

文章编号:1007-2780(2025)09-1321-12

基于改进 EDTER 算法的河流边缘检测方法

高钰栋^{1,2}, 黄凌霄^{1,2*}, 姚新波^{1,2}, 赵苑汝³

(1. 宁夏大学 信息工程学院, 宁夏 银川 750021;

2. 宁夏“东数西算”人工智能与信息安全重点实验室, 宁夏 银川 750021;

3. 宁夏大学 土木与水利工程学院, 宁夏 银川 750021)

摘要: 为了提高遥感图像中河流边缘提取的精度与鲁棒性, 构建了一种基于改进 EDTER 算法的河流边缘检测方法。针对传统方法在复杂自然场景下存在边缘细节丢失、噪声干扰等问题, 提出一种适应性更强的边缘检测方案。采用以 Swin Transformer 为主干网络的两阶段边缘检测结构, 引入残差多层感知机以增强特征表达能力, 并设计轻量化多尺度解码器提升检测效率。第一阶段提取图像全局特征, 生成全局边缘图; 第二阶段进行局部特征提取并与全局特征融合, 生成最终边缘检测结果。在 BSDS500 数据集和自建的跨区域河流边缘数据集上开展对比实验, 结果显示本文方法在 ODS、OIS 和 AP 三个指标上均优于主流方法。其中, 在跨区域河流边缘数据集上的 AP 值达到 0.871, 较 EDTER 方法提升约 3.2%, 说明该方法在复杂遥感场景中表现出更优的边缘提取能力。本文方法能够有效兼顾检测精度与计算效率, 具备良好的泛化性能, 适用于多类型遥感图像中的水体边缘提取任务, 具有较高的工程应用价值。

关键词: 河流边缘; 边缘检测; 遥感图像; Swin Transformer; 水体提取

中图分类号: TP391 文献标识码: A doi:10.37188/CJLCD.2025-0107 CSTR:32172.14.CJLCD.2025-0107

River edge detection based on improved EDTER algorithm

GAO Yudong^{1,2}, HUANG Lingxiao^{1,2*}, YAO Xinbo^{1,2}, ZHAO Yuanru³

(1. School of Information Engineering, Ningxia University, Yinchuan 750021, China;

2. Key Laboratory of Artificial Intelligence and Information Security in the “East Data, West Computing” Initiative of Ningxia, Yinchuan 750021, China;

3. School of Civil and Hydraulic Engineering, Ningxia University, Yinchuan 750021, China)

Abstract: To improve the accuracy and robustness of river edge extraction in remote sensing images, a river edge detection method based on an improved EDTER algorithm is proposed. Addressing issues such as edge detail loss and noise interference in complex natural scenes with traditional methods, a more adaptive edge detection scheme is introduced. The method employs a two-stage edge detection framework with Swin Transformer as the backbone network, incorporating a residual multilayer perceptron to enhance feature representation and designing a lightweight multi-scale decoder to improve detection efficiency. The first stage extracts global image features to generate a global edge map, while the second stage performs

收稿日期: 2025-05-21; 修订日期: 2025-07-09.

基金项目: 国家自然科学基金(No.12462027); 宁夏回族自治区教育厅高等学校科学研究项目(No.NYG2024033)

Supported by National Natural Science Foundation of China (No.12462027); Scientific Research Project of the Education Department of Ningxia Hui Autonomous Region (No.NYG2024033)

*通信联系人, E-mail: huanglx@nxu.edu.cn

local feature extraction and fuses it with global features to produce the final edge detection result. Comparative experiments on the BSDS500 dataset and a self-constructed cross-regional river edge dataset demonstrate that the proposed method outperforms mainstream approaches in ODS, OIS, and AP metrics. Notably, the AP value on the cross-regional river edge dataset reaches 0.871, an improvement of approximately 3.2% over the EDTER method, indicating superior edge extraction capability in complex remote sensing scenarios. The proposed method effectively balances detection accuracy and computational efficiency, exhibiting strong generalization performance and suitability for water body edge extraction tasks in various remote sensing images, with significant engineering application value.

Key words: river edge; edge detection; remote sensing images; Swin Transformer; waterbody extraction

1 引言

河流边缘线作为水体与陆地的分界线,其精准检测对于水资源管理、生态保护及防洪减灾具有重要意义^[1]。随着全球气候变化和人类活动加剧,河流形态动态监测已成为流域综合管理的关键环节。传统人工实地测量虽精度较高,但存在成本高、效率低等局限性,难以满足大范围、高频次监测需求。卫星遥感技术凭借其覆盖范围广、时效性强、成本低等优势^[2],为河流边缘检测提供了重要数据支撑,但如何从复杂遥感影像中稳定、精确地提取河流边缘仍面临诸多挑战。

现有的河流边缘检测方法主要分为3类:基于传统图像处理的方法、基于深度学习的方法和混合方法。传统方法中,梯度算子(如 Sobel、Canny)和数学形态学方法依赖局部强度变化检测边缘。余加俊等人^[3]结合 Sobel 算子和直线拟合技术实现河道线的检测,但其结果易受噪声干扰。黄凌霄等人^[4]使用数学形态方法检测水库边缘,在简单场景下效果良好。基于深度学习的方法中,卷积神经网络(CNN)的边缘检测方法,如 HED(Holistically-Nested Edge Detection)和 RCF(Richer Convolutional Features)通过多层次特征提取显著提升了精度。郑又能等人^[5]通过改进区域生长算法实现水岸线提取,但计算量较大。Vandaele 等人^[6]采用语义分割方法检测水岸线。Yuheng Yin 等人^[7]利用增强型 PSPNet 进行水岸线检测。混合方法通过融合传统图像处理与机器学习技术显著提升了水体边缘检测精度。Gazi 等人^[8]结合图像融合技术和机器学习分类方法,成功检测了 Chalan Beel 地区的水体变化,验证了

混合方法的有效性。Jain 等人^[9]提出了一种基于正则化各向异性扩散和近端支持向量机的混合方法,在含噪遥感图像中实现了高效的边缘检测。尽管这些方法取得了一定进展,但在复杂自然环境中仍存在明显不足:(1)对影像质量敏感,易受噪声、阴影、云层等干扰;(2)边界信息模糊时分割精度下降,传统 CNN 方法易丢失边缘细节^[10-11];(3)依赖人工参数调整,泛化能力有限,难以适应不同类型的遥感影像。特别是在植被覆盖、水面反射等复杂场景下,现有方法往往导致边缘不连续或误检率较高。

