

DOI: 10.19650/j.cnki.cjsi.J2413594

GAE-YOLO: 全局感知增强的输电线路 外破隐患目标检测方法*

刘 敏, 陈 明, 武明虎, 叶永钢

(湖北工业大学太阳能高效利用及储能运行控制湖北省重点实验室 武汉 430068)

摘 要: 超高压架空输电线路在电力系统中至关重要, 但常面临建筑施工、山火等外力因素引发的事故。这不仅损害了国家经济, 影响电网稳定性, 还对电力工作人员的人身安全造成威胁。基于深度学习的目标检测方法为检测外破隐患提供了新方案, 但现有方法往往依赖局部邻域信息执行采样操作, 限制了感知范围和表达能力。为解决这一问题, 提出了一种基于 YOLOv10 的实时全局感知增强方法 GAE-YOLO, 旨在提高超高压架空输电线路外破隐患目标的检测精度。针对传统方法中局部感知的局限, 设计了 2 个新的上下采样模块: 全局感知下采样模块 (GADM) 和全局感知上采样模块 (GAUM)。GADM 通过学习特征图的全局空间信息生成全局感知权重, 优化下采样过程的感知性能; GAUM 则通过利用深层特征图的通道信息生成全局感知权重, 动态增强采样点的隶属关系, 有效突出目标边界。为验证 GAE-YOLO 的有效性, 构建了一个针对超高压架空输电线路外破隐患的大规模数据集, 并在该数据集上取得了 93.05% 的平均精度均值 (mAP), 相较于基线模型 mAP 提升了 5.13%。实验结果表明, GAE-YOLO 能够显著提高外破隐患目标的检测精度, 具有重要的应用价值, 为电网安全运行提供了新的技术支持。

关键词: 目标检测; 全局感知; 下采样; 上采样; 外破隐患; 输电线路

中图分类号: TP391.4 TM773 TH183.3 **文献标识码:** A **国家标准学科分类代码:** 520.20

GAE-YOLO: Global awareness enhanced method for detecting external force damage in power transmission lines

Liu Min, Chen Ming, Wu Minghu, Ye Yonggang

(Hubei Key Laboratory for High-efficiency Utilization of Solar Energy and Operation Control of Energy Storage System, Hubei University of Technology, Wuhan 430068, China)

Abstract: Ultra-high-voltage overhead transmission lines are crucial in power systems. But, they often face accidents triggered by external factors, such as construction activities and wildfires. These incidents not only damage the national economy and affect grid stability, but also pose a threat to the safety of power workers. Deep learning-based object detection methods offer a novel solution for detecting external force damage objects. However, existing methods often rely on local neighborhood information for sampling operations, which limits their perceptual range and expressive capabilities. To address this issue, a real-time global awareness-enhanced method, GAE-YOLO, based on YOLOv10, is proposed to improve the detection accuracy of external force damage objects in ultra-high-voltage overhead transmission lines. To overcome the limitations of local perception in traditional methods, two novel upsampling and downsampling modules are designed, including the global awareness downsampling module (GADM) and the global awareness upsampling module (GAUM). GADM enhances perceptual performance by learning global spatial information from the feature map and generating global perception weights to optimize the downsampling process. GAUM dynamically enhances the membership relationship of sampling points by utilizing channel information from deep feature maps, effectively highlighting object boundaries. To evaluate the effectiveness of GAE-YOLO, a large-scale dataset for detecting external force damages in ultra-high-voltage overhead transmission lines is constructed. The model achieves mAP of 93.05%, a mAP 5.13% improvement over the baseline model. Experimental results show that

收稿日期: 2024-12-16 Received Date: 2024-12-16

* 基金项目: 湖北省自然科学基金 (2022CFA007)、湖北省中央引导地方科技发展资金 (2023EGA027) 项目资助

GAE-YOLO significantly improves the detection accuracy of external-force damage objects, offering substantial application value and providing new technical support for the safe operation of power grids.

Keywords: object detection; global awareness; downsampling; upsampling; external force damage; transmission lines

0 引言

超高压架空输电线路是国家电网和全球能源互联的重要组成部分,常常面临由外力因素引起的各种损害,如建筑施工和山火,尤其是在偏远地区^[1]。这些外破隐患对人员安全和电网运行构成了重大风险。随着计算机视觉技术的快速发展,基于深度学习的目标检测技术为超高压架空输电线路场景中外破隐患目标检测提供了一种新的方法,及时发现这些隐患并排除,对电网安全运行具有重要意义^[2]。实际的输电线路场景存在复杂的背景、不同的照明条件和不同的目标姿态等因素,给外破隐患目标检测带来了巨大的挑战^[3]。

现有深度学习目标检测算法大多利用特征金字塔网络(feature pyramid networks, FPN)^[4]和路径聚合网络(path aggregation networks, PANet)^[5]来捕捉和融合多尺度特征信息,提升模型在面对不同大小的目标时的检测能力。为了增强不同层次特征信息的传播,FPN和PANet通常分别采用池化(或步幅卷积)和插值算法进行特征图的下采样和上采样,从而构建特征金字塔。然而,这些下采样方法是在预定义的卷积核局部窗口内进行的,这导致特征图中的细节信息丢失,从而限制了网络的视觉理解能力^[6]。与此同时,插值算法通过利用像素间的空间距离来引导上采样过程,以放大特征图。然而,这会导致物体边界模糊,进而给确定场景中破隐患目标的位置带来挑战^[7]。

为了解决这些问题,研究人员提出了各种改进方法。对于下采样,混合池化(mixed pooling)^[8]将最大池化和平均池化结合,表现出比单独使用更好的性能。然而,混合池化仍然使用基于规则的采样核在局部区域内实现特征下采样。因此,一些先进的下采样方法逐渐引入了可学习的权重,实现自适应下采样操作。例如,基于反向双边滤波器思想的细节保持池化(detail-preserving pooling, DPP)^[9]专注于重要的结构细节。基于局部重要性的池化(local importance-based pooling, LIP)^[10]使用加权平均池化方法来聚合特征进行下采样。然而,这些方法仍然在小范围的邻域中保留图像细节,这会增强局部噪声的干扰。

至于上采样,也有从不可学习方法向可学习方法转变的趋势。SegNet^[11]在语义分割中采用最大池化索引来保持边缘信息,但这可能引入噪声,并扰乱平滑区域的语义一致性。像素洗牌(PixelShuffle)^[12]将通道中的像素

值重新排列到空间方向上。与前者不同,一些上采样方法在上采样过程中引入了可学习的参数来放大特征图。例如,反卷积(deconvolution)^[13]通过将单个输入映射到多个输出的方法来进行特征上采样。CARAFE^[14](content-aware reassembly of features)通过加权组合在预定义区域内重新组合特征。总体而言,这些可学习的上采样方法有助于恢复物体的边界,但需要学习大量的参数作为权衡。

