

DOI:10.19651/j.cnki.emt.2417668

面向嵌入式设备的绝缘子缺陷检测算法研究^{*}刘梓良¹ 尼鹿帕尔·艾克木² 伊力哈木·亚尔买买提¹ 郭松杰¹

(1.新疆大学电气工程学院 乌鲁木齐 830017;2.国网乌鲁木齐供电公司 乌鲁木齐 830000)

摘要:针对嵌入式设备在资源受限以及雾天复杂环境下绝缘子缺陷检测的高效性与准确性挑战,本文提出了一种新型轻量化检测模型——RNSC-YOLOv7-tiny,并取得了重要的创新性成果与实际应用价值。首先,通过对主干网络中 ELAN 模块的轻量化处理,设计了 RepNCSPELAN 模块,有效降低了模型的参数量与计算复杂度,同时保持了检测精度的显著提升。其次,融入 SGE 模块,使模型能够聚焦与背景重叠的目标区域,显著抑制无关信息的干扰,提升了绝缘子缺陷定位与识别的精确性。此外,引入 NWD 损失函数解决了检测过程中偏差点导致的梯度消失问题,进一步优化了检测精度。最后,CARAFE 上采样算子的引入,使模型在低分辨率图像及复杂雾天环境中依然能够实现精准检测定位。实验结果显示,RNSC-YOLOv7-tiny 模型在绝缘子缺陷检测方面展现出了快速且高精度的性能,其检测精度高达 94.8%。该模型拥有 4298150 个参数,浮点运算次数为 10.5,同时模型内存占用仅为 8.69 MB。与原始 YOLOv7-tiny 模型相比,新提出的模型在多个关键指标上均实现了显著提升:精度提高了 3.4%,参数量减少了 28.5%,浮点运算次数降低了 19.2%,且模型大小缩减了 3.01 MB。这一成果充分验证了该算法在嵌入式设备环境中的高度适用性和实际应用效能。

关键词: 嵌入式设备;绝缘子缺陷检测;YOLOv7-tiny

中图分类号: TP391.41;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.6

Research on insulator defect detection algorithm for embedded devices

Liu Ziliang¹ Nilupaer·Aikemu² Yilihamu·Yaermainaiti¹ Guo Songjie¹

(1. Electrical Engineering College, Xinjiang University, Urumqi 830017, China;

2. State Grid Urumqi Power Supply Company, Urumqi 830000, China)

Abstract: Aiming at the challenges of high efficiency and accuracy of insulator defect detection for embedded devices in resource-constrained and foggy complex environments, this paper proposes a new lightweight detection model, RNSC-YOLOv7-tiny, and achieves important innovative results and practical application value. Firstly, the RepNCSPELAN module is designed by lightweighting the ELAN module in the backbone network, which effectively reduces the number of parameters and computational complexity of the model, while maintaining a significant improvement in detection accuracy. Secondly, the incorporation of the Spatial Group Enhancement module enables the model to focus on the target region overlapping with the background, thereby significantly suppressing the interference of irrelevant information and improving the accuracy of insulator defect localisation and identification. Furthermore, the incorporation of the NWD loss function addresses the issue of gradient vanishing due to deviation points in the detection process, thereby enhancing the overall detection accuracy. Furthermore, the incorporation of the CARAFE upsampling operator enables the model to achieve accurate detection and localisation in low-resolution images and complex foggy environments. The experimental results demonstrate that the RNSC-YOLOv7-tiny model exhibits rapid and highly accurate performance in insulator defect detection, with a detection accuracy of 94.8%. The model comprises 4298150 parameters and 10.5 floating-point operations, yet occupies only 8.69 MB of memory. In comparison to the original YOLOv7-tiny model, the newly proposed model exhibits notable enhancements in several pivotal metrics. The accuracy has been augmented by 3.4%, the number of parameters has been diminished by 28.5%, the number of floating-point operations has been reduced by 19.2%, and the model size has been reduced by 3.01 MB. These outcomes substantiate the algorithm's high applicability in embedded device environments and its efficacy in practical applications.

Keywords: embedded devices; insulator defects detection; YOLOv7-tiny

0 引言

随着电力系统的不断发展,绝缘子的广泛使用导致其

在运行过程中因各种因素产生缺陷的概率增加。如果这些缺陷不能及时处理,可能会引发严重的电力故障。因此,研究高效且准确的绝缘子缺陷检测技术具有重要意义。

收稿日期:2024-12-18

* 基金项目:国家自然科学基金(62362063,61866037)项目资助

在早期,国内研究者尝试使用边缘检测和形态学处理等方法来检测绝缘子缺陷。例如,Liao等^[1]提出了一种多尺度、多特征方法,能够提取局部特征并定位受损绝缘子。Wu等^[2]则提出了一种基于Beltrami框架的半局部算子,用于提取表面不均匀纹理特征。然而,绝缘子图像通常包含各种自然环境背景,不同材质的绝缘子也具有不同的纹理和姿态。这些传统的图像处理手段,其核心在于手工设定的特征,因此在处理复杂或动态变化的图像时,其适应性较为有限。另外,这些手段对于光照条件的变化和复杂背景的识别,缺乏足够的鲁棒性,难以满足复杂缺陷检测任务对精确度和可靠性的高要求。

随着人工智能的发展,研究人员开始将机器学习方法应用于绝缘子缺陷检测。传统的机器学习途径包括细致的特征抽取,随后将这些特征作为输入数据,应用于支持向量机(support vector machine, SVM)等分类器中,以实现精确的缺陷检测与分析。Murthy等^[3]结合小波多分辨率分析与SVM进行绝缘子缺陷检测,Singh等^[4]提出了基于高斯核的SVM,通过几何形态、强度和统计特征对绝缘体缺陷进行分类。然而,依赖复杂特征工程的这些机器学习方法,其特征提取过程既繁琐又耗时,这在某种程度上对检测精度的提升速度和泛化能力的增强效果产生了制约。

近年来,小目标检测领域的深入研究使得YOLO系列算法在绝缘子缺陷检测中展现了广泛的应用前景。He等^[5]通过改进YOLOv4模型识别绝缘子自爆缺陷,将检测准确率提升至88.07%。Chen等^[6]提出了结合注意力机制的INSU-YOLO模型,提高了对小目标缺陷的识别率。李季等^[7]改进了YOLOv3算法,在主干网络中引入了无参数注意力模块(simple attention module, SimAM),提升了检测速度和精度。然而,这些模型通常参数量大、计算资源需求高,难以适用于资源受限的嵌入式设备。YOLOv7-tiny作为轻量化版本,具备较高的检测速度和较小的模型体积。然而,为了在资源受限的嵌入式设备上应用,仍需进一步优化模型。

尽管近年来基于深度学习的目标检测技术在绝缘子缺陷检测领域取得了显著进展,但现有技术仍存在诸多局限性。一方面,现有模型通常具有较大的参数量和高计算复杂度,这使得其难以部署在资源受限的嵌入式设备上,例如低功耗设备和实时处理场景中。此外,嵌入式设备在内存、功耗和计算能力方面的限制,导致轻量化模型在特征提取能力和检测精度上受到折衷,特别是对小目标和复杂背景的处理能力较弱。另一方面,在复杂自然环境下,绝缘子的图像质量通常较低,表现为细节模糊、对比度降低和背景噪声增强,这使得模型难以有效提取关键特征,从而导致检测精度显著下降。同时,现有模型中的主流损失函数在处理小目标时,梯度信息不足,无法有效优化模型性能。这些技术局限性在很大程度上限制了现有绝缘子缺陷检测技术在

