

文章编号:1671-6833(2026)01-0033-08

基于改进 YOLOv8 的遥感影像变电站目标识别

刘润杰^{1,2}, 许慧娜^{1,2}, 胡宇^{1,2}, 王一¹, 谢国钧^{1,3}

(1. 郑州大学 国家超级计算郑州中心, 河南 郑州 450001; 2. 郑州大学 计算机与人工智能学院, 河南 郑州 450001; 3. 中科星图金能(南京)科技有限公司, 江苏 南京 211100)

摘要: 针对现有研究多集中于变电站局部结构检测而缺乏大区域快速发现与动态监测的问题, 通过高分辨率卫星影像实现变电站的高效识别, 提升电网安全隐患排查能力。首先构建了基于高分辨率光学卫星影像的变电站目标检测样本库; 随后提出改进的 YOLOv8 算法, 在骨干网络中嵌入 SimAM 轻量级注意力模块以增强细部特征聚焦能力, 并将颈部结构替换为 Efficient-RepGFPN, 结合 DySample 动态上采样模块设计新型颈部结构 GDFPN, 以解决多层级特征语义错位问题。实验结果表明: 改进方法优于主流检测算法, $mAP75$ 和 $mAP50-95$ 分别提升至 96.8% 和 87.1%, 验证了其在变电站检测任务中的优越性。所提出的改进 YOLOv8 方法可有效支持大区域变电站的快速发现与动态监测, 为电网安全管理提供了可靠的技术支撑。

关键词: YOLOv8; 遥感影像; 目标检测; 变电站; 注意力机制

中图分类号: TP391; TP751; TM63

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2025.04.022

变电站是电力系统的重要基础设施, 在电能输送和配送环节发挥了重要作用。遥感影像具有独特的鸟瞰视角和高分辨率特点, 覆盖范围广、细节丰富, 与深度学习和计算机视觉技术相结合能够准确揭示电力设施时空变化信息, 为电网从业人员提供及时、准确的多方位统计和分析结果, 为电力安全生产和智能调度提供更全面、准确、及时的信息支持。

随着计算机视觉和深度学习技术的快速发展, 目标检测领域涌现出许多高效方法。以 Faster R-CNN^[1] 为代表的双阶段目标检测算法通过区域建议网络 (region proposal network, RPN) 生成初步的目标位置, 然后分别修正目标的精细包围框和类别。这种策略精度较高, 但其两阶段特性也导致处理速度较慢。YOLO^[2-6]、SSD^[7] 等目标检测算法开创了一阶段目标检测的新纪元。其中, YOLO 将目标检测任务简化为一个回归问题, 将图像简单地按照网格划分并直接回归得到目标框坐标, 显著提高了检测速度, 被广泛应用于电力^[8]、农业^[9]、城市建设^[10] 等领域。

尽管目标检测算法已经比较成熟, 但这些算法大多在自然图像上训练和评测, 而卫星影像因其特有的高分辨率和广阔覆盖范围的特点, 为目标检测

任务带来了独特挑战。变电站在卫星影像中包含大量细碎特征, 边界界定较为困难, 常与周边建筑物、耕地等具有相似纹理特征的地物混淆, 在视觉上难以区分。此外, 光照条件、大气效应以及地形特征对目标检测准确性也会造成影响^[11-12]。

对于变电站检测, 已有研究主要集中在利用光学图像或红外图像对变电站中特定结构或组件的检测, 包括变压器、断路器和散热器等^[13-14]。Peng 等^[15] 提出了一种基于深度迁移学习的算法, 用于验证 Google Earth 图像中的变电站位置信息。该算法的核心是利用深度迁移学习模型来自动检测和验证数据库中记录的变电站位置信息是否准确, 然而它并不是专门针对变电站检测任务而设计的。Pérez-Hernández 等^[16] 提出的 DetDSCI 算法能够有效检测遥感影像中的基础设施目标, 包括机场、变电站等典型地物。该算法首先通过一个分类模型判识影像分辨率, 然后选择合适的模型进行设施检测, 但该算法并未针对变电站检测任务专门优化深度学习模型。Mesvari 等^[17] 提出了一种基于深度卷积神经网络的算法, 用于在高分辨率卫星图像中分割变电站, 实现了对变电站相关区域的高精度分割, 但无法提供变

收稿日期: 2025-04-08; 修订日期: 2025-05-10

基金项目: 河南省重大科技专项(221100210600); 河南省高等学校重点科研项目(23A140014)

作者简介: 刘润杰(1969—), 男, 河南安阳人, 郑州大学副教授, 博士, 主要从事遥感图像分析、并行计算研究, E-mail: ierj-liu@zzu.edu.cn。

引用本文: 刘润杰, 许慧娜, 胡宇, 等. 基于改进 YOLOv8 的遥感影像变电站目标识别[J]. 郑州大学学报(工学版), 2026, 47(1): 33-40. (LIU R J, XU H N, HU Y, et al. Remote sensing image substation target recognition based on improved YOLOv8[J]. Journal of Zhengzhou University (Engineering Science), 2026, 47(1): 33-40.)

