

边界引导的气象探测环境障碍物检测^①



李皓辰, 姜淑杨, 缪明榕, 吴嘉伟, 鲍磊磊, 吴锐涛

(南通市气象局 南通市气象探测中心, 南通 226018)

通信作者: 吴锐涛, E-mail: ntqixiang@163.com

摘要: 气象探测环境保护工作是气象观测台站业务的重要一环, 密切关系观测数据精度和气象服务质量. 方法聚焦时间序列图像中树木、作物等自然生长的障碍物边界改变细微的特点, 提出了植入式的边界感知模块, 以监督引导的方式专注于障碍物边界特征信息的提取. 边界特征与主干网络特征通过边界引导模块进行融合, 强化模型对边界的关注程度. 实验选取了轻量化模型 HRNet-W18 作为方法基线, 在真实气象探测环境数据集上进行验证, 边界引导方法在树木、作物障碍物类别上有效提升了 2.29% 的 *mIoU*, 总体分割效果达到 95.82%, 超过其他 14 个主流语义分割方法, 更好地平衡了模型复杂度和效果精度.

关键词: 气象探测; 边界感知; 轻量化语义分割

引用格式: 李皓辰, 姜淑杨, 缪明榕, 吴嘉伟, 鲍磊磊, 吴锐涛. 边界引导的气象探测环境障碍物检测. 计算机系统应用, 2025, 34(7): 272-279. <http://www.c-s-a.org.cn/1003-3254/9937.html>

Boundary-guided Obstacle Detection in Meteorological Detection Environment

LI Hao-Chen, JIANG Shu-Yang, MIAO Ming-Rong, WU Jia-Wei, BAO Lei-Lei, WU Rui-Tao

(Nantong Meteorological Observation Center, Nantong Meteorological Bureau, Nantong 226018, China)

Abstract: The protection of the meteorological detection environment is a pivotal responsibility for meteorological stations, closely related to the accuracy of meteorological data and the quality of meteorological forecasts. Targeting the subtle boundary changes of naturally growing obstacles like trees and crops in time-series images, this study proposes an implanted boundary perception module that focuses on extracting obstacle boundary feature information in a supervised and guided manner. The boundary features are then integrated with backbone network features through the boundary-guided module to enhance the model's attention to the boundaries. The experiment selects the lightweight model HRNet-W18 as the baseline for the method and validates it on a real meteorological detection environment dataset. The boundary guidance method effectively improves the *mIoU* by 2.29% for tree and crop obstacle categories, achieving an overall segmentation accuracy of 95.82%. This performance surpasses 14 other mainstream semantic segmentation methods and achieves a better balance between model complexity and accuracy.

Key words: meteorological detection; boundary-aware; lightweight semantic segmentation

地面气象观测运用精密的观测仪器, 系统监测近地面大气层的多种气象要素, 捕捉自由大气中发生的各类天气现象, 为天气预报^[1]、气象服务^[2]、气候分析^[3]及灾害预警^[4]等提供数据支撑. 为确保观测数据的代表性、准确性、连续性和可比较性, 必须对地面气象观

测设备的安置场所周边环境实施严格的保护措施, 以消除或减轻干扰源和障碍物对观测数据可能造成的负面影响.

针对上述问题, 国务院气象主管部门已颁布《气象设施和气象探测环境保护条例》及国家标准《气象

① 基金项目: 南通市气象局科研资助项目 (NQK202405)

收稿时间: 2024-11-16; 修改时间: 2025-02-12; 采用时间: 2025-03-06; csa 在线出版时间: 2025-05-16

CNKI 网络首发时间: 2025-05-19

探测环境保护规范 地面气象站》. 不仅明令禁止在各级气象观测站(国家基准气候站、国家基准/基本气象站、国家一般气象站)的探测环境保护范围内实施任何可能损害探测环境的行为,也对观测场周边建筑物、树木等障碍物的最近距离和最大高度进行明确规定,以确保气象探测环境的纯净与稳定.

尽管如此,在实际业务中气象探测环境保护工作始终面临诸多挑战. 首先,存在气象站点分布广、探测环境保护区域范围大的客观困难. 以国家基准/基本气象站为例,作为我国气象站网的主体台站,当前我国已建设完成该类气象站点 2000 多个,基本覆盖全国县级行政区域. 相关保护条例要求国家基准/基本气象站以观测场围栏向外延伸 1000 m 的距离作为核心控制区,这一范围内要求障碍物以观测场围栏上沿为参照的高度距离比小于 1/10. 而在城市总体布局中,鉴于社会因素、经济成本及业务需求等多方面权衡,气象站点大多被设置在较为偏远、社会活动较少的区域^[5]. 这种选址方式虽有利于观测业务顺利开展,但一定程度上增加了探测环境保护工作的难度,巡查保护区域、比对区域内各类障碍物的工作变得更为复杂且耗费人力. 其次,探测环境保护工作需要较高的时效性,一部分原因在于环境质量的下降可能对探测结果有即时性的影响. 另一部分原因在于针对保护区域内人为新增或加高建筑类障碍物的问题,事前干预的成本和效果要优于事后监管. 最后,在气象观测业务自动化、集约化程度日益提高,气象站点无人化、智能化改革稳步推进的背景下,现行的气象探测环境保护工作方法已经落后. 一方面树木、作物等障碍物的生长变化难以及时发现,保护工作的效果基本依靠台站值班人员的专业能力和工作经验. 另一方面探测环境会长时间处于较为稳定的状态,台站值守人员容易松懈或是忽视,难以迅速响应外部环境变化,增加了威胁发现滞后的风险.

