

DOI:10.19651/j.cnki.emt.2519585

基于改进 UNeXt 的岩心目标提取算法^{*}

郑方艳¹ 何小海¹ 卿黎波¹ 何海波² 滕奇志¹

(1. 四川大学电子信息学院 成都 610065; 2. 成都西图科技有限公司 成都 610065)

摘要: 岩心区域的准确提取在数字岩心构建与智能油气储层评价等任务中具有重要意义。然而,岩心图像常伴随复杂背景、边缘模糊及多尺度结构分布等问题,给自动化分割带来较大挑战。为此,提出了一种基于改进 UNeXt 架构的岩心目标提取算法,旨在提升模型对岩心区域的分割性能。该方法通过在网络不同层次引入 CBAM 或 EMA 模块,有效增强模型对边缘细节和全局上下文特征的表达;同时,在网络颈部设计并引入多尺度特征增强模块,以进一步提升模型对多尺度结构及复杂纹理的感知能力。此外,考虑到目前缺乏专门用于岩心区域提取的公开数据集,本文自主构建了相关的岩心图像数据集。实验结果显示,与原始 UNeXt 网络相比,所提算法在 mIoU、F1-score 和 mPA 3 项指标上分别提升 1.49%、2.06% 和 0.75%,同时 MSE 下降 78.63%,且通过统计检验证实上述提升均具有显著性。为验证模型的泛化性能,本文还在 BRISC 2025 和 EBHI-Seg 两个公开医学图像数据集上进行了对比实验,结果表明本文算法在自建数据集和公开数据集上均表现优异。

关键词: 岩心图像;图像分割;UNeXt;CBAM;EMA;MSDE

中图分类号: TP391.41; TN919.82 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Core extraction algorithm based on improved UNeXt

Zheng Fangyan¹ He Xiaohai¹ Qing Linbo¹ He Haibo² Teng Qizhi¹

(1. College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China;

2. Chengdu Xitu Technology Co., Ltd., Chengdu 610065, China)

Abstract: The accurate extraction of core regions is of great significance for tasks such as digital core construction and intelligent reservoir evaluation. However, core images often suffer from complex backgrounds, blurred edges, and multi-scale structural distributions, posing significant challenges to automated segmentation. To address these issues, this paper proposes a core object extraction algorithm based on an improved UNeXt architecture, aiming to enhance the model's segmentation performance for core regions. The method effectively strengthens the model's ability to represent edge details and global contextual features by introducing CBAM or EMA modules at different network levels. Simultaneously, a multi-scale feature enhancement module is designed and incorporated into the network neck to further improve the model's perception of multi-scale structures and complex textures. Additionally, given the current lack of publicly available datasets dedicated to core region extraction, this paper independently constructs a relevant core image dataset. Experimental results show that, compared to the original UNeXt network, the proposed algorithm achieves improvements of 1.49%, 2.06% and 0.75% in mIoU, F1-score and mPA, respectively, while the MSE decreases by 78.63%. Statistical tests confirm that these improvements are all significant. To validate the model's generalization capability, comparative experiments were also conducted on two public medical image datasets, BRISC 2025 and EBHI-Seg. The results demonstrate that the proposed algorithm performs excellently on both the self-constructed dataset and the public datasets.

Keywords: core images; image segmentation; UNeXt; CBAM; EMA; MSDE

0 引言

岩心区域的准确提取在数字岩心构建与智能油气储层

评价等任务中具有重要意义。然而,在实际采集过程中,常常会捕捉到大量背景区域,包括噪声、杂质以及非岩石成分等干扰物,这些干扰信息会影响岩心区域的准确识别。目

收稿日期:2025-08-12

* 基金项目:国家自然科学基金(62071315)项目资助

前,岩心区域的提取主要依赖人工操作,效率低下,亟需自动化解决方案提升工作效率。

传统的图像分割方法主要包括阈值分割与边缘检测两类^[1]。在实际应用中,由于相机景深限制导致的边界模糊,以及岩心内部矿物成分复杂、灰度分布不均等因素,该类方法在分割精度与鲁棒性方面均存在明显不足。为改善传统方法的性能,刘庭亭等^[2]提出了一种基于改进斑马优化算法的多阈值图像分割方法,该方法通过引入混沌映射与邻域波动等策略优化搜索过程,解决传统多阈值分割中精度低、效率差的问题。吴京城等^[3]提出了一种融合反背景差分与天津法的图像分割方法,有效解决了因对比度低、背景不均匀及光晕伪影导致的图像分割难题。尽管上述方法在一定程度上提升了分割效果,但在处理复杂岩心图像时,其分割精度与计算效率仍难以满足实际应用需求。

近年来,随着深度学习技术的不断发展,图像分割技术也迎来了新的突破。Shelhamer 等^[4]提出了全卷积网络(fully convolutional network, FCN),该模型将卷积神经网络(convolutional neural network, CNN)的全连接层替换为卷积层,并通过跳跃连接融合了不同深度的特征图,显著提升了语义分割的精度。Chen 等^[5-7]提出的 DeepLab 系列模型,通过对核心模块的创新,实现模型在边界细节和整体精度方面的持续优化。冯哲等^[8]通过将多头自注意力机制融入 YOLOv8x(you only look once version 8x)以获取全局视野,并融合专用小目标检测层以增强细节感知,提出了一种针对暗弱环境小天体岩石的检测模型,使得检测准

确率与鲁棒性得到显著提升。然而,在众多分割模型中,U-Net^[9]凭借其编码器-解码器结构和跳跃连接形成的 U 型对称架构,能够在少量标注样本下实现高精度分割,因此被广泛应用于岩心图像分析任务。此外,研究者们也提出了多种改进方案。例如,曾中华等^[10]提出一种增强型 U-Net 结构,以应对矿石图像识别难、标注成本高的问题。舒小锋等^[11]则在 U²-Net^[12]中融入多尺度特征提取与坐标注意力(coordinate attention, CA),提升了对岩石薄片细小颗粒边缘的捕捉精度。然而,上述改进模型在提升分割精度的同时,也普遍带来了参数数量的增长与计算复杂度的显著上升,使其难以在对实时性要求较高的实际场景中部署。

2022 年,Alanarasu 等^[13]提出一种基于卷积多层感知器的图像分割网络 UNeXt,该模型在 U-Net 基础上进行结构优化,在保持高分割精度的同时显著降低了计算复杂度,其综合性能优于 U-Net 及其多个变体。基于此,本文在 UNeXt 架构的基础上进行改进,在合理控制计算开销的前提下,提升模型对岩心区域的分割精度与鲁棒性,以实现高效、准确的岩心提取。

1 UNeXt 网络介绍

UNeXt 是首个基于卷积的多层感知器(multilayer perceptron, MLP)的轻量级图像分割网络,结合了创新的带有轴向偏移的标记化感知器(Tokenized MLP),有效增强潜在空间的表达能力,能显著提升图像分割的性能,其网络结构如图 1 所示。