近年来, Vision Transformer (ViT)^[12-13]在计算机视觉领域展现出强大性能,基于 Transformer 的边缘检测方法如 EDTER^[14](Edge Detection with Transformer),通过两阶段设计(全局上下文建模和局部精化)在标准数据集上表现优异,特别是在处理复杂场景时展现了较强的鲁棒性。然而,将其直接应用于河流边缘检测存在3个主要问题:(1)缺乏针对水体边缘多尺度特性的专门设计;(2)原始 ViT-large 模型的参数量大,计算复杂度高,特别是在遥感影像中,庞大的计算开销成为限制其实际应用的瓶颈;(3)未针对河流边缘检测任务优化,在植被遮挡或水质变化场景下性能受限。为解决 ViT 的高计算复杂度和多尺度特性不足的问题, Swin Transformer^[15]使用一种高效的 Transformer 架构,通过滑动窗口注意力机制降低计算复杂度,并利用分层特征提取构建多尺度特征金字塔,显著提升了复杂场景下的特征提取能力。Swin Transformer 的轻量化变体(如 Swin-T (Swin Transformer Tiny))在参数量和计算效率上具有明显优势,特别适合高分辨率遥感图像的河流边缘检测任务。

基于Swin Transformer的高效特性^[16-18],针对上述问题,本研究提出一种改进的EDTER算法,主要贡献包括:(1)替换主干网络,将ViT替换为Swin-T,通过分层窗口注意力机制降低计算复杂度并增强多尺度特征提取能力;(2)在Swin Transformer块(Swin Transformer Block)中引入残差多层感知机(Res-MLP),优化特征表达和训练稳定性;(3)将EDTER的双向多层聚合解码器(BiMLA)替换为轻量化多尺度解码器(LiteMSD),减少参数量并提升实时性能;(4)构建大规模河流边缘检测数据集,针对复杂遥感场景优化模型性能。实验结果表明,改进后的算法在保持高精度的同时显著提升了运行效率,为水利智能监控提供了新的技术支撑。

2 改进的EDTER河流边缘检测模型

2.1 基于EDTER的网络结构

EDTER是一种基于ViT^[13]的通用边缘检测框架,该模型创新性地利用Transformer架构的双向注意力机制,实现了对图像全局上下文信息和局部细节特征的协同建模。相较于传统卷积神经网络,EDTER能够更有效地捕捉长距离依赖关系,从而生成具有更高语义一致性的物体边界检测结果。

EDTER采用双阶段编码器-解码器架构:

第一阶段是全局特征提取,通过ViT编码器建立图像全局表征,经双向多层聚合解码器(Bi-directional Multi-Level Aggregation, BiMLA)生成高分辨率特征图;第二阶段是局部特征增强,将输入图像划分为多个不重叠子区域,采用局部Transformer编码器提取各子区域的细粒度特征,并通过BiMLA解码器输出像素级特征图。最终通过特征融合模块对两阶段特征图进行融合并输入到决策头得到边缘概率图。

基于引言中提出的改进方向,本研究优化了EDTER的网络结构。改进的网络结构如图1所示,具体处理流程如下:第一阶段(全局特征提取)将输入图像 X 经图像分块(Patch Partition)划分为不重叠的图像块,再通过块嵌入(Patch Embedding)映射到高维特征空间,生成初始特征序列。随后进入Swin-T编码器,生成4个不同尺度的特征图 $S_1 \sim S_4$ 。这些特征经过轻量级多尺度解码器处理,形成全局特征 f_g ,并通过决策头输出初步边缘图 E_g 。第二阶段(局部特征增强)将原始图像划分为4个不重叠子区域 X_1, X_2, X_3, X_4 ,分别经过图像分块和块嵌入后,通过编码器-解码器提取局部特征 f_l 。局部特征经上采样和拼接形成完整特征图后,与全局特征 f_g 通过特征融合模块(Feature Fusion Module, FFM)结合,最终输出更精确的边缘检测结果 E_r 。上述改进增强了模型的多尺度特征提取能力,同时降低了计算复

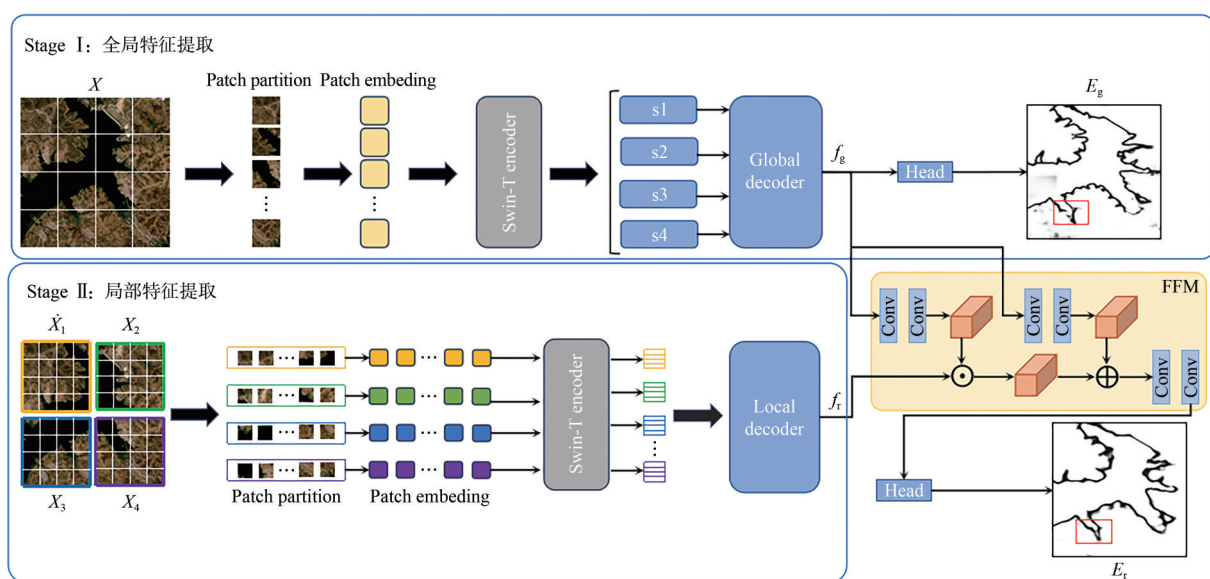


图1 改进的EDTER网络结构

Fig. 1 Network structure of the improved EDTER-based model

杂度,适用于河流边缘检测任务。

2.2 改进的轻量化 Swin-T 主干网络

为了降低计算成本并有效提取图像的多尺度语义信息,本研究选用Swin-T作为特征提取主干网络。如表1所示,与原始 ViT-L(307M 参数量)

