ma 等人 ^[40]	在YOLOv5m网络中引入Swin Transformer和时间信息融合模块	提取不同空间中的所有关键信息;最大限度地减少每帧中白光反射的影响	召回率有待进一步提升

题。该方法在FCN模型的卷积和反卷积阶段应用不同的密集连接模块,并且在实验中设计了不同的层数。结果表明,随着层数的逐渐增加,准确率得到了提高,当层数为78时,Dice系数达到0.891 0。但问题是随着网络变得更深和更复杂,FCN的模型大小和计算复杂性会增加,这可能对部署和实时应用造成一定挑战。

U-Net^[47]是一种主要用于图像分割任务的深度学习模型,有效解决了FCN存在的一系列问题,该模型通过收缩路径(Encoder)和扩展路径(Decoder)的结合,实现了从输入图像到像素级别预测的端到端处理。收缩路径用于提取图像的上下文特征,而扩展路径则通过上采样和特征融合来恢复细节信息,从而在分割任务中取得准确且高质量的分割结果。Sun 等人^[48]借鉴U-Net网络,提出了一种新的深度学习模型,用于结直肠息肉分割。作者使用ResNet-50作为编码器的骨干网络,并且在其最后一个模块中引入扩张卷积来改进编码器的特征表示能力,以便在不降低分辨率的情况下学习高级语义特征。他们还设计了全新的解码器,由四个上采样层以及一个卷积层组成,使得网络参数量变小,节约计算资源。

实验结果表明,该模型的Dice系数达到0.824 8。Mohapatra 等人^[49]提出了一个名为UPolySeg的深度学习分割模型,该模型基于U-Net框架,包括三层编码器-解码器结构,同样使用了扩张卷积来扩大感受野的面积。不同的是,作者实现了几种图像预处理技术,如相干传输和对对比度受限的自适应直方图均衡化(contrast limited adaptive histogram equalization, CLAHE),以应对结肠镜检查图像中的不同挑战。结果表明,使用图像预处理后,模型性能得到大大提升,Dice系数达到0.968 6。虽然上述方法通过使用扩张卷积使模型表现出良好的性能,但是它对于更复杂或更多样的输入图像处理极为有限。为了解决此问题,Dumitru 等人^[50]提出了一种DUCK-Net网络,该网络使用新的DUCK卷积块替换了U-Net中的3×3卷积块,可以进行更深入的特征选择,并结合残差下采样机制,使其能够使用编码器中每个分辨率的初始图像信息。该方法的Dice系数达到0.950 2,并表现出强大的泛化能力。虽然新的卷积块能让网络精确地训练最关键的部分,但它的一个缺点是会破坏后续层的精细细节,此问题有待进一步改进。

在卷积神经网络中,注意力机制是一种重要的

技术,用于增强模型对输入数据中特定区域或特征的关注度。徐昌佳等人^[51]和贾立新等人^[52]都利用注意力机制改进了U-Net分割模型。徐昌佳等人利用DoubleUNet^[53]网络,在两个子网络的解码器部分引入注意力机制,并且加入了密集空洞空间金字塔池化(DenseASPP)模块,最后利用Focal Tversky Loss函数^[54]作为损失函数;而贾立新等人通过将通道分组空间增强注意力、轴向自注意力和逆向注意力三种注意力机制进行融合,增强模型对于小的以及边界模糊的息肉的分割性能,并且提升了模型的泛化能力。经过改进,这两种模型都取得了很好的结果,Dice系数分别为0.954 3和0.921 7。但是这两种模型的缺点为模型非常庞大,需要使用高性能系统进行培训和部署,且不够灵活。为了解决这个问题,Tran等人^[55]提出了一种残差递归U-Net架构,以在确保模型性能的同时降低模型的大小,其网络结构如图3。在这项研究中,作者改变了卷积单元中的卷积核数量,并且利用残差和递归结构在卷积函数重用方面的优势,改进了模型。与传统的U-Net模型和其他模型相比,新模型不仅尺寸更小,而且性能优越。基于以上优点,该模型可以有效地在低性能设备上开发应用程序。

在图像分割领域,除了上述的FCN和U-Net网络之外,SegNet^[56]和DeepLabV3+^[57]也是被经常使用的网络。SegNet是一个深度卷积神经网络,同样是基于FCN的一种改进,被用来将低分辨率的特征与输入分辨率进行映射,以实现像素语义的息肉分割。Eu等人^[58]提出了一种基于SegNet的模型来对结肠镜检查图像进行分割,使用VGG-19^[26]框架作为特征提取网络,网络结构如图4所示。在训练之前,首先使用预定义的阈值对样本进行预处理,其中对于每个红色和绿色通道,图像像素被设置为50和15。作者使用该阈值过滤掉所有值低于预定义阈值的像素,去除所有黑色和非息肉区域。然后将图像输入到

SegNet模型,并通过分割对息肉进行定位。结果表明,所提出的方法是可行的,Dice系数达到0.923 5。但是该模型的局限性在于,对息肉边缘分割表现不佳,无法处理结构更为复杂的息肉。DeepLabV3+是最新提出的用于图像分割的网络,采用了空洞卷积以及空洞空间金字塔池化(atrous spatial pyramid pooling, ASPP),以高分辨率精确捕捉图像语义信息。朱世祺等人^[59]使用DeepLabV3+网络对结肠直息肉进行分割,采用ResNet-50替换了原模型中的Xception^[60]架构对输入图像进行特征提取。实验表明,使用DeepLabV3+网络可以对结肠直息肉进行较为精准的分割,在CVC-ClinicDB数据集上Dice系数达到0.899 0。尽管该模型对息肉边界分割效果很好,但是存在息肉内部分割不充分的现象。针对此问题,王亚刚等人^[61]对DeepLabV3+网络做出改进,在编码器部分,通过引入密集预测单元(dense prediction cell, DPC)架构来对多尺度特征信息进行获取,在解码器部分,使用3层深度可分离卷积替换3×3卷积,以减少上采样过程中的信息丢失的问题。通过改进,模型对于结肠直息肉的分割性能有了一定的提升,Dice系数达到0.936 0。除了上述网络之外,Hu等人^[62]提出了一个基于金字塔池化的息肉分割网络,名为PPNet,该网络首先采用金字塔池化Transformer(pyramid pooling transformer, P2T)作为编码器来提取特征,然后在解码器部分使用结合通道注意机制的金字塔特征融合模块(pyramid feature fusion module, PFFM)学习全局上下文特征,并在编码器的每个分支中引入了记忆保持金字塔池化模块(memory-keeping pyramid pooling module, MPPM),以提取不同尺度的上下文特征。结果表明,模型在处理边缘模糊以及息肉内部有剧烈变化方面具有良好的分割能力。但是该研究的局限性在于,模型对于极小的息肉区域以及高度模糊的边缘可能存在遗漏或错误分割的情况,此问题在其他模型中也同样存在,还有待解决。