复杂环境和嵌入式设备中的广泛应用。

针对上述问题,为保证满足嵌入式设备要求的同时能够有效提高检测效果,本文将对YOLOv7-tiny模型进行改进,具体改进如下:

1)提出了基于改进YOLOv7-tiny的轻量化绝缘子缺陷检测模型(RNSC-YOLOv7-tiny),通过设计轻量化RepNCSPPELAN模块和引入空间组增强模块(spatial group enhancement,SGE)解决YOLOv7-tiny虽然轻量化,但在处理细小目标和复杂场景时,特征提取能力不足的问题。

2)使用损失函数(normalized wasserstein distance, NWD)解决小目标检测的梯度消失问题,在复杂环境中优化检测性能。

3)引入CARAFE上采样算子提升对低分辨率图像的特征恢复能力,使模型在嵌入式设备上实现高效部署。

1 YOLOv7-tiny 算法简介

YOLOv7-tiny是一种目标检测算法,是YOLOv7的一个轻量级版本,旨在实现快速而准确的实时目标检测^[8]。相较于传统的目标检测算法,YOLO系列算法以其卓越的速度和精度脱颖而出,实现了更快速、更准确的检测效果。主要针对嵌入式设备和移动端应用进行优化。在确保高效检测准确性的同时,该模型努力缩减了模型体积与计算负载,以适配资源稀缺设备对即时目标识别功能的迫切需求^[9]。模型的核心架构由三大组件构成:特征提取主干网络(Backbone)、连接颈部网络(Neck)以及检测头部(Head)组件^[10]。

将输入的绝缘子图像经过简化的ELAN模块、SPPCSPC结构和MP模块构成的Backbone进行特征提取后,将来自不同深度特征层的特征图传递到带有特征保留和路径聚合网络的Neck部分,进行多尺度特征融合,最后将特征图送入Head部分进行绝缘子缺陷的定位与检测识别^[11]。YOLOv7-tiny网络如图1所示。

2 RNSC-YOLOv7-tiny 算法

尽管YOLOv7-tiny在参数和计算量上较为轻量,且能够保持较高的检测精度,但在处理绝缘子缺陷等具有特征较弱的目标时,其特征提取能力存在不足之处。因此,为解决上述问题,提出了RNSC-YOLOv7-tiny网络模型。首先,设计一种轻量化的来改造主干网络结构。其次,在主干网络中的RepNCSPPELAN模块引入SGE模块,使模型能够更高效地聚焦于与环境背景重合的目标,增强对缺陷的区分能力。然后,最后采用NWD损失函数来提高模型对小目标的检测能力。最后,在网络特征融合层中引入了CARAFE上采样算子,以实现对低分辨率图像和雾天条件下的精细定位和分割。RNSC-YOLOv7-tiny算法结构如图2所示。

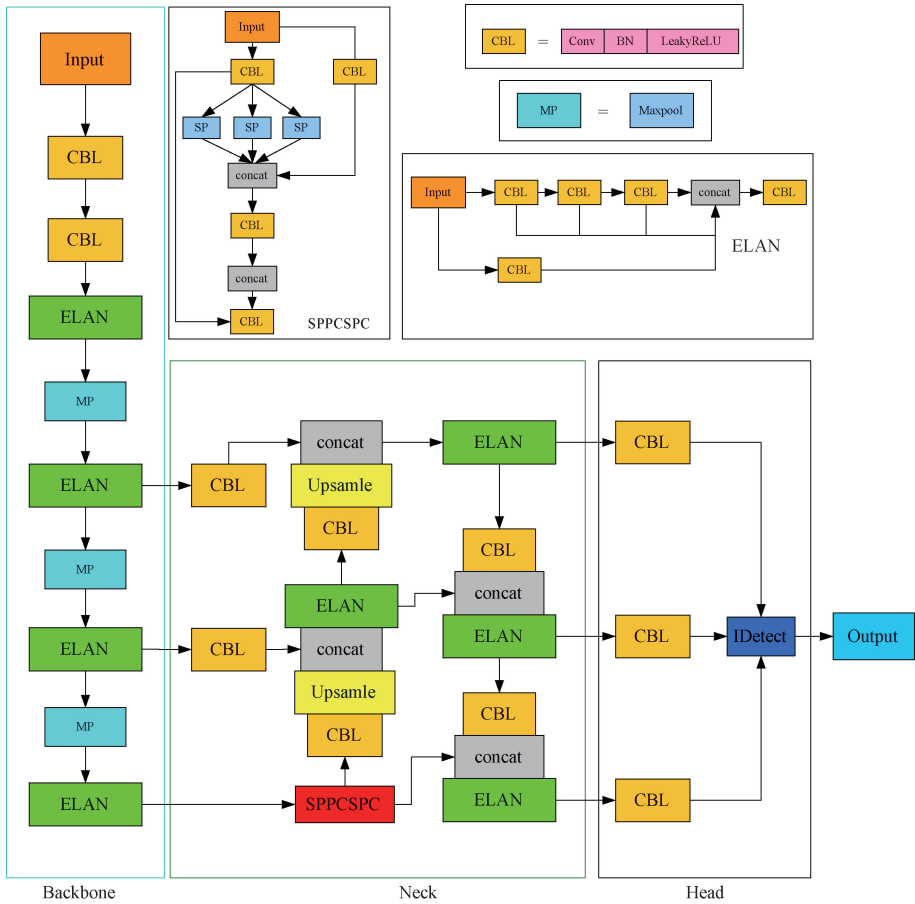


图 1 YOLOv7-tiny 网络
Fig. 1 YOLOv7-tiny network

2.1 改进 ELAN 模块

深度神经网络的参数通过网络输出与目标值之间的差异进行更新。这个过程依赖于损失函数计算出的梯度，以此来调整网络参数。然而，深度神经网络的输出往往难以完整保留预测目标的全部信息，导致网络在训练过程中使用了不完全的信息，可能引发梯度不稳定、收敛性差等问题。

为解决这一问题，广义高效层聚合网络 (generalized efficient layer aggregation network, GELAN) 提出了一种改进方法，相较于基于深度卷积的最新方法，GELAN 仅使用常规卷积算子即可实现更高的参数利用率。同时，可编程梯度信息模块 (programable gradient infomation, PGI) 适用于从轻量级到大型的各种模型，能够获取更完整的预测目标信息^[12]。

PGI 的核心思想是通过辅助可逆分支生成可靠的梯度，确保深层特征在执行目标任务时仍能保留关键信息。该设计避免了传统多路径特征融合与深度监督过程中可能产生的语义损失，从而更好地应对深度网络实现多个目标时的复杂需求。PGI 模块在训练流程中显著降低了网络信息的损耗，同时在推理阶段被有效移除，以确保模型

的大小、计算复杂度及推理速率均不受影响。

此外，GELAN 通过结合 CSPNet^[13] 和 ELAN^[14] 这两种基于梯度路径规划设计的神经网络架构，实现了在轻量化设计上对推理速度与准确性的精妙平衡。GELAN 网络的整体架构如图 3 所示。

RepNCSPELAN 作为轻量级结构，该模块的设计架构是用 CSPNet 模块和 RepConv^[15] 卷积作为计算块，将 ELAN 替换为 GELAN，简化了下采样的过程，同时优化锚定预测头，对于 PGI 辅助损耗部分，遵循 YOLOv7-tiny 的辅助头设置。RepNCSPELAN 结构如图 4 所示。