上述方法主要依赖局部邻域信息执行采样操作,或者需要进行昂贵的计算以执行自适应采样操作。这些方法限制了采样器的感受野,从而限制了其表达能力和性能。在实际场景中,由于环境的复杂性和变化性,需要全局信息来准确捕捉和分析目标。然而,学习全局特征意味着需要处理更多的数据,并进行更复杂的计算,这增加了系统的成本和复杂度。因此,必须寻找更有效的方法,在保持性能的同时处理和利用全局信息。

针对这些局限性,本研究提出了一种基于YOLOv10^[15](you only look once version 10)的实时全局感知方法(global awareness enhanced YOLO, GAE-YOLO),用于输电线路中的外破隐患目标检测,该方法的特色为:

1) 创新地设计了一种全局感知下采样模块(global awareness downsampling module, GADM),通过学习特征图的全局空间信息,生成全局感知权重,优化模型在下采样过程中的感知性能。

2) 创新地设计了一种全局感知上采样模块(global awareness upsampling module, GAUM),该模块通过利用深层特征图通道信息生成全局感知权重,动态增强采样点的隶属关系。有助于突出采样点的层次关系,尤其是目标的边界。

3) 构建了一个针对超高压架空输电线路的外破隐患目标大规模数据集。通过将GADM和GAUM模块集成到YOLOv10检测框架中,在该数据集上实现了93.05%的平均精度均值(mean average precision, mAP)。

1 输电线路外破隐患目标检测相关工作

依据DL/T741—2019《架空输电线路运行规程》^[16],对外力破坏易发区应开展全面的外破隐患排查、制定隐患巡查及管控计划,加装视频(或图像)监控装置等方式,实行外力破坏隐患动态管控。随着计算机视觉和传感器技术的进步,基于深度学习的目标检测不仅能够降

低电力巡检人力成本,还能提供更稳定、高效的巡检模式。

现阶段基于深度学习的外破隐患目标检测方法通常可以分为两大类:二阶段检测器^[17-19]和一阶段检测器^[20-27]。针对二阶段检测器,魏贤哲等^[17]提出了一种基于 Mask R-CNN (region-based convolutional neural network) 算法的目标检测,使用特征映射的迁移学习方法,将边界框标注代替部分掩模标注;Leng 等^[18]开发了一种针对架空线路的防外破系统,提出了一种基于改进 R-FCN(region-based fully convolutional networks)的算法;Qu 等^[19]通过将 E-OHEM(enhanced online hard example mining)集成到 Faster R-CNN 中,减少了检测的误检率。然而以上学者也提到,虽然二阶段目标检测器精度较高,但处理速度较慢,因此难以推广到实际工程应用中,而一阶段目标检测算法在实时性和准确性方面均取得较好的表现。邢志鹏等^[22]在 YOLOv5s 模型 Head 引入 SimAM (simple, parameter-free attention module) 以加强隐患目标在提取特征中的权重,使用 Wise-IoU (intersection over union) 代替模型原始的 Complete-IoU (CIoU) 以提高预测边界框的回归预测精度。Li 等^[23]针对嵌入式设备中的检测速度和功耗问题,提出了一种基于 YOLOv5s 的实时危险检测方法,更高效地检测外破隐患目标。最近,Li 等^[24]改进了 YOLOv7-tiny,以增强对小物体的检测能力,同时减少了计算和存储需求,使其适用于低功耗设备。王彦海等^[25]提出一种改进 YOLOv7 的机械外破隐患目标检测算法,较原始算法在检测精度具有明显优势。

尽管以上方法在检测输电线路上的外破隐患目标均已表现出较好的效果,但在复杂环境中准确检测多尺度物体仍然是一个挑战。本研究旨在通过改进上采样和下采样技术,提升多物体场景中的检测精度和鲁棒性,以应对这些挑战。

2 方法

为了实现外破隐患目标检测在输电线路场景中更高的精度,创新性地提出了 GADM 和 GAUM,分别针对下采样和上采样任务进行优化,并在 YOLOv10-L 框架的基础上实施了一系列改进。

2.1 GAE-YOLO 整体网络结构

YOLOv10 系由清华大学团队提出一种实时端到端的目标检测器,根据用户对于性能及资源约束的不同需求,提供了涵盖 N、S、M、B、L、X 等多个模型规模,进而适配多样化的应用场景。在技术层面,相较于 YOLO 系列前几代,YOLOv10 采用一致的双重标签分配策略,消除

了非极大值抑制训练依赖,从而缩减了推理延迟。同时,团队还提出整体效率-精度驱动模型架构设计策略,涵盖了轻量级分类头、空间通道解耦下采样以及大内核卷积等内容,这些策略对提升模型的效率与准确性均起到了积极作用。

然而同上文所述,该模型仍然依赖于局部邻域信息执行采样操作。其中,SCDown(selective channel downsampling)作为原始 YOLOv10 模型的下采样模块,采用标准卷积和二步长卷积的组合实现了空间-通道分离下采样。上采样则使用了最近邻插值算法。这些基于局部规则的采样技术限制了模型对全局信息的理解。为此,设计了 GADM 替换原始模型 Backbone 上采样模块,GAUM 替换 Head 下采样模块。改进后的模型整体结构如图 1 所示。

2.2 全局感知下采样模块

GADM 是一种新的可学习的下采样器。GADM 通过加权组合和重新排列输入特征图中每个位置的特征,从而在全局感受野中聚合信息,并在保持计算效率的同时,动态地增强目标信息并减少细节信息的丢失。

GADM 由 2 个关键组件组成,分别是全局感知模块(global awareness block, GAB)和降变换模块(down transform block, DTB),如图 2 所示。

1) 全局感知模块

通过对输入特征图变换维度,在不丢失信息的前提下,计算并生成对应的全局感知权重。

具体地,先将输入特征图 X_i 进行维度变换,将原始的三维特征图 $X_i \in \mathbf{R}^{C \times H \times W}$ 变换为二维特征向量 $X \in \mathbf{R}^{HW \times C}$ 。第 1 层线性层对每个特征向量各个元素进行全局统计计算,区别于卷积只对元素局部特征操作,这样的方式可以使得输出的每个元素都来自于全局空间信息。公式表示为:

$$y_i = \sum_{j=1}^{HW} w_{i,j} x_j + b_i, \quad i \in [1, HW/r] \quad (1)$$

$$Y = WX + b, \quad W \in \mathbf{R}^{Hh/r \times Hh} \quad (2)$$

其中, x_j 表示特征向量的第 j 个元素, y_i 表示输出层元素, $w_{i,j}$ 表示连接当前层和下一层的权重, b_i 表示下一层的偏置。同时,还引入一个缩减比 r ,用以控制输出维度,不仅控制模型参数,同时也影响全局信息的聚合。因此寻找参数量和检测精度的最佳平衡也是此研究的任务之一。