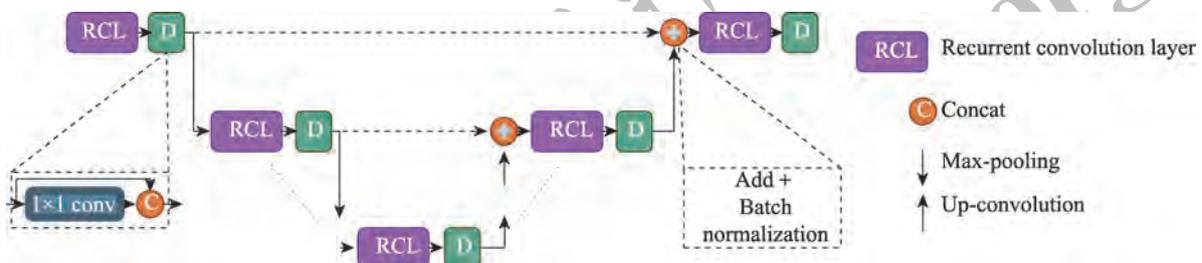


图3 mRR-Unet的网络结构图

Fig.3 Network architecture diagram of mRR-Unet

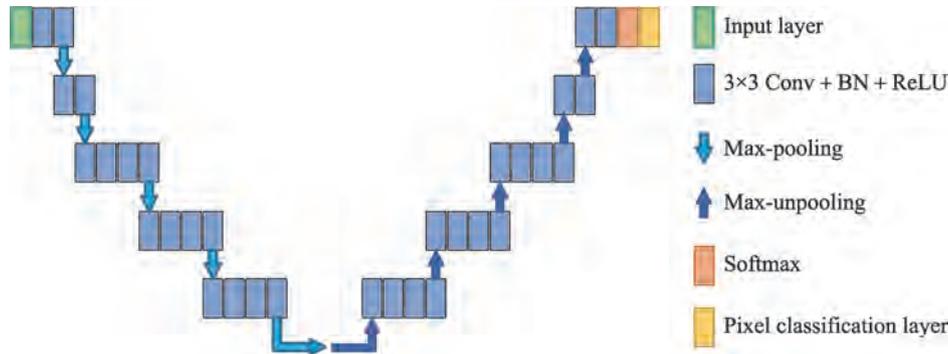


图4 SegNet的网络结构图

Fig.4 Network architecture diagram of SegNet

综上所述,卷积神经网络在结直肠息肉分割任务中已取得了较好的性能,它们能够准确定位和分割目标。尤其是U-Net网络,其采用多尺度的特征融合策略,能够有效检测和分割小目标,并且经过研究人员的改进,通过在网络中添加注意力机制,提高了网络提取图像特征的能力,这不仅提高了小目标的分割精度,也提升了模型的泛化能力。最近新提出的语义分割模型,例如DeepLabV3+网络,在结直肠

息肉分割领域也表现出了良好的性能。但是目前在结直肠息肉分割领域仍然存在一些挑战,例如结直肠息肉与周围正常组织的对比度较低以及结直肠息肉的边界可能模糊不清等问题,而且在临床实际应用中,结直肠息肉分割算法需要具备实时性能,以便在结肠镜检查过程中提供及时的诊断结果。而要解决这些问题,需要研究人员的进一步的研究和创新。表4展示了不同结直肠息肉分割算法的性能比较。

表4 不同结直肠息肉分割算法的性能比较

Table 4 Performance comparison of different colorectal polyp segmentation algorithms

参考文献	分割方法	主要改进	效果	局限性	数据集	Dice系数
Yu 等人 ^[45]	FCN	在卷积和去卷积阶段应用不同的密集连接	解决了空间信息缺失问题;特征可重复使用	随着网络层数增加,计算复杂性会增加	CVC-ColonDB	0.891 0
Mohapatra 等人 ^[49]	U-Net	引入扩张卷积和图像预处理	增加感受野的面积;降低了腺瘤性息肉的漏诊率	对于更复杂或更多样的输入图像处理极为有限	Kvasir-SEG	0.968 6
Dumitru 等人 ^[50]	U-Net	使用 DUCK 卷积块,并结合残差下采样机制	可在编码器中捕捉和处理多种分辨率的图像信息,提升泛化能力	DUCK 卷积块会破坏后续层的精细节	Kvasir-SEG	0.950 2
徐昌佳等人 ^[51]	U-Net	引入注意力机制,Dense-ASPP、Focal Tversky Loss 函数	提高网络提取图像特征的能力;提高小目标的分割精度	模型非常庞大,对设备要求高	CVC-ClinicDB	0.954 3
贾立新等人 ^[52]	U-Net	融合通道分组空间增强注意力、轴向自注意力和逆向注意力	增强了模型的泛化能力	网络参数量大,不够灵活	CVC-ClinicDB	0.921 7
Tran 等人 ^[55]	U-Net	引入残差和递归结构	降低模型尺寸更小,性能更加优越	训练数据量相对较少	CVC-ClinicDB	0.945 9
Eu 等人 ^[58]	SegNet	图像预处理	增强有关信息的可检测性和最大限度地简化数据	对息肉边缘分割表现不佳	CVC-ClinicDB	0.923 5
王亚刚等人 ^[61]	DeepLabV3+	引入 DPC 架构和 3 层深度可分离卷积	进一步获取多尺度特征信息,减少上采样过程中的信息丢失的问题	网络复杂度较高,计算量较大	CVC-ClinicDB	0.936 0
Hu 等人 ^[62]	PPNet	采用 P2T 作为编码器,引入 PFFM、MPPM	提取不同尺度的上下文特征,提高息肉分割的性能	对极小的息肉区域以及高度模糊的边缘表现不佳	Kvasir-SEG	0.920 0

3 CNN在结直肠息肉分类中的应用

结直肠息肉分类算法可以在医学图像中识别和分类结直肠息肉,这有助于早期发现患者是否存在结直肠息肉以及辨别息肉种类,从而及时采取预防和治疗措施。早期的结直肠息肉分类算法是通过人工定义某些组织特征来进行的,这种方法对医生来说不仅耗时繁琐,而且分类准确率也无法得到保障。随着卷积神经网络性能的提高,深度学习在结直肠息肉分类中得到了更广泛的应用,与传统的手工特征提取相比,CNN可以有效地提取抽象的和更高层次的特征,更准确地进行分类。

在早期算法中,研究人员主要使用CNN来提取图像的特征,然后将这些特征传入机器学习算法,来进行最终的分类任务。例如Okamoto等人^[63]提出了一种分类模型,使用预训练的AlexNet^[64]网络作为特征提取器,然后使用SVM进行分类。这种方法虽然可以实现息肉分类,但是因为不能共享梯度,所以分类性能并不出众,限制了模型在临床实践中的广泛使用。因此现阶段研究人员都是单独使用CNN来进行结直肠息肉的分类,使模型具有更高的分类准确性和更快的处理时间。

Jin等人^[65]提出了一种用于结直肠息肉分类的CNN模型,该模型可以对小尺寸的腺瘤性和增生性息肉进行分类。他们使用参数共享和数据增强方法来提升模型性能,并使用包含2 450张息肉图像的私人数据集进行训练和测试,分类准确率达到86.7%。这项研究的主要贡献在于第一次使用了相同的测试集对CNN模型和医生进行了比较,而且对医生使用AI进行辅助诊断后准确率是否提高也进行了测试,结果更加充分表明了AI辅助诊断的可用性。但是该研究的不足在于训练集中不包括无柄锯齿状病变(sessile serrated lesion, SSL),只对腺瘤性和增生性息肉进行分类,无法充分满足临床使用的要求。

