

分类号: TP391

密 级: 公开

学校代码: 10363

学 号: 2210330134



安徽工程大学

Anhui Polytechnic University

硕士学位论文

题 目

基于深度学习的内窥镜

图像智能识别系统设计

论 文 作 者

项诗雨

校 内 导 师

魏利胜(教授)

校 外 导 师

张平改

专业学位类别

电子信息

研究方向(领域)

人工智能与系统集成

论文提交日期: 2024 年 5 月 31 日

分类号：TP391

密 级：公开

单位代码：10363

学 号：2210330134

题 目 基于深度学习的内窥镜图像智能识别系统设计

题 目 Design of an Intelligent Recognition System for Endoscopic
Images Based on Deep Learning

学 生 姓 名： 项诗雨

校 内 导 师： 魏利胜教授

校 外 导 师： 张平改

专业学位类别： 电子信息

研究方向(领域)： 人工智能与系统集成

论文答辩日期： 2024年5月12日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名：项诗雨

日期：2024年5月30日

安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 ☐，在 _____ 年解密后适用本版权书。

本学位论文属于

不保密 ☒。

学位论文作者签名：项诗雨

日期：2024年5月30日

指导教师签名：魏利胜

日期：2024年5月30日

基于深度学习的内窥镜图像智能识别系统设计

摘 要

当今全球结直肠癌、胃癌等常见的消化道疾病发病率逐年增加,息肉、溃疡等与消化道癌变密切相关的疾病可以通过消化道检查来诊断。无线胶囊内窥镜(Wireless capsule endoscope, WCE)是一种有别于传统内窥镜的无痛、无创伤的检测手段,但是每次检查后需要医生高强度阅读大量图像,易出现漏检误检的情况。如何借助计算机诊断系统来提高诊断效率、降低误诊率显得极为迫切。为此,本文将深度学习策略用于 WCE 图像智能识别,结合 WCE 图像预处理、图像分类和图像语义分割三个方面设计了一个智能辅助诊断系统。具体研究内容如下:

1) 针对 WCE 图像中的光斑妨碍算法精确识别病灶的问题,通过融合 VS-UNet 与 Criminisi 算法移除内窥镜图像的光斑,设计了一种 WCE 图像预处理算法。该算法包括光斑分割与修复两个过程,在光斑分割阶段,以 UNet 模型为基本框架定位光斑区域,使用优化的 VGG16 作为主干网络,并且引入 SPPFCSPC 模块增加模型感受野;在光斑修复阶段,选用重建效果较好的 Criminisi 算法,并优化原始优先权函数。实验结果表明该算法提高了光斑分割的精度,并在一定程度上缓解了光斑修复结果的不合理性。

2) 针对 WCE 图像数据量少,类间差异小,类内差异大影响分类准确率的问题,研究了一种图像分类算法 CA-ConvNeXt。该算法结合图像增强和迁移学习思想微调卷积神经网络 ConvNeXt,在 ConvNeXt 网络中引入坐标注意力机制。同时,使用 Polyloss 损失函数优化模型的训练过程。实验结果表明改进的模型在小样本条件下

能实现正常、息肉和溃疡图像的精准分类。

3) 针对目前息肉分割模型较复杂、息肉分割准确率较低的问题, 探讨了一种轻量级息肉分割网络模型 Li-DeepLabV3+。该网络模型以 DeepLabV3+ 为基本框架, 使用优化的 MobileNetV2 网络作为主干网络以降低模型复杂度; 在加强特征提取部分, 使用改进的简易金字塔池化模块, 减少模型参数的同时提高模型训练效率; 为了增强特征表示, 使用改进的统一注意力融合模块融合浅层特征和深层特征, 该模块分别利用空间注意力和通道注意力产生的权值与特征图融合。实验结果表明, Li-DeepLabV3+ 模型兼顾了息肉分割精度和实时性, 同时具有一定的泛化能力。

4) 结合最优 WCE 图像分类算法和 WCE 图像息肉分割算法, 设计了一个 WCE 图像智能辅助诊断系统。该系统不仅提供高精度的识别结果, 还具备用户友好的控制界面。

综上所述, 本文将深度学习与医学图像处理相结合, 设计了一个 WCE 图像智能识别辅助诊断系统, 在减轻医生负担的同时提高诊断的效率和准确率, 并为 WCE 图像智能识别领域的深入探索提供了一定的理论支撑。

关键词: 深度学习; 图像修复; 图像分类; 语义分割; 智能辅助诊断系统; 无线胶囊内窥镜

Design of an Intelligent Recognition System for Endoscopic Images Based on Deep Learning

ABSTRACT

The incidence of common digestive diseases such as colorectal cancer and gastric cancer is increasing year by year globally today. Diseases such as polyps and ulcers, which are closely related to cancerous changes in the digestive tract, can be diagnosed through GI tests. Wireless capsule endoscope (WCE) is a painless and non-invasive detection method different from traditional endoscopy, but it requires doctors to read a large number of images with high intensity after each examination, which makes it easy to miss and misdiagnose. How to improve the diagnosis efficiency and reduce the misdiagnosis rate with the help of computer diagnosis system is extremely urgent. For this reason, this paper uses deep learning strategies for WCE image intelligent recognition, combining three aspects of WCE image preprocessing, image classification and image semantic segmentation to design an intelligent auxiliary diagnosis system. The specific research content is as follows:

1) Aiming at the problem of highlight in WCE images preventing the algorithm from accurately identifying lesions, a WCE image preprocessing algorithm is designed by fusing the VS-UNet and Criminisi algorithms to remove the highlight in endoscopic images. The algorithm includes two processes of highlight segmentation and repair, in the highlight segmentation stage, the UNet model is used as the basic framework to locate the highlight area, the optimised VGG16 is used as the backbone network, and the SPPFCSPC module is introduced to increase the model

receptive field; in the highlight repair stage, the Criminisi algorithm, which has a better reconstruction effect, is selected, and the original priority function algorithm is optimised. The experimental results show that the algorithm improves the accuracy of highlight segmentation and alleviates the irrationality of the highlight repair results to a certain extent.

2) To address the problem of small amount of WCE image data, small inter-class differences and large intra-class differences affecting classification accuracy, an image classification algorithm CA-ConvNeXt is studied. The algorithm combines image enhancement and transfer learning ideas to fine-tune the convolutional neural network ConvNeXt, and introduces a coordinate attention mechanism in the ConvNeXt network. Meanwhile, the Polyloss loss function is used to optimise the training process of the model. The experimental results show that the improved model can achieve accurate classification of normal, polyp and ulcer images under small sample conditions.

3) A lightweight polyp segmentation network model Li-DeepLabV3+ is explored to address the problems of complex polyp segmentation models and low accuracy of polyp segmentation. The network model takes DeepLabV3+ as the basic framework and uses the optimised MobileNetV2 network as the backbone network to reduce the model complexity; in the enhanced feature extraction part, an improved simple pyramid pooling module is used to reduce the model parameters while increasing the model training efficiency; to enhance the feature representation, shallow and deep features are fused using an improved unified attention fusion module, which uses the weights generated by spatial and channel attention, respectively, to fuse with the feature map. The experimental results show

that the Li-DeepLabV3+ model balances polyp segmentation accuracy and real-time performance with some generalisation ability.

4) Combining the optimal WCE image classification algorithm and the WCE image polyp segmentation algorithm, a WCE image intelligent assisted diagnosis system is designed. The system not only provides high-precision recognition results, but also has a user-friendly control interface.

In summary, this paper combines deep learning and medical image processing to design a WCE image intelligent recognition assisted diagnosis system, which reduces the burden of doctors while improving the efficiency and accuracy of diagnosis, and provides certain theoretical support for the in-depth exploration of the field of WCE image intelligent recognition.

Xiang Shiyu (Electronic Information)

Supervised by Wei Lisheng

Key words: Deep learning; Image restoration; Image classification; Semantic segmentation; Intelligent assisted diagnosis system; Wireless capsule endoscopy

目 录

摘 要.....	I
ABSTRACT.....	III
目 录.....	VI
第 1 章 绪论.....	1
1.1 研究背景和意义.....	1
1.2 国内外研究现状.....	2
1.2.1 内窥镜图像预处理研究现状.....	2
1.2.2 内窥镜图像分类研究现状.....	3
1.2.3 内窥镜息肉图像分割研究现状.....	4
1.3 主要研究内容.....	5
1.4 本章小结.....	7
第 2 章 基于深度学习的内窥镜图像预处理.....	8
2.1 传统的光斑移除算法.....	8
2.2 融合 UNet 与 Criminisi 的内窥镜光斑移除算法原理	9
2.2.1 UNet 模型	10
2.2.2 VS-UNet 模型.....	11
2.2.3 Criminisi 算法的改进.....	13
2.3 实验分析.....	14
2.3.1 光斑分割实验.....	16
2.3.2 光斑修复实验.....	17
2.4 本章小结.....	18
第 3 章 基于深度学习的内窥镜图像分类.....	19
3.1 基于 CA-ConvNeXt 的内窥镜图像分类模型	19
3.1.1 图像增强.....	19
3.1.2 ConvNeXt 模型	20
3.1.3 CA-ConvNeXt 模型	21
3.1.4 迁移学习.....	24
3.1.5 损失函数.....	25
3.2 实验分析.....	25
3.2.1 不同损失函数对于改进模型的影响.....	26
3.2.2 不同注意力机制对于改进模型的影响.....	28
3.2.3 与经典深度学习模型对比试验.....	30

3.3 本章小结.....	32
第 4 章 基于深度学习的内窥镜图像分割.....	33
4.1 基于 Li-DeepLabV3+的轻量级息肉分割模型.....	33
4.1.1 主干网络的改进.....	35
4.1.2 特征加强模块的改进.....	36
4.1.3 特征融合模块的改进.....	37
4.1.4 损失函数.....	39
4.2 实验分析.....	39
4.2.1 迁移学习对于改进模型的影响.....	40
4.2.2 不同损失函数对改进模型的影响.....	42
4.2.3 消融实验.....	45
4.2.4 定量分析.....	48
4.3 本章小结.....	50
第 5 章 基于深度学习的内窥镜图像智能识别系统设计.....	52
5.1 系统需求分析.....	52
5.2 系统应用环境需求.....	53
5.3 系统功能设计.....	53
5.4 界面设计及功能实现.....	54
5.4.1 设计思路.....	54
5.4.2 效果实现与测试.....	55
5.5 本章小结.....	58
第 6 章 总结与展望.....	59
6.1 总结.....	59
6.2 展望.....	60
参考文献.....	61
附录.....	66
攻读学位期间发表的学术论文.....	68
攻读学位期间获得的奖励.....	68
致 谢.....	69

插图清单

图 1-1 论文主要研究框架	6
图 2-1 传统算法的光斑移除结果	9
图 2-2 整体模型结构图	9
图 2-3 UNet 模型结构图	10
图 2-4 VS-UNet 模型结构图	11
图 2-5 VGG16 模型结构图	12
图 2-6 SPPFCSPC 模块	12
图 2-7 改进的 Criminisi 算法流程图	14
图 2-8 不同模型的光斑分割效果图	16
图 2-9 改进前后算法光斑修复效果图	18
图 3-1 CA-ConvNeXt 模型结构图	19
图 3-2 数据增强	20
图 3-3 ConvNeXt-T 模型结构图	21
图 3-4 SE 注意力机制	22
图 3-5 CBAM 注意力机制	22
图 3-6 CA 集成到 ConvNeXt 网络中结构图	23
图 3-7 微调后的网络最后几层结构图	24
图 3-8 混淆矩阵	26
图 3-9 不同损失函数的损失曲线	27
图 3-10 Grad-CAM 可视化	29
图 3-11 对比实验的混淆矩阵	30
图 4-1 DeepLabV3+模型结构图	33
图 4-2 Li-DeepLabV3+模型结构图	34
图 4-3 倒残差结构	35
图 4-4 加强特征模块	37
图 4-5 改进 UAFM 模块	38
图 4-6 迁移学习对改进模型的影响	42
图 4-7 不同损失函数的损失函数曲线	44
图 4-8 消融实验	47
图 4-9 消融实验模型参数以及 FPS 比较	47
图 4-10 不同数据集效果图	48
图 4-11 Li-DeepLabV3+和基准模型参数以及 FPS 比较	50
图 5-1 系统整体架构图	54

图 5-2 系统总体流程图	55
图 5-3 系统登录界面	56
图 5-4 患者基本信息界面	56
图 5-5 患者病理信息界面	57
图 5-6 患者 WCE 图像分类界面.....	57
图 5-7 患者 WCE 图像息肉分割界面.....	58

插 表 清 单

表 2-1 改进前后模型光斑分割性能对比.....	16
表 2-2 改进前后光斑修复结果对比	18
表 3-1 不同损失函数对模型的影响	28
表 3-2 添加不同注意力机制后的比较	29
表 3-3 与经典深度学习模型对比试验	31
表 3-4 与其他现有方法对比试验	31
表 4-1 剪枝后的 MobileNetV2 网络结构	36
表 4-2 迁移学习对改进模型的影响	42
表 4-3 不同损失函数在改进模型上的性能对比.....	44
表 4-4 消融实验	46
表 4-5 定量分析	49
表 5-1 运行环境	53

第 1 章 绪论

1.1 研究背景和意义

肠胃道疾病在人体消化系统中越来越常见，息肉、溃疡等常见消化道疾病与消化道肿瘤的发生和发展有着密切的关系，而消化道检查在诊断和评估消化道疾病方面发挥着关键作用，若能及早发现并治疗病变，将极大地减少消化道疾病的致死率^[1]。所以，通过早期检查及时发现并治疗消化道疾病，对于保证病人的健康状况和完善卫生保健系统都是非常重要的。

一般来说，内镜检查是诊断疾病的必要条件，也是识别胃肠道疾病的第一步^[2]。这些内镜检查可以改善对病变临床特征的分析，以确定病变的类型和严重程度，并做出正确的诊断。结肠镜检查、上消化道内镜检查及推进式小肠镜检查是治疗肠胃道疾病的有效方法，但其受内镜专家数量的限制并且检查耗时较长。WCE 是一种有别于传统内窥镜的无痛、无创伤的检测手段^[3]。WCE 是用于检查消化道的医疗设备，它采用无线技术和微型化设计，可以在人体内部进行无创检查。WCE 与传统的内窥镜可视化技术相比，具有高效率、高准确性和高度适应性等多项优势。WCE 通常由胶囊形状的摄像头、电池和无线传输模块组成。患者可以轻松地吞下胶囊，胶囊内的摄像头会自动拍摄消化道的图像，并通过无线传输模块将图像传输到外部设备供医生进行观察和诊断。在胃肠道的检查中，内窥镜可以捕捉 50000 多帧画面，同时储存在便携设备中。然而，目前临床上对海量影像进行人工诊断的方法仍受到外界影响，其结果与医生的技术水平、经验及精神状态密切相关，存在较强的主观性，易造成误诊漏诊^[4]。因此，需要一种功能强大且精确的计算机辅助检查系统，能够自动分辨出异常图像，在降低检查医生工作量的同时提高检查准确率^[5]。WCE 图像自动分类技术能够为医生提供可靠、客观的多种类 WCE 图像识别结果，从而节省临床医生宝贵时间的同时降低误诊率。

随着深度学习技术的发展，肠胃道病变的识别方法正从机器学习转向深度学习^[6]。因此，基于深度学习构建一个完善的 WCE 图像智能辅助诊断系统非常有意义。该系统能在保证精确识别的前提下，具备用户友好的控制界面，在降低医生压力的同时提高诊断效率。

1.2 国内外研究现状

近年来,利用计算机断层扫描、磁共振成像、正电子发射断层成像等多种先进的医学成像分析技术,为临床疾病诊断提供了巨大帮助。医学成像分析一般是由放射科医师、内科医师等医疗专家来进行。鉴于病理学领域的复杂性和差异性巨大,以及医疗专业人员精力分配有限,研究工作已经开始转向使用计算机辅助技术。然而,由于 WCE 图像中病变区域的纹理、颜色、形状等各不相同,计算机辅助诊断仍然面临很大挑战。针对 WCE 肠胃道图像的处理,国内外研究人员展开了深入的研究工作,主要涉及图像预处理、分类以及分割等方面。下面将对国内外相关领域的研究成果进行总结,并分析其不足。

1.2.1 内窥镜图像预处理研究现状

WCE 在拍摄图像时易产生细小的光斑^[7],这些光斑除了影响算法对病灶的识别以外,还会影响医生的诊断,增加漏检误检的风险。因此,如何在消除镜面光斑的同时,还能保持器官原有的色彩与纹理,具有重要的医学价值和研究意义。内窥镜光斑移除算法包括两个步骤:光斑分割和光斑修复。光斑分割是光斑修复的前提,通过光斑分割会生成坐标掩膜,其精确程度直接关系到光斑修复的结果。光斑修复技术就是将光斑分割后的区域重新构建出来,实现光斑的去除。国内外相关专家学者对镜面光斑的移除进行了大量研究,文献[8]通过用相邻区域填充光斑像素并执行高斯滤波获得模糊图像,然后融合模糊图像与自适应权重下的原始图像以实现光斑修复。文献[9]提出了一种仅从单个图像中分离和恢复镜面反射区域的新方法,该方法将高光像素的分离视为基于监督学习分类算法的二值化问题;同时,提出了一种基于多尺度动态图像扩展和融合的方法来恢复光斑区域。文献[10]提出了一种内窥镜光斑移除算法,在光斑定位方面,使用一种基于暗区域内生长和 scharr 滤波结合的方法定位相对光斑;在光斑填充方面,选用重建效果较好的 Criminisi 算法做出针对性的改进。文献[11]提出全自动内窥镜图像恢复模型 ReTouching,首先基于颜色变化和渐变信息条件相结合的方法分割光斑,其次引入图形数据结构的局部上下文的修复方法恢复光斑。文献[12]提出了一种基于亮度分类的内窥镜图像镜面反射区域分割和去除方法,在光斑分割阶段,使用随图像亮度变化的自适应阈值函数分割绝对光斑;在光斑恢复阶段,修改了基于样本的图像修复算法的优先级函数,

以确保合理正确的修复。

上述方法都是基于传统的方法，而在深度学习技术的不断完善下，一些学者开始尝试结合深度学习处理镜面光斑。文献[13]提出深度学习框架 EasySpec 适当地重建血管和精细的细节，该框架使用 U-Net 变体轻量级分割模型 Scaled-Unet 分割光斑，以及使用深度修复模型 GatedResUNet 恢复光斑。文献[14]在光斑分割阶段使用基于照明补偿的亮度阈值将内窥镜图像划分为高亮掩模和非高亮区域；光斑修复阶段使用集成了注意力机制的部分注意力网络(partial attention network, PatNet)，其专注于恢复有缺陷的像素。基于以上文献，大多数 WCE 图像高光修复方法需要人为干预，操作较复杂具有一定的局限性。

1.2.2 内窥镜图像分类研究现状

WCE 图像的特殊成像模式和病理的多样性使得 WCE 图像之间存在类间差异小和类内差异大等问题，因此，对自然图像有良好效果的分类模型可能并不适用于 WCE 图像。最初，机器学习技术如朴素贝叶斯、决策树、随机森林等被用于 WCE 图像分类^[15]。文献[16]基于归一化方差位置约束，挖掘了局部纹理特征，提出了一种基于正样本的息肉识别方法。机器学习模型的主要局限性在于需要领域专家来正确识别用于分类的重要特征。大多侧重于使用手工制作的特征(如纹理、颜色、形状等)来训练分类器^{[17][18]}，给 WCE 医师的诊断提供了一定帮助，但在传统的特征提取过程中，针对不同的数据集需设计不同的特征，而这些特征的适用性在一定程度上是有限的，主要依赖于设计者的实践经验。而深度学习算法不需要人工提取图像特征，在协助放射科医生检查和帮助诊断疾病方面发挥着至关重要的作用^[19]。深度学习技术能够自动提取特征，有助于提高模型的性能，其预测的准确性由模型架构、超参数以及数据集的质量和大小决定。文献[20]结合全局平均池化和经典深度学习网络 ResNet，设计了一种 ResNet-GAP 网络，用于在白光下息肉和正常 WCE 图像的精确分类。文献[21]首先利用平均滤波器去除 WCE 图像的噪声，然后使用 DenseNet-121 网络提取形状和位置等特征，并使用增强鲸鱼优化算法选择基本特征；最后，提出了一种 SEResNet 技术，用秃鹰搜索优化方法对八类胃肠道疾病进行分类，以获得更好的分类结果。文献[22]通过在经典深度神经网络 DNN 中引入一种新的类不平衡损失函数，用于解决 WCE 图像分类中数据集不平衡的问题。以上方法可以有效地分类 WCE 图像的病变，为进一步提高模型对病变区域的注意力，研究人员试

图将注意力机制与深度学习网络结合起来。文献[23]对残差网络 ResNet 进行改进并嵌入注意力机制 SE 模块实现胃炎的分类。文献[24]通过融入混合注意力机制的卷积神经网络实现胃肠道疾病精确分类, 该网络由多个注意力层组成, 通过堆叠层来生成注意力特征以优化网络, 同时使用数据增强技术克服数据不平衡问题。

上述研究取得了较好的 WCE 图像分类结果, 但在 WCE 图像分类领域仍存在一些挑战, 主要表现在以下两个方面: 1) WCE 在成像过程中会产生气泡、混浊和伪影等干扰因素影响模型注意力; 2) 病灶的分布和大小导致类内差异较大, 增加了区分各种类型病灶的难度。因此, WCE 图像的数据预处理以及 WCE 图像分类模型的注意力还有待改进。

1.2.3 内窥镜息肉图像分割研究现状

准确的息肉分割不仅有助于检查中发现病变, 还有利于手术中切除异常组织, 在一定程度上降低肠胃道癌症的发生。但由于 WCE 检查中息肉的大小、形状、颜色、质地和外观各不相同, 息肉与其周围黏膜高度相似, 异常组织的边界不明确, 类间相似性和类内变异性高, 息肉精确分割的实现仍然具有挑战性。国内外许多相关专家学者对息肉分割相关方向进行了研究, 息肉分割方法主要分为基于手工制作特征和基于深度学习两类。传统的息肉分割方法有区域生长、基于阈值、统计形状、边缘检测等^[25]。然而, 这些传统方法往往存在分割效果差、效率低、复杂度高等问题。传统方法的息肉分割能力有限, 难以满足医学图像分割的实际应用需求^[26]。深度学习技术极大地改善了息肉分割的发展, 文献[27]通过应用条件随机场和测试时间增强来扩展 ResUNet++架构进一步提高息肉分割能力。文献[28]等提出了一种称为增强型多尺度网络用于精确地分割息肉, 采用一系列多尺度方法增强特征提取并利用局部上下文注意模块来增加对息肉边界的关注。文献[29]介绍了一种基于编码器和解码器的架构用来生成息肉的掩模和边缘图像, 构建了对象和边缘反向的注意力块作为编码器和解码器之间的桥梁。文献[30]在边界增强模块中提出了边界感知注意力模块, 该模块采用编码器特征和部分解码器输出以创建息肉掩码。文献[31]为了更加精确分割息肉的同时满足临床上对轻量级的需求, 提出了使用空洞空间金字塔池化的 Li-SegPNet 模型, 通过注意力门控的跳跃连接来解决编码器和解码器之间的语义差距。文献[32]提出了卷积多层感知器息肉分割网络, 以实现更准确的息

肉分割，采用卷积 MLP 编码器，利用并行自注意力模块增强低级特征；引入了级联上下文聚合模块聚合深层语义特征和低级局部特征；同时，利用通道引导组反向注意力，通过挖掘区域和边界线索之间的关系增强结构和纹理细节。文献[33]为了解决模型泛化和分割性能低的挑战，提出一种上下文特征细化 (Context Feature Refinement, CFR) 模块，通过使用多个并行卷积层从特征图中提取上下文信息，并逐渐增加内核大小，促使网络能够精准识别和有效分割 WCE 图像中的细节部分。文献[34]提出了一种基于编码器-解码器框架的级联边界引导注意力网络，将细粒度边界信息引入编码器和解码器层的跳跃连接中，以实现精确的息肉分割；此外，将级联细化策略用于边界特征的多阶段增强以逐步产生更好的预测。

基于上述文献，尽管息肉分割算法的性能从不同方面得到了加强，并取得了一定的积极成果，但由于息肉分割在临床应用中具有较高要求，如何进一步提高息肉分割算法的精确性、轻量性和泛化性等值得探索。

1.3 主要研究内容

在 WCE 检查过程中，医生需深入地阅读大量图像，花费时间较久且容易错过有价值的信息。因此借助计算机诊断系统降低误诊率尤为重要。本文基于深度学习结合 WCE 图像预处理、WCE 图像分类和 WCE 图像语义分割三个方面建立一个 WCE 图像智能识别辅助诊断系统。首先，融合 UNet 与 Criminisi 算法对 WCE 图像预处理，实现光斑的检测和修复；其次，结合数据增强和迁移学习策略研究一种 CA-ConvNeXt 内窥镜图像分类模型，实现正常、息肉和溃疡的分类；最后，探讨了一种轻量级结肠息肉分割网络模型 Li-DeepLabV3+，实现息肉肿瘤类和非肿瘤类的语义分割。论文主要研究框架如图 1-1 所示，具体内容如下：

第 1 章：详细阐述了本课题的研究背景和意义，分别对 WCE 图像光斑检测与修复，WCE 图像分类以及 WCE 息肉图像语义分割的国内外研究现状进行了概述，并给出本文的主要研究内容。

第 2 章：探讨了一种融合 UNet 与 Criminisi 算法移除 WCE 图像光斑预处理算法。在光斑分割阶段，以 UNet 网络为基本框架定位光斑区域，使用优化的 VGG16 作为主干网络，并且引入 SPPFCSPC 模块增加模型感受野。在光斑修复方面，使用优化原始优先权函数的 Criminisi 算法实现光斑区域的修复。实验证

