



计算机科学

COMPUTER SCIENCE

基于局部和全局特征表示的小样本绝缘子缺陷检测

崔克彬, 胡真真

引用本文

崔克彬, 胡真真. [基于局部和全局特征表示的小样本绝缘子缺陷检测](#)[J]. 计算机科学, 2025, 52(6): 286-296.

CUI Kebin, HU Zhenzhen. [Few-shot Insulator Defect Detection Based on Local and Global Feature Representation](#) [J]. Computer Science, 2025, 52(6): 286-296.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于显著性掩模混合的小样本图像分类](#)

Saliency Mask Mixup for Few-shot Image Classification

计算机科学, 2025, 52(6): 256-263. <https://doi.org/10.11896/jsjcx.240600123>

[面向工业品缺陷检测的对比表示学习](#)

Contrastive Representation Learning for Industrial Defect Detection

计算机科学, 2025, 52(1): 210-220. <https://doi.org/10.11896/jsjcx.240100202>

[小样本飞机生产质量偏差数据分析与预测方法研究](#)

Study on Analysis and Prediction Method of Small Sample Aircraft Production Quality Deviation Data

计算机科学, 2024, 51(11A): 240300123-8. <https://doi.org/10.11896/jsjcx.240300123>

[一种EO-YOLOX输电线绝缘子检测方法](#)

EO-YOLOX Model for Insulators Detection in Transmission Lines

计算机科学, 2024, 51(11A): 240200107-6. <https://doi.org/10.11896/jsjcx.240200107>

[基于对称卷积块网络和原型校准的小样本学习方法](#)

Few-Shot Learning Method Based on Symmetric Convolutional Block Network and Prototype Calibration

计算机科学, 2024, 51(11): 182-190. <https://doi.org/10.11896/jsjcx.230900022>

基于局部和全局特征表示的小样本绝缘子缺陷检测

崔克彬^{1,2} 胡真真¹

1 华北电力大学计算机系 河北 保定 071003

2 复杂能源系统与智能计算教育部工程研究中心 河北 保定 071003

(ncepuckb@163.com)

摘要 为解决绝缘子缺陷样本数量少且缺陷目标小导致目前绝缘子缺陷检测精度偏低这一问题,提出一种结合 CNN 与 Transformer 的小样本目标检测模型(C-TFSIDD),通过融合图像局部和全局特征来更有效地实现绝缘子缺陷检测。首先,采用融合 CNN 局部细节捕捉能力与 Transformer 全局信息整合能力的 Next-ViT 作为特征提取模块,精准捕获绝缘子图像局部和全局特征信息;其次,采用改进路径聚合特征金字塔网络(Path Aggregation Feature Pyramid Network,PAFPN)进行双向多尺度特征融合,增强底层特征表示,以改善小目标的检测效果;最后,提出一个基于度量的判别性损失函数,在微调阶段优化分类器学习更具判别性的特征表示,以增加类别之间的可分性,减少类内变化的影响。在两个公开的绝缘子缺陷数据集上进行训练和评估,实验结果表明,与基线模型 TFA 相比,C-TFSIDD 在样本为 5shot,10shot,20shot 的检测结果分别提升 28.7%,35.5%,47.7%;与小样本目标检测模型 FSCE 相比,C-TFSIDD 分别提升 21.8%,26.7%,21.1%。结果表明,C-TFSIDD 能有效提升小样本条件下的绝缘子缺陷检测精度。

关键词: 缺陷检测;绝缘子;小样本;CNN-Transformer;度量学习

中图分类号 TP391;TM216

Few-shot Insulator Defect Detection Based on Local and Global Feature Representation

CUI Kebin^{1,2} and HU Zhenzhen¹

1 Department of Computer, North China Electric Power University, Baoding, Hebei 071003, China

2 Engineering Research Center of Intelligent Computing for Complex Energy Systems, Ministry of Education, Baoding, Hebei 071003, China

Abstract In order to solve the problem that the small number of insulator defect samples and small defect targets lead to the current low accuracy of insulator defect detection, this paper proposes a few-shot object detection model(C-TFSIDD) combining CNN and Transformer, which fuses local and global features of images to realize insulator defect detection more effectively. Firstly, Next-ViT, which integrates the local detail capture capability of CNN and the global information integration capability of Transformer, is used as the feature extraction module to accurately capture local and global feature information of insulator images. Secondly, the improved path aggregation feature pyramid network(PAFPN) is used for bidirectional multi-scale feature fusion to enhance the underlying feature representation and improve the detection effect of small targets. Finally, a metric-based discriminative loss is proposed to optimize the classifier in the fine-tuning stage to learn more discriminative feature representations to increase the separability between classes and reduce the effect of intra-class variations. Trained and evaluated on two public insulator defect datasets, the experimental results show that C-TFSIDD improves the detection results with samples of 5shot, 10shot, and 20shot by 28.7%, 35.5%, and 47.7%, respectively, compared to the baseline model TFA, and compared with the few-shot object detection model FSCE, the proposed method improved by 21.8%, 26.7%, and 21.1%, respectively. The results show that C-TFSIDD can effectively improve the defect detection accuracy of few-shot insulator samples.