接着通过归一化层 Norm 和激活函数 Relu,然后再次引入线性层,对特征向量进行维度还原。因此,其结构可以简化为:

$$\text{Linear}(HW, HW/r) \longrightarrow \text{Norm \& Relu} \longrightarrow \text{Linear}(HW/r, HW) \quad (3)$$

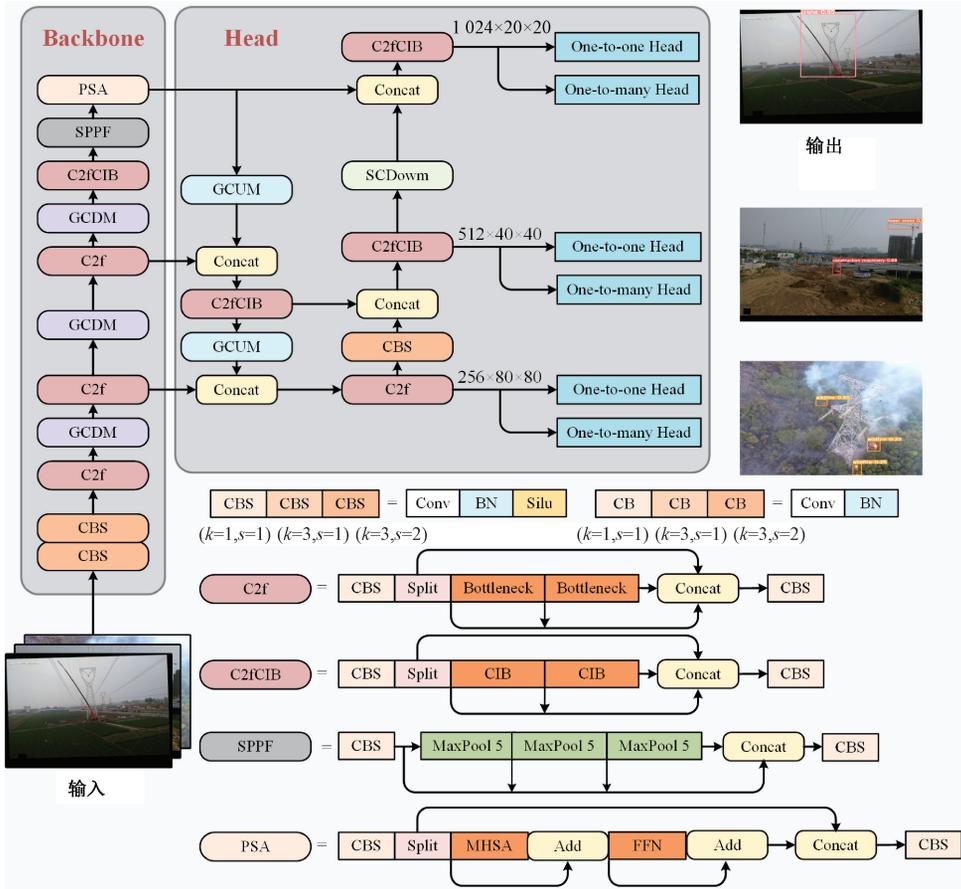


图1 GAE-YOLO的整体结构

Fig. 1 The overall architecture of GAE-YOLO

经过这一系列的操作,特征向量的每个元素都具备全局空间性。最后在使用 Softmax 函数对每个特征向量进行归一化,将二维还原为三维,并将其定义为全局感知权重 \mathbf{W}_A 。该权重通过自适应聚合全局空间信息,使每个采样点的特征值受所有位置的影响,从而克服局部窗口的局限性。

全局感知权重是输入特征图的抽象化表示,因此再将全局感知权重与输入特征图 \mathbf{X}_i 进行逐元素相乘,得到采样具象化表示 \mathbf{X}_A ,如式(4)所示。

$$\mathbf{X}_A = \mathbf{W}_A \odot \mathbf{X}_i \quad (4)$$

2) 降变换模块

通过 PixelUnShuffle 变换空间, 1×1 卷积用于改变通道,两者结合实现了下采样变换,公式为:

$$\mathbf{X}_0 = \text{Conv}_{1 \times 1, c}(\text{PixelUnshuffle}(\mathbf{X}_A)) \quad (5)$$

其中, PixelUnShuffle 将输入特征图按块重排为低分辨率输出,是一种无参数的空间信息重组,避免传统池化的不可逆信息丢失,可以达到特征图 2 倍下采样的目的,如图 3 所示。

通过 GAB 和 DTB 的配合,将全局空间特征聚合到

每个位置中,可以缓解信息丢失的影响,使模型在 Backbone 处理下采样特征时能够保持信息的丰富性。

2.3 全局感知上采样模块

传统上采样仅依赖局部像素距离,导致边界模糊。GAUM 是一种新的可学习的上采样器,如图 4 所示。GAUM 通过学习采样点的通道隶属关系,动态生成契合全局上下文的新像素值,让后续操作能基于更优质的特征进行,进而提升整个神经网络在检测任务中的准确率和鲁棒性。

本研究提出的 GAUM 包括 GAB 和升变换模块(up transform block, UTB)。虽然 GADM 和 GAUM 的 GAB 均是生成全局感知权重,但二者在设计阶段存在差别。

在 GADM 架构里,其 GAB 模块获取全局感知权重的方式基于输入特征图的空间信息计算。然而,考虑到上采样发生于 Head 部分,此时特征图已积累了较为丰富的上下文信息,再次引入空间维度的全局信息不仅对提升效果助力有限,反而容易造成信息冗余。因此这也是保留原模型 Head 下采样模块的原因。而反观常见的上采样手段,多聚焦于从当前特征图的局部邻域获取信息,这