明该算法可以有效分割光斑区域并实现光斑修复。

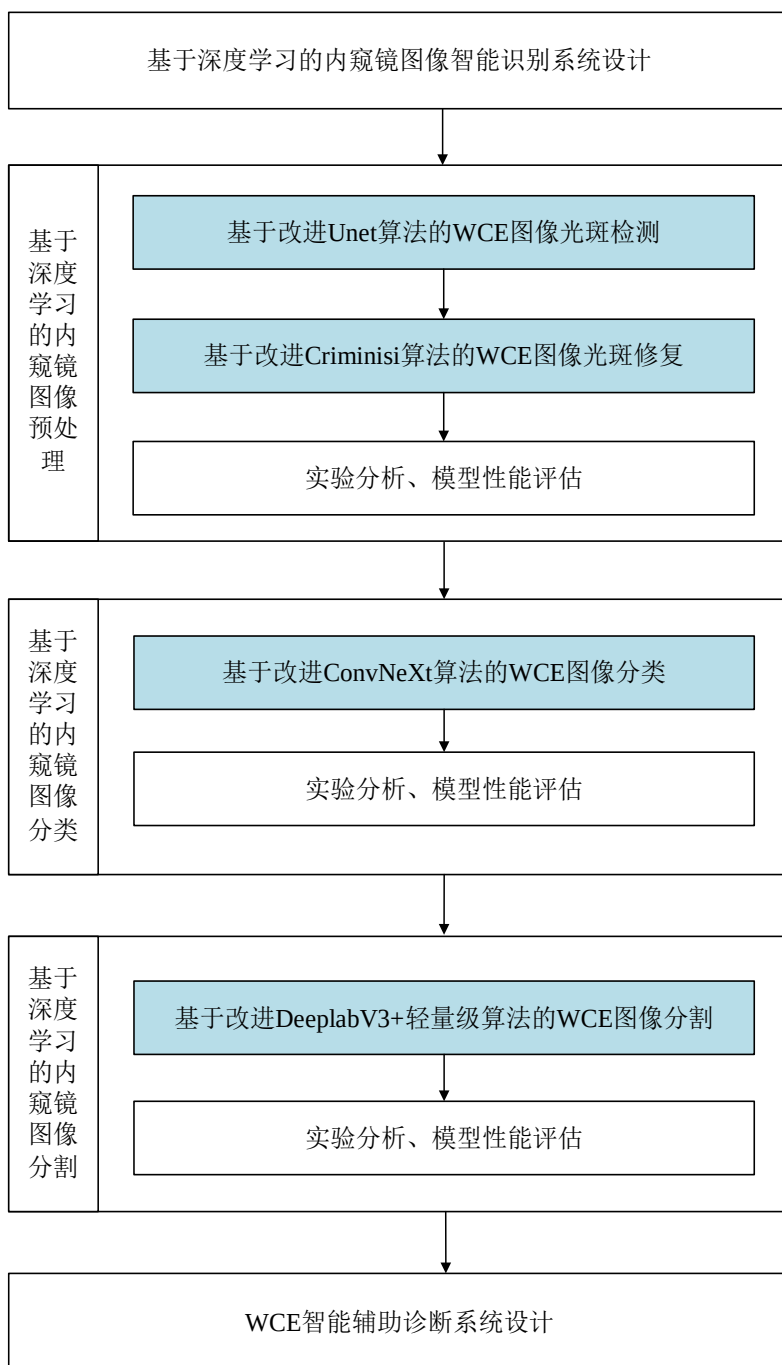


图 1-1 论文主要研究框架

第 3 章：设计了一种图像分类算法 CA-ConvNeXt，首先在 ConvNeXt 模型中引入坐标注意力机制(CoordAttention, CA)，结合迁移学习并使用 Polyloss 损失函数优化模型的训练过程。实验验证了该模型在小样本条件下可以实现正常、息肉以及溃疡的精确分类，为溃疡和息肉的诊断提供了帮助。

第 4 章：研究了一种轻量级息肉分割网络模型 Li-DeepLabV3+。该网络模型

以 DeepLabV3+ 为基本框架，首先，使用优化的 MobileNetV2 网络作为主干网络；其次，采用改进的简易金字塔池化模块(Simple Pyramid Pooling Module, SPPM)加强特征提取；最后，为增强特征表示，使用改进的统一注意力融合模块(Unified Attention Fusion Module, UAFM)融合浅层特征和深层特征，该模块同时利用空间注意力和通道注意力产生的权值与特征图融合。实验结果表明，Li-DeepLabV3+模型在分割精度和分割速度方面表现优越，并具有一定的泛化能力。

第 5 章：结合前面章节的研究设计了一种基于深度学习的 WCE 图像智能识别系统。该系统对以上 WCE 图像识别算法进行了良好的集成，系统实现了 WCE 图像分类和 WCE 息肉图像语义分割的功能，并且提供了精确的诊断结果辅助医生诊断。

第 6 章：对本文的主要研究内容进行概括总结，并在此基础上分析下一步工作方向。

1.4 本章小结

本章首先对课题的研究背景和意义进行介绍，WCE 图像诊断需要耗费医生大量精力，亟需计算机技术辅助诊断，这一问题促进了基于深度学习 WCE 图像智能识别系统的研究；然后分别从 WCE 图像预处理，WCE 图像分类以及 WCE 息肉图像语义分割方面介绍当前国内外研究现状；最后概括全文的主要研究框架，并简要阐述各章节内容，为后续的详细研究做铺垫。

第2章 基于深度学习的内窥镜图像预处理

在 WCE 拍摄图像时，由于光源和相机的角度几乎相同，再加上消化道环境较特殊，镜面反射始终存在于 WCE 图像中，这些高亮的点在视觉上表现为细小的光斑，并可能隐藏血管和病变等有价值的信息。这些光斑易干扰算法识别病灶，而且还会造成操作医师的视觉疲劳，从而降低诊断的准确性。迄今为止，多数研究都集中在识别和抑制内窥镜数据中的镜面反射上，并取得了令人满意的结果。然而，现有的图像预处理大多数需要人为进行操作，操作繁琐且具有局限性。

2.1 传统的光斑移除算法

光斑移除包含光斑的分割和修复两个阶段，传统的光斑分割主要通过阈值分割定位光斑的位置；传统的光斑修复大部分通过已有像素信息还原被光斑覆盖的像素信息，例如 FMM 算法^[35]。该算法在确定光斑区域以后，为确保像素值的连续性，利用周围已知像素信息估计未知像素值。算法迭代运算由外向内进行，直到所有待修复区域的像素值被替代。边界法线和轮廓上的像素会得到更多权重，修复过程较连续。

本章节先对 WCE 图像的光斑部分进行阈值分割，然后利用周围的像素值填充需要修复的区域，从而实现光斑的移除。具体步骤如下：

- 1) 光斑分割：使用 200-255 的阈值范围对 WCE 图像的光斑区域进行分割，该过程会批量生成对应的 mask 掩膜；
- 2) 光斑修复：将光斑分割得到的 mask 掩膜与原图一同输入 FMM 修复算法进行 WCE 图像修复。

图 2-1 展示了传统算法的光斑移除结果。由图 2-1 可见，对于光斑分割部分，存在部分面积较小的光斑区域未分割完全的现象，这是由于阈值分割存在一定的局限性，容易忽略亮度较低的光斑导致分割不完全。对于光斑修复部分，修复后的区域与未修复区域存在明显的边界与色差，未能有效重建图像。这说明传统的修复算法不适用于 WCE 图像光斑移除领域，无论是光斑分割还是光斑修复算法都有待改善。

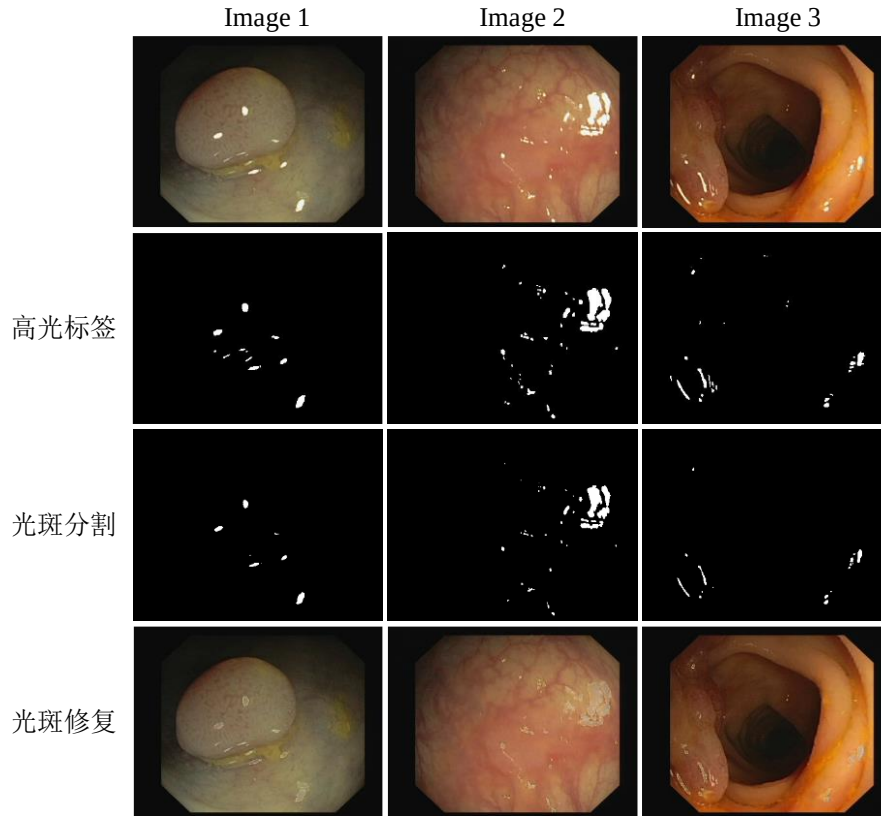


图 2-1 传统算法的光斑移除结果

2.2 融合 UNet 与 Criminisi 的内窥镜光斑移除算法原理

为了缓解传统图像移除算法存在明显的边界与色差问题，本章节探讨了一种融合 Unet^[36]与 Criminisi^[37]的 WCE 图像光斑移除算法，模型结构如图 2-2 所示。WCE 图像光斑移除算法主要分为光斑分割和光斑修复两个步骤。光斑分割是光斑修复的前提，光斑分割会生成一个坐标掩膜，其精确程度直接关系到光斑修复的效果；光斑修复是对光斑分割后的区域进行重建，实现光斑的去除。

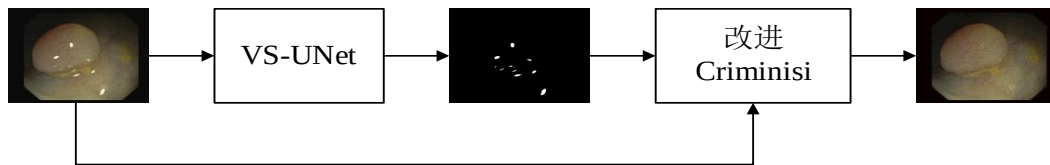


图 2-2 整体模型结构图

WCE 光斑图像先进入 VS-UNet 模块实现光斑区域分割并生成一个光斑掩膜图，然后将存在光斑的原图和光斑掩膜图同时输入改进的 Criminisi 模块得到修复光斑后的 WCE 图像。在光斑分割阶段，使用 VGG16^[38]提取图像特征，并引

入空间金字塔池化快速跨阶段局部通道(Spatial Pyramid Pooling Fast and Cross Stage Partial Channel, SPPFCSPC)模块提高模型的感受野和特征表达能力, 结合迁移学习训练模型使其准确分割光斑区域。在光斑修复阶段, 使用重建效果较好的 Criminisi 算法, 并优化原始优先权函数增强 WCE 图像修复能力。

2.2.1 UNet 模型

Unet 是 FCNN 的改进网络, UNet 包括对称的编码器和解码器, 编码器负责提取特征并捕获图像的上下文信息; 解码器则负责将特征图上采样到原始输入图像的尺寸, 并通过跳跃连接将编码器中的浅层和深层特征进行融合, 从而保留更多的细节信息^[39]。这种架构使得 UNet 在处理边缘和细节时非常有效。图 2-3 是 UNet 模型结构图。

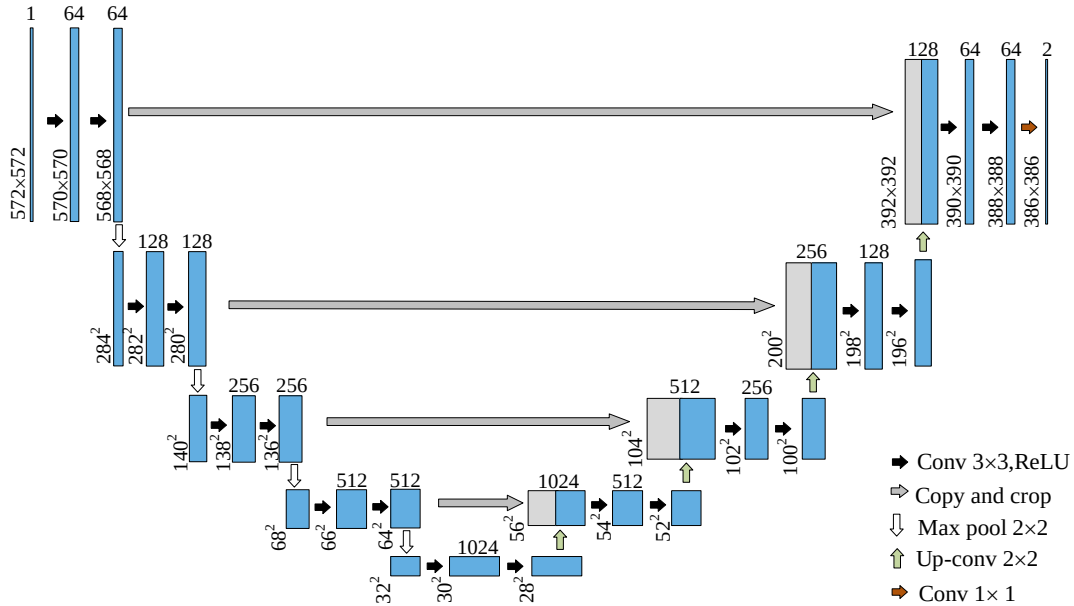


图 2-3 UNet 模型结构图

不同层通过一个 2×2 的最大池化层相互连接, 该层由图 2-3 中的白色箭头标记。特征图的大小在每次通过 max-pool 时都会减半。由于未设置填充, 当特征图大小为奇数时, 某些像素信息会存在丢失现象。因此, 需要仔细设置输入图像的大小, 并将图像的长度和宽度保持为偶数个像素点。当卷积层中通道数量增多时, 图像中的特征通道数量也随之增多, 这些层由 2×2 转置卷积上采样层连接。每次上采样后, 图像的尺寸都会增加一倍。特征增强网络的每一层都将融合左侧特征提取部分的特征, 如图 2-3 中灰色箭头所示。但是, 左侧的图像

比右侧的图像大，因此在特征融合之前需要进行一些剪切。两部分之间的每一层在进行 3×3 卷积操作后经过 Relu 激活函数层，如图 2-3 中的黑色箭头标记，卷积运算是步幅为 1，卷积核为 3×3 。由于未设置填充，因此每次卷积操作后图像高和宽将减少两个单位。输出的最后一层使用 1×1 卷积实现分类。

在 Unet 的原始结构的下采样方法中，卷积和池化可能会导致目标空间信息和详细特征丢失，因此，对于 WCE 图像中小像素目标的特征提取并不理想；同时，由于 Unet 结构是一个具有固定层(通常为五层)的神经网络，只能提取固定深度的特征，难以适应 WCE 图像中不同深度的信息特征。

2.2.2 VS-Unet 模型

为了提高 WCE 图像特征提取的效率，选择 VGG16 作为 Unet 分割模型的主干网络。同时，为解决原有 Unet 结构中卷积、池化等操作导致目标空间信息和细节信息丢失的问题，在 Unet 中引入 SPPFCSPC 模块扩大模型感受野，从而减少信息丢失。改进后的模型被称为 VS-Unet 模型，VS-Unet 的结构如图 2-4 所示。

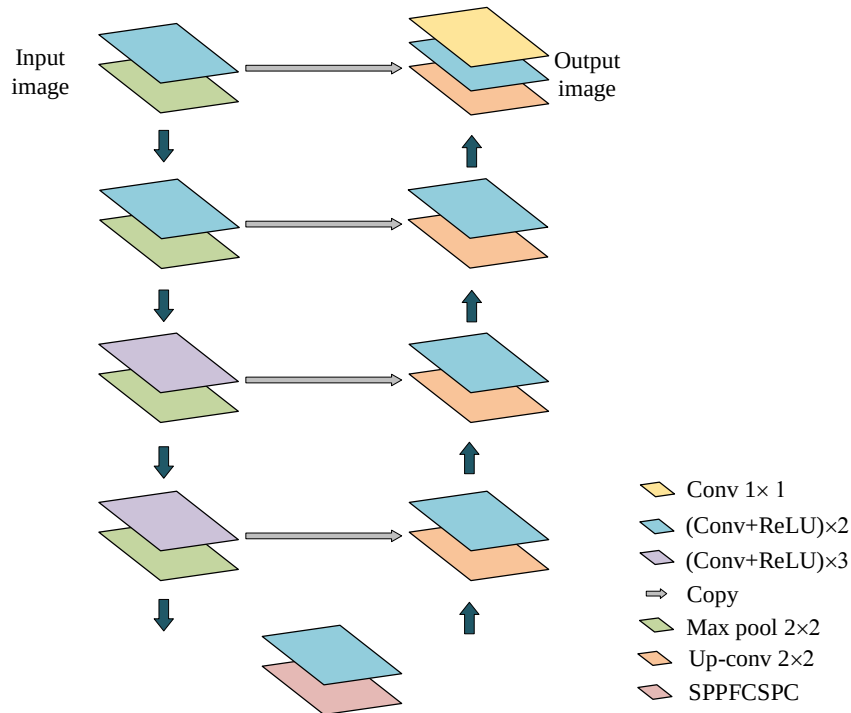


图 2-4 VS-Unet 模型结构图

图 2-4 左侧是网络的编码器部分，包括五个卷积和四个池化操作。在获得第五个有效特征层后，加入 SPPFCSPC 模块以聚合不同区域的上下文信息，实现对面积较小的光斑区域的精确分割。图 2-4 右侧是解码器部分，包括四个 2 倍上

采样操作和 8 个 3×3 的卷积操作。中间是跳跃连接部分，去除了原先的剪切操作以保证输入和输出图像大小一致。最后使用 1×1 卷积调整通道数为 2，即光斑和背景两个种类。

Unet 分割模型的主干网络选择 VGG16 用于内窥镜图像的特征提取。VGG16 共有 16 层深度，其中 13 个卷积层用于从输入图像中提取不同层次的特征，3 个全连接层负责将特征映射到输出类别上。VGG16 以其简单而有效的架构闻名，采用了小尺寸的 3×3 卷积核和多个重复的卷积层。该模型在图像分类、识别和定位等方面都有很好的效果，并得到广泛应用。其模型结构如图 2-5 所示。

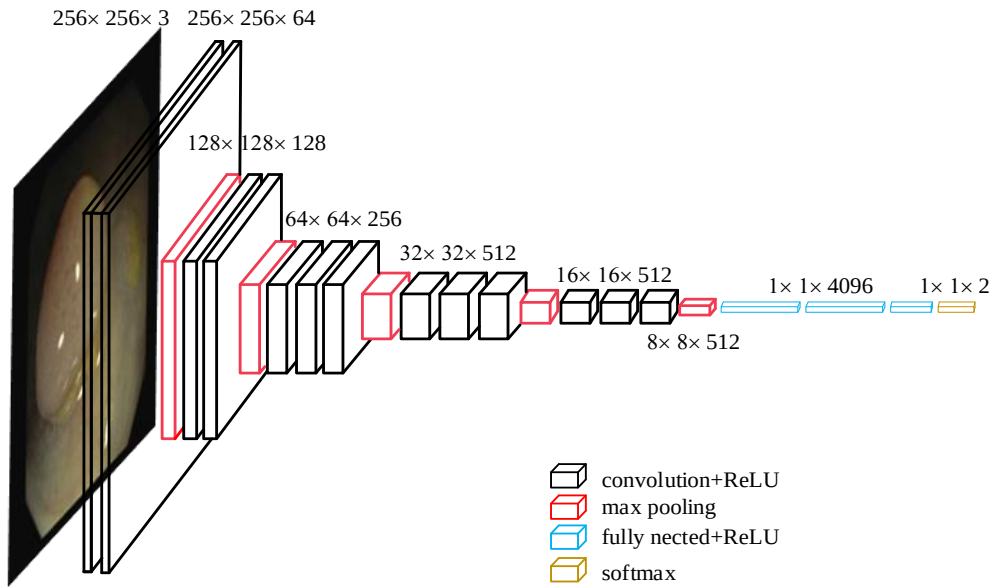


图 2-5 VGG16 模型结构图

为了在不增加过多参数的情况下提高模型的感受野，在 Unet 网络中引入 SPPFCSPC 模块，其结构如图 2-6 所示。

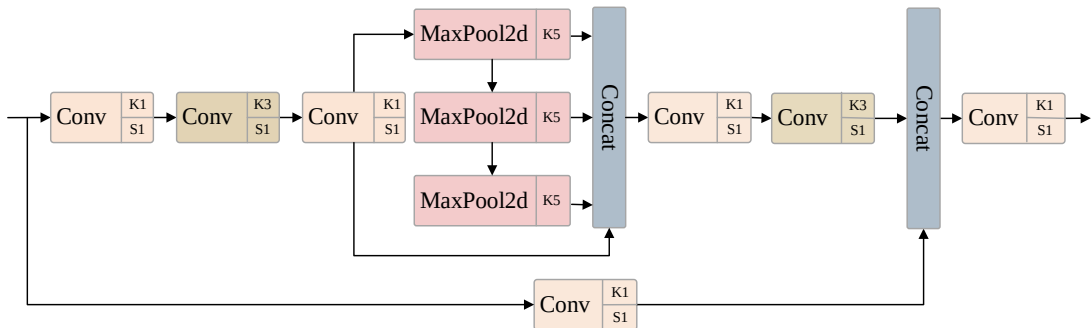


图 2-6 SPPFCSPC 模块

在图 2-6 中，SPPFCSPC 模块的作用是对输入特征图进行多尺度的池化操作，从而提取更加丰富的特征信息。具体来说，SPPFCSPC 模块会对输入进行四种

不同大小的池化操作，然后将这些池化结果拼接在一起，作为下一层的输入。在不增加模型参数的情况下，提高其感受野，从而提升网络对光斑区域的分割性能。

2.2.3 Criminisi 算法的改进

由于 Criminisi 算法在处理大面积区域修复时拥有较好的重构能力，并且在结构还原上效果优越，因此选用 Criminisi 算法来进行光斑修复。Criminisi 将受损边缘的像素作为中心，通过匹配图像中已有的像素块来填充待修复区域，从而修复图像。该算法包含以下三个步骤：

- 1) 优先权的计算：Criminisi 算法的优先级由置信度决定，该置信度表示待修复块中原始图像中信息的比例，以及显示线性结构强度的数据项。
- 2) 搜索并填充最佳匹配模板：在确定优先修复的像素后，基于 SSD 原理在完整区域进行全局最优匹配模块的搜索和填充。
- 3) 更新置信度：当选择出了最佳的匹配模板并将其填充到相应的位置后，待修复区域的边缘将会进行相应的更新。

连续重复上述三个步骤，直到填充所有待修复区域。为避免 Criminisi 算法因优先级计算不准确导致修复有误，其优先级的计算方法被改进，具体流程图如图 2-7。其中： n 具体计算式为 $n = \min(H, W) / 30$ ， H 、 W 分别表示图像的高度和宽度。

