引文格式:许国伟,吴亮,翟晗,等.耦合置换注意力和可形变卷积的输电线金具目标检测[J].武汉大学学报(信息科学版),2026,51(3):544-556.DOI:10.13203/j.whugis20240409

Citation:XU Guowei, WU Liang, ZHAI Han, et al.A Model Coupled Shuffle Attention and Deformable Convolution for Overhead Transmission Line Hardware Object Detection[J].Geomatics and Information Science of Wuhan University,2026,51(3):544-556.DOI:10.13203/j.whugis20240409

耦合置换注意力和可形变卷积的输电线金具目标检测

许国伟[1] 吴 亮[1] 翟 晗[2] 李 欣[3]

1 广东电网有限责任公司汕头供电局,广东 汕头,515041

2 中国地质大学(武汉)地理与信息工程学院,湖北 武汉,430074

3 武汉大学遥感信息工程学院,湖北 武汉,430079

摘 要:金具部件的准确识别是架空输电线路安全巡检的一项关键性基础工作。针对金具目标纹理特征多变、形态复杂等问题,提出一种耦合置换注意力和可形变卷积的YOLOv8模型,用于架空输电线路无人机高分辨率可见光影像金具目标识别。通过置换注意力捕获通道和像素间的依赖关系,增强模型空间通道特征挖掘能力;耦合可形变卷积网络,精细挖掘各类金具目标的复杂空间形态特征,实现金具目标的高精度定位;采用基于动态非单调聚焦机制的边界框损失进行网络模型训练,提高模型的样本学习能力;采用无人机实地采集汕头架空输电线路高分辨率可见光影像作为数据源,构建金具目标数据集进行实验,并与多种先进的金具目标检测模型进行对比。实验结果表明,所提模型金具目标检测性能明显优于现有主流方法,在交并比阈值为0.5时,平均精度达到84.4%,所提模型能够快速、高精度识别各类金具目标,保障输电基础设施的安全维护。

关键词:金具目标检测;无人机;高分辨率可见光影像;YOLOv8;可形变卷积;置换注意力

中图分类号:P237 **文献标识码**:A **收稿日期**:2025-01-26

DOI:10.13203/j.whugis20240409 **文章编号**:1671-8860(2026)03-0544-13

A Model Coupled Shuffle Attention and Deformable Convolution for Overhead Transmission Line Hardware Object Detection

XU Guowei[1] *WU Liang*[1] *ZHAI Han*[2] *LI Xin*[3]

1 Shantou Power Supply Bureau, Guangdong Power Grid Co., Ltd., Shantou 515041, China

2 School of Geography and Information Engineering, China University of Geosciences(Wuhan), Wuhan 430074, China

3 School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

Abstract: **Objectives**: The accurate identification of hardware components is a crucial foundational task for the safety inspection of overhead transmission lines. This paper proposes a novel YOLOv8 model coupling shuffle attention and deformable convolution (YOLOv8-SADC) to address the issues of variable texture features and complex shapes of hardware targets, which is used for overhead transmission line hardware object detection based on high-resolution visible images by unmanned aerial vehicles. **Methods**: The shuffle attention is constructed to capture the dependencies between channels and pixels to enhance spatial channels feature mining ability of the model. The deformable convolutional network is used to finely mine the complex spatial morphological features of various hardware targets to achieve high-precision positioning of hardware targets. An intersection over union (IoU)-based loss with a dynamic non-monotonic focusing mechanism based bounding box loss (named Wies-IoU) is adopted for network training to improve sample learning

基金项目:南方电网公司科技项目(030500KK52220007);国家自然科学基金(42271386,42171434,42001313)。

第一作者:许国伟,高级工程师,研究方向为输电智能运维与遥感航测大数据应用。1231468@qq.com

通信作者:翟晗,博士,副教授。zhaihan@cug.edu.cn

李欣,博士,教授。xli2126@whu.edu.cn

ability of the model. **Results**: The high-resolution visible light images of Shantou overhead transmission lines were collected by unmanned aerial vehicles (UAVs) as data sources, and a hardware target dataset was constructed for experiments, and compared with several advanced hardware target detection models. (1) Comparative analysis was carried out on CPLID, 2511 and self-built datasets, and the CPLID dataset included 600 images of normal insulators taken by UAVs; the 2511 dataset uses a blue mat as the background, and shoots four types of fittings from multiple angles: insulators, bolt pins, anti-vibration hammers, and tension clamps, with a total of 2 511 images. The self-built dataset includes 6 types of fittings including connecting fittings, plate fittings, fan-shaped fittings, triangular fittings, short connecting fittings and long fittings, with a total of 516 images and 1 745 labels. Experimental results show that YOLOv8-SADC has excellent performance on CPLID and 2511 datasets, and can fully adapt to scenarios with simple backgrounds and single object categories. (2) The comparative experimental results show that the target detection performance of the proposed model is significantly better than that of the existing mainstream algorithms, and the mean average precision (mAP) reached 84.4% when the IoU threshold is 0.5, indicating that the method demonstrates good robustness to various types of fittings and can effectively detect all kinds of fittings. Among them, the mAP50 of the connecting category reached 90.52%, and the mAP50 of the long category reached 93.52%, indicating that the method has good detection ability for all kinds of fittings, especially the rectangular shape of the target fittings. For the sector fittings with small targets and a large number of occlusions, the detection accuracy is also improved. (3) There were 116 samples with blue sky background, 646 samples with cloud background, 495 samples with buildings as background, and 488 samples with natural environment as background. There is less interference in the blue sky background, and the target characteristics of the fittings are prominent, and it is easy to extract key information. The natural environment background is mostly trees and grasslands, and the detection accuracy decreases slightly. The cloud texture in the cloudy background may introduce some background interference, resulting in the reduced feature extraction efficiency of the model. There is a certain similarity between the structural characteristics of the building (such as straight lines, edges, etc.) and the shape of the fittings, which can easily lead to false detection. The performance improvement of YOLOv8-SADC varies in different backgrounds, with the natural environment and building background being relatively complex, and the detection accuracy is significantly higher in simple backgrounds (such as blue sky and cloudy background). **Conclusions**: YOLOv8-SADC, based on high-resolution visible light images captured by drones, enhances the extraction capability of spatial channels features through coupled shuffle attention, and introduces deformable convolution to delve deeper into the complex morphological features of metal hardware. Additionally, a Wise-IoU loss function based on a dynamic non-monotonic focusing mechanism is employed to optimize localization accuracy. Experiments conducted on a drone-acquired dataset along the coastline of Shantou, Guangdong demonstrate that this method significantly improves the detection performance of various metal fastener targets in complex scenes, effectively ensuring the integrity and detail of recognition results. This represents an efficient solution for metal fastener identification.