Keywords Defect detection, Insulator, Few-shot, CNN-Transformer, Metric learning

到稿日期:2024-03-20 返修日期:2024-07-05

基金项目:国家电网有限公司总部管理科技项目(5700-202340289A-1-1-ZN)

This work was supported by the State Grid Corporation of China Headquarters Management Science and Technology Project(5700-202340289A-1-1-ZN).

通信作者:胡真真(zzjtcbmlm@163.com)

1 引言

作为电网系统中的关键组成部分,绝缘子的性能直接影响着电网的安全稳定运行^[1]。然而,由于长期暴露在复杂的自然环境中,绝缘子受到温湿度、雷电、强电场、污秽等自然灾害的影响,常常出现掉串、破损、闪络、覆冰等问题^[2]。这些绝缘子的缺陷会对电力系统的安全性和稳定性造成严重影响,甚至引发火灾和重大事故。因此,对绝缘子缺陷进行及时检测和维修具有十分重要的意义。

近年来,随着无人机技术的飞速发展,无人机巡检已经成为一种有效的绝缘子缺陷检测方式^[3],其高清晰度的绝缘子图像数据为基于深度学习的目标检测提供了优质的数据基础。基于候选框和基于回归框的目标检测算法在绝缘子检测任务中均取得了良好的结果。Ma等^[4]采用一阶段目标检测算法YOLO系列^[5]优化损失函数检测大小不一、背景复杂的绝缘子,并通过自适应旋转操作全面学习绝缘子的特征,从而提高识别精确度。Zhao等^[6]在两阶段目标检测算法Faster-RCNN^[7]的基础上加入注意力机制并改进区域建议网络,以满足无人机巡检下的绝缘子检测需求。然而,无人机拍摄的大量绝缘子图像存在着不同的背景干扰,并且绝缘子缺陷图像的数量相对较少,很难建立一个包含各种缺陷类型的完整数据集,从而导致绝缘子缺陷检测算法受限于训练数据的多样性和数量^[8]。此外,无人机拍摄的绝缘子图像通常尺寸大、分辨率高,但缺陷目标较小,这就需要更高的检测精度和准确性。针对视觉检测技术对图像质量和数据量的依赖,Zhao等^[9]通过全聚焦方法(TFM)对复合绝缘子内部缺陷进行超声检测,直接从物理信号层面定位缺陷,避免了背景干扰和数据集限制问题,为绝缘子内部缺陷检测提供了新的技术路径。但是,在绝缘子缺陷图像样本量不足的情况下,准确检测出绝缘子缺陷仍然是一个具有挑战性的任务。

针对绝缘子缺陷样本数据稀缺的问题,提出一种基于小样本学习技术的目标检测方法,其仅使用少量绝缘子缺陷样本进行模型训练,旨在实现对绝缘子缺陷的准确检测。基于迁移学习方法的小样本目标检测模型TFA^[10](Two-stage Fine-tuning Approach)具有简单直观的特点,只需要少量样本对模型进行微调,就能取得超过元学习方法的效果。因此,本文选择TFA模型来实现小样本绝缘子缺陷检测。TFA模型虽然在公共数据集上的检测效果良好,但在绝缘子缺陷数据集上的检测效果并不理想。主要原因有:首先,公共数据集中图像尺寸均在600像素左右,而由无人机拍摄的绝缘子缺陷图像的尺寸大、分辨率高;其次,公共数据集中的目标占比适中,而在绝缘子图像中掉串、破损、闪络等缺陷目标在原图中的占比小,多为小目标;此外,绝缘子图像背景复杂,多为森林、农田、城市等,对模型抗背景干扰能力要求较高。上述原因都会影响TFA模型在绝缘子缺陷检测上的精确度。因此,本文对TFA模型进行改进,根据绝缘子缺陷图像的特点提出一种新的小样本目标检测模型C-TFSIDD(CNN-Transformer based Few Shot Insulator Defect Detection),使模型深入学习

到绝缘子缺陷目标的局部和全局特征,减少绝缘子缺陷的漏检、错检问题。

本文的主要工作有:

1)引入一种融合CNN与Transformer的Next-ViT网络^[11]替换传统的ResNet101作为特征提取网络,充分挖掘缺陷目标的全局和局部特征信息,提升特征提取网络的学习能力;

2)采用改进的路径聚合特征金字塔PAFPN网络^[12]融合低层次的空间信息,进一步加强绝缘子图像中缺陷小目标的图像特征,提升检测效果;