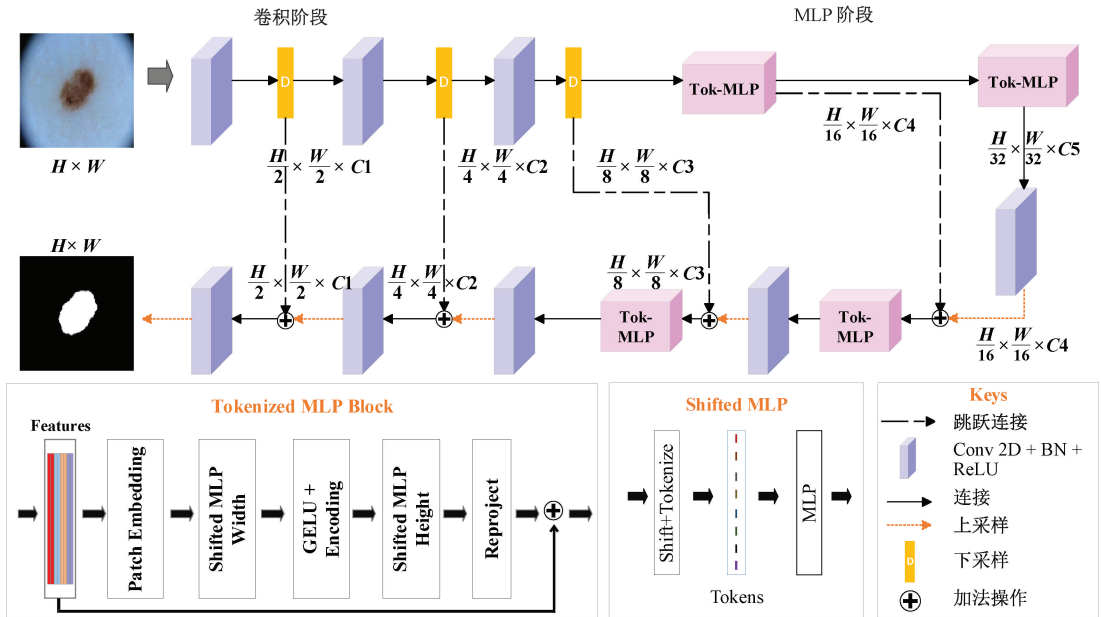


图 1 UNeXt 算法的网络结构

Fig. 1 Network architecture of UNeXt algorithm

在 UNeXt 中,编码器和解码器各由 3 个卷积块和 2 个 Tokenized MLP 模块组成。每个卷积块包含 3×3 卷积、批归一化(batch normalization, BN)和线性整流函

数(rectified linear unit, ReLU)。编码器通过 3 次最大池化和 2 次 Tokenized MLP 实现 5 次下采样,将特征图缩小至原始的 $1/32$ 。解码器采用双线性插值逐步上采

样,并通过跳跃连接融合不同阶段的特征,提升信息传递能力。

UNeXt 通过引入 Tokenized MLP 模块来增强深层特征建模能力。该模块首先将卷积特征通过 3×3 卷积进行块嵌入,生成标记并映射至维度 E。随后,通过层归一化后输入到移位 MLP 模块,依次在宽度和高度方向上进行偏移、划分、融合,并结合深度卷积和高斯误差线性单元激活函数(Gaussian error linear unit, GELU)提取空间特征,提

升表达能力。

2 UNeXt 网络改进

2.1 整体结构

本文以 UNeXt 网络为基础框架,针对岩心图像矿物成分复杂、纹理结构多变、目标边界模糊等挑战,在其编码器-解码器架构的关键节点进行了系统性增强,其整体结构如图 2 所示。具体改进包括以下 4 个方面:

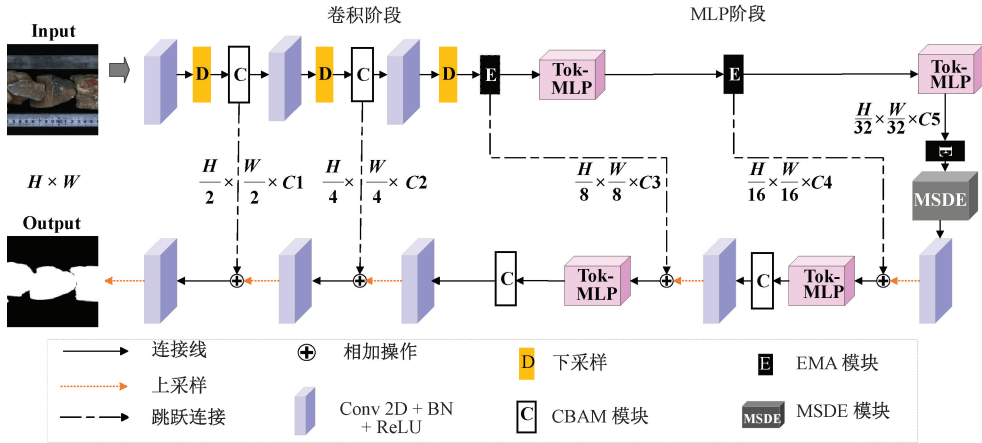


图 2 改进 UNeXt 算法的网络结果图

Fig. 2 Network architecture of the improved UNeXt algorithm

1)在编码器的前两个卷积阶段引入卷积块注意力模块(convolutional block attention module, CBAM)^[14]。UNeXt 在编码器的浅层阶段分辨率较高,富含岩心边缘、矿物信息等空间细节。CBAM 可通过其通道注意力筛选重要特征通道,再通过空间注意力聚焦关键空间位置,强化对岩心边缘和纹理的敏感度,为后续深度特征提取提供高质量的细节基础。

2)在编码器的第 3 个卷积层及 Tokenized MLP 模块之后集成高效的多尺度注意力模块(efficient multi-scale attention, EMA)^[15]。在卷积层之后引入 EMA,可利用其多尺度特征提取能力,强化对岩心中等尺度结构的感知,并捕获跨空间的上下文依赖。而在 Tokenized MLP 模块之后部署的 EMA,可对经过标记化处理和空间交互后的特征进行跨尺度上下文再校准,弥补纯 MLP 结构在空间细节保持上的不足,形成从卷积特征空间到标记序列空间的全方位多尺度增强。

3)在颈部 EMA 模块之后,设计多尺度空洞编码模块(multi-scale dilated encoder, MSDE),实现岩心特征增强的递进与互补。EMA 模块虽能高效捕获跨空间依赖关系,但其感受野与特征变换方式相对有限。MSDE 则通过 4 个并行的多尺度卷积与空洞卷积,以更强的显式编码能力进一步拓宽了感受野的多样性,针对岩心图像中尺度差异悬殊的目标进行深度特征增强,为解码器提供兼具丰富语义与精确空间细节的特征表示。