鉴于上述背景,提升气象探测环境保护工作效率和业务自动化水平^[6]显得尤为迫切. 探索利用计算机视觉方法对气象探测环境障碍物进行自动快速检测,不仅是为了应对当前面临的挑战,更是为了提升探测环境保护工作效能、长远保障气象观测业务的稳定与可持续发展. 然而在真实应用场景下,障碍物检测方法需要重点考虑以下因素. 一是真实应用场景存在算力有限的现实制约,气象台站业务值班电脑型号老旧、性能有限,方法需要谨慎平衡检测精度和算法复杂度. 二是真实应用场景对于分辨率有较高的要求,探测环境

保护区内情况复杂,树木、建筑等各类障碍物相互遮挡,树木、作物等障碍物末端尖细,这些图像细节都需要较高的分辨率才能提供. 三是相比建筑物、构筑物因人为活动带来的变化,树木、作物等植物的自然生长变化更加常见,而这种变化往往需要一段时间的积累才会体现在障碍物边界附近,因此需要模型更加关注障碍物的轮廓分割.

因此本文利用观测场内 10 m 风塔上架设的高清球型摄像机采集周边探测环境图像数据,制作高分辨率真实数据集. 结合真实场景客观存在算力有限的问题和高分辨率检测的需求,本文选取轻量化高分辨率模型 HRNet-W18 作为方法基线,在整个主干网络中保持高分辨率特征表示,有效捕捉更精细的结构和边界. 本文提出了 PnP (plug-and-play) 的轻量化边界感知模块,可以通用地嵌入采用多阶段编码结构的主流模型. 轻量化边界感知模块 BAM 融合编码器第一阶段的高分辨率图像特征和最后一阶段的深层次语义特征,形成对目标局部边缘特征的关注,并在边界特征引导模块 BGM 中将目标边界特征以注意力引导的方式与主干特征进行深度融合. 实验在真实数据集上与 14 种代表性方法对比效果, BAM 和 BGM 模块有效提升了 1.03% 的 *mIoU* 分割精度,总体分割效果达到最优, *mIoU* 为 95.82%. 最后以 HRNet-W18 和 SegFormer-B0 分别作为基线方法进行消融实验,在不同方法轻量化的边界感知和特征引导模块上均以较小代价有效提升了分割效果,证明了模块的有效性和通用性.

1 相关工作

真实场景下的语义分割算法一直是研究者关注重点, FCN^[7]首次运用卷积神经网络完成语义分割任务,使用反卷积层替代全连接层进行上采样,使得模型能够接收任意大小的输入. SegNet^[8]为了更好地保留边界特征信息,创新性地提出了带索引的最大池化方式进行上采样. PSPNet^[9]提出金字塔池化模块来捕捉不同尺度上的上下文信息,并在训练中引入辅助损失函数. RefineNet^[10]进一步谨慎设计了编码器-解码器结构,将编码器下的高分辨率特征与前序解码器模块的低分辨率特征级联. DeepLabV3^[11]设计了级联或并行的空洞卷积结构进一步地扩充了空洞空间金字塔池化模块. HRNet^[12]创新性地采用并行连接从高分辨率到低分辨率子网络输出的模型架构,在模型始终保持高分辨率表征,更好地保留了图像细节.

随着计算资源的进一步解放,越来越多的语义分割方法涌现,语义分割模型不再拘泥于卷积神经网络,基于 Transformer 模型的语义分割方法成为近年的研究热点. SETR^[13]首先提出以纯 Transformer 结构的编码器来代替 CNN 编码器. SETR 对输入图像先做分块处理,然后为每个图像分块嵌入位置编码,将图像转换为向量序列,最后输入 Transformer block 里进行编码. SegFormer^[14]框架重新设计了编码器-解码器,提出了一种多尺度层次结构且无位置编码的 Transformer 编码器,避免了测试分辨率差异带来的位置编码问题.同时采用轻量的全连接多层感知机解码器设计,有效提升模型效率.

伴随着物联网和移动设备的发展与普及,越来越多的图像资源集中在移动终端、车载摄像头、监控摄像头等小型设备,能够部署在计算资源受限环境的轻量级语义分割与实际应用紧密结合,成为另一研究热点.轻量级语义分割更加需要考虑分割方法在实际应用中存在的困境,谨慎平衡性能、耗时和计算成本多方面因素. ENet^[15]设计注重减少模型参数量,与早期的对称编码器-解码器构造不同,ENet 采用了大型编码器和小型解码器的组合,同时设计了初始化输入模块大大减小了输入图像的大小. CGNet^[16]设计了全新的上下文引导块,将局部特征与周围上下文信息巧妙结合,并通过全局上下文进一步增强联合特征,优化了模型参数量和运行内存.

