

# 基于改进 YOLOv8n 的变电设备 红外图像实例分割算法\*

李冰 杜喜英 王玉莹 翟永杰

(华北电力大学自动化系 保定 071003)

**摘要:** 变电设备是电网输变电过程的重要组成部分,为保证电网的正常运行,需对变电设备进行故障诊断,红外图像中变电设备的精确分割是故障诊断的关键步骤。针对红外图像复杂场景中变电设备分割时存在的分割精度低和漏分割的问题,提出一种基于改进 YOLOv8n 的变电设备红外图像实例分割算法。首先设计一种上下文引导的特征增强下采样块替换 YOLOv8n 中的下采样卷积层,充分利用上下文信息和全局信息,增强模型对复杂场景的理解能力;然后引入可变形卷积重构 Backbone 中的 C2f 模块,增强对不规则设备特征的提取能力;最后用 Wise-IoUv2 对损失函数进行优化,提高模型的泛化性和分类能力。使用变电设备红外图像数据集对该模型进行实验验证,实验结果表明,相较于 YOLOv8n 基准模型,本文所提方法的 mAP50 和 mAP50:95 分别提高了 4.2% 和 3.5%,所提方法能够较好地解决复杂场景下设备漏分割的问题,有效提高变电设备实例分割的准确率。

**关键词:** 变电设备实例分割;YOLOv8n;可变形卷积;Wise-IoUv2

**中图分类号:** TP391.4;TN081 **文献标识码:** A **国家标准学科分类代码:** 520.6040

## Instance segmentation algorithm for infrared images of substation equipment based on improved YOLOv8n

Li Bing Du Xiying Wang Yuying Zhai Yongjie

(Department of Automation, North China Electric Power University, Baoding 071003, China)

**Abstract:** The substation equipment is an important component of the power transmission and transformation process of the power grid. To ensure the normal operation of the power grid, it is necessary to diagnose faults in substation equipment, and accurate segmentation of substation equipment in infrared images is a key step in fault diagnosis. In response to the low segmentation accuracy and missed segmentation issues faced when segmenting substation equipment in complex scenes of infrared images, an instance segmentation algorithm for substation equipment in infrared images based on improved YOLOv8n is proposed. Firstly, a context-guided feature-enhanced down-sampling block is designed to replace the down-sampling convolutional layer in YOLOv8n, fully utilizing contextual and global information to enhance the model's understanding of complex scenes. Then, the C2f module is introduced to reconstruct the deformable convolution in the Backbone, enhancing the extraction capability of irregular equipment features. Finally, the loss function is optimized using Wise-IoUv2 to improve the model's generalization and classification ability. Using the infrared image dataset of substation equipment to experimentally verify the model, the experimental results show that compared to the YOLOv8n baseline model, the proposed method in this paper has increased mAP50 and mAP50:95 by 4.2% and 3.5% respectively. The proposed method can better solve the problem of missed segmentation of equipment in complex scenes and effectively improve the accuracy of substation equipment instance segmentation.

**Keywords:** substation equipment instance segmentation; YOLOv8n; deformable convolution; Wise-IoUv2

## 0 引言

变电站作为电网输变电过程的一个重要环节,站内变

电设备的故障诊断对维护电力系统的安全运行有着至关重要的作用<sup>[1]</sup>。变电设备出现故障时若不能及时排除,容易引起起火、爆炸等危险情况,影响电网正常运行,因此需要

收稿日期:2024-04-16

\* 基金项目:国家自然科学基金面上项目(62373151)、国家自然科学基金联合基金项目重点支持项目(U21A20486)、河北省自然科学基金面上项目(F2023502010)、中央高校基本科研业务费专项资金(20237488)资助

对变电设备进行状态监测。变电设备发生故障时一般会伴随着温度的急剧上升,通常借助红外热成像技术来监测变电设备温度异常进行故障诊断<sup>[2-3]</sup>。传统的故障诊断对操作人员的专业技能要求较高且效率较低,无法实现变电设备故障在线诊断<sup>[4]</sup>。

随着图像处理技术的发展,故障诊断朝着基于红外图像智能诊断的方向发展<sup>[5]</sup>。通过红外热成像技术获取的图像中背景温度信息复杂、干扰因素较多,难以直接进行设备的故障诊断,因此可以利用图像分割方法将目标设备分割提取出来,再结合目标的温度信息进行故障诊断,能够有效提高诊断结果的准确度<sup>[6]</sup>。目标设备分割的精确度对故障诊断结果影响较大,因此研究如何提高红外图像中变电设备分割的准确度非常必要。

为了实现红外图像中目标设备的准确分割,国内外研究人员基于不同方法开展了大量研究工作。张秋铭等<sup>[7]</sup>通过引入边缘能量项、径向基函数和内部能量项对 Chan-Vese 模型进行改进完成了对红外图像中电力设备的有效分割。Li 等<sup>[8]</sup>使用两个并行特征编码器提取 RGB 图像和红外图像的特征,利用多模态特征和多尺度特征,完成对背光场景下绝缘子的分割工作。Wang 等<sup>[9]</sup>设计了一种交叉指导网络,在学习语义特征时构建交叉引导模块进行边缘细节的有效利用,根据语义特征和边缘特征获取更准确的变电设备边缘。Zhao 等<sup>[10]</sup>将变电设备显著性视觉特征作为先验知识,设计基于电力领域专业知识和视觉特征推理的双分支实例分割模型,实现对变电设备较高精度分割。