相比,Swin-T(29M 参数量)实现了90.9%的参数量压缩,并且FLOPs降低了97.6%(从190.7G降至4.5G)。Swin-T在224×224输入下达到了81.3%的Top-1准确率,在参数量和精度间取得优异平衡,适合高分辨率遥感图像的边缘检测任务。

表 1 不同主干网络在 ImageNet-1K 分类上的对比

Tab. 1 Comparison of different backbone networks on ImageNet-1K classification

模型	图片尺寸/像素	参数量/M	FLOPs/G	吞吐率/(张·s ⁻¹)	准确率/%
ViT-B/16 ^[13]	384×384	86	55.4	85.9	77.9
ViT-L/16 ^[13]	384×384	307	190.7	27.3	76.5
Swin-T ^[13]	224×224	29	4.5	755.2	81.3
Swin-S ^[14]	224×224	50	8.7	436.9	83.0
Swin-B ^[14]	224×224	88	15.4	278.1	83.5
Swin-B ^[14]	224×224	88	47.0	84.7	84.5

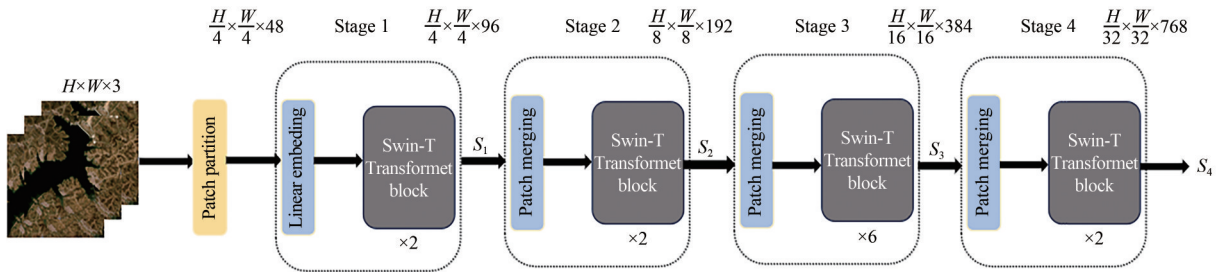


图 2 Swin-T 网络结构

Fig. 2 Network architecture of Swin-T

为进一步优化 Swin-T 的特征表达能力,本研究在 Swin Transformer 块中引入残差多层感知机(Res-MLP)^[19],以提升深层特征提取和训练稳定性。Swin-T 的网络结构如图2所示,其关键机制包括图像块合并和窗口注意力机制的优化。

2.2.1 图像块合并

为了能够对图像内的上下文有更深理解, Swin Transformer 通过图像块合并获取分层特征图来有效表示图像中的不同特征。如图3所示,图像块合并(Patch Merging)会对相邻的两个2×2图像块进行合并,每隔一个位置取一个像素,从而组成4个特征图。将这4个特征图在通道维度上进行拼接,再经过一个归一化层,最后通过一个全连接层在特征图的通道维度上做线性变换,特征图的深度由4×C变成2×C。在图像块合并后,特征图的高和宽会减半,通道数翻倍,达到下

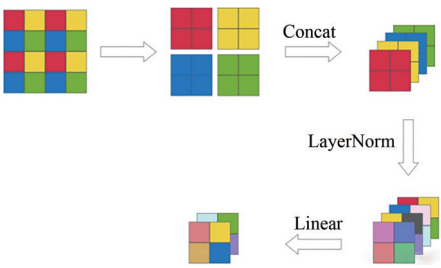


图 3 图像块合并示意图

Fig. 3 Diagram of patch merging process

采样的目的。

2.2.2 改进的窗口注意力机制

为提升 Swin Transformer 在河流边缘检测中的特征表达能力,本研究基于其窗口自注意力机制(Window Multi-Head Self-Attention, W-MSA)引入残差多层感知机(Residual MLP, Res-MLP)^[19],以增强深层特征提取和训练稳定性,从而进一步

优化水体边界检测任务的性能。Swin Transformer的核心特性之一是采用窗口自注意力机制,它通过将输入图像划分为多个固定大小的窗口,并在每个窗口内独立计算自注意力,使得计算量仅与窗口大小相关,而非整张图像的尺寸。如图4所示,输入的特征图会被类似图4(a)分割成多个不重叠的窗口,每个窗口包含 $M \times M$ 个图像块,每个图像块由一个固定大小的块嵌入转换为维度为 C 的特征向量。接着对窗口内的注意力进行计算,在每个窗口使用标准的多头自注意力(MSA)计算所有图像块之间的关系:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{SoftMax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d} + \mathbf{B}}\right)\mathbf{V}, \quad (1)$$

其中: $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 为查询、键、值向量,通过线性变换从输入特征生成; d 是特征维度, $\sqrt{d}$ 用于缩放注意力得分; $\mathbf{B}$ 为相对位置偏置矩阵,沿每个坐标轴在 $[-M+1, M-1]$ 范围内定义,通过一个偏置矩阵 $\hat{\mathbf{B}} \in \mathbb{R}^{(2M-1) \times (2M-1)}$ 进行参数化,最终的偏置矩阵 $\mathbf{B}$ 中的元素由 $\hat{\mathbf{B}}$ 中的对应值获取。为增强窗口间的信息交互,Swin Transformer采用循环移位窗口(Shifted Window Multi-Head Self-Attention, SW-MSA),如图4(b)所示,通过将窗口偏移 $\lfloor M/2 \rfloor$ 个图像块实现相邻窗口的特征交换。这种交替机制在保持计算效率的同时,部分缓解了固定窗口分割带来的特征割裂问题。然而,在河流边缘检测任务中,SW-MSA的局部性限制仍可能导致边缘细节丢失,尤其在多尺度边界或噪声干扰下。