电站的位置信息,不符合对变电站快速发现与动态监测的需求。

针对现有技术的局限性,本文基于 YOLOv8 进行改进,提出遥感影像变电站检测算法,主要贡献包括:构建变电站目标检测样本库,包含 10 192 张 2 m 分辨率卫星影像,覆盖城市、郊区、农田等各类场景;在骨干网络嵌入 SimAM 注意力模块,通过能量函数优化特征通道权重,提升对变电站细部特征的聚焦能力;设计梯度感知特征金字塔颈部 GDFPN,结合 Efficient-RepGFPN 的跨层连接优势与 DySample 的动态上采样特性,缓解特征金字塔中的语义鸿沟问题。

1 YOLOv8 遥感影像变电站目标识别存在的问题

本文实验数据为高分辨率光学遥感卫星影像,通常包含更复杂的背景和环境信息,且可能受到云层遮挡和传感器噪声等影响,进一步降低了图像质量,导致目标与背景之间的对比度较低,变电站在遥感影像中的细节信息更难捕捉。此外,遥感影像普遍存在同谱异物和同物异谱问题,中小尺度下变电站因其特有的矩形结构,在图像中与众多类似的人造结构,如建筑物、耕地上的灌溉网格、操场等在几何形态上高度相似,容易造成混淆或识别位置不准确。

在变电站检测数据集上训练并测试 YOLOv8 的实验结果表明,耕地、低矮平房容易被错检为变电站,部分测试图像上出现明显漏检,典型的误检问题包括:①耕地与变电站双向误判造成的错检;②耕地与变电站边界混淆造成的定位不准确;③将低矮平房错误分类为变电站的误检;④在广阔耕地区域中无法准确识别变电站目标的问题。

因此,需要针对各种误检问题对该模型进行改进优化,更好地提取遥感影像中变电站的细节特征,提高遥感影像中变电站检测的准确率。

2 YOLOv8 改进方案

对比 YOLOv8 下 s、m、l 这 3 组不同大小模型的检测精度和效率后发现,m 组取得了最优的效率和精度平衡,后续改进均基于该模型进行。改进后的遥感影像变电站检测网络详细结构如图 1 所示,在网络骨干不同层级处引入了 SimAM 注意力模块引导模型聚焦于变电站不同尺度上细部结构,提出基于动态上采样的颈部结构 GDFPN,通过多级跨层设计实现了不同层级间的信息交换。

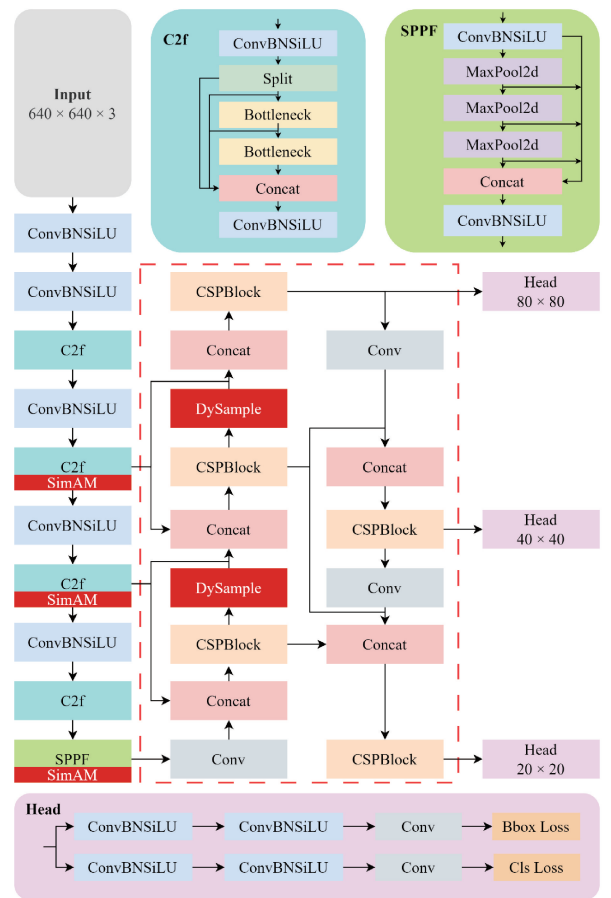


图 1 改进后的遥感影像变电站检测网络详细结构

Figure 1 Detailed architecture of the improved remote sensing image substation detection network