上述文献采用不同的方法对红外图像中的变电设备进行了分割提取,都取得了一定的成果,但是准确度和实时性没有同时兼顾,在实际应用中仍有不足<sup>[11]</sup>。YOLO 系列实例分割算法在分割精度和速度方面取得了较好的平衡,因此具有广泛的应用场景。Kang 等<sup>[12]</sup>基于 YOLO 分割框架,采用尺度序列特征融合模块和三重特征编码器模块进行不同尺度特征的提取和融合,对细胞实现了较高精度和速度的实例分割。Bai 等<sup>[13]</sup>先用 FocalNext block 替换了 neck 中的 C2f 模块,再将 EMA 注意力机制加入主干网络与 neck 之间,最后用一个上下文聚合模块对 C2f 和 SPPF 优化,实现对 YOLOv8 的改进,完成了建筑工地图像的实例分割。王译崧等<sup>[14]</sup>用融合 ECA 注意力的 RepVGG 模块代替 YOLOv8 骨干网络中的卷积层,同时在 C2f 模块中加入 eSE 模块对 YOLOv8 进行改进,提出 Rep-YOLOv8 车辆行人检测分割算法。以上对 YOLOv8 的改进方法都是基于特定场景的特点完成的,而在变电设备实例分割领域缺乏对 YOLO 系列算法的研究工作。变电设备数据集为红外图像,其特点与上述文献中研究的可见光图像和医学图像具有明显差异,故需针对变电设备数据集的特点进行 YOLOv8 分割工作的研究。

目前变电设备实例分割主要存在以下问题:1)变电站现场环境复杂,设备密集,采集样本时由于拍摄角度和距离

等原因,不可避免地会出现遮挡、设备不完整等复杂情景,模型不能充分利用上下文信息和全局信息导致漏分割或分割不完整;2)红外图像中变电设备外形不规则且与背景对比度低,模型不能充分提取其有效特征反而会引入背景噪声,导致准确度下降。

针对上述问题,本文提出了一种基于改进 YOLOv8n 的变电设备实例分割算法,先利用传统卷积提取局部特征,同时用空洞卷积提取周围上下文信息并结合全局信息对 YOLOv8n 直接用卷积进行下采样的方式升级改进,建立局部特征、上下文特征和全局特征之间的依赖关系;再引入可变形卷积对特征提取模块进行改进,可变形卷积通过引入偏移量可以灵活调整采样位置,提取到的特征能更精准地描述设备特点,从而增强对形状不规则变电设备的特征提取能力;最后引入损失函数 Wise-IOUv2(WIOUv2)替换 CIOU, WIOUv2 能够聚焦于困难示例,提高模型的分类能力和泛化性,进而提升对变电设备的分割精度。

## 1 YOLOv8 原理

YOLOv8 是 Ultralytics 公司于 2023 年发布的 YOLO 系列的其中一员<sup>[15]</sup>,目前支持图像分类、物体检测和实例分割任务,YOLOv8n 能在取得较高准确度的同时拥有更快的速度,本文选择在 YOLOv8n 模型基础上进行改进。YOLOv8 实例分割模型由主干网络 Backbone、颈部网络 Neck 和分割层 Head 三部分组成。在 YOLOv8 中,主干网络通过卷积层和 C2f 模块完成对输入图像中目标特征的提取,空间金字塔池化模块(spatial pyramid pooling fast, SPPF)提取不同感受野特征,使网络更好地适应不同尺度的目标;然后将特征送入颈部网络,颈部网络是特征金字塔网络和路径聚合网络结构,实现了自上而下和自下而上的特征金字塔,完成多尺度特征融合;最后将 3 个融合特征送入分割层,分割层在检测分支的基础上增加了一个语义分割分支,将检测结果与分割结果结合完成对目标设备的分割。

## 2 改进 YOLOv8 模型

YOLOv8n 只采用传统卷积进行下采样的方式无法充分利用上下文信息,导致模型对复杂场景理解力有限,造成设备漏分割或分割不完整;另外,YOLOv8n 的 C2f 模块对形状不规则设备的特征提取能力不足,导致分割精度较低。针对 YOLOv8n 在复杂红外背景下变电设备漏分割和分割精度低的问题,本文对 YOLOv8n 的改进主要有三点:首先设计上下文引导的特征增强下采样块,保持模型中第一个下采样卷积层不变,其他下采样卷积层用该模块进行替换,通过空洞卷积提取上下文信息,同时有效利用全局信息,提高模型对像素分类的准确度和对复杂场景的理解能力;然后结合可变形卷积建立不规则特征提取模块 C2f\_DCN,用