Key words: hardware object detection; unmanned aerial vehicles; high-resolution visible images; YOLOv8; deformable convolution; shuffle attention

架空输电线路是电力能源传输的基础设施，其状态稳定是电力系统安全运行的重要保障[1-2]。随着电网规模的不断扩大和电力设备的不断激增，电力设备日常安全维护面临更高的要求。金具作为架空输电线路上广泛应用的金属附件，起到支撑、固定和连续裸导线及其他导体的作用，是十分重要的组成部分[3-4]。然而，由于输电线路通常暴露在野外环境中，金具部件极易受到锈蚀，导致破损、脱落，严重威胁电力系统的安全稳定运行[5]。金具检测是金具状态检测的前提和基础，相较于金具检测，状态检测需要进一步分析金具的微观细节（如锈蚀程度、表面裂纹等），需要高分辨率影像和复杂的图像处理，在实际输电线路维护上，先确定金具位置，快速筛查潜在问

题,再通过更精细的检查方式进行金具状态检测。因此,金具部件检测成为输电线路正常运行的一项关键保障措施,具有十分重要的意义。

传统的金具部件检测主要依靠作业人员攀爬杆塔方式进行逐个排查,费时费力、效率低下。随着无人机遥感技术的快速发展,大区域高分辨率金具影像可以快速获取,其能详细记录金具部件的形态结构和空间细节信息,为金具部件检测提供了新的解决途径。近年来,无人机遥感技术结合人工智能算法成为架空输电线路巡检的研究热点[6-7]。然而,由于输电线路连接金具结构复杂、形态多变,如何从无人机获取的高分辨率遥感影像快速、准确地识别各类金具零部件,成为困扰架空输电线路巡检的一个难题。

目前,基于机器学习的架空输电线路金具目标检测已取得一定的进展。例如,考虑到防振锤、绝缘子等部件均具有特定的圆形边缘,王伟等[8]提出一种基于边缘提取的目标识别方法,取得了较好的检测效果。基于金具部件特定的几何形状,陈晓娟等[9]设计了Hough变换的防震锤检测方法。基于色区的不同,宋伟等[10]利用形态学处理和RGB彩色模型进行防振锤锈蚀检测。黄乐等[11]通过自适应同态滤波,改进局部方向模式及卡方距离法,实现弱光条件下的金具识别。刘进等[12]提出可适应目标多尺度特性并提高Canny边缘检测自适应性的“特征簇”目标定位算法。尽管这些机器学习算法在处理单一金具时表现较好,但面对复杂背景下多类部件检测时存在一定的局限性,通常检测精度有限,难以满足实际架空输电线路金具部件检测需求。

随着人工智能技术的快速发展,深度学习模型通过层次化特征提取和抽象,表现出强大的信息提取和非线性拟合能力,并在计算机视觉、模式识别、目标检测等众多领域得到广泛应用。近年来,基于深度学习的金具目标检测算法逐渐增多[13-14]。目前,大多数方法将公共数据集表现良好的目标检测模型直接迁移到金具目标检测,并加以优化后,取得了优异的表现。杨罡等[15]提出一种多尺度特征融合检测算法,用于输电线路关键部件识别,有效提升了识别效率和精度。ZHAO等[16]基于卷积神经网络(convolutional neural network,CNN)模型提取绝缘子多维特征,并利用支持向量机进行绝缘子状态检测。赵振兵等[17-18]提出一种嵌入关系模型,有效提升了小目标金具的检测精度,并基于深度学习构建了输电线路危险源智能监控系统,能够准确识别危险源,保障电力系统安全运行。ZHANG等[19]提出了一种高分辨率实时探测网络,通过迭代聚合逐步融合高分辨率影像多级特征,有效提高了输电线路中金具的检测精度。WANG等[20]提出一种基于多尺度几何变换和注意力掩蔽机制的金具检测方法,利用同源变换从不同视图中获取图像特征,融合多尺度特征,并利用注意力掩码机制降低模型复杂度。ZHANG等[21]提出一种基于输电线路知识推理的端到端螺栓缺陷检测方法,结合位置知识和属性知识辅助螺栓缺陷分类器推理螺栓缺陷,有效识别了视觉无法区分的螺栓缺陷。CHEN等[22]提出一种结合YOLOv5(you only look once version 5)和DeblurGANv2(deblur generative adversarial network)目标实时检测网络,有效增强了模型对输电线路硬件的关注度。WANG等[23]在YOLOv5网络的基础上,采用*K*-means聚类算法优化锚框的大小,通过递归门控卷积提高关键目标的特征提取能力,有效提高了高压输电线路关键目标缺陷的检测性能。汤寓麟等[24]使用遗传算法A+*k*-means算法优化检测框,并对损失函数进行改进,从而检测沉船目标的侧扫声呐。ZHAO等[25]针对螺栓提出一种基于位置关系的输电线路螺栓及其缺陷检测方法,在Faster R-CNN中增加空间注意力模块,以候选区域的相对几何特征为输入,对螺栓及其缺陷的空间位置关系进行建模,有效提高了检测精度。张瑞倩等[26]提出一种多尺度空洞卷积的无人机影像目标检测方法,在现有的目标检测算法的基础上,增加多尺度的空洞卷积模块,加大视野感知域,提高网络对无人机影像中的目标分布情况、尺寸差异等特点的学习能力。张宇等[27]提出了一种基于逐层特征选择的多层部件模型目标检测算法,通过提取目标多特征并结合局部判别式模型的建模方法同样也可用于金具目标检测。

虽然基于深度学习的金具目标检测取得了一定的进展,但现有模型通常仅关注目标局部信息,忽略了影像蕴含的丰富背景信息和各金具目标间的相互关系,在一定程度上限制了检测精度。此外,金具目标通常存在纹理特征多变、形态复杂等问题,进一步增加了金具目标检测的挑战性。针对上述问题,本文以YOLOv8网络为基础,提出一种新颖的耦合置换注意力(shuffle attention,SA)和可形变卷积的YOLOv8模型(YOLOv8 coupled SA and deformable convolution,YOLOv8-SADC),用于架空输电线路无人机高分辨率可见光影像金具目标检测。首先,在特征

提取部分,通过耦合SA捕获通道和像素间的相关性信息,深入挖掘各类金具目标间的相互关系,生成更具判别性的空-谱特征。其次,在网络模型中耦合可形变卷积,灵活建模形态多变的各类金具目标,提升模型的特征表达能力。然后,采用基于动态非单调聚焦机制的边界框损失函数(Wise-IoU(intersection over union))指导模型训练,平衡高质量锚框和低质量锚框对样本学习的影响,引导模型充分挖掘各类金具部件特征,实现复杂场景的金具部件准确识别。

1　方　法

1.1　总体架构

为提高复杂场景的金具目标检测能力,减少错检、漏检,本文充分考虑金具纹理特征多变、形态复杂等特点,提出一种耦合SA和可形变卷积的YOLOv8-SADC,实现复杂环境下无人机高分辨率可见光影像各类金具部件的准确识别。

YOLOv8-SADC模型网络总体架构如图1所示。与YOLOv8模型[28]类似,YOLOv8-SADC模型同样包括4个主要部分:输入部分、Backbone部分、Neck部分和Head部分。其中,输入部分、Neck部分和Head部分沿用YOLOv8模型的改进,以保持高效的特征增强和目标预测。在Backbone部分,考虑到YOLOv8模型采用常规卷积进行目标特征提取,仅关注目标局部信息,忽略了影像蕴含的丰富背景信息和各金具目标间的相互关系,本文耦合SA模块,深入挖掘通道间和像素间的依赖关系,增强模型特征提取能力,以学习判别性更强、更加鲁棒的空间通道语义特征,更准确地表示各类金具目标。考虑到金具目标形态结构的不规则性,本文耦合可形变卷积,深入挖掘复杂多变的各类金具目标的准确特征表示,提高特征的鲁棒性,避免因形变带来的漏检和误检。此外,还采用先进的Wise-IoU损失函数替代常规损失函数CIoU(complete IoU),指导整个网络模型的高效学习训练,通过动态非单调聚焦机制,增强模型的样本学习能力,从而提高各类金具位置和类别的预测精度。