Criminisi 算法修复块的原始优先权的计算方法如式(2-1)-(2-3)所示，其中： $C(p)$ 用于衡量目标块 Ψ_p 中已知信息的比重， ϕ 是已知的图像信息， p 为 $\partial\Omega$ 上的一点； Ψ_p 是以 p 点为中心点 $h \times w$ 邻域的目标块； $D(p)$ 称为数据项，数据项的目的是优先填充有较强纹路结构的点， $\nabla I_p^\perp$ 为 p 点的等照度线向量，它的方向与点 p 的梯度向量呈相互垂直关系， n_p 是 p 点的法向量， α 为归一化因子^[40]。

$$P(p) = C(p) \times D(p) \quad (4-1)$$

$$C(P) = \frac{\sum_{p \in \Psi_p \cap \phi} C(P)}{|\Psi_p|} \quad (4-2)$$

$$D(p) = \frac{|\nabla I_p^\perp \times n_p|}{\alpha} \quad (4-3)$$

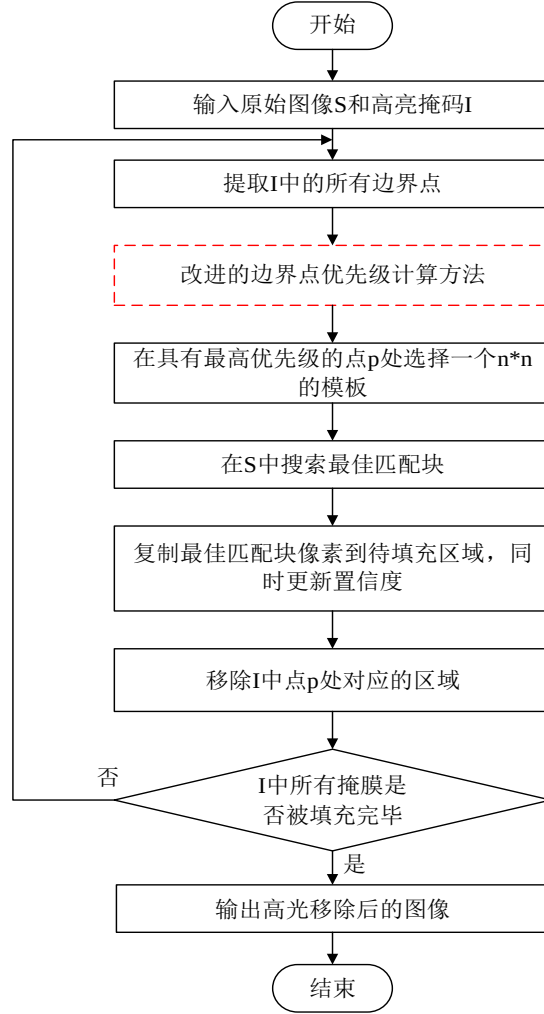


图 2-7 改进的 Criminisi 算法流程图

在迭代过程中，置信值 $C(p)$ 会出现下降为零的情况，导致计算出的优先值不具有可靠性，从而影响修复顺序^[40]。由式(2-1)可知，当 $C(p)$ 为 0 时，无论数据项 $D(p)$ 为何值，优先权 $P(p)$ 的值均为零，从而使该目标块失去了优先修复的机会。为了解决上述问题，对优先权 $P(p)$ 做出修改，计算方式如式(2-4)所示，其中， α ， β ， γ 三个系数分别取 0.5、0.4、0.1。

$$P(p) = \alpha C(p) + \beta D(p) + \gamma e^{C(p) \cdot D(p)} \quad (4-4)$$

2.3 实验分析

为验证 VS-Unet 模型以及改进 Criminisi 算法的有效性，本章节使用 CVC-ClinicSpec 数据集进行实验，该数据集包含 612 张结肠镜检查图像，带有由经验丰富的放射科医生注释的镜面真实标签^[41]。在实验中，训练集、验证集和测试

集的选取比例为 8:1:1。

由于使用的数据集较小，所提模型结合迁移学习思想利用 VOC 数据集的预训练权重进行训练。训练方式分为冻结阶段和解冻阶段两个部分，其中，冻结训练的迭代次数设置为 20，batch_size 设置为 8；解冻训练迭代次数设置 30，batch_size 设置为 4；使用 AdamW 优化器其初始学习速率设置为 0.0001，weight_decay 设置为 0.02；并使用余弦退火的学习率下降方式。

为了对光斑分割性能进行评估和分析，选用交并比(IoU)、Dice 系数(Dice)、精确率(Precision, Pre)、召回率(Recall, Re)、准确率(Accuracy, Acc)五个指标进行衡量。上述指标的计算式如下：

$$IoU = \frac{x_{TP}}{x_{TP} + x_{FP} + x_{FN}} \quad (4-5)$$

$$Dice = \frac{2x_{TP}}{2x_{TP} + x_{FP} + x_{FN}} \quad (4-6)$$

$$Pre = \frac{x_{TP}}{x_{TP} + x_{FP}} \quad (4-7)$$

$$Re = \frac{x_{TP}}{x_{TP} + x_{FN}} \quad (4-8)$$

$$Acc = \frac{x_{TP} + x_{TN}}{x_{TP} + x_{FP} + x_{FN} + x_{TN}} \quad (4-9)$$

其中： x_{TP} 表示真阳性像素点的个数； x_{FP} 表示假阳性像素点的个数； x_{FN} 表示假阴性像素点的个数^[42]。

光斑修复部分采用峰值信噪比(Peak Signal-To-Noise Ratio, PSNR)和结构相似性(Structure Similarity Index Measure, SSIM)对光斑修复方法的性能进行评估。PSNR 和 SSIM 如式(2-10)-(2-14)。

$$PSNR = 10 \log_{10} \frac{255^2}{\frac{1}{H \cdot W} \sum_{k=1}^H \sum_{l=1}^W (\hat{x}_{k,l} - x_{k,l})^2} \quad (4-10)$$

其中： x 和 $\hat{x}$ 分别表示原始图像和修复后图像， H 和 W 分别表示图像的高度和宽度。

$$L(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (4-11)$$

$$C(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (4-12)$$

$$S(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (4-13)$$

$$SSIM = L(x, y) \cdot C(x, y) \cdot S(x, y) \quad (4-14)$$

其中： μ_x 和 μ_y 分别是图像 X 与图像 Y 的均值； σ_x 、 σ_y 是图像 X 和 Y 的标准差； σ_x^2 、 σ_y^2 是它们的方差； $\sigma_x\sigma_y$ 代表图像 X 和 Y 的协方差； C_1 、 C_2 、 C_3 均为常数，起维持分式稳定的作用。为了避免分母为 0 的情况，通常取 $C_1 = (K_1 \cdot L)^2$, $C_2 = (K_2 \cdot L)^2$, $C_3 = C_2 / 2$ ，其中： $K_1 = 0.01$, $K_2 = 0.03$, $L = 255$ 。

2.3.1 光斑分割实验

为验证 VS-Unet 模型的优越性，在相同的实验条件下将 VS-Unet 与经典的分割算法 Hrnet、Unet 进行对比实验。图 2-8 展示了光斑分割结果，其中 Ground Truth 表示的是光斑标签。

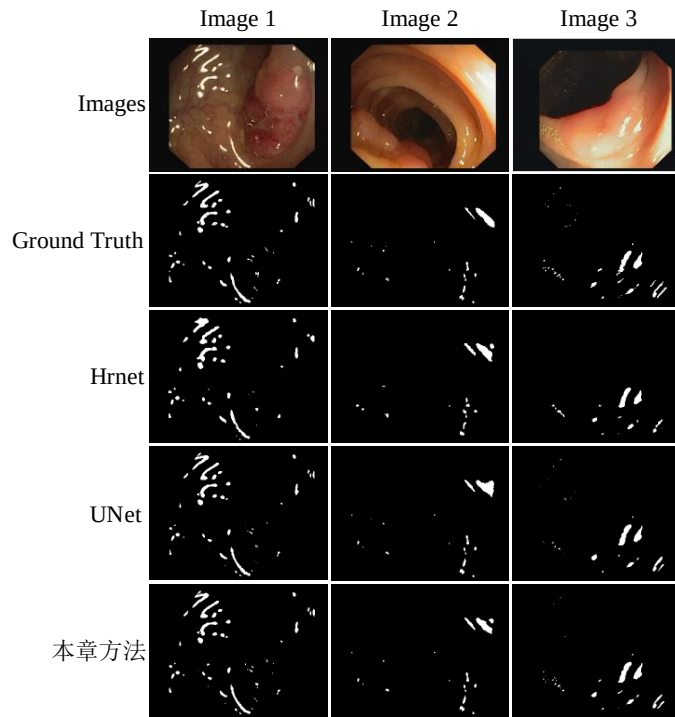


图 2-8 不同模型的光斑分割效果图

从图 2-8 可以看出，相对于 Hrnet 和原始的 Unet 网络，VS-UNet 模型的光斑

分割结果更接近 Ground truth，漏检误检的可能性更小。具体来说，对于 Image 2，改进后的光斑分割效果更优越，误检的区域较小；对于 Image 1 和 Image 3，可以看出 Image 1 中间靠下区域以及 Image 3 左下角区域的小面积光斑分割更全面，漏检的光斑区域更小。这些结果表明，SPPFCSPC 模块通过提高模型感受野，加强模型的特征提取能力，从而使改进的算法能够准确定位内窥镜图像中的大部分光斑反射区域。三个模型评价指标如表 2-1 所示。

表 2-1 改进前后模型光斑分割性能对比

方法	Dice(%)	IoU(%)	Acc(%)	Pre(%)	Re(%)
HrNet	64.68	47.81	99.45	68.14	61.75
UNet	80.72	67.67	99.70	88.13	74.46
本章方法	81.54	68.85	99.71	84.78	78.56

由表 2-1 可见，Unet 模型的光斑分割性能优于 Hrnet 网络。同时，改进 Unet 模型的 Dice、IoU、准确率和召回率均高于原模型，分别得到了 0.82%、1.18%、0.01%和 4.1%的提升，这表明本文改进算法的有效性。由于精确率和召回率是相互影响的关系，改进模型的精确率较 UNet 模型低 3.35%，但是其召回率达到 78.56%，表明该模型的漏检的可能性更小，具有较强的临床应用潜力。同时也验证本文所加模块的有效性，能够提高网络获取全局信息的能力从而加强光斑区域分割能力。

2.3.2 光斑修复实验

为验证改进 Criminisi 算法的有效性，在光斑分割测试集的 61 张图像上进行光斑修复实验，并与改进前的 Criminisi 算法作对比，修复结果如图 2-9 所示。

从图 2-9 中可以看出，原 Criminisi 算法大致能完成修复。但是对于 Image 3，右下角产生了黑色失真，出现修复不合理现象，改进优先级的 Criminisi 算法修复结果更加真实可靠。同时，无论待修复区域是否存在纹理，都实现了较好的重构，毛细血管和人体肠道组织轮廓得到了很好的连接。改进前后 Criminisi 算法光斑修复评价指标如表 2-2 所示。

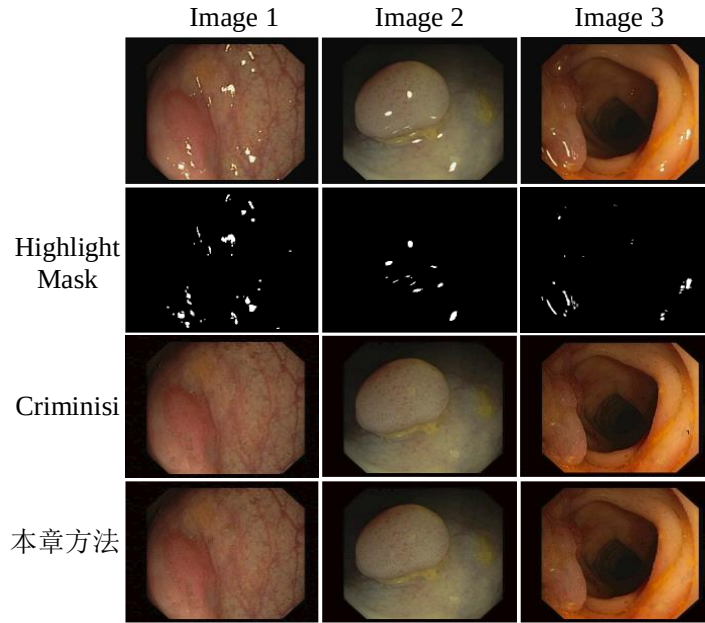


图 2-9 改进前后算法光斑修复效果图

表 2-2 改进前后光斑修复结果对比

方法	PSNR	SSIM
Criminisi	27.09	0.775
本章方法	27.87	0.780

由表 2-2 可以得到改进后的 Criminisi 比原始算法的 PSNR 和 SSIM 值分别高出 0.78dB 和 0.005。这些数值的提升不仅体现了改进算法的有效性和可行性，同时也说明了所提算法在实际应用中具有更高的修复能力。

2.4 本章小结

本章节首先介绍了传统 WCE 图像的光斑移除算法，分析其移除结果存在修复不合理的情况。为了解决此问题，探讨了一种融合 VS-UNet 与 Criminisi 算法的光斑移除算法。该算法包括光斑分割与修复两个过程，在光斑分割方面，以 UNet 网络为基本框架定位光斑区域，使用优化的 VGG16 作为主干特征提取网络，并且在 VGG16 获得的第五个有效特征层后加入 SPPFCSPC 模块增加模型感受野。在光斑修复方面，使用优化原始优先权函数的 Criminisi 算法实现光斑区域的修复。实验结果表明所提算法提高了光斑分割的精度，并且在一定程度上缓解了光斑修复结果的不合理性，为后续 WCE 图像处理工作奠定了基础。

第3章 基于深度学习的内窥镜图像分类

结直肠癌、胃癌等是常见的消化道疾病，对人类健康构成重大威胁。息肉、溃疡等常见消化道疾病与消化道癌变密切相关。如果能够在早期发现这些癌前病变，可以提高患者的生存概率，这就体现了准确分类出这些病变的重要性。现有的 WCE 图像分类算法实现了较高的准确率，然而大部分模型的训练需要大量数据做支撑，并且只实现单一病变的分类。本章节将在第 2 章节图像预处理的基础上，结合数据增强策略探讨一种改进的 ConvNeXt 模型，实现较少样本条件下正常、溃疡和息肉图像的精确分类。

3.1 基于 CA-ConvNeXt 的内窥镜图像分类模型

针对 WCE 图像数据量少，类间差异小，类内差异大而影响分类准确率的问题，构建了一个基于数据增强与迁移学习的 CA-ConvNeXt 内窥镜图像分类模型。该模型以 ConvNeXt 为研究对象，采用坐标注意力机制强化模型对病灶区域的注意，并通过 Polyloss 损失函数对模型进行优化，最终达到小样本情况下的准确分类^[42]。该模型实现的具体结构图如图 3-1 所示。

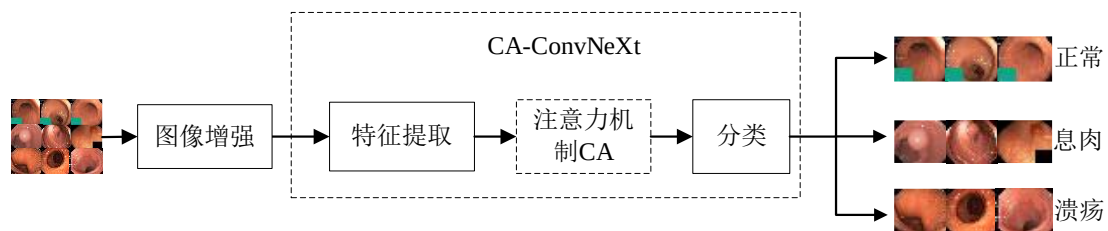


图 3-1 CA-ConvNeXt 模型结构图

由图 3-1 可知，在小样本数据条件下，首先使用数据增强对 WCE 图像数据集进行扩充，然后结合迁移学习使用 CA-ConvNeXt 模型实现正常、息肉、溃疡三个种类的分类。

3.1.1 图像增强

深度学习的训练过程中一般需要大量的样本，在其他条件相同的情况下，若使用的样本数量越大，越有利于模型的训练。为此，使用一种数据增强方法来防止因数据集较小而可能导致的过拟合。首先，对 WCE 图像周围存在的黑色边框进行裁剪处理。这些黑色边框不仅不能反映图像本身的特征，反而会对需

要识别的目标区域产生干扰，影响模型的注意力。然后，再对裁剪后的 WCE 图像进行数据增强。数据增强方法包括水平翻转，垂直翻转以及水平垂直翻转相结合这三种。数据增强效果图如图 3-2 所示。

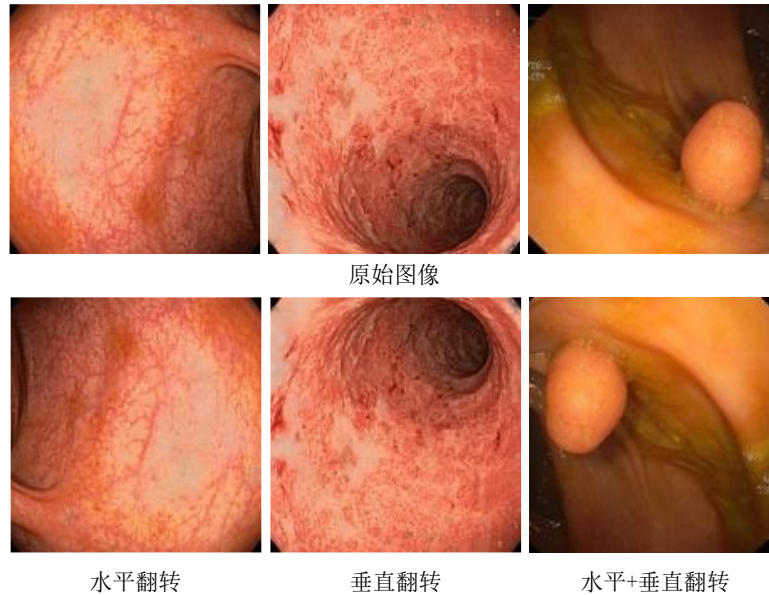


图 3-2 数据增强

3.1.2 ConvNeXt 模型

ConvNeXt 是一种基于卷积神经网络的模型架构，具有良好的分类性能和运算效率^[43]。ConvNeXt 结合了 Inception 和 ResNet 等网络结构的优点，采用分组卷积和逐步增强的思想，从而实现更高的分类准确性和更低的计算复杂度。该模型通过一系列卷积层捕获图像中的复杂图案的特征，展示了出色的特征提取能力^[44]。此外，它还具有强大的泛化能力，能够捕捉各种尺度图像的特征，增强其对不同大小和不同形状图像的适应性。

ConvNeXt 纯卷积神经网络模型主要有 ConvNeXt-T, ConvNeXt-S, ConvNeXt-B, ConvNeXt-L, ConvNeXt-XL 这五个版本。综合考虑临床上对模型精确度以及轻量性的需要，本章节选择 ConvNeXt-T 作为预训练模型。ConvNeXt-T 具有强大的特征提取能力并且参数量较小，在训练过程中对硬件条件的要求较低。ConvNeXt-T 结构如图 3-3 所示。

如图 3-3 所示，ConvNeXt-T 模型主要由卷积、Layer Norm、ConvNeXt Block 模块、下采样、全局平均池化和全连接层等串联而成。其中，ConvNeXt Block 模块是由卷积层，Layer Norm 和 GELU 激活函数等组成的倒残差结构。

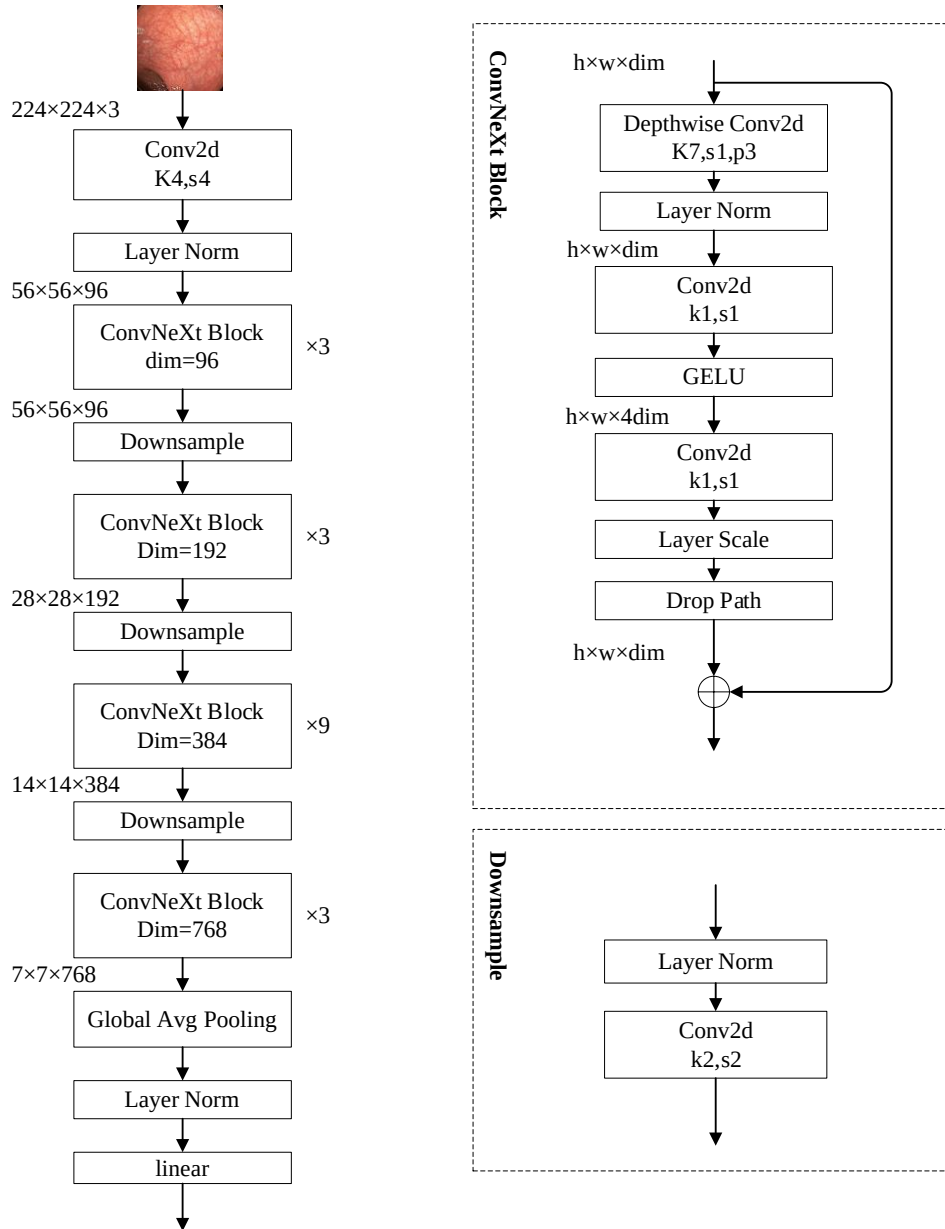


图 3-3 ConvNeXt-T 模型结构图

3.1.3 CA-ConvNeXt 模型

注意力机制能够使网络在处理输入图片时，对于重要的特征进行更集中的关注，忽略不相关或次要的特征，有助于提升网络在复杂任务上的表现。因此，在 CNN 中添加注意力机制可以提高网络的表现和识别准确率，使其更加智能和灵活。

传统的注意力机制有压缩激励注意力机制(Squeeze and Excitation, SE)^[45]，卷积模块的注意力机制(Convolutional Block Attention Module, CBAM)^[46]等。其中，SE 注意力机制通过学习每个通道的重要性加权来增强有用的特征。其名称中的 Squeeze 和 Excitation 分别指代了两个关键步骤：首先，通过全局池化操作

来挤压特征图，将每个通道的重要性进行汇总；然后，通过一个多层感知器结构，经过一系列非线性变换后生成最终的特征权重，SE 注意力机制如图 3-4 所示。

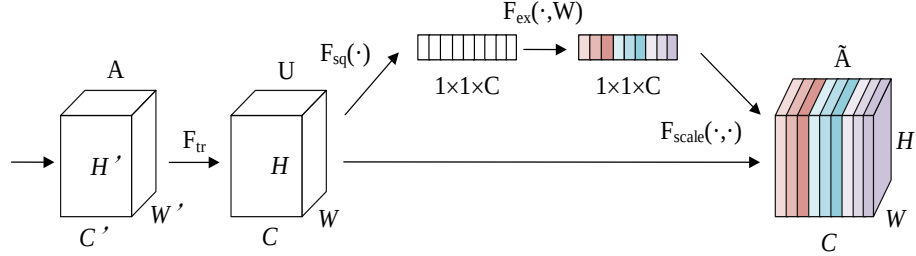


图 3-4 SE 注意力机制

图中的 F_{tr} 表示传统的卷积结构； $F_{sq}(\cdot)$ 表示执行全局平均池化操作； $F_{ex}(\cdot, W)$ 表示两级全连接操作；然后使用 sigmoid 激活函数得到权值，利用乘法操作将其作用到 U 的 C 个通道上。虽然 SE 模块在通道注意力方面表现出色，但其忽略了对空间维度的关注。相比之下，CBAM 模型则将通道与空间注意力相结合，使其既能考虑通道特性，又能考虑空间特性，其结构图如图 3-5 所示。

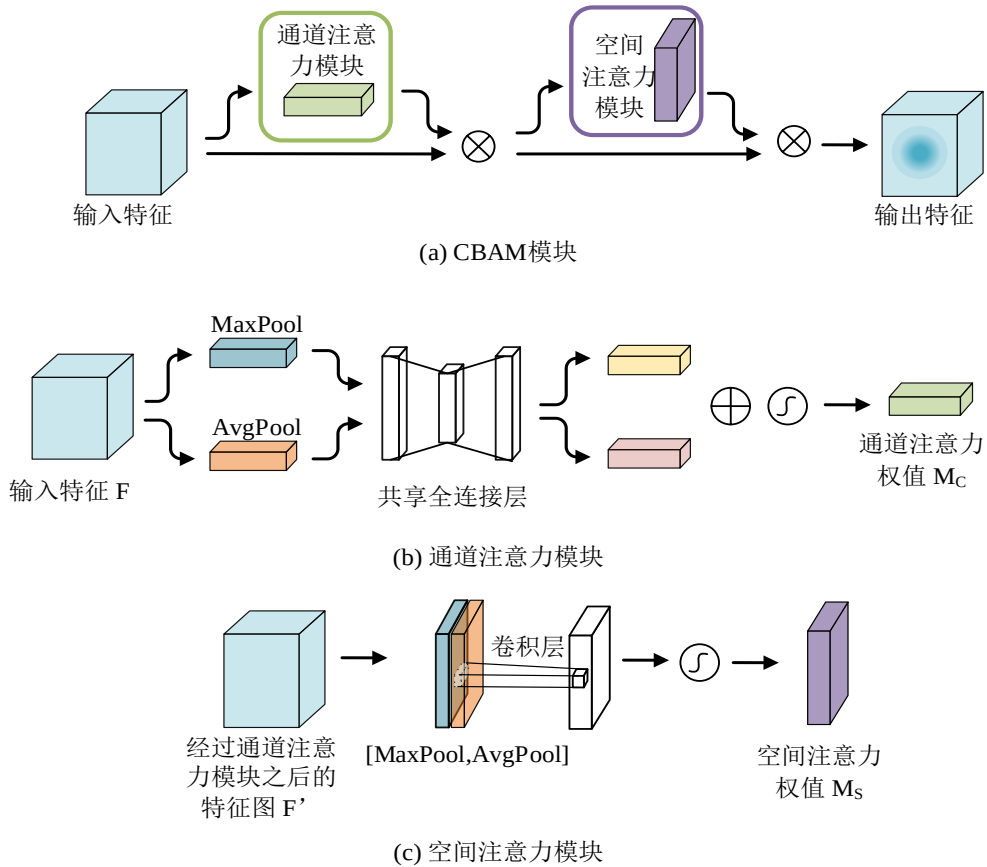


图 3-5 CBAM 注意力机制

CBAM 通过结合通道注意力和空间注意力，使网络更好地捕获图像特征之间的相互关系，从而提高网络在各种计算机视觉任务上的性能。通道注意力旨在对输入特征图的不同通道进行加权，以便网络可以更好地关注重要的特征通道，其权值 M_c 通过对特征图的每个通道进行平均池化和最大池化，然后经过共享全连接层后进行合并操作，最后再送入激活函数得到。空间注意力旨在对特征图的不同空间位置进行加权，以便网络可以更好地关注重要的空间区域。其权值 M_s 通过对输入特征图的每个通道分别应用平均池化和最大池化后合并，再经过 7×7 的卷积和函数激活得到。