C2f\_DCN 模块替换 Backbone 中的 C2f 模块,增强模型对形状不规则变电设备的特征提取能力;最后用 WIOUv2 替

换 CIOU,对损失函数进行优化,提高模型泛化性和分类能力。改进后的网络结构如图 1 所示。

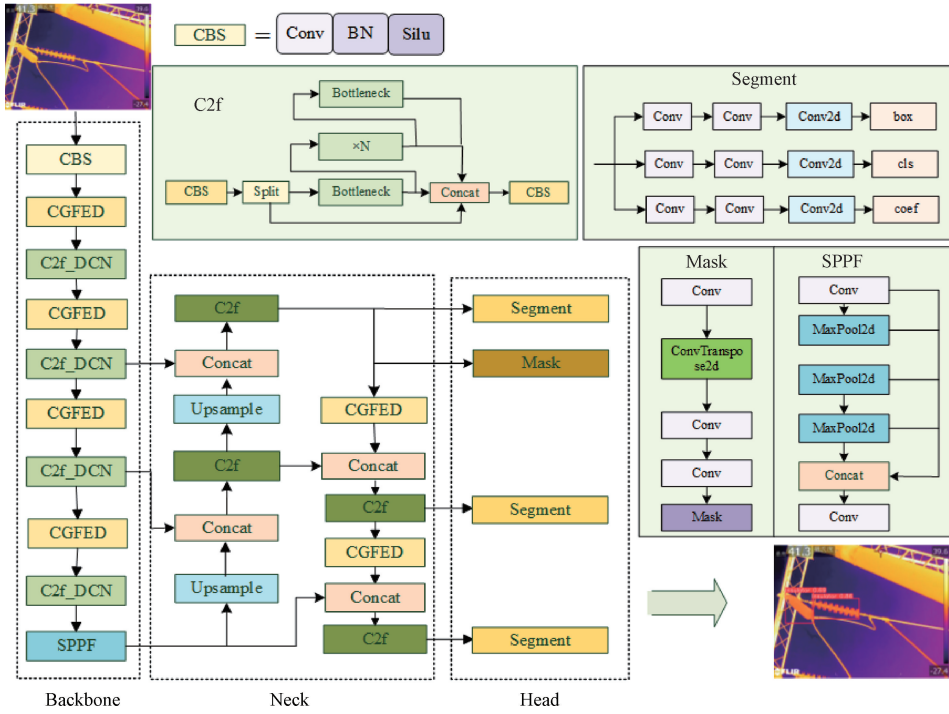


图 1 改进 YOLOv8 结构

## 2.1 上下文引导的特征增强下采样块原理

变电站环境复杂,巡检时存在变电设备遮挡、拍摄不完整等情况,由于 YOLOv8n 中的下采样部分采用传统的  $3 \times 3$  卷积层, $3 \times 3$  卷积感受野有限,无法充分提取上下文信息和全局信息,对局部特征、上下文特征和全局特征之间的联系缺乏有效建模,导致出现漏分割和分割不完整的问题<sup>[16]</sup>。受 CGNet<sup>[17]</sup> 和 SENetv2<sup>[18]</sup> 的启发,本文设计了适用于红外图像中变电设备分割的上下文引导特征增强下采样块(context guided feature enhancement down-sampling block, CGFED),充分利用上下文信息,建立局部特征和上下文特征之间的依赖关系,同时结合全局信息增强目标设备的特征并抑制背景噪声。该模块的结构如图 2 所示。

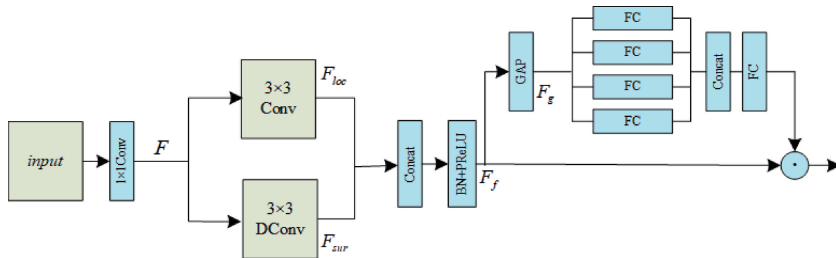


图 2 上下文引导的特征增强下采样块(CGFED)

在下采样块中,首先,输入  $input$  通过一个  $1 \times 1$  的卷积层进行特征提取得到特征图  $F$ ,  $F$  的尺寸为原来输入特征图尺寸的一半、通道数翻倍,公式如下:

$$F = \text{PReLU}(\text{BN}(\text{Conv}(input))) \quad (1)$$

其次,特征图  $F$  通过并联的局部特征采样单元和一个周围上下文特征采样单元分别提取到局部特征  $F_{loc}$  和周围上下文特征  $F_{sur}$ 。局部特征采样单元通过一个  $3 \times 3$  的卷积层实现,周围上下文特征采样单元利用一个  $3 \times 3$  的空洞卷积来增加感受野的大小,捕获更宽广的上下文信息。

然后,将局部特征和周围上下文特征送入特征融合单元后得到融合特征图  $F_f$ ,特征融合单元具体实现公式如下:

$$F_f = \text{PReLU}(\text{BN}(\text{Concat}(F_{loc}, F_{sur}))) \quad (2)$$

其中,局部特征和周围上下文特征通过 Concat 操作形成一个综合特征,综合特征再依次进行批量归一化(batch normalization, BN)和参数化修正线性单元(parametric relu, PReLU)激活,学习更加复杂的非线性关系,以增强特征表示的能力。

最后,为利用全局信息,设计一个全局信息采样器,通过学习全局信息来增强有用特征,减弱背景噪声的干扰。全局信息采样器首先通过全局平均池化(global average pooling, GAP)对整个特征图的全局信息进行聚合,得到全局特征向量  $\mathbf{F}_g$ , 全局特征向量  $\mathbf{F}_g$  输入到多分支全连接层(fully connected layer, FC),提高特征表达的精细度并增强网络对全局信息的整合能力,多分枝 FC 层的输出拼接后送入到 FC 层生成通道权重  $F'_g$ 。最后将通道权重和融合特征相乘实现对融合特征的缩放,凸显有效特征信息。实现过程如下:

$$\begin{cases} \mathbf{F}_g = \text{GAP}(\mathbf{F}_f) \\ \mathbf{F}_g^1 = \text{FC}(\mathbf{F}_g), \\ \mathbf{F}_g^2 = \text{FC}(\mathbf{F}_g), \\ \mathbf{F}_g^3 = \text{FC}(\mathbf{F}_g), \\ \mathbf{F}_g^4 = \text{FC}(\mathbf{F}_g) \\ \mathbf{F}'_g = \text{FC}(\text{Concat}(\mathbf{F}_g^1, \mathbf{F}_g^2, \mathbf{F}_g^3, \mathbf{F}_g^4)) \\ \mathbf{F}'_f = \mathbf{F}'_g \odot \mathbf{F}_f \end{cases} \quad (3)$$