3)设计基于度量学习的判别性损失函数^[13],提高同类缺陷目标的紧密性以及不同缺陷目标的区分度,使得同类样本更接近而异类样本更远离。

2 相关工作

目前,小样本目标检测的主要方法大致分为4种:基于元学习(Meta Learning)的方法、基于迁移学习(Transfer Learning)的方法、基于数据增强的方法,以及基于度量学习的方法^[14]。基于元学习的方法被广泛认为是一种具有潜力的技术,与传统的学习方法所关注的学习结果不同,它的关注点在于如何学习,即从学习的过程中获取最佳的学习策略。Shi等^[15]提出了一种基于元学习的输电线路小样本缺陷检测方法(Meta PowerNet),该方法采用元注意RPN和特征重构模块,既能关注到小型缺陷目标特征,又能利用元学习器捕获到与缺陷特征相关的特征通道。基于迁移学习的方法通过在大规模基类数据集上进行预训练,并在包含新类和基类数据的数据集上微调部分模块参数,来提高有限新类数据集的泛化能力。微调方法是基于迁移学习的核心方法,采用比较简单的两阶段训练策略。TFA模型同样采用此方法,在第一阶段,该模型采用基类数据进行预训练,以提高模型的泛化性能和特征抽取能力;在第二阶段中,冻结图像特征提取器和RPN网络的参数,微调更新其分类网络和回归网络,以提高针对特定目标的检测任务的检测性能。Zhai等^[16]利用迁移学习在大量无缺陷的绝缘子数据集上训练Faster-RCNN目标检测算法,再使用训练后的模型参数在少量样本的缺陷绝缘子图像上进行训练,从而实现小样本条件下的绝缘子缺陷的分类及定位。基于数据增强的方法就是通过对样本进行旋转、裁剪、颜色变换、添加噪声等生成新的样本数据,但实际上并没有生成新目标。Cui等^[17]利用生成模型CycleGAN在原图训练过程中学习原始数据的分布,并添加随机噪声生成接近样本真实数据的绝缘子缺陷图像,实现对绝缘子缺陷样本数量的扩增。基于度量学习的方法将目标样本映射到一个紧凑的度量空间,计算相似度或者度量距离对目标进行分类,能够更好地处理目标之间的差异和变化。Karlinsky等^[18]提出一种新的度量学习方法,并将其应用于目标检测模型的分头,通过检测输入图像的特征与目标特征的相似度,完成目标分类。

3 C-TFSIDD模型

本文采用基于迁移学习的小样本目标检测模型TFA作

为基线模型,目标检测算法为 Faster R-CNN,该算法主要由特征提取模块、RPN 模块以及分类和定位模块 3 部分组成。模型采用两阶段训练法,为了能够学习到绝缘子的相关特征,在第一阶段选取正常绝缘子图像和 Pascal VOC 数据集^[19]中的 15 类图像共计 16 类作为基类数据集;第二阶段使用部分基类样本与少量新类绝缘子缺陷样本微调所提网络模型。本模型中的特征提取模块使用结合 CNN 与 Transformer 优点的 Next-ViT 网络替代 ResNet-101 作为主干网络提取图像

特征。针对绝缘子缺陷目标尺寸小的问题,提出一个改进的特征金字塔网络(PAFPN)模块。将主干网络的前两层特征层与其同尺度自底向上的特征层进行跳跃连接,最大程度地获取基类和新类数据的多尺度特征信息,增强小目标的信息表示,并全面获取多尺度上下文信息。除此之外,对模型损失函数进行优化,在小样本微调阶段使用基于度量的判别性损失,加强样本集中类内紧凑性和类间可分性,提升小样本场景中的检测性能。模型的总体结构如图 1 所示。