2.2 SGE 注意力机制

由于复杂背景的干扰、小目标检测困难以及噪声等影响导致卷积神经网络对绝缘子无法准确定位和识别。此外，作为图像中的小目标，绝缘子细节容易在高层特征中被忽略，且周围环境中的噪声会进一步干扰网络的特征提取过程，导致识别和定位不准确。添加注意力机制能够显著提升卷积神经网络对绝缘子的定位和识别能力。它通过聚焦关键区域，忽略无关背景信息，增强对复杂场景中目标的区分能力。同时，注意力机制极大地改善了小目标的检测能力，确保了关键细节不被忽视，并有效抑制了噪

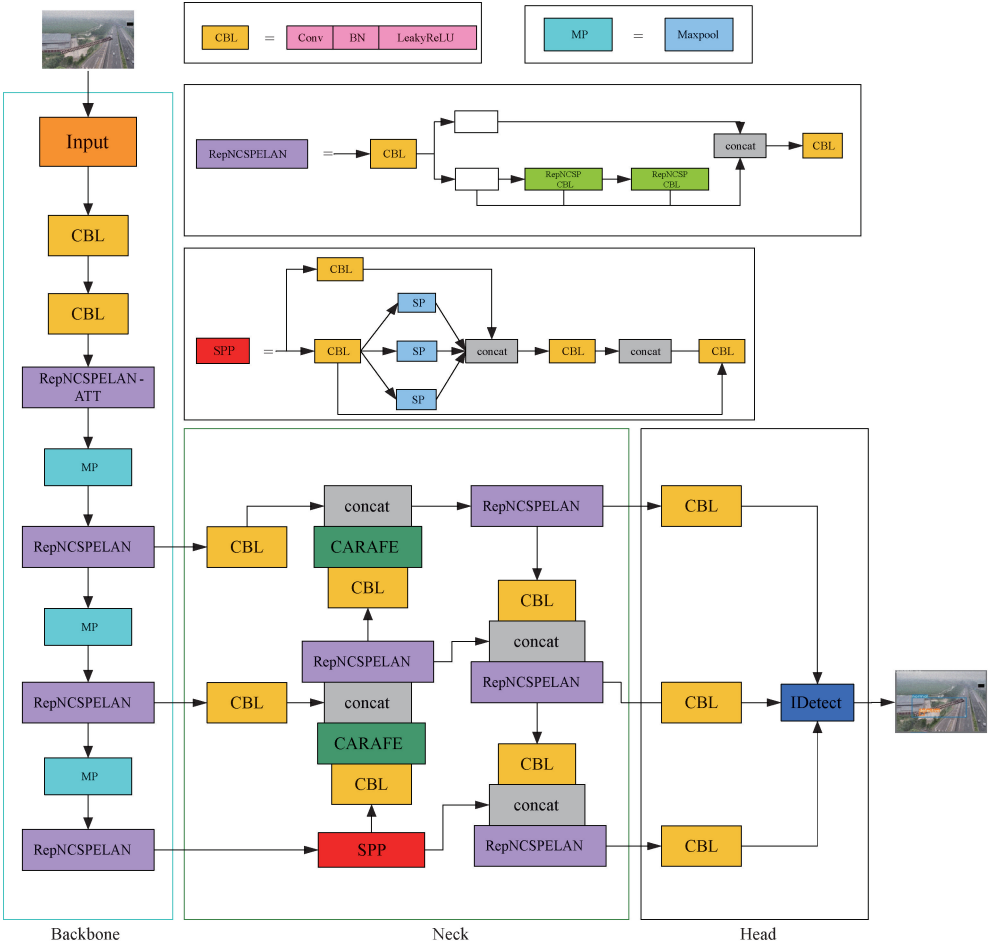


图 2 RNSC-YOLOv7-tiny 算法结构
Fig. 2 Network chart of RNSC-YOLOv7-tiny

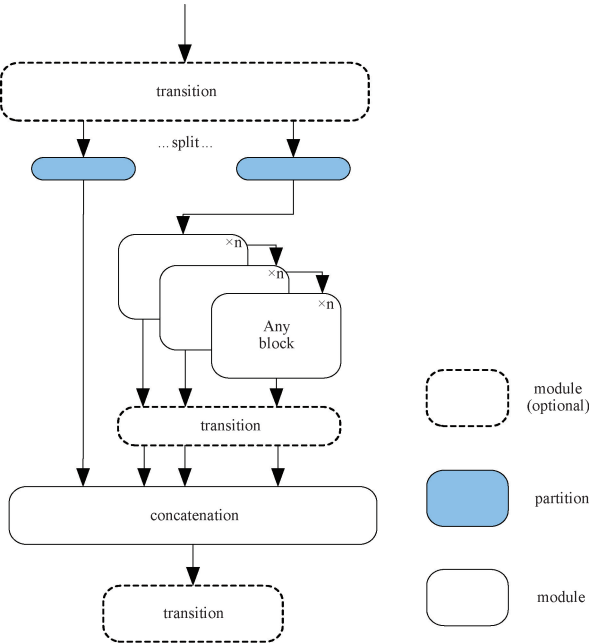


图 3 GELAN 结构
Fig. 3 GELAN structure

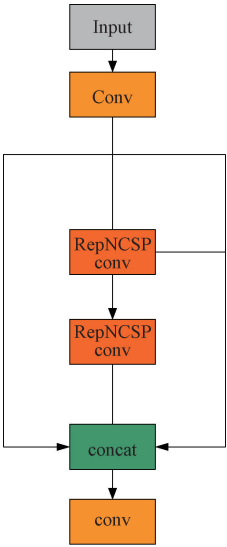


图 4 RepNCSPPELAN 结构
Fig. 4 RepNCSPPELAN structure

声的干扰,从而显著提高了绝缘子定位和识别的准确度与效率。

绝缘子通常位于自然场景中,容易与其他类似物体混淆,导致检测过程中难以实现准确的定位和识别。在主干网络中添加 SGE 模块能够显著提升绝缘子在复杂自然环境中的定位和识别准确性。SGE 模块通过为不同空间区域生成注意力权重,帮助网络聚焦关键区域,抑制无关背景的干扰,增强特征表达能力,特别是在绝缘子与类似物体难以区分的场景中。此外,SGE 模块还有效抑制噪声干扰,并提升了对小目标的检测精度,从而在复杂环境中实现更准确的绝缘子定位和识别^[16]。SGE 注意力模块如图 5 所示。