## 2.2 不规则特征提取模块原理

由于部分变电设备形状不规则,而标准卷积层在进行特征提取时感受野形状是固定的矩形,因此导致特征提取不充分。可变形卷积的采样形状是可变的,感受野能够更加贴合变电设备的不规则形状,相比传统的卷积它能够较好的排除背景噪声干扰,得到更多有效信息。结合可变形卷积,本文构建了不规则特征提取模块,增强对变电设备

的特征提取能力,同时减少背景的干扰。

可变形卷积(DCNv1)<sup>[19]</sup>在标准卷积计算的基础上,通过学习额外的偏移量  $\Delta p_n$ ,卷积核在每个中心采样位置上动态调整大小和位置,采样点位置根据当前图像内容进行自适应的变化,从而适应形状不规则的目标。可变形卷积计算式如下:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (4)$$

式中: $x$  为输入特征图, $y$  为输出特征图, $w(p_n)$  表示  $p_n$  位置的权值, $p_0$  为中心采样位置, $p_n$  表示采样区域  $R$  中以  $p_0$  为中心的所有采样位置。

由于 DCNv1 采样范围依旧超出目标范围,提取的特征仍然存在无关内容干扰,Zhu 等<sup>[20]</sup>提出 DCNv2 对 DCNv1 进行了改进,DCNv2 通过增添每个采样点的权重系数  $\Delta m_n$  来调节不同采样点对特征的影响,在调整感受野贴合目标形状的同时减少噪声干扰,增强有效特征提取能力,计算公式如下:

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \cdot \Delta m_n \quad (5)$$

通过用可变形卷积模块(DCNv2 BN SiLu, DBS)替换 C2f 模块 Bottleneck 中的传统卷积层构建不规则特征提取模块 C2f\_DCN,增强对变电设备的特征提取能力,提高对有关目标特征的关注度,减少背景噪声,以实现对形状不规则的变电设备目标特征的有效提取。C2f\_DCN 结构如图 3 所示。

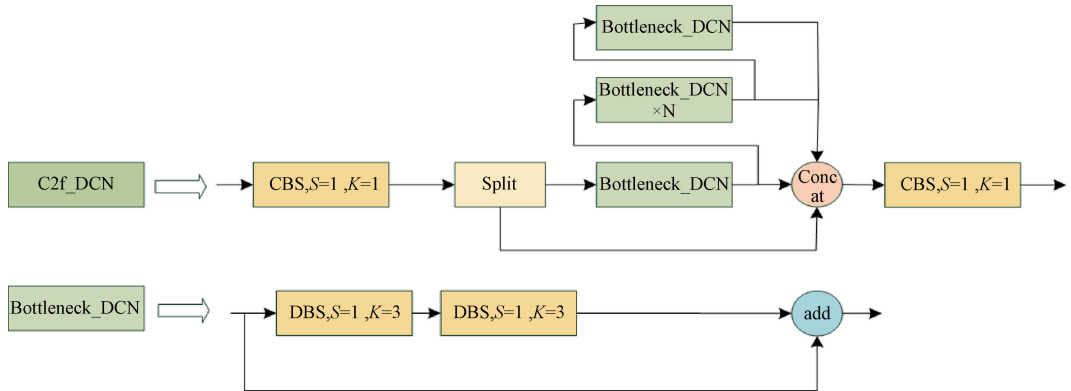


图 3 不规则特征提取模块

构建不规则特征提取模块后,将 YOLOv8n 原始结构中 Backbone 部分的 C2f 模块用 C2f\_DCN 模块替换得到适应不规则形状的变电设备的特征提取网络,增强对变电设备特征的有效提取能力,提升分割精度。

## 2.3 损失函数优化

YOLO8n 实例分割的损失函数由分类损失、边界框回归损失和分割损失组成。在计算分割损失时,仅对每个检测框内有效的 mask 做分割损失的计算,边界框准确度会对分割结果产生影响。YOLOv8n 采用 CIOU<sup>[21]</sup>作为回归的损失函数,CIOU 综合考虑了图像重叠面积、中心点距离

和纵横比,在目标检测任务中表现良好。但由于数据集中不可避免地会存在低质量标注样本,CIOU 中距离和纵横比几何度量会加剧对低质量标注样本的惩罚进而会导致模型泛化性低。为此,本文引入 WIOUv2<sup>[22]</sup>替换 CIOU 完成对损失函数的优化,可以削弱几何度量的惩罚,提升模型泛化性,同时聚焦于困难样本,提升分类性能。WIOUv2 是在 WIOUv1 基础上构建的,WIOUv1 计算公式如下:

$$\begin{cases} L_{WIOUv1} = R_{WIOU} L_{IOU} \\ R_{WIOU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \end{cases} \quad (6)$$