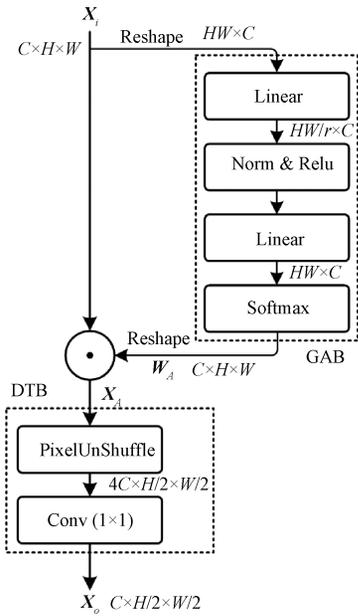


图 2 GADM 结构
Fig. 2 GADM architecture

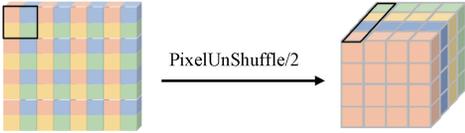


图 3 PixelUnShuffle 操作
Fig. 3 PixelUnShuffle operation

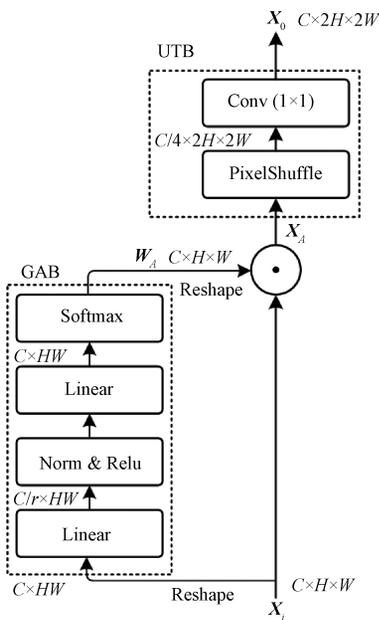


图 4 GAUM 结构
Fig. 4 GAUM architecture

一过程忽视了通道间的内在联系与隶属关系,导致通道维度的信息整合与利用不足。针对这一问题,所设计的 GAUM 中的 GAB 采用输入特征图的通道信息来计算全局感知权重。通过这种方式,GAUM 能够深度挖掘通道间的潜在关联,打破传统上采样在通道维度的局限性,使得生成的特征信息在通道层面得以有效融合与优化。

因此 GAUM 的 GAB 大致为:将输入特征图变换为二维特征向量 $\mathbf{X} \in \mathbf{R}^{C \times HW}$,第 1 层线性层对每个元素位置的各个通道进行全局统计计算,即:

$$y_m = \sum_{n=1}^C w_{m,n} x_n + b_m, \quad m \in [1, C/r] \quad (6)$$

$$\mathbf{Y} = \mathbf{W}\mathbf{X} + \mathbf{b}, \quad \mathbf{W} \in \mathbf{R}^{C/r \times C} \quad (7)$$

后续的操作和 2.2 节的 GAB 类似,最终该模块生成全局感知权重 W_A ,并与输入特征图 X_i 进行逐元素相乘,得到具有全局通道性的特征图 X_A 。

3) 升变换模块

通过 PixelShuffle 和 1×1 卷积来进行上采样特征变换,用公式表示为:

$$\mathbf{X}_o = \text{Conv}_{1 \times 1, c}(\text{PixelShuffle}(\mathbf{X}_A)) \quad (8)$$

其中,PixelShuffle 操作利用超分辨率技术,通过重排像素来提高图像分辨率,而不是简单的插值方法。这种方式能够更好地保留图像的语义关联,避免插值过程中可能引入的伪影和失真。由于网络的上采样因子为 2,因此 PixelShuffle 层的缩放因子设置为 2,达到特征图 2 倍上采样的目的,如图 5 所示。

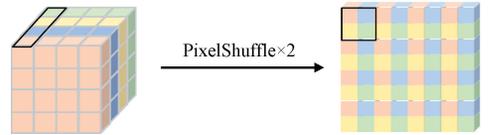


图 5 PixelShuffle 操作
Fig. 5 PixelShuffle operation

在 PixelShuffle 操作之后,使用 1×1 卷积进行通道变换,以确保与插值算法输出特征图的形状一致。

通过利用 GAB 和 UTB 进行的全局感知上采样,减轻了插值算法引起的边界模糊和采样点归属模糊等问题,从而有效地学习特征图的全局通道信息。

3 实 验

3.1 数据集介绍

本研究中使用的原始图像数据集是通过陕西省某电网公司部署的监测设备收集的。初始数据集较为杂乱,经过整理后,将带有外破隐患的图像进行了分类和整合。最终的外破隐患数据集包含 4 种常见类型:起重机、塔吊、施工机械和山火,具体信息如表 1 所示。

表1 数据集的样本统计
Table 1 Dataset sample statistics

数据集	类别	目标数量	图像数量
输电线路外 破隐患数据集	施工机械	2 475	7 035
	吊车	2 211	
	塔吊	2 754	
	山火	4 338	

3.2 实施细节与评价指标

按照迁移学习的常见做法,首先在 ImageNet1k 图像分类数据集上对所有主干网络进行了预训练,然后在现有数据集的训练集上进行了微调。所有模型均在 PyTorch 深度学习框架下使用 Nvidia RTX 4090 GPU 进行训练、评估和测试。超参数设置如表 2 所示。

对于输电线路外破隐患数据集,将数据集按 6:2:2 的比例划分为训练集 4 221 张、验证集 1 407 张和测试集 1 407 张。

性能评估指标包括平均精度(average precision, AP)和 mAP。AP 计算不同类别的平均精度,mAP 则对所有类别的精度进行平均,是衡量目标检测模型整体性能的

表2 超参数设置

Table 2 Hyperparameter settings

超参数	值
训练轮次	300
优化器	SGD
权重衰减	5×10^{-4}
动量因子	0.937
批量大小	16
初始学习率	1×10^{-2}
最终学习率	1×10^{-4}

常用指标。每秒帧数(frames per second, FPS)用于评估模型在批量大小为 1 的推理速度。

3.3 实验结果及分析

1)与当前流行的目标检测器的对比

该研究对比了不同流行目标检测器在外破隐患数据集上的表现,具体包括 Faster-RCNN^[28]、SSD^[29](single shot multibox detector)、YOLO 系列^[30-35]以及 RT-DETR^[36](real-time detection transformer),实验结果如表 3 所示。