4)在解码器 Tokenized MLP 模块之后部署 CBAM 模块作为语义校准器。当特征经移位 MLP 处理并开始上采样恢复时,CBAM 通过对显著岩心区域的注意力聚焦,优化特征恢复的精度与语义一致性,有效改善分割边界和纹层形态的表达效果。

综上所述,本文构建了一个层次化的特征增强体系,让 3 个模块在特征处理的不同阶段发挥协同作用,实现了从局部细节、上下文信息到跨尺度特征的协同感知。

2.2 CBAM 模块

CBAM 通过串联通道注意力模块(channel attention module, CAM)与空间注意力模块(spatial attention module, SAM),以轻量高效的方式引入注意力机制,其结构如图 3 所示。该模块通过对输入特征图进行自适应的通道和空间注意力建模,能够强化岩心区域的关键信息,同时抑制背景和噪声,从而显著提升特征的判别能力和模型的鲁棒性。

1) 通道注意力模块

CAM 的核心在于挖掘不同通道岩心特征的重要性,并通过全局池化与共享全连接层(Shared MLP)自适应地为各通道分配权重,从而增强关键通道的响应、抑制无关信息。首先,对输入特征图 $F \in \mathbf{R}^{C \times H \times W}$ 进行全局平均池化(Average Pooling)和全局最大池化(Max Pooling)操作,生成两个大小为 $C \times 1 \times 1$ 的通道描述向量 F_{avg}^c 和 F_{max}^c 。然后,将 F_{avg}^c 和 F_{max}^c 送入到 Shared MLP,该网络由全连接

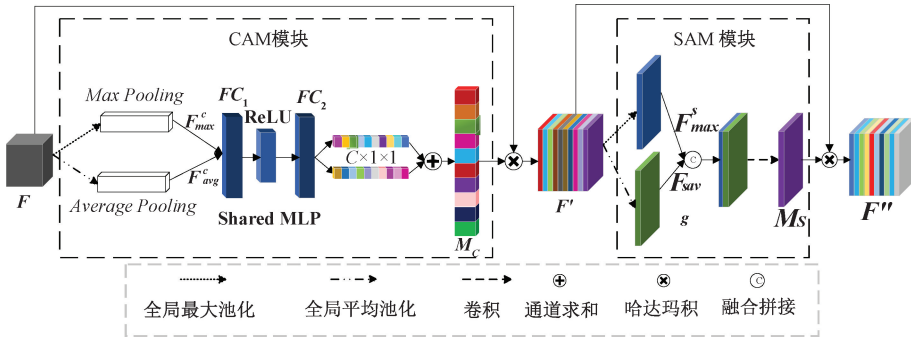


图 3 CBAM 结构

Fig. 3 Structure of CBAM

层 FC_1 和 FC_2 以及一个 ReLU 组成,将两个向量分别映射为两个一维特征表示。最后,将两个输出进行逐元素求和并通过 S 型函数(Sigmoid)进行归一化,得到每个通道的注意力权重图 $M_c \in \mathbf{R}^{C \times 1 \times 1}$ 。该权重图与输入特征图 F 逐通道相乘得到 F' , 实现通道维度上的加权特征增强。其公式如式(1)~(4)所示。

$$F_{avg}^c = AvgPool(F) \tag{1}$$

$$F_{max}^c = MaxPool(F) \tag{2}$$

$$M_c(F) = \sigma(MLP(F_{avg}^c) + MLP(F_{max}^c)) \tag{3}$$

$$F' = M_c \otimes F \tag{4}$$

式中: $AvgPool$ 表示全局平均池化; $MaxPool$ 表示全局最大池化; σ 表示 Sigmoid 函数; \otimes 表示哈达玛积。

2) 空间注意力模块

SAM 通过池化与卷积操作生成空间注意力图,挖掘特征图中关键的空间位置信息,从而增强岩心显著区域的响应能力。该模块以通道注意力模块输出的特征图 F' 作为输入,首先进行 Max Pooling 和 Average Pooling 操作,生成两个大小为 $1 \times H \times W$ 的空间注意力图 F_{avg}^s 和 F_{max}^s 。

随后,将特征图 F_{avg}^s 和 F_{max}^s 在通道维度上进行拼接,得到一个大小为 $2 \times H \times W$ 的联合特征图。该特征图随后通过一个卷积核大小为 7×7 的卷积层进行特征融合,进一步提取局部空间上下文信息。最终,融合结果通过 Sigmoid 激活函数生成空间注意力权重图 $M_s \in \mathbf{R}^{1 \times H \times W}$ 。该权重图与输入特征图 F' 逐元素相乘得到最终的输出 F'' , 实现空间维度上的加权增强。计算公式如式(5)~(8)所示。

$$F_{avg}^s = AvgPool(F') \tag{5}$$

$$F_{max}^s = MaxPool(F') \tag{6}$$

$$M_s(F) = \sigma(f^{7 \times 7}([F_{avg}^s, F_{max}^s])) \tag{7}$$

$$F'' = M_s \otimes F' \tag{8}$$

式中: $f^{7 \times 7}$ 表示卷积核大小为 7×7 的卷积层。

2.3 EMA 模块

EMA 是一种轻量级的多尺度注意力机制,旨在高效捕获图像中的局部与全局上下文信息。该模块通过并行的多尺度分支和注意力融合策略,对不同空间尺度的特征进行自适应建模,从而增强模型对岩心图像中复杂结构、边界细节以及小目标的感知能力。其网络结构如图 4 所示。

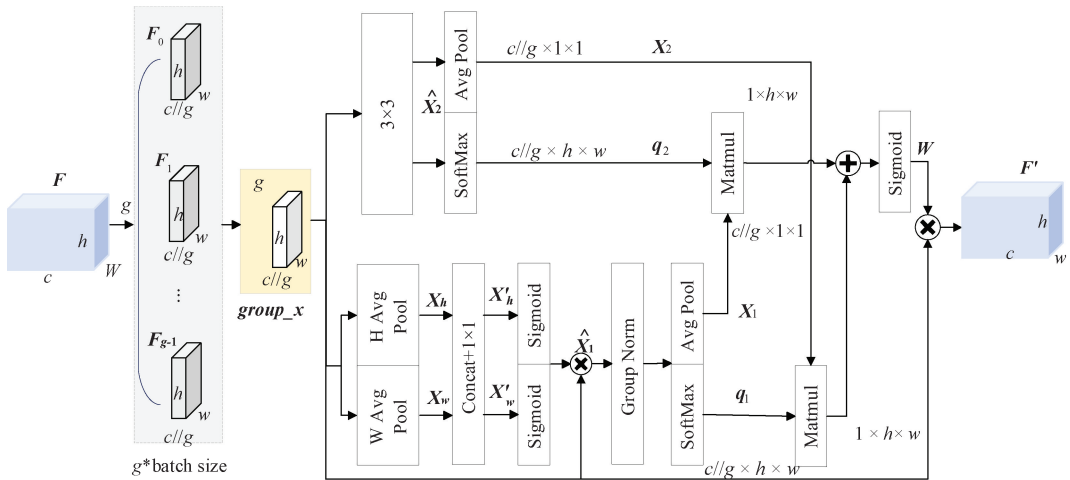


图 4 EMA 结构

Fig. 4 Structure of EMA

首先,对于输入特征图 $F \in \mathbf{R}^{C \times H \times W}$, EMA 将其划分为 g 个子特征组,表示为: $F = \{F_0, F_1, F_2, \dots, F_{g-1}\}$,