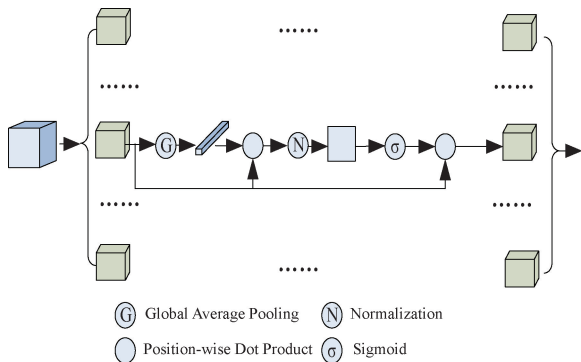


图 5 SGE 注意力模块

Fig. 5 SGE attention module

2.3 NWD 损失函数

微小物体在现实世界中广泛存在,然而,对于绝缘子这类外观信息极其有限的微小物体来说,网络模型在学习其特征时面临较大挑战,这往往导致绝缘子检测的失败率较高。此外,IoU 对不同尺度物体的灵敏度存在较大差异,尤其对于像绝缘子这样的小物体,即使位置出现轻微偏差,也会导致 IoU 显著下降,进而导致标签分配不准确。而对于正常大小的物体,IoU 的变化较小,位置偏差几乎没有影响。这一现象表明,IoU 度量在面对具有位置偏差的不同尺度对象时存在不一致性,最终导致标签分配不准确^[17]。因此,在训练微小目标检测器时,由于微小目标的特征信息不足,网络无法有效地获取这些特征反馈,导致模型更多地关注中大型目标的特征,完全忽视了对微小目标特征的学习。

为解决上述问题,将 CIoU 替换为 NWD 作为模型的损失函数,采用最优输运理论中的 Wasserstein 距离^[18]来衡量分布间的差异。对于两个二维高斯分布 $\mu_1 = N(m_1, \Sigma_1)$ 和 $\mu_2 = N(m_2, \Sigma_2)$,其二阶 Wasserstein 距离定义为:

$$W_2^2(\mu_1, \mu_2) = \|m_1 - m_2\|_2^2 + \text{Tr}(\Sigma_1 + \Sigma_2 - 2(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2})^{1/2}) \quad (1)$$

将式(1)简化为:

$$W_2^2(\mu_1, \mu_2) = \|m_1 - m_2\|_2^2 + \|\Sigma_1^{1/2} - \Sigma_2^{1/2}\|_F^2 \quad (2)$$

式中: $\|\cdot\|_F$ 为 Frobenius 规范。

对于由边界框 $A = (cx_a, cy_a, w_a, h_a)$ 和 $B = (cx_b,$

$cy_b, w_b, h_b)$, 经过式(2)可进一步简化为:

$$W_2^2(N_a, N_b) = \left\| \left([cx_a, cy_a, \frac{w_a}{2}, \frac{h_a}{2}]^T, [cx_b, cy_b, \frac{w_b}{2}, \frac{h_b}{2}]^T \right) \right\|_2^2 \quad (3)$$

由于 $W_2^2(N_a, N_b)$ 是一个距离度量,不能直接用作相似性度量。因此,使用它的指数形式的归一化,对式(3)进行指数形式的归一化,得到归一化 Wasserstein 距离^[18]:

$$NWD(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{C}\right) \quad (4)$$

式中: C 是与数据集密切相关的常数。

绝缘子缺陷通常属于小目标检测范畴,在改进模型中使用 NWD 损失函数,可以更好地捕捉这些细微的缺陷,提高检测的可靠性和准确性。同时, NWD 损失函数对不同大小和形状的目标有较好的适应性,可以在多样化的环境条件下依然保持较高的检测效果,这对雾天场景下的绝缘子缺陷检测非常有帮助。

2.4 轻量级上采样算子 CARAFE

在目标检测任务中,传统的上采样算子如双线性插值和转置卷积存在明显不足,包括特征细节的丢失、固定插值权重无法动态调整以及容易引入伪影。这些问题导致放大特征图时,尤其是在处理复杂背景中的小目标如绝缘子时,难以准确恢复其边缘和细节,增加了误报和漏检的风险,限制了检测的精度。

相比之下,轻量级 CARAFE 上采样算子^[19](如图 6 所示)通过动态生成插值核,增强了特征细节的重构能力,使网络能够更精确地定位小目标。它通过提升空间感知能力,保留了目标的边缘信息,有效减少了背景噪声的干扰。此外, CARAFE 具有轻量化设计,特别适合在资源有限的设备上高效地执行绝缘子缺陷检测任务。

在绝缘子缺陷检测中,尤其是在低分辨率或大雾等复杂环境下,更精细的定位和分割需求显得尤为重要。CARAFE 通过内容感知的内核重组操作能够实现更精确的上采样。其两步操作包括根据目标位置预测重组核,随后利用该核对特征图进行重组。结合通道注意力和可变形卷积操作, CARAFE 能够显著提升上采样的质量和精度,满足高精度检测的需求。

3 实验与结果分析

3.1 实验配置及其参数选择

本文的实验配置采用 Windows 11 家庭中文版操作系统,配备有 13 代 Intel Core i9-13980HX 处理器, GPU 选用 NVIDIA GeForce RTX 4060 显卡, 辅以 16 GB 的运行内存。实验过程中,运用了 CUDA 12.1 和 Pytorch 2.1.0 这两个深度学习框架,并基于 Python 3.9 版本进行开发。特别地,本实验未采用预训练权重进行迁移学习。在实验设置上,将初始学习率设定为 0.01,权重衰减 0.000 5, batch

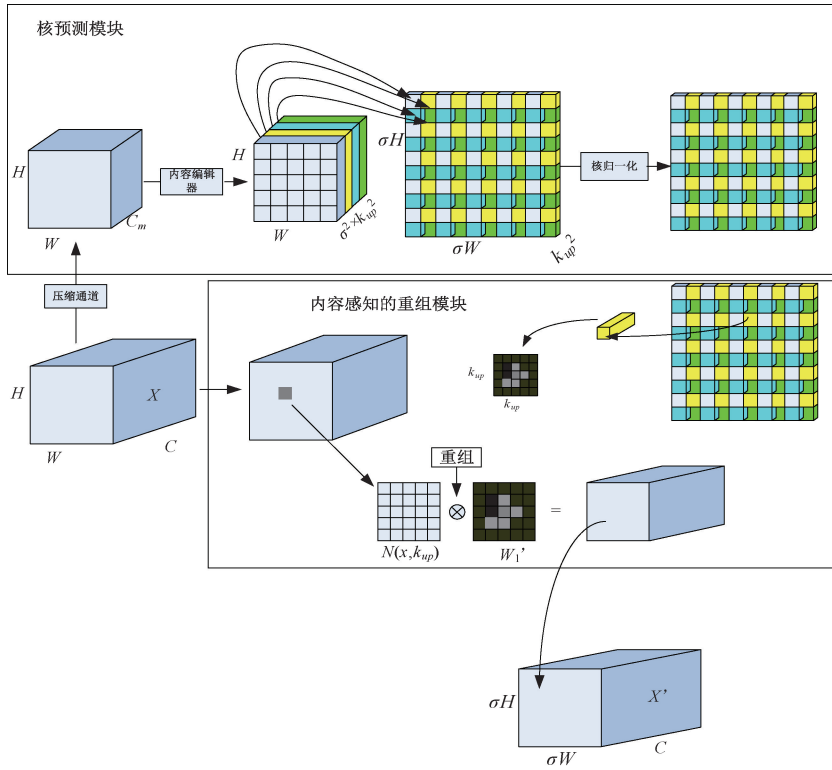


图6 CARAFE过程

Fig. 6 CARAFE process

size 设定为 16, 采用 Adam 优化器进行模型优化, 学习率调度采用 Cosine Annealing, 动态调整学习率以加速收敛, 训练轮数为 200 轮。

3.2 数据集

1) 数据集构建过程

本实验使用了 2 743 张输电线路绝缘子图片, 数据集经过标注工具 (如 LabelImg) 精确标注。数据分为训练集和测试集, 比例为 4:1。图片中包含正常和有缺陷的绝缘子实例, 确保了类别平衡性。为模拟真实环境中的复杂条件, 数据集中增加了 1 138 张 (训练集) 和 266 张 (测试集) 带有雾霾条件的图像, 确保模型在低对比度条件下的鲁棒性。