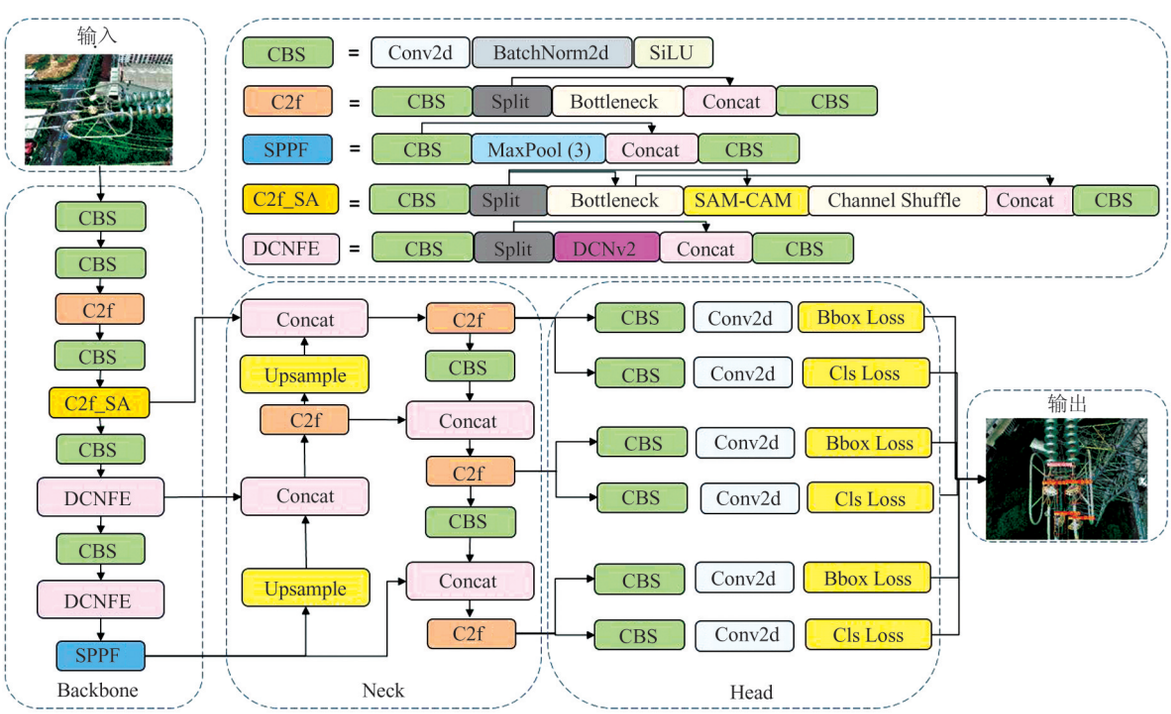


图1　YOLOv8-SADC总体架构

Fig. 1　Overall Architecture of YOLOv8-SADC

1.2　SA

传统的C2f模块虽然通过将输入特征分为两部分,一部分直接残差连接传递信息,另一部分经瓶颈块挖掘深层特征,最后融合,在一定程度上实现了特征的多层次处理,但对于特征之间的交互和融合主要局限于局部和固定的路径,缺乏对特征全局关系的充分挖掘和利用。为增强网络的特征提取能力,本文在Backbone部分耦合SA,并将其嵌入C2f模块中。C2f模块是CSP(cross stage partial)模块的变体之一,通过引入更少的计算量和参数,实现对特征更精细的处理。SA的核心思想是对输入特征图进行打乱操作,

然后在打乱后的特征上应用注意力机制,最后再将特征恢复到原始顺序。通过这种方式,模型能够打破特征之间的固有位置关系,迫使模型学习到更具全局性和鲁棒性的特征表示,而不仅仅依赖于局部的位置信息。

SA通过置换单元组合空间注意力和通道注意力,相比传统混合注意力更加高效,是一种非常轻量的注意力结构[29]。通道注意力用于捕获通道间的依赖关系,空间注意力用于提取像素间的相关性信息。通过不同的聚合策略、转换方法和强化函数,从所有位置聚合相同的特征来强化原始特征。SA模块可以有效捕获通道注意力和空间注意力之间的相关性,将输入按照通道进行分组,对每组子特征使用置换单元构建通道注意力和空间注意力。

SA的结构如图2所示,其主要包括3个部分:特征分组、混合注意力和特征聚合。首先,将输入特征分为多组,每组对应一个子特征,每个子特征对应一种特定的语义信息。其次,子特征沿通道维度分为两个分支,一个分支利用通道注意力获取通道间的依赖关系,另一分支捕获特征之间的空间依赖生成空间注意力图。然后,将两分支拼接到一起,同时完成语义和位置信息的捕获。最后,将各组子特征打乱通道顺序,重新组合,促进不同子特征之间的交互,确保特征之间的信息能够充分结合,从而生成与输入特征图维度相同的输出。

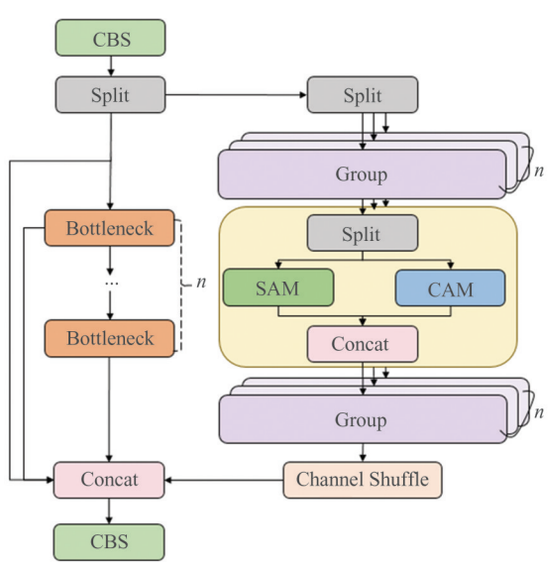


图2 置换注意力结构

Fig. 2 Structure of Shuffle Attention

在很多情况下,模型容易过度依赖特征的原始位置信息,导致对数据的微小变化(如目标的平移、旋转等)较为敏感,泛化能力受限。SA的引入使得网络不再仅仅依据特征的绝对位置来进行处理和决策,而是更注重特征的内在语义和全局关系。当面对不同场景的输入数据时,即使目标的位置发生了变化,经过SA处理后的C2f模块能够更好地适应这种变化,因为它已经学习到了不依赖于特定位置的特征表示。例如,在目标检测任务中,对于同一类目标出现在图像不同位置的情况,带有SA的C2f模块能够更稳定地提取出目标的关键特征,从而提高检测的准确性和鲁棒性,减少因位置变化导致的漏检或误检。SA机制在C2f模块中的引入为特征的流动和梯度的反向传播提供了新的路径和方式。打乱和重组特征的过程在一定程度上可以缓解梯度消失或梯度爆炸的问题,使得网络在训练过程中能够更高效地学习到有用的特征。通过SA,模型在特征提取阶段可以实现空间通道信息交互,有效捕获通道间和像素间的依赖关系,充分挖掘场景信息和目标间的对应关系,提高特征判别性,获取复杂多变的各类金具目标的精确特征表达,从而更加准确地预测各类金具目标的位置和类别,提高模型的检测精度。

1.3 可形变卷积

金具目标形状结构通常呈现一定的不规则性,导致传统固定窗口的卷积操作难以准确地捕获形态各异的金具目标的边缘与结构细节,在一定程度上导致特征表示不精确、判别性信息丢失等问题,在很大程度上限制了目标检测的精度。对此,本文在YOLOv8模型基础上,在特征提取的Backbone部分引入可形变卷积,构建可变形特征提取模块,以进一步增强模型的特征提取能力。

传统卷积操作具有固定的感受野和采样位置,对于具有复杂几何形变的物体(如形状不规则的物体、因视角变化而产生形变的物体等),难以准确捕捉其特征。例如,在金具检测任务中,长方形、正方形、三角形、凹凸不平、金具的形状多样,传统卷积可能无法很好地适应这些变化,可形变卷积网络(deformable convolutional networks v2,DCNv2)在处理具有几何形变物体时,表现出出色的性能[30]。DCNv2通过在标准卷积基础上引入可学习的偏移量,使得卷积核能够根据输入特征的内容自适应地调整采样位置,通过偏移量来调整卷积核的采样点,从而灵活建模不同形态金具目标,学习准确的目标特征表示。其次,通过引入位移增强了感兴趣区域池化的灵活