为解决上述问题,本研究引入Res-MLP优化Swin Transformer块,如图5所示。传统的MLP层在通道维度上增强特征表达,但深层网

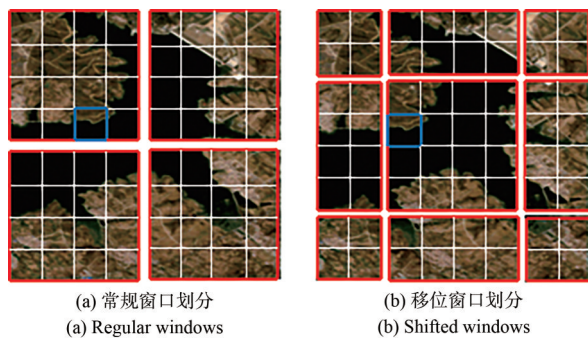


图4 计算注意力的移位窗口示意图

Fig. 4 Illustration of shifted window attention

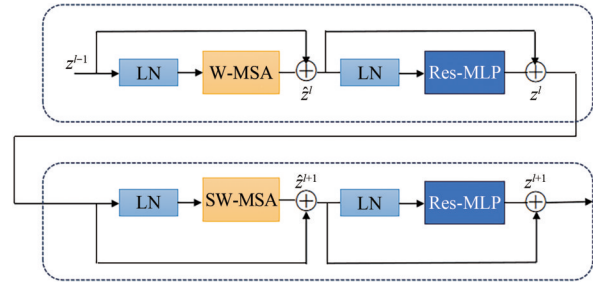


图5 改进的Swin Transformer块结构图

Fig. 5 Architecture of the improved Swin Transformer Block

络易受梯度消失和训练不稳定的影响,尤其在处理高分辨率遥感图像的复杂边缘特征时表现不佳。Res-MLP通过残差连接缓解深层特征提取过程中的梯度消失问题,有效增强了特征的非线性表达能力,从而提升了边缘检测的精度。

整个特征的提取过程包含两个窗口自注意力机制(W-MSA和SW-MSA)与两个MLP模块,均采用残差连接。输入特征首先经过归一化后输入W-MSA模块,通过残差结构得到中间特征。随后,该特征继续经过归一化后输入MLP模块,再次通过残差结构输出特征。接着,采用移位窗口注意力机制增强窗口间信息交互,经过归一化与残差连接生成特征。最终再次通过MLP模块和残差结构输出块特征。如图6所示,MLP部分采用双层线性变换,内部嵌入GELU激活函数与Dropout以提升特征的非线性表达能力,其计算公式如下:

$$\hat{Z}^l = \text{W-MSA}(\text{LN}(Z^{l-1})) + Z^{l-1}, \quad (2)$$

$$Z^l = \text{ResMLP}(\text{LN}(\hat{Z}^l)) + \hat{Z}^l, \quad (3)$$

$$\hat{Z}^{l+1} = \text{SW-MSA}(\text{LN}(Z^l)) + Z^l, \quad (4)$$

$$Z^{l+1} = \text{ResMLP}(\text{LN}(\hat{Z}^{l+1})) + \hat{Z}^{l+1}, \quad (5)$$

$$w = \text{Dropout}(\text{GELU}(\text{Linear}(x))), \quad (6)$$

$$y = \text{LN}(x + \text{Dropout}(\text{Linear}(w))), \quad (7)$$

其中: $\hat{Z}^l$ 和 Z^l 分别表示第 l 个块中W-MSA模块和MLP模块的输出特征;W-MSA和SW-MSA分别表示基于常规窗口划分和偏移窗口划分配的窗口多头自注意力机制; w, y 分别表示Res-MLP模块中非线性变换的中间输出和最终输出,其中Linear操作将输入特征维度从 d 扩展到 $4d$ 。

这种设计显著提升了模型对河流边缘细粒

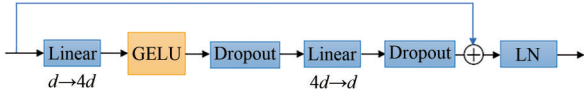


图6 Res-MLP结构

Fig. 6 Architecture of Res-MLP

度特征的捕捉能力。例如,在植被覆盖或水质变化的复杂场景中,Res-MLP通过残差连接保留了浅层边缘信息,同时利用深层非线性变换增强了语义一致性。

2.3 轻量级多尺度解码器 LiteMSD

与一般的图像边缘检测任务不同,遥感影像中的河流边缘具有显著的多尺度特征。河流宽度从几米到数百米不等,导致边缘特征在不同分辨率下的表现差异较大。窄小河流依赖高分辨率的细节特征以捕捉细微边缘,而宽阔河流则需深层语义信息以确保边界连贯性。然而,单一尺度的特征提取难以同时兼顾细粒度边缘和全局语义。因此,融合多尺度特征是提高河流边缘检测效果的重要手段。

EDTER模型的解码器部分是在多级特征聚合解码器(Multi-Level Aggregation, MLA)^[20]基础上增加双向信息流动,利用BiMLA对多层特征进行融合。但是,这样大幅增加了计算成本,不

适用于对高分辨率遥感图像进行的河流边缘检测任务。为降低BiMLA的高计算成本,类似于在SegFormer解码器中只采用MLP层的All-MLP^[21],本文提出了一种轻量化多尺度解码器(Lightweight Multi-Scale Decoder, LiteMSD),充分利用Swin-T编码器4个阶段输出的不同尺度特征(S_1, S_2, S_3, S_4),通过统一通道映射、反卷积上采样、拼接融合和卷积降维等步骤,有效整合不同层次的特征信息。相较于现有的多尺度特征融合模块(如BiMLA、FPN、BiFPN等),LiteMSD在结构设计上不仅更加轻量,有效降低了计算开销,更通过对Swin-T编码器多尺度特征的充分整合,实现了局部细节与全局结构的协同表达。这一特性特别适用于遥感影像中河流边缘这类形态复杂、结构连续、尺度跨度大的检测任务,能够显著缓解细长边缘检测中常见的断裂与误检问题。

如图7所示, LiteMSD模块由统一维度、上采样、特征拼接、卷积降维输出4个部分组成。在编码器提取到4个不同尺度和通道数的特征后,通过 1×1 卷积映射将其通道数统一,使得不同尺度的特征能够在通道维度上进行有效的拼接。