$F_i \in \mathbf{R}^{C//g \times H \times W}$ 。这种分组策略有助于增强局部语义区域的建模能力,同时有效抑制背景噪声干扰,提升特征表达的鲁棒性。

随后,EMA 通过 3 个并行分支对每个子特征组 $group_x$ 进行多尺度注意力建模。首先,将每个子特征图在水平和垂直方向上分别进行 *Average Pooling*,生成两个单维度的空间注意力向量,可表示为: $\mathbf{x}_w = AvgPool(\mathbf{x}) \in \mathbf{R}^{C//g \times 1 \times W}$, $\mathbf{x}_h = AvgPool(\mathbf{y}) \in \mathbf{R}^{C//g \times H \times 1}$, 然后进行拼接,并通过 1×1 卷积融合生成权重特征 h_w 。接着,将其切分为方向权重 \mathbf{x}'_h 和 \mathbf{x}'_w ,并通过 *Sigmoid* 激活后与 $group_x$ 相乘,从而生成方向增强的特征 $\hat{\mathbf{x}}_1$ 。除方向增强外,EMA 还使用一个 3×3 卷积分支提取局部空间结构特征,表示为 $\hat{\mathbf{x}}_2$,以增强模型对小目标和细节的感知能力。接下来,EMA 将方向增强特征 $\hat{\mathbf{x}}_1$ 和局部卷积特征 $\hat{\mathbf{x}}_2$ 然后通过 *Average Pooling* 得到 \mathbf{x}_1 和 \mathbf{x}_2 ,然后又通过归一化指数函数得到每一路径的注意力权重 q_1 和 q_2 ,然后将 (q_1, \mathbf{x}_2) 、 (q_2, \mathbf{x}_1) 做点积操作,进行权重进行融合,生成最终的注意力权重图 $\mathbf{W} = sigmoid(q_1 \cdot \mathbf{x}_2 + q_2 \cdot \mathbf{x}_1)$ 。最后,将原始特征 $group_x$ 与权重 \mathbf{W} 进行逐点相乘,恢复原尺寸。

2.4 MSDE 模块

MSDE 通过融合多尺度卷积、轻量级注意力机制与残差连接策略,增强岩心图像中多尺度特征的表达能力及语义的一致性。其结构如图 5 所示。主要包括 4 个步骤:

1) 多尺度特征提取

首先,输入特征 $\mathbf{F} \in \mathbf{R}^{C \times H \times W}$ 分别通过 4 个为 1×1 卷积、 3×3 卷积、 5×5 卷积和膨胀卷积(dilated convolution)的并行卷积分支,每个支路使用不同的感受野对岩心特征

进行提取。随后,将这 4 个分支在通道维度上进行拼接得到 \mathbf{F}_{cat} , 计算公式为:

$$\mathbf{F}_{cat} = Concat(\mathbf{F}_1, \mathbf{F}_3, \mathbf{F}_5, \mathbf{F}_d) \quad (9)$$

2) 通道注意力融合

为自适应增强岩心多尺度特征中的关键结构响应并抑制无关特征,本文设计了一种轻量级注意力模块对拼接后的特征进行加权调节。该模块融合了压缩与激励模块(squeeze-and-excitation networks, SENet)^[16]和 CBAM 的思想,针对岩心图像进行了简化设计,仅保留 SENet“特征重标定”的核心思想与“降维—激活—升维”的瓶颈结构,舍弃了全局池化,直接对含空间信息的 \mathbf{F}_{cat} 进行操作,既减少计算量,又通过保留空间维度提升对局部特征的自适应能力。同时,简化 CBAM 空间注意力实现方式,不单独设空间分支,而是通过多尺度卷积引入空间感受野差异,结合通道注意力实现对“空间—通道”联合特征的隐式加权。具体而言,特征先经 1×1 卷积降维,再通过 ReLU 激活与 1×1 卷积恢复通道数,最终经 *Sigmoid* 生成通道注意力权重,与输入特征逐通道相乘,得到加权输出 \mathbf{F}_{att} 。计算公式为:

$$\mathbf{F}_{att} = \mathbf{F}_{cat} \otimes \sigma(\text{Conv}_{1 \times 1}(\text{ReLU}(\text{Conv}_{1 \times 1}(\mathbf{F}_{cat})))) \quad (10)$$

3) 残差连接

为避免多尺度特征提取与注意力加权过程中的信息丢失,本部分引入残差连接来保留原始输入特征。本文使用 1×1 卷积对输入 \mathbf{F} 进行通道调整,并与注意力加权后的特征图相加,实现残差增强。计算公式为:

$$\mathbf{F}_{res} = \mathbf{F}_{att} + \text{Conv}_{1 \times 1}(\mathbf{F}) \quad (11)$$

4) 最终特征融合

为融合残差增强特征与输入信息,本文将残差特征与原始输入特征在通道维度上进行拼接,并通过一个 1×1 卷积压缩通道数至目标维度。然后,配合批归一化、ReLU 和正则化处理来提升稳定性和鲁棒性。计算公式为:

$$\mathbf{F}' = \text{Dropout}(\text{ReLU}(\text{Conv}_{1 \times 1}(\text{Concat}(\mathbf{F}_{res}, \mathbf{F})))) \quad (12)$$

3 实验与分析

3.1 数据集介绍

鉴于目前缺乏用于岩心区域提取的公开数据集,本文自主构建了一个相关数据集。该数据集由数字化岩心图像采集系统获取,包含大量背景信息的白光平扫、白光滚扫和荧光图像,共分别采集了 135、92 和 85 张图像。本文对这些图像进行了筛选、降采样和分块处理,最终得到容量为 3 949 张、尺寸为 $512 \text{ pixel} \times 512 \text{ pixel}$ 的岩心图像。然后使用 Labelme 工具进行了标注,并将标注结果转化为二值化掩模,形成了最终的岩心数据集。为进行模型的训练与评估,本文按 8:2 的比例随机划分为训练集和测试集。