性,该模型可以对不同形状的候选区域进行更好的处理。此外,DCNv2的卷积核不仅可以学习位置偏移,还可以学习不同采样点的权重,使得卷积能够聚焦于重要的区域或特征。

基于此,本文将DCNv2耦合到C2f模块中,构建可形变卷积特征提取模块(deformable convolutional network based feature extraction module,DCNFE),以增强模型对各类形态金具目标的适应能力,其结构如图3所示。

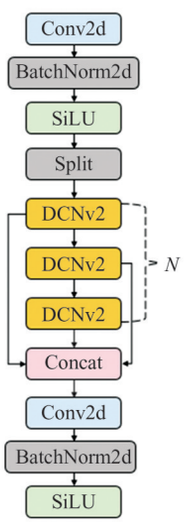


图3　可形变卷积特征提取模块结构

Fig. 3　Architecture of DCNFE

具体而言,输入特征首先经过一个普通卷积块,进行初步特征提取。经过卷积后的特征被分为两部分,一部分特征依次进入多个DCNv2模块进行处理,这些DCNv2模块能够自适应调整卷积核的位置,从而更好地处理形变特征。另一部分特征跳过DCNv2模块,直接传递到后续模块。然后,将经过DCNv2模块处理的特征与跳过处理的特征进行特征融合,从而整合普通卷积和可形变卷积学习得到的互补信息,生成更具判别性的特征。最后,将融合特征经过一个卷积块处理进行进一步特征精化和信息整合,生成最终特征。

一方面,DCNv2的引入增强了C2f模块对特征的提取和处理能力,使其能够更好地应对各种复杂情况;另一方面,C2f模块的结构特点有助于更好地发挥DCNv2的优势,例如通过特征的分流和融合,使得可形变卷积学习到的特征能够在不同分支之间进行交互和整合,进一步优化特征表示,从而提升整个网络的性能和效率,使网络在各种视觉任务中表现更出色。通过在Backbone部分嵌入多个DCNFE,模型可以有效学习复杂多变的各类金具目标的形态表达,从而提高金具目标的定位精度,生成更加准确的金具目标检测结果。

1.4　Wise-IoU边界框损失函数

边界框损失函数作为目标检测模型损失函数的重要组成部分,其直接决定了模型的训练效果和最终的预测性能。传统边界框损失函数单纯强化边界框对低质量示例的回归,会导致模型性能下降。对此,Wise-IoU损失函数使用离群度替代IoU评估锚框质量,采用梯度增益分配策略进行动态调整[31]。对于高质量锚框,降低其梯度增益,避免模型过度关注这些样本;对于低质量锚框,分配较小的梯度增益,减少其对训练的负面影响,从而平衡不同质量锚框对模型学习的贡献。CIoU虽然考虑了边界框的重叠面积、中心点距离和宽高比等几何因素,但在某些复杂场景下对目标的定位精度仍有待提高。例如,当目标物体存在遮挡、变形或尺度变化较大时,CIoU可能无法准确地反映预测框与真实框之间的差异。Wise-IoU通过动态调整梯度增益,能够更灵活地适应不同质量的样本,使模型在训练过程中更加关注那些对定位精度有重要影响的样本,从而提高目标的定位精度。例如,在处理小目标或形状不规则的目标时,Wise-IoU可以给予这些目标更多的关注,减少定位误差。基于此,本文采用先进的Wise-IoU损失函数替代常规损失函数CIoU,指导整个网络模型的高效学习训练,通过动态非单调聚焦机制,增强模型的样本学习能力。动态非单调聚焦机制使具有最高梯度增益的边界框条件处于动态变化中,并用离群度描述锚框质量,其定义如下:

$$\beta=\frac{L_{\mathrm{IoU}}^{*}}{\bar{L}_{\mathrm{IoU}}}\in[0,+\infty)\tag{1}$$

式中,L_{IoU}^{*}、$\bar{L}_{\mathrm{IoU}}$分别表示具有动量m的指数移动值、平均值;β表示离群度。离群度小意味着锚框质量高,分配的梯度增益小,以使边界框回归聚焦到普通质量的锚框上,对离群度较大的锚框分配较大的梯度增益,利用β构造一个非单调聚焦系数,并将其应用到$L_{\text{Wise-IoUv1}}$上:

$$\beta L_{\text{Wise-IoUv3}}=\gamma L_{\text{Wise-IoUv1}},\gamma=\frac{\beta}{\delta\alpha^{\beta-\alpha}}\tag{2}$$

式中,γ表示梯度增益,由超参数α和δ控制;$L_{\text{Wise-IoUv1}}$表示边界框IoU为某一指定值时,最高增益的静态聚焦机制。

2 实验结果与分析

2.1 数据集

本文将大疆御3无人机搭载的可见光相机实地拍摄汕头±160 kV青汇线、±160 kV塑汇线、110 kV莱雄线、110 kV湾金线、220 kV汕红甲乙线和500 kV胪祯甲乙线的架空输电线路金具影像作为数据源,构建金具目标检测数据集,对本文提出的方法进行实验测试。该无人机具备长续航、高精度定位、专业级成像传感器、全向避障及远距离高清图传功能,裸机重量(包含螺旋桨和实时动态技术模块)约为951 g,最大起飞重量为1 050 g,最大起飞海拔高度可达6 000 m(空载飞行),在普通挡位下,最大水平飞行速度为15 m/s(54 km/h);运动挡位下,向前飞行速度可达21 m/s(75.6 km/h)。在无风环境下,最长飞行时间可达43 min,最长悬停时间约为37 min,最大续航里程可达32 km。配备4/3英寸大底CMOS传感器,有效像素为2 000万,24 mm等效焦距,f/2.8至f/11可变光圈,支持机械快门(最高1/2 000 s)和RAW格式拍摄,满足高质量正射影像和倾斜摄影需求。该传感器获取的高分辨率遥感影像包含RGB三波段影像,大小为5 280×3 956像素。首先针对无人机巡检拍摄到的输电线路航拍图像进行筛选,去除不包含金具目标或金具目标在图像中占比极小的图像。然后对所有图像进行重命名、格式转换和批量入库的操作,即对筛选后的所有图像批量重命名并转为JPEG格式,保证来自不同巡检线的巡检图像格式统一,便于后期标注。实验使用的数据集包括连接金具、板状金具、扇形金具、三角金具、短连接金具和长形金具6类金具目标,共516张图像,1 745个标签。与公共数据集图像不同,航拍输电线路图像具有背景复杂、金具目标形态多样、尺度差异大、密集程度高、遮挡现象严重、目标占比小等特点,其中包含存在锈蚀、损坏等缺陷的金具样本。蓝天背景的样本116个,云背景的样本646个,建筑物背景的样本495个,自然环境背景的样本488个。

本文选择各类金具的典型场景构建金具目标数据集,如图4所示。利用LabelImg平台,通过目视解译,对影像中的金具进行目标级标注,得到真实标签数据。在影像中为每个金具目标绘制边界框,边界框以金具位置作为中心坐标,依据金具的形状大小确定边界框的宽和高,同时按照金具常见形态结构对金具目标进行类别划分,为其分配类别标签。根据巡检需求,本文主要关注6类常见金具,分别为连接金具(边界框用绿色表示)、板状金具(边界框用浅粉色表示)、扇形金具(边界框用藕粉色表示)、三角金具(边界框用紫色表示)、短连接金具(边界框用桃粉色表示)和长形金具(边界框用黄色表示),生成最终的YOLO格式标签文件。其格式如下:〈object-class〉〈x_center〉〈y_center〉〈width〉〈height〉,其中object-class是目标类别索引,其他4个参数是边界框的位置信息,如图5所示。基于构建的金具目标检测数据集,按照8:1:1将其划分训练集、验证集、测试集,其样本分布如表1所示。