表3 与先进的目标检测器的比较结果

Table 3 Comparison results with advanced object detectors

目标检测器	主干网络	参数量/ M	模型大小/ MB	FLOPs/ G	AP/%				mAP@ 0.5/%	FPS
					施工机械	吊车	塔吊	山火		
Faster-RCNN	VGG-16	137.09	333.52	368.24	81.56	86.32	72.49	85.67	81.51	72.36
Faster-RCNN	ResNet-101	41.53	109.27	251.43	83.83	87.69	74.00	85.27	82.70	80.87
SSD300	VGG16	26.28	50.47	116.20	71.35	81.97	74.29	85.70	78.33	190.30
SSD512	VGG16	-	-	-	73.62	82.15	73.31	86.58	78.92	153.27
YOLOv3	Darknet-53	61.63	120.48	193.92	80.12	83.56	73.31	85.62	81.15	168.32
YOLOv4	CSP-Darknet53	64.36	100.83	119.95	84.26	85.38	75.21	87.46	83.07	150.58
YOLOv5-L	Modified CSP v5	46.56	89.18	109.13	84.95	86.02	76.01	87.54	83.63	178.45
YOLOv7	ELAN	37.21	72.03	104.85	86.32	86.45	77.80	87.47	84.51	185.47
RT-DETR-L	HGNetv2	32.02	63.38	110.08	85.78	88.32	84.21	87.92	86.55	194.63
YOLOv9-C	GELAN	25.45	49.30	102.77	86.39	89.22	85.04	88.70	87.40	186.21
YOLOv10-L	Enhanced CSPNet	24.44	48.98	118.11	87.04	89.56	86.21	88.86	87.92	210.58
YOLO11-L	Enhanced CSPNet	25.32	49.03	86.89	86.95	89.80	87.10	87.92	87.94	204.59
GAE-YOLO(所提方法)	Enhanced CSPNet	29.91	54.40	120.26	96.66	94.08	91.08	90.40	93.05	188.33

从表 3 可以看出, Faster-RCNN 使用 VGG (visual geometry group) 和 ResNet (residual network) 作为主干网络 mAP 达到了 81.51% 和 82.70%。SSD300 和 SSD512 的 mAP 分别达到了 78.33% 和 78.92%。YOLO 系列目标

检测器在检测精度上表现更好,同时在检测速度上也具有一定优势。YOLOv3 的 mAP 达到了 81.15%,FPS 达到了 168.32。YOLOv4、YOLOv5-L 和 YOLOv7 进一步提高了检测精度。端到端目标检测器的表现也非常优秀,其

中 RT-DETR-L 的 mAP 达到了 86.55%,YOLOv9-C 的 mAP 达到了 87.40%,而最新的 YOLO11-L 的 mAP 则达到了 87.94%,检测速度为 204.59 fps。新的目标检测器 YOLOv10-L 在参数量最低的条件下,检测精度上 mAP 达到了 87.92%,最快检测速度为 210.58 fps。该研究以 YOLOv10-L 作为基准,改进的方法 GAE-YOLO 达到了最佳的检测精度, mAP 为 93.05%,比 YOLOv10-L 高出 5.13%。 FPS 达到了 188.33,满足实时性要求。

总的来说,这些先进的目标检测方法利用局部操作,如池化(或步幅卷积)和插值算法进行特征图的下采样和上采样。与之不同,该研究所提方法在上下采样过程中引入了全局信息,以捕捉采样点之间的相互作用,从而增强特征图的表达能力。该方法不仅能有效保留重要的上下文信息,还能有效恢复图像细节,特别是在物体边界和关键特征处。实验结果验证了所提方法在目标检测领域的有效性和应用潜力。

2) 与输电线路外破隐患目标检测算法的对比

该研究将所提出的方法与近年来提出的几种性能较好的输电线路外破隐患目标检测算法进行对比,分别来自文献[22-25]。在相同的数据集上,使用精确率(precision, P)、召回率(recall, R)、 IoU 阈值为 0.5 的平均检测精度($mAP@0.5$)等指标对比了各方法的检测性能,结果如表 4 所示。

表 4 与输电线路外破隐患目标检测算法的对比结果

Table 4 Comparison with transmission line external force damage detection algorithms (%)

算法	P	R	$mAP@0.5$
改进 YOLOv5s ^[22]	89.85	85.36	88.41
YOLOv5s ^[23]	88.32	83.76	86.79
YOLO-RANGE ^[24]	92.74	89.53	91.50
改进 YOLOv7 ^[25]	91.25	87.94	90.34
所提方法	94.57	91.20	93.05

结果表明,提出的方法在各项评价指标上均表现出色,尤其是在 $mAP@0.5$ 上达到 93.05%,超越了所有对比方法。此外,该方法在保持高精度的同时,精确率和召回率分别达到了 94.57% 和 91.20%,证明所提方法在输电线路外破隐患检测任务中的强大能力。

为了更加直观地对比输电线路外破隐患目标检测同类算法和所提方法在检测任务中的表现,在数据集上进行了定性分析,结果如图 6 所示。实验结果表明,在隐患目标姿态变化和背景复杂的情况下,文献[22-23]提出的方法出现了较为严重的漏检情况,检测精度也不如其他方法。文献[24]提出方法的检测精度稍逊于本研究所提方法,但表现依然出色,尤其对于大部分清晰可见的隐

患目标具有较高的召回率,但是对于遮挡物体出现漏检情况。文献[25]所提方法在检测精度上略有下降,对隐患目标的置信度随之下降,对于遮挡物体的漏检现象较为明显。本研究提出的方法检测精度较高,能够在复杂的场景、不同的物体姿态、遮挡情况下准确识别输电线路外破隐患目标,并且没有出现误检和漏检的情况。这些定性分析结果进一步证明了本研究所提出方法的优越性能。



图 6 不同检测算法的定性分析结果
Fig. 6 Qualitative analysis results of different detection algorithms

3) 消融实验

不同的缩减比 r 。为了研究不同缩减比对所提模块性能的影响,在 CADM 和 CAUM 中测试了缩减比分别为 8~128 和 2~32 的情况,并分析了性能的变化。如表 5 所示,对于 GADM,使用不同的 r 值对模型性能的提升效果不同。实验结果表明,随着 r 值的增加,模块参数量减少,模型精度略微下降。当 r 增长到 64 前,模型精度下降幅度较低。当 r 设置为 128 时,模型精度下降幅度突然增大,这意味着压缩的信息在一定程度上影响着感知权重的聚合。因此,综合参数量和精度的考虑,选择了缩减比为 16 的 GADM。表 6 显示,GAUM 也通过调节缩减比 r 以调节模型的性能。与 GADM 的表现不同,当 r 设置为 16 时,模型

表5 不同缩减比下 GADM 的性能

Table 5 Performance of GADM under different reduction ratios

YOLOv10-L	r	参数量/M	FLOPs/G	AP/%				$mAP@0.5/\%$
				施工机械	吊车	塔吊	山火	
baseline		24.44	118.11	87.04	89.56	86.21	88.86	87.92
+GADM	8	33.69	123.80	95.42	93.16	89.81	89.91	92.08
+GADM	16	31.45	121.49	94.92	92.89	89.55	90.22	91.89
+GADM	32	30.33	120.86	94.50	92.15	89.43	89.51	91.40
+GADM	64	29.78	120.04	94.21	92.52	89.86	89.85	91.61
+GADM	128	29.49	119.31	93.86	89.95	88.23	88.91	90.24