式中:  $L_{IOU}$  为交并比损失,  $R_{WIOU}$  为距离注意力,  $(x, y)$  和  $(x_{gt}, y_{gt})$  分别为预测框和为真实框中心点坐标,  $W_g$  和  $H_g$  分别为最小包围框的宽和高,  $*$  表示分离操作, 使其变为不带梯度的常量, 该方法没有引入长宽比且有效地消除了阻碍收敛的因素。为了使模型更能聚焦于困难样本, 从而获得分类性能的提升, 利用单调聚焦系数  $L_{IOU}^*$  构建 WIOUv2 计算公式如下:

$$L_{WIOUv2} = \left( \frac{L_{IOU}^*}{L_{IOU}} \right)^\gamma L_{WIOUv1} \quad (\gamma > 0) \tag{7}$$

式中:  $\overline{L_{IOU}}$  是归一化因子;  $\gamma$  为调节因子, 模型对低质量样本的聚焦度与调节因子呈正相关性。

3 实验结果与分析

本文实验所采用的数据集包含绝缘子、套管、避雷器、电流互感器、电压互感器五类变电设备, 总共有 1 519 张变电设备图像。数据集中训练集与测试集数量之比约为 7:3, 其中训练集 1 064 张, 测试集 455 张。

在实验过程中, 本文设置的输入图像尺寸为  $640 \times 640$ , 训练批次 Batchsize 大小为 4, 训练周期一共为 100 个 epoch, 采用 SGD 优化器对模型进行优化, 并将初始学习率设置为 0.01, 动量设置为 0.937, 权重衰减率设置为 0.000 5。实验平台参数如表 1 所示。

表 1 实验平台参数

参数	配置
操作系统	Ubuntu18.04
深度学习框架	Pytorch 1.11.3
CPU 型号	Intel(R) Xeon(R) Gold 6148 CPU
显卡(GPU)型号	NVIDIA GeForce RTX 3090
显卡内存	24G
编程语言	Python3.8

3.1 评价指标

本文采用平均精度均值(mean average precision, mAP)、参数量(Params)、每秒帧率(frames per second, FPS)、计算量(giga floating-point perations per second, GFLOPs)作为评估模型实例分割性能的指标。特别的, mAP50 表示取计算预测掩膜和标注掩膜的交并比阈值为 0.5, 认为交并比在阈值以上时预测掩膜为正确结果, 并计算对应 mAP 值; mAP50:95 表示以 0.05 为步长, 计算交并比阈值从 0.5 到 0.95 的 mAP, 然后取平均值。

$$P = \frac{TP}{TP + FP} \times 100\% \tag{8}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{9}$$

$$AP = \int_0^1 P(R) dR \tag{10}$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP(i) \times 100\% \tag{11}$$

式中:  $P$  和  $R$  分别为精确率和召回率,  $n$  是变电设备的类别数量, TP(true positive)为正例且预测正确, FP(false positive)为正例但预测错误, FN(false negative)为负例且预测错误。

3.2 模块对比实验

1) 下采样模块对比

为了验证本文所提出的下采样模块 CGFED 的有效性和先进性, 对原 YOLOv8n 下采样、上下文引导下采样(context guided block, CGB)和本文设计的 CGFED 下采样进行对比实验, 实验结果如表 2 所示。

表 2 下采样模块对比实验

模块	mAP50/%	mAP50:95/%	Params/M
Baseline	85.7	65.3	3.25
CGB	87.4	66.3	3.87
CGFED	<b>88.3</b>	<b>66.8</b>	<b>3.93</b>

通过实验结果可以发现, 相比原 YOLOv8n 下采样和 CGB 下采样, 本文设计的 CGFED 下采样具有更高的分割精度, 证明 CGFED 下采样的设计更加有效和先进。

2) 卷积改进模块对比

本小节分别测试了用不同卷积对 C2f 模块进行改进的效果, 对可切换空洞卷积 SAConv、感受野注意卷积 RFACnv 和可变形卷积 DCNv2 改进的 C2f 模块优化骨干网络的结果进行对比, 实验结果如表 3 所示。

表 3 卷积改进模块对比实验

模块	mAP50/%	mAP50:95/%	Params/M
Baseline	85.7	65.3	3.25
SAConv	87.3	66.0	3.56
RFACnv	86.5	66.1	3.30
DCNv2	<b>88.5</b>	<b>67.7</b>	<b>3.42</b>

通过对比实验结果发现, 相比 Baseline, SAConv 和 RFACnv 均可以提升模型的分割性能, 但 DCNv2 能获得更优秀的分割性能提升。SAConv 对不同尺度的特征具有良好的提取能力, RFACnv 能够更有效地处理图像中的细节, 而 DCNv2 的感受野更加灵活, 对不同尺度和不规则特征的提取能力更为优秀, 能在提取有效特征的同时减少红外背景噪声的引入, 相比之下, DCNv2 更适合变电设备红外图像分割任务。

3) 损失函数对比

为了选取合适的损失函数, 本小节对比了 CIOU、SIOU、GIOU 和 WIOUv2 对模型的影响, 对比实验结果如表 4 所示。

表 4 损失函数对比实验

模块	mAP50/%	mAP50:95/%	Params/M
CIOU	85.7	65.3	3.25
SIOU	86.6	65.0	3.25
GIOU	86.9	65.3	3.25
WIOUv2	<b>87.6</b>	<b>66.9</b>	<b>3.25</b>

对比表 4 结果可以发现, SIOU 和 GIOU 相比原

YOLOv8 采用的 CIOU 表现出了一定的优势, 但二者带来的模型分割能力的提升均不及 WIOUv2, WIOUv2 在变电设备分割任务中更具有有效性。

3.3 消融实验

为了验证本文所提出改进方案对红外图像中变电设备实例分割的有效性, 设计了消融实验, 以 YOLOv8n 为基线, 分别验证了上下文引导的特征增强下采样模块、不规则特征提取模块和损失函数的影响。实验结果如表 5 所示。