CBAM 相对于 CA 注意力机制在一些特定任务中可能不能够捕捉到足够的位置敏感信息，尤其是对于需要更精确定位和识别的任务。CA 注意力机制能够同时捕捉到方向感知和位置敏感信息，这有助于提高模型对目标区域的关注程度，从而提高分类准确性^[47]。因此，在 WCE 图像分类任务中，添加 CA 注意力机制能够促使模型更加关注病变区域，减少不必要的背景关注，进而提高模型整体分类性能。

为了提高模型注意力的同时不破坏网络主干部分的权值，CA 模块被添加到 ConvNeXt 的非主干部分。具体而言，CA 模块被添加在最后一个 ConvNeXt Block 后，其结构如图 3-6 所示。其中：C 代表通道数；H 代表高度；W 代表宽度；r 是缩减系数，常取 16。

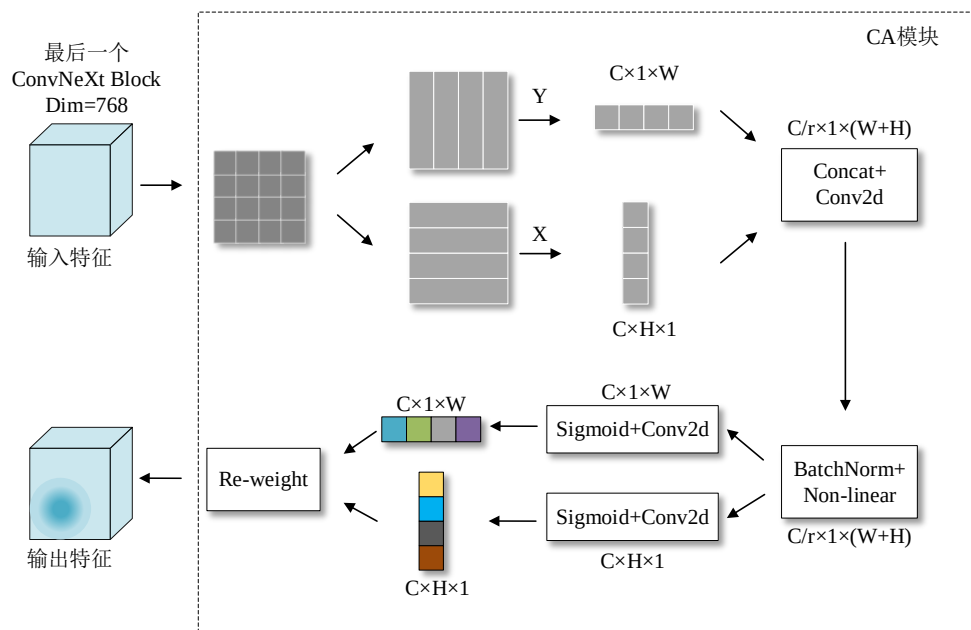


图 3-6 CA 集成到 ConvNeXt 网络中结构图

图 3-7 中 CA 模块主要实现步骤如下：1) 对最后一个 ConvNeXt Block 的特征图分别沿着水平和垂直方向进行一维平均全局池化，生成 $H \times 1 \times C$ 、 $1 \times W \times C$ 尺寸的特征向量组，将 $1 \times W \times C$ 转置并与 $H \times 1 \times C$ 进行拼接，使用 1×1 卷积对通道进行 r 倍压缩操作；2) 通过批归一化(Batch Normalization, BN)和非线性激活函数，对特征图进行水平和垂直的分割，使用 1×1 卷积进行 r 倍升维操作；3) 采用 Sigmoid 激活函数，提取注意力权值，并将其乘以输入的特征图，从而获得附加了关注权值的特征图。通过该特征图可以定位目标对象的位置，提升模型对物体的关注程度的同时提升分类精度。

3.1.4 迁移学习

CNN 的主要局限性是需要大量的数据构建稳健的模型。考虑到医学领域可用于模型训练和测试数据量是有限的，而迁移学习技术可提取通用特征，帮助提高模型的泛化能力和性能。因此，结合 WCE 数据集的样本量较小的实际情况，基于迁移学习训练所提模型。

在迁移学习中，使用预训练模型非常有效，由于不需要从头进行训练，节省了计算开销和训练时间。如果源数据集和目标数据集完全不同，则可以通过在原始层基础上添加或修改若干层微调预训练模型，以匹配数据集的特定特征。本章节将使用 ImageNet-1k 数据集^[48]上的预训练权重，同时使用更适合内窥镜图像分类的全连接层，即设置成 3 类别，微调后网络最后几层如图 3-7 所示。

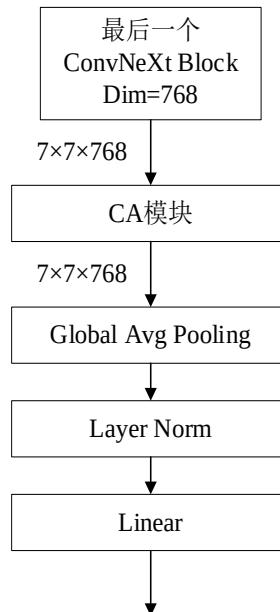


图 3-7 微调后的网络最后几层结构图

3.1.5 损失函数

为优化模型的训练过程，选用 Polyloss 损失函数训练模型。Polyloss 通过泰勒展开思想将损失函数表示为多项式函数的线性组合形式，结合交叉熵损失和多项式因子形成最终的损失函数。这样可以在训练过程中平衡各个类别的重要性，进而改善模型对样本量较少类别的识别能力^[42]。Polyloss 损失函数通过修改交叉熵损失函数(Cross Entropy loss)得到，其表达式如式(3-1)所示：

$$loss_{ce} = -y \log y' - (1-y) \log(1-y') \quad (5-1)$$

其中： y 是标签， y' 是预测值。

交叉熵损失函数的泰勒展开式如(3-2)所示。

$$L_{CE} = -\log(P_t) = \sum_{j=1}^{\infty} \frac{1}{j(1-P_t)^j} = (1-P_t) + \frac{1}{2}(1-P_t)^2 + \dots \quad (5-2)$$

其中： P_t 表示目标标签预测的概率。

Polyloss 损失函数仅仅修改了交叉熵损失中的第一个多项式系数，如式(3-3)所示：

$$L_{poly-1} = (1+\varepsilon_1)(1-P_t) + \frac{1}{2}(1-P_t)^2 + \dots = -\log(P_t) + \varepsilon_1(1-P_t) \quad (5-3)$$

其中： ε_1 是交叉熵损失中的首个多项式系数，为了获得最优的训练结果，需要根据不同的任务调参得到^[42]。

3.2 实验分析

为了验证 CA-ConvNeXt 模型在 WCE 图像分类领域的有效性，本章节使用 Hyper Kvasir 数据集中的部分图片，其中包含正常、息肉、溃疡各 500 张，共计 1500 张图像，划分训练集和验证集比例为 8:2。由于选取数据集较小，为避免过拟合情况的出现，对训练集使用数据增强技术，即将正常、息肉、溃疡图像数量扩增为各 600 张，共 1800 张图像用于模型训练。模型训练迭代次数设置为 50，批量大小设置为 4，使用的 AdamW 优化器其初始学习速率设置为 0.0005，weight_decay 设置为 0.005，使用的 Polyloss 损失函数中 ε_1 的值设置为 1。

混淆矩阵是评估和分析图像分类结果的常用方法，其结构如图 3-8 所示。在深度学习领域，混淆矩阵是一种常见的精确度评估方法，可将其划分为真实分类和预测分类两种。

		真实值	
		True	Flase
预测值	Positive	TP	FP
	Negative	FN	TN

图 3-8 混淆矩阵

其中：TP 表示被模型预测为正类的正样本；TN 表示被模型预测为负类的负样本；FP 被模型预测为正类的负样本；FN 被模型预测为负类的正样本。

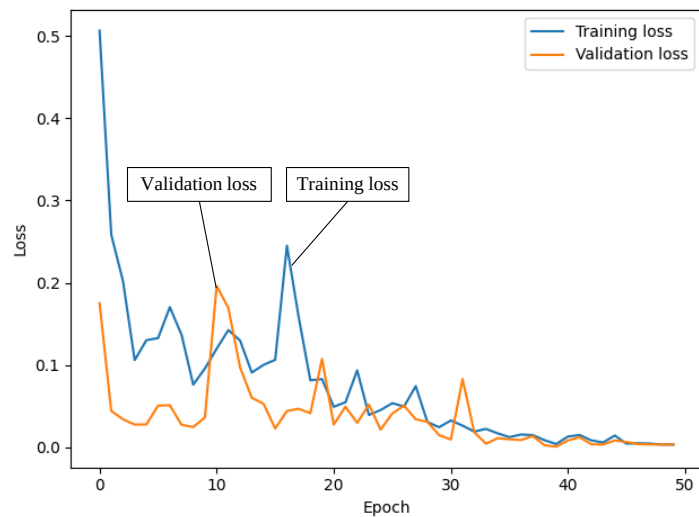
为了对内窥镜图像分类性能进行评估和分析，选用准确率(Acc)、精确率(Pre)、召回率(Re)和 F1 分数四个指标进行评价。其中精确率、召回率、准确率如式(2-7)、(2-8)、(2-9)所示，F1 分数评价指标计算式如(3-4)所示：

$$F1\text{分数} = 2 \times \frac{Pre \times Re}{Pre + Re} \quad (5-4)$$

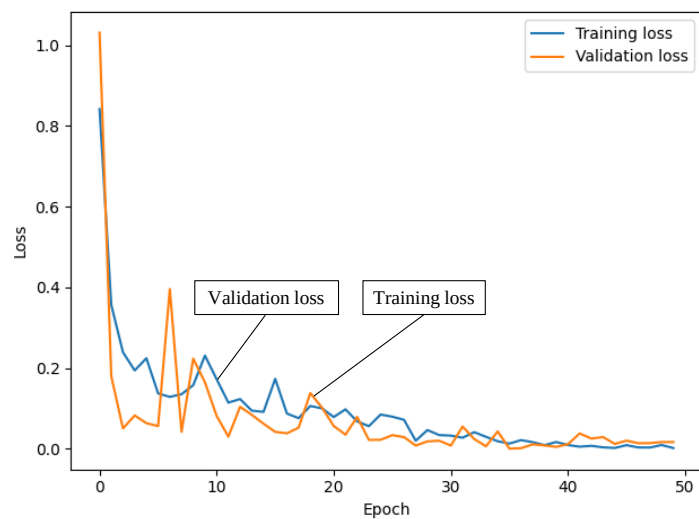
其中，准确率反映模型整体的分类准确程度；精确率反映模型预测为正例的准确性；召回率反映模型对于正样本的覆盖程度，F1分数由精确率和召回率二值计算而来，其反映模型在识别正例的情况下保持良好的准确性和覆盖范围，可更加全面的评估模型的识别能力。

3.2.1 不同损失函数对于改进模型的影响

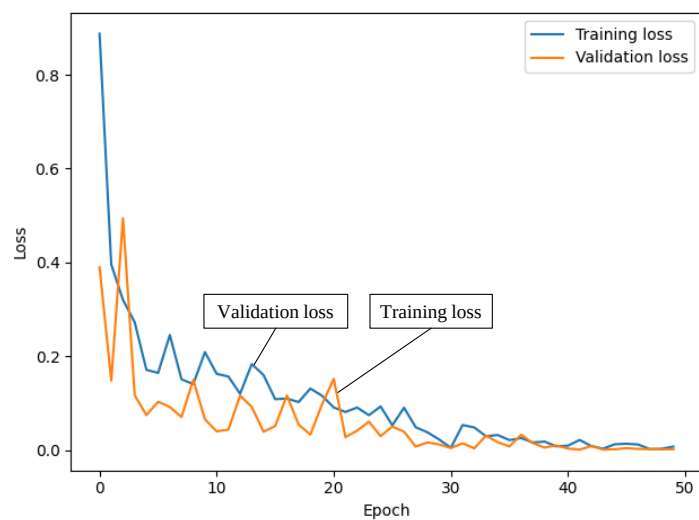
为了验证本章节所使用损失函数的优越性，在其他实验条件相同的情况下，分别采用 Polyloss 损失函数与经典的损失函数 Cross Entropy loss 和焦点损失(Focal loss)训练 CA-ConvNeXt 模型，并观察不同损失函数对模型训练过程的影响，损失函数曲线如图 3-9 所示。



(a) Cross Entropy loss



(b) Focal Loss



(c) Polyloss

图 3-9 不同损失函数的损失曲线

从图 3-9 可以看出，当模型收敛到稳定状态时，Polyloss 损失函数需要迭代 41 次左右，其他两个损失函数均需要迭代 45 次左右。在相同的训练条件下，Polyloss 损失函数用更少的迭代次数就能达到较好的拟合效果。由于 Polyloss 的设计结构，可能在较少的迭代次数后就能够找到最优解，因此只需要相对较少的迭代次数。另外，与 Cross Entropy loss 和 Focal loss 损失函数相比，Polyloss 损失函数在训练时损失鲁棒性更强。不同损失函数训练的评价指标如表 3-1 所示。

表 3-1 不同损失函数对模型的影响

损失函数		Cross Entropy loss	Focal loss	本章方法
总体	Acc	0.99	0.977	0.993
	Pre	0.99	0.99	1.00
	Re	1.00	1.00	1.00
正常	F1 分数	0.99	0.99	1.00
	Pre	0.98	0.98	1.00
	Re	1.00	0.96	0.98
息肉	F1 分数	0.99	0.97	0.99
	Pre	1.00	0.96	0.98
	Re	0.97	0.97	1.00
溃疡	F1 分数	0.98	0.96	0.99

从表 3-1 可得，polyloss 损失函数的各个评价指标均处于最佳水平，取得了最佳的分类结果，其中总体的分类准确率分别比 Cross Entropy loss 和 Focal loss 损失函数高出 0.3%和 1.6%。综上，采用 polyloss 损失函数可保证模型稳定训练的同时兼顾识别的准确率。

3.2.2 不同注意力机制对于改进模型的影响

为验证改进模型中 CA 模块的优越性，在相同训练条件下，使用 ConvNeXt 原模型与分别添加注意力机制 SE、CBAM 和 CA 的模型进行对比实验，各评价指标结果如表 3-2 所示。为可视化模型的重点关注区域，使用 Grad-CAM 可视化工具生成热力图来显示模型在决策过程中对输入 WCE 图像的不同区域的关注程度^[49]。Grad-CAM 可视化图能够更好地理解模型的决策依据和预测结果。四个模型的 Grad-CAM 可视化图如图 3-10 所示。

表 3-2 添加不同注意力机制后的比较

模型		ConvNeXt	ConvNeXt+SE	ConvNeXt+CBAM	本章方法
总体	Acc	0.977	0.977	0.983	0.993
	Pre	0.98	0.97	0.98	1.00
	Re	1.00	1.00	1.00	1.00
	F1 分数	0.99	0.98	0.99	1.00
正常	Pre	1.00	1.00	0.99	1.00
	Re	0.94	0.95	0.97	0.98
	F1 分数	0.97	0.97	0.98	0.99
	Pre	0.95	0.96	0.98	0.98
息肉	Re	0.99	0.98	0.98	1.00
	F1 分数	0.97	0.97	0.98	0.99
	Pre	0.95	0.96	0.98	0.98
	Re	0.99	0.98	0.98	1.00
溃疡	F1 分数	0.97	0.97	0.98	0.99
	Pre	0.95	0.96	0.98	0.98
	Re	0.99	0.98	0.98	1.00
	F1 分数	0.97	0.97	0.98	0.99

由表 3-2 可得知，相比于原模型，CA-ConvNeXt 模型总体准确率得到 1.6% 的提升，单独每个种类的各评价指标也得到一定的提升，这验证了改进模型在 WCE 图像分类任务上的有效性。添加 CA 注意力机制相对于添加 SE 和 CBAM 准确率分别提高了 1.6%、1.0%，这证明了 CA 模块的优越性。同时，所提模型在能够准确区分验证集上的健康和患病肠道，未出现将病变被识别为正常的情况，减少了漏检的可能性。

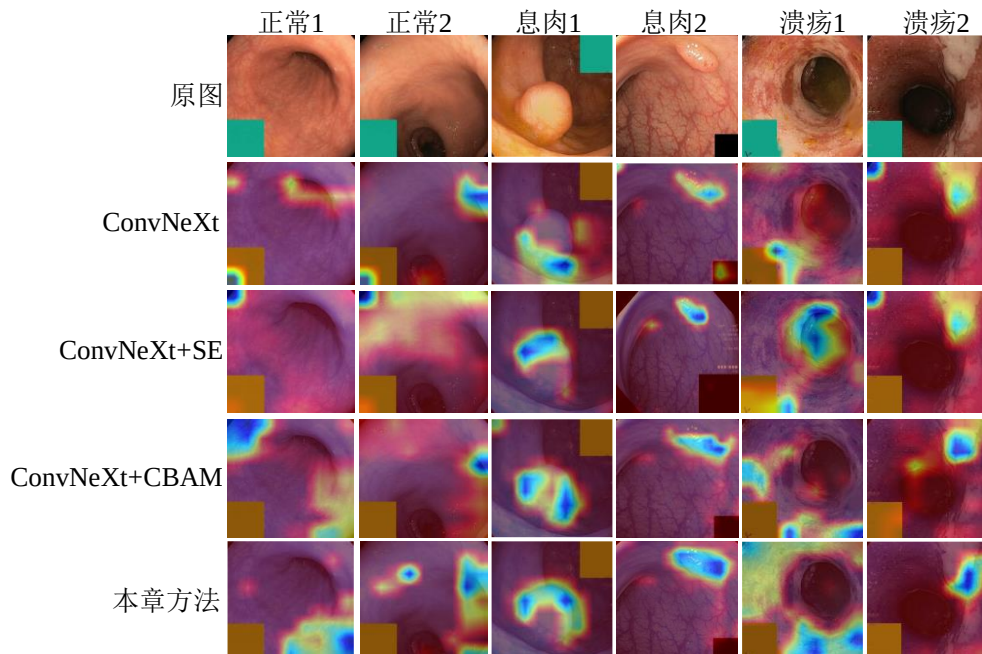


图 3-10 Grad-CAM 可视化

由图 3-10 可以看出相对于 CA-ConvNeXt 模型，其他三个模型对病变区域的关注度不够，并且更易受到不相关区域的影响，而 CA-ConvNeXt 模型可更准确地定位病灶，对较大范围的溃疡病灶也能较为精确地聚焦。Grad-CAM 可视化图进一步体现了所提模型的优越性。

3.2.3 与经典深度学习模型对比试验

为验证该模型的分类性能，将经典的分类网络模型 ResNet34 和 MobileNetV3 与 CA-ConvNeXt 模型进行了对比实验。使用相同条件训练模型，混淆矩阵如图 3-11 所示。

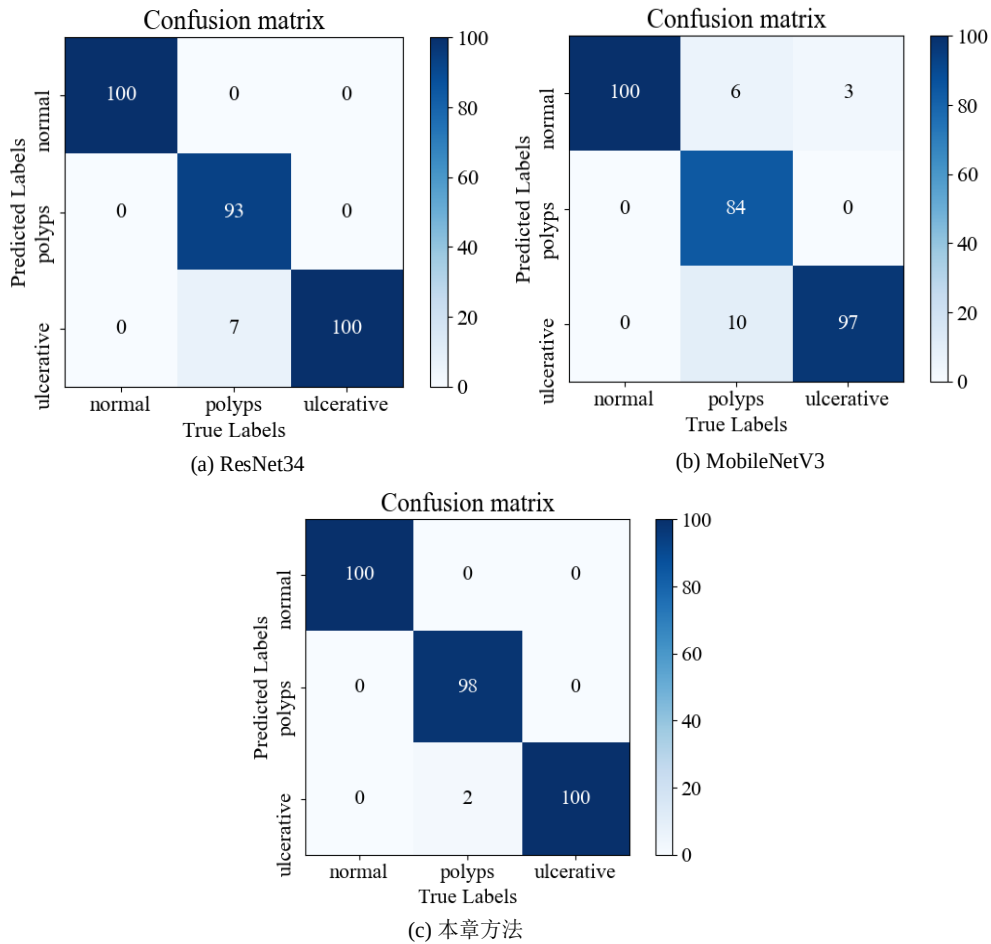


图 3-11 对比实验的混淆矩阵

由图 3-11 可见，ResNet34 模型将 7 张息肉图像误检为溃疡，不存在漏检的情况；MobileNetV3 模型将 10 张息肉图像误检为溃疡，同时漏检 6 张息肉图像和 3 张溃疡图像；所提模型仅仅将 2 张息肉图像误检为溃疡，不存在漏检的情况。说明了相对于 ResNet34 和 MobileNetV3 模型，本章节提出的 CA-ConvNeXt

模型漏检误检概率更低，更加满足临床上对分类精度的需要，具有一定的优越性。各评价指标如表 3-3 所示。

表 3-3 与经典深度学习模型对比试验

模型		ResNet34	MobileNetV3	本章方法
总体	Acc	0.977	0.937	0.993
	Pre	1.00	0.917	1.00
	Re	1.00	1.00	1.00
	F1 分数	1.00	0.96	1.00
正常	Pre	1.00	1.0	1.00
	Re	0.93	0.84	0.98
	F1 分数	0.96	0.91	0.99
	Pre	0.935	0.907	0.98
息肉	Pre	1.00	0.97	1.00
	Re	0.93	0.84	0.98
	F1 分数	0.96	0.91	0.99
	Pre	0.935	0.907	0.98
溃疡	Pre	1.00	0.97	1.00
	Re	0.93	0.84	0.98
	F1 分数	0.96	0.91	0.99
	Pre	0.935	0.907	0.98

从表 3-3 可以看出，CA-ConvNeXt 比其它两种经典的分类方法在 WCE 分类任务上表现更好，其准确率比 ResNet34 提高了 1.6%，比 MobileNetV3 提高了 5.6%。同时，对于各病灶的其它评价指标，CA-ConvNeXt 模型总体优于其它模型，这进一步说明 CA-ConvNeXt 模型在 WCE 图像分类中的应用中优势较大。

为了检验该模型的普适性，利用一个具有八个类别的 WCE 图像分类数据集 Kvasir^[50]对模型进行训练。该数据集的每个类别各有一千张图像，划分训练集和验证集比例为 8:2。分别使用 Accuracy 以及其他三个指标的宏平均作为 Kvasir 数据集的评价指标。同时，将 CA-ConvNeXt 模型与已有的基于深度学习的 WCE 图像分类模型作对比，得到的结果如表 3-4 所示。

表 3-4 与其他现有方法对比试验

方法	Acc	Pre	Re	F1 分数
文献[51]	0.974	0.972	0.973	0.972
文献[52]	0.946	0.876	0.874	0.872
文献[53]	0.934	0.937	0.934	0.934
本章方法	0.983	0.982	0.983	0.982