表6 不同缩减比下 GAUM 的性能

Table 6 Performance of GAUM under different reduction ratios

YOLOv10-L	r	参数量/M	FLOPs/G	AP/%				$mAP@0.5/\%$
				施工机械	吊车	塔吊	山火	
baseline		24.44	118.11	87.04	89.56	86.21	88.86	87.92
+GAUM	2	25.02	119.58	87.59	90.02	86.80	89.12	88.38
+GAUM	4	24.76	118.41	89.63	89.80	87.32	88.95	88.93
+GAUM	8	24.63	118.36	90.32	89.86	88.43	89.75	89.59
+GAUM	16	24.56	118.20	91.26	90.17	88.58	89.91	89.98
+GAUM	32	24.54	118.15	90.23	89.50	86.59	88.61	88.73

mAP 达到了最高的 89.98%。以上 2 个实验均反映了缩减比 r 对所提方法 GADM 和 GAUM 的重要性。

基于 YOLOv10 的消融研究。为了验证所提各个组

件的有效性,以 YOLOv10-L 为基准进行了消融实验,并逐步添加了 GADM 和 GAUM。表 7 展示了各个组件的消融实验数据。

表7 提出的方法在 YOLOv10 上的消融实验

Table 7 Ablation study of the proposed method on YOLOv10

YOLOv10-L	GADM	GAUM	参数量/M	FLOPs/G	AP/%				$mAP@0.5/\%$
					施工机械	吊车	塔吊	山火	
Enhanced CSPNet-PANet			24.44	118.11	87.04	89.56	86.21	88.86	87.92
Enhanced CSPNet-PANet	✓		29.78	120.04	94.21	92.52	89.86	89.85	91.61
Enhanced CSPNet-PANet		✓	24.56	118.20	93.26	90.17	87.58	88.91	89.98
Enhanced CSPNet-PANet	✓	✓	29.91	120.26	96.66	94.08	91.08	90.40	93.05

总体来看,每个模块互补,GADM 相比基准模型 mAP 提高了 3.69%,验证了 GADM 的有效性。而加入 GAUM 使得 mAP 从 87.92% 提升至 89.98%,进一步证明了 GADM 和 GAUM 能够缓解上采样和下采样过程中信息的损失。通过评估指标可以看出,嵌入这两个模块后,模型的整体性能得到了提升。该方法提高了外破隐患目标的检测精度, mAP 达到了 93.05%,高出基线模型 5.13%。

4) 影响因素分析

输电线路的检测图像通常具有较大范围的背景,尤

其是超高压架空输电线路,其通道宽广且覆盖的区域非常广泛。在此背景下,4 类隐患目标的实际尺寸较小,尤其是山火这一类目标,其在图像中占据的像素区域较少,造成检测方法对其检测精度和置信度相对较低。

为了定量分析小目标对检测性能的影响,在图 7 展示了目标相对于整幅图的高宽比例图,明确说明了小目标的数量之多以及其在整体目标中的比例。由图 7 可以看到,在图像中,小目标占据了较大比例,这一特点在大范围监控场景中尤为突出。

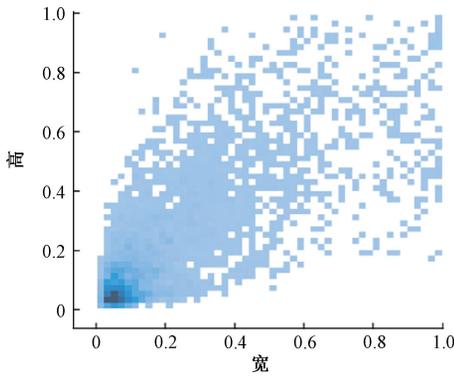


图 7 隐患目标相对于整幅图像的高宽比例
Fig. 7 Aspect ratio of the damaged objects relative to the entire image

为了进一步验证小目标对检测精度的影响,采用了小、中、大目标的平均精度 AP_s 、 AP_m 、 AP_l 作为定量分析手段。具体来说,分别计算了小目标、中目标和大目标的 AP 值,并对比了这些不同尺度目标的检测精度。实验结果如表 8 所示。

表 8 模型对不同尺度目标的检测结果

Table 8 Detection results of the model for objects of different scales (%)

算法	$AP_s@0.5:0.95$	$AP_m@0.5:0.95$	$AP_l@0.5:0.95$
基线模型	49.69	67.35	76.43
GAE-YOLO	58.34	73.50	80.61

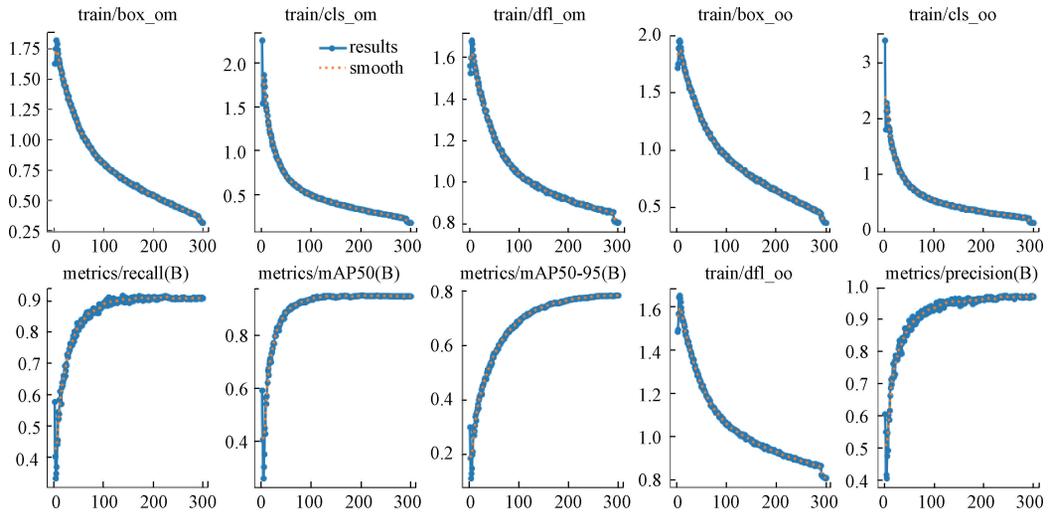


图 8 模型训练的损失曲线
Fig. 8 Loss curve of model training

结果表明,小目标的检测精度明显低于中、大目标,而小目标的存在是对模型在实际应用中产生较大影响的主要因素。

5) 可视化实验结果

在历经 300 轮迭代训练的过程中,本研究模型于训练集与验证集上所展现的 box loss、cls loss 及 dfl loss 变化趋势,以及包括 precision、recall、 $mAP@0.5$ 和 $mAP@0.5:0.95$ 在内的性能评估指标的收敛情况,如图 8 所示。该图以迭代次数为横轴坐标。观察图 8 可发现,大约在 290 次迭代之后,各个损失函数曲线逐渐逼近零点,标志着模型开始进入稳定收敛阶段。