针对这种不足,Zachariah等人^[66]设计了一种基于Inception-ResNet-v2算法的深度学习模型,对小尺寸的腺瘤性和锯齿状息肉(增生性和无柄锯齿状息肉)进行分类。他们使用了6 223张图片(包括白光图像(white light imaging, WLI)和窄带图像(narrow band imaging, NBI))进行训练和测试,实验结果表明,该模型在区分腺瘤和锯齿状息肉的总体准确率为94%,其中WLI和NBI的准确率分别为92%和94%。出现这种情况的原因可能是使用WLI的光学诊断通常比

NBI检查的信息量少。Patino-Barrientos等人^[67]提出了一种基于VGG-16架构的深度学习模型,该模型在特征提取模块增加密集连接,用于对结肠息肉进行良恶性分类。他们根据Kudo分类方案^[68],如图5所示,因前四类含有癌变的风险很低,所以将其归为良性,将第五类归为恶性。并且使用微调来提高模型的性能,结果表明深度学习模型可以有效地用于癌症的早期诊断,分类准确率达到83%。该研究的不足在于分类速度较慢,还无法满足实时诊断的要求,并且由于数据集缺乏,导致模型对于相似性息肉的分类精度不够。

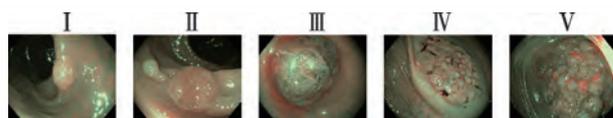


图5 Kudo息肉分类图

Fig.5 Polyp classification chart of Kudo

上述研究都是针对二分类,但是在实际应用中,简单的二分类无法满足医生的需求,随着研究深入,复杂的多分类模型开始出现。Liu等人^[69]提出了一个名为Polyp-DedNet的卷积神经网络,用于结直肠息肉的四类和六类分类任务。他们根据世界卫生组织(World Health Organization, WHO)和WASP(Workgroup Serrated Polyps and Polyposis)工作组的息肉分类标准,将息肉分为四类:常规腺瘤、增生性息肉、无柄锯齿状腺瘤/息肉(sessile serrated adenoma/polyp, SSA/P)和正常,其中常规腺瘤可进一步分为管状腺瘤、绒毛状腺瘤和绒毛管状腺瘤三个亚类。该模型使用ResNet-50作为特征提取网络,其中主要由两个残差模块组成,并在残差模块的不同阶段加入扩张卷积以及高效通道注意力(efficient channel attention, ECA)模块,以此来保留更多的高维空间信息和进一步提高分类性能,并且为了消除扩张卷积的影响,使用了传统的卷积层代替最大池化层。结果表明,所提出的Polyp-DedNet对结直肠息肉的四类和六类分类的平均准确率分别为89.91%和85.13%。Sharma等人^[70]提出了一种独特的网络LPNet,用于对结直肠息肉进行分类。作者利用小波池化替换了普通池化层,减少了在池化过程中的数据丢失,在感受野大小和计算效率之间实现更好的权衡。该模型将息肉分为八个类别,准确率达到93.55%,并且因为参数量少,所以对设备要求较低。虽然现在基于卷积神经

网络的算法被广泛用于结直肠息肉分类,然而现有的方法并不包含任何对超参数的设置信息,但是这对模型优化至关重要。因此 Younas 等人^[71]提出了一种建立优化的超参数的方法,通过建立优化的超参数,将单个弱学习者的力量结合起来,形成一个加权的集合模型,以实现准确的分类模型。该研究评估和比较了六种 CNN 模型,来对增生性息肉、腺瘤和锯齿状腺瘤进行分类,结果表明,提出的方法获得更好的评价指标结果。但是由于数据集的不足,分类性能还有待进一步提高。

在结直肠息肉分类问题上,研究人员越来越倾向于采用多个 CNN 模型的集成方法,以提高分类的准确率。Rahman 等人^[72]提出了一种用于提高息肉分类性能的堆叠集成方法,网络结构如图 6。该方法由三个卷积神经网络组成: Xception、ResNet-101 和 VGG-19。这些网络经过微调,然后用 Softmax 函数来计算每个网络的可能结果,最终的结果衡量了三个网络的值,因此拥有更高的准确率。所提出的模型对增生型、锯齿型、腺瘤型息肉和非息肉的分类准确性达到 98.53%。结果表明,所提出的方法比单一的网络表现更好,集合模型不仅弥补了单一模型的不足,提高了综合性能,而且还大大减少了处理时间,并产生了几乎实时的息肉分类。因此,为了更好地决策,以后的研究方向应该实验更多的架构。

为了加强计算机辅助诊断在临床实践中的集成,研究人员开始研究能够执行结直肠息肉检测和分类的完整工作流程模型。Ozawa 等人^[73]使用 SSD 架构来检测和分类结直肠息肉,它可以将息肉分类为腺瘤、增生性息肉、无柄锯齿状腺瘤/息肉、癌性和其他息肉。他们使用了一个私人数据集,16 418 张图片被用于训练 CNN 算法,训练后的 CNN 模型灵敏度

为 92%, 阳性预测值 (positive predictive value, PPV) 为 86%, 在正确检测出的息肉中,分类准确率为 83%。由于数据集分布不均匀,该模型对于腺瘤的识别能力并不出众,准确率有待提升,并且模型存在过度拟合和梯度消失问题。但是他们的方法表明,CNN 在结直肠息肉检测和分类的表现是成功的。Tanwar 等人^[74]针对上述模型进行了改进,他们首先使用引导图像过滤器和动态直方图均衡化方法来过滤和增强结肠镜图像,然后使用 SSD 来检测和分类结直肠息肉。值得注意的是,他们使用带有 dropout 的全连接层来对息肉进行分类,该方法有利于防止模型过拟合,提高了模型的泛化能力。实验结果表明,所提出的模型可以从结肠镜图像中检测出息肉并进行分类,准确率达到 92%。但是该模型也存在同样的问题,对于腺瘤性息肉和增生性息肉的分类效果欠佳。为了解决这个问题,杨昆等人^[75]提出了一种基于 Faster R-CNN 的改进模型,该模型使用 ResNeSt^[76]作为特征提取网络,实现了不同通道之间的信息交互,并且引入边缘感知特征的边界定位 (side-aware boundary localization, SABL) 模块,对息肉的位置信息进行更精准的预测。不仅如此,作者还将非极大抑制 (non-maximum suppression, NMS) 替换为 Soft-NMS,对相邻息肉的信息进行最大程度的保留。通过改进,模型对于腺瘤性息肉和增生性息肉的分类准确率大幅提升,平均精度达到 86.8%。

综上所述,在结直肠息肉分类领域,基于 CNN 的模型因为可以进行端到端训练,非常受到研究人员的青睐,其中,多分类模型比二分类模型更满足临床使用的需求。基于多网络的集成模型因为高准确率且时间更快,应该是未来分类模型的研究方向。基于多分类的结直肠息肉检测和分类模型,因为其完