为验证模型的泛化能力,本文在公开医学图像数据集

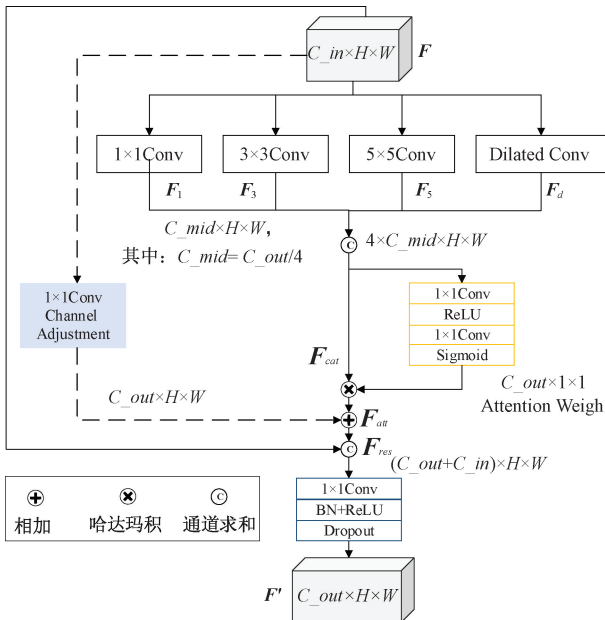


图 5 MSDE 结构

Fig. 5 Structure of MSDE

BRISC 2025^[17]和 EBHI-Seg^[18]上进行对比实验。

3.2 实验环境与评价指标

实验在 Ubuntu 22.04.5 LTS 操作系统环境下,采用 Python 编程语言和深度学习框架进行实验。硬件平台包括 Intel (R) Core (TM) i5-11500 处理器和 NVIDIA GeForce RTX 2080 Ti 显卡,软件环境为 CUDA 12.4、Python 3.9.18 以及 PyTorch 1.12.1。

本文采用 Adam 优化器,初始学习率为 1×10^{-4} ,权重衰减为 1×10^{-4} ,结合余弦退火调度策略,最大学习周期为 100,最小学习率为 1×10^{-5} 。在岩心和 BRISC 2025 数据集上,批次大小为 8,训练轮数为 80;在 EBHI-Seg 数据集上,批次大小为 16,训练轮数为 100。在损失函数方面,采用 Dice 损失与二值交叉熵损失的加权组合形式,其数学表达式为:

$$L = 0.5 \cdot BCE(\hat{y}, y) + Dice(\hat{y}, y) \quad (13)$$

3.3 评价指标

1) 性能评价

为全面评估模型的性能,本文采用了多个评价指标,包括平均交并比(mean intersection over union, mIoU)、F1 分数(F1-score)、平均像素准确率(mean pixel accuracy, mPA)和均方误差(mean squared error, MSE)。此外,还引入了模型参数量(parameters)、推理时间(inference time)以及每秒帧数(frames per second, FPS)等指标。

mIoU 是衡量模型对所有类别预测的结果和真实值的平均交并比。mPA 是指各类别中被正确分类的像素所占比例的平均值;F1-score 是指精确率和召回率的调和平均值。MSE 是预测图与真实图像之间每个像素差值的平方的平均值。各指标表达如式(14)~(17)所示。

$$mIoU = \frac{1}{2} \left(\frac{TP}{TP + FP + FN} + \frac{TN}{TN + FN + FP} \right) \times 100\% \quad (14)$$

$$F1 = 2 \frac{\frac{TP}{TP + FP} \times \frac{TP}{TP + FN}}{\frac{TP}{TP + FP} + \frac{TP}{TP + FN}} \times 100\% \quad (15)$$

$$mPA = \frac{1}{2} \left(\frac{TP}{TP + FP} + \frac{TN}{TN + FN} \right) \times 100\% \quad (16)$$

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I_{pred}(i, j) - I_{true}(i, j))^2 \quad (17)$$

式中: TP 表示模型预测和实际均为正例; TN 表示模型预测和实际均为反例; FP 表示模型预测为正例、实际为反例; FN 表示模型预测为反例、实际为正例; $I_{pred}(i, j)$, $I_{true}(i, j)$ 分别表示像素 (i, j) 处的预测值和真实值。

parameters 表示模型中可训练的参数总数。inference time 表示模型处理单张图像的平均时间。FPS 是衡量系统处理速度的指标,表示每秒能够完成完整推理的帧数。

2) 统计学方法

本文采用 SPSS 25.0 软件进行统计分析。为验证改进模型与各对比模型性能差异的统计显著性,采用配对样本 t 检验进行比较,以 $P < 0.05$ 为差异具有统计学意义。

3.4 实验结果及分析

为验证改进模型的有效性和泛化性能,本文在相同实验环境下,在自建岩心数据集和公开数据集上同其他算法进行了对比实验。本文所采用的对比算法包括 UNeXt, Unet, TransAttUnet^[19], Tyche^[20], SegNeXt^[21], SegFormer^[22]和 DDRNet^[23]。此外,本文还设计了消融实验,以验证各个改进部分的有效性和合理性。

1) 岩心数据集上的对比实验结果与分析

在本文的所有表格中,加粗数据为最优值;“↓”表示数值越低性能越好,“↑”表示数值越高性能越好。

表 1 的实验结果表明,本文算法在 mIoU、F1-score、mPA 和 MSE 4 项指标上分别达到了 99.01%、98.90%、99.60% 和 0.001 87,均优于其他对比方法。与原始 UNeXt 架构相比,本文算法在前 3 项指标上分别提升了 1.49%、2.06% 和 0.75%,MSE 指标降低了 78.63%,充分验证了改进方案在提升分割精度方面的有效性。在模型复杂度方面,尽管参数量略有增加,但仍低于部分对比模型。在推理速度方面,模型的 FPS 达到 121.9,仅次于 UNeXt 和 DDRNet,显著高于 Tyche 和 TransAttUnet 等方法,具备较强的实时处理能力。

表 1 岩心数据集上的对比实验结果

Table 1 Compare results on the core datasets

方法	mIoU/% ↑	F1-score/% ↑	mPA/% ↑	MSE/($\times 10^{-3}$) ↓	params/M ↓	time/ms ↓	FPS ↑
Unet	96.37*	96.90*	98.42*	13.92*	13.39	24.56±0.06	40.7
DDRNet	93.67*	94.52*	96.27*	27.01*	5.69	5.72±0.14	174.9
TransAttUnet	98.18*	97.80*	99.20*	5.27*	22.65	57.13±0.10	17.5
Tyche	96.92*	97.33*	98.63*	11.23*	1.11	34.58±0.10	28.9
SegNeXt	97.12*	92.73*	98.36*	10.25*	4.22	13.09±0.34	76.4
SegFormer	98.29*	98.12*	99.07*	4.72*	7.71	9.14±0.22	109.4
UNeXt	97.52*	96.84*	98.85*	8.75*	1.47	4.24±0.10	235.9
本文算法	99.01	98.90	99.60	1.87	2.42	8.20±0.24	121.9