2) 数据预处理

在数据预处理步骤中, 为了确保模型输入数据的一致性和鲁棒性, 所有原始图片首先被调整为统一的尺寸 640 pixel × 640 pixel。同时, 为增强模型在不同环境条件下的适应性, 对数据进行了多种增强操作, 包括水平翻转和垂直翻转来模拟视角变化, 随机填充以增强局部特征的多样性, 添加高斯噪声用于提升模型对噪声干扰的鲁棒性, 以及随机调整亮度、对比度和饱和度来模拟不同光照条件。为进一步增加数据的多样性, 增强操作在 40% 的训练样本上随机应用。此外, 考虑到雾天对图像清晰度的影响, 特别添加了人工模拟的有雾条件图片, 以确保模型能够在低对比度和细节模糊的情况下仍能有效检测目标。

3.3 评价指标

为全面评估模型的性能与轻量化程度, 本文精选了以下关键指标作为评价标准: 精度 P (Precision)、参数量 (Parameters)、浮点运算次数 (GFLOPs) 以及模型大小。这些指标共同构成了评估模型效能与效率的综合框架。式(5)用于获得精度的值。

$$P = \frac{TP}{TP + FP} \quad (5)$$

其中, TP 代表预测为阳性且实际也为阳性的样本数量, 而 FP 则指的是预测为阳性但实际上为阴性的样本数量。

3.4 对比实验

为了评估改进模型在检测效能上的提升, 本实验集成了当前主流的多种目标检测算法, 并在同一数据集上进行了详尽的测试与对比, 包括 YOLOv3-tiny、YOLOv5s、YOLOv5n、YOLOv5-GhostNet^[20]、YOLOv7-tiny、YOLOv8m^[21]。选取 Precision、Parameters、GFLOPs 以及模型大小作为评价指标。对比结果如表 1 所示。

通过对比实验证明, RSNC-YOLOv7-tiny 改进算法的精度达到 94.8%, 与 YOLOv3-tiny、YOLOv5s、YOLOv5n、YOLOv7-tiny 和 YOLOv8m 相比, 精度分别提升了 13.1%、2.4%、4.3%、3.4% 和 0.8%。虽然本实验算法在参数量、浮点运算次数、模型大小要略高于 YOLOv5n, 但是在精度上有了明显的提升。与 YOLOv5-

表 1 对比实验结果

Table 1 Comparative experimental results

算法	Precision/ %	Parameters	GFLOPs	模型 大小/MB
YOLOv3-tiny	81.7	8 669 002	12.9	16.6
YOLOv5s	92.4	7 015 519	15.8	13.7
YOLOv5n	90.5	1 761 871	4.1	3.69
文献[20]	95.0	15 573 199	33.0	30.4
YOLOv7-tiny	91.4	6 010 302	13.2	11.7
YOLOv8m	94.0	11 126 358	28.4	21.4
ours	94.8	4 298 150	10.5	8.67

GhostNet 相比,RSNC-YOLOv7-tiny 算法,尽管精度略低 0.2%,但在参数量、浮点运算次数和模型大小上均表现出显著优势。这种性能权衡使得本文算法在资源受限的实际应用场景中具有明显的优势。具体而言,较低的数量和浮点运算次数使得模型在嵌入式系统和移动设备上运行更加高效,能够显著降低计算资源的需求和能耗。综上所述,RSNC-YOLOv7-tiny 改进算法在各方面指标均有明显提升,证明本文的改进算法用于复杂环境下的绝缘子缺陷检测可以发挥较好的效果。

3.5 引入不同损失函数的对比实验

为检验 NWD 损失函数在绝缘子缺陷识别任务中的实际效能,本文将 R-YOLOv7-tiny 模型配置为采用多种损失函数,包括 CIOU、Focal_EIoU、WIoU、SIOU、MPDIOU、Shape-IoU 以及 NWD,进行效果对比。在本文中,利用提供的数据集进行了深入的训练,设计相关的对比实验。对比损失函数实验结果如表 2 所示。

表 2 损失函数对比实验结果

Table 2 Comparative experimental results of loss function

方法	Precision/%	Parameters	GFLOPs
CIOU	93.0	4 266 430	10.4
Focal_EIoU	90.2	4 266 430	10.4
WIoU	92.9	4 266 430	10.4
SIOU	91.9	4 266 430	10.4
MPDIOU	92.0	4 266 430	10.4
Shape-IoU	92.2	4 266 430	10.4
NWD	93.8	4 266 430	10.4