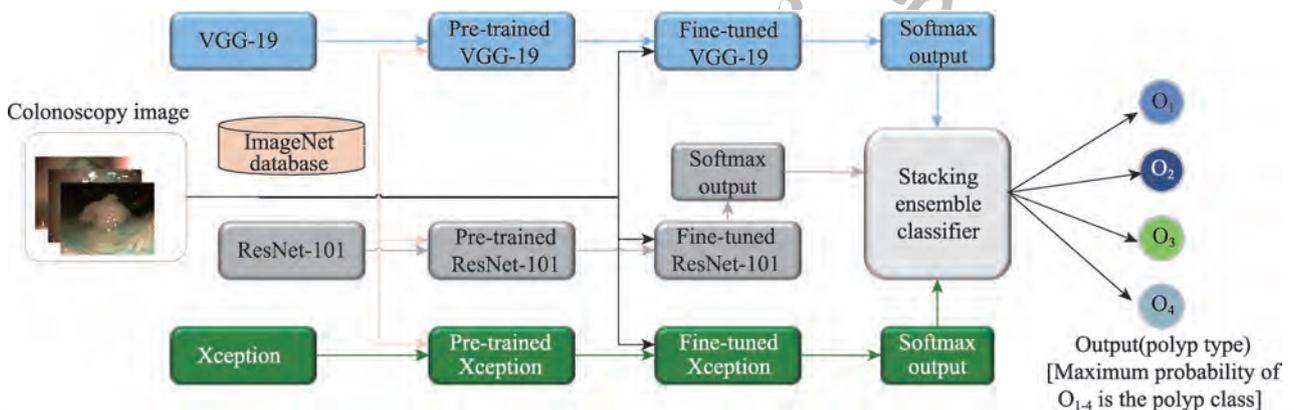


图 6 网络堆叠集成技术结构图

Fig.6 Architecture diagram of network stack ensemble technology

整的流程,更能满足临床诊断,所以也应该是未来的研究方向。并且通过研究,在特征提取模块使用密集连接以及注意力机制等方法,有利于提高模型提取特征的能力,提升分类性能。但是用于结直肠息肉分类的数据集数量太少且质量残次不齐等问题仍亟需解决。表5为不同结直肠息肉分类算法的性能比较。表6为结直肠息肉辅助诊断中的CNN架构总结。

4 模型可解释性

目前,基于卷积神经网络(CNN)的模型在结直肠息肉的诊断方面表现出先进的性能,但是这种基于CNN的模型不仅要保证高精度,还需要对预测的结果提供解释,以便在医疗背景下进行使用。前面提到的结直肠息肉诊断模型大多侧重于提高最终结果的准确性,尽管性能提高的好处显而易见,但基于深度学习的模型难以提供对预测结果的解释。这种对深度模型的基本机制缺乏直观的理解,导致基于深度学习的模型经常被称为“黑盒子”。

最近的多项研究提出了一些方法,在一定程度上解决了模型透明度不足的问题。Wickstrøm等人^[77]

利用导向反向传播(guided-backpropagation)技术生成显著图,为结直肠息肉分割模型的预测提供可视化解释,该方法主要利用反向传播中梯度的大小来判断输入图像中像素的重要程度,通过将梯度更加集中在激活的区域,从而更好地反映模型对输入特征的敏感性。虽然它能产生高分辨率的显著图,但是不能区分类别。针对分类模型,Song等人^[78]使用类激活映射方法(class activation mapping, CAM)对分类模型进行可视化解释,该方法通过将网络中的特征图和权重结合起来,对输入图像生成热力图。如图7所示,数字代表每个类别的预测的概率,颜色越红的地方表明特征越有效,对网络预测结果越重要。虽然这种方法可以对模型的工作原理进行一定的解释,但是它需要将最后的全连接层改为全局平均池化层并重新训练模型,导致在实际应用中非常不方便。后来,基于上述方法所改进的梯度加权类激活映射方法(Grad-CAM)被提出,在一定程度上克服了以上缺陷。Jin等人^[65]使用Grad-CAM的方法来解释卷积神经网络的判断原因,该方法通过梯度信息来引导特征图的加权过程,将每个通道的权重乘以对应的卷积层的输出,然后将这些加权的特征图

表5 不同结直肠息肉分类算法的性能比较

Table 5 Performance comparison of different colorectal polyp classification algorithms

参考文献	主要思想	局限性	准确率/%	备注
Okamoto等人 ^[63]	使用预训练的 AlexNet 网络提取特征,然后使用 SVM 进行分类	不能实现端到端的分类	>90.00	二分类
Jin等人 ^[65]	通过参数共享进行有效的神经架构搜索,第一次使用了相同的测试集对 CNN 模型和医生进行了比较	训练集中不包括无柄锯齿状息肉	86.70	二分类
Zachariah等人 ^[66]	基于 Inception-ResNet-v2 架构,并使用迁移学习,对小尺寸的腺瘤性和锯齿状息肉进行分类	未使用外部数据集进行验证	94.00	二分类
Patino-Barrientos等人 ^[67]	基于 VGG-16 架构,并在特征提取模块增加密集连接	分类速度较慢,无法满足实时诊断的要求	83.00	二分类
Liu等人 ^[69]	提出 Polyp-DedNet 网络,在 ResNet-50 基础上,结合扩张卷积和 ECA 模块	分类准确率有待进一步提升	四类:89.91 六类:85.13	多分类
Sharma等人 ^[70]	提出了 LPNet 网络,用小波池化代替了普通池化层	由于数据集的缺乏,模型的泛化性能还有待提高	93.55	多分类
Rahman等人 ^[72]	采用结合 Xception、ResNet-101 和 VGG-19 三种网络的集成模型	未使用其他数据集进行验证	98.53	多分类
Ozawa等人 ^[73]	提出一种基于 SSD 架构的结直肠息肉检测和分类模型,使用 Caffe 深度学习框架来训练和验证模型	模型存在过度拟合和梯度消失问题	83.00	检测、分类
Tanwar等人 ^[74]	基于 SSD 架构,使用图像预处理,并在全连接层中加入 dropout	对于腺瘤性息肉和增生性息肉的分类效果欠佳	92.00	检测、分类
杨昆等人 ^[75]	在 Faster R-CNN 模型中引入 SABL 模块和 Soft-NMS,并且使用 ResNeSt 作为主干网络	未使用不同数据集证明其泛化性	86.80	检测、分类