在表 3-4 中，可以清楚地看到，所提方法在四个关键评价指标上都展现出了卓越的性能。具体来说，CA-ConvNeXt 模型在准确率方面表现尤为突出，通过将其与文献[51]、文献[52]以及文献[53]进行比较，能够直观地看到其准确率分别得到 0.9%、3.7%和 4.9%的提升。这验证了该模型具有一定的普适性。

3.3 本章小结

针对 WCE 图像数据量少，类间差异小，类内差异大而影响分类准确率的问题，探讨了一种 CA-ConvNeXt 图像分类模型。首先，采用数据增强的方法对原数据集进行扩展；其次，结合迁移学习微调卷积神经网络 ConvNeXt，引入 CA 注意力机制加强模型注意力；最后，使用 Polyloss 损失函数优化模型的训练过程。通过对实验数据的分析，验证了该方法的有效性，为溃疡以及息肉的诊断提供了帮助。

第 4 章 基于深度学习的内窥镜图像分割

结直肠癌是一种潜在危害极大的癌症，通常起源于结肠息肉。结直肠癌的最初阶段是结肠内形成良性息肉，这些息肉一般是可逆的，并且在早期可以通过 WCE 检查发现。若良性息肉未得到及时检查和治疗，会增加发展成肿瘤的可能性。因此，定期进行筛查、早期发现和干预非常重要。准确的息肉分割可以帮助医生发现和切除异常组织，传统息肉分割需要临床医师对着 WCE 图像手动分割，操作起来费时费力，因此迫切需要人工智能的辅助诊断。

4.1 基于 Li-DeepLabV3+的轻量级息肉分割模型

近年来，尽管息肉分割算法的性能从不同方面得到了加强，并取得了一定的积极成果，但由于息肉分割在临床中具有较高要求，如何进一步提高息肉分割算法的精确性、轻量性和泛化性等值得探索。因此，将探讨一种轻量级息肉分割网络模型 Li-DeepLabV3+。该网络模型以 DeepLabV3+[54]为基本框架，优化的 MobileNetV2 网络作为主干网络；采用改进的简易金字塔池化模块 SPPM 加强特征提取；使用改进的统一注意力融合模块 UAFM 融合浅层特征和深层特征。基本框架 DeeplabV3+采用编码器-解码器结构实现分割，其模型结构如图 4-1 所示。

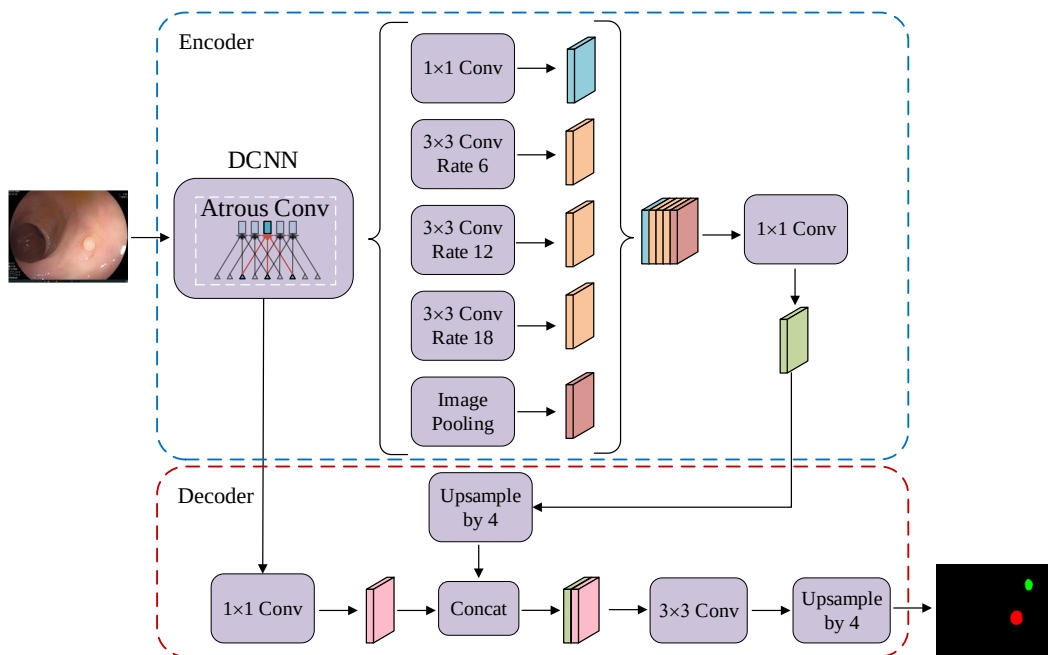


图 4-1 DeepLabV3+模型结构图

编码器负责提取浅层和深层语义信息，解码器进一步融合浅层和深层特征，提高分割边界的准确性，对不同像素的语义信息进行分类。在编码器中，DeepLabv3+模型使用 Xception^[55]作为骨干网络，从 Xception 中提取浅层和深层特征，并将深层特征输入到空洞空间金字塔池化(Atrous spatial pyramid pooling, ASPP)模块中。ASPP 模块由四个膨胀因子为 1、6、12 和 18 的卷积层和一个全局平均池化组成，其中 ConvX 表示卷积、batch norm 和 relu 操作，rate 表示空洞卷积的扩张率。ASPP 模块的膨胀卷积可以扩大感受野，以增强 Xception 深层特征的特征信息。在此基础上，Concat 操作将在通道维度上合并这五个特征，并通过 1×1 卷积层对特征进行压缩，得到深层特征图。在解码器中，使用 1×1 卷积层调整被压缩两次的浅层特征的通道数，并与上采样四次的深层特征图串联。堆叠后，通过 3×3 卷积对特征进行细化。最后，通过 4×4 线性插值得到具有原始图像分辨率的预测图像。

由于息肉分割在临床应用中模型的轻量性和精确性要求较高，将探讨一种轻量级息肉分割算法 Li-DeeplabV3+，其模型结构图由图 4-2 所示。

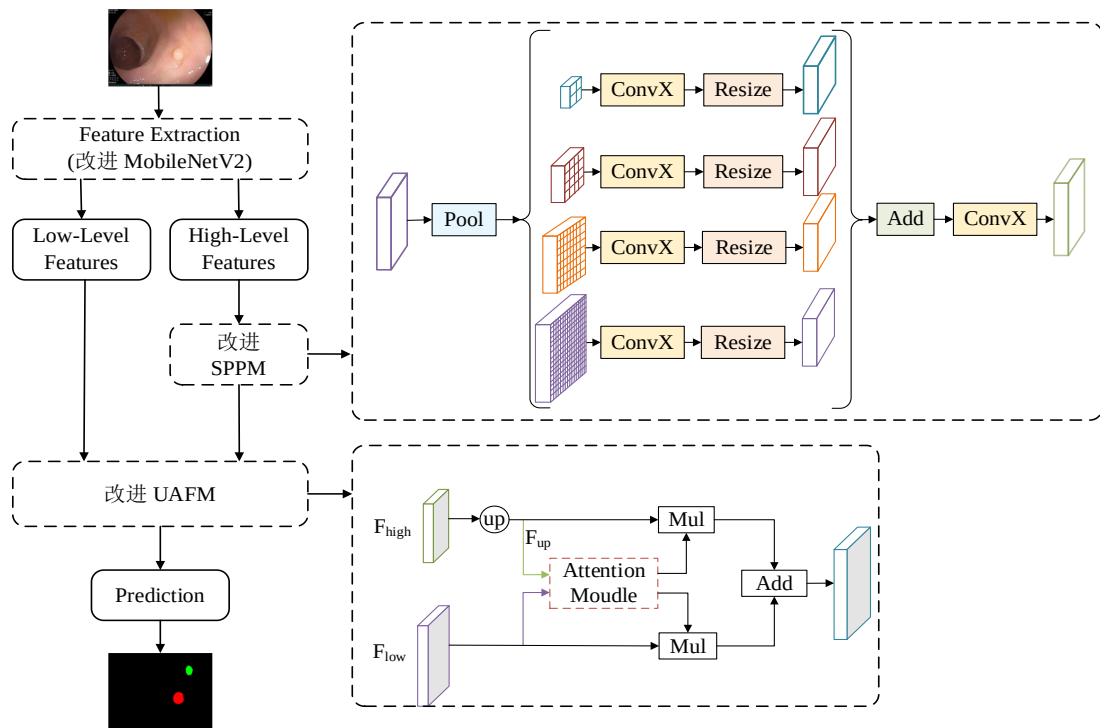


图 4-2 Li-DeeplabV3+模型结构图

如图 4-2 所示，WCE 息肉图像首先进入优化的 MobileNetV2 网络进行特征提取得到一个深层特征和一个浅层特征，深层特征被输入改进 SPPM 模块实现

加强特征提取；其次，浅层特征和经过改进 SPPM 的深层特征一同被输入改进的 UAFM 模块用以增强特征表示；最后，得到的特征图被输入预测模块从而实现息肉区域的分割。

4.1.1 主干网络的改进

由于传统 DeepLabV3+模型中使用的 Xception 骨干网络复杂度较高，不太适用于息肉分割。息肉分割在临床上除了需要较高的识别精度，对模型处理图片的速度也有一定的要求。基于此，引入轻量级的 MobileNetV2^[56]作为骨干网络实现特征提取。MobileNetV2 主要由倒残差结构(Inverted Residual Block)组成，在保持模型大小和计算效率的同时提供了更好的精度。倒残差结构如图 4-3 所示，首先利用 1×1 卷积升维，然后利用 3×3 深度卷积(Depthwise Convolution)进行特征提取，再利用 1×1 卷积降维；右侧是残差部分，输入和输出直接相连^[57]。

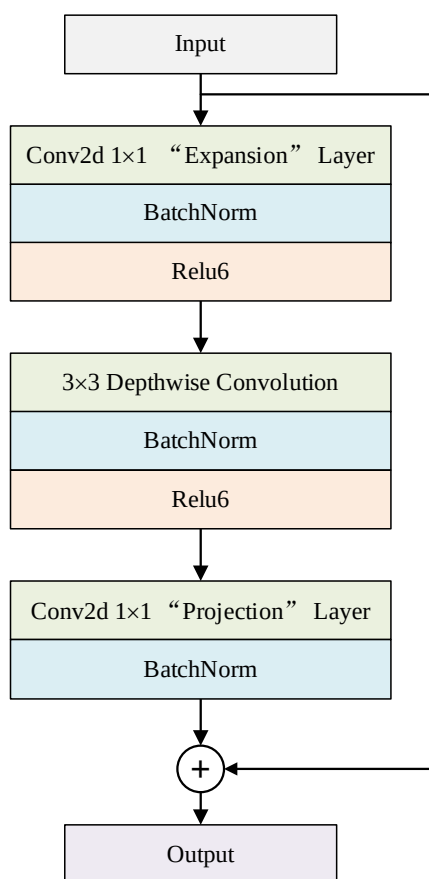


图 4-3 倒残差结构

图 4-3 中，深度卷积单独对输入数据的每个通道进行卷积运算，这种操作可以有效降低算法的复杂度，提高计算效率，并在保持良好性能的同时实现模型

的轻量化。**Relu6** 激活函数可在降低模型参数的前提下，增加数值分辨率，以减少有用网络信息的丢失。这些增强和调整提高了息肉语义分割的准确性和适应性，使模型更加实用。

MobileNetV2 原网络主要用于图像分类领域，若将其作为 **DeeplabV3+** 的主干网络需要进行剪枝操作，即删除原网络用于分类的最后三层；同时，为了保证图像分辨率和分割精度，修改原网络参数只进行四次下采样^[57]。以 $512 \times 512 \times 3$ 图像作为输入，剪枝后的 **MobileNetV2** 网络结构如表 4-1 所示。

表 4-1 剪枝后的 **MobileNetV2** 网络结构

不同网络模块层	升维倍数	输出数	使用次数	卷积步长	输出特征图尺寸
Conv2d 3×3	-	32	1	2	256×256×32
Bottleneck	1	16	1	1	256×256×16
Bottleneck	6	24	2	2	128×128×24
Bottleneck	6	32	3	2	64×64×32
Bottleneck	6	64	4	2	32×32×64
Bottleneck	6	96	3	1	32×32×96
Bottleneck	6	160	3	1	32×32×160
Bottleneck	6	320	1	1	32×32×320

4.1.2 特征加强模块的改进

由于 **DeeplabV3+** 模型的原有的 **ASPP** 模块计算量比较大，将采用改进的简易金字塔池化模块 **SPPM** 替换原有的空洞空间金字塔池化 **ASPP** 结构。**SPPM** 模块使用自适应池化操作收集上下文信息，使全局池化算子分别输出大小为 1×1 、 2×2 、 4×4 尺寸的特征图^[58]，其结构图如图 4-4(a)所示。为了在不增加太多参数量和尽可能少的减少图像细节损失的情况下收集上下文信息，原始的 **SPPM** 模块被改进，使全局池化算子分别输出大小为 2×2 、 4×4 、 8×8 和 16×16 的特征图，改进后的 **SPPM** 模块流程如图 4-4(b)所示。

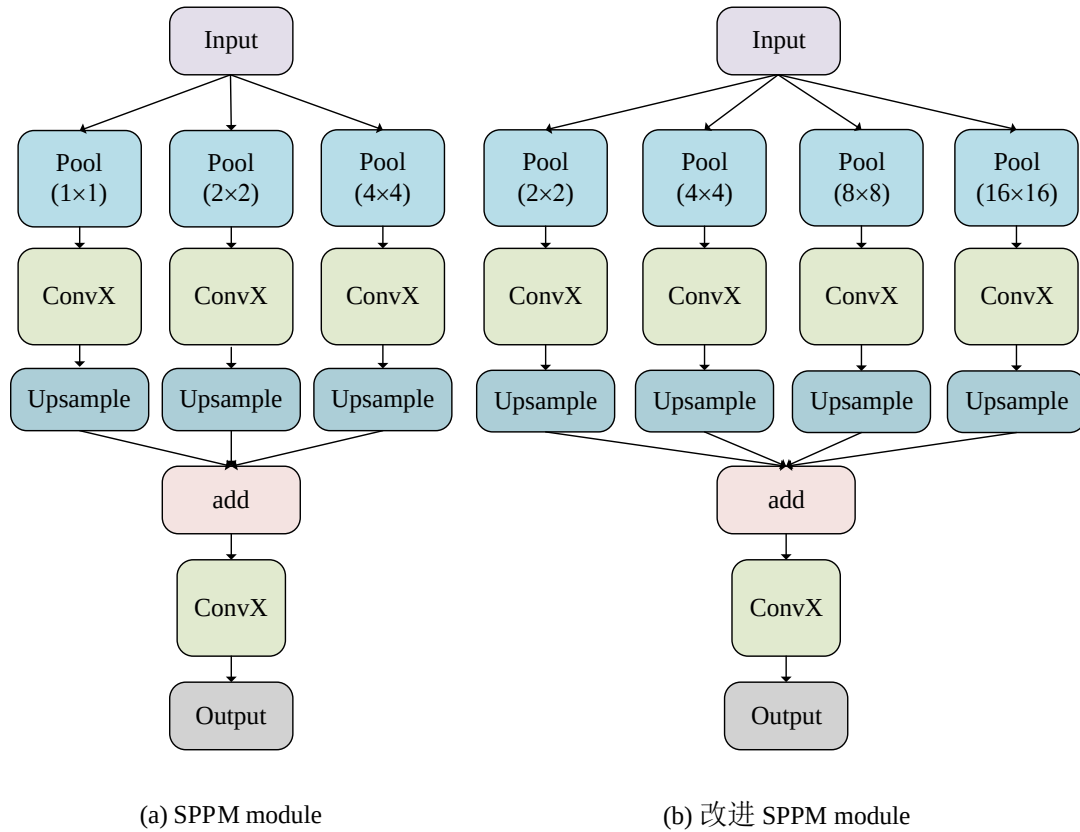


图 4-4 加强特征模块

改进的 SPPM 模块首先通过四个全局平均池化操作得到尺寸分别为 2×2 、 4×4 、 8×8 和 16×16 的特征图；然后分别对得到的四个特征图进行卷积和上采样操作；最后，将这些特征用 Add 操作结合起来，并应用卷积运算来生成细化的特征。与原始 SPPM 相比，改进的 SPPM 使图像细节损失的更少，因此息肉分割效果更好；与 ASPP 模块相比，改进的 SPPM 模块减少了参数量，更加满足模型对实时性的需要。

4.1.3 特征融合模块的改进

为了加强特征表示能力的同时进一步减少模型的参数，使用改进 UAFM 模块融合深层和浅层特征。传统的 UAFM 模块单独利用空间或通道注意力来产生权重，然后将输入特征与权重融合。为了更高效的融合特征，传统的 UAFM 模块被改进，同时利用空间与通道注意力，进一步增强特征的表示能力，实现息肉边界清晰分割。具体改进特征融合模块如图 4-5 所示。

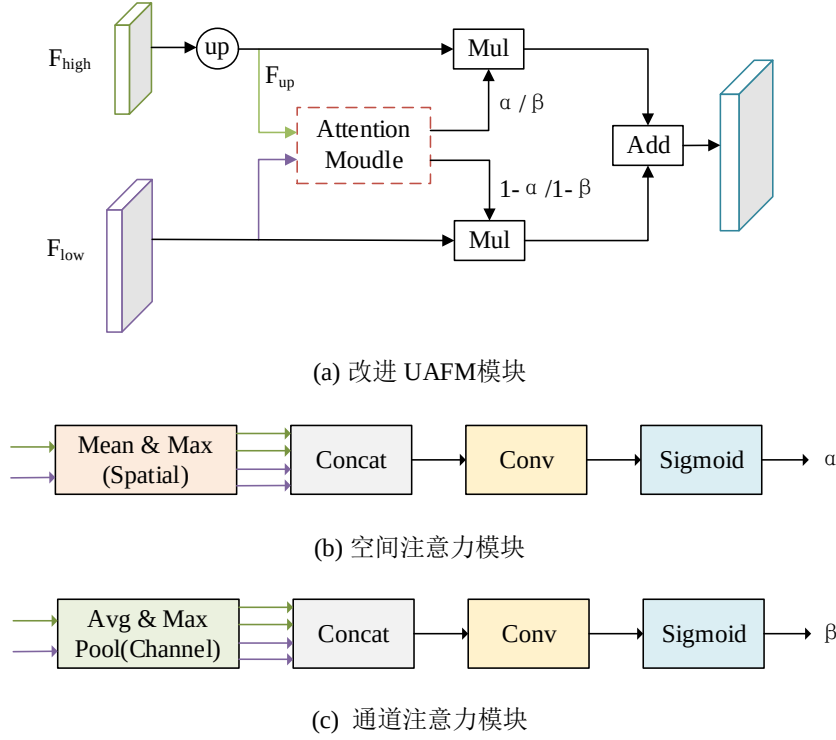


图 4-5 改进 UAFM 模块

图 4-5(a)中 F_{high} 是深层语义特征， F_{low} 是浅层语义特征，改进 UAFM 中的 Attention Moudle 有空间注意力模块和通道注意力模块两个模块，分别用图 4-5(b) 和图 4-6(c)表示。首先，对输入的 F_{high} 使用双线性插值运算使其与 F_{low} 大小相同，并用 F_{up} 表示；其次，空间注意力模块和通道注意力模块分别将 F_{up} 以及 F_{low} 作为输入生成权重 α 和 β ，并将得到的权重 α 与 F_{up} 以及 F_{low} 进行 Mul 操作后进行 Add 操作得到一个特征图，同理使用权重 β 得到另一个特征图；最后，将得到的两个特征图进行 Mul 操作。上述过程可以表示为式(4-1)-(4-4)。

$$F_{up} = \text{Upsample}(F_{high}) \quad (6-1)$$

$$\alpha = SP_Attention(F_{up}, F_{low}) \quad (6-2)$$

$$\beta = CH_Attention(F_{up}, F_{low}) \quad (6-3)$$

$$F_{out} = (F_{up} \cdot \alpha + F_{low} \cdot (1-\alpha)) \cdot (F_{up} \cdot \beta + F_{low} \cdot (1-\beta)) \quad (6-4)$$

其中： $SP_Attention$ 代表空间注意模块，利用空间中的关系来产生权重，该权重表示输入特征中每个像素的重要性，其式如(4-5)所示； $CH_Attention$ 代表通道注意力模块，利用通道之间的关系来生成权重，该权重表示每个通道在输入

特征中的重要性，其式如(4-6)所示。

$$\alpha = \text{Sigmoid}(\text{Conv}(\text{Concat}(\text{Mean}(F_{up}), \text{Max}(F_{up}), \text{Mean}(F_{low}), \text{Max}(F_{low})))) \quad (6-5)$$

$$\beta = \text{Sigmoid}(\text{Conv}(\text{Concat}(\text{AvgPool}(F_{up}), \text{MaxPool}(F_{up}), \text{AvgPool}(F_{low}), \text{MaxPool}(F_{low})))) \quad (6-6)$$

4.1.4 损失函数

为了缓解息肉和背景像素点分布不均匀给模型训练带来的影响，Focal Loss 和 Dice Loss 混合的损失函数被用来优化模型^[59]。混合函数结合了 Focal Loss 的类别加权和 Dice Loss 的像素级相似性度量的优点，使模型能够更好地处理类别不平衡问题，提高像素级预测准确性。其式如(4-7)-(4-10)所示。

$$TP_p(c) = \sum_{n=1}^N p_n(c) g_n(c) \quad (6-7)$$

$$FN_p(c) = \sum_{n=1}^N (1 - p_n(c)) g_n(c) \quad (6-8)$$

$$FP_p(c) = \sum_{n=1}^N p_n(c) (1 - g_n(c)) \quad (6-9)$$

$$\begin{aligned} L &= L_{Dice} + L_{Focal} \\ &= C - \sum_{c=0}^{C-1} \frac{TP_p(c)}{TP_p(c) + \alpha FN_p(c) + \beta FP_p(c)} \\ &\quad - \frac{1}{N} \sum_{c=0}^{C-1} \sum_{n=1}^N g_n(c) (1 - p_n(c))^2 \log(p_n(c)) \end{aligned} \quad (6-10)$$

其中： $TP_p(c)$ 、 $FN_p(c)$ 、 $FP_p(c)$ 是用预测概率分别计算类别 c 的真阳性、假阴性和假阳性； $p_n(c)$ 表示像素 n 为类别 c 的预测概率； $g_n(c)$ 表示像素 n 为类别 c 的 ground truth； C 表示需要分割的种类数，本章节需要分割息肉和背景，所以 C 取 2； α 和 β 是权衡假阴性和假阳性的参数，实验中设置均为 0.5； N 是输入 WCE 图像中的像素总数。

4.2 实验分析

为了验证 Li-DeepLabV3+模型的优越性以及泛化性，使用 CVC-ClinicDB^[60]、Kvasir-SEG^[61]和 BKAI-NeoPolyp^[62]三个数据集进行实验分析。其中，CVC-ClinicDB 和 Kvasir-SEG 两个数据集提供了息肉分割标签；BKAI-NeoPolyp 数据

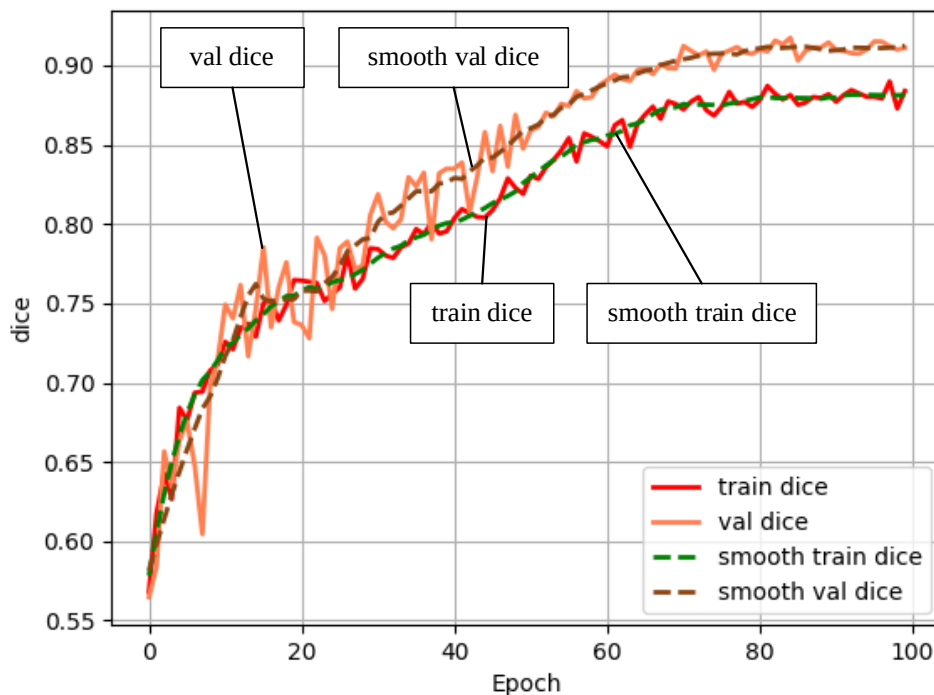
集提供了息肉肿瘤类和非肿瘤类的语义分割标签，分别用红色和绿色表示。

CVC-ClinicDB 数据集包含 612 张图像，其分辨率都是 384×288 像素，划分训练集、验证集、测试集的比例为 8:1:1。Kvasir-SEG 数据集包含 1000 张息肉图像以及对应的标签，其分辨率从 332×487 到 1920×1072 像素不等，训练时遵循官方拆分，使用了 880 张图像及其掩码来训练模型，而其余 120 张图像及其掩码用于测试。BKAI-NeoPolyp 数据集由 1200 张息肉图像组成，其中只有 1000 张拥有成对的标签，为便于实验对比，本章节使用 1000 张拥有成对标签的数据集进行实验，划分训练集，验证集，测试集的比例为 8:1:1。