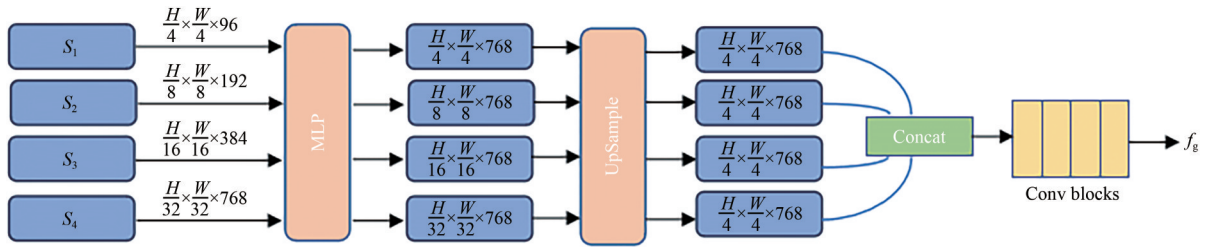


图7 解码器结构

Fig. 7 Decoder architecture

在上采样阶段,对除最高分辨率外的特征采用不同大小反卷积核进行上采样,同时每个反卷积块后引入归一化与激活函数,进一步优化特征表示。经过上采样后,所有特征被调整为相同的空间尺寸。随后,将这些特征在通道维度上拼接,形成融合特征图。在第一阶段中, LiteMSD输出包括全局边缘检测图像与全局特征图。其中,全局特征图由拼接后的融合特征经多层卷积提炼后得到,再通过 1×1 卷积压缩通道数并经过非线性激活函数生成最终的全局

边缘检测图。在第二阶段,由于输入图像被划分为多个序列和更小的图像块分别提取特征, LiteMSD在输出阶段将每个块的边缘特征图填充回全图对应位置,形成完整特征图。其计算过程如下:

$$F'_i = \text{Conv}_{1 \times 1}(S_i), i = 1, 2, 3, 4, \quad (8)$$

$$\hat{F}_i = \text{Upsample}(F'_i), i = 2, 3, 4, \quad (9)$$

$$F_{\text{fusion}} = \text{Concat}(\hat{F}_1, \hat{F}_2, \hat{F}_3, \hat{F}_4), \quad (10)$$

$$F_{\text{output}} = \text{Conv}_{3 \times 3}(F_{\text{fusion}}), \quad (11)$$

其中: $S_1 \sim S_4$ 表示编码器4个阶段输出特征图; F'_i 和 $\hat{F}_i$ 分别表示通道统一后的特征图和上采样后的特征图,且 $\hat{F}_1$ 和 F'_1 相同。

2.4 损失函数

本文借鉴了DiffusionEdge^[22]的损失函数来训练模型,分别用于两个阶段的边缘图。假设边缘图为 E ,对应的真实标签图为 Y ,损失函数计算公式为:

$$\ell(E, Y) = - \sum_{i,j} (Y_{i,j} \alpha \log(E_{i,j}) + (1 - Y_{i,j})(1 - \alpha) \log(1 - E_{i,j})), \quad (12)$$

其中, $E_{i,j}$ 和 $Y_{i,j}$ 分别表示矩阵 E 和 Y 中第 (i,j) 个元素。 $\alpha = |Y^-| / (|Y^-| + |Y^+|)$ 参数表示负样本像素所占的比例,其中, $|Y^-|$ 、 $|Y^+|$ 分别表示负样本和正样本像素的数量。由于BSDS500数据集由多位标注者提供注释,我们首先将多标注信息归一化为取值范围为 $[0,1]$ 的边缘概率图。如果某像素的概率值大于阈值 η ,则被标记为正样本;否则标记为负样本。

将第一阶段的生成的全局边缘图表示为 E_g ,同时中间还生成了4个不同尺度的特征 S_1 、 S_2 、 S_3 、 S_4 ,第一阶段是通过最小化每个边缘图与真实标签之间损失来实现的。第二阶段与第一阶段类似,本文将第二个阶段的最终边缘图表示为 E_r ,两个阶段的损失函数分别如下:

$$\mathcal{L} = \mathcal{L}^E + \lambda \mathcal{L}^S = \ell(E_g, Y) + \lambda \sum_{k=1}^4 \ell(S^k, Y), \quad (13)$$

$$\mathcal{L} = \mathcal{L}^E + \lambda \mathcal{L}^S = \ell(E_r, Y) + \lambda \sum_{k=1}^4 \ell(S^k, Y). \quad (14)$$

其中: $\mathcal{L}^E$ 和 $\mathcal{L}^S$ 分别表示预测边缘图和辅助边缘图的损失项, E_g 和 E_r 分别代表两个阶段的边缘图。参数 λ 表示 $\mathcal{L}^E$ 和 $\mathcal{L}^S$ 之间的权重系数,同时根据HED^[15]实验设置,本文将 λ 设置为0.4。

3 实验与分析

3.1 数据集与环境设置

3.1.1 数据集

本研究选用经典的Berkeley分割数据集(BSDS500)作为基准数据集。该数据集由加州

大学伯克利分校计算机视觉团队构建,广泛用于图像分割和边缘检测任务评估。该数据集包含500张自然图像,其中200张训练图,100张验证图,200张测试图,每张图片均经过5位标注者的手工标注(以.mat格式存储标注结果)。为增强模型训练的鲁棒性,本研究基于原始BSDS500构建扩充数据集HED-BSDS500,通过对训练集和验证集的图像进行不同角度旋转、翻转和缩放,得到共28800张训练图。

针对河流边缘检测任务,本研究构建了跨区域河流边缘数据集(Cross-Regional River Edge Dataset, CRRE)。该数据集的遥感图像通过谷歌地球(Google Earth)以及谷歌地球引擎(Google Earth Engine, GEE)平台获取,主要选用哨兵-2(Sentinel-2)多光谱成像卫星作为数据来源。CRRE数据集涵盖宁夏、山东、河南、内蒙古等多个地区的河流与水库遥感影像,充分考虑了地形多样性,所选图像涉及城市、平原以及黄河流域等多种地貌类型,能够较好地体现遥感场景中复杂水体分布特征。

为保证数据质量,图像筛选过程中首先剔除云量大于5%的遥感图像,并人工删除对比度低、存在强雾霾遮挡或语义模糊的无效影像。同时,所有图像均经过Sen2Cor工具进行大气校正和反射率校正,确保图像光谱特征一致性。标注阶段采用labelme工具对遥感图像中的水体边缘进行像素级标注,标注形式为闭合多边形边界,生成的JSON文件经过转换得到对应的PNG格式标签图像。为提升标注精度,采用双人复核机制,确保边缘轮廓准确清晰。

最终得到385张分辨率为 512×512 的RGB三通道图像及其对应的标签图像。考虑到数据集规模有限,为避免数据划分过细影响训练与评估效果,本文采用训练-测试划分方案,将其中308张图像作为训练集,77张图像作为测试集。训练过程中,测试集同时作为验证集使用,以评估模型在不同训练轮次下的性能变化。CRRE数据集在样本多样性、场景复杂性和标注精度等方面均具有良好的代表性,可为后续河流边缘检测模型的训练与评估提供可靠支持。图像样本如图8所示。