2 提出的方法

2.1 边界感知和特征引导

- 边界感知模块 (boundary-aware module, BAM) 采用轻量化的设计结构,可以方便得嵌入多阶段编码结构的主干网络.该模块同时接收主干网络提取出的底层边界细节特征 F_L 和高层全局语义特征 F_H .底层特征 F_L 通常从第 1 个编码阶段的末端输出,高层特征 F_H 则通常由最后一个编码阶段的末端输出.边界细节特征用来弥补主干网络下采样过程中容易损失的图像细节,全局语义特征则用来指导目标主体位置,抑制无关噪声. BAM 以监督学习的方式,利用 F_L 和 F_H 主动感知障碍物的轮廓细节,有效增强物体边界的分割精度.模块结构如图 1,左侧立方体表示分辨率较高、维度较浅的底层特征 F_L ,右侧立方体表示分辨率较低、维度较深的高层特征 F_H .

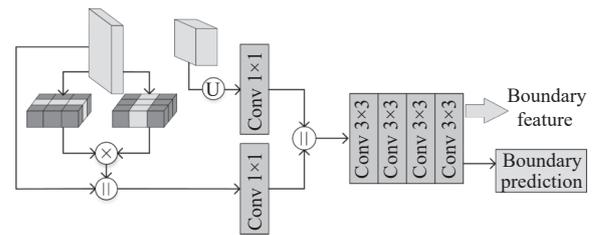


图 1 边界感知模块结构图

底层特征 F_L 先输入两个参数固定的 3×3 卷积 C_X 和 C_Y , 在水平方向和垂直方向上应用 Sobel 算子分别得到含有丰富边缘信息的水平梯度图 G_X 和垂直梯度图 G_Y . C_X 和 C_Y 卷积核参数如式 (1) 所示:

$$C_X = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, C_Y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (1)$$

水平梯度图 G_X 和垂直梯度图 G_Y 分别进行 Sigmoid 函数归一化后与底层特征 F_L 进行逐元素相乘得到水平方向和垂直方向上的边界特征加权图,函数归一化是为了消除边界特征在水平方向和垂直方向上可能存在的量纲差异,突出边缘信息增益.最后再将两个方向上的边界特征加权图与底层特征 F_L 进行通道拼接后得到增强边界信息的特征图 F_E .计算方法如式 (2) 所示,其中“ \otimes ”表示逐元素相乘,“+”表示维度相加,“sig”表示 Sigmoid 归一化函数.

$$F_E = F_L + F_L \otimes \text{sig}(G_X) + F_L \otimes \text{sig}(G_Y) \quad (2)$$

高层特征 F_H 先通过上采样操作恢复至与特征 F_E 相同的分辨率大小,图 1 中用“U”表示上采样操作.之后,将用于指导全局语义信息的高层特征 F_H 与边界信息强化后的底层特征 F_E 分别通过一个 1×1 的卷积层.通过 1×1 卷积操作跨通道整合特征信息,并将 F_E 维度向上与 F_H 维度对齐.最后,将 F_E 和 F_H 拼接并输入由 4 个 3×3 卷积层组成的 Basic block 卷积块,提取出有效的边界细节特征 F_B 输出边界感知模块. BAM 通过额外的辅助分割头参与模型训练,训练时需对真实标注进行处理,提取出适当的障碍物边界标注再对模块参数进行监督学习,具体细节见第 2.3 节.

- 边界特征引导模块 (boundary-guided module, BGM) 用来接收 BAM 输出的边界细节特征 F_B 和主干网络编码器末端特征 F ,并将两者以注意力引导的方式进行融合,在主干特征 F 中进一步增强对障碍物边界附近的关注度,具体结构如图 2 所示.

首先边界细节特征 F_B 通过一个 1×1 卷积层与主

干特征 F 进行维度对齐, 然后对 F_B 和 F 采用先逐元素相乘再逐元素相加的融合方式. 通过这种融合方式既有效增强特征 F 中的边界信息权重, 又有效整合了不同源的特征信息. 之后通过一个 1×1 卷积进行初步特征融合得到特征 F .

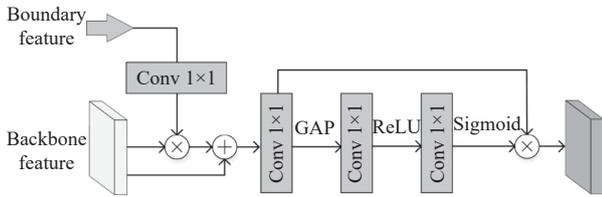


图2 边界特征引导模块结构图

初始融合特征 F 继续通过全局平均池化层 (global average pooling, GAP) 来聚合全局上下文信息, 之后依次通过 ReLU 和 Sigmoid 两个非线性激活函数层及两个 1×1 卷积来进一步估计层相关性, 生成通道注意力权重 F_w .

最后将通道注意力权重 F_w 和初始融合特征 F 逐元素相乘, 输出最终含有通道权重的融合特征. 融合边界强化特征后的主干特征与初始主干特征 F 维持相同的分辨率和维度大小, 并参与后续的解码阶段得到最终结果.