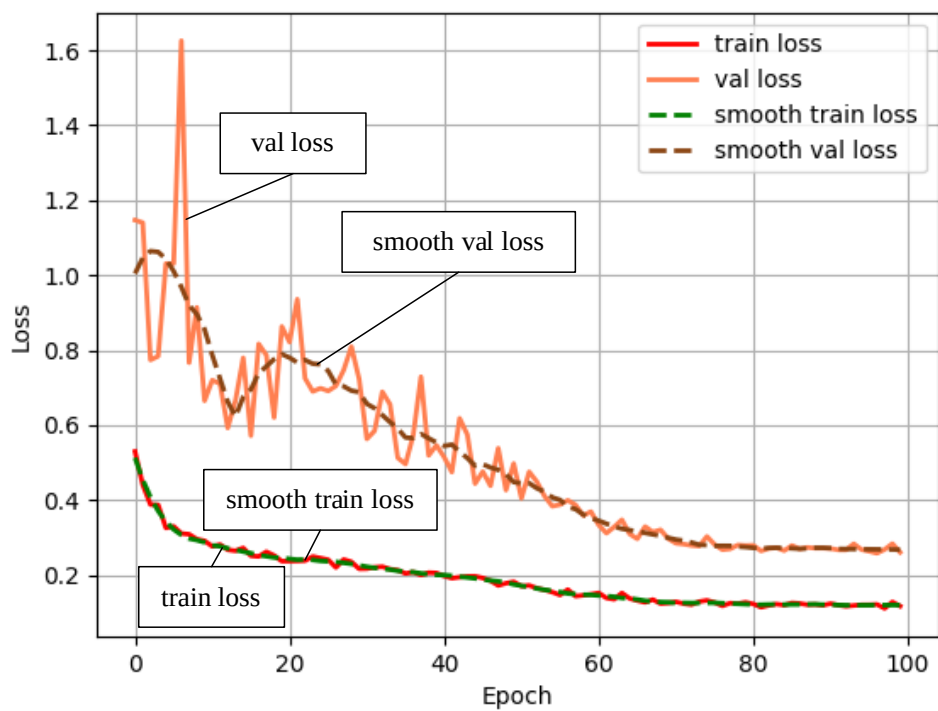
考虑到所用数据集中图像数量有限，迁移学习有助于解决训练时间过长和新任务标注数据不足的难题。因此，本章模型仍然结合迁移学习思想进行训练，即在 VOC 数据集^[63]的预训练权重的基础上进行训练。模型的训练次数设置 100，batch_size 设置为 4；使用 AdamW 优化器其初始学习速率设置为 0.0005；采用余弦退火的学习率下降方式。

4.2.1 迁移学习对于改进模型的影响

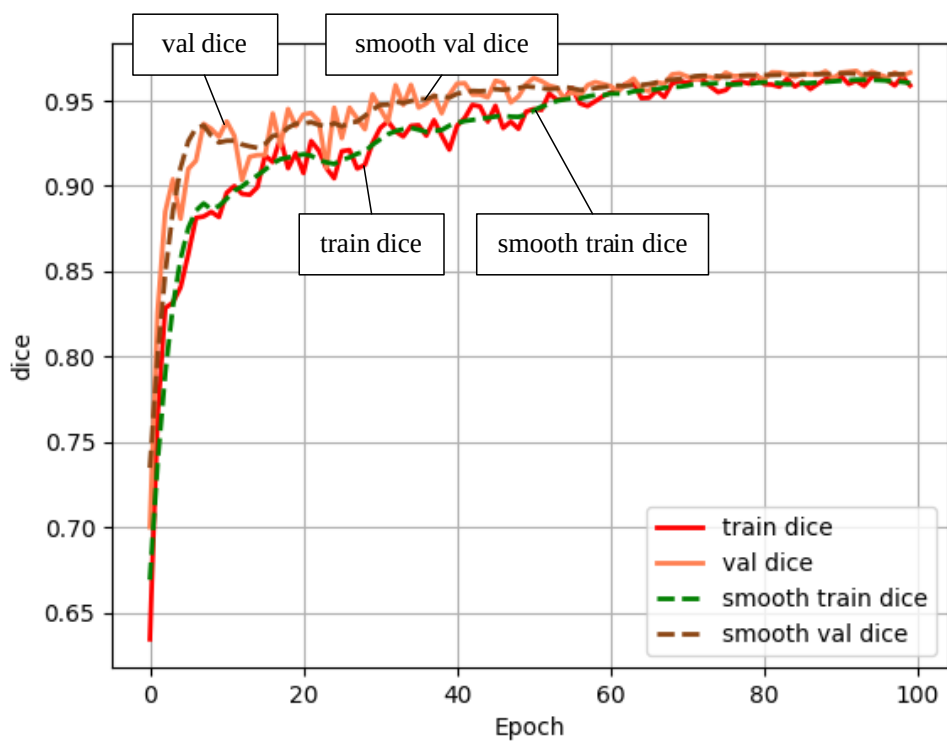
为了研究迁移学习对改进模型的影响，分别基于迁移学习和不使用迁移学习的两种训练方式在 CVC-ClinicDB 数据集上训练模型。不同训练方式的 Dice 曲线以及损失函数曲线如图 4-6 所示。



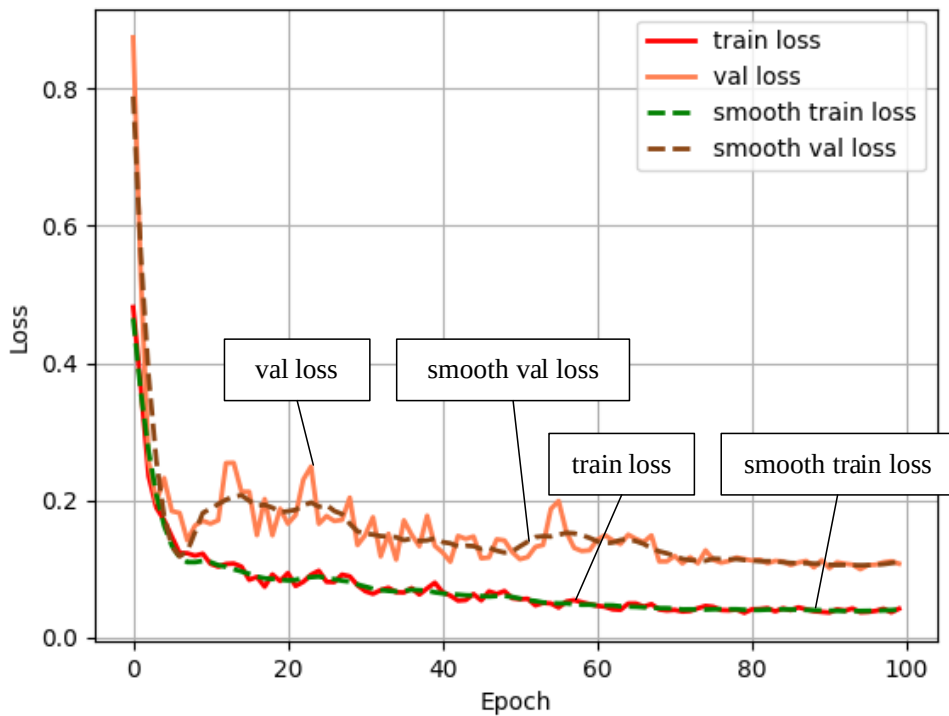
(a) 未使用迁移学习的dice曲线



(b) 未使用迁移学习的loss曲线



(c) 本章方法的dice曲线



(d) 本章方法的loss曲线

图 4-6 迁移学习对改进模型的影响

从图 4-6 可以看出，不使用迁移学习思想训练模型需迭代约 80 次，损失函数才收敛至稳定值，而基于迁移学习思想训练模型，损失函数在训练约 70 次时就趋于稳定。除此之外，使用迁移学习的思想在训练过程中损失值的变化更加稳定，说明通过利用预训练模型的权重，可以加速目标任务的训练过程，从而节省时间和计算资源。这是由于基于迁移学习的模型已学会一些通用的特征和模式，更易泛化到新的数据集上。表 4-2 展示了不同训练方式的分割性能。

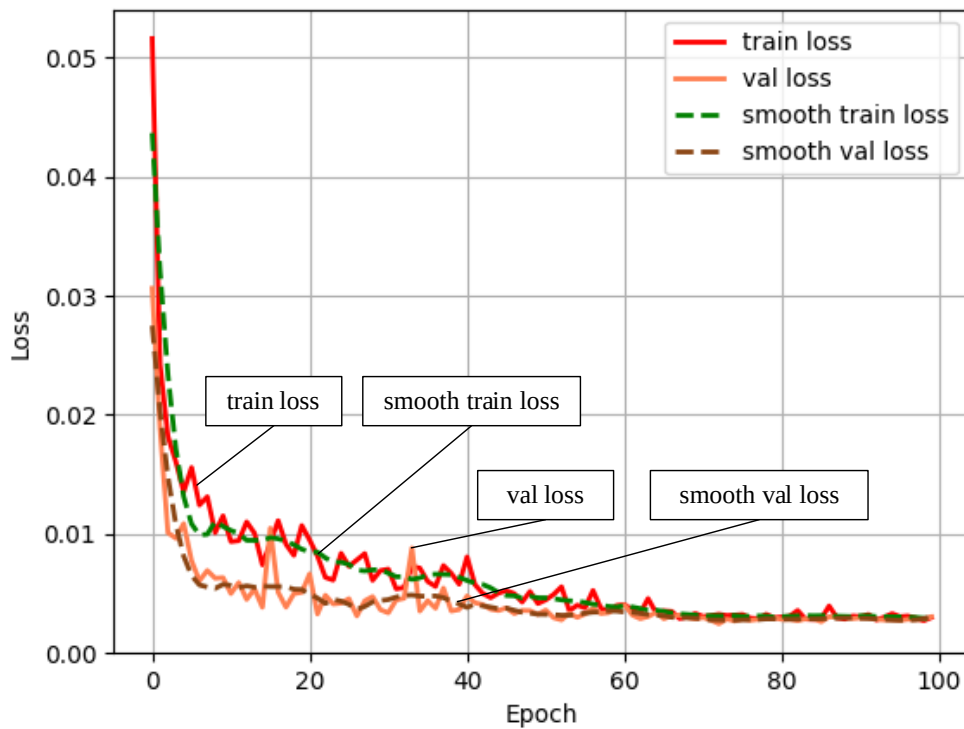
表 4-2 迁移学习对改进模型的影响

方法	IoU(%)	Dice(%)	Pre(%)	Re(%)	Acc(%)
未使用迁移学习	82.27	90.27	90.09	90.45	98.26
本章方法	87.42	93.28	93.05	93.54	98.80

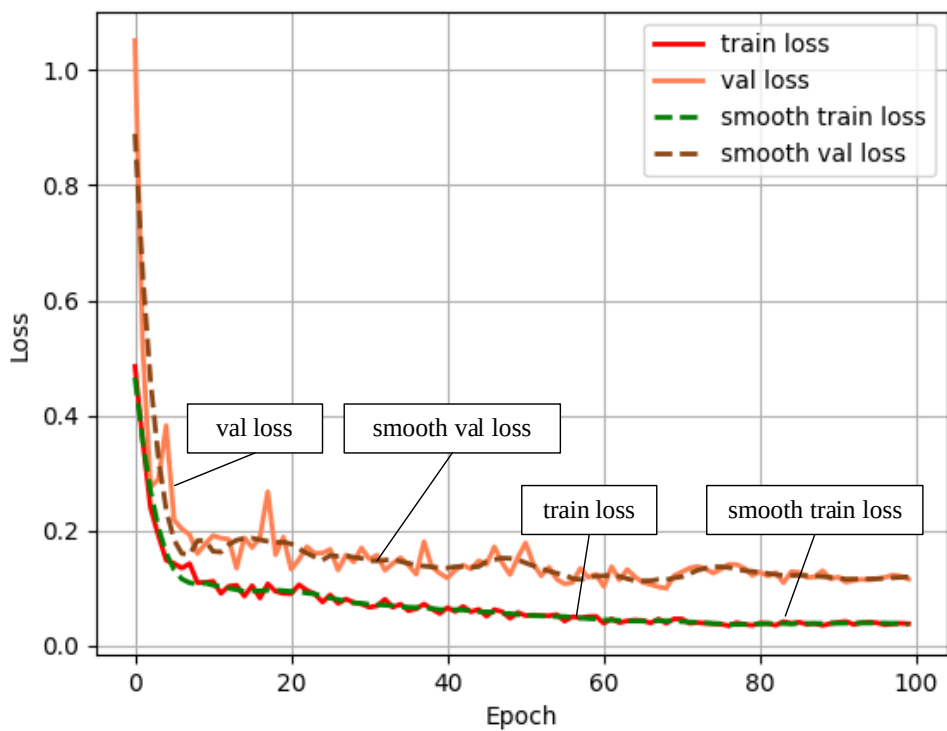
从表 4-2 可以看出，当训练方式为基于迁移学习时，各评价指标相对于未使用迁移学习都有较大的提升，从而进一步验证基于迁移学习不仅使训练效率得到提升，息肉分割效果也得到了极大改善。

4.2.2 不同损失函数对改进模型的影响

为了探索不同损失函数对 Li-DeeplabV3+模型分割性能以及训练过程的影响,在相同实验条件下分别使用 Focal loss^[64]、Dice loss^[65]、和 Dice loss + Focal loss 三种不同损失函数进行对比实验。各损失函数曲线如图 4-7 所示。



(a) Focal Loss



(b) Dice Loss

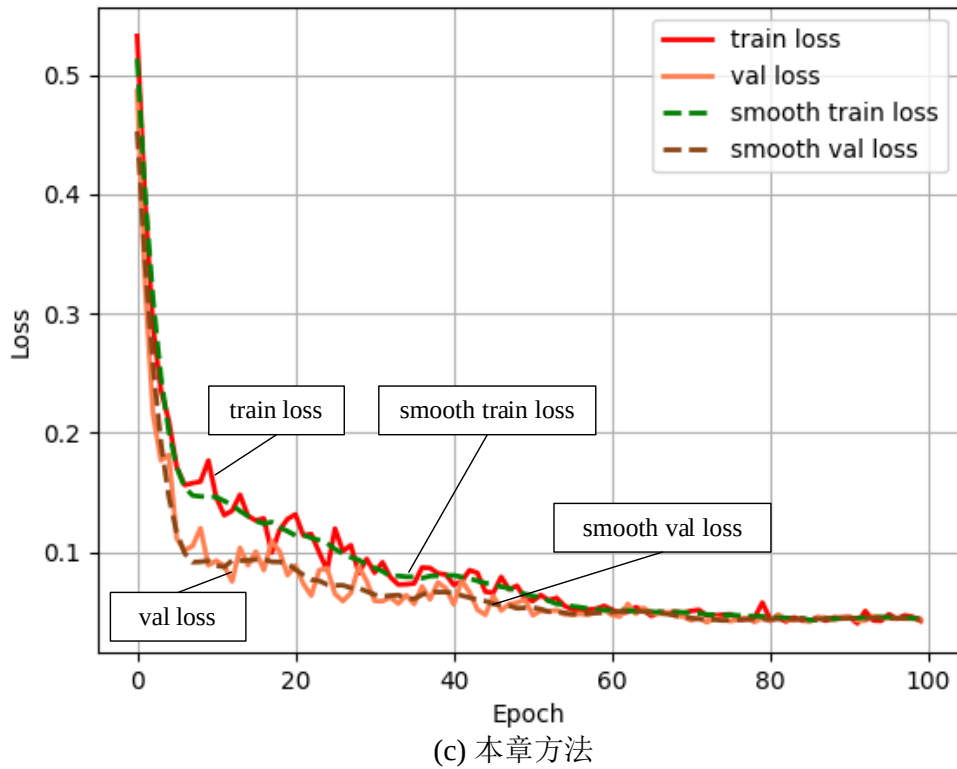


图 4-7 不同损失函数的损失函数曲线

从图 4-7 可以看出，使用混合损失函数和单独使用 Focal loss 时，模型在训练 60 个 epoch 时就趋于稳定，而使用 Dice loss 模型需要训练 80 个 epoch 才能达到稳定。由于本文数据集中样本的实际情况为：息肉和背景像素点数量不平衡。Focal loss 专注于处理类别不平衡的问题，但是训练过程中 dice 值波动较大。在使用 Dice loss 训练模型时，可重点关注息肉区域信息的挖掘，但由于息肉区域面积相对背景面积较小，分割结果并不十分优越。而混合损失函数结合了这两个损失函数各自的特性，对网络训练时不易学习到的样本使用稳定的、针对性强的方法进行优化，更有效地解决了样本不平衡，目标样本较小等问题。各评价指标如表 4-3 所示。

表 4-3 不同损失函数在改进模型上的性能对比

Methods	IoU(%)	Dice(%)	Pre(%)	Re(%)	Acc(%)
focal[64]	86.21	92.59	92.56	92.63	98.68
dice[65]	84.93	91.85	91.37	92.34	98.54
本章方法	87.42	93.28	93.05	93.54	98.80

从表 4-3 可以看出, 使用 Dice loss+Focal loss 损失函数的评价指标均高于两种单独使用的损失函数, 其中 IoU 较 Focal loss 和 Dice loss 分别提升 1.21%、2.49%, 说明混合损失函数更适用于 WCE 息肉图像分割任务, 可以有效解决正负样本不平衡的问题。

4.2.3 消融实验

消融实验是指通过系统地剥离模型中的某些组件或特性, 以便更好地理解其对模型整体性能的贡献。消融实验有助于验证模型的假设和设计选择, 确保模型的各个部分都对最终性能有正面影响。同时, 通过消融实验, 可以深入了解模型在不同条件下的实验效果, 从而揭示模型内部的决策过程和推理机制。

为了验证本文改进主干网络、特征加强模块以及特征融合模块的有效性, 分别选取息肉分割数据集集中的 CVC-ClinicDB 数据集以及息肉语义分割 BKAI-NeoPolyp 数据集作为代表进行训练, 分别反映本章节模型在普通分割和语义分割任务上消融实验的结果。主要观察主干网络、特征加强模块以及特征融合模块对分割性能和速度的影响。性能评价指标结果如表 4-4; 分割性能的可视化结果如图 4-8 (其中: Image 1-3 属于 CVC-ClinicDB 数据集, Image 4-6 属于 BKAI-NeoPolyp 数据集); 模型的参数量以及 FPS 如图 4-9 (其中: Params 表示网络模型中需要训练的参数总数, 单位 M 代表 Million、FPS 表示每秒传输帧数)。

从表 4-4 中可以看出与使用骨干网络 Xception 相比, 当骨干网络使用 MobileNetV2 时, 虽然在两个数据集上网络的 IoU 以及 Dice 参数略低, 但是模型训练参数少了将近十倍的同时 FPS 也增加大约两倍。考虑到息肉分割对模型分割速度的需要, 骨干网络 MobileNetV2 更加适用。同时, 从图 4-8 可以看出方法 2 的 Image 5 息肉分割边缘出现毛刺, 存在分割不准确的问题, 因此模型需要进一步改进。

对于特征加强模块, 当原有的特征加强模块 ASPP 被替换为 SPPM 模块时, 模型的参数再次减少 1.87M, 同时 FPS 达到 89.92, 但是模型的分割效果较差。因此 SPPM 模块被改进, 虽然使用改进 SPPM 模块的 FPS 减少 2.28, 但是模型在参数量仅仅增加了 0.082M。对于分割性能, 在 CVC-ClinicDB 数据集上模型的 IoU 以及 Dice 参数分别提升了 8.56%, 5.21%, 在 BKAI-NeoPolyp 数据集上模型的 IoU 以及 Dice 参数分别提升了 6.19%, 5.00%。图 4-8 中也能看出方法 4 的分割效果优于方法 3, 这体现出改进 SPPM 的有效性。

表 4-4 消融实验

数据集	方法	IoU(%)	Dice(%)	Pre(%)	Re(%)	Acc(%)
CVC-ClinicDB	方法 1 Xception+ASPP	86.64	92.84	92.17	93.52	98.71
	方法 2 MobileNetV2+ASPP	85.97	92.46	92.02	92.89	98.65
	方法 3 MobileNetV2+SPPM	77.05	87.04	85.48	88.66	97.64
	方法 4 MobileNetV2+ 改进 SPPM	85.61	92.25	90.98	93.54	98.63
	方法 5 MobileNetV2+ASPP+UAFM	86.81	92.91	93.15	92.73	98.74
	方法 6 MobileNetV2+ASPP+改进 UAFM	87.05	93.07	92.82	93.33	98.76
	本章方法	87.42	93.28	93.05	93.54	98.80
BKAI-NeoPolyp	方法 1 Xception+ASPP	77.78	86.19	85.45	87.58	98.88
	方法 2 MobileNetV2+ASPP	76.88	85.29	84.96	85.87	98.87
	方法 3 MobileNetV2+SPPM	71.77	81.32	82.56	80.58	98.41
	方法 4 MobileNetV2+改进 SPPM	77.96	86.32	87.74	84.97	98.85
	方法 5 MobileNetV2+ASPP+UAFM	75.55	84.33	84.33	84.52	98.81
	方法 6 MobileNetV2+ASPP+改进 UAFM	77.50	85.91	87.20	84.68	98.91
	本章方法	79.12	87.09	85.92	88.78	98.98

对于特征融合模块，当模型原有的特征融合模块被 UAFM 模块替换，模型参数下降了 0.482M，并且准确率有所提升。为了进一步提升分割效果，对 UAFM 模块进行改进，使用改进的 UAFM 在参数量不变的情况下，模型在 CVC-ClinicDB 数据集上的 IoU 以及 Dice 系数分别提高到 87.05% 和 93.07%，在 BKAI-NeoPolyp 数据集上的 IoU 以及 Dice 系数分别提高到 77.50% 和 85.91%，这证明改进 UAFM 的优越性及有效性。

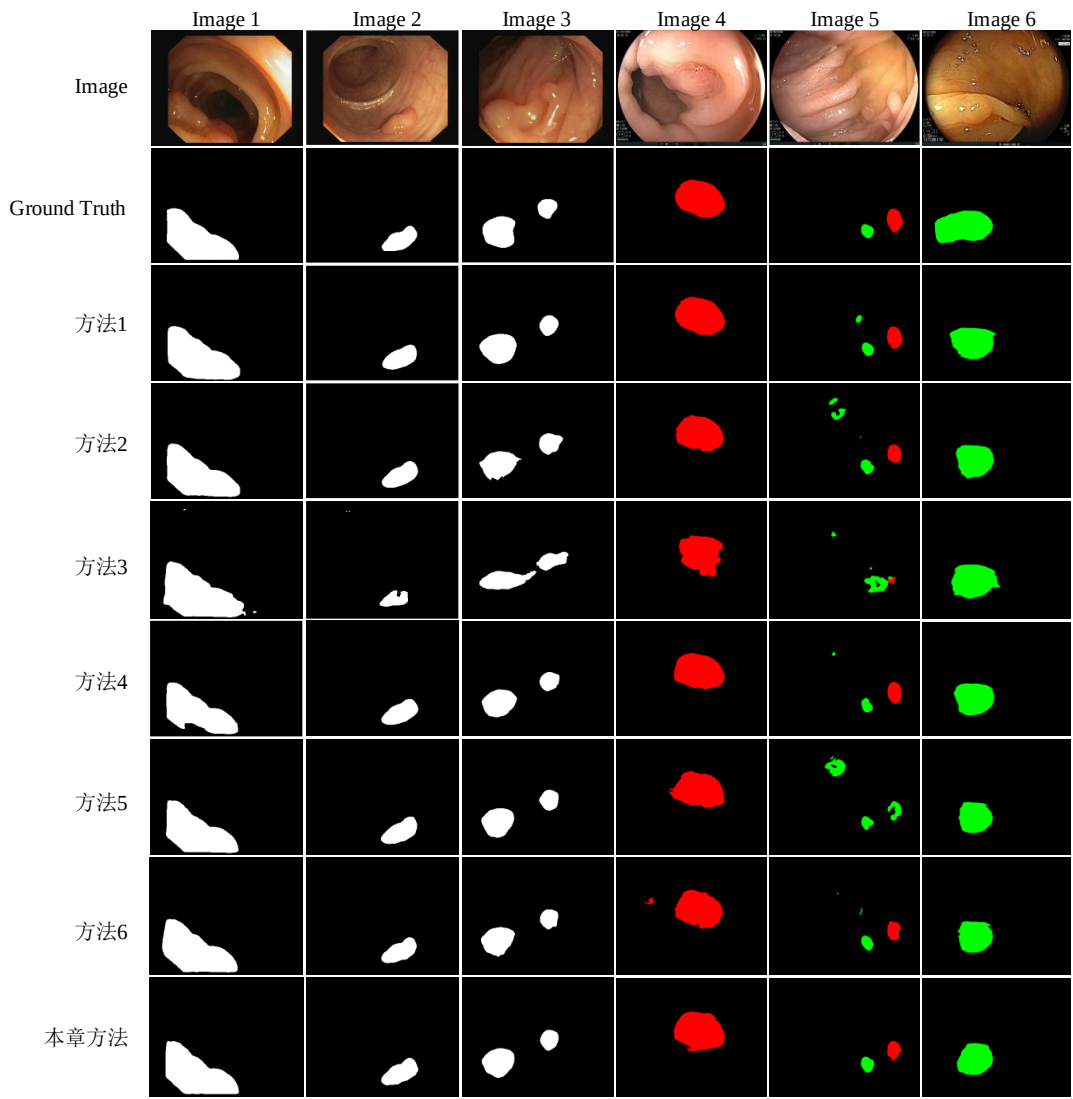


图 4-8 消融实验

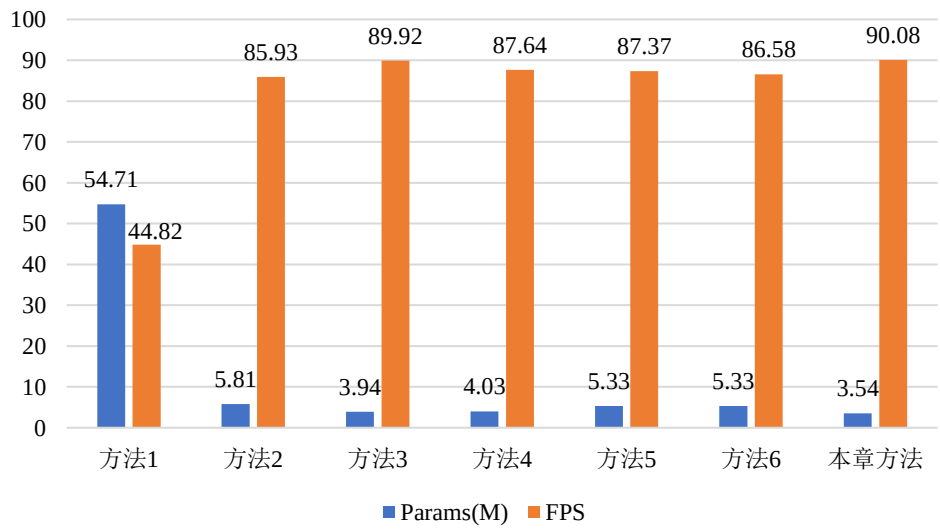


图 4-9 消融实验模型参数以及 FPS 比较

当改进 UAFM 和改进 SPPM 模块同时被使用时，改进模型较原始模型参数减少了 15.4 倍，FPS 增加了两倍，模型在 CVC-ClinicDB 数据集的 IoU、Dice 系数以及准确率分别达到了 87.42%、93.28%、98.80%，Recall 和 Precision 达到了均衡状态，同时在 BKAI-NeoPolyp 数据集上各评价指标也基本达到最佳状态。从图 4-8 中仍然能体现所提模型的优越性，分割轮廓更加接近 Ground Truth，因此 Li-DeeplabV3+在满足息肉分割速度需求的同时能实现更高的分割精度。