表6 结直肠息肉辅助诊断中的CNN架构总结

Table 6 Summary of CNN architecture in assisted diagnosis of colorectal polyps

任务	模型	主要思想	优点	缺点
检测	R-CNN	首先生成候选区域,然后通过CNN提取特征,最后使用分类器进行检测	对不同尺寸和形状的目标有较好的适应性	计算速度慢,不适用于实时检测
	Faster R-CNN	引入RPN网络,将候选框提取和特征提取合并为一个端到端的网络	相对于R-CNN有更快的检测速度	对小目标检测效果一般
	Mask R-CNN	在Faster R-CNN的基础上添加了目标分割分支,实现实例分割	同时实现目标检测和语义分割	计算量大,速度较慢
	SSD	直接在图像上预测目标的边界框和类别,使用多个预设的固定大小的卷积核对图像进行扫描	单次前向传播即可完成目标检测,适用于实时目标检测	相对于Faster R-CNN,检测精度略低
	YOLO	将目标检测任务看作回归问题,通过卷积神经网络直接预测边界框和类别	高速实时目标检测,适用于多种目标检测场景	对小目标检测的效果相对有限
分割	FCN	将全连接层替换为卷积层,通过上采样还原特征图大小,实现像素级别的预测	可以处理任意尺寸的输入图像,实现端到端训练	没有考虑上下文信息,可能导致分割不准确
	U-Net	使用U形状结构,包括编码器和解码器,利用跳跃连接保留更多信息	分割精度高,适用于小样本学习	对于复杂背景和噪声的处理能力有待提高
	SegNet	使用编码器-解码器结构,通过最大池化的索引进行解码	结构相对简单,易于理解和训练	对大尺寸图像可能存在计算效率问题
	DeepLabV3+	使用空洞卷积提高感受野,使用ASPP处理不同尺度的信息	能够处理不同尺度的目标,具有较高的精度	计算量较大,需要更多的计算资源
	AlexNet	加深CNN的深度,引入ReLU激活函数以及Dropout正则化	结构相对简单,易于理解和实现	卷积核尺寸过大
分类	VGG-16	通过连续的小型卷积核(如3×3)代替大型卷积核	增强了非线性表达能力	使用了计算成本高的全连接层
	VGG-19	在VGG-16的基础上进一步增加了网络深度	可以学习到更复杂、抽象的特征	计算量大
	Xception	引入了深度可分离卷积	使学习更加高效,并提高了性能	计算成本较高
	ResNet-50	引入残差网络结构以及恒等映射	解决了深度神经网络中的梯度消失问题,增强信息传递	结构复杂,对于小数据集可能过拟合
	ResNeSt	在ResNet基础上引入Split-Attention机制	利用注意力机制提高网络对特征的关注度	参数量大,结构复杂

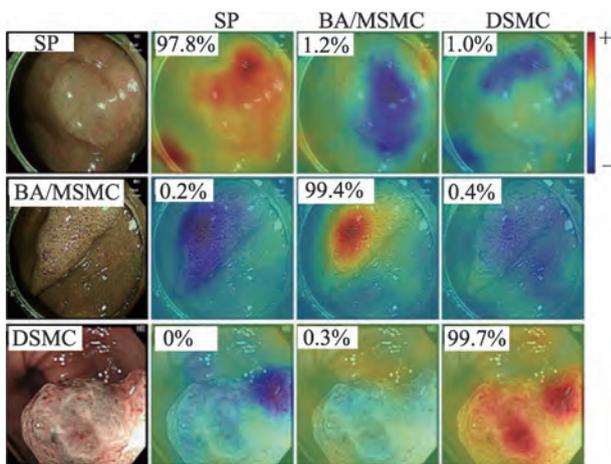


图7 可视化的类激活图

Fig.7 Visualization of class activation maps

求和,进而将模型预测的类别重要性映射回输入图像,实现对模型决策的解释。Krenzer等人^[79]针对生成的热图中存在大量噪声的问题,在使用Grad-CAM方法时加入Eigen smooth,该方法使用主成分降低噪声,可以有效减少热力图的噪声,更加突出对象。他们使用该方法生成热图,来查看网络对息肉进行分类的关键区域。虽然该方法可以对网络进行一定的解释,但是它缺乏突出显示细微细节的能力。对此,Wang等人^[80]使用沙普利加和解释(Shapley additive explanations, SHAP)来直观地看到模型的内部运行机制,该方法利用博弈论中的Shapley值,量化了每个特征对模型所做预测的贡献。SHAP的最大优势是可以反映出每一个样本中每个特征的影响力,而且

还表现出影响的正负性。然而,随着网络结构的扩大,精确计算沙普利值的时间急剧增加。因此,需要一种快速和近似的算法来计算深度学习模型的沙普利值。针对此问题,可以使用沙普利值作为层间表示,并结合 Shapley 模块进行计算压缩。上述研究主要关注推理后的解释,这不仅忽略了模型推理过程,而且增加了模型的计算复杂度。为此,Wang 等人^[81]通过引入 Shapley 解释网络(SHAPNet)解决了上述问题,该方法使每个特征的 Shapley 值计算只取决于模型中的部分输入,减少了计算量,并且作者尝试在模型的不同位置添加 SHAPNet,并将结果进行可视化。结果表明,在特征提取部分添加 SHAPNet 可以解释模型对更细粒度像素的推理过程,有助于提高模型的可解释性。

在其他医学图像辅助诊断领域,研究人员还发现了一些针对深度学习模型的可解释方法。例如,Böhle 等人^[82]针对阿尔茨海默病分类问题,采用分层相关性传播(layer-wise relevance propagation, LRP)方法生成脑部热力图对模型进行解释,该方法根据模型的结构和权重,将预测结果的重要性逐层传播回输入特征,以确定每个输入特征对于预测结果的贡献。Wollek 等人^[83]利用 Transformer 多模态可解释性(transformer multi-modal explainability, TMME)创建了基于注意力的显著图,用于解释气胸分类模型,

该方法通过计算注意力权重对输入图像进行量化。研究人员还将此方法与 Grad-CAM 方法进行对比,结果表明,基于注意力的显著图比梯度加权类激活图具有更高的可解释性。这些方法目前在结直肠息肉诊断领域应用不多,可以为研究人员在模型可解释性上提供一定启发。

上述各种方法都旨在提供对深度学习模型预测结果的解释和可视化,帮助理解模型的决策依据和关注的区域,它们生成的热力图或显著图具有直观性,能够直观地显示模型对输入的关注程度或特征的重要性,但是这些方法也存在计算复杂度高、解释不够详细等问题亟需解决。表 7 为深度学习模型的可解释性方法总结。

5 总结与展望

本文对卷积神经网络在结直肠息肉目标检测、分割和分类等方面的应用分别进行阐述,在结直肠息肉目标检测领域,研究人员对 Faster R-CNN、SSD 和 YOLO 等常用算法进行改进,实现对结直肠息肉的定位和边界框预测;在结直肠息肉分割领域,本文对基于 U-Net、SegNet 等分割算法进行总结,这些算法可以实现精确的结直肠息肉分割,提供更详细的形态和位置信息;在结直肠息肉分类领域,对现阶段二分类及多分类模型进行分析,并总结了不同模型