表 5 消融实验

YOLOv8n	CGFED	C2f_DCN	WIOUv2	mAP50/%	mAP50:95/%	Params/M	GFLOPs	FPS/(f·s <sup>-1</sup> )
✓				85.7	65.3	3.25	12.0	88
✓	✓			88.3	66.8	3.93	13.1	74
✓		✓		88.5	67.7	3.42	11.6	67
✓			✓	87.6	66.9	3.25	12.0	104
✓	✓	✓		88.9	68.2	4.09	12.7	57
✓	✓	✓	✓	<b>89.9</b>	<b>68.8</b>	<b>4.09</b>	<b>12.7</b>	<b>60</b>

通过表 5 可以看出, 仅用 CGFED 模块替换 YOLOv8n 中的下采样卷积层时, mAP50 和 mAP50:95 分别提高 2.6% 和 1.5%, 参数量和计算量只有少量增长。该实验结果说明, 本文提出的下采样模块能有效利用局部信息、上下文信息和全局信息, 增强模型对复杂场景的理解能力, 提升分割精度。

在 YOLOv8n 的基础上使用不规则特征提取模块后, mAP50 和 mAP50:95 分别提高 2.8% 和 2.4%, 结果表明, 结合可变形卷积的不规则特征提取模块对变电设备的特征提取能力更强, 能有效缓解传统卷积不能灵活地适应变电设备的不规则形状导致的特征提取不充分的问题, 也减少了背景噪声的引入。将 CIOU 替换为 WIOUv2 后, mAP50 和 mAP50:95 分别提高 1.9% 和 1.6%, 损失函数的优化改善了模型的泛化能力, 增强了模型的分类能力从

而提升了分割的准确度。

当将 3 种改进策略同时引入后, 相较于基线模型, mAP50 提高 4.2%, mAP50:95 提高 3.5%, 模型分割精度达到最高, 说明本文所提出的 3 种改进策略能提升红外图像中变电设备的实例分割精度, 同时, 参数量和计算量上涨较少, FPS 能达到 60 帧/s, 仍可满足实际应用中的实时性要求。

3.4 对比试验

为了进一步验证本文所提出的变电设备实例分割算法的有效性和先进性, 选取经典实例分割模型 YOLACT<sup>[23]</sup>、Mask R-CNN<sup>[24]</sup>、YOLOv5n、YOLOv5s、YOLOv7<sup>[25]</sup>、YOLOv9<sup>[26]</sup> 以及文献[12]、文献[13]和文献[14]改进的 YOLOv8 实例分割模型进行对比, 在同一数据集和同样的条件下进行对比试验, 实验结果如表 6。

表 6 对比试验

模型	mAP50/%	mAP50:95/%	Params/M	GFLOPs	FPS/(f·s <sup>-1</sup> )
YOLACT	82.2	56.5	—	—	7
Mask R-CNN	85.9	60.6	—	—	—
YOLOv5n	80.1	53.6	1.88	6.8	98
YOLOv5s	86.3	59.6	7.41	25.7	89
YOLOv7	88.0	64.8	37.86	142.0	57
YOLOv8n	85.7	65.3	3.25	12.0	88
YOLOv9	88.8	68.8	57.48	368.7	12
文献[12]	86.6	63.6	48.46	154.7	27
文献[13]	87.4	67.3	4.83	14.9	23
文献[14]	80.6	54.6	3.72	12.3	32
本文	<b>89.9</b>	<b>68.8</b>	<b>4.02</b>	<b>12.7</b>	<b>60</b>

由表 6 实验结果可知,相比于两阶段实例分割算法 Mask R-CNN,本文所提算法 mAP50 领先 4%,mAP50:95 领先 8.2%,分割精度具有明显优势。相较于一阶段实例分割算法 YOLACT,本文模型的精度和速度都要更高。YOLOv5n 实例分割算法具有最少的参数量和计算量,但 mAP50 和 mAP50:95 比本文模型分别低 9.8%和 15.2%,同系列算法 YOLOv5s 模型参数量和计算量更大,分割效果相较 YOLOv5n 也有所提升,但精度仍不及本文算法且本文算法的参数量和计算量更少。YOLOv7 取得了与本文算法比较接近的 mAP50,仅比本文算法低 1.9%,但本文算法的 mAP50:95 比 YOLOv7 高出了 4.4%,且模型参数量和计算量分别只是 YOLOv7 的 10.62%和 8.94%。相较于原始 YOLOv8n 模型,本文所提改进模型较好地平衡了模型精度和模型复杂度,在 mAP50 和 mAP50:95 分别提升 4.2%和 3.5%的情况下,模型参数量和计算量仅有

少量增长。YOLOv9 在分割精度方面获得了除本文所提算法外最佳结果,但同时也有最大的模型参数量和计算量,在实际应用中不及 YOLOv8n 高效。在与其他学者改进的 YOLO 系列算法的对比中,三者效果均不及本文所提算法,虽然 3 种算法对基准模型的各部分做出了改进,但改进的出发点未从变电设备实例分割任务特性角度做考虑,因此在处理变电设备分割任务时,效果均不及本文所提算法。综上分析对比,本文的改进模型不仅在变电设备分割精度方面达到了最优,同时保持了较少的模型参数量和计算量,有一定实时性,具有较高的应用价值。