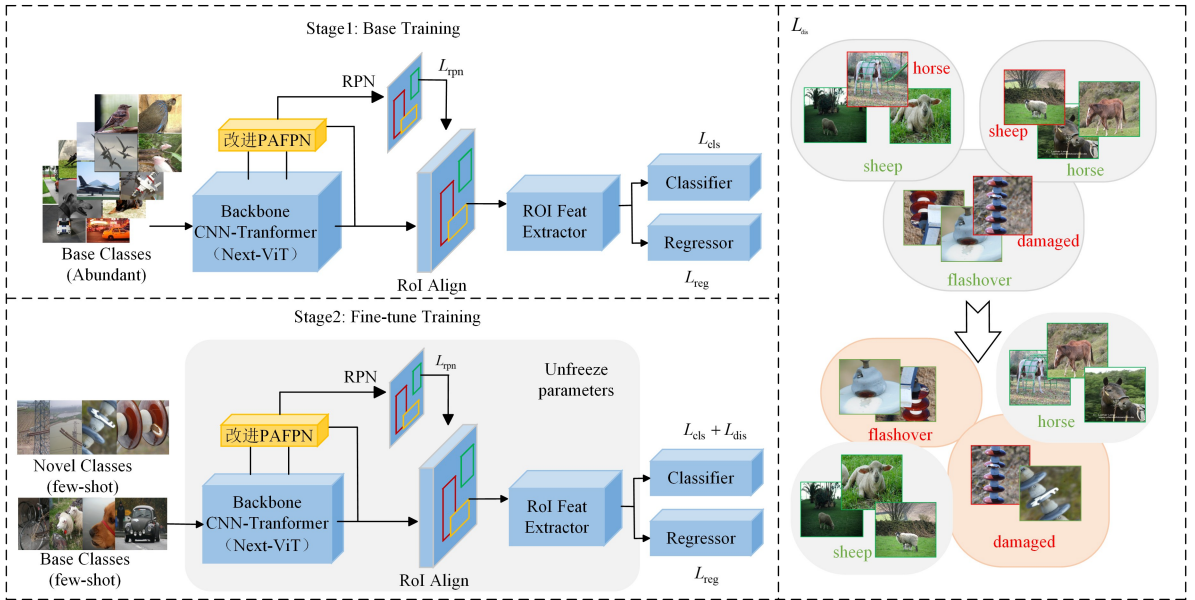


图 1 C-TFSIDD 的整体结构

Fig. 1 Overall structure of C-TFSIDD

3.1 特征提取模块

近几年里,卷积神经网络 CNN 作为特征提取网络被广泛用于绝缘子缺陷检测任务^[20]。然而,CNN 结构的感受野小,更加关注局部信息,对全局信息的关注较弱,导致长距离上下文信息容易丢失^[21]。例如,针对同一绝缘子图像中两处相隔较远的闪络缺陷,基于 CNN 的 ResNet 网络在特征提取时可能无法捕捉到共同特征。为解决这一问题,一些研究人员开始探索使用 Transformer 模型^[22]替代传统的卷积神经网络。Transformer 模型采用自注意力机制,能够有效地捕捉输入序列中的长距离依赖关系。在绝缘子检测任务中,使用 Transformer 模型能够更好地建模目标之间和目标与背景之间的长距离依赖关系,提升模型在全局信息上的捕获能力。Xi 等^[23]使用 Swin-Transformer 取代 ResNet101 作为新的特征提取网络来检测输电线路中的绝缘子,以获得更丰富的特征;Guo 等^[24]将单阶段目标检测算法 YOLOv5 的主干网络替换成 Transformer-CSP 网络,通过跨不同深度的网络特征图捕获信息来提升网络的能力。尽管 Transformer 在目标检测领域有一定的发展,但复杂的注意力机制和模型设计,使得它在效率方面难以像 CNN 那样高效。因此,本文选取基于 CNN 与 Transformer 混合结构的主干网络 Next-ViT,旨在提高模型计算速度的同时,充分利用图像中的局部与全局信息,提升小样本条件下绝缘子缺陷检测的精度。

分层金字塔体系结构的 Next-ViT 的网络结构如图 2 所示。由于传统的基于 Transformer 的特征提取网络只是在网络最后一层或者几层加入 Transformer 模块,在浅层阶段无法捕捉全局信息,因此 Next-ViT 在网络结构的设计中采用新的混合策略,在不同的阶段 stage,顺序叠加 N 个 NCB (Next Convolution Block) 模块和 1 个 NTB (Next Transformer Block) 模块来提升局部和全局特征的提取能力。其中,NCB 完成局部特征的提取,NTB 对更大范围的上下文信息进行捕获并完成局部与全局信息的融合,以此提高模型对缺陷目标的感知能力。

NCB 由 MHCA (Multi-Head Convolutional Attention) 和 MLP 组成。MHCA 为多头卷积注意力模块,其通过整合不同通道的信息,学习到不同头所关注的特征子空间,以有效捕捉局部的特征信息,提升对绝缘子缺陷区域的关注能力。具体来说,在分组卷积 (Group Convolution) 中,输入特征图的通道被分成多个组,每组内的通道只与该组内的卷积核进行卷积操作。这种分组策略允许不同组的通道独立学习,相当于不同的“头”各自关注输入数据的不同特征。根据卷积操作的特性,每个组内的卷积核专注于图像的局部窗口进行特征提取。这种局部感受野内的独立学习更加适宜捕获小尺寸缺陷目标的细节特征,使得网络能够专注于图像中最有信息量的部分,从而使得模型能够更加精确地识别和响应缺陷区域。

进一步地,卷积核的权重系数在训练过程中的动态更新,使得网络能够更加精确地调整其对局部特征的敏感度,提升模型对微小缺陷目标的关注度。不同的组可能专注于学习不同类型的特征,如纹理、边缘或形状特征,所以在分组卷积之后,使用 1×1 的点卷积来整合不同通道的信息,使得整个网络能够更全面地捕捉到缺陷目标的特征信息,并且有助于模型从多角度理解缺陷区域,增强对缺陷区域的关注度。MHCA模块的计算过程总结为:

$$MHCA(z) = \text{Concat}(\text{GroupConv}(z_1, z_2, \dots, z_n))W^P \quad (1)$$

其中, z_1, z_2, \dots, z_n 为输入特征 z 的 n 个子组, W^P 表示对拼接后的特征进行 1×1 卷积操作。

NTB是一个轻量级模块,以高效的多头自注意力模块EMHSA(Efficient Multi-Head Self Attention)为核心。由于绝缘子缺陷可能会在图像中的离散位置上出现,因此,利用该模块的自注意力计算能够有效地建模并捕捉到缺陷的上下文关系。自注意力机制通过独立生成每个序列元素的Query(Q),Key(K),Value(V)向量,使模型能够捕捉序列内元素间的复杂关系。通过计算 Q 与 K 的点积得到注意力分数,这些分数反映了元素间的相互依赖性。利用这些分数对 V 向量进行加权求和,模型构建出序列的全局关联视图。这个过程允许模型在序列的不同位置间建立动态的联系,无须考虑

它们之间的物理距离,允许模型区分哪些部分是当前上下文中更为重要的,具有全局感受野。该机制的多头并行处理能力进一步丰富了模型的表示,使得Next-ViT能够并行学习输入数据的多个子空间表示,有效地建模并捕捉到绝缘子缺陷的上下文关系。NTB将E-MHSA每个头的输出和MHCA的输出特征进行拼接,实现局部细节与全局上下文信息的有效融合。这种策略不仅增强了模型对图像中关键信息的捕捉能力,而且显著提升了图像中有效信息的利用率,能更好地捕捉到绝缘子缺陷的关键特征。