2.1 轻量级注意力机制 SimAM

SimAM^[18]是一种基于三维权重分配的注意力机制,其核心思想是通过分析特征图的局部自相似性来生成注意力权重。该方法的理论基础为在视觉处理过程中,那些能够产生显著空间抑制效果的神经元通常具有更高的重要性。

为了高效地构建三维注意力权重,SimAM 采用能量函数来评估单个神经元的重要性。具体而言,该方法通过测量目标神经元与其周围神经元的线性可分性来量化其显著性。基于这一原理,SimAM 定义了如下能量函数:

$$E_t(w_t, b_t, \mathbf{y}, u_t) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_t u_i + b_t))^2 + (1 - (w_t t + b_t))^2 + \lambda w_t^2. \quad (1)$$

式中: w_t 和 b_t 分别为权重和偏置; t 为目标神经元; u_t 为同一通道中除 t 外的其他神经元; M 为通道中的神经元总数; λ 为正则化项。

能量越低,神经元 t 与周围神经元的区别越大,其重要性越高。根据注意力机制的定义,最终的特征增强可以描述为

$$\hat{X} = X \odot \text{sigmoid}\left(\frac{1}{E}\right). \quad (2)$$

式中: E 为包含所有最小能量的集合。

针对遥感影像变电站图像细节复杂且易与周围相似地物混淆的问题,在 YOLOv8 骨干网络的 3 个位置分别加入了 SimAM 注意力机制。具体而言,第 1 次引入位于骨干网络的第 5 层后,旨在利用注意力机制对经过下采样的特征图进行精细化处理,增强模型对小目标的检测能力;第 2 次引入位于第 7 层后,在更高层次的特征抽象中进一步提升模型对大目标的检测性能,同时保持对小目标检测的敏感性;第 3 次引入位于第 10 层后,经过 SPPF (spatial pyramid pooling fast) 处理之后,在最粗糙的特征级别上进一步提炼模型的注意力分配,确保在全局特征上也能实现对目标物体的精确识别。

2.2 动态上采样颈部 GDFPN

YOLOv8 的颈部采用 PAFAN 结构,旨在融合从主干网络中提取的浅层与深层特征,并输出 3 组多尺度特征。相比 PAFPN,GFPN 通过多级跨层设计实现了不同层级间的信息交换^[19],但存在计算开销大的问题。Efficient-RepGFPN^[20] 对 GFPN 中的 Queen-Fusion 进行了加速,并使用 CSPNet^[21] 替换原始基于 3×3 卷积的特征融合,结合 ELAN 结构,兼顾效率与精度。

为进一步提升遥感影像变电站检测效果,基于 Efficient-RepGFPN,引入基于点采样的动态上采样技术 DySample^[22],提出了 GDFPN (generalized dynamic feature pyramid network) 颈部结构,如图 2 所示。

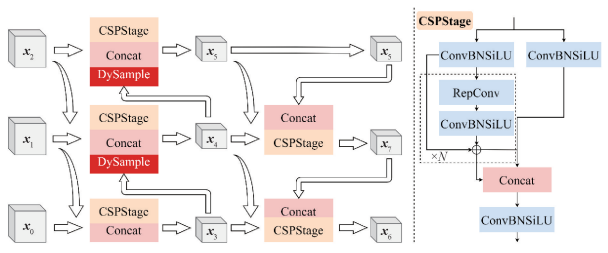


图 2 GDFPN 网络结构图

Figure 2 GDFPN network structure

设 x_0, x_1, x_2 为 GDFPN 输入,在计算 x_4 时,动态上采样操作将 x_3 特征图尺寸放大,以便与 x_1 和 x_2 经过卷积后的特征图进行拼接;在计算 x_5 时,动态上采样操作再次被使用,将 x_4 的特征图尺寸放大,与 x_2 进行拼接。通过这种方式,模型能够在不同的尺度上捕获和融合特征信息。最后,GDFPN 输出 3 个不同尺度的特征图 x_5, x_6, x_7 。颈部的输出特征图被送入网络检测头用于目标位置和类别预测。

GDFPN 通过动态上采样技术构建一个多尺度的特征金字塔,允许模型在不同的层次上捕捉图像的细节和上下文信息;此外,利用 DySample 技术根据目标的大小和形状动态调整特征图的上采样过程,使得模型能够更加关注于目标区域的特征增强;颈部结构包含多个阶段的特征融合,有助于模型在保持细节信息的同时,也能够获取更加丰富的上下文信息。

