

基于改进 YOLOv9 的电力设备红外 目标检测模型

贺强^{1*}, 刘洋¹, 游鑫¹, 张瑞亮¹, 王玉峰¹, 黄谦¹, 何雨非²

(1. 中国南方电网有限责任公司超高压输电公司贵阳局, 贵州 贵阳 550000;

2. 华北电力大学(保定) 电子与通信工程系, 河北 保定 071003)

摘要:在电力巡检领域, 电力设备红外图像检测的重要性不言而喻, 其准确性直接关系到电力系统的稳定运行与安全性。针对当前复杂场景下红外图像目标检测存在的识别精度不足、分类混淆以及漏检误检等问题, 该文提出了一种基于 YOLOv9 的改进检测方法。该方法首先通过引入空间到深度卷积层 (SPD-Conv), 实现了模型主干网络的轻量化设计, 有效提升了计算效率与实时性。随后, 结合卷积块注意力模块 (CBAM), 进一步优化了模型对关键特征的提取能力, 增强了检测的准确性。此外, 还采用 Focal-IoU 作为损失函数, 强化了模型的判别能力, 减少了定位误差。实验结果表明, 在南方电网电力设备红外数据集上, 该算法取得了显著的成效, mAP 值达到了 94.4%, 同时推理速度也达到了 59.5 f/s, 充分展示了该方法在检测精度与计算效率方面的双重优势。与现有方法相比, 该模型在各类电力设备上的检测能力均有所提升, 为电力巡检领域提供了一种更为可靠、高效的解决方案。

关键词:电力设备; 红外图像检测; YOLOv9 算法; 注意力机制; 轻量化模型; 损失函数

中图分类号: TP183

文献标识码: A

文章编号: 1673-629X(2025)07-0182-08

doi: 10.20165/j.cnki.ISSN1673-629X.2025.0112

Target Detection Model for Infrared Image of Power Equipment Based on Improved YOLOv9

HE Qiang^{1*}, LIU Yang¹, YOU Xin¹, ZHANG Rui-liang¹, WANG Yu-feng¹,

HUANG Qian¹, HE Yu-fei²

(1. Guiyang Bureau of China Southern Power Grid Co., Ltd., Ultra High Voltage Transmission Company,
Guiyang 550000, China;

2. Department of Electronic and Communication Engineering, North China Electric Power University
(Baoding), Baoding 071003, China)

Abstract: In the field of power inspection, the importance of infrared image detection of power equipment is self-evident, and its accuracy is directly related to the stable operation and safety of the power system. In order to solve the problems of insufficient recognition accuracy, classification confusion and missed detection and false detection in infrared image target detection in complex scenes, an improved detection method based on YOLOv9 is proposed. Firstly, by introducing the space-to-deep convolutional layer (SPD-Conv), the lightweight design of the model backbone network is realized, which effectively improves the computational efficiency and real-time performance. Subsequently, combined with the Convolutional Block Attention Module (CBAM), the model's ability to extract key features was further optimized, and the accuracy of detection was enhanced. In addition, Focal-IoU is used as the loss function, which strengthens the discriminant ability of the model and reduces the positioning error. The experimental results show that the proposed algorithm has achieved remarkable results on the infrared dataset of power equipment of China Southern Power Grid, with the mAP value reaching 94.4% and the inference speed reaching 59.5 f/s, which fully demonstrates the dual advantages of the proposed method in terms of detection accuracy and computational efficiency. Compared with the existing methods, the detection ability of the proposed model on various power equipment has been improved, which provides a more reliable and efficient solution for the field of power inspection.

Key words: power equipment; infrared target detection; YOLOv9; attention mechanism; light-weight model; loss function

0 引言

维持电力设备的安全稳定是电力系统运行维护的重要任务,需要及时监控电力设备的运行状态,从而有效防止由于设备缺陷及故障所导致的安全事故问题。可见光图像能够清晰直观地呈现出电力设备的形态特征,便于工作人员快速识别。然而,可见光图像对环境有一定的要求,如光线及天气因素都会对可见光图像的成像质量造成影响。红外热成像技术则具有全天候监测的特点,在夜间或恶劣的环境下,红外图像可以提供较好的目标检测及识别能力,具有较好的鲁棒性。同时,红外图像对于热源目标敏感,可以检测电力设备的温度信息,对于设备发热故障有着较好的监测作用。因此基于红外图像的目标检测已经成为当前电力视觉研究的热点,而如何对红外图像中电力设备进行精确的目标定位成为了电力智能巡检的关键问题。

目前针对电力设备目标检测的计算机视觉技术则主要分为两类:一类是基于图像处理的检测方法,另一类则是基于深度学习的检测方法。基于图像处理的检测方法主要从图像中提取与目标检测相关的关键特征,如形状、颜色、纹理等,随后进行相应特征的检测定位。曾军等^[1]提出了结合改进的K-means算法,实现了对电力图像的有效分割和设备定位。郭文诚等^[2]通过提取Zernlike矩特征,利用相关向量机分类器来实现电力设备的分类识别。而随着深度学习技术的兴起,其在电力巡检中的应用也逐渐受到关注。刘子全等^[3]在Mask-RCNN方法的基础上进行了改进,应用于红外图像中电力设备及其部件的语义分割,从而有效提高了模型检测的精度。