2.2 基于 HRNet-W18 的边界引导模型

传统的编码解码结构中, 编码阶段通过连续的卷积和池化操作来减少特征图的分辨率, 从而提取高层次的语义信息, 但这一过程往往伴随着细节特征的丢失. HRNet^[10]摒弃了先下采样后上采样的处理范式, 转

而采用一种包含多个并行运行且分辨率各异的多分支网络框架, 其核心理念在于维护高分辨率特征图的同时, 巧妙融合多尺度特征. HRNet 提出了 3 个不同复杂度的版本, 分别是 HRNet-W18、HRNet-W32、HRNet-W48, 三者之间的结构一致, 区别仅在于采用了不同宽度的特征维度, 继而带来了模型量级的差异. HRNet 的独特设计使该方法在图像分割任务中表现出色, 更契合探测环境障碍物检测的应用场景. 一方面只有高分辨检测方法才能提供复杂的障碍物细节, 如树梢、天线等. 另一方面真实应用场景存在算力有限的现实制约. 因此本节选择最轻量级的 HRNet-W18 作为主干网络, 采用 HRNet 和 OCR^[17]的组合作为基线方法. OCR 专注于利用像素间的对象上下文关系来增强特征表示, 通过引入对象区域的概念, 计算每个像素与其所属对象区域间的关系, 来优化不同对象的精度.

考虑到探测环境检测任务中, 树木、作物等障碍物具备自然生长的特性, 在不同时序图像间这些植物易在生长边界处发生微小改变, 因而模型需要加倍关注目标边界的细节处理, 基于 HRNet 融入 BAM、BGM 模块提出本文的模型结构, 如图 3 所示. HRNet 主干可以分为 4 个阶段, 第 1 阶段通过步长为 2 的卷积下采样到原始图像 1/4 的尺寸, 提取初始图像特征. 随后, 每一个阶段的开头都在前一阶段的基础上进行 2 倍下采样得到更低分辨率子网, 不同分辨率的子网都会在各异的尺度上通过连续 4 个卷积块进一步提取该分辨率的特征. 每个阶段的末端都会进行多尺度特征融合, 使得不同分辨率的子网之间可以交换互补的特征信息.

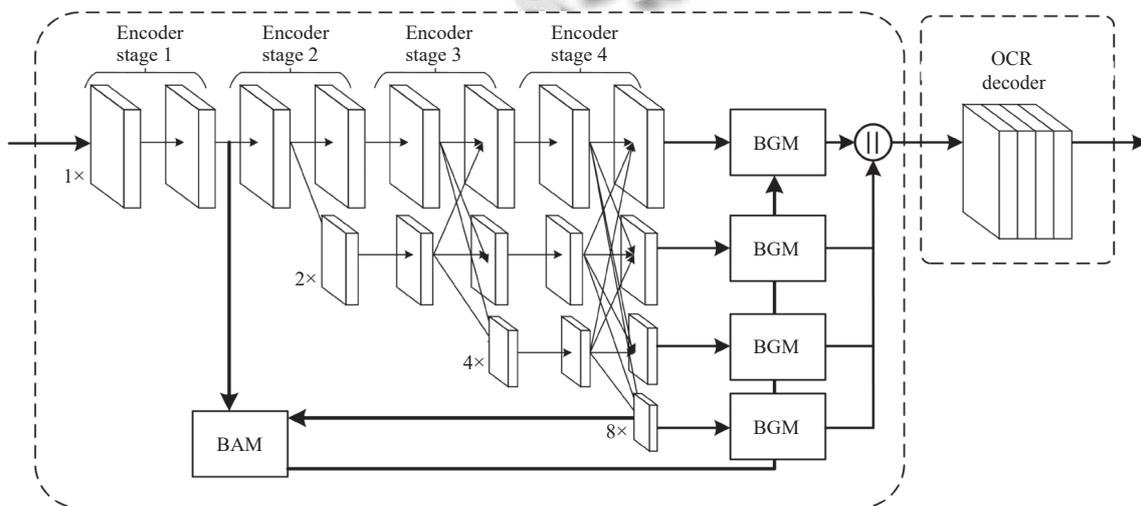


图3 基于 HRNet 的边界引导模型结构图

将 HRNet-W18 第 1 阶段的高分辨率底层图像特征 N_1 和最后一阶段的高层语义特征 N_4 输入 BAM, 经辅助头监督学习的边界感知后输出尺寸为 $(144, H/4, W/4)$ 的边界强化特征, N_1 和 N_4 特征维度大小分别为 $(18, H/4, W/4)$ 和 $(144, H/32, W/32)$. 边界强化特征自底向上同时输入并行的 BGM, 在 4 个不同尺度上引导主干特征强化边界信息, 经 BGM 融合后的主干特征与原特征尺寸一致. 最后将所有的并行的低分辨率子网上采样至高分辨率子网, 即 $(H/4, W/4)$ 的尺寸大小. 然后通过连接所有的子网特征获取到具有丰富信息的高分辨率表征, 特征大小为 $(270, H/4, W/4)$, 并输入 OCR 解码器得到最后的分割结果.