2.2 实验设置

本文提出的YOLOv8-SADC模型和使用的对比方法Faster R-CNN[32]、SSD[33]、YOLOR[34]、YOLOv5、YOLOv8、YOLOv10[35]、LiteYOLO-ID[36]、IDD-YOLO[37]均采用随机梯度下降优化器对损失函数进行优化,初始学习率设置为0.01,训练批次为16。训练周期上限设为300,并使用余弦退火调度策略动态调整学习率。在连续10个周期精度不提升时,启动早停法终止训练。

模型的实现基于PyTorch框架[38],所有实验在NVIDIA RTX 3090 GPU服务器上进行,其显存为24 GB。为保证模型的充分训练,提高模型的泛化能力,本文对数据集进行了多种数据增强操作,包括随机水平翻转、颜色抖动、缩放、随机裁剪以及马赛克数据增强。

2.3 评价指标

为综合评估金具目标检测效果,本文采用目视效果和定量评价两种策略。其中,定量评价使用目标识别通用指标,即精确率(P)、召回率(R)和单类平均精确率(average precision,AP)和类平均精度(mean average precision,mAP)[39-40],计算公式如下:

$$P=\frac{N_{\mathrm{TP}}}{N_{\mathrm{TP}}+N_{\mathrm{FP}}}\times 100\% \tag{3}$$

$$R=\frac{N_{\mathrm{TP}}}{N_{\mathrm{TP}}+N_{\mathrm{FN}}}\times 100\% \tag{4}$$

$$P_{\mathrm{mAP}}=\sum_{i=1}^{C}P_{\mathrm{AP}_i}\Big/C,P_{\mathrm{AP}_i}=\int_0^1 P(R_i)\mathrm{d}R_i \tag{5}$$

式中,TP(true positive)、FN(false negative)、FP(false positive)、TN(true negative)由混淆矩阵求得,分别表示真实为正且预测为正、真实为正但预测为负、真实为负但预测为正和真实为负且预测为负的样本;N表示样本数。P-R曲线下的面积即为AP,用来衡量单类别的



图 4　金具目标数据集

Fig. 4　Hardware Target Dataset

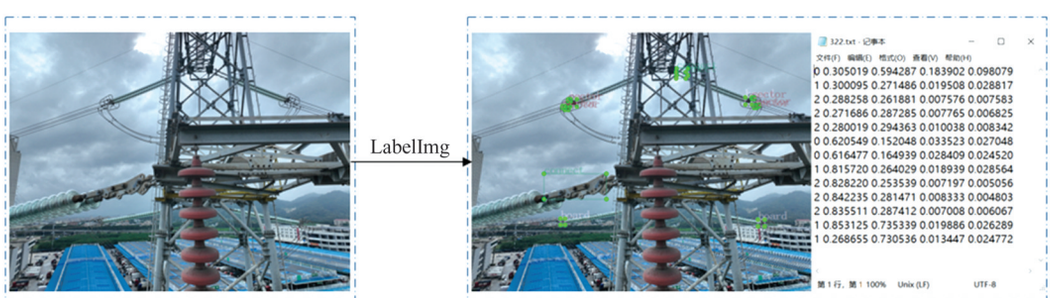


图 5　数据集标注结果

Fig. 5　Dataset Annotation Result

表 1　金具目标检测数据集样本分布

Table 1　Sample Distribution of the Hardware Object Detection Dataset

类型	训练集/个	验证集/个	测试集/个
连接金具	278	49	40
板状金具	159	27	20
扇形金具	328	44	50
三角金具	643	66	72
短连接金具	73	9	14
长形金具	62	7	5
总体	1 745	202	201

模型平均精确度,mAP将多个类别的AP累加后求平均。mAP50和mAP50-95是目标检测中两种常用指标,其中,mAP50指在IoU为0.5的阈值下的mAP,mAP50-95指IoU从0.5～0.95多阈值下按照一定步长计算得到的mAP,能更全面反映模型在不同匹配条件下的检测性能。

2.4　识别结果比较

本文进一步在公开数据集上评估了YOLOv8-SADC的性能,并与自建的输电线路金具目标数据集进行了对比分析。与公共金具数据集相比,本文构建的输电线路金具目标数据集具有背景复杂、金具目标形态多样、尺度差异大、密集程度高、遮挡现象严重、目标占比小等特点。CPLID数据集[41]包含无人机拍摄的正常绝缘子图像600张,只包含绝缘子一个目标类别;2511数

据集使用LabelImg标注,以蓝色垫子为背景,对绝缘子、螺栓销钉、防振锤、耐张线夹4类金具进行多角度拍摄的数据集,共2 511张图片。本文分别在CPLID、2511和自建金具数据集上进行了性能比较,见表2,其中每个数据集包括YOLOv8和YOLOv8-SADC两种模型训练的结果。

表2 3种数据集的实验结果/%

Table 2 Experimental Results on Three Datasets/%

数据集	训练模型	*P*	*R*	mAP50	mAP50-95
CPLID	YOLOv8	90.23	95.33	95.83	82.26
	YOLOv8-SADC	91.90	93.50	96.66	80.13
2511	YOLOv8	96.40	96.40	97.03	69.60
	YOLOv8-SADC	96.63	96.40	97.46	69.66
YOLOv8-SADC数据集	YOLOv8	86.81	75.98	80.54	51.90
	YOLOv8-SADC	86.50	79.96	84.40	56.10

公开数据集CPLID仅检测绝缘子这一单一目标,目标特征较为单一;2511数据集虽然包含4种金具目标类型,但其背景为单一的蓝色垫子,环境相对简单,目标间的区分难度较低;本文自建的数据集复杂性更高,包括背景复杂、金具目标形态多样、尺度差异大、密集程度高、遮挡现象严重以及目标占比小等特点。表2实验结果表明,YOLOv8-SADC在CPLID和2511数据集上的性能表现优异,能够充分适应简单背景和单一目标类别的场景,mAP达到较高水平。然而,在本文复杂背景和多样目标的自建数据集上,YOLOv8-SADC模型的优势更加突出,表现出显著优于传统目标检测模型的检测能力,尤其是在小目标检测、密集目标区域识别和遮挡目标定位等方面,检测精度的提升幅度更加显著。

表2对比结果充分表明,YOLOv8-SADC模型不仅在公开数据集的常规目标检测任务中具有较强的适用性,还在复杂环境中的高难度检测任务中表现出了优越的性能,为输电线路金具目标检测提供了更加可靠的解决方案。

为验证本文方法的有效性,选择6种代表性目标检测模型作为对比方法,包括Faster R-CNN、SSD、YOLOR、YOLOv5、YOLOv8、YOLOv10和绝缘子检测网络Lite-YOLO-ID、IDD-YOLO,在构建的金具目标检测数据集上进行性能比较,如表3所示。所有方法采用源代码,参数设置采用推荐设置。从表3可以看出,本文提出的YOLOv8-SADC模型取得了最优的定量评价结果,表明架空输电线路金具目标识别结果中存在较少的漏检,且对各类金具识别均取得较好结果。相较于其他方法,YOLOv8-SADC模型具有明显优势。由于金具目标在整个场景中占比较少,因此在准确率方面,本文方法优势并不明显。但在考虑了类别不均衡性的综合指标mAP方面,本文方法相比于其他方法体现出明显的优势,表明本文方法能较好地适应复杂场景金具目标检测任务,可以较好地满足金具识别巡检需求。