2.3 基于点采样的动态上采样器

DySample 基于点采样,通过以下几个关键步骤实现上采样。

(1) 采样点生成。给定一个上采样尺度因子 s 和一个大小为 $C \times H \times W$ 的特征图 X , 使用一个输入和输出通道分别为 C 和 $2s^2$ 的线性层,生成大小为 $2s^2 \times H \times W$ 的偏移量 O 。然后,通过像素洗牌 (pixel shuffling)^[23] 重塑为 $2 \times sH \times sW$ 。采样集 S 是偏移量 O 和原始采样网格 g 的和,用于从连续的特征图中重新采样。

$$O = \text{linear}(X); \quad (3)$$

$$S = g + O. \quad (4)$$

(2) 偏移量作用范围调整。为了减少采样点之间的重叠,分别引入了静态作用范围因子 O_{sta} 和动态作用范围因子 O_{dyn} 调节偏移量的大小,结构如图 3 所示。

$$O_{sta} = 0.25 \text{linear}(X); \quad (5)$$

$$O_{dyn} = 0.5 \text{sigmoid}(\text{linear}_1(X)) \cdot \text{linear}_2(X). \quad (6)$$

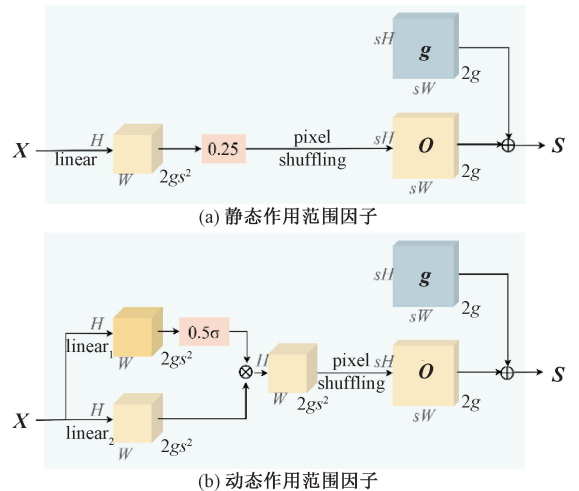


图 3 DySample 中的采样点发生器

Figure 3 Sampling point generator in DySample

(3) 连续插值。给定大小为 $C \times H_1 \times W_1$ 特征图,以及一个大小为 $2 \times H_2 \times W_2$ 的采样集 S , 其中第一维的数字 2 对应每个采样点的 x 和 y 坐标。grid_sample 函数使用双线性插值将输入特征图插值到一个连续的空间中,生成大小为 $C \times H_2 \times W_2$ 的 X' , 如图 4 所示,这意味着在原始离散的像素点之

间生成了连续的像素值。计算公式如下所示:

$$X' = \text{grid_sample}(X, S)。 \quad (7)$$

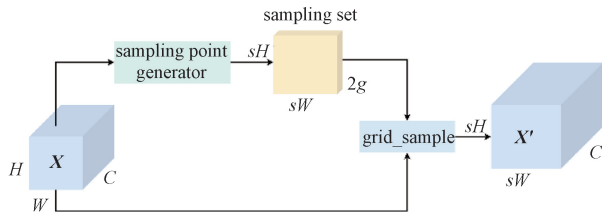


图4 基于点采样的动态上采样

Figure 4 Dynamic upsampling based on point sampling

3 实验结果与分析

3.1 数据集介绍

实验数据为高分辨率遥感卫星影像,空间分辨率为2 m,标注工具选择 LabelImg,标签数据类型保存为 VOC 数据集数据类型后转化为 YOLO 格式进行训练。数据集中变电站范围分布广泛,覆盖了从城市到郊区以及偏远地区的各种地形和环境,比如密集城区、房屋周围、农田区域和交通基础设施附近等。训练集和测试集中图像为 $2\ 753 \times 1\ 647$ 像素,共 10 192 张,变电站目标共 10 448 个,其中小尺寸目标(小于 500 像素)数量为 2 641 个,中尺寸目标(500~1 000 像素)数量为 6 858 个,大尺寸目标(超过 1 000 像素)有 949 个。训练集数量与测试集数量比例为 9:1,所有图像进入网络前按照等比缩放原则重采样到 960×960 像素。

3.2 模型训练细节

本文在 16 GB 的 NVIDIA Tesla P100 GPU 上进行训练和测试,初始加载 YOLOv8m 预训练权重,批大小为 16,训练轮数为 100。使用平均精度(mAP)、精确度($Precision$)和召回率($Recall$)作为性能评估指标。由于变电站数据集是单分类目标检测,使用更严格的 $mAP75$ 作为评估指标,同时采用 0.5~0.9 倍 IoU 阈值范围模型的平均精度 $mAP50-90$ 评价变电站位置准确性,计算公式如下所示:

$$Precision = \frac{TP}{TP + FP}; \quad (8)$$

表1 SimAM 注意力机制与动态上采样颈部 GDFPN 消融实验

Table 1 SimAM attention mechanism and dynamic upsampling neck GDFPN ablation experiment