2.3 训练设计

近期提出的通过边界特征辅助分割的方法从设计上可以分为两类: 一类利用辅助头参与特征引导^[18,19], 另一类则在网络内部利用特征图估算出二值边界再计算分割结果^[20]. 本文采用第 1 类思路, 需要对真实标注进行边界提取, 生成掩码边界标注. 首先计算真实标注 I 中每一个像素点 p 到背景点 q (背景点 q 的标注值均为 0) 的最小像素距离 $L_{bg}(p)$. 计算方式如式 (3):

$$L_{bg}(p) = \min_{q \in I, q=0} dist(p, q) \quad (3)$$

根据每个像素点的 $L_{bg}(p)$ 值对轮廓附近一定宽度内的像素标注进行保留, 生成边界标注图 G_{bt} . 采用式 (4) 进行计算.

$$G_{bt} = \{p \mid p \in I \text{ and } L_{bg}(p) \leq \alpha\} \quad (4)$$

保留 G_{bt} 中距离背景最小像素距离不大于预设值 α 的像素点 (本文将 α 值默认设为 5), 其余像素的值均置为 0. 最后将多语义类别标注的边界二值化, 生成用来监督辅助分割头学习的掩码边界标注.

在训练阶段中, 需要同时对主干网络末端和边界感知辅助分割头同时进行监督训练, 主干网络末端为多类分割, 辅助分割头为二类分割. 训练过程中整体损失函数 L 如式 (5) 所示:

$$L = \beta \cdot L_{Focal} + (1 - \beta) \cdot L_{CCE} \quad (5)$$

考虑到掩码边界标注图中掩码边界样本数和背景像素点数量量级上的不平衡, 采用 Focal Loss 损失函数对边界感知辅助分割头进行监督, 以有效平衡边界感知模块对目标边界分布的学习. 在主干网络的末端则采用常规的 CCE Loss 损失函数对网络整体进行监

督, 以便与其他语义分割方法进行实验比较.

3 实验分析

3.1 数据集

本文实验采用真实的气象探测环境图像数据集, 采集站点为南通大气探测中心国家基本气象站. 通过地面观测场内架设于 10 m 测风塔上的高清摄像球机对四周探测环境进行拍摄采样. 为了准确体现观测场周边探测环境的季节变化、植物生长、人为活动等因素带来的障碍物变化, 进行了跨度为期 12 个月的图像采集, 所采集图像共包含 10 个不同的拍摄角度. 每个拍摄角度依照采集时间均匀抽取 100 张原始图像, 依照相关规范定义的建筑物、构筑物、树木、作物等主要障碍物, 在图像标注中将障碍物分为建筑和植物两类进行标注, 制成 1000 张图像规模的真实数据集.

3.2 评价指标

本文选用平均交并比 (mean intersection over union, $mIoU$) 作为主要的性能评价指标. $mIoU$ 作为语义分割算法性能的重要评估指标之一, 可以综合考虑预测结果的准确率和召回率, 比较不同算法之间的优劣. 从混淆矩阵的角度看, 多类别平均交并比 $mIoU$ 的计算方式如式 (6) 所示:

$$mIoU = \frac{1}{k+1} \cdot \sum_{i=0}^k \left(\frac{TP}{TP+FP+FN} \right) \quad (6)$$

其中, k 为语义类别标签数量, TP 为真正例像素个数即被正确分为该类别的像素点个数、 FP 为假正例像素个数即被错误分为该类别实际语义标签为其他类别的像素点个数、 FN 为假反例像素个数即被错误分为其他类别实际语义标签为该类别的像素点个数.

同时, 通过参数量 Params 和计算量 FLOPs 对不同模型的空间复杂度和时间复杂度进行比较. Params 代表模型中需要学习的参数总数, 该指标的大小决定了模型运行时占用存储和显存的大小. FLOPs 即每秒浮点运算次数, 代表模型在一次前向传播过程中需要执行浮点运算的量. 在实际应用中, 这两个性能指标可以用来衡量运行模型需要的硬件环境, 即芯片算力和显存大小.

3.3 实验参数配置

实验训练过程中对输入图像采用随机缩放、裁剪、水平翻转、旋转、加灰度条、高斯模糊和色域变

换的数据增强方式. 增强处理后的图像被统一缩放到 512×512 分辨率, 再输入网络模型. 本文选取 U-Net^[21]、SegNet^[8]、PSPNet^[9]、RefineNet^[10]、DeepLabV3^[11]、ENet^[15]、CGNet^[16]、CCNet^[22]、HRNet^[12]、SETR^[13]、SegFormer^[14]等方法参与实验比较, 并设置一致的训练参数. 实验将训练时的批大小设为 4, 训练轮数上限设为 500, 训练采用参数 β_1 和 β_2 分别设置为 0.9 和 0.999 的 Adam 优化器进行随机梯度下降. 整个训练过程中, 初始学习率被设置为 1×10^{-3} , 并在训练过程中逐渐下降至 1×10^{-5} . 训练中每 5 个训练周期对模型参数进行保存, 训练结束选取在测试集上的最佳效果参与方法比较. 除了本文方法外, 其余所有对比方法的实验都采用经典的分类交叉熵 (categorical cross-entropy) 作为损失函数. 实验设置在 NVIDIA GeForce RTX 4060 Ti 单卡环境中, 参与对比实验的模型均采用 PyTorch 1.10.1 框架实现.