如图 9 所示,使用下采样特征层生成了大尺度隐患目标的热力图,在上采样特征层上生成了小尺度隐患目标的热力图。结果表明,与基线模型相比,该研究的全局感知上下采样模块更关注隐患目标本身,并具有更清晰

的目标轮廓。

图 10 展示了基线模型与所提方法在不同现实场景下的检测结果对比,包括复杂背景、亮度变化和不同的姿态。可以观察到,与基线模型相比,所提方法在更复杂背景和暗场景的环境下,能够更准确地定位物体的边界框,减少了漏检的情况,展现出更强的鲁棒性和泛化能力。而基线模型在这些场景中的表现则出现了下降,特别是在该案例中的暗场景无法检测到隐患目标,在不同姿态环境下则出现了漏检情况。

图 11 显示了所提方法在实际场景中对 4 种类型的外破隐患目标的检测结果:施工机械 (construction machinery)、吊车 (crane)、塔吊 (tower crane) 和山火 (wildfire)。图 11 表明所提方法能有效地检测到隐患目标,证明了所提方法在实际环境中的有效性和可行性。

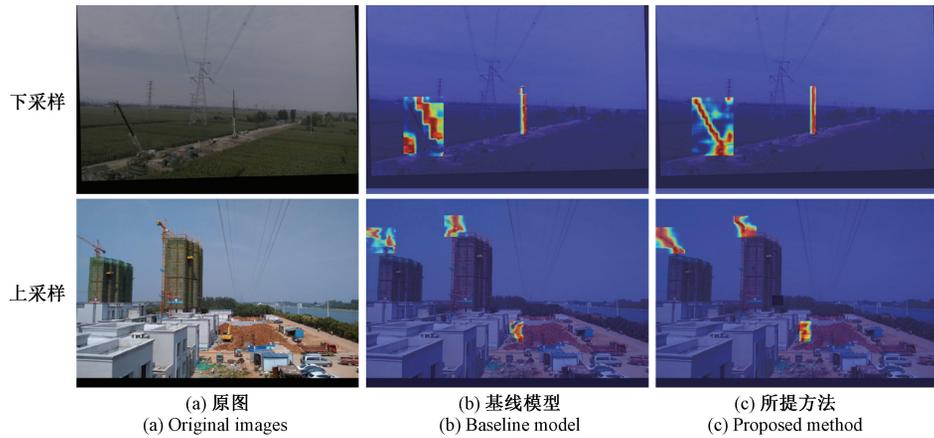


图 9 使用来自不同层的特征图生成的热力图

Fig. 9 Heatmaps computed using feature maps from different layers

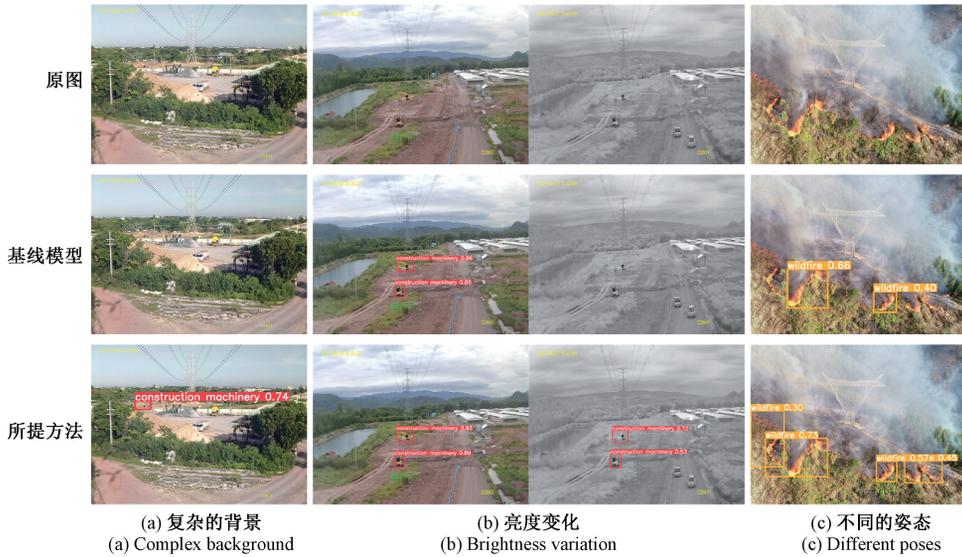


图 10 基线模型和所提方法在不同场景下的检测结果比较

Fig. 10 Comparison of detection results between the baseline model and the proposed method in different scenarios



图 11 所提方法对 4 类隐患目标的检测结果

Fig. 11 Detection results for four types of external force damage objects

4 结 论

本研究提出的 GAE-YOLO 输电线路外破隐患目标检测方法,通过引入 GADM 和 GAUM 模块,解决了传统上下采样操作忽略全局信息的问题。实验结果表明, GAE-YOLO 在适度增加参数量的情况下,显著提升了检测精度,具有较高的实际应用价值。该方法不仅为超高压架空输电线路的安全运行提供了新的思路,也可拓展至其他目标检测领域。未来的研究将扩展数据集类别、优化模块设计、提升推理速度与精度,并通过与电网公司的合作进行现场验证。

参考文献

[1] YANG L, FAN J F, LIU Y H, et al. A review on state-of-the-art power line inspection techniques [J]. IEEE