同时,为了更直观地理解训练过程的动态变化,绘制了包括精度、框回归损失、目标置信度损失、类别分类损失的训练曲线,如图 7(a)~(d)所示。

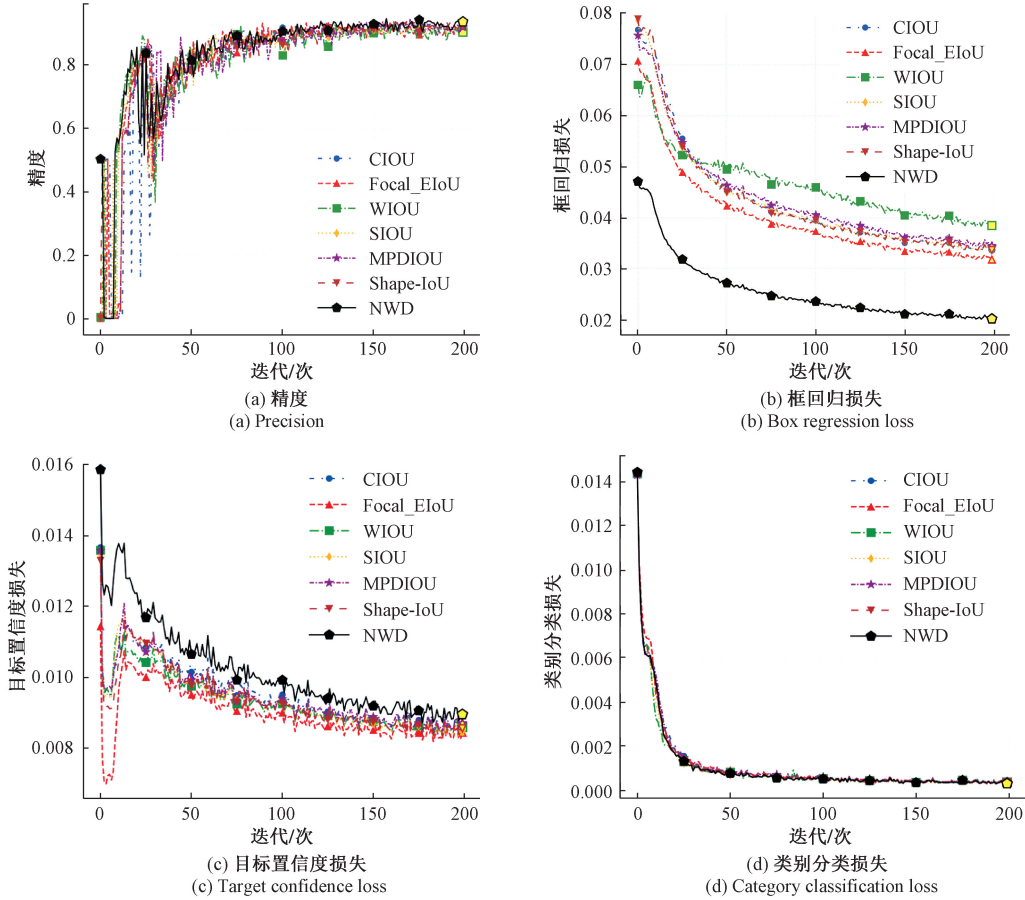


图 7 损失函数训练结果对比
Fig. 7 Comparison of loss function training results

通过损失函数数据对比,NWD 损失函数在无需额外增加参数量及计算负担的条件下,能够在检测精度上带来一定程度的优化与提升。原因在于 NWD 损失函数在处理目标检测中的不同情况时表现出更强的鲁棒性。相比之下,其他的 IoU 类损失函数在框的重叠区域较小时,梯度信息较少,容易导致误差积累,而 NWD 通过衡量更全面的距离信息,减少了这种误差的积累。此外,NWD 的全面衡量方式和丰富的梯度信息使得它在复杂检测任务中表现出色,显著提升了模型的检测性能。

在绘制损失函数对比曲线中,NWD 能够显著提升模型边界框损失的收敛效果,这一优势源自其基于 Wasserstein 距离的独特设计。该损失函数深入考量了两个分布间的几何距离与质量分布差异,因此能更全面、精确地捕捉预测框与真实框之间的细微差别。尤为值得一提的是,NWD 并不仅仅聚焦于框的重叠区域,而是同时兼顾了框的中心位置与形状差异,这一特性使得模型在训练阶段能够接收到更为有效、更具针对性的梯度更新信号,进而加速了模型向最优解收敛的进程。

3.6 引入注意力机制实验

为进一步验证 SGE 模块在绝缘子缺陷检测任务中的优越性,选取高效通道注意力(efficient channel attention, ECA)模块、无参数注意力模块(simple attention module, SimAM)、三重注意力模块(triplet attention)3 种不同模块,在相同嵌入位置进行对比实验,如表 3 所示。

表 3 不同注意力模块对比实验

Table 3 Comparison experiment of different attention modules

模型	精度 P/%	参数量	浮点运算次数
YOLOv7-tiny-RepNCSPELAN	93.0	4 266 430	10.4
+ SimAM	89.8	4 266 433	10.4
+ EMA	92.7	4 266 430	10.4
+ Triplet Attention	90.2	4 266 430	10.4
+ SGE	93.1	4 266 494	10.4

在绝缘子检测任务中,SGE 虽然在参数略微上上升,但是在精度上要优于 SimAM、ECA 和 Triplet Attention 的原因为:1)SGE 通过空间分组增强特征,使得特征在空间位置和通道之间有更好的交互;2)SGE 计算效率较高,而 SimAM、ECA 和 Triplet Attention 在处理复杂场景时可能效率不如 SGE;3)SGE 能够更好地捕捉局部特征,适合绝缘子检测这种对局部特征要求高的任务。综上所述,SGE 在绝缘子检测任务中的优越表现可能是由于其更好地捕捉了局部特征,同时保持了较高的计算效率。

为了探讨在主干网络的不同节点嵌入 SGE 模块对模型效能的潜在影响,实验将用到的 SGE 注意力模块对于不同

的位置进行 4 组对比实验。SGE 注意力机制添加的位置如图 8 所示,其中 a、b、c、d 分别代表主干网络的 2、4、6、8 层。

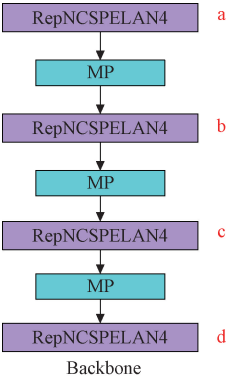


图 8 SGE 注意力机制添加的不同位置

Fig. 8 Different positions added by SGE attention mechanism

在不同位置添加 SGE 注意力机制的检测性能如表 4 所示,在不同的位置引入 SGE 注意力机制模块均会是网络模型的精度产生影响。虽然方法 1、4 加入 SGE 注意力机制模块后,对于绝缘子缺陷检测精度最高,但是方法 1 相比于方法 4 在参数量方面有所下降,更符合本文轻量化的目标。

表 4 不同位置的检测性能对比

Table 4 Comparison of detection performance at different positions

方法	添加位置	精度 P/%	参数量
1	(a)	93.1	4 266 494
2	(d)	91.6	4 266 942
3	(a)、(d)	92.5	4 267 006
4	(a)、(b)、(c)、(d)	93.1	4 267 398

3.7 消融实验

为了证明针对基线模型 YOLOv7-tiny 所改进的算法 RSNC-YOLOv7-tiny 的有效性。针对所加入的模块,采用组合依次叠加的方式进行了消融实验。具体实验结果如表 5 所示。其中 A 表示引入 RepNCSPELAN 模块,B 表示采用 SGE 注意力模块,C 表示使用 NWD 损失函数,D 表示使用 CARAFE 上采样算子。

表 5 给出的 5 组数据,在引入 RepNCSPELAN 模块后,模型的识别准确率有着明显的提升,精度 P 提升了 1.6%,同时由于 RepNCSPELAN 模块作为一种轻量化结构,模型的参数量相比于原模型降低了 29%,模型大小也有一定幅度的下降;在 RepNCSPELAN 的基础上进一步引入 SGE 注意力机制,使模型更高效地关注与环境重合的目标,提高模型对缺陷区分能力,进而提高检测精度,使精度 P 提升了 0.1%,但是参数量有了小幅度的增加;在上述基础上引入 NWD 损失函数,在目标检测时即使素点产生

表 5 消融实验结果
Table 5 Ablation experimental results

方法	A	B	C	D	精度 P/%	参数量	浮点运算次数	模型大小/MB
原模型	×	×	×	×	91.4	6 010 302	13.0	11.7
1	✓	×	×	×	93.0	4 266 430	10.4	8.61
2	✓	✓	×	×	93.1	4 266 494	10.4	8.61
3	✓	✓	✓	×	94.0	4 266 494	10.4	8.61
4	✓	✓	✓	✓	94.8	4 298 150	10.5	8.69

注：“✓”表示引用该模块，“×”表示为引用该模块。

偏差的情况依然具有梯度,进而提高检测精度,模型的精度提高了 0.9%,但是在参数量上维持不变。将网络特征融合层中原有的最近邻上采样方法,优化为采用基于内容感知的轻量级特征重建模块 CARAFE,以实现更高效、精准的特征处理,更加丰富了上采样后特征图的语义信息,进一步增强了利于网络的识别能力,使得精度 P 在改进三的基础上增加了 0.8%,参数量和模型大小只有小幅度增加。总体而言,4 种改进的融合达到了良好效果,在消融实验中结果最佳,证实了改进算法 RSNC-YOLOv7-tiny 对复

杂环境下的绝缘子缺陷的识别与检测能力有明显的提升。

3.8 检测结果分析

图 9(a)、(b)、(c)分别为原图、原模型对其检测性能以及改进后模型对其检测性能结果图,其中 7 幅图从左到右分别是正常、填充、噪声、有雾、填充+噪声、填充+翻转、有雾+填充+翻转。如图 9 所示,相较于 YOLOv7-tiny 算法,经过改进之后的网络对绝缘子的缺陷识别能力有着明显的提升,识别精度也有所提高。