深度学习技术的发展为电力设备目标检测提供了新的思路。深度学习方法不仅提高了精度,同时也大大节省了算法的运行时间。因此,如何更好地提升算法的精度和速度则成为了研究者们的主要方向。在业界上,YOLO算法因其高效且精准的效果得到了广泛的应用。郑含博等^[4]运用改进YOLOv3模型,实现了对电力设备红外图像的目标检测及快速定位。舒朗等^[5]提出Dense-YOLOv5的网络模型,针对红外目标尤其是特征不明显的小目标的检测进行了优化。Huang X^[5]、高伟^[6]采用了PP-YOLOv2模型,在模型的检测速度和精度上实现了巨大的优化。王小栋等^[7]则针对有遮挡的电力红外图像,提出改进YOLOv5模型,在提升模型精度的同时,实现了网络模型的轻量化。陈佳^[8]、杨立新等^[9]通过在YOLOv7上增加注意力机制,并改进损失函数,使得算法在电力设备图像分类检测上取得了更加优异的性能。李坤等^[10]在YOLOv8n的基础上,添加坐标注意力CA与可变形卷积网络替换残差模块中的标准卷积,提高了电力设备

红外图像识别的精度,并在一定程度上减少了计算量。Wang等^[11]提出YOLOv9模型,使得YOLO模型在精确率、速率及轻量化上有了显著的提升。该文章将YOLOv9模型运用到电力视觉领域,并对其进行了一定的改进,使其更加适配于电力设备红外图像的目标检测与定位。

该文在研究过程中提出如下几个问题:YOLOv9模型的训练往往需要大量的训练样本,然而对电力设备进行红外数据采集难度较大,现有数据多以小样本数据集为主;YOLOv9原始模型是根据自然图像公开数据集进行训练得到的,而红外图像与自然图像在成像原理上有较大差异,原始模型的检测效果需要优化。针对以上几个问题,首先针对南方电网电力设备红外图像数据集进行一定的图像增强处理,扩充数据集。在模型优化上,融合注意力机制,优化主干网络,并运用新的损失函数,提高模型在电力设备红外图像上的识别能力。

1 电力设备红外图像检测模型

1.1 YOLOv9模型

YOLOv9模型由三部分构成:骨干网、颈部网、检测头。骨干网核心为ResNSCPELAN4,融合CSP、RepVGG、ELAN模块,分别通过跨阶段连接减计算、多分支结构增强特征学习、自适应加权关注难识别目标,提升准确率和效率。颈部网连接骨干与检测头,融合并增强特征,含卷积、池化层,用FPN、PAN及注意力机制提升多尺度特征利用和检测的能力。检测头则基于骨干和颈部网络的特征进行目标检测。

1.2 改进的YOLOv9模型

YOLOv9模型能够高速、高精度、多尺度地检测图像目标,但红外图像因角度多变、温差小、噪声多、背景复杂等因素,使得检测难度增大。针对上述问题,提出的改进方案如下:(1)用SPD-Conv替换传统层,优化小物体及低分辨率处理,减参数不降精度;(2)引入CBAM注意力机制,增强特征表达,关注重要特征提取精度;(3)用Focaler-IoU替代原损失函数,提升检测性能。改进后的模型如图1所示。

1.2.1 SPD-Conv空间深度转换卷积

自AlexNet以来,卷积神经网络(CNN)^[12]在计算机视觉任务中表现出色。但处理电力设备红外图像时,因数据集分辨率低且目标小,传统CNN易丢失信息,影响模型精度。该文提出用SPD-Conv^[13]改进YOLOv9的CNN结构,保留细粒度信息,提高特征提取能力,减少计算量。

SPD-Conv由SPD层和无步长卷积层组成,SPD-Conv结构如图2所示。图中:

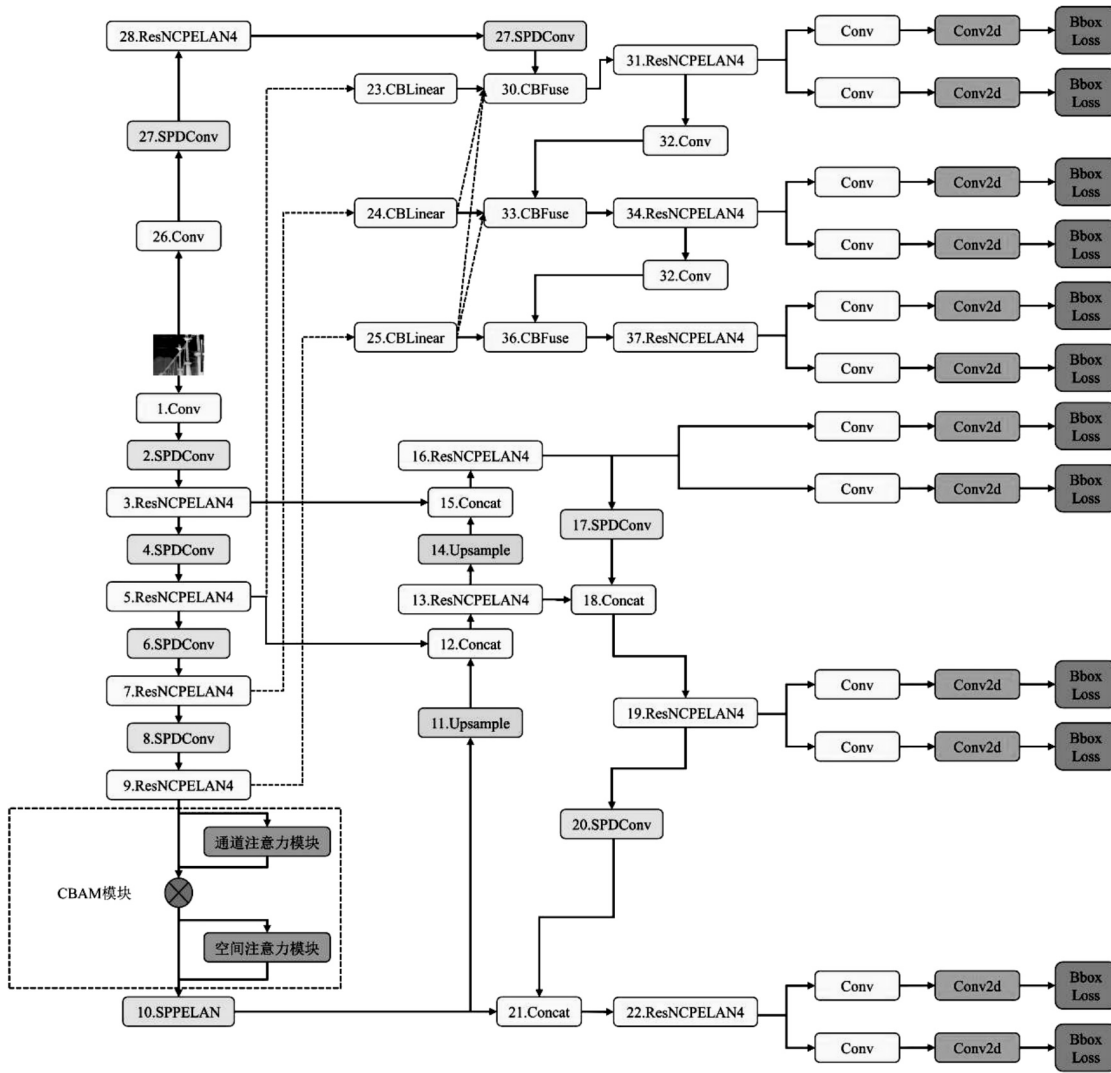


图 1 改进的 YOLOv9 网络结构

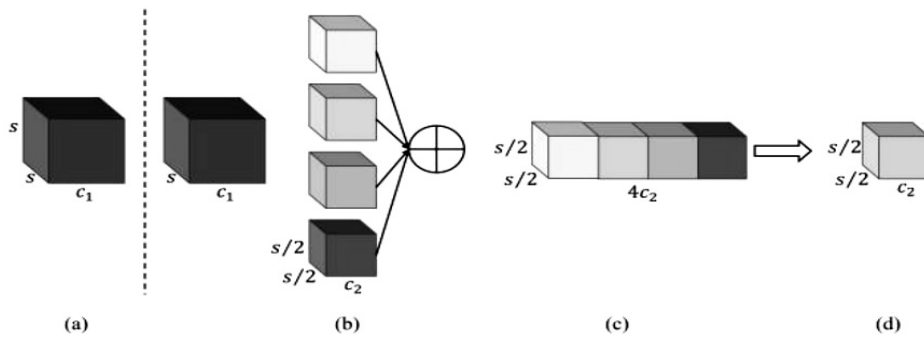


图 2 SPD-Conv 图解 (scale = 2)

(a) 传统特征图, 通道数为 c_1 , 且具有高度及宽度。

(b) 将 (a) 中特征图进行切片, 得到子特征图 $f_{0,0}, f_{scale-1,0}, f_{0,scale-1}, f_{scale-1,scale-1}$, 由于图 2 中 $scale = 2$, 故得到四个子特征图, 且每个子特征图的空间维度缩小为原来的 $1/2$ 。

(c) 将得到的子特征图在通道维度上进行合并, 此时获得的特征图通道数为 $4c_2$, 空间维度为原来的

一半。

(d) 对 (c) 中特征图进行无跨步卷积 (步长为 1, 卷积核的大小为 $1 \times 1 \times c_2$), 无跨步卷积可以尽可能地保留所有的判别特征信息, 最终将大小为 (s, s, c_1) 的特征图输出大小为 $(s/2, s/2, c_2)$ 的特征图。