表3 各模型金具目标识别结果定量评价/%

Table 3 Quantitative Evaluation of Hardware-Object Detection Results by Various Models/%

模型	mAP	AP					
	50	连接	板状	扇形	三角	短连接	长形
Faster R-CNN	64.20	71.36	78.70	34.05	76.02	64.71	60.28
SSD	40.57	44.71	41.02	13.10	49.73	35.18	59.90
Lite-YOLO-ID	66.93	69.20	81.60	31.07	71.97	75.67	72.07
IDD-YOLO	62.60	72.00	71.90	23.10	71.70	71.50	65.20
YOLOR	78.75	73.52	86.21	51.03	89.33	94.55	77.92
YOLOv5	80.26	88.14	88.44	53.08	84.60	79.46	87.90
YOLOv8	80.54	88.08	83.40	53.54	84.30	81.78	92.10
YOLOv10	77.92	83.30	86.38	47.84	82.50	77.28	90.28
YOLOv8-SADC	84.40	90.52	88.88	57.42	87.84	88.26	93.52

从表3也可看出,Faster R-CNN、SSD、Lite-YOLO-ID和IDD-YOLO在该数据集上表现较差,绝缘子通常具有较为统一的特征形态,Lite-YOLO-ID和IDD-YOLO针对绝缘子缺陷进行了优化,但金具目标包括多种形态和功能,外观复杂且差异显著,复杂的自然背景也会干扰模型对金具的准确识别,导致大量漏检,其mAP50仅为65%左右。相比之下,YOLO系列模型通过小样本学习显著提升了金具检测精度,其中,YOLOR和YOLOv10取得了较好的检测效果,mAP相较于LiteYOLO-ID和IDD-YOLO模型均有大幅提升,表明其可以有效减少各类金具的错检和漏检问题。尤其是YOLOR在三角金具和短连接金具两个类别取得了最优平均精度。YOLOv5和YOLOv8模型进一步提升了模型的检测性能,mPA50均超过80%,表明其能够较好适应各类金具目标的检测。然而,这两种方法仍然存在较多漏检,在一定程度上限制了模型的目标检测性能。本文提出的YOLOv8-SADC模型有效克服了

上述问题,进一步提升了各类金具的检测精度,mPA50达到84.40%,表明该方法对各类金具变化具有较好的鲁棒性,能够有效检测各类金具。其中连接金具的mAP50达到90.52%,长形金具的mAP50达到93.52%。对于金具目标较小、且存在大量遮挡的扇形金具,同样提升了检测精度,相较于YOLOv8,YOLOv8-SADC提高了近4%,相较于其他方法,优势更加明显,证明了本文方法在复杂场景金具目标检测中的有效性。

图6显示了本文提出的YOLOv8-SADC模型在金具目标数据集上16个典型场景中的可视化结果。从图6可以看出,在各场景中,本文方法均表现出较好的检测性能,能够更准确地区分金具目标和非金具背景,误检和漏检较少,对不同类别金具具有较好的区分度和检测效果,进一步表明了本文方法的鲁棒性。

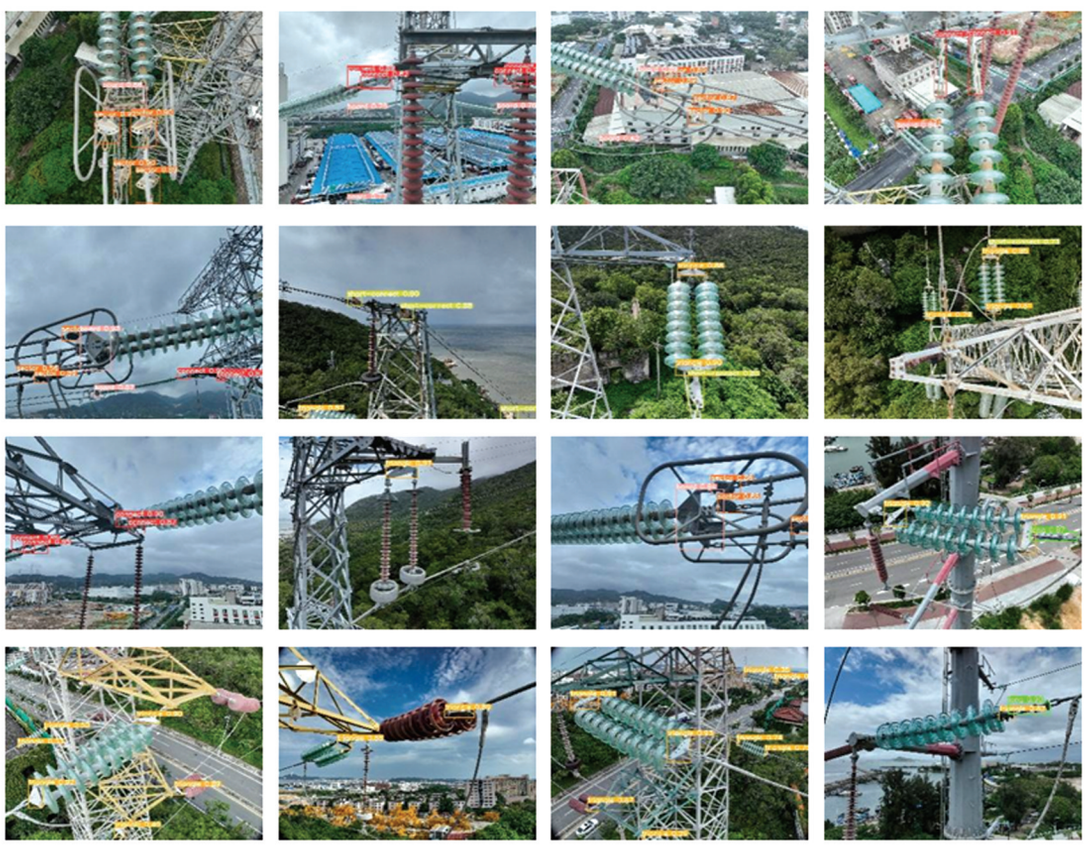


图6　YOLOv8-SADC金具目标检测的可视化结果

Fig. 6　Visualization Results of Hardware Target Detection by YOLOv8-SADC

为了全面评估YOLOv8-SADC在输电线路金具检测任务中的性能,本文对不同背景下的样本检测精度进行了实验分析。通过将数据集分为蓝天、多云、建筑物和自然环境4种背景,并对每种背景下的数据进行独立实验,进一步探讨了算法的适用性和局限性,实验结果如表4所示。从表4可以看出,模型在蓝天背景下的检测性能最佳,mAP50达到97.37%。由于蓝天背景干扰少,金具目标特征突出,易于提取关键信息。数据集中的自然环境背景绝大多数是树木、草地,检测精度略有下降,mAP50为84.67%。多云背景下的mAP50为78.43%。这是由于云层纹理可能引入一定的背景干扰,导致模型的特征提取效率降低。在建筑物背景下,模型检测性能进一步下降,mAP50为76.63%。建筑物的结构特征(如直线、边缘等)与金具的形态存在一定相似性,且建筑物背景复杂多样,包含各种形状、颜色

表4　不同背景的实验结果/%

Table 4　Experimental Results Under Different Backgrounds/%

背景	训练模型	*P*	*R*	mAP50	mAP50-95
蓝天	YOLOv8	94.63	91.17	96.17	66.83
	YOLOv8-SADC	94.70	91.23	97.37	69.83
多云	YOLOv8	86.73	70.40	77.00	41.60
	YOLOv8-SADC	87.33	69.63	78.43	42.10
建筑物	YOLOv8	76.07	65.03	71.33	41.90
	YOLOv8-SADC	86.43	67.57	76.63	49.33
自然环境	YOLOv8	85.20	75.70	80.90	54.40
	YOLOv8-SADC	85.00	79.36	84.67	53.97