E-MHSA模块中自注意力计算的定义如式(2)所示:

$$\text{Attention}(Q, K', V') = \text{softmax}\left(\frac{QK'^T}{\sqrt{d_k}}\right)V' \quad (2)$$

具体过程为:上一模块输入特征经过线性变换映射成 Q, K, V 向量, Q, K, V 经过平均池化再次线性变换映射成 K', V' 向量;再将查询向量 Q 与键向量的转置 K'^T 相乘,并通过除以键向量维度 d 的平方根对结果进行缩放,消除数值影响;最后对缩放后的结果进行softmax函数归一化,再与值向量 V' 进行相乘得到加权向量,获得最终的上下文向量表示。通过自注意力机制对整个图像建立全局关注,并对细节和上下文进行准确的建模,以捕捉缺陷与周围环境之间的关联。

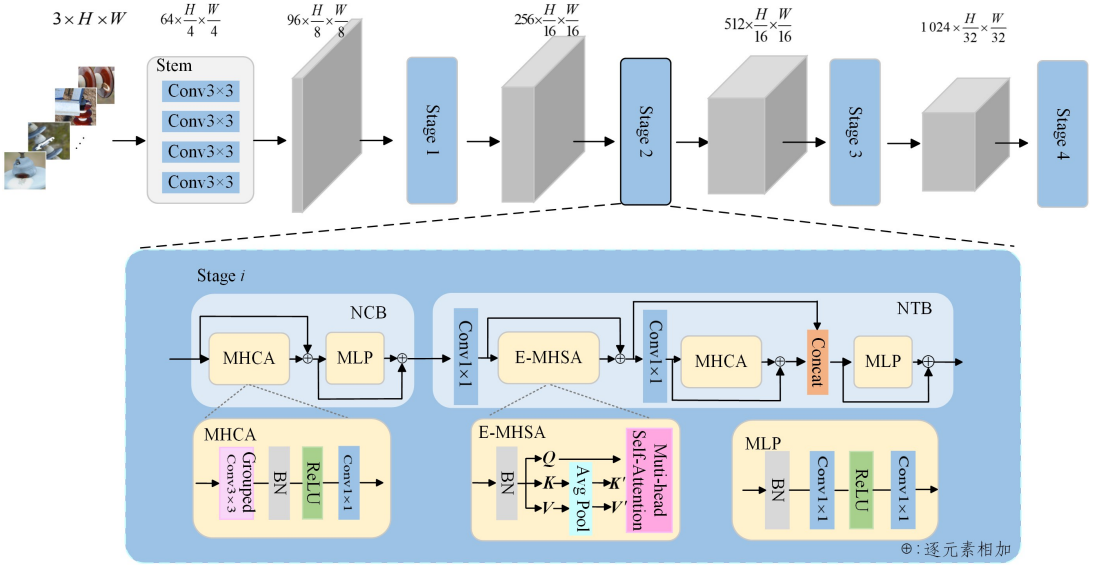


图2 特征提取网络结构

Fig. 2 Structure of feature extraction network

3.2 改进PAFPN模块

在绝缘子缺陷检测任务中,背景的复杂性和缺陷目标的小尺寸特性,使得精确识别缺陷区域面临重大挑战。传统的特征金字塔网络(Feature Pyramid Network, FPN)虽然在多尺度特征提取方面表现出色,但随着网络层数的加深,特征图尺寸递减,细小目标的重要细节信息可能丢失,而这些细小目标对于缺陷目标的精确定位至关重要。改进PAFPN模块是一种针对绝缘子缺陷检测任务的改进型多尺度特征金字塔网络。PAFPN模块对特征提取网络各阶段的特征图采用两层融合策略。首先,通过自上向下的连接,将包含丰富语义信息

的深层特征图进行上采样后与下一层特征图相加融合,在保留浅层特征细节信息的同时,引入深层的语义信息,提升网络对小目标的识别能力。在这个跨层级连接的特征融合过程中,不断将更高层级的语义信息传递给低层级特征图,使得低层级特征既包含自身拥有的目标细节信息,又获得更丰富的目标上下文信息,增强对绝缘子缺陷目标的语义表示能力。然后,通过自下向上的连接,将首次融合后的浅层特征图进行下采样后与其上一层特征图进一步相加融合,增强深层特征图的定位信息,缓解深层特征图中小目标细节信息易丢失的问题。经过自上向下的语义增强和自下向上的细节传播这两

次特征融合,改进 PAFPN 模块增强了不同特征图之间的互补性,为网络提供了一个更加全面和丰富的特征表示,能够同时捕获到信息丰富的多尺度特征表示,有助于网络对图像内容的整体理解,这对于精确检测在不同尺度上表现的绝缘子缺陷目标至关重要。