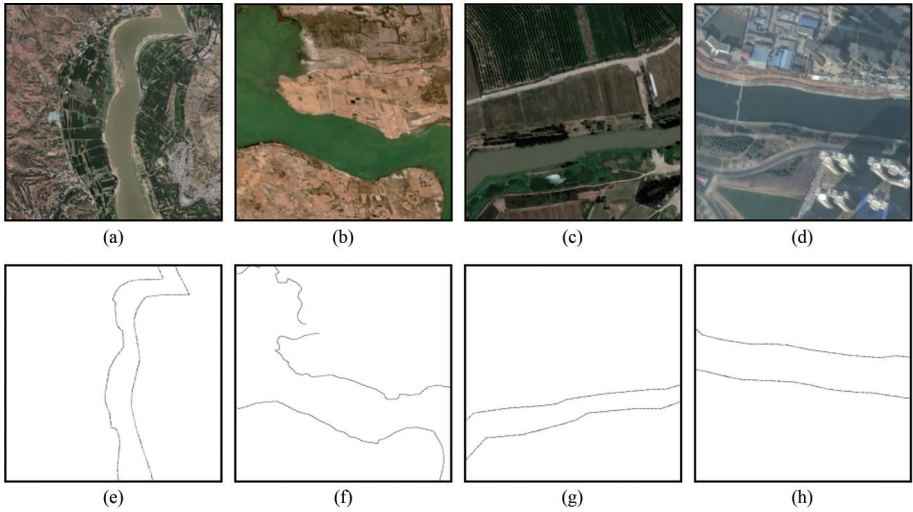


图8 训练样本示例。(a)~(d)原图;(e)~(h)对应的标注图。
Fig. 8 Examples of training samples. (a)~(d) Original images; (e)~(h) Corresponding label maps.

3.1.2 环境设置

本研究使用基于Pytorch的深度学习框架，操作系统为unbntu 20.04，GPU为NVIDIA GeForce RTX 4060Ti-16G，CPU为Intel Core i5-12400F Processor。采用SGD (Stochastic Gradient Descent,SGD)优化器对模型训练,初始

学习率设置为 $1e-6$,动量为0.9,权重衰减为0.000 5,批尺寸大小(Batch Size)为2,共进行80 000次迭代。在训练前使用Swin-T预训练权重初始化模型的编码器,同时对进入编码器的每张图像随机裁剪成分辨率为 224×224 的图像。

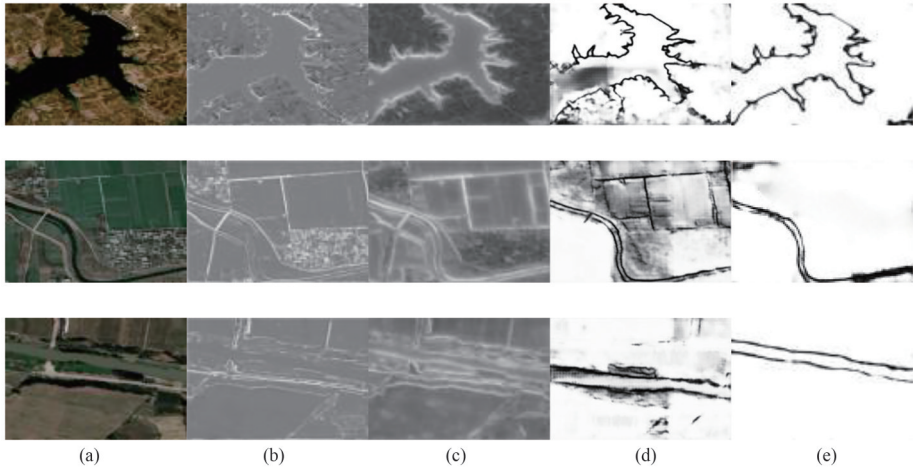


图9 CRRE数据集上模型对比效果图。(a) 原始图像;(b) BDCN模型;(c) RCF模型;(d) EDTER模型;(e) 本文模型。
Fig. 9 Visual comparison of model performance on CRRE dataset. (a) Original images; (b) BDCN; (c) RCF; (d) EDTER; (e) Ours.

3.1.3 评价指标

实验采用通用的边缘检测性能评价指标对检测结果进行评估,分别是全局最佳阈值(ODS)、单图最佳阈值(OIS)和平均精度(AP),公式如式(10)~(12)所示。其中,ODS是指在整个数据集上使用同一个最佳阈值来计算F1-score,然后取

所有图像的F1-score的平均值作为最终的评价指标。OIS则是针对每张图像找到其最佳阈值以最大化F1-score,然后对所有图像的F1-score求平均。

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{15}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (16)$$

$$\text{F1-score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}, \quad (17)$$

其中:TP表示被模型预测为正类的正样本,FP表示被模型预测为正类的负样本,FN表示被模型预测为负类的正样本^[23]。此外,在评估前,需要对模型预测的边缘图像进行非极大值抑制(Non-Maximum Suppression,NMS)操作。

3.2 实验结果分析

3.2.1 消融实验

为验证本研究提出改进方法的可行性,在BSDS500数据集上分别对轻量化主干网络Swin-T、残差多层感知模块Res-MLP、轻量化多尺度解码器LiteMSD进行了消融实验,实验结果如表2所示,打勾表示采用该改进点。

由表2可知,基准模型的ODS为0.824,OIS为0.841,AP为0.880。引入Swin-T后,模型的ODS提升了0.3%,AP提升了0.1%,表明Swin-T在轻量化特征提取和计算效率方面发挥了重要作用,有助于提升整体的边缘检测效果。引入Res-MLP后,模型的ODS略微下降了0.1%,AP下降了0.3%,而OIS提升了0.4%。这一结果表明,尽管Res-MLP在增强边缘细节和优化边缘提取方面起到了显著作用,但它对全局特征的影响相较于Swin-T稍逊一筹,因此导致ODS和AP略有下降。接下来,LiteMSD模块的引入使得模型的多尺度解码能力和细节增强效果得到了显著提升。去除LiteMSD后,模型的ODS和AP分别下降了0.2%和0.1%,显示出该模块在细节保留和提升边缘检测精度方面的关键作用。当Swin-T、Res-MLP和LiteMSD 3个模块同时引入时,模型的性能达到了最佳,ODS为0.835,OIS为0.848,AP为0.884。综合结果表明,3个