模型	$Precision/\%$	$Recall/\%$	$mAP75/\%$	$mAP50-95/\%$	帧率/ (帧·s ⁻¹)	计算量/ GFLOPs	参数量 /10 ⁶
YOLOv8m	98.7	96.9	94.5	85.4	72.7	78.7	25.8
YOLOv8m-SimAM	98.7	97.6	95.9	86.5	72.6	78.7	25.8
YOLOv8m-RepGFPN	97.9	97.5	95.7	85.9	63.8	81.5	29.1
YOLOv8m-DySample	99.1	97.8	95.3	86.0	70.6	78.7	25.9
YOLOv8m-GDFPN	99.0	97.9	96.0	86.2	65.5	81.6	29.1

$$Recall = \frac{TP}{TP + FN}; \quad (9)$$

$$mAP = AP = \sum_{k=1}^n P(k) \cdot \Delta R(k)。 \quad (10)$$

式中: TP 为模型正确检测到的目标数量; FP 为模型错误地将背景检测为目标的数量; FN 为模型未能检测到的目标数量; $P(k)$ 为在第 k 个召回率水平下的精确度; $\Delta R(k)$ 为在第 k 个召回率阈值和第 $k+1$ 个召回率阈值之间的召回率变化量; n 为评估中使用的召回率水平的数量。

3.3 消融实验

为验证注意力机制以及高效特征融合策略与动态上采样 DySample 技术的有效性,进行消融实验,测试结果如表 1 所示。利用 Efficient-RepGFPN 的多尺度特征融合策略的颈部替换 YOLOv8 基于 PAFAN 的颈部后,召回率提升了 0.6 百分点, $mAP75$ 提升了 1.2 百分点, $mAP50-95$ 提升了 0.4 百分点。通过动态采样技术替换 YOLOv8 中使用的传统上采样操作后,精确度提升了 0.4 百分点,召回率提升了 0.9 百分点, $mAP75$ 提升了 0.8 百分点, $mAP50-95$ 提升了 0.6 百分点。而使用基于动态采样技术和高效特征融合策略改进的 GDFPN 结构作为 YOLOv8 的颈部后,召回率提升了 1.0 百分点, $mAP75$ 提升了 1.5 百分点, $mAP50-95$ 提升了 0.8 百分点。实验结果表明,使用动态上采样技术与高效特征融合策略改进后的网络提升了变电站检测精度,更加适合遥感场景下的变电站检测任务。

同时,为证明注意力机制在特征筛选中的重要作用,本文测试了几种常用的注意力机制在变电站检测场景下的效果,测试结果如表 2 所示。实验结果表明,加入 CBAM^[24] 注意力模块不能充分发挥注意力机制的优势,没有提高网络的特征提取能力; DA^[25]、LocalWindow^[26] 注意力机制对网络特征的提取效果有一定改善,但同时引入了额外的参数,降低了网络效率;而 SimAM 注意力机制对网络的改善最明显,能够充分筛选出变电站目标设备内的主要特征,同时不会增加网络的参数量。

表 2 不同注意力机制的对比实验结果

Table 2 Comparative experiment results of different attention mechanisms

注意力机制	$mAP75/\%$	参数增量/ 10^6
CBAM ^[24]	93.5	0.065
DA ^[25]	95.2	0.267
LocalWindow ^[26]	95.5	0.093
SimAM	95.9	0

3.4 实验结果

本文在变电站检测数据集上对比测试了多个目标检测算法,各方法检测精度和效率对比如表 3 所示。由表 3 可知,本文方法取得了最优检测精度,而且在推理速度和参数量之间实现了良好的平衡。各对比方法中,Faster R-CNN^[1]和 Mask R-CNN^[27]作

为经典的两阶段检测模型,在检测精度上有一定的保证,但误检现象明显,且在速度和参数量方面相对较差。YOLOv5m 和 YOLOv6m 平均精度明显低于本文方法。DetDSCI^[16]为两阶段方法,将设施划分为小型类别,实验中直接采用 CI-SS_Det_stable 模型,效果明显不如本文方法。YOLOv8s 速度最快,参数量最小,但是精度较低。YOLOv8l 相较于 YOLOv8m,参数量增加了 50%,但是精度没有明显提升。虽然 YOLO11m 在评估指标上表现良好,但整体上精度与效率均略逊于本文方法。RTMDet^[28]和 DINO^[29]两个最新的非 YOLO 系列模型虽具有较强的竞争力,但它们的推理速度和参数量都逊于本文方法。图 5 展示了部分典型方法在变电站检测数据上的表现,可以看出,本文方法在郊区、城区、农田等不