和纹理,以及可能存在的阴影、遮挡等情况,容易导致误检。实验结果表明,YOLOv8-SADC在不同背景下的性能提升程度有所不同,本文引入注意力来提高复杂特征的提取能力,自然环境和建筑物背景相对复杂,性能提升较为显著。在简单背景(例如蓝天和多云背景)中,检测精度明显较高,但是模型对金具目标的检测能力提升有限。

综上所述,本文方法在该数据集中取得了最优的金具目标检测结果,能够适应复杂场景电力设施金具目标检测,具有较大的应用潜力。

2.5 消融实验

本文方法以YOLOv8网络为骨干,通过耦合SA和DCNFE提高模型的特征提取和复杂目标表示能力,并采用Wise-IoU指导模型训练,提高模型的样本学习能力。为验证本文所提各网络模块的有效性,通过消融实验,对不同网络配置在构建数据集上进行金具目标检测性能测试,实验结果如表5所示。

表5 消融实验结果/%

Table 5 Ablation Experiment Results/%

方法	*P*	*R*	mAP50	mAP50-95
Baseline	86.81	75.98	80.54	51.90
Baseline+SA	87.00	77.30	82.47	56.10
Baseline+SA+DCNFE	87.17	77.70	83.37	54.77
Baseline+SA+DCNFE+ Wise-IoU	86.50	79.96	84.40	56.10

从表5可以看出,通过耦合SA,大部分指标得到了明显提升,其中,mAP50-95从51.90%提升至56.10%,提升近4%;*P*和*R*分别提升约0.2%和2%,mAP50提高了约2%,说明通过耦合SA,模型可以有效捕获通道和像素间的依赖关系,提升空间特征的判别性。在此基础上,通过耦合DCNFE,大部分指标同样得到了一定的提升,其中,mAP50提升约1%,*P*提升了0.17%,*R*提升了0.4%,进一步引入Wise-IoU,大部分指标得到进一步的提升,其中,*R*提升了2.26%,mAP50和mAP50-95分别提升了1.03%和1.33%,说明Wise-IoU损失函数可以有效引导模型训练,平衡不同质量锚框贡献,提高模型对各类金具目标的定位能力。通过组合上述模块,本文方法取得了最优的精度,表明了所提方法各模块的有效性。

3 结 语

本文针对金具目标纹理特征多变、形态复杂等问题,基于无人机遥感平台,提出一种有效的耦合SA和可形变卷积的YOLOv8模型(YOLOv8-SADC),用于高分辨率可见光影像金具目标检测。首先,通过耦合SA建立通道和空间之间的相关性,增强网络模型空间特征提取能力。其次,构建可形变卷积特征提取模块,深入挖掘金具目标的复杂形态表示,提升模型对各类金具目标的判别能力。然后,采用Wies-IoU指导模型训练,增强模型的样本学习能力,提高模型对各类金具目标的定位能力。最后,利用无人机实地采集广东汕头沿海地区架空输电线路高分辨率可见光影像,构建金具目标检测数据集进行实验验证。结果表明,本文方法相较于其他网络模型具有明显优势,对复杂场景的适应能力更强,可以有效保证各类金具目标识别结果的完整性和细节信息,是一种有效的金具目标识别方法。

参 考 文 献

[1] 方苏, 李立学, 郑益慧, 等. 基于激光测距成像和图像处理的输电线路防护技术[J]. 电气自动化, 2017, 39(3): 6-8.
FANG Su, LI Lixue, ZHENG Yihui, et al. Protection Technology for Transmission Lines Based on Laser Range Imaging and Image Processing [J]. *Electrical Automation*, 2017, 39(3): 6-8.

[2] 王炜, 袁奇, 顾俊杰, 等. X射线无损探伤技术在检测输电线路压接金具中的应用[J]. 上海交通大学学报, 2018, 52(10): 1189-1194.
WANG Wei, YUAN Qi, GU Junjie, et al. Application of X-Ray Nondestructive Testing Technology in Detecting the Pressure Welding Fittings of Transmission Lines [J]. *Journal of Shanghai Jiao Tong University*, 2018, 52(10): 1189-1194.

[3] 张秋雁,杨忠,姜遇红,等. 一种基于CNN的航拍输电线路图像分类方法[J]. 应用科技,2019,46(6):41-45.
ZHANG Qiuyan, YANG Zhong, JIANG Yuhong, et al. CNN-Based Aerial Image Classification Method for Aerial Transmission Lines [J]. *Applied Science and Technology*, 2019, 46(6): 41-45.

[4] 赵文清, 程幸福, 赵振兵, 等. 注意力机制和Faster RCNN相结合的绝缘子识别[J]. 智能系统学报, 2020, 15(1): 92-98.
ZHAO Wenqing, CHENG Xingfu, ZHAO Zhenbing, et al. Insulator Recognition Based on Attention Mechanism and Faster RCNN[J]. *CAAI Transactions*

on Intelligent Systems, 2020, 15(1): 92-98.

[5] DENG C, WANG S W, HUANG Z, et al. Unmanned Aerial Vehicles for Power Line Inspection: A Cooperative Way in Platforms and Communications[J]. *Journal of Communications*, 2014, 9(9): 687-692.

[6] 王万国, 张晶晶, 韩军, 等. 基于无人机图像的输电线断股与异物缺陷检测方法[J]. 计算机应用, 2015, 35(8): 2404-2408.
WANG Wanguo, ZHANG Jingjing, HAN Jun, et al. Broken Strand and Foreign Body Fault Detection Method for Power Transmission Line Based on Unmanned Aerial Vehicle Image[J]. *Journal of Computer Applications*, 2015, 35(8): 2404-2408.

[7] 赵建坤, 王璋奇, 刘世钊. 基于灰度共生矩阵纹理特征的输电导线识别[J]. 云南电力技术, 2015, 43(2): 126-129.
ZHAO Jiankun, WANG Zhangqi, LIU Shizhao. Research on Transmission Line Recognition Based on GLCM Texture Feature[J]. *Yunnan Electric Power*, 2015, 43(2): 126-129.

[8] 王伟, 刘国海. 绝缘子图像的边缘检测[J]. 微计算机信息, 2008, 24(27): 308-309.
WANG Wei, LIU Guohai. Image Edge Detection of the Insulator[J]. *Microcomputer Information*, 2008, 24(27): 308-309.

[9] 陈晓娟, 吴英石, 赵亮. 基于随机Hough变换的OPGW防震锤识别[J]. 黑龙江电力, 2010, 32(1): 1-2.
CHEN Xiaojuan, WU Yingshi, ZHAO Liang. Identification of OPGW Vibration Damper Based on Random Hough Transformation[J]. *Heilongjiang Electric Power*, 2010, 32(1): 1-2.

[10] 宋伟, 左丹, 邓邦飞, 等. 高压输电线防震锤锈蚀缺陷检测[J]. 仪器仪表学报, 2016, 37(S1): 113-117.
SONG Wei, ZUO Dan, DENG Bangfei, et al. Corrosion Defect Detection of Earthquake Hammer for High Voltage Transmission Line[J]. *Chinese Journal of Scientific Instrument*, 2016, 37(S1): 113-117.

[11] 黄乐, 吴功平, 叶旭辉. 输电线巡检机器人弱光条件下的障碍物识别研究[J]. 光学学报, 2018, 38(9): 0915006.
HUANG Le, WU Gongping, YE Xuhui. Research on Obstacle Recognition for Transmission Line Inspection Robots Under Low-Light Conditions[J]. *Acta Optica Sinica*, 2018, 38(9): 0915006.