表 7 深度学习模型的可解释性方法总结

Table 7 Summary of interpretability methods for deep learning models

方法	主要思想	优点	缺点
导向反向传播	通过分析模型中神经元的激活情况,将激活信号从输出层传播回输入层,从而计算出输入特征的梯度,进而理解模型对这些特征的敏感性	实现简单,能够可视化特征重要性,有助于直观理解模型	仅提供相对特征重要性排名,忽略特征相互作用,无法完全解释模型内部机制
分层相关性传播	根据模型的结构和权重,将预测结果的重要性逐层传播回输入特征,以确定每个输入特征对于预测结果的贡献	通过结合模型结构和权重,在一定程度上准确地解释了深度神经网络的决策	该方法的有效性依赖于选取适当的传播规则;在非常复杂的模型中,解释可能不够详细
类激活映射	通过分析网络中的特征图和权重,将每个类别的激活信号映射回输入图像,以可视化模型对不同类别的关注区域	能够直观地将模型对于不同类别的关注区域可视化在输入图像上,帮助解释模型的决策	需要改变网络结构,无法用于预先训练的网络
梯度加权类激活映射	通过梯度信息来引导特征图的加权过程,将模型预测的类别重要性映射回输入图像,以实现模型决策的解释	解决了类激活映射需要改变网络结构的问题,提供了可视化的类别关注区域	缺乏显示细微细节的能力,难以解释多实例以及定位不够精确
SHAP	基于博弈论中的 Shapley 值概念,旨在为每个特征分配一个重要性分数,揭示该特征对于模型预测的贡献	能够为每个特征分配合理的贡献权重,保证了特征重要性的全局一致性,适用于各种深度学习模型	计算复杂度随着特征数量的增加而指数增长,在局部决策上的解释仍然不够详细
注意力机制	利用模型内部的注意力权重来解释模型对输入数据的关注程度,从而为人们提供关于模型预测的直观解释	通过可视化注意力权重,能够直观地展示模型对输入数据的关注区域	注意力权重并不能准确地量化特征的重要性;无法提供全局特征重要性的解释

的优缺点,最后对深度学习模型的可解释性进行论述。尽管卷积神经网络在结直肠息肉诊断中取得了一定的成功,但仍然存在一些问题需要解决:(1)卷积神经网络需要大量的数据进行训练,然而结直肠息肉数据集相对稀缺,这会限制模型的性能。(2)在结直肠息肉的数据集中,某些类别的数据可能远多于其他类别,导致数据集的不平衡。(3)卷积神经网络等深度学习模型通常被视为“黑盒子”,其内部决策过程难以解释和理解。(4)模型在不同数据集和不同应用场景中的性能表现可能存在差异,因此,算法的泛化能力需要进一步提高。(5)将结直肠息肉计算机辅助诊断系统应用于实际临床中仍面临一些挑战。

针对以上问题,对结直肠息肉辅助诊断未来的研究方向做出以下展望:(1)利用一系列数据增强技术扩充数据集,或者可以使用预训练的模型,将其迁移到结直肠息肉辅助诊断任务上。其次,可以使用生成对抗网络(generative adversarial network, GAN)合成更多的医学图像,来解决数据不足的问题。(2)针对数据不平衡问题,可以考虑使用数据增强技术,增加少数类别样本数量。其次,可以对数据集进行重采样,使用过采样(增加少数类样本)和欠采样(减少多数类样本)方法平衡数据。或者,可以在损失函数中为不同类别赋予不同的权重,以便让模型更关注少数类别。(3)通过使用可视化技术、模型蒸馏等方法解释模型的决策过程,或者可以开发将逻辑规则与卷积神经网络进行结合的新模型,把医生的经验加入模型的决策过程中,使模型的决策过程更加透明和可解释。(4)为了提高算法的泛化能力,首先可以通过增加训练数据、使用迁移学习等方法,使模型学习到更多的一般化特征。其次可以使用一系列正则化技术,防止网络对特定的输入过于依赖。另外还可以采用集成学习方法,结合多个模型的预测结果,减少过拟合,或者通过加入注意力机制,使模型能够更专注于输入的关键部分,提高模型泛化能力。(5)在临床应用中,模型不仅要为决策过程提供可视觉解释,对预测结果的不确定性估计也很重要,因此应该开发新的模型,来对预测结果的不确定性进行评估。其次,进行更多的临床实验和验证,将结直肠息肉计算机辅助诊断系统与实际临床工作流程进行集成和评估。

综上所述,卷积神经网络在结直肠息肉辅助诊

断中呈现出巨大的潜力和价值,通过深度学习技术,可以从医学图像中提取复杂的特征,实现对结直肠息肉的自动检测和诊断。目前尽管已经取得了令人鼓舞的进展,但在临床实践中仍然需要面对一系列挑战,而解决这些挑战就是未来的研究方向。

参考文献:

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA: A Cancer Journal for Clinicians*, 2021, 71(3): 209-249.
- [2] 王彩霞. 结直肠息肉的内镜诊治进展[J]. *中国城乡企业卫生*, 2023, 38(6): 21-23.
WANG C X. Progress in endoscopic diagnosis and treatment of colorectal polyps[J]. *Chinese Journal of Urban and Rural Enterprise*, 2023, 38(6): 21-23.
- [3] ZHAO S, WANG S, PAN P, et al. Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: a systematic review and meta-analysis[J]. *Gastroenterology*, 2019, 156(6): 1661-1674.
- [4] 张松兰. 基于卷积神经网络的图像识别综述[J]. *西安航空学院学报*, 2023, 41(1): 74-81.
ZHANG S L. A review of image recognition based on convolutional neural networks[J]. *Journal of Xi'an Aeronautical Institute*, 2023, 41(1): 74-81.
- [5] SANCHEZ-PERALTA L F, BOTE-CURIEL L, PICON A, et al. Deep learning to find colorectal polyps in colonoscopy: a systematic literature review[J]. *Artificial Intelligence in Medicine*, 2020, 108: 101923.
- [6] 孙福艳, 王琼, 吕宗旺, 等. 深度学习在结肠息肉分割中的应用综述[J]. *计算机工程与应用*, 2023, 59(23): 15-27.
SUN F Y, WANG Q, LYU Z W, et al. Review of the application of deep learning in colon polyp segmentation[J]. *Computer Engineering and Applications*, 2023, 59(23): 15-27.
- [7] AZER S A. Challenges facing the detection of colonic polyps: what can deep learning do?[J]. *Medicina*, 2019, 55(8): 473.
- [8] 王啸, 黄鉴, 吉祥, 等. 人工智能在结肠息肉检测与分类中的应用[J]. *南方医科大学学报*, 2021, 41(2): 310-313.
WANG X, HUANG J, JI X, et al. Application of artificial intelligence for detection and classification of colon polyps [J]. *Journal of Southern Medical University*, 2021, 41(2): 310-313.
- [9] 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.
- [10] BERNAL J, SÁNCHEZ J, VILARINO F. Towards automatic polyp detection with a polyp appearance model[J]. Pattern Recognition, 2012, 45(9): 3166-3182.
- [11] JHA D, SMEDSRUD P H, RIEGLER M A, et al. Kvasir-seg: a segmented polyp dataset[C]//Proceedings of the 26th International Conference on Multimedia Modeling, Daejeon, Jan 5-8, 2020. Cham: Springer, 2020: 451-462.
- [12] VÁZQUEZ D, BERNAL J, SÁNCHEZ F J, et al. A benchmark for endoluminal scene segmentation of colonoscopy images[J]. Journal of Healthcare Engineering, 2017: 4037190.
- [13] ANGERMANN Q, BERNAL J, SÁNCHEZ-MONTES C, et al. Towards real-time polyp detection in colonoscopy videos: adapting still frame-based methodologies for video sequences analysis[C]//Proceedings of the 4th International Workshop on Computer Assisted and Robotic Endoscopy and Clinical Image-Based Procedures, and 6th International Workshop on Clinical Image-Based Procedures, Held in Conjunction with International Conference on Medical Image Computing and Computer-Assisted Intervention 2017, Québec City, Sep 14, 2017. Cham: Springer, 2017: 29-41.
- [14] TAJBAKSH N, GURUDU S R, LIANG J. Automated polyp detection in colonoscopy videos using shape and context information[J]. IEEE Transactions on Medical Imaging, 2015, 35(2): 630-644.
- [15] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2014: 580-587.
- [16] GIRSHICK R. Fast R-CNN[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Washington: IEEE Computer Society, 2015: 1440-1448.
- [17] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Advances in Neural Information Processing Systems 28, Montreal, Dec 7-12, 2015: 91-99.
- [18] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2016: 779-788.
- [19] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of the 14th European Conference on Computer Vision, Amsterdam, Oct 11-14, 2016. Cham: Springer, 2016: 21-37.
- [20] ZHU R, ZHANG R, XUE D. Lesion detection of endoscopy images based on convolutional neural network features[C]//Proceedings of the 2015 8th International Congress on Image and Signal Processing. Piscataway: IEEE, 2015: 372-376.
- [21] SORNAPUDI S, MENG F, YI S. Region-based automated localization of colonoscopy and wireless capsule endoscopy polyps[J]. Applied Sciences, 2019, 9(12): 2404.
- [22] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2016: 770-778.
- [23] LIU X, LI Y, YAO J, et al. Classification of polyps and adenomas using deep learning model in screening colonoscopy [C]//Proceedings of the 2019 8th International Symposium on Next Generation Electronics. Piscataway: IEEE, 2019: 1-3.
- [24] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, Inception-ResNet and the impact of residual connections on learning[C]//Proceedings of the 2017 AAAI Conference on Artificial Intelligence. Menlo Park: AAAI, 2017: 4278-4284.
- [25] CHEN B L, WAN J J, CHEN T Y, et al. A self-attention based faster R-CNN for polyp detection from colonoscopy images[J]. Biomedical Signal Processing and Control, 2021, 70: 103019.
- [26] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv:1409.1556, 2014.
- [27] WITTENBERG T, ZOBEL P, RATHKE M, et al. Computer aided detection of polyps in whitelight-colonoscopy images using deep neural networks[J]. Current Directions in Biomedical Engineering, 2019, 5(1): 231-234.
- [28] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2017: 7263-7271.
- [29] LIU M, JIANG J, WANG Z. Colonic polyp detection in endoscopic videos with single shot detection based deep convolutional neural network[J]. IEEE Access, 2019, 7: 75058-75066.
- [30] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2016: 2818-2826.