模块的协同作用显著提升了模型在边缘检测任务中的能力,最终的性能优于基准模型,验证了每个模块在优化边缘检测精度、增强细节提取和提高计算效率方面的有效性。

Swin-T、Res-MLP和LiteMSD模块在特征提取、边缘细节表达和多尺度特征融合方面分别发挥了重要作用。Swin-T主干采用层次化设计,相较于ViT类模型,能够更好地保留浅层空间细节信息,提升边缘区域特征的表达能力。Res-MLP模块通过双重残差连接增强了非线性特征表达能力,尤其在复杂背景区域对边缘细节信息具有较强的增强效果。LiteMSD解码器通过统一通道映射、空间尺寸对齐与多尺度特征融合,实现了不同尺度特征的有效集成,进一步强化了边缘检测能力。

在消融实验中引入Res-MLP后,虽然ODS和OIS均有所提升,但AP指标略有下降。这主要是由于Res-MLP强化了局部细节特征表达,导致部分弱边缘区域响应增强,从而引入了一定比例的伪边缘。伪边缘的增加在一定程度上影响了整体的精确率(AP),但同时也提升了模型对复杂结构边缘的捕获能力。后续通过LiteMSD的全局特征融合进一步平衡了细节与全局信息,有效抑制了伪边缘的干扰,使最终模型的3项指标均达到最优水平。

综合来看,3部分改进模块在边缘检测任务中形成互补,通过细节增强、全局特征整合与多尺度上下文信息建模,实现了在公开数据集和自构建数据集上的性能全面提升。

3.2.2 对比实验

本研究分别在公开数据集BSDS500和CRRE数据集上进行对比实验。在BSDS500数据集上与BDCN^[24]、HED^[25]、RCF^[26]、EDTER^[15]、DeepContour^[27]、DeepEdge^[28]、PiDiNet^[29]进行了对比实验,结果如表3所示。基于ODS、OIS和AP这3个指标,所提模型分别达到了0.835、0.848和0.884,相较于其他方法均有提升。考虑到模型的参数量和检测精度,本文模型优于其他方法。这充分证明所提方法在边缘检测任务中具有较高的准确性和鲁棒性。

为进一步验证所提模型在实际应用中的有效性,本研究在CRRE的河流边缘检测数据集上开展了对比实验。对比方法包括当前主流的边

表2 BSDS500数据集上的消融实验

Tab. 2 Ablation experiments on BSDS500 dataset

No.	Swin-T	Res-MLP	Lite MSD	ODS	OIS	AP
1				0.824	0.841	0.880
2	✓			0.827	0.845	0.881
3		✓		0.826	0.843	0.878
4			✓	0.825	0.842	0.882
5	✓	✓	✓	0.835	0.848	0.884

表 3 BSDS500数据集上的模型性能对比结果

Tab. 3 Comparison of model performance on BSDS500 dataset

模型	ODS	OIS	AP
BDCN ^[24]	0.806	0.826	0.890
HED ^[25]	0.788	0.808	0.840
EDTER ^[15]	0.824	0.841	0.880
DeepContour ^[27]	0.757	0.776	0.800
DeepEdge ^[28]	0.753	0.772	0.807
PiDiNet ^[29]	0.789	0.803	—
RCF ^[26]	0.811	0.830	—
Ours	0.835	0.848	0.884

缘检测模型 BDCN^[24]、HED^[25]、EDTER^[15]和 RCF^[26]。从表 4 可以看出,所提模型在 3 个评价指标上均取得了最优结果,ODS 达到 0.841,OIS 达到 0.849,AP 提升至 0.871,相较于其他模型均有显著提升。其中,AP 指标相比表现最优的对比模型 BDCN 提升了约 2.6%,说明本文方法在边缘区域的精确识别和稳定性方面具有较强优势。相比 EDTER 模型,所提模型的 ODS 提升了 1.6%,OIS 提升了 1.8%,进一步表明本文引入的轻量化主干、残差多层感知结构与多尺度解码策略在增强边缘信息提取方面发挥了良好作用。此外,本文还对比了不同主干网络在 ImageNet-1K 上的参数量与计算复杂度,如表 1 所示。结果表明,Swin-T 相较于 ViT-L 在参数量上压缩了约 90.9%,FLOPs 降低了约 97.6%,显著降低了模型训练与推理过程中的计算资源消

表 4 CRRE数据集上的模型性能对比结果

Tab. 4 Comparison of model performance on CRRE dataset

模型	ODS	OIS	AP
BDCN ^[24]	0.821	0.835	0.845
HED ^[25]	0.812	0.824	0.834
EDTER ^[15]	0.825	0.831	0.839
RCF ^[26]	0.817	0.833	—
Ours	0.841	0.849	0.871

耗。综上,本文模型不仅在公开数据集上表现优异,在具有挑战性的 CRRE 数据集上同样展现出良好的泛化能力与实际应用价值。

4 结 论

本文针对河流边缘检测任务提出了一种基于改进 EDTER 的边缘检测算法,解决了通用边缘检测模型参数量大、河流边缘检测效果不佳的问题。本文利用两个阶段分别提取全局特征和局部特征,同时引入了 Swin-T 编码器、残差多层感知模块 Res-MLP 和轻量化多尺度解码器改进策略,实现河流边缘的高效精准提取。在 BSDS500 数据集和 CRRE 的河流边缘数据集上的实验结果表明,所提方法在 ODS、OIS 和 AP 等性能指标上均优于现有主流方法,特别是在 AP 指标上有 2.6% 的提升。然而所提方法对于一些细小的河流边缘检测仍存在一定的漏检、误检,未来的研究还需要对复杂图像的检测做进一步的探索。

参 考 文 献:

[1] 徐依婷,穆月英,侯玲玲. 水资源稀缺性、灌溉技术采用与节水效应[J]. 农业技术经济,2022(2):47-61.
XU Y T, MU Y Y, HOU L L. Water scarcity, adoption of irrigation technologies and water conservation [J]. *Journal of Agrotechnical Economics*, 2022(2): 47-61. (in Chinese)

[2] 张继贤,顾海燕,杨懿,等. 高分辨率遥感影像智能解译研究进展与趋势[J]. 遥感学报,2021,25(11):2198-2210.
ZHANG J X, GU H Y, YANG Y, *et al.* Research progress and trend of high-resolution remote sensing imagery intelligent interpretation [J]. *National Remote Sensing Bulletin*, 2021, 25(11): 2198-2210. (in Chinese)

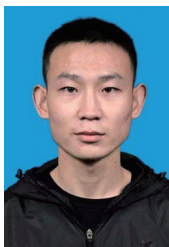
[3] 余加俊,骆文韬,徐飞鸿,等. 智能清漂船河道边界图像识别算法[J]. 电子设计工程,2018,26(13):132-136.
YU J J, LUO W T, XU F H, *et al.* River boundary recognition algorithm for intelligent float-garbage ship [J]. *Electronic Design Engineering*, 2018, 26(13): 132-136. (in Chinese)