除此之外,考虑到浅层网络特征图的空间分辨率较高,在图像中编码了更细粒度的信息,可以更好地定位和识别小目标。还在同一尺度的浅层和深层特征之间增加了跳跃连接线,用于实现第一阶段和第二阶段输出特征的再次融合。此举的目的在于增加底层信息的利用度,使网络的输出节点更大程度地保留浅层小目标的重要信息,并增强微小缺陷特征的表达力^[25]。改进 PAFPN 的具体结构如图 3 所示,其中 $C_1 - C_4$ 为浅层特征, $P_1 - P_4$ 为中间层特征, $N_1 - N_4$ 为深层特征,虚线为浅层与深层特征的跳跃连线。

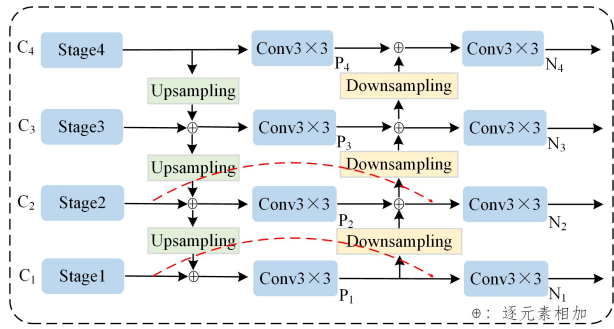


图 3 改进 PAFPN 模块

Fig. 3 Improved PAFPN module

3.3 基于度量的判别性损失

在输电线路场景中,一些目标具有类内方差大的特点,即同一个类别的不同样本之间可能存在一些差异。例如,同属破损缺陷的绝缘子可能会呈现不同的损坏程度,闪络缺陷出现的位置及大小也有所差异。这些都会导致绝缘子缺陷呈现出多样性的外观,因此模型需要抓住它们的共同特征并准确地区分不同的绝缘子缺陷。为减小这种目标间巨大的类内方差,TFA 将传统的基于 FC 的分类器替换为基于余弦相似性的分类器(见式(3)),但仍会因为样本数据稀缺存在偏差。

$$s_{i,k} = \frac{\alpha \mathbf{F}(x)_i^T \mathbf{w}_k}{\|\mathbf{F}(x)_i\| \|\mathbf{w}_k\|}, p_{i,y_i} = \frac{e^{s_{i,y_i}}}{\sum_{k=0}^C e^{s_{i,k}}} \quad (3)$$

其中, $\mathbf{F}(x)$ 为输入样本的特征; \mathbf{w}_k 代表类 k 的权重向量; 固定比例因子 $\alpha = 20$, 其值为原 TFA 模型经过多次实验后得出, 本文沿用该值; $\|\cdot\|$ 为范数操作。将第 i 个 RoI 特征向量与类的权重向量作内积, 并除以各自范数得到余弦相似性分数。 p_{i,y_i} 表示第 i 个 RoI 特征属于类 k 的概率, $C = C_{\text{基类}} + C_{\text{新类}}$ 表示数据集中的类别总数量。

考虑到度量样本之间的相似度值是提高绝缘子缺陷检测精度的关键, 提出一种基于度量的判别性损失函数。在模型微调阶段, 通过参数更新学习一个更具判别性的分类器, 使得样本与其对应的类原型的相似度大于与其他类原型的相似度^[26], 帮助模型从有限的样本中学习到更泛化的特征表示, 从而提高模型在识别属于同类但形态相差大的缺陷目标时的

精度, 减少误检情况。换句话说, 相似的样本更接近, 不同的样本更易区分, 从而达到预期性能。

对于每个样本, 该损失函数将分类器的学习权重向量 \mathbf{w}_k 重新定义为对应类别原型向量, 将样本 x 属于类 c 的预测概率作为样本 x 与类原型之间的余弦相似度得分, 以期达到每个样本与所对应的类原型的相似度大于与其他类原型的相似度。这里的 x 和 c 均为泛指。利用上述余弦相似性分数对损失函数进行描述, 如式(4)所示:

$$L_{dis_i} = \sum_{k=0, k \neq y_i}^C -\log(\max\{p_{x_i, y_i} - p_{x_i, k}, 0\} + \epsilon) \quad (4)$$

其中, $\epsilon = 1 \times 10^{-7}$ 以保持数值的稳定性; p_{x_i, y_i} 和 $p_{x_i, k}$ 由式(3)替换 \mathbf{w}_k 后计算得出, 分别表示样本 x_i 与所属类 y_i 的原型之间的余弦相似性得分以及样本 x_i 与除所属类别 y_i 之外的原型的余弦相似性得分, 两者的差值将映射到一个负无穷到 0 的区间内, 结果越接近 0 (即同类样本相似度越高), 损失就越大, 从而鼓励模型提高同类样本的得分。