3.4 对比实验

对比实验将本文方法与其他方法在 *mIoU*、Params、FLOPs 上进行定量比较. 其中 U-Net、SegNet、PSPNet、RefineNet、DeepLabV3 都是采用编码器-解码器结构的卷积神经网络方法, DeepLabV3-M 采用 MobileNet 作为主干网, DeepLabV3-X 采用 Xception 作为主干网. ENet、CGNet 是实时语义分割方法, CCNet 是注意力机制代表方法, SETR、SegFormer 是基于 Transformer 的语义分割方法, 其中 SegFormer-B0 和 SegFormer-B1 代表不同的模型量级. HRNet 方法是本文方法的基线. 对比实验定量数据如表 1 所示, 每个指标的最优结果加粗表示, 次优结果下划线表示.

表 1 本文方法与其他方法性能和计算成本的比较

Methods	<i>mIoU</i> (%)			FLOPs (G)	Params (M)
	All	Tree	Build		
U-Net	85.77	74.87	83.98	218.97	31.04
SegNet	85.86	75.75	84.63	160.83	29.44
PSPNet	88.12	80.36	86.89	258.33	65.57
RefineNet	89.17	81.75	88.69	253.55	154.28
DeepLabV3-M	86.57	79.31	83.47	26.44	5.81
DeepLabV3-X	91.26	86.33	89.8	83.42	54.71
CCNet	93.07	89.41	91.16	309.46	70.83
ENet	91.15	85.75	90.58	2.11	0.36
CGNet	92.64	88.18	92.26	<u>3.56</u>	<u>0.49</u>
HRNet-W18	94.79	90.42	94.19	24.67	10.02
HRNet-W32	<u>95.17</u>	<u>91.69</u>	<u>94.24</u>	67.16	30.94
SETR	88.38	80.5	87.06	88.27	86.21
SegFormer-B0	91.52	86.78	91.83	6.77	3.71
SegFormer-B1	92.67	88.49	92.37	13.25	13.68
Ours	95.82	92.71	94.57	27.53	10.21

表 1 可以看出本文提出方法在 *mIoU* 性能指标上有着明显提升, 与基线方法 HRNet-W18 相比, 总体 *mIoU* 提升了 1.03 个百分点, 在建筑物、构筑物上提升了 0.38 个百分点, 尤其是在对树木、作物等植物类别的分割精度上提升了 2.29 个百分点, 这与本文方法设计初衷相契合. 通过边界感知模块, 用强化边界特征引导的方式优化目标掩码边界处的分割精度, 使模型对目标边界处的变化更加敏感. 本文设计方法在总体分割精度和各类别分割精度上, 都取得了最佳效果. 同时与基线方法 HRNet-W18 相比我们仅提升了 1.89% 的参数量. 本文方法在 FLOPs、Params 和 *mIoU* 上均优于 HRNet-W32 方法, 意味着相比单纯增加模型宽度和复杂度来说, 本文方法更好针对了快速分割的具体应用领域.

对比实验可以发现实时语义分割网络如 ENet、CGNet 的分割效果反而比结构更加复杂的常规语义分割网络的效果来得更好, 分别达到了 91.15% 和 92.64% 的总体 *mIoU*. 这是因为气象探测环境中需要检测识别的障碍物类别不多, 树木、农作物和建筑物、构筑物之前的底层特征差别也很大, 不需要复杂的网络来理解大量语义类别的图像特征. Transformer 方法的代表模型 SegFormer 也取得了不错的效果. SegFormer-B0 和 B1 是系列模型里参数量最小的两种, SegFormer-B0 比较轻便且效果不错, *mIoU* 达到了 91.52%, SegFormer-B1 模型参数略大于 HRNet-W18 但计算复杂度较小, 总体效果 *mIoU* 为 92.67%.

本节进一步选择 DeepLabV3-M、CGNet、ENet 这 3 个方法和本文方法进行可视化比较. 这些方法均采用轻量化网络模型, 整体参数量在 10 M 以下. 图 4 中前两行为原始图像和相应的标注, 分割掩码和 RGB 图像融合后的可视效果如后 4 行所示. 可以看出总体上本文方法和 CGNet 有较好的分割效果. ENet 则受模型分辨率小的制约, 掩码边界多呈锯齿状, 轮廓精度较差. DeepLabV3-M 在语义类别上错误判别较多、分割效果欠佳. 如第 1 列中, 4 种方法虽然都分割出树林掩映中的房屋, 但是 DeepLabV3-M 标记的建筑类别偏少, CGNet 标记的建筑类别偏多. 在第 2 列中, 除本文方法外其他 3 种方法都未能准确分割出高塔的轮廓和灌木林之间的空洞. 第 3 列和第 4 列则重点比较建筑物存在部分遮挡时的分割效果, 可以看出 ENet 方法没有成功区分两张图像的障碍物变化. DeepLabV3-M 和