通过 SPD-Conv 层后, 原特征图的特征信息获得了尽可能的保留, 且对特征图大小进行了优化, 可以进一步减少模型的参数量, 提升模型的效率。

1.2.2 CBAM 注意力模块

在 YOLOv9 模型中,采用通道注意力机制(Squeeze and Excitation Network, SENet)为各种深层卷积架构带来性能提升^[14]。然而,SENet 通常只考虑通道内部信息,忽略了空间位置对特征提取的重要性。与一般图像相比,复杂电力设备图像场景下红外目标的检测对空间位置预测的要求更高^[15]。CBAM(Convolutional Block Attention Module)注意力机制在学习通道位置权重的同时,引入了空间注意力的模块,该模块通过学习每个空间位置的重要性权重,使得网络可以关注当前对于某目标检测任务最重要的区域,增强关键区域的特征表达能力。通过结合空间注意力机制与通道注意力机制,使得模型能够更全面地捕捉图像中的关键特征,从而提高模型的整体表现力,同时引入空间注意力可以更好地区分目标与复杂背景,使得模型有更好的鲁棒性。引入注意力机制通常会增加一定的计算量,然而空间注意力与通道注意力的相互结合使得模型更加聚焦于关键特征,减少对非关键区域的关注,从而在一定程度上均衡了计算资源。CBAM 模块结构如图 3 所示。

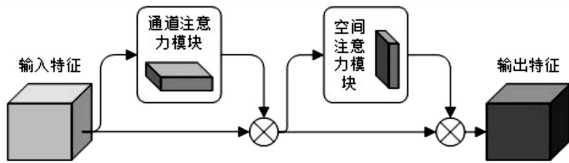


图 3 CBAM 模块结构

若给定输入特征图 $F \in R^{C \times H \times W}$,通过 CBAM 依次推断出二维空间注意力图 $M_s \in R^{1 \times H \times W}$ 和一维通道注意力图 $M_c \in R^{C \times 1 \times 1}$ 。如图 3 所示,整体过程可表示为:

$$F' = M_c(F) \otimes F \quad (1)$$

$$F'' = M_s(F') \otimes F' \quad (2)$$

其中, \otimes 表示逐元素乘法。在乘法过程中,注意力特征图内权值被相应地广播,通道注意值随空间维度传播,而空间注意值随通道维度传播。

(1)通道注意力模块。

通道注意力模块结构示意图如图 4 所示。

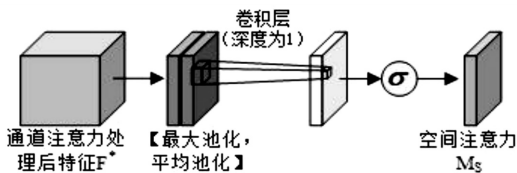


图 4 通道注意力模块结构示意图

通过利用特征通道间的关系,生成一个通道注意力图,一般一个通道被认为是一个特征检测器^[16],通道注意图的目的就是让模型将注意力集中至相应任务中更需要关注的特征。通道注意图的生成方式可阐述为:首先通过平均池化以及最大池化对每个通道的空

间维度进行压缩得到不同空间的描述符 F_{avg}^c 与 F_{max}^c ,对于空间信息的聚合,通常会采用平均池化,而最大值同样是空间信息的另一个重要表征,故 CBAM 采用了二者兼备的方式;之后将两个描述符传入共享网络,并通过该网络生成通道注意力图 M_c ,该共享网络由多层感知器(MLP)和一个隐藏层组成。为了节约算力,CBAM 将每个隐藏层的激活输出大小设置为 $R_{c/r} \times 1 \times 1$,其中 R_c 是输入特征图的通道数, r 则是降维比例。每个通道描述符在通过共享网络后,最终得到通道注意力图,通道注意力可表示为:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) = \sigma(w_1(w_0(F_{avg}^c)) + w_1(w_0(F_{max}^c))) \quad (3)$$

其中, σ 表示 sigmoid 函数, $w_0 \in R_{c/r} \times C$, $w_1 \in R_c \times C/r$ 。 w_1 与 w_0 是 MLP 网络中两个共享的权重。

(2)空间注意力模块。

空间注意力模块结构示意图如图 5 所示。

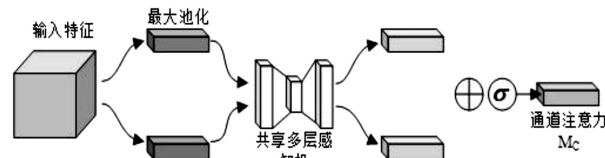


图 5 空间注意力模块结构示意图

通过利用特征图空间位置关系,生成空间注意力图。通过空间注意力图的目的是为了让模型将注意力集中至相应检测任务中更需要注意的区域。为计算空间注意力,CBAM 模型沿通道轴应用平均池化及最大池化操作,从而生成一个空间特征描述符。具体步骤如下:首先通过平均池化操作和最大池化操作聚合特征图的通道信息,以得到两个 2D 特征图: $F_{avg}^s \in R^{1 \times H \times W}$ 与 $F_{max}^s \in R^{1 \times H \times W}$,即每个通道的平均池化特征与最大池化特征。之后通过卷积层对此特征图进行处理,生成空间注意力特征图,空间注意力可表示为:

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{Maxpool}(F)])) = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (4)$$

其中, σ 表示 sigmoid 函数, $f^{7 \times 7}$ 表示卷积核大小为 7×7 的卷积操作。

通过以上操作得到通道注意力图与空间注意力图,后将两个模块依次顺序排列组成 CBAM 注意力模块(见图 3)。

1.2.3 损失函数

在目标检测任务中,模型的损失函数通常由分类误差、边界框坐标预测误差、置信度误差、用于学习预测框与真实框的损失函数四部分组成^[17]。自建数据集往往存在正负样本与难易样本不均的问题,如在文中数据集中,避雷器、电压互感器与电流互感器轮廓特

征相近,使得分类任务较为困难;而隔离开关与避雷器则差异较大,容易分类。故在训练过程中,对于正负样本差异较小的样本,易出现分类错误。

针对以上问题,引入 FocalerIoU 损失函数,针对不同难度的回归样本,动态调整权重,使得回归器能够更加准确地预测目标边框的位置以及大小。FocalerIoU 是一种结合了 IoU (Intersection over Union) 与 Focal Loss^[17] 思想的边界框回归损失函数,通过引入一个可学习的关注因子来调整不同回归样本的权重。在模型训练过程中,关注因子会根据回归结果动态调整其参数,从而使回归器更加关注对定位精度影响较大的样本。针对目标不同的正负样本与难易样本,优化模型的判别能力,其计算过程如下:

(1) 最常见的目标检测评价标准 IoU^[18] 定义如下:

$$V_{IoU} = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (5)$$

其中, V_{IoU} 表示 IoU 的值, B 和 B^{gt} 分别表示预测框和真实框。

(2) 在 IoU 的基础上,一步步增加新的约束,提出 GIoU^[19]、DIoU^[20]、CIoU、EIoU^[21]、SIoU^[22] 等,该文所使用的 FocalerIoU 主要采用 Focaler Loss 与 CIoU, CIoU 定义如下:

$$V_{CIoU} = V_{IoU} - \frac{\rho^2(b, b^{gt})}{c^2} - \alpha v \quad (6)$$

$$\alpha = \frac{v}{(1 - V_{IoU}) + v} \quad (7)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (8)$$

其中, V_{CIoU} 表示 CIoU 的值, b 与 b^{gt} 分别表示预测框与真实框的中心点, ρ 为欧几里得距离, c 为 b 与 b^{gt} 之间最小围合框的对角线距离, w^{gt} 与 h^{gt} 为真实框的宽度和高度, w 和 h 为预测框的宽度和高度。

(3) 专注于不同的样本,从而专注于不同的应用场景,故采用线性区间映射方法来重建 IoU。重建的 IoU 表达式如下:

$$V_{IoUfocaler} = \begin{cases} 0, & V_{IoU} < d \\ \frac{V_{IoU} - d}{u - d}, & d \leq V_{IoU} \leq u \\ 1, & V_{IoU} > u \end{cases} \quad (9)$$

其中, $V_{IoUfocaler}$ 为重建的 IoU,其中 IoU 是通过(2)计算的 SIoU 初始值, $[d, u] \in [0, 1]$, IoUfocaler 可以通过调整式中 d 与 u 的值来适应于不同的回归样本,其损失定义如下:

$$L_{Focaler-IoU} = 1 - V_{IoUfocaler} \quad (10)$$

(4) 将式(9)应用于基于 IoU 的边界框回归损失

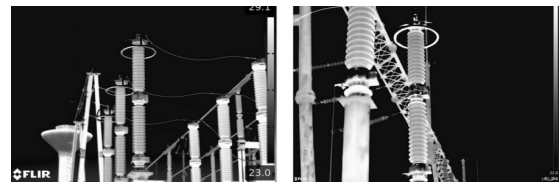
函数,可以得到最终的损失函数表达式。

$$L_{Focaler-SIoU} = L_{SIoU} + V_{IoU} - V_{IoUfocaler} \quad (11)$$

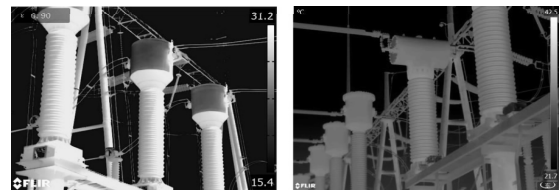
2 实验设计与结果分析

2.1 南方电网电力设备红外图像数据集

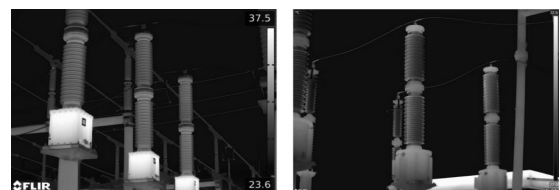
数据集来源于南方电网电力设备红外图像样本数据库,由运检部门历次采用 FLIR—P630 红外热像仪所采集的红外图像数据组成,包括避雷器、电流互感器、电压互感器、隔离开关四类数据,图像分辨率为 640×512。将四类图片数据进行筛选,得到总样本数为 400 幅,每幅图像中包含多个设备。由于该数据集样本数较小,无法充分满足工业领域检测模型的训练要求。为避免模型出现过拟合问题,使用翻转、旋转以及增加高斯噪声等图像增强方法对数据集进行了一定的扩充,提高模型的泛化能力及鲁棒性。数据增强后所得图像数量共 1 200 幅,包括避雷器、电流互感器、电压互感器与隔离开关。参考 PASCAL VOC 数据集的构建方法,对数据集进行划分,且训练集、验证集、测试集占比分别为 70%、20%、10%。图 6 展示了部分样本数据。



(a) 避雷器红外图像



(b) 电流互感器红外图像



(c) 电压互感器红外图像



(d) 隔离开关红外图像

图 6 部分样本图

2.2 实验环境与评价指标

该文采用的操作系统为 linux version 3. 10. 0 - 1160. e17. x86_64, 采用 CUD + A 12. 1 等环境搭载模

型,运用搭载2块GeForce GTX 4090显卡的服务器进行模型训练。利用平均精确度均值(mean Average Precision at 50% IoU, mAP50)以及每秒帧数(Frame Per Second, FPS)作为评估模型的定量指标。mAP常用于评价模型的准确性,从召回率以及准确率两个角度共同衡量模型的精确度;而FPS用来评判模型在单块4090 GPU上的推理速度。mAP50的计算过程如下所示:

计算精度(Precision)、召回率(Recall)定义如下:

$$S_{\text{Precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}} \times 100\% \quad (12)$$

$$S_{\text{Recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}} \times 100\% \quad (13)$$

其中, $S_{\text{Precision}}$ 和 S_{Recall} 表示 Precision 和 Recall, N_{TP} 、 N_{FP} 和 N_{FN} 表示 TP、FP 和 FN 的数量。

将 Precision 作为垂直轴, Recall 作为水平轴, 可以得到 P-R 曲线, 则:

$$S_{\text{AP}} = \int p(r) dr \quad (14)$$

其中, S_{AP} 表示精确度。

按照如上方法计算所有类的 AP, 最后取平均值即为 mAP(mean Average Precision)。

$$S_{\text{mAP}} = \frac{\sum_{i=1}^C S_{\text{AP}_i}}{C} \quad (15)$$

其中, S_{mAP} 为平均精确度, C 为检测的类别数, 文中 $C = 4$ 。

2.3 训练策略

在模型训练过程中发现, 不同模型以及不同优化方式所得模型的收敛速度略有差别。该文旨在评估不同模型在电力设备红外图像监测场景下的准确性以及推理速度, 故选取各模型最佳性能时的模型权重进行比较。训练策略中批大小定为 8, 优化器采用 SGD 优化器, 初始及截止学习率分别设置为 $1e-2$ 、 $1e-5$, 根据训练过程将训练 epoch 定为 600。随着迭代次数增加, 模型损失总体呈下降趋势, 且收敛性较好, 无明显振荡。

2.4 实验结果与分析

2.4.1 与主流电力设备红外图像检测方法对比

电力设备红外图像检测的主流识别方法一般为 Faster R-CNN^[23] 以及 YOLO 系列模型。为验证文中方法的有效性, 将其与主流识别方法进行对比, 同时选取部分面向电力设备的改进 YOLO 算法与该文提出的模型进行比较。如表 1 所示, 文中模型无论在准确性指标 mAP 还是预测速度指标 FPS 上, 都优于其他模型。同时与现有用于电力设备红外图像的目标检测算法 YOLOv7-C^[8]、YOLO-S-Superior^[9]、YOLOv8-

DCSW^[10] 相比, 文中模型在检测精度和推理速度 FPS 上是最高的, 同时参数量较小, 实现了检测精度、推理速度与模型轻量化上的平衡。而与基准模型相比, 由于增加了计算量, 推理速度略有减缓, 而在推理速度 FPS 仅有 0.2 的差异下, 模型精确度提升了 2.6%, 在满足推理速度的条件下提升了准确性, 满足当前电力设备红外图像检测对高精度高时效性的要求。同时, 文中模型在保证检测精度的同时, 在参数量上较基准模型减少 12.6%。

表 1 主流电力设备红外图像检测方法对比

模型	mAP50	FPS	参数量/M
Faster R-CNN	0.880	21	137.2
YOLOv5	0.892	29	31.5
PP-YOLOv2	0.907	31	54
YOLOv7 ^[24]	0.904	36	35.6
YOLOv7-C	0.916	36	31.4
YOLO-S-Superior	0.921	43	20.2
YOLOv8-DCSW	0.931	53.8	6.2
YOLOv9	0.918	59.7	25.3
OURS	0.944	59.5	22.1