此外, 在微调阶段, 模型的学习偏差可能导致基类样本的分类相对容易, 而对于新增的新类样本, 模型的预测效果可能存在偏差, 使得这些新类样本的损失贡献较低。为了解决这个问题, 在损失函数中引入聚焦因子来动态调整每个样本的损失贡献。聚焦因子的定义如式(5)所示:

$$factor_i = (1 - p_i)^\gamma \quad (5)$$

其中, p_i 表示第 i 个样本对应真实类别的预测概率, $\gamma = 4$ 是聚焦参数。

通过聚焦因子, 模型可以更关注于具有挑战性的样本, 包括那些属于新类的目标。这个调整过程可以理解为让损失函数更加专注于难以分类的样本, 因为它们相对于其他样本更有价值。在损失函数的优化过程中, 模型会更加重视难以处理的样本, 将相似样本的距离尽可能缩小, 学习其共同特征; 同时将不同样本尽可能远离, 增强它们之间的差异性, 以帮助模型更准确地识别和分类绝缘子上的缺陷。因此, 基于度量的判别损失最终表示为:

$$L_{dis} = \frac{1}{N} \sum_{i=1}^N (L_{dis_i} \cdot factor_i) \quad (6)$$

其中, N 表示训练样本数。

3.4 损失函数优化

目标检测模型的损失函数通常由两部分组成: 分类损失和回归损失。分类损失用于评估模型对目标类别的准确性, 分类损失越小, 代表模型的分类准确性越高。回归损失用于评估模型在目标定位上的准确性, 回归损失越小, 代表模型的定位准确性越高。在绝缘子缺陷检测任务中, 网络在小样本微调阶段仅使用来自每个类的 K 个样本进行训练, 绝缘子缺陷样本的数量相对较少, 会存在正负样本之间类别不平衡的问题, 导致模型可能会偏向于多数负样本, 而难以学习到少数正样本的特征。

为此, 在分类损失函数中使用 Focal Loss 来替代传统的交叉熵损失函数, 通过动态调整损失函数的权重, 巧妙地将模型的焦点从易分类的样本转移到位于决策边界、较难分类的样本上, 有效地增强模型对难以识别样本的学习, 从而提高缺

陷目标分类的准确性,进而有效应对类别不平衡问题,提高难例样本的训练效果^[27]。Focal Loss 的数学表达式如下:

$$FL = \begin{cases} -\sigma(1-p)^\psi \log p, & \psi=1 \\ -(1-\sigma)p^\psi \log(1-p), & \text{其他} \end{cases} \quad (7)$$

其中, σ 用于平衡正负样本的权重, 本文取 0.5; ψ 用于控制降低易分类样本损失的速度, 本文取 2; p 表示模型对样本的预测概率。

因此, 模型在两阶段训练中的损失函数为:

$$L_1 = L_{rpn} + L_{cls} + L_{reg} \quad (8)$$

$$L_2 = L_{rpn} + L_{cls} + \beta \times L_{reg} + L_{dis} \quad (9)$$

其中, L_1 为第一阶段的损失函数, L_2 为第二阶段的损失函数; β 表示分类损失在回归损失中所占的权重比例, 由于在微调中额外引入基于度量的辅助损失函数用于分类任务, 此时设置 $\beta=2$ 来平衡这两个任务。

4 实验结果与分析

4.1 数据集

本文将数据集分为基类数据集和新类数据集, 并且基类、新类的类别无交叉。模型在微调阶段训练时, 只使用新类数据集中 K 个已标注的新类实例 ($K=5 \text{ shot}, 10 \text{ shot}, 20 \text{ shot}$), 其余实例作为测试集用于模型测试阶段。

4.1.1 基类数据集

基类数据集选取 Pascal VOC 2007+2012 数据集中的 15 个类别图像和中国电力线路绝缘子数据集 (CPLID)^[28] 中的正常绝缘子图像。其中, VOC 数据集中的 15 个类别图像用于捕捉各种不同类别目标的共性特征, 学习有效的特征表示; 正常绝缘子图像加入基类数据集, 使模型能够学习到相关绝缘子的特征信息, 辅助绝缘子缺陷目标检测。基类数据集的类别和各类实例数量如表 1 所列。

表 1 基类数据集的各类统计信息

Table 1 Statistical information of the base class dataset

Class	Number of instances
aeroplane	1285
bicycle	1208
boat	1397
bottle	2116
car	4008
cat	1616
chair	4338
diningtable	1057
dog	2079
horse	1156
person	15576
pottedplant	1724
sheep	1347
train	984
tvmonitor	1193
insulator	1073
total	42157

4.1.2 新类数据集

将 CPLID 数据集中的缺陷图像和公开绝缘子缺陷数据集 (ID) 作为新类数据集, 缺陷类型包括掉串、破损、闪络, 掉串缺陷图像 248 张, 破损缺陷图像 312 张, 闪络缺陷图像 400 张。

其中, 掉串缺陷图像选自 CPLID 数据集, 破损和闪络缺陷图像选自网络上的公开数据集并自行标注。新类数据集中包含的种类及对应数量如表 2 所列。正常绝缘子图像和各类绝缘子缺陷图像示例如图 4 所示, 其中破损和闪络缺陷展示的是对原图裁剪后的图片。

表 2 绝缘子缺陷数据集信息

Table 2 Insulator defect dataset information

Class	Number of instances
defect(掉串)	248
damaged(破损)	383
flashover(闪络)	862
total	1538



(a)正常 (b)掉串 (c)破损 (d)闪络

图 4 绝缘子缺陷样本图像

Fig. 4 Insulator defect sample images

4.2 实验配置

表 3 列出了实验中具体使用的硬件和环境。

表 3 实验硬件和环境

Table 3 Experimental hardware and environment

配置	参数
操作系统	Ubuntu 18.04
GPU	Nvidia GeForce RTX 2080Ti 显卡(11GB)
实验环境版本	Python3.8, PyTorch1.8, Cuda10.2
实验框架	MMDetection2.26.0