表 3 不同模型实验结果对比

Table 3 Experimental result comparison of different models

模型	Precision/ $\%$	Recall/ $\%$	$mAP75/\%$	$mAP50-95/\%$	帧率/(帧·s ⁻¹)	计算量/GFLOPs	参数量/ 10^6
Faster R-CNN (resnet-50) ^[1]	97.2	95.8	91.0	80.5	10.2	112.4	40.5
Mask R-CNN (resnet-50) ^[27]	97.5	96.0	91.5	81.0	8.5	120.1	42.0
YOLOv5m	98.3	96.4	93.6	84.5	76.3	64.0	25.0
YOLOv6m	98.9	97.0	94.0	85.1	52.8	161.1	51.9
DetDSCI ^[16]	97.5	96.1	91.7	83.6	58.6	164.0	51.4
YOLOv8s	98.3	96.7	92.6	82.4	143.5	28.4	11.1
YOLOv8m	98.7	96.9	94.5	85.4	72.7	78.7	25.8
YOLOv8l	98.4	97.4	94.5	85.3	43.2	164.8	43.6
YOLO11m	98.5	97.6	95.2	86.1	58.3	145.6	38.2
RTMDet ^[28]	98.2	97.0	94.0	85.0	30.6	120.0	42.5
DINO ^[29]	98.6	97.3	94.7	86.5	25.0	130.2	48.0
本文方法	98.9	98.0	96.8	87.1	64.9	81.6	29.1



图 5 不同模型检测结果可视化对比

Figure 5 Visual comparison of detection results by different models

同地物背景条件下均有效改善了错检和漏检问题,定位准确性高,不易受周边地物干扰。

为了更直观地展示网络效果,单独对比了YOLOv8m和本文改进方案,如图6所示,YOLOv8m易将耕地、低矮平房错检为变电站,部分场景下还存在明显漏检,本文方法则无此类问题。

为进一步探究本文改进的有效性,以热力图的形式可视化了模型的注意力分布,如图7所示。从热力图来看,SimAM能够强化与变电站相关的特征并抑制无关背景,使网络更能注意到变电站区域内的细节信息,减少模型对非目标物体的响应,从而与周围地物,如建筑物、耕地、操场等区分开来,降低误检率。GDFPN颈部结构利用多尺度特征融合策略

和动态上采样技术可以为后续的检测头提供更丰富的信息,进一步提升了检测精度。

实验结果表明,本文对YOLOv8m模型进行优化,在处理遥感影像中远距离、复杂背景下的变电设备方面,相较于原始模型展现出了更强的抗干扰性能和更高的精确度。这种提升主要体现在模型能够更有效地区分与变电站目标相似的干扰物,同时保持了高准确率的检测结果。此外,本文所采用的改进措施并未对原模型的主干网络和检测模块进行改动。因此,改进后的模型保持了与原始模型相同的广泛适用性。这种优化在不牺牲通用性的前提下,提高了特定场景下的目标检测能力。