CGNet 在轮廓处理上欠佳, 在移动板房和主体大楼的分割掩码之间有很多粘连, 不如本文方法精确。

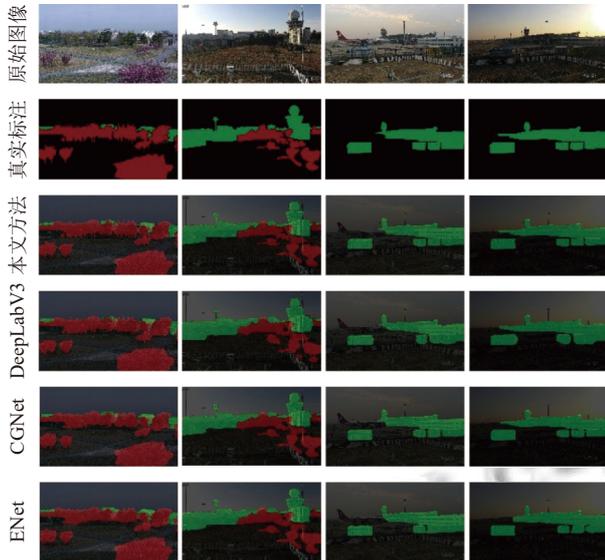


图4 本文方法与其他方法分割掩码可视化比较

3.5 消融实验

本节在不同的基线方法上进行消融实验, 选择 SegFormer-B0 和 HRNet-W18 分别作为 Transformer 方

法和 CNN 方法的代表来验证本文模块的有效性. 实验对边界感知模块 BAM 和边界特征引导模块 BGM 进行拆分比较. 对于 HRNet-W18, 将只采用 BAM 的记为“HRNet+A”, 用特征连接进行融合. 同时采用 BAM 和 BGM 的记为“HRNet+A+G”.

对于 SegFormer-B0, 如图 5 我们采用和本文模型思路一样的构造方式, 将主干网络第 1 阶段高分辨特征和最后一阶段高语义特征输入 BAM, 经 BGM 对不同尺度的进行边界特征引导后, 特征融合后流入解码阶段. 用“SegFormer+A”和“SegFormer+A+G”分别代表只采用 BAM 和同时采用 BAM、BGM 的模型。

表 2 展示了消融实验的结果, 结果显示, 相比于基线方法 HRNet-W18 和 SegFormer-B0, 采用边界感知模块的模型效果更优, 分别提升了 0.66% 和 0.58%. 与维度相连的特征融合相比, 采用注意力机制的 BGM 能够有效地强化主干特征在目标边界附近的关注度, 引导感知到的边界特征细化分割结果, 在 HRNet-W18 和 SegFormer-B0 基线上 BAM 和 BGM 模块仅增加了 0.19M 和 0.4M 的参数量. 与特征直接相连融合的相比, 采用 BGM 模块不仅减少了计算复杂度, 还进一步提升了分割精度。

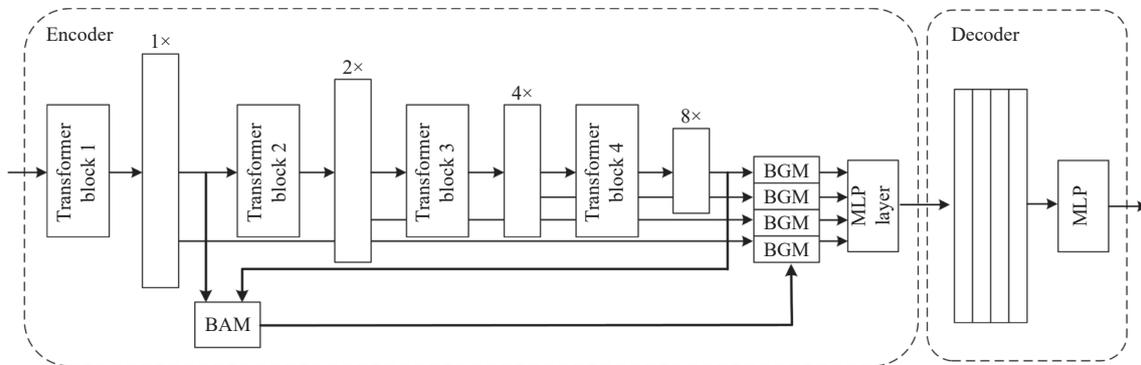


图5 基于 SegFormer 的边界引导模型结构图

表 2 消融实验

Methods	<i>mIoU</i> (%)	FLOPs (G)	Params (M)
HRNet	94.79	24.67	10.02
HRNet+A	95.45	27.98	10.18
HRNet+A+G	95.82	27.53	10.21
SegFormer	91.52	6.77	3.71
SegFormer+A	92.1	8.71	3.82
SegFormer+A+G	92.59	8.59	4.14

4 结论

本文针对地面观测气象站点探测环境保护工作存

在的困难, 探索采用计算机视觉方法以利用站点统一布局的监控设备, 对观测场周围探测环境进行图像采集和障碍物检测. 方法设计以轻量应用和高分辨率检测需求为背景, 以树木、作物等障碍物边界处易发生微小自然变化的难点为引导, 提出边界感知模块以学习高分辨率的目标边界特征和边界特征引导模块以注意力机制强化边界特征与主干特征有效融合. 在真实气象探测环境数据集上进行对比实验, 本文方法与主流的语义分割模型进行比较, 取得最优效果, *mIoU* 达到了 95.82%.