2.4.2 消融实验

基于 YOLOv9 模型, 针对电力设备红外图像检测任务, 该文提出了三项改进策略: 一是使用 SPDCConv 替换原模型主干网络中的 CNN, 以增强多尺度特征提取能力; 二是在模型中增加 CBAM 注意力模块, 以提高模型对关键特征的关注度; 三是引入 FocalerIoU 损失函数, 以解决类别不平衡问题并提升检测精度。

表 2 消融实验测试结果

SPDCConv	CBAM	FocalerIoU	mAP50	FPS
			0.918	59.7
✓			0.924	59.6
	✓		0.933	59.5
		✓	0.933	59.7
✓	✓		0.936	59.5
✓		✓	0.926	59.6
	✓	✓	0.935	59.6
✓	✓	✓	0.944	59.5

为探究不同改进方法对原模型检测精度的提升, 设计了消融实验进行对照, 结果如表 2 所示。通过消融实验验证, 每项策略单独应用时均能在一定程度上提升模型性能。而当这些策略联合应用时, 模型的检测准确度得到了显著提升, 同时推理速度的下降非常小(仅 0.2 f/s)。这表明, 该文提出的优化方案在保持高效推理速度的同时, 有效提高了电力设备红外图像检测的精度, 能够适用于电力巡检部门在日常巡检过

程中对检测精度和速度均有较高要求的场景,有着较为可观的实用价值。

2.4.3 与基准方法对比

改进模型与原始模型检测所得 4 类设备的 mAP50 结果如表 3 所示,可进一步对比改进模型与基准模型的效果差异。通过实验结果可得,改进方法在对不同设备的检测效果上均优于原始模型。而对于避雷器、电流互感器与电压互感器这类轮廓近似的设备,改进模型也获得了较大的提升效果,说明改进算法可以较好地处理正负样本与难以样本不均的问题,相比基准模型能更好地应对电力设备红外图像检测中的复杂场景,提升了模型的准确性和鲁棒性。

表 3 不同设备检测结果

设备类别	YOLOv9	OURS
避雷器	0.874	0.928
电流互感器	0.925	0.936
电压互感器	0.919	0.927
隔离开关	0.952	0.984

在此基础上,为了更好地体现改进模型的优化效果,将模型的检测效果进行可视化,展示部分测试集推理结果,如图 7 所示。

通过检测结果可以发现,两种模型的检测精确度均较高,而对于较易样本如隔离开关的检测,改进模型的置信率处在更高的水平且较少发生错检漏检现象。但在较难样本时,基准模型发生错检现象的情况明显多于改进,如将黑色杆体错检成避雷器,将正方形箱体错检成电流互感器。而在处理图像中有多个样本的场景时,两种模型均表现出不错的检测效果,而改进检测出的目标表现出更高的置信度水平。说明在引入 SPDCConv 网络结构, CBAM 注意力模块以及 FocalerIoU 后,模型不仅在准确度上有所提高,且能在面对正负样本差异较小以及复杂图像背景的检测场景下,表现出更好的检测水平。总之,改进方法可以在保障检测效果的前提下运用于更丰富的电力图像目标检测任务。