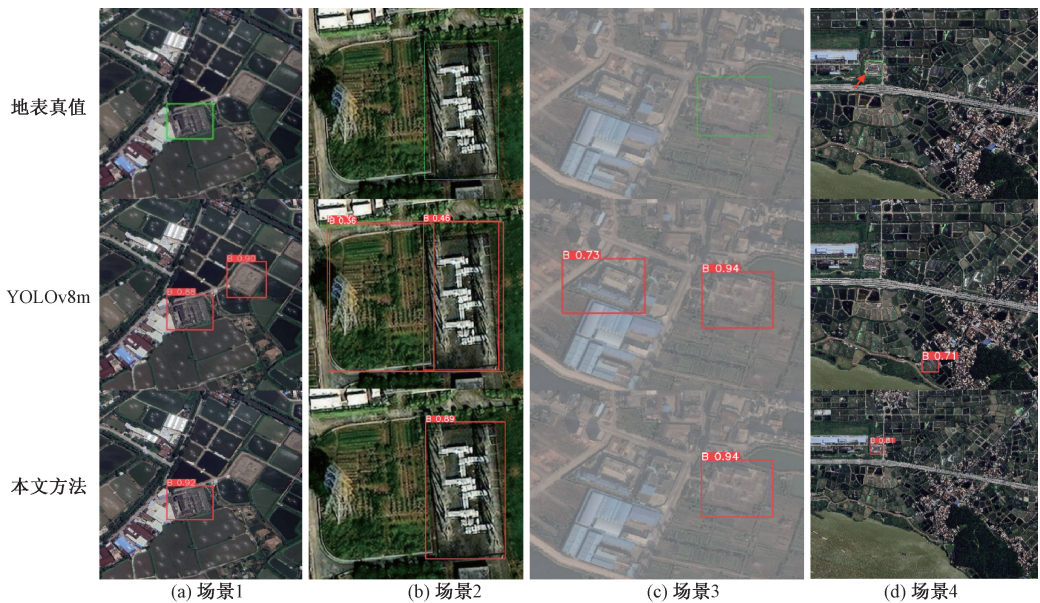


图6 误检情况对比

Figure 6 Comparison of false detection cases

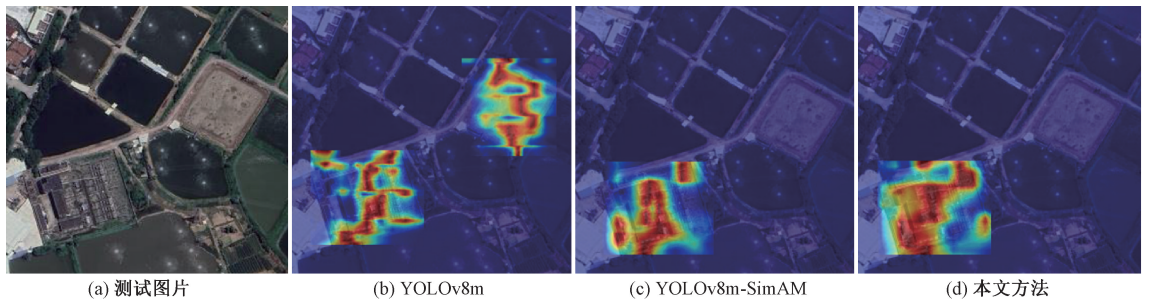


图7 热力图可视化

Figure 7 Heatmap visualization

4 结论

YOLOv8因其高效和实时的目标检测能力在自然图像目标检测中得到了广泛应用。本文为适应遥感场景下变电站检测任务,在其基础上引入SimAM

注意力机制,并提出了基于多尺度高效特征融合策略及动态上采样技术的改进颈部结构GDFPN,有效减少了误检情况。相比YOLOv8m,改进方法平均检测精度 mAP_{75} 由94.5%提升至96.8%,提高了2.3个百分点; mAP_{50-95} 由85.4%提升至87.1%,提高了

1.7百分点。改进方法能够生成准确的变电站边界且不易受周围相似地物干扰。此外,本文所提出的方法对于遥感影像大型设施识别,如机场、港口等,具有一定启发的意义。

(本文受到蓝天实验室无人机警务应用科研项目“基于互联网的无人机监测定位与反制系统”(LTJW-24-02-03)的支持。)

参考文献:

- [1] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [2] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 779-788.
- [3] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 6517-6525.
- [4] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08) [2025-03-29]. <https://arxiv.org/abs/1804.02767>.
- [5] BOCHKOVSKIY A, WANG C Y, LIAO H M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2025-03-29]. <https://arxiv.org/abs/2004.10934>.
- [6] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023: 7464-7475.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Computer Vision-ECCV 2016. Cham: Springer, 2016: 21-37.
- [8] SUN H B, SHEN Q C, KE H C, et al. Power transmission lines foreign object intrusion detection method for drone aerial images based on improved YOLOv8 network[J]. *Drones*, 2024, 8(8): 346.
- [9] SHI J Y, BAI Y H, ZHOU J, et al. Multi-crop navigation line extraction based on improved YOLO-v8 and threshold-DBSCAN under complex agricultural environments[J]. *Agriculture*, 2024, 14(1): 45.
- [10] XUE F C, CHENG Y N, SHEN Y F, et al. Urban flood inundation area detection using YOLOv8 model[J]. *Water Resources Management*, 2025, 39: 5443-5460.
- [11] 邝先验,程福军,吴翠琴,等. 基于改进YOLOv7-tiny的高效轻量遥感图像目标检测方法[J]. *电子测量与仪器学报*, 2024, 38(7): 22-33.
KUANG X Y, CHENG F J, WU C Q, et al. Efficient and lightweight target detection method for remote sensing images based on improved YOLOv7-tiny[J]. *Journal of Electronic Measurement and Instrument*, 2024, 38(7): 22-33.
- [12] 魏明军,王镛涵,刘亚志,等. 基于特征融合和混合注意力的小目标检测[J]. *郑州大学学报(工学版)*, 2024, 45(3): 72-79.
WEI M J, WANG M H, LIU Y Z, et al. Small object detection based on feature fusion and mixed attention[J]. *Journal of Zhengzhou University (Engineering Science)*, 2024, 45(3): 72-79.
- [13] 鲁玲,李明明,熊威,等. 基于改进的YOLOv8变电设备红外图像检测[J]. *红外技术*, 2025, 47(9): 1135-1141.
LU L, LI M L, XIONG W, et al. Infrared image detection of substation equipment based on improved YOLOv8[J]. *Infrared Technology*, 2025, 47(9): 1135-1141.
- [14] 李斌,李亚霖,朱新山,等. 基于注意力机制与特征平衡的变电站多目标检测[J]. *电网技术*, 2022, 46(6): 2122-2132.
LI B, LI Y L, ZHU X S, et al. Multi-target detection in substation scene based on attention mechanism and feature balance[J]. *Power System Technology*, 2022, 46(6): 2122-2132.
- [15] PENG J, LIU Y B, MA Y, et al. Deep transfer learning for power substation recognition with Google Earth[C]//2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP). Piscataway: IEEE, 2018: 54-57.
- [16] PÉREZ-HERNÁNDEZ F, RODRÍGUEZ-ORTEGA J, BENHAMMOU Y, et al. CI-dataset and DetDSCI methodology for detecting too small and too large critical infrastructures in satellite images: airports and electrical substations as case study[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 12149-12162.
- [17] MESVARI M, SHAH-HOSSEINI R. Segmentation of electrical substations using deep convolutional neural network[J]. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2023, 10(4): 495-500.
- [18] YANG L X, ZHANG R Y, LI L D, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks[C]//International Conference on Machine Learning. ICML: San Diego, 2021: 18-24.