参考文献

- 1 师雨豪, 王昊亮, 刘玉宝. 联合同化闪电资料和地面常规观测资料对强对流天气预报的影响. 气象, 2024, 50(10): 1231–1242. [doi: [10.7519/j.issn.1000-0526.2024.021701](https://doi.org/10.7519/j.issn.1000-0526.2024.021701)]
- 2 张全军, 宏观, 吴东丽, 等. 中国农业气象观测业务的发展及未来展望. 中国农业气象, 2023, 44(8): 735–749. [doi: [10.3969/j.issn.1000-6362.2023.08.008](https://doi.org/10.3969/j.issn.1000-6362.2023.08.008)]
- 3 赵煜飞, 廖捷, 张强, 等. 1991~2020年中国地面气候值数据集研制. 大气科学, 2024, 48(2): 555–571. [doi: [10.3878/j.issn.1006-9895.2204.22010](https://doi.org/10.3878/j.issn.1006-9895.2204.22010)]
- 4 黄晶, 蔡思琴, 庞甜甜, 等. 基于主体建模的城市暴雨洪涝灾害预警策略仿真研究. 地球信息科学学报, 2024, 26(5): 1151–1165. [doi: [10.12082/dqxxkx.2024.230311](https://doi.org/10.12082/dqxxkx.2024.230311)]
- 5 姜天姣, 陈杉, 韩莉莉. 地面气象台站探测环境保护规划初探——以天津市为例. 人民城市, 规划赋能——2022中国城市规划年会论文集. 武汉: 中国城市规划学会, 2023.
- 6 钟静, 张沪生, 林小杰, 等. 气象探测环境保护监测系统设计与实现. 气象水文海洋仪器, 2022, 39(2): 99–101. [doi: [10.3969/j.issn.1006-009X.2022.02.029](https://doi.org/10.3969/j.issn.1006-009X.2022.02.029)]
- 7 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 3431–3440.
- 8 Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481–2495. [doi: [10.1109/TPAMI.2016.2644615](https://doi.org/10.1109/TPAMI.2016.2644615)]
- 9 Zhao HS, Shi JP, Qi XJ, *et al.* Pyramid scene parsing network. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 6230–6239.
- 10 Lin GS, Milan A, Shen CH, *et al.* RefineNet: Multi-path refinement networks for high-resolution semantic segmentation. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 5168–5177.
- 11 Chen LC, Zhu YK, Papandreou G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 833–851.
- 12 Wang JD, Sun K, Cheng TH, *et al.* Deep high-resolution representation learning for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(10): 3349–3364. [doi: [10.1109/TPAMI.2020.2983686](https://doi.org/10.1109/TPAMI.2020.2983686)]
- 13 Zheng SX, Lu JC, Zhao HS, *et al.* Rethinking semantic segmentation from a sequence-to-sequence perspective with Transformers. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 6877–6886.
- 14 Xie EZ, Wang WH, Yu ZD, *et al.* SegFormer: Simple and efficient design for semantic segmentation with Transformers. Proceedings of the 35th International Conference on Neural Information Processing Systems. IEEE, 2021. 924.
- 15 Lieman-Sifry J, Le M, Lau F, *et al.* FastVentricle: cardiac segmentation with ENet. Proceedings of the 9th International Conference on Functional Imaging and Modeling of the Heart. Toronto: Springer, 2017. 127–138.
- 16 Wu TY, Tang S, Zhang R, *et al.* CGNet: A light-weight context guided network for semantic segmentation. IEEE Transactions on Image Processing, 2021, 30: 1169–1179. [doi: [10.1109/TIP.2020.3042065](https://doi.org/10.1109/TIP.2020.3042065)]
- 17 Yuan YH, Chen XL, Wang JD. Object-contextual representations for semantic segmentation. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 173–190.
- 18 He YL, Yi YG, Zheng CX, *et al.* BGF-Net: Boundary guided filter network for medical image segmentation. Computers in Biology and Medicine, 2024, 171: 108184. [doi: [10.1016/j.combiomed.2024.108184](https://doi.org/10.1016/j.combiomed.2024.108184)]
- 19 Li S, Tang XH, Cao B, *et al.* Boundary guided network with two-stage transfer learning for gastrointestinal polyps segmentation. Expert Systems with Applications, 2024, 240: 122503. [doi: [10.1016/j.eswa.2023.122503](https://doi.org/10.1016/j.eswa.2023.122503)]
- 20 Zhou Q, Wang LJ, Gao GW, *et al.* Boundary-guided lightweight semantic segmentation with multi-scale semantic context. IEEE Transactions on Multimedia, 2024, 26: 7887–7900. [doi: [10.1109/TMM.2024.3372835](https://doi.org/10.1109/TMM.2024.3372835)]
- 21 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention. Munich: Springer, 2015. 234–241.
- 22 Wenzek G, Lachaux MA, Conneau A, *et al.* CCNet: Extracting high quality monolingual datasets from Web crawl data. Proceedings of the 12th Conference on Language Resources and Evaluation. Marseille: ACL, 2020. 4003–4012.

(校对责编: 张重毅)