注: * 表示与本文改进算法组比较, $P < 0.05$

为验证结果的统计显著性,本文采用配对样本 t 检验对改进算法与 7 个对比模型进行系统比较。实验结果表明,本文算法在 mIoU、F1-score、mPA 和 MSE 4 个指标上均显著优于所有对比模型(均 $P < 0.05$),证实了性能提升的统计学意义。

综上,本文算法在显著提升分割性能的同时,依然保持了较低的计算开销,兼具高精度、轻量化和高效率。

图 6 展示了多种方法在岩心图像数据集上的预测效果,选取了 4 张形态不规则、类型不同的岩心图像作为对比样本。

在第 1 幅图像中,UNet 和 SegNeXt 存在大量背景误识别现象,而 DDRNet 未能完整提取全部岩心区域;在第 2 幅图像中,由于岩心与背景灰度接近,仅 SegNeXt 和本文方法准确分割出了目标区域;在第 3 幅图像中,DDRNet 仅提取了岩心的轮廓部分,分割不完整;在第 4 幅荧光图像中,SegNeXt 出现明显漏检,未识别出岩心主体结构。综合来看,对比算法在面对岩心图像中边界模糊、多尺度结构与复杂背景干扰等挑战时,均存在不同程度的识别偏差,表现出一定的局限性。而本文方法能够实现岩心目标的完整提取。

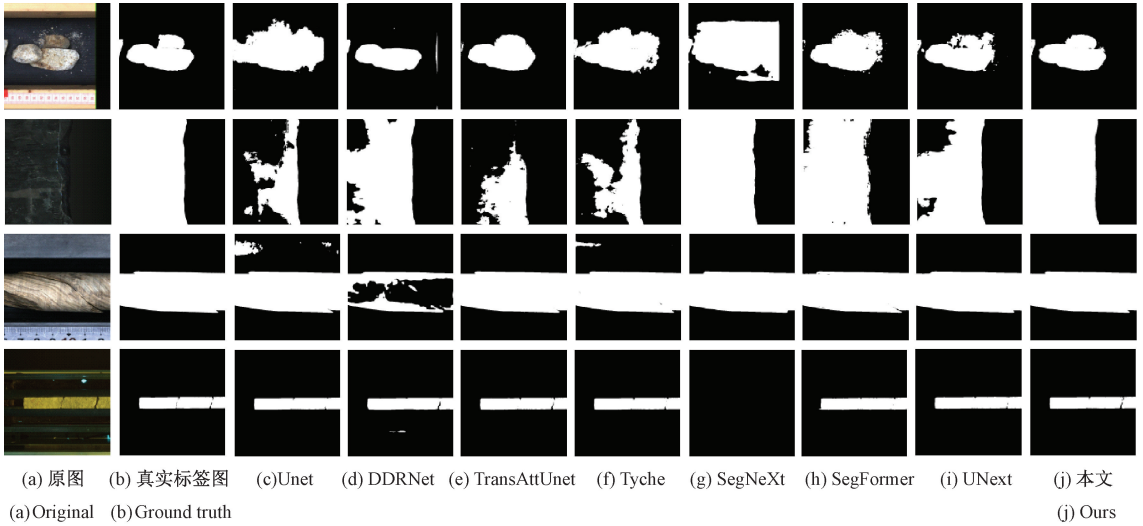


图 6 不同算法在岩心数据集上的分割效果对比

Fig. 6 Comparison of segmentation results on the core datasets

2) 公开数据集上的对比实验结果与分析

为了验证本文算法的泛化性能,本文在 BRISC 2025 和 EBHI-Seg 两个公开数据集上进行了对比实验,实验结果如表 2 所示。在分割精度方面,本文算法在 BRISC 2015 数据集上取得了 mIoU 86.65%、F1-score 81.20%、mPA 89.04% 与 MSE 0.006 12 的优异表现,全面优于其他对比模型。在 EBHI-Seg 数据集上,算法同样表现优异,各项指标分别为 89.29%、96.33%、95.26% 与 0.049 55,仅 mPA

略低于 TransAttUnet,展现出良好的跨场景适应能力。此外,统计显著性分析进一步证实了上述优势的可靠性,实验结果显示,本文算法在绝大多数比较中均达到统计显著水平(均 $P < 0.05$)。具体而言,在 BRISC 2015 数据集中,除与 UNeXt 在 MSE 指标上无显著差异外,其余比较均具统计学意义;在 EBHI-Seg 数据集中,除与 TransAttUnet 在 mPA 指标上无显著差异外,其余比较亦具有统计学意义。

表 2 公开数据集上的对比实验结果

Table 2 Compare results on public datasets

方法	BRISC 2015				EBHI-Seg			
	mIoU/% ↑	F1-score/% ↑	mPA/% ↑	MSE/($\times 10^{-3}$) ↓	mIoU/% ↑	F1-score/% ↑	mPA/% ↑	MSE/($\times 10^{-3}$) ↓
Unet	78.79*	65.78*	80.05*	9.69*	88.80*	96.12*	95.32*	52.29*
DDRNet	84.75*	78.14*	86.25*	6.94*	82.35*	93.38*	91.29*	85.83*
TransAttUnet	76.04*	69.95*	83.91*	13.49*	89.06*	96.23*	95.30	50.86*
Tyche	80.08*	71.81*	82.23*	9.33*	87.51*	95.63*	94.12*	58.62*
SegNeXt	82.73*	73.18*	85.52*	8.10*	61.10*	85.30*	72.90*	205.86*
SegFormer	81.87*	73.63*	85.17*	8.70*	78.52*	92.04*	88.76*	107.11*
UNeXt	86.32*	79.87*	88.41*	6.22	85.66*	94.89*	93.49*	68.47*
本文算法	86.65	81.20	89.04	6.12	89.29	96.33	95.26	49.55

注: * 表示与本文改进算法组比较, $P < 0.05$

综上,实验结果表明,本文算法在不同场景下均能保持稳定的性能优势,具备良好的跨域适应能力与鲁棒性。

图 7 展示了两个公开数据集部分样本的可视化分割结果。其中,第 1 和 2 幅图为 BRISC 2015 数据集样本,整体分割效果与表 2 实验结果一致。UNeXt 与本文方法表现最佳,但在边界细节与局部区域的处理上,本文算法更

为精准。第 3 和 4 幅图来自 EBHI-Seg 数据集。其中,SegNeXt 分割效果最差,尤其在第 4 幅图中几乎未能识别目标区域。并且多数对比模型在目标内部区域存在空洞或漏检现象,而本文算法在目标提取的完整性与边界精度方面表现更优,分割结果与真实标签图更为接近,边缘过渡自然,细节保持更完整。