- [19] JIANG Y Q, TAN Z Y, WANG J Y, et al. GiraffeDet: a heavy-neck paradigm for object detection [EB/OL]. (2022-02-09) [2025-03-29]. <https://arxiv.org/abs/2202.04256>.
- [20] XU X Z, JIANG Y Q, CHEN W H, et al. DAMO-YOLO: a report on real-time object detection design [EB/OL]. (2022-11-23) [2025-03-29]. <https://arxiv.org/abs/2211.15444>.
- [21] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSP-Net: a new backbone that can enhance learning capability of CNN [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2020: 390-391.
- [22] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample [C] // 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2023: 6004-6014.
- [23] SHI W Z, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 1874-1883.
- [24] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C] // Computer Vision-ECCV 2018. Cham: Springer, 2018: 3-19.
- [25] XIA Z F, PAN X R, SONG S J, et al. Vision transformer with deformable attention [C] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 4784-4793.
- [26] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [C] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2021: 9992-10002.
- [27] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C] // 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017: 2980-2988.
- [28] LYU C Q, ZHANG W W, HUANG H A, et al. RTM-Det: an empirical study of designing real-time object detectors [EB/OL]. (2022-12-14) [2025-03-29]. <https://arxiv.org/abs/2212.07784>.
- [29] ZHANG H, LI F, LIU S L, et al. DINO: DETR with improved denoising anchor boxes for end-to-end object detection [EB/OL]. (2022-07-11) [2025-03-29]. <https://arxiv.org/abs/2203.03605>.

Remote Sensing Image Substation Target Recognition Based on Improved YOLOv8

LIU Runjie^{1,2}, XU Huina^{1,2}, HU Yu^{1,2}, WANG Yi¹, XIE Guojun^{1,3}

(1. National Supercomputing Center in Zhengzhou, Zhengzhou University, Zhengzhou 450001, China; 2. School of Computer Science and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China; 3. Zhongke Xingtu Jinneng (Nanjing) Technology Co., Ltd., Nanjing 211100, China)

Abstract: Aiming at the limitation in existing studies focused on the detection of substation local structures, such as lacking methods for rapid discovery and dynamic monitoring over large areas, the capability of identifying potential safety hazards in power grids was enhanced through high-resolution satellite imagery. Firstly, a substation object detection dataset based on high-resolution optical satellite imagery was constructed. Subsequently, an improved YOLOv8 algorithm was proposed, embedding the SimAM lightweight attention module into the backbone network to enhance the ability to focus on detailed features, and replacing the neck with an Efficient-RepGFPN, combined with a DySample dynamic upsampling module to design a novel neck named GDFPN, addressing issues of multi-level feature semantic misalignment. Experimental results demonstrated that the improved method outperformed mainstream detection algorithms, with $mAP75$ and $mAP50-95$ increasing to 96.8% and 87.1%, respectively, confirming its superiority in substation detection tasks. The improved YOLOv8 approach proposed could effectively support the rapid discovery and dynamic monitoring of substations over large areas, providing reliable technical support for the safety management of power grids.

Keywords: YOLOv8; remote sensing image; object detection; substation; attention mechanism