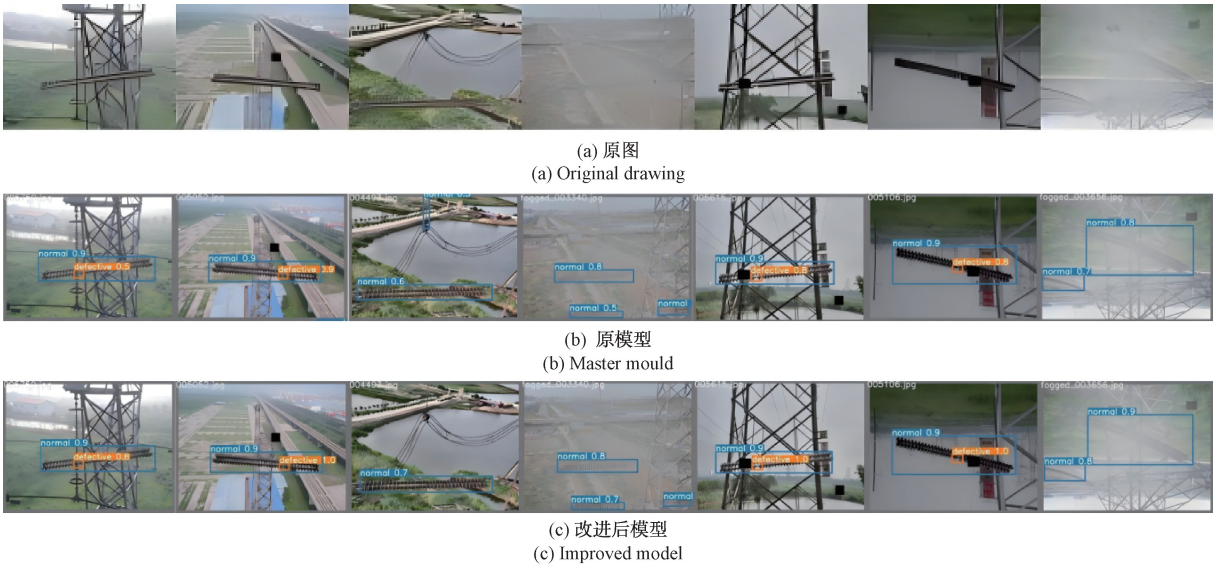


图 9 实际检测效果图

Fig.9 Actual test effect drawing

3.9 不同缺陷类型实验

为验证优化模型在不同类型缺陷检测上的表现、模型的泛化能力以及在未见过的缺陷类型上的表现。使用 3 652 张输电线路绝缘子图片,经过标注工具(如 Labellmg)精确标注,包含 4 种标签类型,分为绝缘子、绝缘子自爆、绝缘子破损、绝缘子闪络,确保了类别平衡性。数据分为训练集和测试集,比例为 4:1。将改进后的算法与原始算法在数据集上的训练结果进行对比分析。实验结果如表 6 所示。

从表 6 给出的数据可知,改进后的模型在不同类型缺

陷的检测上表现出了较好的性能,尤其是在自爆类型的缺陷上,准确率从 81.8% 提升至 87.5%,提升幅度达到 5.7%,表明改进后模型更好地捕捉到了该类型缺陷的特征。对于闪络和破损类型,改进后的模型未能显著提高准确率,可能因为这些缺陷的特征已经被原模型充分学习。模型的改进提高了泛化能力,特别是在自爆类型和其他较难检测的缺陷上显示了较好的适应性和鲁棒性。此外,改进后的模型通过减少参数量和优化模型大小,在保证性能的同时降低了计算资源消耗,适应了更为高效的部署要求。总体而言,改进后的模型不仅提升了对特定缺陷类型

表 6 不同绝缘子类型检测性能对比表

Table 6 Comparative performance of different insulator types

不同 类型	模型	P/%	Parameters	模型大 小/MB
闪络	原模型	81.3	6 015 714	11.7
	改进后模型	81.3	4 303 626	8.7
破损	原模型	78.2	6 015 714	11.7
	改进后模型	78.1	4 303 626	8.7
自爆	原模型	81.8	6 015 714	11.7
	改进后模型	87.5	4 303 626	8.7
绝缘子	原模型	71.7	6 015 714	11.7
	改进后模型	72.8	4 303 626	8.7

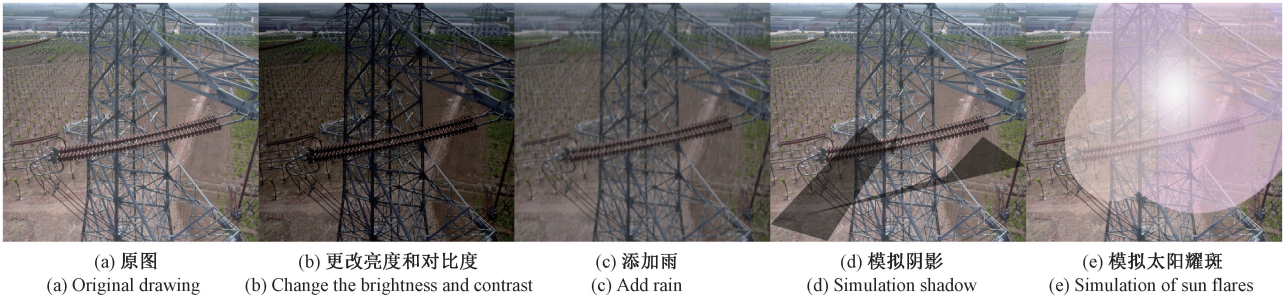


图 10 数据增强效果图

Fig. 10 Data augmentation effect images

示了光照变化对模型的影响;而背景复杂度的增加也导致了检测性能的下降,尤其在太阳耀斑时,模型的准确度显著降低。此外,模拟雨的效果时,模型的表现也有所下降,但在光照和背景复杂度变化下,增强后的模型显示了较强的适应性。整体来看,数据增强显著提升了模型在多种环境条件下的鲁棒性,表明模型在面对复杂光照、背景和天气因素时,仍需进一步优化其鲁棒性和适应能力,未来的工作可以通过多任务学习和网络架构优化来进一步提升其性能。

表 7 影响因素检测性能表

Table 7 Performance table of influence factors detection

增强类型	P/%	Parameters	模型大小/MB
更改图像亮度和对比度	91.6	4 298 150	8.69
添加雨效果	89.0	4 298 150	8.69
模拟阴影效果	94.5	4 298 150	8.69
模拟太阳耀斑	80.2	4 298 150	8.69

3.11 VOC 公共数据集实验

为了衡量经过优化的算法在不同数据集环境下的检测效果,选定 PASCAL VOC 公共数据集^[22]作为验证基准,该数据集是一个视觉对象分类挑战赛,包含 20 类图像用于分类和检测。本文将提出的算法应用于 PASCAL

的检测能力,也在模型效率和泛化能力上有所增强,展现出较强的适应性。

3.10 其他影响因素

本文实验考虑了有雾条件下对模型性能的影响,但对于其他可能影响检测性能的因素如光照、天气、背景复杂度等,实验数据和分析相对不足。为此,对现有测试集上进行数据增强,通过随机调整图像亮度 and 对比度、添加雨效果、模拟阴影和太阳耀斑,全面评估模型在不同复杂场景下的表现。这些增强手段不仅有助于验证模型的鲁棒性,还能提升其在光照变化、天气因素和复杂背景下的适应能力。数据增强效果如图 10 所示。