- [31] ZHENG Y, ZHANG R, YU R, et al. Localisation of colorectal polyps by convolutional neural network features learnt from white light and narrow band endoscopic images of multiple databases[C]//Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Piscataway: IEEE, 2018: 4142-4145.
- [32] WAN J, CHEN B, YU Y. Polyp detection from colorectum images by using attentive YOLOv5[J]. *Diagnostics*, 2021, 11(12): 2264.
- [33] 司丙奇, 王志武, 姜萍萍, 等. 基于改进 YOLOv5 网络的内窥镜息肉检测[J]. *中国新通信*, 2022, 24(11): 123-128.
- SI B Q, WANG Z W, JIANG P P, et al. Endoscope polyp detection based on improved YOLOv5 network[J]. *China New Telecommunications*, 2022, 24(11): 123-128.
- [34] JOCHER G, STOKEN A, CHAURASIA A, et al. Ultralytics/YOLOv5: v6.0-YOLOv5n Nano models, roboflow integration, TensorFlow export, OpenCV DNN support[Z]. Zenodo, 2021.
- [35] MA Y, LI Y, YAO J, et al. Polyp location in colonoscopy based on deep learning[C]//Proceedings of the 2019 8th International Symposium on Next Generation Electronics. Piscataway: IEEE, 2019: 1-3.
- [36] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//Proceedings of the 32nd International Conference on Machine Learning, Lille, Jul 6-11, 2015: 448-456.
- [37] 薛钦原, 胡珊珊, 胡新军, 等. 改进 YOLOv7 的结直肠息肉检测算法[J/OL]. *计算机工程与应用*(2023-11-28)[2023-12-02]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20231127.1356.012.html>.
- XUE Q Y, HU S S, HU X J, et al. Improved YOLOv7 algorithm for colorectal polyp detection[J/OL]. *Computer Engineering and Applications*(2023-11-28)[2023-12-02]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20231127.1356.012.html>.
- [38] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 7464-7475.
- [39] MO X, TAO K, WANG Q, et al. An efficient approach for polyps detection in endoscopic videos based on faster R-CNN[C]//Proceedings of the 2018 24th International Conference on Pattern Recognition. Piscataway: IEEE, 2018: 3929-3934.
- [40] MA C, JIANG H, MA L, et al. A real-time polyp detection framework for colonoscopy video[C]//Proceedings of the 2022 Chinese Conference on Pattern Recognition and Computer Vision. Cham: Springer, 2022: 267-278.
- [41] NADIMI E S, BUIJS M M, HERP J, et al. Application of deep learning for autonomous detection and localization of colorectal polyps in wireless colon capsule endoscopy[J]. *Computers & Electrical Engineering*, 2020, 81: 106531.
- [42] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]//Proceedings of the 13th European Conference on Computer Vision, Zurich, Sep 6-12, 2014. Cham: Springer, 2014: 818-833.
- [43] ALBUQUERQUE C, HENRIQUES R, CASTELLI M. A stacking-based artificial intelligence framework for an effective detection and localization of colon polyps[J]. *Scientific Reports*, 2022, 12(1): 17678.
- [44] LONG J, SHEHMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2015: 3431-3440.
- [45] YU J, PAN H, YIN Q, et al. Fully convolutional densenets for polyp segmentation in colonoscopy[C]//Proceedings of the 2019 IEEE 35th International Conference on Data Engineering Workshops. Piscataway: IEEE, 2019: 306-311.
- [46] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2017: 4700-4708.
- [47] 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, Munich, Oct 5-9, 2015. Cham: Springer, 2015: 234-241.
- [48] SUN X, ZHANG P, WANG D, et al. Colorectal polyp segmentation by U-Net with dilation convolution[C]//Proceedings of the 2019 18th IEEE International Conference on Machine Learning and Applications. Piscataway: IEEE, 2019: 851-858.
- [49] MOHAPATRA S, PATI G K, MISHRA M, et al. UPolySeg: a U-Net-based polyp segmentation network using colonoscopy images[J]. *Gastroenterology Insights*, 2022, 13(3): 264-274.
- [50] DUMITRU R G, PETELEAZA D, CRACIUN C. Using DUCK-Net for polyp image segmentation[J]. *Scientific Reports*, 2023, 13(1): 9803.
- [51] 徐昌佳, 易见兵, 曹锋, 等. 采用 DoubleUNet 网络的结直