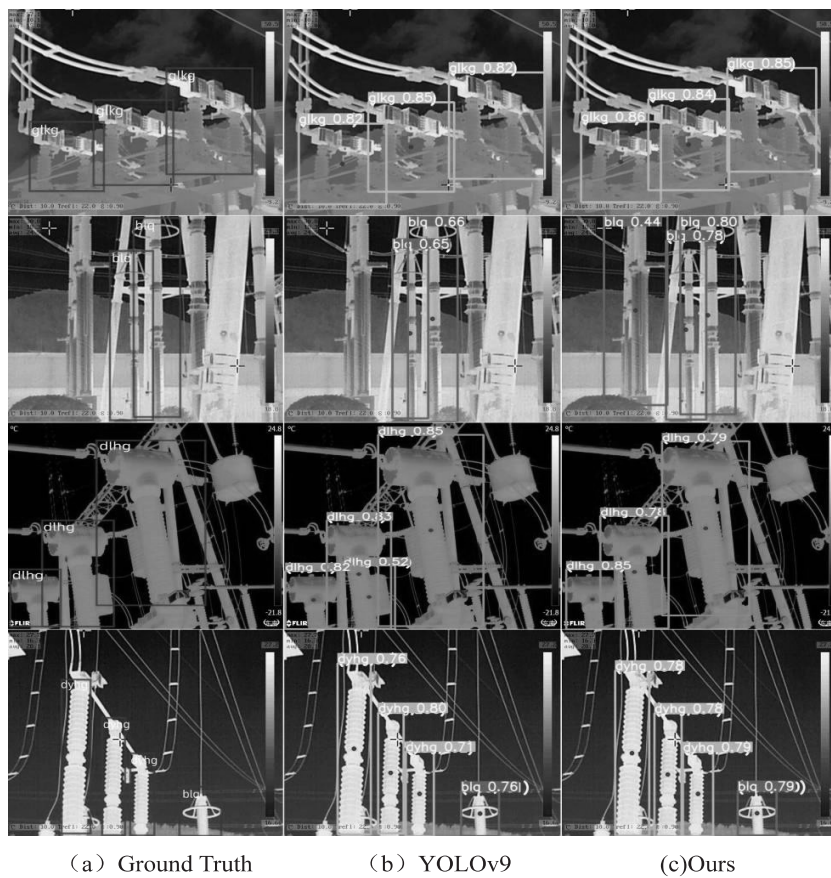


图 7 不同模型对同组图像的检测效果对比

3 结束语

针对电力设备红外图像在复杂环境下的目标检测挑战,诸如识别精度不足、分类错误频发及错检漏检等

问题,创新性地提出了一种优化的 YOLOv9 检测框架。该方案通过三重策略显著提升了检测性能:一是引入 SPDCConv 技术,有效减轻了基准模型主干网络的计算负担,实现了模型的轻量化设计;二是集成

CBAM卷积块注意力模块,该模块能够智能地增强模型对图像中关键特征的关注与提取能力,从而提升检测精度;三是采用FocalIoU作为新的损失函数,不仅优化了目标框的回归精度,还强化了模型在复杂场景下的判别能力,减少了误检与漏检现象。

实验验证显示,该改进算法在电力设备红外图像数据集上表现卓越,mAP值高达94.4%,同时保持了高效的推理速度,达到了59.5 f/s,完美平衡了检测精度与实时性需求。相较于现有技术,该方法展现了更强的综合检测能力,尤其是在处理各类电力设备目标时,其识别准确性与稳定性均显著提升,为电力设备红外图像的智能检测与监测领域带来了实质性的技术突破,具有较高的实际应用价值与推广前景。

参考文献:

- [1] 曾军,王东杰,范伟,等.基于红外热成像的电气设备组件识别研究[J].红外技术,2021,43(7):679-687.
- [2] 郭文诚,崔昊杨,马宏伟,等.基于Zernik矩特征的电力设备红外图像目标识别[J].激光与红外,2019,49(4):503-506.
- [3] 刘子全,付慧,李玉杰,等.基于Mask-RCNN迁移学习的红外图像电力设备检测[J].数据采集与处理,2021,36(1):176-183.
- [4] 郑含博,李金恒,刘洋,等.基于改进YOLOv3的电力设备红外目标检测模型[J].电工技术学报,2021,36(7):1389-1398.
- [5] HUANG X, WANG X, LV W, et al. PP-YOLOv2: a practical object detector[J]. arXiv:2104.10419, 2021.
- [6] 高伟,郭美青,张兴忠,等.基于改进PP-YOLOv2的红外图像电力设备检测[J].计算机仿真,2023,40(9):81-87.
- [7] 王小栋,吕通发,鲍明正,等.基于改进YOLO v5方法的电力设备红外图像识别方法[J].红外技术,2024,46(6):722-727.
- [8] 陈佳.基于YOLOv7的电力设备红外图像分类检测[D].重庆:重庆理工大学,2024.
- [9] 杨立新.基于改进YOLOv7算法的变电站设备缺陷检测模型研究[D].西安:西安理工大学,2024.
- [10] 李珅,杜科,李舟演,等.一种改进YOLOv8n的电力设备红外图像识别网络[J].无线电工程,2024,54(10):2362-2370.
- [11] WANG C Y, YE H I, LIAO M H Y. Yolov9: learning what you want to learn using programmable gradient information[J]. arXiv:2402.13616, 2024.
- [12] CHUA L O. CNN: a vision of complexity[J]. International Journal of Bifurcation and Chaos, 1997, 7(10): 2219-2425.
- [13] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects[C]//Joint European conference on machine learning and knowledge discovery in databases. Berlin: Springer, 2022: 443-459.
- [14] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City: IEEE. 2018: 7132-7141.
- [15] CHIEN C T, JU R Y, CHOU K Y, et al. YOLOv9 for fracture detection in pediatric wrist trauma X-ray images[J]. Electronics Letters, 2024, 60(11): e13248.
- [16] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]//Proceedings of the European conference on computer vision(ECCV). Berlin: Springer, 2018: 3-19.
- [17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. Venice: IEEE, 2017: 2980-2988.
- [18] YU J, JIANG Y, WANG Z, et al. Unitbox: an advanced object detection network[C]//Proceedings of the 24th ACM international conference on multimedia. Amsterdam: ACM, 2016: 516-520.
- [19] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Long Beach: IEEE, 2019: 658-666.
- [20] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//Proceedings of the AAAI conference on artificial intelligence. Menlo Park: AAAI 2020: 12993-13000.
- [21] ZHANG H, ZHANG S. Focaler-IoU: more focused intersection over union loss[J]. arXiv:2401.10525, 2024.
- [22] GEVORGYAN Z. Siou loss: more powerful learning for bounding box regression[J]. arXiv:2205.12740, 2022.
- [23] REN S, HE K, 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, 2016, 39(6): 1137-1149.
- [24] 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. Vancouver: IEEE, 2023: 7464-7475.