如表 7 所示,模型在不同光照条件下的表现有所波动,特别是在强光或低光照环境下,检测精度有所下降,显

VOC 数据集进行训练,并使用其测试集进行验证。随后,将改进后的 RSNC-YOLOv7-tiny 算法与原始 YOLOv7-tiny 算法在 PASCAL VOC 数据集上的训练结果进行对比,对比结果如表 8 所示。

表 8 PASCAL VOC 数据集检测性能对比表

Table 8 PASCAL VOC dataset detection performance comparison

算法	精度 P/%	Parameters	GFLOPs	大小/ MB
YOLOv7-tiny	62.8	6 059 010	13.2	11.7
RSNC-YOLOv7-tiny	65.0	4 346 858	10.7	8.78

通过对比 PASCAL VOC 数据集检测结果,改进后的算法 RSNC-YOLOv7-tiny 较原算法 YOLOv7-tiny,在精度上高了 2.2%,在模型大小及其参数量方面有了明显的下降,证明了文中对于改进的轻量化算法的适用性。

4 结 论

针对嵌入式设备在资源受限和雾天复杂环境下绝缘子缺陷检测的难题,本文提出了轻量化检测模型 RSNC-YOLOv7-tiny。通过轻量化处理 ELAN 模块并设计 RepNCSPELAN,有效降低参数量与计算复杂度,同时提

升检测精度;引入 SGE 模块 聚焦目标区域,抑制背景干扰,增强缺陷定位与识别精度;结合 NWD 损失函数 解决梯度消失问题,进一步优化检测性能;最终通过 CARAFE 上采样算子,实现低分辨率与雾天环境下的精准检测定位。在与现有文献的对比中,本研究的模型在精度、轻量化设计以及实际部署能力上均取得了领先优势,为绝缘子缺陷检测领域提供了新的解决方案。

虽然模型在雾天环境下表现良好,但在强光照、雨雪等极端天气下的鲁棒性仍需提升,且针对复杂缺陷类型的泛化能力有待提升。未来研究将着重优化模型的鲁棒性与泛化性,结合更多场景数据与策略,推动模型的广泛应用。

参考文献

- [1] LIAO SH L, AN J B. A robust insulator detection algorithm based on local features and spatial orders for aerial images [J]. IEEE Geoscience and Remote Sensing Letters, 2014, 12(5): 963-967.
- [2] WU Q G, AN J B. An active contour model based on texture distribution for extracting inhomogeneous insulators from aerial images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 52(6): 3613-3626.
- [3] MURTHY V S, TARAKANATH K, MOHANTA D K, et al. Insulator condition analysis for overhead distribution lines using combined wavelet support vector machine (SVM) [J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2010, 17(1): 89-99.
- [4] SINGH L, ALAM A, KUMAR K V, et al. Design of thermal imaging-based health condition monitoring and early fault detection technique for porcelain insulators using Machine learning [J]. Environmental Technology & Innovation, 2021, 24: 102000.
- [5] HE H, HUANG X L, SONG Y X, et al. An insulator self-blast detection method based on YOLOv4 with aerial images [J]. Energy Reports, 2022, 8: 448-454.
- [6] CHEN W X, LI Y N, ZHAO ZH G. Missing-sheds granularity estimation of glass insulators using deep neural networks based on optical imaging [J]. Sensors, 2022, 22(5): 1737.
- [7] 李季,刘乐,牛雨潇,等. 融入注意力的 YOLOv3 绝缘子串识别方法[J]. 高压电器, 2022, 58(11): 67-74.
LI J, LIU L, NIU Y X, et al. YOLOv3 insulator string recognition method incorporating attention[J]. High Voltage Electrical Appliances, 2022, 58(11): 67-74.
- [8] WANG Q Y, ZHANG ZH, CHEN Q G, et al. Lightweight transmission line fault detection method based on leaner YOLOv7-Tiny [J]. Sensors, 2024, 24(2): 565.
- [9] 张光华,李聪发,李钢硬,等. 基于改进 YOLOv7-tiny 的无人机航拍图像小目标检测算法[J]. 工程科学与技术, 2025, 57(3): 235-246.
ZHANG G H, LI C F, LI G Y, et al. A small target detection algorithm for unmanned aerial vehicle images based on improved YOLOv7-tiny [J]. Engineering Science and Technology, 2025, 57(3): 235-246.
- [10] 薛珊,安宏宇,吕琼莹,等. 复杂背景下基于 YOLOv7-tiny 的图像目标检测算法 [J]. 红外与激光工程, 2024, 53(1): 269-280.
XUE SH, AN H Y, LYU Q Y, et al. Image object detection algorithm based on YOLOv7-tiny in complex backgrounds [J]. Infrared and Laser Engineering, 2024, 53(1): 269-280.
- [11] 刘凌峰,陈洪刚,卿粼波,等. 基于改进 YOLOv7-Tiny 的工业缺陷检测研究[J]. 智能计算机与应用, 2023, 13(11): 69-76.
LIU L F, CHEN H G, QING L B, et al. Research on industrial defect detection based on improved YOLOv7-Tiny [J]. Intelligent Computers and Applications, 2023, 13(11): 69-76.
- [12] 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. Cham: Springer Nature Switzerland, 2024: 1-21.
- [13] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020: 390-391.
- [14] WANG C Y, LIAO H Y M, YE H I H. Designing network design strategies through gradient path analysis[J]. ArXiv preprint arXiv:2211.04800, 2022.
- [15] 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[J]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR), 2023: 7464-7475.
- [16] LI X, HU X L, YANG J. Spatial group-wise enhance: Improving semantic feature learning in convolutional networks [J]. ArXiv preprint arXiv: 1905.09646, 2019.
- [17] REZATOFIGHI H, TSOI N, GWAK J Y, 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, 2019: 658-666.

[18] WANG J W, XU CH, YANG W, et al. A normalized Gaussian Wasserstein distance for tiny object detection[J]. ArXiv preprint arXiv:2110.13389, 2021.

[19] WANG J Q, CHEN K, XU R, et al. CARAFE: Content-aware reassembly of features[C]. IEEE/CVF International Conference on Computer Vision, 2019: 3007-3016.

[20] 贾玉进,张振程,李浠铭,等. 基于轻量化 YOLOv5 网络的输电线路绝缘子缺陷检测[J]. 电力学报, 2024, 39(1):36-44.

JIA Y J, ZHANG ZH CH, LI X M, et al. Defect detection of transmission line insulators based on lightweight YOLOv5 network [J]. Electric Power Journal, 2024, 39(1): 36-44.

[21] VARGHESE R, SAMBATH M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems(ADICS). IEEE, 2024: 1-6.

[22] EVERINGHAM M, WINN J. The pascal visual object classes challenge 2010 (voc2010) development kit contents[C]. International Conference on Machine Learning, 2011: 117-176.

作者简介

刘梓良,硕士研究生,主要研究方向为目标检测。

E-mail:623441197@qq.com

尼鹿帕尔·艾克木,助理工程师,主要研究方向为电力设备维护。

E-mail:2496216464@qq.com

伊力哈木·亚尔买买提(通信作者),硕士,教授,主要研究方向为人工智能、模式识别、图像缺陷识别、篡改图像检测等。

E-mail:65891080@qq.com

郭松杰,硕士研究生,主要研究方向为目标检测。

E-mail:1945726067@qq.com