3.5 分割结果可视化

为了直观观察本文所提改进策略的有效性,对 YOLACT、Mask R-CNN、YOLOv5n、YOLOv7、YOLOv8n 和本文算法对 5 类变电设备实例分割结果进行了可视化处理。结果如图 4 所示。

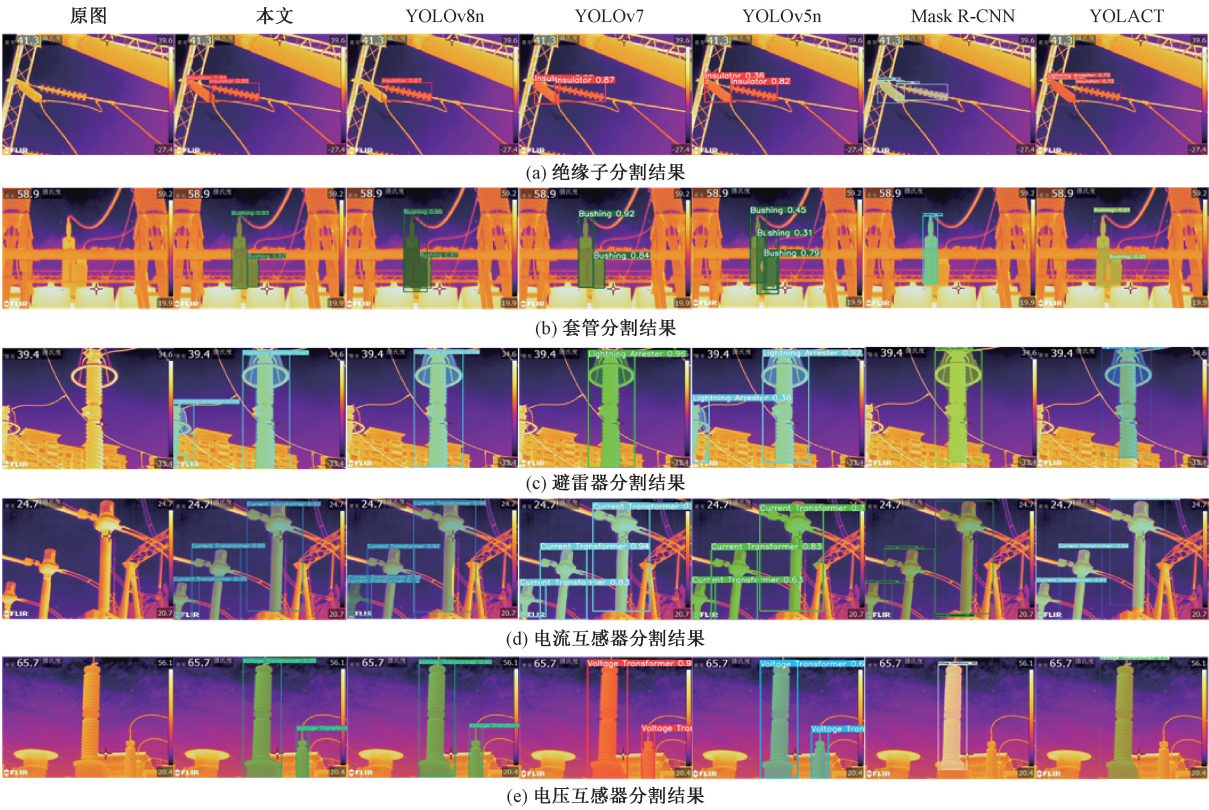


图 4 分割结果可视化

从图 4(a)可以看出,在分割绝缘子时,其中一个绝缘子外观形态和位置与其他绝缘子差别较大属于复杂情况,YOLOv8n 出现了漏分割现象。从图 4(b)可以看出,在对套管的分割中,由于 2 个套管部分被遮挡且较密集,YOLOv8n 和 YOLOv5n 误将 2 个套管分割为一体,Mask R-CNN 出现漏分割现象。从图 4(c)可以看出,对避雷器分割时,其中一个避雷器只出现部分结构,除改进模型和 YOLOv5n 其余模型均没有完成这部分避雷器的分割,而

YOLOv5n 的在分割完整避雷器时效果不佳。从图 4(d)可以看出,在分割电流互感器时,5 种对比模型对部分像素分类错误,导致无法完整分割出设备。从图 4(e)可以看出,对电压互感器分割的过程中,由于目标设备不规则且部分结构与背景对比度低,Mask R-CNN 和 YOLACT 只分割出其中一个设备且准确度较低,另外 3 种对比模型的分割完整度均有所欠缺。在以上 5 种设备分割中,改进后的模型在复杂情况下均可以准确分割出目标设备,未出现漏分



割和分割不完整的问题。以上结果表明,本文引入的改进策略有三方面的效果:一是可以有效利用上下文信息和全局信息,提高对复杂场景的理解能力,增强对像素的分类能力,提升对目标设备分割的完整度;二是可以增强对有效特征的提取能力,抑制背景噪声,提升分割精度;三是可以增强模型泛化性和分类准确性,提升对困难样本的分割能力。