4.2.4 定量分析

为了验证 Li-DeepLabV3+算法的优越性，使用 Kvasir-SEG 和 CVC-ClinicDB 数据集在三个分割算法 Unet^[36]、SegFormer^[67]、ConvSegNet^[68]进行了对比实验，同时还进行了两个跨数据集实验探索改进模型的泛化能力，图像大小调整为 256×256，并且在相同条件下进行训练。分割性能的可视化结果如图 4-10 所示。

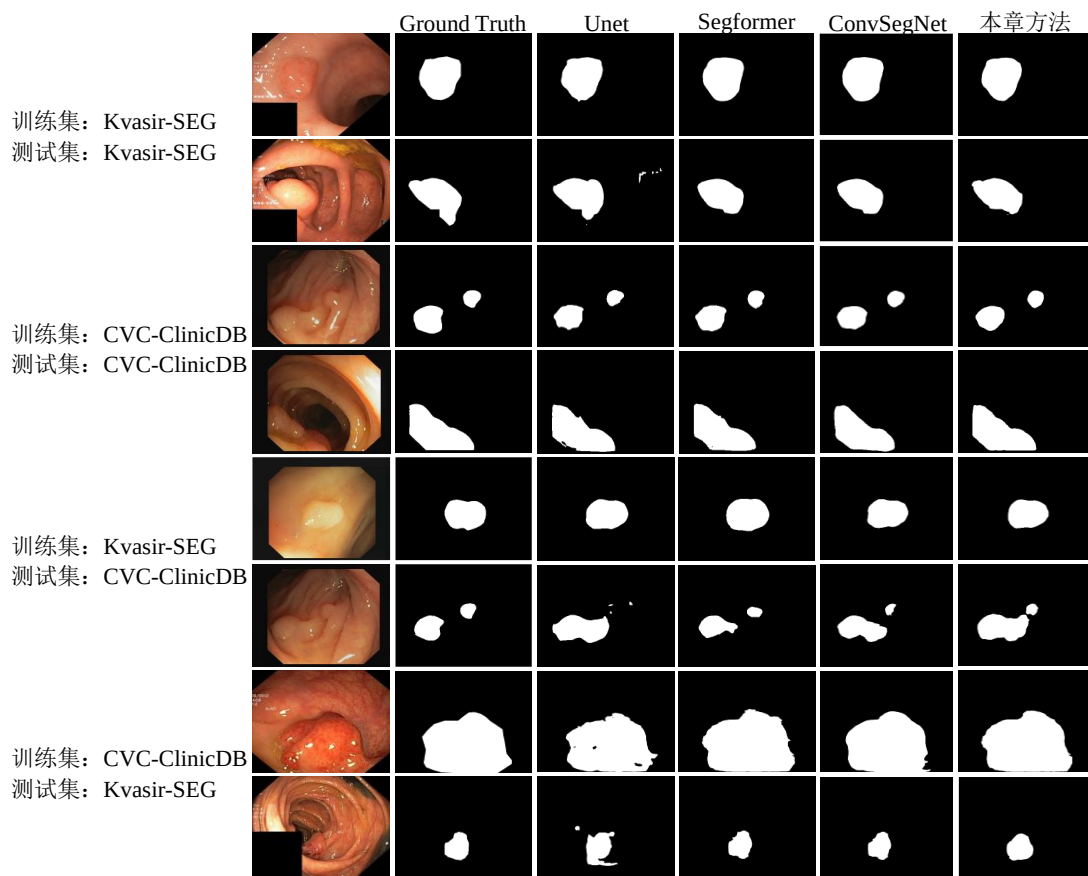


图 4-10 不同数据集效果图

从图 4-10 可以看出 Li-DeepLabV3+模型在两个数据集上的分割结果比其他模型更接近 Ground Truth，该模型面对相似的背景区域和噪声表现更出色，能够

更彻底地抑制噪声。同时还展现出更强的鲁棒性，不受目标区域的形状和大小变化影响。对于跨数据集试验，改进模型较其他模型而言息肉边界的分割更加细腻，具有更强的泛化性。实验的分割评价指标如表 4-5 所示。

表 4-5 定量分析

数据集	方法	IoU(%)	Dice(%)	Pre(%)	Re(%)	Acc(%)
训练集: Kvasir-SEG 测试集: Kvasir-SEG	Unet	73.84	84.95	88.17	81.11	95.38
	SegFormer	76.94	86.97	86.07	87.89	96.51
	ConvSegNet	79.44	86.30	86.23	91.03	95.72
	本章方法	79.88	88.82	88.42	89.22	96.56
训练集: CVC-ClinicDB 测试集: CVC-ClinicDB	Unet	84.67	91.70	94.50	89.06	98.74
	SegFormer	77.58	87.37	89.55	85.30	97.80
	ConvSegNet	83.45	90.27	89.64	92.58	98.79
	本章方法	87.42	93.28	93.05	93.54	98.80
训练集: Kvasir-SEG 测试集: CVC-ClinicDB	Unet	61.18	75.92	80.08	72.16	96.44
	SegFormer	67.69	80.73	83.38	78.25	96.66
	ConvSegNet	68.35	77.87	79.64	81.41	96.98
	本章方法	70.95	83.00	83.84	82.20	96.99
训练集: CVC-ClinicDB 测试集: Kvasir-SEG	Unet	66.66	79.99	81.52	78.53	94.38
	SegFormer	67.20	80.38	82.13	78.71	94.50
	ConvSegNet	61.63	71.55	69.18	87.89	90.13
	本章方法	68.58	81.36	83.74	79.35	94.79

由表 4-5 可以看出，对于 Kvasir-SEG 数据集，改进模型的 IoU、Dice 系数、召回率和准确率指标均为最佳结果，分别达到了 79.88%、88.82%、88.42%和 96.56%的水平。由于精确率和召回率是相互影响的关系，改进模型的召回率较 ConvSegNet 模型低 1.81%，但是其精确率达到 88.42%，表明该模型的误检的可能性更小。对于 CVC-ClinicDB 数据集，改进模型的各项评价指标优于所有的基准测试模型，精确率与召回率也达到了平衡状态，具有较强的临床应用潜力，其中 IoU 比 Unet、SegFormer，ConvSegNet 分别高出 2.75%、9.84%、3.97%，体现

改进算法的优越性。

对于从 Kvasir-SEG 到 CVC-ClinicDB 数据集的跨数据集实验，与最具竞争力的 ConvSegNet 模型相比，改进的模型 IoU、Dice 系数、精确率、召回率、准确率分别高出 2.6%、5.13%、4.2%、0.79%、0.01%，从图 4-10 的分割结果可以看出漏检的可能性较其他模型更小。对于在 CVC-ClinicDB 上训练模型后在 Kvasir-SEG 上测试的实验，改进的模型准确率和 Dice 分别达到 94.79%、81.36%。由于召回率和精确率互相影响的关系，召回率要比 ConvSegNet 低 8.35%，但是精确率比其高 14.56%，体现所提模型具有一定的泛化能力。

为了探索改进模型的轻量性和实时性，在相同的实验条件下将以上模型的参数量以及 FPS 做出对比，结果如图 4-11 所示。

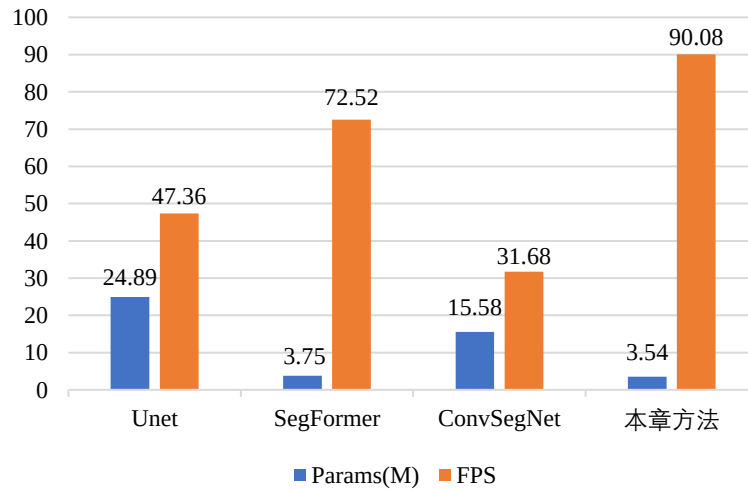


图 4-11 Li-DeepLabV3+和基准模型参数以及 FPS 比较

从图 4-11 可以看出，本章方法 Li-DeepLabV3+模型在参数最少的情况下具有最快的分割速度，其参数分别比 Unet、SegFormer 和 ConvSegNet 少 21.35M、0.21M 和 12.04M，同时模型的 FPS 优于所有基准测试模型，其中，与最有竞争力的 SegFormer 模型相比，Li-DeepLabV3+的 FPS 提升了 17.56。这进一步证明了所提方法兼顾了息肉分割模型的轻量性和实时性。

4.3 本章小结

本章节提出一种端到端的轻量级息肉分割网络模型 Li-DeepLabV3+。首先，使用优化的 MobileNetV2 网络为主干网络以降低模型复杂度；其次，采用改进的 SPPM 模块加强特征提取，在减少模型参数的同时提高模型的训练效率；最

后，使用改进的 UAFM 融合浅层特征和深层特征，增强模型的特征表示能力。Li-DeepLabV3+模型较原始模型参数量减少了 15 倍，在 Kvasir SEG、CVC-ClinicDB 以及 BKAI-NeoPolyp 数据集上分别获得了 88.82%、93.28%、87.09%的 Dice 系数。与实验中的基准模型相比，Li-DeepLabV3+模型兼顾了息肉分割精度和实时性，同时具有一定的泛化能力。

第 5 章 基于深度学习的内窥镜图像智能识别系统设计

本章节根据实际需求，结合第 3 章和第 4 章的最优 WCE 图像分类算法和 WCE 图像息肉分割算法，设计了一个 WCE 图像智能识别辅助诊断系统。该系统以提高医生诊断肠胃道疾病的效率为设计目标，致力于减轻医生负担，提供高精度的识别结果。本章将 WCE 图像智能识别系统分为三个部分进行详细阐述，包括系统需求分析、系统应用环境和界面设计及功能实现。

5.1 系统需求分析

肠胃道病变是一种常见的消化道疾病，每年都有上百万人患上不同类型的消化道疾病，甚至有许多人因此而丧命。然而大部分肠胃道疾病是可治愈的，若能及时发现这种不正常现象，就有可能拯救生命。WCE 检查是一种先进的诊断技术，可用于检测肠胃道异常。患者通过吞咽一颗微型无线胶囊来完成该检查，该胶囊内置摄像头和传感器，能够拍摄消化道内部的图像并传输到外部设备供医生分析。这种非侵入性的手术方法对于一些无法进行常规内窥镜检查的患者可能是一种更为舒适的选择，并且有助于提高肠胃道异常的诊断准确性。尽管 WCE 检查是全球最首选的肠胃道异常诊断方法，但由于诊断需要医生高强度的阅读大量图像，即使是训练有素的医生也需要大量时间进行诊断且容易错过有价值的信息。因此，如何借助计算机诊断系统来提高诊断效率、降低误诊率显得极为迫切。

深度学习技术可减轻医生阅读 WCE 图像的工作量，快速识别各种疑似病变，提高 WCE 的诊断效率。临床上对基于深度学习的 WCE 智能辅助诊断系统提出了以下需求：1) 数据收集和预处理：系统需要大量的 WCE 图像数据进行训练，这涉及到从医院或数据库中获取数据，根据实际情况对数据进行预处理，包括去噪、标准化、增强等。2) 模型选择和训练：需要选择合适的深度学习模型进行训练，以确保模型具备较高识别准确性和良好的泛化能力。3) 实时性要求：根据临床应用的需求，系统可能需要在实时或接近实时的情况下处理内窥镜图像。因此，需要考虑选择轻量级模型或部署优化策略来满足实时性要求。4) 系统的用户界面需要符合医护人员的使用习惯，提供友好的操作界面，便于医生快速准确地进行诊断。

5.2 系统应用环境需求

系统应用环境的设置对于确保系统整体的稳定运行和性能表现至关重要，其是决定系统是否能够顺畅工作的基础，同时也是保证系统在各种环境中能够正常响应的重要条件。正确配置系统运行环境可以确保系统顺利运行，并提高系统的可靠性。所提系统运行环境如表 5-1 所示。

表 5-1 运行环境

名称	说明
操作系统	Window11
CPU	AMD Ryzen 7 5800H with Radeon Graphics
GPU	NVIDIA RTX3060 6GB
数据库	MySql
语言	Python
开发环境	Pycharm
深度学习框架	Pytorch
计算机视觉库	OpenCv
环境依赖	PyQt5
架构依赖	C/S

5.3 系统功能设计

为防止用户信息泄露，本章将设计医生用户登录界面保证信息安全。结合系统需求分析，设计系统整体结构如图 5-1 所示。首先是医生验证登录界面，通过比对输入的账号和密码是否为数据库中存储的信息判定访问者的合法性，验证正确即可进入系统的内部界面。

进入系统内部会有两个功能界面分别为患者基本信息、患者病理信息。其中，患者基本信息界面可实现患者信息预览和患者信息查询；患者病理信息除了上述信息还可访问数据库中存储的对应患者病理信息，并实现 WCE 图像分类和息肉图像病灶区域语义分割，同时，系统会将辅助诊断的结果传给数据库，以便医生下一次查询。

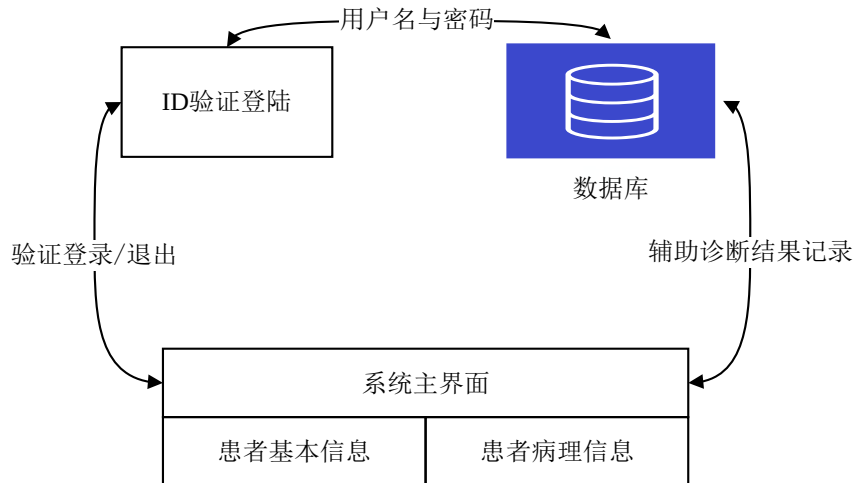


图 5-1 系统整体架构图

5.4 界面设计及功能实现

本文将开发一种基于深度学习的系统，该系统可以从内窥镜图像中自动诊断多种肠胃道异常，与人工诊断相比，可以大大减少所需的时间，从而协助医生诊断。所设计的 WCE 图像识别系统的主要功能：1) 实现医生界面的登录；2) 实现患者信息预览，包括患者姓名、性别、年龄等；3) 实现每个患者 WCE 图像的分类，具体分成正常、息肉和溃疡三个种类；4) 实现 WCE 息肉图像病灶区域语义分割，准确分割出肿瘤和非肿瘤区域。

5.4.1 设计思路

PyQt5 是一个 Python 类库，以 Qt 应用框架为基础，用来创建桌面应用。PyQt5 可用于设计、开发和部署交互式 and 可视化的应用程序。PyQt5 包括各种组件，如按钮、标签、文本框等，以及布局管理器和事件处理等功能。此外，PyQt5 还支持与数据库、多媒体和网络通信等方面的集成。Qt Designer 是 PyQt5 的一个可视化的工具，用于创建用户界面。通过使用 PyQt5 加载并使用这些设计好的界面。所设计的系统通过 Qt Designer 工具实现 GUI 控件的布局与设计，然后通过 PyUIC 工具将 Qt Designer 创建的用户界面文件转换为可在 PyQt5 应用程序中使用的 Python 代码。

图 5-2 展示了该系统的总体流程，医护人员通过主界面登录系统，登录成功后可查看患者基本信息，同时，可调用患者的 WCE 图像病理信息并查看对应患者图像分类和息肉图像语义分割结果，从而结合系统提供的结果辅助诊断。

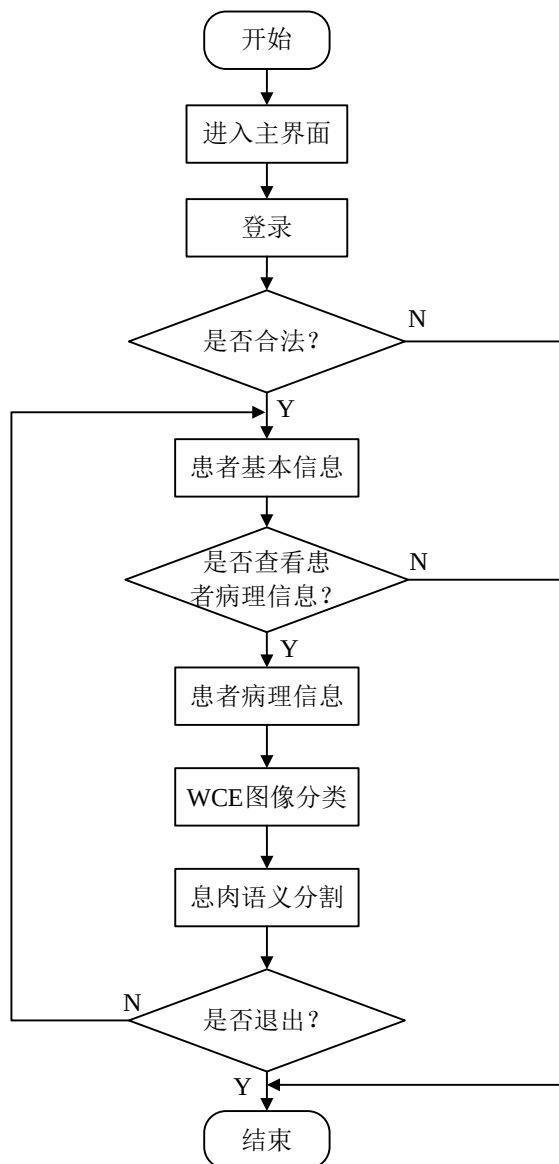


图 5-2 系统总体流程图

5.4.2 效果实现与测试

WCE 图像智能识别系统包括登录界面、患者基本信息、患者病理信息以及对应患者的 WCE 图像分类和息肉图像语义分割界面。

首先是医生登录界面，需要医生在登录界面输入账号和密码，若输入正确可以直接进入医生界面；若输入有误，会弹窗提示重新输入。验证成功后可以进入系统主界面。图 5-3 是系统的登录界面。



图 5-3 系统登录界面

医生登录系统进入主界面后可查看患者基本信息和病理信息，其中，患者基本信息界面可以查看患者的基本信息，包括患者编号、患者姓名、性别和年龄。医生可以通过患者编号或者姓名对患者进行查询，同时可通过显示所有按钮查看数据库中所有的患者信息。图 5-4 展示了查询单个患者的基本信息界面。



图 5-4 患者基本信息界面

对于患者病理信息界面，医生除了可以查看上述患者基本信息以外，还可以查看报告的医生和检测时间。同时还可以选中单个患者查看患者的 WCE 数据，如图 5-5 所示。



图 5-5 患者病理信息界面

在点击查看患者 WCE 数据按钮后，系统会读取患者对应的 WCE 图像数据的地址，同时调用第 3 章节 CA-ConvNeXt 模型训练后的权重对对应患者的 WCE 图像数据进行分类，从而实现正常、溃疡和息肉 WCE 图像的分类，图 5-6 展示了患者 WCE 图像分类界面。

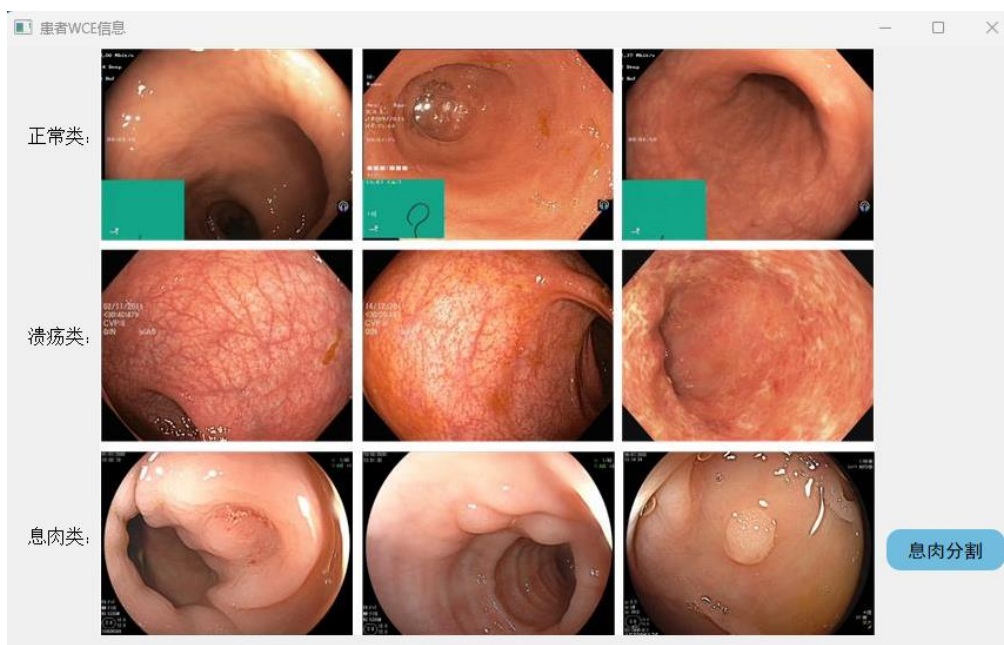


图 5-6 患者 WCE 图像分类界面

在查看患者 WCE 图像分类后，针对分类出的息肉类图像可以实现语义分割。通过调用第 4 章节的 Li-DeepLabV3+ 算法实现息肉图像病灶区域的语义分割，并生成对应的分割标签，如图 5-7 所示，其中绿色区域代表非肿瘤类，红色区域代