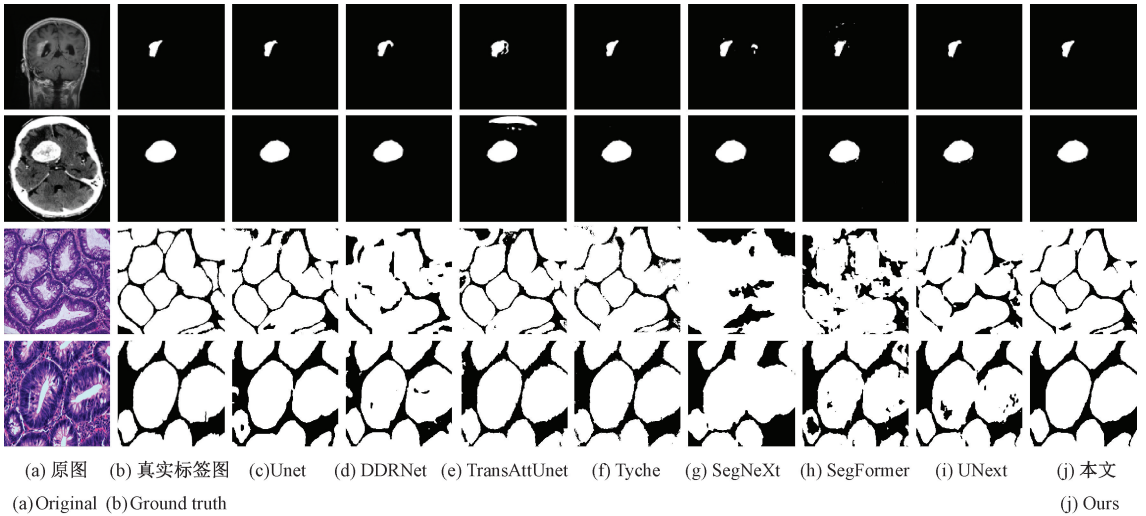


图 7 公开数据集上的分割效果对比

Fig. 7 Comparison of segmentation results on public datasets

3) 消融实验结果与分析

为验证本文算法各模块的有效性 with 合理性,本文在相

同实验条件下进行了 8 组消融实验,实验结果如表 3 所示,其中“N”表示不加入对应模块,“Y”表示加入对应模块。

表 3 消融实验结果

Table 3 Results of ablation experiments

CBAM	EMA	MSDE	mIoU/% ↑	F1-score/% ↑	mPA/% ↑	MSE/($\times 10^{-3}$) ↓	params/M ↓	time/ms ↓	FPS ↑
N	N	N	97.52	96.84	98.85	8.75	1.47	4.24 ± 0.10	235.9
Y	N	N	98.42	97.25	99.13	4.31	2.08	6.67 ± 0.04	149.9
N	Y	N	98.19	97.20	99.04	5.45	2.13	5.78 ± 0.17	173.0
N	N	Y	98.28	97.17	99.10	5.02	2.34	5.16 ± 0.12	193.9
Y	Y	N	98.76	98.46	99.31	2.52	2.14	7.44 ± 0.20	134.5
Y	N	Y	98.53	98.27	99.26	3.61	2.36	7.35 ± 0.08	136.0
N	Y	Y	98.65	98.43	99.24	3.05	2.41	6.75 ± 0.17	148.2
Y	Y	Y	99.01	98.90	99.60	1.87	2.42	8.20 ± 0.24	121.9

实验结果表明,与原始 UNeXt 相比,单独引入 CBAM、EMA 或 MSDE 模块均能提升模型性能,mIoU、F1-score 和 mPA 最高分别提升 0.90%、0.41% 和 0.28%,MSE 最多下降 50.74%,说明各模块在局部细节、上下文建模和多尺度融合方面均具积极作用。此外,如表 3 所示,模块组合能带来更显著的性能提升。在第 6 组实验中,CBAM 与 EMA 联合使用使 mIoU 提升至 98.76%,F1-score 达 98.46%,mPA 达 99.31%,MSE 降低 71.2%。在第 8 组实验中,3 个模块共同作用下性能进一步优化,验证了各模块的协同优势。

为验证消融实验结果和展示各模块在特征关注区域的差异,采用 Grad-CAM 对模块进行了注意力可视化分析,结果如图 8 所示。如图 8(c)所示,CBAM 主要聚焦于岩心边缘与纹理区域,强化空间细节提取,因而在单模块使用中效果最佳。如图 8(d)所示,EMA 响应主要集中于岩心内部,侧重于中等尺度结构与跨空间上下文信息的建模。而如图 8(e)所示,MSDE 能同时捕捉边缘与内部信息,覆盖多尺度区域,其单模块使用性能仅次于 CBAM。在多模块组合中,如图 8(f)所示,CBAM 与 EMA 联合实现了边缘与内部信息的协同增强,因此性能改善显著。如

图 8(g)和(h)所示, CBAM 与 MSDE 的组合虽强化了边缘信息,但对内部上下文捕捉不足,而 MSDE 与 EMA 的组合则更侧重于内部信息,边缘感知较弱。而如图 8(i)所