[12] 刘进, 杨洁, 庞瑞帆, 等. 一种基于多尺度特征簇的舰船目标快速定位与识别方法[J]. 武汉大学学报(信息科学版), 2016, 41(1): 111-116.
LIU Jin, YANG Jie, PANG Ruifan, et al. A Method for Fast Ship Detection and Recognition in Sea-Sky Background Based on Multi-scale Feature Cluster[J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(1): 111-116.

[13] 杨洪朝, 席燕辉, 项胜, 等. 基于深度学习的输电线路巡检无人机图像分类方法[J]. 湖南电力, 2024, 44(1): 53-58.
YANG Hongzhao, XI Yanhui, XIANG Sheng, et al. Image Classification Method Based on Deep Learning for Inspection Unmanned Aerial Vehiclesin Transmission Line[J]. *Hunan Electric Power*, 2024, 44(1): 53-58.

[14] 谢小瑜, 周俊煌, 张勇军. 深度学习在泛在电力物联网中的应用与挑战[J]. 电力自动化设备, 2020, 40(4): 77-87.
XIE Xiaoyu, ZHOU Junhuang, ZHANG Yongjun. Application and Challenge of Deep Learning in Ubiquitous Power Internet of Things[J]. *Electric Power Automation Equipment*, 2020, 40(4): 77-87.

[15] 杨罡, 孙昌雯, 张娜, 等. 基于多尺度特征融合的输电线路关键部件检测[J]. 电测与仪表, 2020, 57(3): 54-59.
YANG Gang, SUN Changwen, ZHANG Na, et al. Detection of Key Components of Transmission Lines Based on Multi-scale Feature Fusion[J]. *Electrical Measurement & Instrumentation*, 2020, 57(3): 54-59.

[16] ZHAO Z B, XU G Z, QI Y C, et al. Multi-Patch Deep Features for Power Line Insulator Status Classification from Aerial Images[C]//International Joint Conference on Neural Networks, Vancouver, BC, Canada, 2016.

[17] 赵振兵, 江爱雪, 戚银城, 等. 嵌入遮挡关系模块的SSD模型的输电线路图像金具检测[J]. 智能系统学报, 2020, 15(4): 656-662.
ZHAO Zhenbing, JIANG Aixue, QI Yincheng, et al. Fittings Detection in Transmission Line Images with SSD Model Embedded Occlusion Relation Module[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(4): 656-662.

[18] 赵振兵, 马迪雅, 丁洁涛, 等. 基于SAW-PCL的输电线路缺销螺栓弱监督检测方法[J]. 北京航空航天大学学报, 2024, 50(11): 3319-3326.
ZHAO Zhenbing, MA Diya, DING Jietao, et al. Weak Supervision Detection Method for Pin-Missing Bolts of Transmission Lines Based on SAW-PCL[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2024, 50(11): 3319-3326.

[19] ZHANG K, ZHAO K, GUO X W, et al. HRM-CenterNet: A High-Resolution Real-Time Fittings Detection Method[C]// IEEE International Conference on Systems, Man, and Cybernetics, Melbourne, Australia, 2022.

[20] WANG N, ZHANG K, ZHU J W, et al. Fittings

Detection Method Based on Multi-Scale Geometric Transformation and Attention-Masking Mechanism [J]. *Sensors*, 2023, 23(10): 4923.

[21] ZHANG K, LOU W S, WANG J C, et al. PA-DETR: End-to-End Visually Indistinguishable Bolt Defects Detection Method Based on Transmission Line Knowledge Reasoning[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 5016914.

[22] CHEN G Q, WANG S K, LIU Y K, et al. Research on Transmission Line Hardware Identification Based on Improved YOLOv5 and DeblurGANv2 [J]. *IEEE Access*, 2023, 11: 133351-133362.

[23] WANG S S, TAN W W, YANG T F, et al. High-Voltage Transmission Line Foreign Object and Power Component Defect Detection Based on Improved YOLOv5[J]. *Journal of Electrical Engineering & Technology*, 2024, 19(1): 851-866.

[24] 汤寓麟, 边少锋, 翟国君, 等. 侧扫声呐检测沉船目标的改进YOLOv5法[J]. 武汉大学学报(信息科学版), 2024, 49(6): 977-985.
TANG Yulin, BIAN Shaofeng, ZHAI Guojun, et al. An Improved YOLOv5 Method for Shipwreck Target Detection by Side-Scan Sonar Images [J]. *Geomatics and Information Science of Wuhan University*, 2024, 49(6): 977-985.

[25] ZHAO Z B, XIONG J, HAN Y, et al. Transmission Line Bolts and Their Defects Detection Method Based on Position Relationship[J]. *Frontiers in Energy Research*, 2023, 11: 1269087.

[26] 张瑞倩, 邵振峰, Aleksei Portnov, 等. 多尺度空洞卷积的无人机影像目标检测方法[J]. 武汉大学学报(信息科学版), 2020, 45(6): 895-903.
ZHANG Ruiqian, SHAO Zhenfeng, PORTNOV A, et al. Multi-scale Dilated Convolutional Neural Network for Object Detection in UAV Images [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(6): 895-903.

[27] 张宇, 何楚, 石博, 等. 逐层特征选择的多层部件模型用于遥感图像飞机目标检测[J]. 武汉大学学报(信息科学版), 2014, 39(12): 1406-1411.
ZHANG Yu, HE Chu, SHI Bo, et al. Multi-layer Feature Selection Based Hierarchal Component Model for Aero-Plane Detection on Remote Sensing Image [J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(12): 1406-1411.

[28] LI Y, FAN Q, HUANG H, et al. A Modified YOLOv8 Detection Network for UAV Aerial Image Recognition[J]. *Drones*, 2023, 7(5): 304.

[29] ZHANG Q L, YANG Y B. SA-Net: Shuffle Attention for Deep Convolutional Neural Networks[C]// IEEE International Conference on Acoustics, Speech and Signal Processing, Toronto, ON, Canada, 2021.

[30] ZHU X Z, HU H, LIN S, et al. Deformable ConvNets V2: More Deformable, Better Results[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2020.

[31] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism [EB/OL]. 2023: arXiv: 2301.10051, https://arxiv.org/abs/2301.10051.

[32] 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.

[33] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multibox Detector [C]//The 14th European Conference on Computer Vision, Amsterdam, The Netherlands, 2016.

[34] WANG C Y, YEH I H, LIAO H P. You Only Learn One Representation: Unified Network for Multiple Tasks[J]. *Journal of Information Science and Engineering*, 2021, 39: 691-709.

[35] WANG A, CHEN H, LIU L H, et al. YOLOv10: Real-Time End-to-End Object Detection[C]// The 38th Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 2024.

[36] LI D H, LU Y, GAO Q, et al. LiteYOLO-ID: A Lightweight Object Detection Network for Insulator Defect Detection [J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73: 5023812.

[37] LU Y, LI D H, LI D, et al. A Lightweight Insulator Defect Detection Model Based on Drone Images [J]. *Drones*, 2024, 8(9): 431.

[38] HUJ, SHENL, SUNG. Squeeze-and-Excitation Networks[C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018.

[39] PADILLA R, PASSOS W L, DIAS T L B, et al. A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit [J]. *Electronics*, 2021, 10(3): 279.

[40] REZATOFIGHI H, TSOI N, GWAK J, et al. Generalized Intersection over Union: A Metric and a Loss for Bounding Box Regression[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2020.

[41] RAIMUNDO A. Insulator Data Set — Chinese Power Line Insulator Dataset (CPLID)[DS/OL]. [2024-10-12] https://ieee-dataport.org/open0access/insulator-data-set-chinese-power-line-insulator-dataset-cplid.