表肿瘤类。

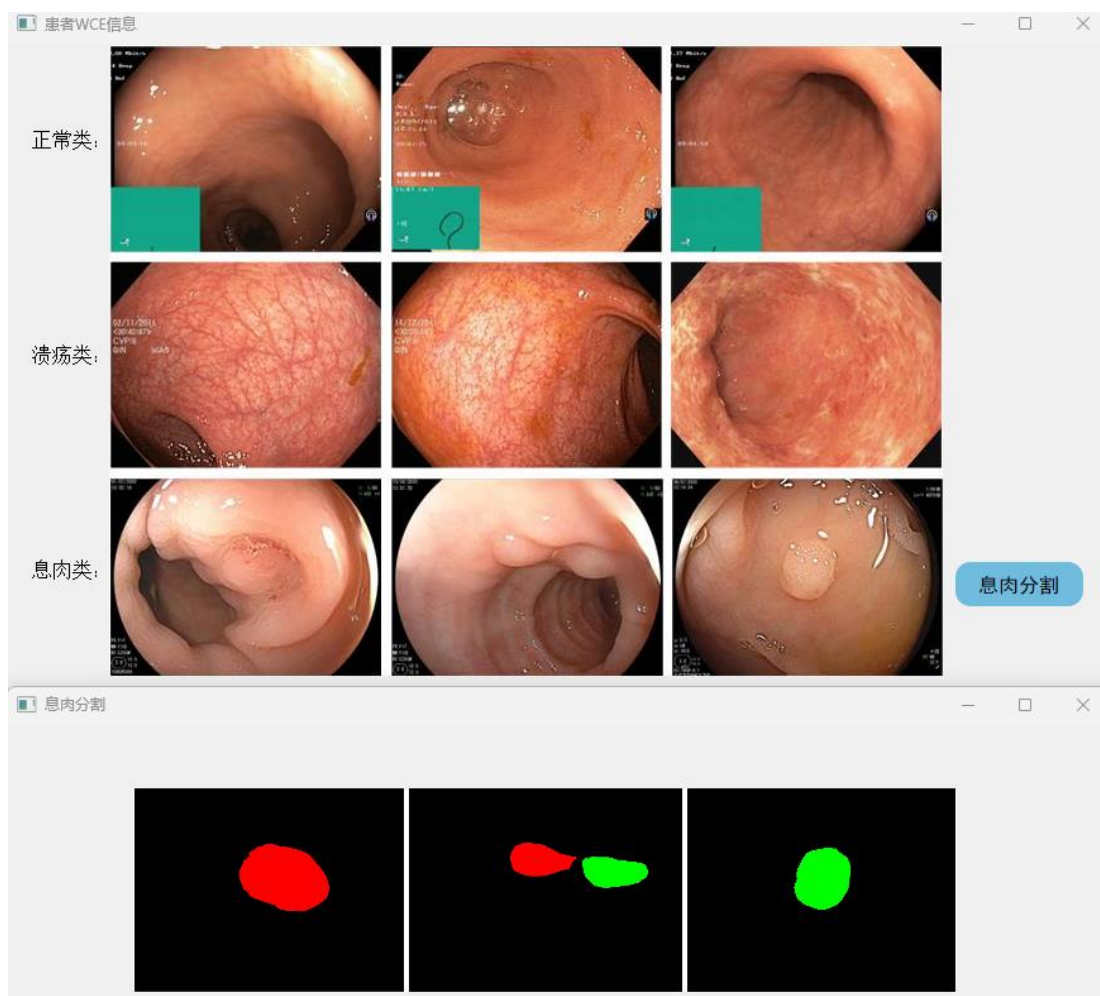


图 5-7 患者 WCE 图像息肉分割界面

5.5 本章小结

本章节开发了一个 WCE 图像智能识别辅助诊断系统，首先介绍了 WCE 图像识别系统的需求分析；然后介绍了系统应用环境需求；最后介绍了整个系统的设计思路以及界面展示。该系统结合前两章的优异算法提供高精度的识别和语义分割结果，同时提供友好的控制界面，在一定程度上减轻医护人员的负担。

第 6 章 总结与展望

6.1 总结

WCE 技术的核心优势在于其无痛和无创伤，单次检测可产生数万张图像。面对如此庞大的数据量，即便是资深医生都有可能出现漏检误检的情况。因此，如何借助计算机诊断系统来提高诊断效率、降低误诊率显得极为迫切。本文基于深度学习结合 WCE 图像预处理、WCE 图像分类和 WCE 图像语义分割三个方面建立一个 WCE 图像智能辅助诊断系统。首先，针对 WCE 图像中的光斑妨碍算法精确识别病灶的问题，融合 VS-UNet 与 Criminisi 算法对 WCE 图像预处理，从而实现光斑移除；其次，针对 WCE 图像数据存在数据量少，类间差异小，类内差异大的问题，结合数据增强和迁移学习策略设计了一种 CA-ConvNeXt 内窥镜图像分类模型；最后，针对目前息肉分割模型较复杂、模型泛化性较弱、息肉分割准确率较低的问题，探讨了一种轻量级息肉分割网络模型 Li-DeepLabV3+。具体研究内容如下：

1) 设计了一种融合 VS-UNet 与 Criminisi 算法移除 WCE 图像光斑的预处理算法。该算法包括光斑分割与修复两个过程，在光斑分割阶段，以 UNet 网络为基本框架定位光斑区域，使用优化的 VGG16 作为主干网络，并且引入 SPPFCSPC 模块增加模型感受野；在光斑修复方面，使用优化原始优先权函数的 Criminisi 算法实现光斑区域的修复。

2) 研究了一种图像分类算法 CA-ConvNeXt。结合迁移学习微调卷积神经网络 ConvNeXt，在 ConvNeXt 网络中引入 CA 模块加强注意力，并在网络训练时使用 Polyloss 损失函数优化模型。结果表明改进的模型在小样本条件下能实现正常、息肉和溃疡图像的精准分类，验证了算法的优越性。

3) 探讨了一种轻量级息肉分割网络模型 Li-DeepLabV3+。该网络模型以 DeepLabV3+ 为基本框架，首先，使用优化的轻量级 MobileNetV2 网络作为主干网络；其次，采用改进的 SPPM 模块加强特征提取；最后，为增强特征表示，使用改进的统一注意力融合模块 UAFM 融合浅层特征和深层特征。模型结合迁移学习在数据集 CVC-ClinicDB、Kvasir SEG 和 BKAI-NeoPolyp 上进行训练和验证，并跨数据集验证模型的泛化性。实验结果表明，Li-DeepLabV3+ 模型在分

割精度和分割速度方面具有优越的优势，并具有一定的泛化能力。

最后，结合最优 WCE 图像分类算法和 WCE 图像息肉分割算法，设计了一个 WCE 图像智能识别辅助诊断系统。该系统提供友好的用户界面以及精确的识别结果，使得诊断效率与准确性得到提升。

6.2 展望

本文设计的 WCE 图像智能识别系统实现了较高的识别准确率，在一定程度上减轻医生的负担。但是尚有值得改进之处需要深入剖析探讨，具体如下：

1) 本文主要研究二维 WCE 图像，未来可进一步研究三维 WCE 图像，将改进的分类分割算法应用在三维图像上，实现三维 WCE 图像的智能识别。

2) WCE 图像智能识别系统在应对常见肠胃道疾病如息肉和溃疡时展现出了较高的识别准确率，然而，该系统所能识别的肠胃道疾病种类相对有限。未来希望收集到更多种类的数据集，实现对不同类型病变的识别。

3) 系统界面比较简单，未来可继续优化系统功能，进一步提高用户界面的友好性。

参考文献

- [1] 中华人民共和国国家卫生健康委员会. 中国结直肠癌诊疗规范(2020 年版)[J]. 中华外科杂志, 2020, 58(8): 561-585.
- [2] Yang S, Lemke C, Cox B F, et al. A learning-based microultrasound system for the detection of inflammation of the gastrointestinal tract[J]. IEEE Transactions on Medical Imaging, 2020, 40(1): 38-47.
- [3] Iddan G, Meron G, Glukhovskiy A, et al. Wireless capsule endoscopy[J]. Nature, 2000, 405(6785), 417-418.
- [4] Leggett C L, Wang K K. Computer-aided diagnosis in GI endoscopy: looking into the future[J]. Gastrointestinal endoscopy, 2016, 84(5): 842-844.
- [5] 朱霆威, 李胜, 何熊熊. 先验权重共享码本下内窥镜图像大肠病变分类[J]. 中国图象图形学报, 2021, 26(09): 2270-2280.
- [6] Muhammad K, Khan S, Kumar N, et al. Vision-based personalized wireless capsule endoscopy for smart healthcare: taxonomy, literature review, opportunities and challenges[J]. Future Generation Computer Systems, 2020, 113: 266-280.
- [7] Shrestha R, Mohammed S K, Hasan M M, et al. Automated adaptive brightness in wireless capsule endoscopy using image segmentation and sigmoid function[J]. IEEE transactions on biomedical circuits and systems, 2016, 10(4): 884-892.
- [8] Arnold M, Ghosh A, Ameling S, et al. Automatic segmentation and inpainting of specular highlights for endoscopic imaging[J]. EURASIP Journal on Image and Video Processing, 2010, 2010:1-12.
- [9] Gao Y, Yang J, Ma S, et al. Dynamic searching and classification for highlight removal on endoscopic image[J]. Procedia Computer Science, 2017, 107: 762-767.
- [10] 池月, 李正平, 徐超等. 医用内窥镜图像的光斑移除算法[J]. 计算机应用, 2023, 43(04): 1278-1283.
- [11] Alsaleh S M, Aviles-Rivero A I, Hahn J K. ReTouchImg: Fusioning from-local-to-global context detection and graph data structures for fully-automatic specular reflection removal for endoscopic images[J]. Comput Med Imaging Graph. 2019, 73: 39-48.
- [12] Nie C, Xu C, Li Z, et al. Specular reflections detection and removal for endoscopic images based on brightness classification. Sensors (Basel) [J]. 2023, 23(2): 974.
- [13] Monkam P, Wu J, Lu W, et al. EasySpec: Automatic specular reflection detection and suppression from endoscopic images[J]. IEEE Transactions on Computational Imaging, 2021, 7: 1031-1043.
- [14] Zhang C, Liu Y, Wang K, et al. Specular highlight removal for endoscopic images using partial attention network[J]. Physics in Medicine & Biology, 2023, 68(22): 1-15.

- [15] Naz J, Sharif M, Yasmin M, et al. Detection and classification of gastrointestinal diseases using machine learning[J]. *Current Medical Imaging*, 2021, 17(4): 479-490.
- [16] Yang J, Chang L, Li S, et al. WCE polyp detection based on novel feature descriptor with normalized variance locality-constrained linear coding[J]. *International Journal of Computer Assisted Radiology and Surgery*, 2020, 15(8): 1291-1302.
- [17] Souaidi M, Ansari M E. Multi-scale analysis of ulcer disease detection from wce images[J]. *IET Image Processing*, 2019, 13(12): 2233-2244.
- [18] Charfi S, El Ansari M, Balasingham I. Computer-aided diagnosis system for ulcer detection in wireless capsule endoscopy images[J]. *IET Image Processing*, 2019, 13(6): 1023-1030.
- [19] Ramzan M, Raza M, Sharif M, et al. Gastrointestinal tract infections classification using deep learning[J]. *Comput. Mater. Contin*, 2021, 69: 3239-3257.
- [20] Wang W, Tian J, Zhang C, et al. An improved deep learning approach and its applications on colonic polyp images detection[J]. *BMC Medical Imaging*, 2020, 20: 1-14.
- [21] Padmavathi P, Harikiran J. Se-Resnet: A novel method for gastrointestinal (GI) diseases classification from wireless capsule endoscopy (WCE) images[J]. *Traitement du Signal*, 2023, 40(4): 1341-1353.
- [22] Yue G, Wei P, Liu Y, et al. Automated endoscopic image classification via deep neural network with class imbalance loss[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-11.
- [23] 巩稼民, 马豆豆, 蒋杰伟等. 基于深度学习的慢性萎缩性胃炎诊断[J]. *中国医学物理学杂志*, 2020, 37(5): 649-655.
- [24] Lonseko Z M, Adjei P E, Du W, et al. Gastrointestinal disease classification in endoscopic images using attention-guided convolutional neural networks[J]. *Applied Sciences*, 2021, 11(23): 1-12.
- [25] Qin Y, Xia H, Song S. RT-Net: Region-enhanced attention transformer network for polyp segmentation[J]. *Neural Processing Letters*, 2023, 55(9): 11975-11991.
- [26] Xia H, Qin Y, Tan Y, et al. BA-Net: Brightness prior guided attention network for colonic polyp segmentation[J]. *Biocybernetics and Biomedical Engineering*, 2023, 43(3): 603-615.
- [27] Jha D, Smedsrud P H, Johansen D, et al. A comprehensive study on colorectal polyp segmentation with ResUNet++, conditional random field and test-time augmentation[J]. *IEEE journal of biomedical and health informatics*, 2021, 25(6): 2029-2040.
- [28] Wang M, An X, Li Y, et al. EMS-Net: Enhanced multi-Scale network for polyp segmentation[C]//2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, Mexico, 01-05 November, 2021: 2936-2939.
- [29] Bhattacharya D, Betz C, Eggert D, et al. Dual parallel reverse attention edge network: DPRA-EdgeNet[J]. *Nordic Machine Intelligence*, 2021, 1(1): 8-10.
- [30] Ta N, Chen H, Lyu Y, et al. Ble-net: boundary learning and enhancement network for polyp segmentation[J]. *Multimedia Systems*, 2023, 29(5): 3041-3054.

- [31] Sharma P, Gautam A, Maji P, et al. Li-SegPNet: Encoder-Decoder mode lightweight segmentation network for colorectal polyps analysis[J]. IEEE Transactions on Biomedical Engineering, 2022, 70(4): 1330-1339.
- [32] Jin Y, Hu Y, Jiang Z, et al. Polyp segmentation with convolutional MLP[J]. The Visual Computer, 2023, 39(10): 4819-4837.
- [33] Ige A O, Tomar N K, Aranuwa F O, et al. ConvSegNet: automated polyp segmentation from colonoscopy using context feature refinement with multiple convolutional kernel sizes[J]. IEEE Access, 2023, 11: 16142-16155.
- [34] Lai H, Luo Y, Zhang G, et al. Toward accurate polyp segmentation with cascade boundary-guided attention[J]. The Visual Computer, 2023, 39(4): 1453-1469.
- [35] Telea A. An image inpainting technique based on the fast marching method[J]. Journal of graphics tools, 2004, 9(1): 23-34.
- [36] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Springer International Publishing, Munich, Germany, 5-9 October, 2015: 234-241.
- [37] Criminisi A, Pérez P, Toyama K. Region filling and object removal by exemplar-based image inpainting[J]. IEEE Transactions on image processing, 2004, 13(9): 1200-1212.
- [38] 张国标, 李信, 陆伟. 基于深度卷积神经网络的生物医学混合图像检测[J]. 现代情报, 2020, 40(07): 74-81+103.
- [39] 潘英杰, 程远志, 刘豪等. 基于视觉变换器的级联多阶层医学影像配准方法[J]. 生物医学工程学杂志, 2022, 39(05): 876-886.
- [40] 王琛, 闵永军, 张天慈. 基于改进 Criminisi 算法的自动驾驶仿真道路场景图像修复[J]. 软件工程, 2022, 25(03): 44-47.
- [41] Sánchez F J, Bernal J, Sánchez-Montes C, et al. Bright spot regions segmentation and classification for specular highlights detection in colonoscopy videos[J]. Machine Vision and Applications, 2017, 28(8): 917-936.
- [42] 项诗雨, 魏利胜. 基于卷积神经网络的内窥镜图像分类[J]. 安徽工程大学学报, 2023, 38(03): 29-36+57.
- [43] 李建威, 吕晓琪, 谷宇. 基于改进 ConvNeXt 的皮肤镜图像分类方法[J]. 计算机工程, 2023, 49(10): 239-246+254.
- [44] 黄文博, 黄钰翔, 姚远等. 融合注意力的 ConvNeXt 视网膜病变自动分级[J]. 光学精密工程, 2022, 30(17): 2147-2154.
- [45] Hu J, Shen L, Albanie S, et al. Squeeze-and-Excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [46] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). Springer, Cham, Munich, Germany, 10-13 September, 2018: 3-19.

- [47] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, Nashville, TN, USA, 20-25 June, 2021: 13708-13717.
- [48] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database[C]//2009 IEEE conference on computer vision and pattern recognition. IEEE, Miami, FL, USA, 20-25 June, 2009: 248-255.
- [49] Ramprasaath R, Selvaraju, Michael Cogswell, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]//2017 IEEE International Conference on Computer Vision (ICCV), IEEE, Venice, Italy, 22-29 October, 2017: 618-626.
- [50] Pogorelov K, Randel K R, Griwodz C, et al. Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection[C]//Proceedings of the 8th ACM on Multimedia Systems Conference. Association for Computing Machinery, New York, NY, United States, 20-23 June, 2017: 164-169.
- [51] Gamage C, Wijesinghe I, Chitraranjan C, et al. GI-Net: Anomalies classification in gastrointestinal tract through endoscopic imagery with deep learning[C]//Moratuwa Engineering Research Conference. IEEE, Moratuwa, Sri Lank, 03-05 July, 2019: 66-71.
- [52] Escobar J P, Gomez N, Sanchez K, et al. Transfer learning with convolutional neural network for gastrointestinal diseases detection using endoscopic images[C]//2020 IEEE Colombian Conference on Applications of Computational Intelligence, IEEE, Cali Colombia, 07-08 August, 2020: 1-6.
- [53] Demirbaş A A, Üzen H, Fırat H. Spatial-attention ConvMixer architecture for classification and detection of gastrointestinal diseases using the Kvasir dataset[J]. Health Inf Sci Syst, 2024, 1(12): 1-15.
- [54] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European conference on computer vision (ECCV). Springer Cham, Munich, Germany, 8-14 September, 2018: 801-818.
- [55] Chollet F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, Honolulu, HI, USA, 21-26 July, 2017: 1251-1258.
- [56] Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, Salt Lake City, UT, USA, 18-23 June, 2018: 4510-4520.
- [57] 闫志蕊, 王宏伟, 耿毅德. 基于改进 DeeplabV3+和迁移学习的煤岩界面图像识别方法[J]. 煤炭科学技术, 2023, 51(S1): 429-439.
- [58] 于浩, 张震, 张雷, 等. 基于图像语义分割的药用玻璃瓶外轮廓尺寸检测方法研究[J]. 河北省科学院学报, 2023, 40(1): 19-25.
- [59] Zhu W, Huang Y, Zeng L, et al. AnatomyNet: deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy[J]. Medical physics, 2019, 46(2): 576-589.

- [60] 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.
- [61] Jha D, Smedsrud P H, Riegler M A, et al. Kvasir-seg: A segmented polyp dataset[C]//MultiMedia Modeling: 26th International Conference. Springer International Publishing, Daejeon, South Korea, 5–8 January, 2020: 451-462.
- [62] Ngoc Lan P, An N S, Hang D V, et al. NeoUNet: Towards accurate colon polyp segmentation and neoplasm detection[C]//Advances in Visual Computing: 16th International Symposium (ISVC 2021). Springer International Publishing, 4-6 October, 2021: 15-28.
- [63] Everingham M, Eslami S M A, Van Gool L, et al. The pascal visual object classes challenge: A retrospective[J]. International journal of computer vision, 2015, 111: 98-136.
- [64] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Trans Pattern Anal Mach Intell. 2020, 42(2): 318-327.
- [65] 陈英, 张伟, 林洪平, 等. 医学图像分割算法的损失函数综述[J]. 生物医学工程学杂志, 2023, 40(02): 392-400.
- [66] 殷晓航, 王永才, 李德英. 基于 U-Net 结构改进的医学影像分割技术综述[J]. 软件学报, 2021, 32(02): 519-550.
- [67] Xie E, Wang W, Yu Z, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers[J]. Advances in Neural Information Processing Systems, 2021, 34: 12077-12090.
- [68] Ige A O, Tomar N K, Aranuwa F O, et al. ConvSegNet: Automated polyp segmentation from colonoscopy using context feature refinement with multiple convolutional kernel sizes[J]. IEEE Access, 2023, 11: 16142-16155.

附录

1. CA-ConvNeXt 算法的部分代码

```

class ConvNeXt(nn.Module):
    def __init__(self, in_chans: int = 3, num_classes: int = 3, depths: list = None, dims: list = None,
drop_path_rate: float = 0., layer_scale_init_value: float = 1e-6, head_init_scale: float = 1.):
        super().__init__()
        self.downsample_layers = nn.ModuleList()
        stem = nn.Sequential(nn.Conv2d(in_chans, dims[0], kernel_size=4, stride=4),
                             LayerNorm(dims[0], eps=1e-6, data_format="channels_first"))
        self.downsample_layers.append(stem)
        for i in range(4):
            stage=nn.Sequential([Block(dim=dims[i],drop_rate=dp_rates[cur+j],
layer_scale_init_value=layer_scale_init_value)
                                for j in range(depths[i])])
            self.stages.append(stage)
            cur += depths[i]
        self.feat_attention = CA_Block(channel=768,h=7,w=7)
        self.norm = nn.LayerNorm(dims[-1], eps=1e-6)
        self.head = nn.Linear(dims[-1], num_classes)
        self.apply(self._init_weights)
        self.head.weight.data.mul_(head_init_scale)
        self.head.bias.data.mul_(head_init_scale)
    def _init_weights(self, m):
        if isinstance(m, (nn.Conv2d, nn.Linear)):
            nn.init.trunc_normal_(m.weight, std=0.2)
    def forward_features(self, x: torch.Tensor) -> torch.Tensor:
        for i in range(4):
            x = self.downsample_layers[i](x)
            x = self.stages[i](x)
        x = self.feat_attention(x)
        return self.norm(x.mean([-2, -1]))
    def forward(self, x: torch.Tensor) -> torch.Tensor:
        x = self.forward_features(x)
        x = self.head(x)
        return x

```

2. Li-DeepLabV3+算法的部分代码

```

class DeepLab(nn.Module):
    def __init__(self, num_classes, backbone="mobilenet", pretrained=True,
downsample_factor=16):
        super(DeepLab, self).__init__()
        self.backbone=MobileNetV2(downsample_factor=downsample_factor, pretrained=pretrained)
        in_channels = 320
        low_level_channels = 24
        self.sppm=SPPM(in_channels=in_channels,inter_channels=256,
out_channels=256,bin_sizes=[2,4,8,16])
        self.shortcut_conv = nn.Sequential(
            nn.Conv2d(low_level_channels, 48, 1),
            nn.BatchNorm2d(48),
            nn.ReLU(inplace=True)
        )
        self.cat_conv = nn.Sequential(
            nn.Conv2d(48 + 256, 256, 3, stride=1, padding=1),
            nn.BatchNorm2d(256),
            nn.ReLU(inplace=True),
            nn.Dropout(0.5),
            nn.Conv2d(256, 256, 3, stride=1, padding=1),
            nn.BatchNorm2d(256),
            nn.ReLU(inplace=True),
            nn.Dropout(0.1),
        )
        self.cls_conv = nn.Conv2d(256, num_classes, 1, stride=1)
        self.uafl1 = UAFL(24,256,256)
    def forward(self, x):
        H, W = x.size(2), x.size(3)
        low_level_features, x = self.backbone(x)
        x = self.sppm(x)
        x = self.uafl1(x,low_level_features)
        x = self.cls_conv(x)
        x = F.interpolate(x, size=(H, W), mode='bilinear', align_corners=True)
        return x

```


攻读学位期间发表的学术论文

- [1] **Xiang Shiyu**, Wei Lisheng, Hu Kgaifeng. Lightweight colon polyp segmentation algorithm based on improved DeepLabV3+[J]. Journal of Cancer, 2024, 15(1): 41-53. (SCI:001119285000020)
- [2] **Xiang Shiyu**, Wei Lisheng, Gao Gang, et al. Colonic Polyp Segmentation Algorithm Based on Improved SegFormer Network[C]//2023 3rd International Conference on Computer, Control and Robotics (ICCCR). IEEE, 2023: 101-105. (EI: 20233414619461)
- [3] **Xiang Shiyu**, Wei Lisheng, Tang Shaoyu. Fusion of UNet and Criminisi algorithms for endoscopic specular highlight removal[C]//2023 38th Youth Academic Annual Conference of Chinese Association of Automation (YAC). IEEE, 2023: 1093-1097. (EI:20240815614624)
- [4] **项诗雨**,魏利胜.基于卷积神经网络的内窥镜图像分类[J].安徽工程大学学报,2023,38(03):29-36+57.

攻读学位期间获得的奖励

- [1] 获得 2024 年安徽省普通高等学校优秀毕业生
- [2] 获得 2023-2024 学年校国元证券奖学金
- [3] 获得 2023-2024 学年校研究生一等奖学金
- [4] 获得 2022-2023 学年校研究生二等奖学金
- [5] 获得 2021-2022 学年校研究生二等奖学金
- [6] 获得 2022 年校第八届“互联网+”大学生创新创业大赛研究生创意组三等奖
- [7] 获得 2023 年中国(国际)传感器创新创业大赛(华东赛区)一等奖
- [8] 获得 2023 年中国(国际)传感器创新创业大赛决赛三等奖
- [9] 获得 2023 年第九届安徽省互联网+大学生创新创业大赛总决赛铜奖

致 谢

连雨不知春去，一晴方觉夏深。行文至此，将为自己的学生生涯划上句点。满是不舍，满怀感恩。

饮水思源，学成念师。首先要感谢我的导师魏利胜教授和校外导师张平改老师，感谢魏老师在学术研究中对我的悉心栽培与耐心教导，能够成为魏老师的学生，倍感荣幸。魏老师学识渊博，治学严谨，从论文开题到最终成文，字斟句酌，详细批改，提出宝贵的修改意见。三载求学路，无论是学业还是生活，都给予我充分的支持与包容。在此，衷心祝愿诸位老师工作顺遂，万事顺意。

山水一程，三生有幸。其次要感谢一起在实验室学习的大家，在求学的路上一路同行。感谢张倩、汪瑞、马晓东等师兄师姐向我分享学习与生活中的经验；感谢我的同门高港、王瑞仁、唐绍语、刘瑞给予的帮助和鼓励；感谢王勤、李明、陶新杰、朱圣博、汪启良等师弟师妹的真诚陪伴。感谢谢雨欣、吴培、郝泽强等朋友的关心陪伴和时常挂念。祝愿各位朋友一切尽意，百事从欢。

亲人之教，得以成才。感谢父母的悉心栽培与无私奉献。是父母的支持让我能够无所顾忌选择自己想要的人生。感谢父母一如既往的疼爱，让我在求学路上衣食无忧无惧人海。感谢我的弟弟妹妹，陪我一起长大，在成长的道路上给予我无限的陪伴和最纯真的爱。如今我将离开校园，步入社会，我知道家是我的前进路上的永远能够停靠的港湾。

最后，向在百忙之中对我的论文进行评阅的各位教授和专家致以真诚的感谢。

项诗雨

安徽工程大学电气工程学院

二〇二四年四月

尚德敏学 唯实惟新