- 肠息肉分割算法[J]. 光学精密工程, 2022, 30(8): 970-983.
- XU C J, YI J B, CAO F, et al. Colorectal polyp segmentation algorithm using DoubleUNet network[J]. Optics and Precision Engineering, 2022, 30(8): 970-983.
- [52] 贾立新, 胡奕标, 金燕, 等. 融合多种注意力机制的结直肠息肉分割神经网络[J]. 计算机辅助设计与图形学学报, 2023, 35(3): 463-473.
- JIA L X, HU Y B, JIN Y, et al. Polyp segmentation network combined with multi-attention mechanism[J]. Journal of Computer-Aided Design & Computer Graphics, 2023, 35(3): 463-473.
- [53] JHA D, RIEGLER M A, JOHANSEN D, et al. DoubleUNet: a deep convolutional neural network for medical image segmentation[C]//Proceedings of the 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems. Piscataway: IEEE, 2020: 558-564.
- [54] JADON S. A survey of loss functions for semantic segmentation[C]//Proceedings of the 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology. Piscataway: IEEE, 2020: 1-7.
- [55] TRAN S T, NGUYEN M H, DANG H P, et al. Automatic polyp segmentation using modified recurrent residual Unet network[J]. IEEE Access, 2022, 10: 65951-65961.
- [56] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [57] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the 15th European Conference on Computer Vision. Cham: Springer, 2018: 801-818.
- [58] EU C Y, TANG T B, LU C K. Automatic polyp segmentation in colonoscopy images using single network model: SegNet[C]//Proceedings of the 2020 International Conference on Artificial Intelligence for Smart Community, Ho Chi Minh City, Oct 14-15, 2020. Singapore: Springer, 2022: 717-723.
- [59] 朱世祺, 徐昶, 周鑫, 等. 基于DeepLab V3+深度神经网络的结直肠息肉内镜图像分割[J]. 中国医学物理学杂志, 2023, 40(8): 944-949.
- ZHU S Q, XU C, ZHOU X, et al. Colorectal polyp segmentation in endoscopic images using DeepLab V3+[J]. Chinese Journal of Medical Physics, 2023, 40(8): 944-949.
- [60] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2017: 1251-1258.
- [61] 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.
- 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.
- [62] HU K, CHEN W, SUN Y Z, et al. PPNet: pyramid pooling based network for polyp segmentation[J]. Computers in Biology and Medicine, 2023, 160: 107028.
- [63] OKAMOTO T, ODAGAWA M, KOIDE T, et al. Feature extraction of colorectal endoscopic images for computer-aided diagnosis with CNN[C]//Proceedings of the 2019 2nd International Symposium on Devices, Circuits and Systems. Piscataway: IEEE, 2019: 1-4.
- [64] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [65] JIN E H, LEE D, BAE J H, et al. Improved accuracy in optical diagnosis of colorectal polyps using convolutional neural networks with visual explanations[J]. Gastroenterology, 2020, 158(8): 2169-2179.
- [66] ZACHARIAH R, SAMARASENA J, LUBA D, et al. Prediction of polyp pathology using convolutional neural networks achieves “resect and discard” thresholds[J]. The American Journal of Gastroenterology, 2020, 115(1): 138.
- [67] PATINO-BARRIENTOS S, SIERRA-SOSA D, GARCIA-ZAPIRAIN B, et al. Kudo's classification for colon polyps assessment using a deep learning approach[J]. Applied Sciences, 2020, 10(2): 501.
- [68] KUDO S E, TAMURA S, NAKAJIMA T, et al. Diagnosis of colorectal tumorous lesions by magnifying endoscopy [J]. Gastrointestinal Endoscopy, 1996, 44(1): 8-14.
- [69] LIU S, LIU X, CHANG S, et al. Multi-classification of polyps in colonoscopy images based on an improved deep convolutional neural network[J]. Computers, Materials & Continua, 2023, 75(3): 5837-5852.
- [70] SHARMA P, DAS D, GAUTAM A, et al. LPNet: a light-weight CNN with discrete wavelet pooling strategies for colon polyps classification[J]. International Journal of Imaging Systems and Technology, 2023, 33(2): 495-510.
- [71] YOUNAS F, USMAN M, YAN W Q. A deep ensemble learning method for colorectal polyp classification with optimized network parameters[J]. Applied Intelligence, 2023, 53(2): 2410-2433.
- [72] RAHMAN M M, WADUD M A H, HASAN M M. Computerized classification of gastrointestinal polyps using stack-

- ing ensemble of convolutional neural network[J]. *Informat-ics in Medicine Unlocked*, 2021, 24: 100603.
- [73] OZAWA T, ISHIHARA S, FUJISHIRO M, et al. Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks[J]. *Therapeutic Advances in Gastroenterology*, 2020, 13: 1756284820910659.
- [74] TANWAR S, VIJAYALAKSHMI S, SABHARWAL M, et al. Detection and classification of colorectal polyp using deep learning[J/OL]. *BioMed Research International* [2023-08-06]. <https://doi.org/10.1155/2022/2805607>.
- [75] 杨昆, 原嘉成, 高聪, 等. 基于改进的 Faster R-CNN 的息肉目标检测和分类方法[J]. *河北大学学报(自然科学版)*, 2023, 43(1): 103-112.
YANG K, YUAN J C, GAO C, et al. Object detection and classification of polyps based on improved Faster R-CNN [J]. *Journal of Hebei University (Natural Science Edition)*, 2023, 43(1): 103-112.
- [76] ZHANG H, WU C, ZHANG Z, et al. ResNeSt: split-attention networks[C]//*Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 2022: 2736-2746.
- [77] WICKSTRØM K, KAMPFFMEYER M, JENSSEN R. Uncertainty and interpretability in convolutional neural networks for semantic segmentation of colorectal polyps[J]. *Medical Image Analysis*, 2020, 60: 101619.
- [78] SONG E M, PARK B, HA C A, et al. Endoscopic diagnosis and treatment planning for colorectal polyps using a deep-learning model[J]. *Scientific reports*, 2020, 10(1): 30.
- [79] KRENZER A, HEIL S, FITTING D, et al. Automated classification of polyps using deep learning architectures and few-shot learning[J]. *BMC Medical Imaging*, 2023, 23(1): 59.
- [80] WANG S, YIN Y, WANG D, et al. An interpretable deep neural network for colorectal polyp diagnosis under colonoscopy[J]. *Knowledge-Based Systems*, 2021, 234: 107568.
- [81] WANG D, WANG X, WANG S, et al. Explainable multitask Shapley explanation networks for real-time polyp diagnosis in videos[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(6): 7780-7789.
- [82] BÖHLE M, EITEL F, WEYGANDT M, et al. Layer-wise relevance propagation for explaining deep neural network decisions in MRI-based Alzheimer's disease classification [J]. *Frontiers in Aging Neuroscience*, 2019, 11: 194.
- [83] WOLLEK A, GRAF R, ČEČATKA S, et al. Attention-based saliency maps improve interpretability of pneumothorax classification[J]. *Radiology: Artificial Intelligence*, 2022, 5(2): e220187.



考文涛(1999—),男,山东潍坊人,硕士研究生,主要研究方向为医学图像处理。

KAO Wentao, born in 1999, M.S. candidate. His research interest is medical image processing.



李明(1978—),男,山东淄博人,硕士,教授,主要研究方向为中医药智能化。

LI Ming, born in 1978, M.S., professor. His research interest is the intelligence of traditional Chinese medicine.



马金刚(1980—),男,山东潍坊人,硕士,副教授,主要研究方向为中医药智能化。

MA Jingang, born in 1980, M.S., associate professor. His research interest is the intelligence of traditional Chinese medicine.