示,三者共同作用时,形成了对岩心边缘与内部细节的全面覆盖,展现出最优的协同增强效果。因此,从视觉可视化角度进一步证明了本文模型改进的有效性。

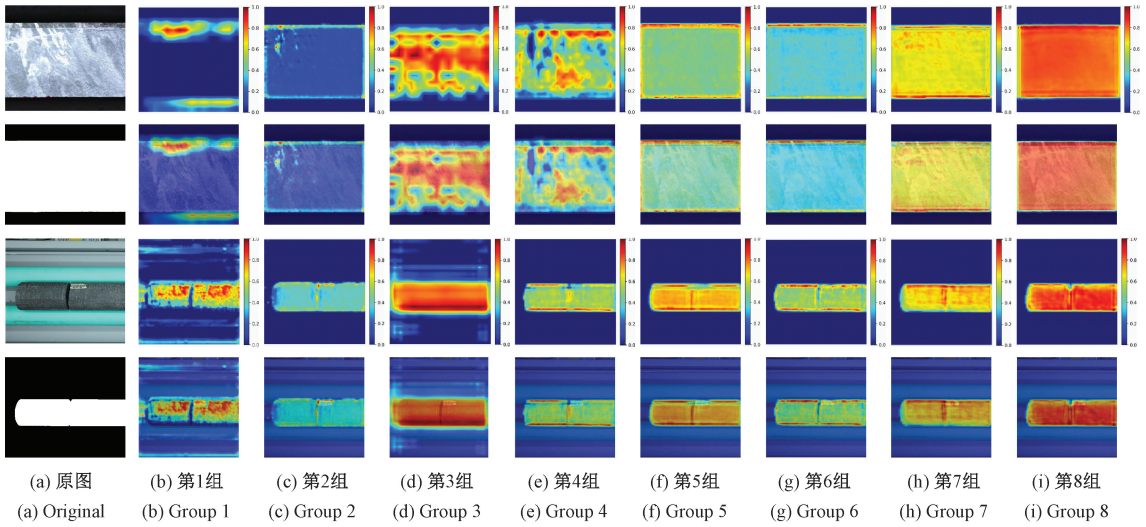


图 8 改进模块的消融 Grad-CAM 热力图

Fig. 8 Rad-CAM heatmap of the ablation of the improved module

4 结 论

本研究基于轻量化的 UNeXt 网络,提出一种改进的岩心目标提取算法,以应对岩心图像中的边界模糊、纹理复杂及背景干扰等问题。该算法引入多种注意力机制和结构增强模块,有效提升了分割精度与鲁棒性。首先,在编码器的浅层引入了 CBAM 模块,增强低级特征对边缘和纹理的感知能力;然后,在编码器的中高层引入 EMA 模块,加强对全局上下文建模与细粒度区域的感知;此外,在网络颈部设计并引入 MSDE 模块,用于提升多尺度特征融合与复杂结构的辨析能力;最后,在解码器深层再次引入 CBAM 模块,增强模型对目标区域和结构细节的关注,进一步提高分割精度与结构连续性。实验结果表明,本文算法在自建岩心数据集和公开数据集上均表现优异,具有良好的跨场景泛化能力,且消融实验验证了各模块的有效性。在未来的研究中,将在保证良好分割性能的前提下,进一步探索模型的轻量化设计,以实现更高效的岩心图像处理。

参考文献

- [1] 姚成武,常琛. 基于像素强度顺序变换和 UNetFormer 的裂缝分割模型[J]. 电子测量技术, 2024, 47(11): 151-159.
YAO CH W, CHANG CH. Pavement crack segmentation model based on pixel-intensity order transform and UNetFormer[J]. Electronic Measurement Technology, 2024, 47(11): 151-159.
- [2] 刘庭亭,何志琴. 基于 IZOA 结合最小交叉熵的图像

- 分割算法[J]. 电子测量技术, 2025, 48(16):40-53.
LIU T T, HE ZH Q. Based on IZOA combined with minimum cross-entropy image segmentation algorithm[J]. Electronic Measurement Technology, 2025, 48(16): 40-53.
- [3] 吴京城,洪欢欢,施露露,等. 反背景差分结合 Otsu 的细胞图像分割方法[J]. 电子测量与仪器学报, 2021, 35(4):82-89.
WU J CH, HONG H H, SHI L L, et al. Cell image segmentation method using anti-background difference combined with Otsu [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(4): 82-89.
- [4] SHELHAMER E, LONG J T, DARRELL T. Fully convolutional networks for semantic segmentation[C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [5] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. Computer Science, 2014(4): 357-361.
- [6] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834-848.
- [7] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for

- semantic image segmentation [C]. European Conference on Computer Vision, 2018: 833-851.
- [8] 冯哲,王彬,黄鹏程,等. 小天体表面着陆区岩石目标检测算法[J]. 仪器仪表学报, 2024, 45(4):194-205. FENG ZH, WANG B, HUANG P CH, et al. Algorithm of detection rock object in landing zone of smallcelestial body surface [J]. Chinese Journal of Scientific Instrument, 2024, 45(4): 194-205.
- [9] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015: 234-241.
- [10] 曾中华,曹东. 基于改进 Unet 的矿石图像分割[J]. 电子测量技术, 2023,46(21):176-182. ZENG ZH H, CAO D. Core image segmentation based on improved Unet[J]. Electronic Measurement Technology, 2023, 46(21): 176-182.
- [11] 舒小锋,吴晓红,卿粼波,等. 基于改进 U²Net 的岩石薄片图像分割[J]. 计算机系统应用, 2024, 33(2): 159-165. SHU X F, WU X H, QING L B, et al. Rock slice image segmentation based on improved U²Net[J]. Computer Systems and Applications, 2024, 33(2): 159-165.
- [12] QIN X B, ZHANG Z CH, HUANG CH Y, et al. U²-Net: Going deeper with nested U-structure for salient object detection [J]. Pattern Recognition, 2020, 106: 107404.
- [13] ALANARASU J M J, PATEL V M. UNeXt: MLP-based rapid medical image segmentation network[C]. International Conference on Medical Image Computing and Computer-Assisted Intervention, 2022: 23-33.
- [14] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. European Conference on Computer Vision. Munich, 2018: 3-19.
- [15] OUYANG D L, HE S, ZHANG G ZH, et al. Efficient multi-scale attention module with cross-spatial learning [C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing, 2023: 1-5.
- [16] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. 2018 IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [17] FATEH A, REZVANI Y, MOAYEDI S, et al. BRISC: Annotated dataset for brain tumor segmentation and classification with swin-HAFNet [J]. ArXiv preprint arXiv:2506.14318, 2025.
- [18] SHI L Y, LI X Y, HU W M, et al. EBHI-Seg: A novel enteroscope biopsy histopathological hematoxylin and eosin image dataset for image segmentation tasks[J]. Frontiers in Medicine, 2023, 10: 1114673.
- [19] CHEN B ZH, LIU Y SH, ZHANG ZH, et al. TransAttUnet: Multi-level attention-guided U-net with transformer for medical image segmentation[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2024, 8(1): 55-68.
- [20] RAKIC M, WONG H E, GONZALEZ ORTIZ J J, et al. Tyche: Stochastic in-context learning for medical image segmentation [C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 11159-11173.
- [21] GUO M H, LU CH Z, HOU Q B, et al. SegNeXt: Rethinking convolutional attention design for semantic segmentation [J]. ArXiv preprint arXiv: 2209.08575, 2022.
- [22] XIE EN Z, WANG W H, YU ZH D, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers [C]. 35th International Conference on Neural Information Processing Systems, 2021: 12077-12090.
- [23] PAN H H, HONG Y D, SUN W CH, et al. Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes [J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(3): 3448-3460.

作者简介

郑方艳, 硕士, 主要研究方向为图像处理与识别、计算机视觉等。

E-mail: ZHENGfangyanab@outlook.com

何小海(通信作者), 博士, 教授, 主要研究方向为图像处理与网络通信、人工智能与大数据分析、图像通信等。

E-mail: hxxh@scu.edu.cn

卿粼波, 博士, 教授, 主要研究方向为多媒体通信与信息系系统、人工智能与计算机视觉、嵌入式系统等。

E-mail: qing_lb@scu.edu.cn

何海波, 学士, 高级工程师, 主要研究方向为图像处理、计算机应用等。

E-mail: cdxtkj@126.com

滕奇志, 博士, 教授, 主要研究方向为数字图像处理、模式识别、三维图像重建及分析等。

E-mail: qzteng@scu.edu.cn