[4] 黄凌霄,李春光,郑兰香,等. 鸭子荡水库地形前处理及二维水流运动模拟[J]. 水利水电科技进展,2017,37(4):47-52.
HUANG L X, LI C G, ZHENG L X, *et al.* Terrain pretreatment and two-dimensional flow simulation in Yazidang Reservoir [J]. *Advances in Science and Technology of Water Resources*, 2017, 37(4): 47-52. (in Chinese)

- [5] 郑又能,肖宇峰,寸超,等. 基于改进区域生长的水岸线提取方法研究[J]. 计算机应用研究, 2020, 37(6): 1876-1881.
ZHENG Y N, XIAO Y F, CUN C, *et al.* Research on waterfront line extraction based on improved region growing method [J]. *Application Research of Computers*, 2020, 37(6): 1876-1881. (in Chinese)
- [6] VANDAELE R, DANCE S L, OJHA V. Deep learning for automated river-level monitoring through river-camera images: an approach based on water segmentation and transfer learning [J]. *Hydrology and Earth System Sciences*, 2021, 25(8): 4435-4453.
- [7] YIN Y H, GUO Y G, DENG L W, *et al.* Improved PSPNet-based water shoreline detection in complex inland river scenarios [J]. *Complex & Intelligent Systems*, 2023, 9(1): 233-245.
- [8] AREFIN R, MESHRAM S G, SANTOS C A G, *et al.* Hybrid modelling approach for water body change detection at Chalan Beel area in northern Bangladesh [J]. *Environmental Earth Sciences*, 2020, 79(19): 442.
- [9] JAIN S K, KUMAR D, THAKUR M, *et al.* Proximal support vector machine-based hybrid approach for edge detection in noisy images [J]. *Journal of Intelligent Systems*, 2019, 29(1): 1315-1328.
- [10] ZOU N, XIANG Z Y, CHEN Y M, *et al.* Boundary-aware CNN for semantic segmentation [J]. *IEEE Access*, 2019, 7: 114520-114528.
- [11] WALDNER F, DIAKOIANNIS F I. Deep learning on edge: extracting field boundaries from satellite images with a convolutional neural network [J]. *Remote Sensing of Environment*, 2020, 245: 111741.
- [12] LIU X Y, PENG H W, ZHENG N X, *et al.* EfficientViT: memory efficient vision Transformer with cascaded group attention [C]//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023: 14420-14430.
- [13] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, *et al.* An image is worth 16×16 words: Transformers for image recognition at scale [C]//*Proceedings of the 9th International Conference on Learning Representations*. Online: ICLR, 2021: 1-21.
- [14] PU M Y, HUANG Y P, LIU Y M, *et al.* EDTER: edge detection with Transformer [C]//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Orleans: IEEE, 2022: 1392-1402.
- [15] LIU Z, LIN Y T, CAO Y, *et al.* Swin Transformer: hierarchical vision Transformer using shifted windows [C]//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021: 9992-10002.
- [16] 徐旺旺,许良凤,李博凯,等. TransAS-UNet:融合Swin Transformer和UNet的乳腺癌区域分割[J]. 中国图象图形学报, 2024, 29(3): 741-754.
XU W W, XU L F, LI B K, *et al.* TransAS-UNet: regional segmentation of breast cancer Swin Transformer and of UNet algorithm [J]. *Journal of Image and Graphics*, 2024, 29(3): 741-754. (in Chinese)
- [17] BI K F, XIE L X, ZHANG H H, *et al.* Accurate medium-range global weather forecasting with 3D neural networks [J]. *Nature*, 2023, 619(7970): 533-538.
- [18] XU X K, FENG Z J, CAO C Q, *et al.* An improved Swin Transformer-based model for remote sensing object detection and instance segmentation [J]. *Remote Sensing*, 2021, 13(23): 4779.
- [19] PACAL I. A novel Swin Transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images [J]. *International Journal of Machine Learning and Cybernetics*, 2024, 15(9): 3579-3597.
- [20] ZHENG S X, LU J C, ZHAO H S, *et al.* Rethinking semantic segmentation from a sequence-to-sequence perspective with Transformers [C]//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021: 6877-6886.
- [21] XIE E Z, WANG W H, YU Z D, *et al.* SegFormer: simple and efficient design for semantic segmentation with Transformers [C]//*Proceedings of the 35th International Conference on Neural Information Processing Systems*. Online: Curran Associates Inc., 2021: 924.
- [22] YE Y F, XU K, HUANG Y H, *et al.* DiffusionEdge: diffusion probabilistic model for crisp edge detection [C]//*Proceedings of the 38th AAAI Conference on Artificial Intelligence*. Vancouver: AAAI Press, 2024: 6675-6683.
- [23] 肖扬,周军. 图像边缘检测综述[J]. 计算机工程与应用, 2023, 59(5): 40-54.
XIAO Y, ZHOU J. Overview of image edge detection [J]. *Computer Engineering and Applications*, 2023, 59(5): 40-54. (in Chinese)

- [24] HE J Z, ZHANG S L, YANG M, *et al.* Bi-directional cascade network for perceptual edge detection [C]// *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019: 3823-3832.
- [25] XIE S N, TU Z W. Holistically-nested edge detection [C]// *Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago: IEEE, 2015: 1395-1403.
- [26] LIU Y, CHENG M M, HU X W, *et al.* Richer convolutional features for edge detection [C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 5872-5881.
- [27] SHEN W, WANG X G, WANG Y, *et al.* DeepContour: a deep convolutional feature learned by positive-sharing loss for contour detection [C]// *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015: 3982-3991.
- [28] BERTASIUS G, SHI J B, TORRESANI L. DeepEdge: a multi-scale bifurcated deep network for top-down contour detection [C]// *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015: 4380-4389.
- [29] SU Z, ZHANG J H, WANG L G, *et al.* Lightweight pixel difference networks for efficient visual representation learning [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(12): 14956-14974.

作者简介:



高钰栋,男,硕士研究生,2021 年于内蒙古农业大学获得学士学位,主要从事图像处理方面的研究。E-mail: 15929148943@139.com



黄凌霄,男,博士,副教授,2017 年于宁夏大学获得博士学位,主要从事计算机视觉和计算水力学的研究。E-mail: huanglx@nxu.edu.cn