## 4 结 论

针对 YOLOv8n 在分割红外图像中变电设备时出现的漏分割和分割精度低的问题,本文通过设计上下文引导的特征增强下采样块和引入可变形卷积重构 C2f 模块对 YOLOv8n 模型进行了改进优化,并引入损失函数 WIOUv2 增强了模型的泛化能力,提出了一种基于改进的 YOLOv8n 的变电设备红外图像实例分割算法,实验结果表明:上下文引导的特征增强下采样块能有效利用上下文信息,建立局部特征和上下文特征之间的联系,加强模型对复杂场景的理解能力,同时结合全局信息增强设备特征;不规则特征提取模块所提取特征图的感受野能随变电设备的不规则外形灵活调整,提高了模型对变电设备特征的有效提取能力,同时抑制了背景噪声的干扰;损失函数 WIOUv2 替换 CIOU 后,模型的分类性能和泛化性有所提升,提高了对困难样本的分割能力。本文所提出的改进方法提升了对红外图像中变电设备实例分割的准确度,改善了分割效果,同时 FPS 满足实时性要求,能在实际应用中发挥作用。

## 参考文献

- [1] 高玮,李轩,李灿. 变电设备在线监测技术在智能变电站的应用研究[J]. 光源与照明, 2023(6):159-161.
- [2] 郑文杰,杨祎,乔木,等. 基于改进 YOLO 和 Resnet 的变电设备热缺陷识别及诊断方法[J]. 重庆理工大学学报(自然科学), 2023, 37(9):261-269.
- [3] 张瑞强,徐贵,经权,等. 基于红外热成像检测技术的变电设备异常发热故障检测[J]. 制造业自动化, 2022, 44(9):171-174.
- [4] 崔克彬,李宝树,徐雪涛,等. 基于图像增强技术的电气设备热故障自动诊断与定位[J]. 红外技术, 2014, 36(2):162-167.
- [5] 侍海将. 人工智能技术在电力设备运维检修中的应用[J]. 电子技术, 2023, 52(4): 352-353.
- [6] 张志超,左雷鹏,邹捷,等. 基于多模态图像信息的变电设备红外分割方法[J]. 红外技术, 2023, 45(11): 1198-1206.
- [7] 张秋铭,李云红,罗雪敏,等. 改进 Chan-Vese 模型的电力设备红外图像分割算法[J]. 红外技术, 2023, 45(2): 129-136.
- [8] LI B, WANG T, ZHAI Y J, et al. RFINet: RGB-thermal feature interactive enhancement network for semantic segmentation of insulator in backlight scenes[J]. Measurement, 2022, 205: 112177.
- [9] WANG F T, GUO Y, LI C L, et al. Electrical thermal image semantic segmentation: Large-scale dataset and baseline [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71:1-13.
- [10] ZHAO Z B, FENG S, ZHAI Y J, et al. Infrared thermal image instance segmentation method for power substation equipment based on visual feature reasoning[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72:1-13.
- [11] 齐冬莲,韩译锋,周自强,等. 基于视频图像的输变电设备外部缺陷检测技术及其应用现状[J]. 电子与信息学报, 2022, 44(11):3709-3720.
- [12] KANG M, TING C M, TING F F, et al. ASF-YOLO: A novel YOLO model with attentional scale sequence fusion for cell instance segmentation[J]. Image and Vision Computing, 2024, 147: 105057.
- [13] BAI R H, WANG M K, ZHANG Z P, et al. Automated construction site monitoring based on improved YOLOv8-SEG instance segmentation algorithm [J]. IEEE Access, 2023, 11: 139082-139096.
- [14] 王译崧,华杭波,孔明,等. Rep-YOLOv8 车辆行人检测分割算法[J]. 现代电子技术, 2024, 47(9):143-149.
- [15] 方涛涛,王池社,王洁,等. 基于 YOLOv8n 的探地雷达图像管线定位方法[J]. 国外电子测量技术, 2023, 42(11):170-177.
- [16] 杨海燕,王凤随,张兴旺. 基于上下文增强和特征融合的目标检测算法[J/OL]. 重庆工商大学学报(自然科学版):1-8[2024-04-11]. <http://kns.cnki.net/kcms/detail/50.1155.N.20231218.1736.006.html>.
- [17] WU T Y, TANG S, ZHANG R, et al. Cgnet: A light-weight context guided network for semantic segmentation [J]. IEEE Transactions on Image Processing, 2020, 30: 1169-1179.
- [18] Narayanan M. SENetV2: Aggregated dense layer for channelwise and global representations [J]. ArXiv preprint arXiv:2311-10807, 2023.
- [19] DAI J F, QI H Z, XIONG Y W, et al. Deformable



- convolutional networks [J]. ArXiv preprint arXiv: 764-773, 2017.
- [20] ZHU X Z, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
- [21] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [22] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [J]. ArXiv preprint arXiv: 2301-10051, 2023.
- [23] BOLYA D, ZHOU C, XIAO F Y, et al. YOLACT: Real-time instance segmentation[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019.
- [24] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, 2017.
- [25] WANG C Y, BOCHKOVSKIY A, LIAO H Y. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. Vancouver: IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
- [26] WANG C Y, YEI I H, LIAO H Y. YOLOv9: Learning what you want to learn using programmable gradient information[J]. ArXiv preprint arXiv: 2402-13616, 2024.

## 作者简介

**李冰**, 副教授, 主要研究方向为模式识别与电力视觉。

E-mail: li\_bing\_hb@126.com

**杜喜英**, 硕士研究生, 主要研究方向为模式识别与计算机视觉。

E-mail: dxy3185311962@163.com

**王玉莹**, 硕士研究生, 主要研究方向为模式识别与计算机视觉。

E-mail: wyyemail@ncepu.edu.cn

**翟永杰**(通信作者), 教授, 主要研究方向为模式识别与电力视觉。

E-mail: zhaiyongjie@ncepu.edu.cn