- Transactions on Instrumentation and Measurement, 2020, 69(12): 9350-9365.
- [2] SHAKIBA F M, AZIZI S M, ZHOU M CH, et al. Application of machine learning methods in fault detection and classification of power transmission lines: A survey[J]. Artificial Intelligence Review, 2023, 56(7): 5799-5836.
- [3] LI H, REN ZH Y, FAN M, et al. A review of scenario analysis methods in planning and operation of modern power systems: Methodologies, applications, and challenges[J]. Electric Power Systems Research, 2022, 205: 107722.
- [4] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 936-944.
- [5] LIU SH, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [6] HE L L, WANG M. SliceSamp: A promising downsampling alternative for retaining information in a neural network[J]. Applied Sciences, 2023, 13(21): 11657.
- [7] LIU M, LIN K, HUO W J, et al. Feature enhancement modules applied to a feature pyramid network for object detection[J]. Pattern Analysis and Applications, 2023, 26(2): 617-629.
- [8] YU D J, WANG H L, CHEN P Q, et al. Mixed pooling for convolutional neural networks[C]. Rough Sets and Knowledge Technology: 9th International Conference, 2014: 364-375.
- [9] SAEEDAN F, WEBER N, GOESELE M, et al. Detail-preserving pooling in deep networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 9108-9116.
- [10] GAO Z T, WANG L M, WU G SH. Lip: Local importance-based pooling[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 3354-3363.
- [11] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [12] SHI W ZH, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1874-1883.
- [13] NOH H, HONG S, HAN B. Learning deconvolution network for semantic segmentation[C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1520-1528.
- [14] WANG J Q, CHEN K, XU R, et al. Carafe++: Unified content-aware reassembly of features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(9): 4674-4687.
- [15] WANG AO, CHEN H, LIU L H, et al. Yolov10: Real-time end-to-end object detection[J]. ArXiv preprint arXiv:2405.14458, 2024.
- [16] 国家能源局. 架空输电线路运行规程: DL/T 741—2019[S]. 北京: 中国电力出版社, 2019.
- National Energy Administration. Operating code for overhead transmission line: DL/T 741—2019[S]. Beijing: China Electric Power Press, 2019.
- [17] 魏贤哲, 卢武, 赵文彬, 等. 基于改进 Mask R-CNN 的输电线路防外破目标检测方法研究[J]. 电力系统保护与控制, 2021, 49(23): 155-162.
- WEI X ZH, LU W, ZHAO W B, et al. Target detection method for external damage of a transmission line based on an improved Mask R-CNN algorithm[J]. Power System Protection and Control, 2021, 49(23): 155-162.
- [18] LENG X M, DAI J, GAO Y, et al. Overhead transmission line anti-external force damage system[C]. IOP Conference Series: Earth and Environmental Science, 2022, 1044(1): 012006.
- [19] QU L, LIU K X, HE Q, et al. External damage risk detection of transmission lines using E-OHEM enhanced faster R-CNN[C]. Pattern Recognition and Computer Vision: First Chinese Conference, 2018: 260-271.
- [20] 龙乐云, 周腊吾, 刘淑琴, 等. 改进 YOLOv5 算法下的输电线路外破隐患目标检测研究[J]. 电子测量与仪器学报, 2022, 36(11): 245-253.
- LONG L Y, ZHOU L W, LIU SH Q, et al. Identification of hidden damage targets by external forces based on domain adaptation and attention mechanism[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(11): 245-253.
- [21] 苏怡萱, 李智, 盘书宝. 基于改进 YOLOv8 的输电线路绝缘子缺陷检测方法[J/OL]. 电子测量与仪器学报, 1-12[2025-02-21].
- SU Y X, LI ZH, PAN SH B. Improved YOLOv8-based insulator defect detection method for transmission lines[J/OL]. Journal of Electronic Measurement and

- Instrumentation, 1-12 [2025-02-21].
- [22] 邢志鹏, 郑含博, 崔明慧, 等. 基于改进 YOLOv5s 的输电通道隐患目标检测方法[J/OL]. 中国测试, 1-9 [2024-12-11].
XING ZH P, ZHENG H B, CUI M H, et al. Detection method of hidden danger targets in transmission channels based on improved YOLOv5s[J/OL]. China Measurement & Test, 1-9 [2024-12-11].
- [23] LI H, JIANG F, GUO F, et al. A real-time detection method of safety hazards in transmission lines based on YOLOv5s [C]. International Conference on Artificial Intelligence and Intelligent Information Processing, 2022: 124561C.
- [24] LI J H, ZHENG H B, CUI ZH M, et al. Intelligent detection method with 3D ranging for external force damage monitoring of power transmission lines [J]. Applied Energy, 2024, 374: 123983.
- [25] 王彦海, 郭宸昕, 吴德强. 基于改进 YOLOv7 的输电线路机械外破隐患目标检测方法[J/OL]. 电测与仪表, 1-10 [2024-12-11].
WANG Y H, GUO CH X, WU D Q. Hidden target detection method for mechanical external damage of transmission line based on improved YOLOv7 [J/OL]. Electrical Measurement & Instrumentation, 1-10 [2024-12-11].
- [26] 徐业东, 蔡亚恒, 李严, 等. 基于 YOLOv5s 的轻量化架空输电线路鸟巢检测网络[J]. 电子测量技术, 2024, 47(7): 138-148.
XU Y D, CAI Y H, LI Y, et al. Lightweight overhead transmission line bird's nest detection network based on YOLOv5s [J]. Electronic Measurement Technology, 2024, 47(7): 138-148.
- [27] 熊伟, 路鑫, 邱维进, 等. 基于改进 YOLOv8 的输电线路绝缘子表面缺陷识别算法[J]. 电子测量技术, 2025, 48(2): 178-188.
XIONG W, LU X, QIU W J, et al. Surface defect detection algorithm of transmission line insulators based on improved YOLOv8 [J]. Electronic Measurement Technology, 2025, 48(2): 178-188.
- [28] REN SH 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.
- [29] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C]. Computer Vision-ECCV 2016; 14th European Conference, 2016: 21-37.
- [30] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. ArXiv preprint arXiv: 1804.02767, 2018.
- [31] BOCHKOVSKIY A, WANG CH Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv preprint arXiv:2004.10934, 2020.
- [32] JOCHER G, CHAURASIA A, STOKEN A, et al. Ultralytics/YOLOv5:v7.0-YOLOv5 sota realtime instance segmentation[J]. Zenodo, 2022.
- [33] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [34] WANG C Y, YE H I H, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information [C]. European Conference on Computer Vision, 2024: 1-21.
- [35] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [J]. ArXiv preprint arXiv:2410.17725, 2024.
- [36] ZHAO Y AN, LYU W Y, XU SH L, et al. Detsr beat YOLOs on real-time object detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

作者简介



刘敏, 2004 年于郑州大学获得学士学位, 2007 年于华中科技大学获得硕士学位, 2013 年于华中科技大学获得博士学位, 现为湖北工业大学副教授, 主要研究方向为机器学习、图像处理和人工智能。
E-mail: liu_min@hbut.edu.cn

Liu Min received her B. Sc. degree from Zhengzhou University in 2004, received her M. Sc. degree from Huazhong University of Science and Technology in 2007, and received her Ph. D. degree from Huazhong University of Science and Technology in 2013. She is currently an associate professor at Hubei University of Technology. Her main research interests include machine learning, image processing, and artificial intelligence.



陈明(通信作者), 2022 年于湖北工业大学获得学士学位, 现为湖北工业大学硕士生, 主要研究方向为机器视觉、图像处理和人工智能。
E-mail: 102210266@hbut.edu.cn

Chen Ming (Corresponding author) received his B. Sc. degree from Hubei University of Technology in 2022. He is currently a master student at Hubei University of Technology. His main research interests include machine vision, image processing, and artificial intelligence.