特征提取网络模块使用 ImageNet 数据集上的预训练 Next-ViT 模型。针对两阶段训练中数据集图像尺寸大小不同的问题, 基训练时, 将图片调整为 640×640 像素; 微调训练时, 将图片调整为 1333×800 像素。具体训练参数设置如表 4 所列。

表 4 训练参数设置

Table 4 Training parameter settings

参数	设置
基训练学习率(lr)	0.0001
微调学习率(lr)	1×10^{-5}
衰减权重	0.0001
优化器	AdamW
基训练批量大小(Batch Size)	8
微调批量大小(Batch Size)	2
基训练 epoch	12
微调 Iter	15000

在基训练阶段, 将第 8 和第 11 个 epoch 学习率乘 0.1 进行学习率的调整; 在微调阶段, 将网络所有模块参数解冻, 重新参与模型训练并随机初始化新类分类头权重^[29], 分类头的类别数由 16 改成 19, 通过 15000 次迭代微调基训练后的模型参数; 在推理阶段, 将非最大抑制(NMS)的置信阈值和 IoU 阈值分别设置为 0.05 和 0.5。

4.3 评价指标

为充分评估 C-TFSIDD 算法在目标检测任务中的整体性能,采用目标检测中常用的平均精度值(AP)指标来评估特定类别的检测效果。为获得更全面的评价,使用平均 AP 值(mean Average Precision, mAP)这一综合指标来总结所有类别目标检测的平均精度表现。将基类的平均检测精度称为 bAP,新类的平均检测精度称为 nAP。

4.4 实验结果对比分析

4.4.1 复杂背景下绝缘子缺陷检测对比分析

为验证 C-TFSIDD 模型在小样本绝缘子缺陷检测中的有效性,针对综合多种不同背景的绝缘子缺陷图像数据集进行实验。该数据集具有丰富的背景信息,并且包含多种绝缘子缺陷情况。本研究选择 Pascal VOC 数据集集中的 15 个类别以及 CPLD 数据集集中的正常绝缘子作为基训练样本。随后,重塑 bbox head 权重。在新的 bbox head 中,基础类别的权重直接使用基训练阶段的权重进行初始化,而新类别的权重则进行随机初始化。最后,在 3 类绝缘子缺陷数据集上进行微调,并与 FSCE^[30], FsDetView^[31], Meta-RCNN^[32], MSPR^[33] 和 TFA^[10] 模型进行对比,其中 K 取 5 shot, 10 shot, 20 shot。表 5 列出了在阈值为 0.5 时数据集中基类、新类以及所有类别的小样本检测性能。

表 5 不同模型在绝缘子缺陷数据集上的检测结果对比

Table 5 Detection results comparison of different models on insulator defect dataset

Shot	Method	Backbone	(%)		
			bAP50	nAP50	mAP
5	FSCE ^[30]	ResNet101	57.2	16.8	50.8
	FsDetView ^[31]	ResNet101	31.2	22.5	29.8
	Meta-RCNN ^[32]	ResNet101	37.1	20.5	34.5
	MSPR ^[33]	ResNet101	31.6	9.5	28.1
	TFA ^[10]	ResNet101	49.2	9.9	43.0
	C-TFSIDD(Ours)	Next-ViT	46.9	38.6	45.6
10	FSCE ^[30]	ResNet101	60.3	25.6	54.8
	FsDetView ^[31]	ResNet101	41.5	26.5	39.1
	Meta-RCNN ^[32]	ResNet101	42.9	27.3	39.9
	MSPR ^[33]	ResNet101	36.6	22.3	34.4
	TFA ^[10]	ResNet101	63.0	16.8	55.7
	C-TFSIDD(Ours)	Next-ViT	64.3	52.3	62.4
20	FSCE ^[30]	ResNet101	59.3	45.4	57.1
	FsDetView ^[31]	ResNet101	46.7	35.1	44.8
	Meta-RCNN ^[32]	ResNet101	44.7	36.1	43.4
	MSPR ^[33]	ResNet101	40.4	32.4	39.2
	TFA ^[10]	ResNet101	68.1	18.8	60.3
	C-TFSIDD(Ours)	Next-ViT	68.4	66.5	68.1

在检测绝缘子缺陷时,C-TFSIDD 模型的新类检测结果(nAP50)最高,性能较其他模型有显著提升。由表 5 中数据可知,在 5 shot 训练集的设置下,C-TFSIDD 模型新类的平均检测精度达到 38.6%,高于次优结果 16.1%;在 10 shot 和 20 shot 训练集的设置下,随着 K 的数量的翻倍增长,新类的平均检测精度大幅提升,分别达到 52.3% 和 66.5%。然而,在同等样本量训练下,次优结果只有 27.3%, 45.4%。C-TFSIDD 模型在此基础上分别提升 25.0%, 21.1%, 综合性能明显优于其他模型,这一结果印证了本文模型在绝缘子缺陷检测上的可行性。

除此之外,实验结果表明,10 shot 训练下,C-TFSIDD 模型在新类的检测结果上取得了最大的提升。为了直观展示这一结果,将测试集的检测结果进行可视化,如图 6 所示。为保证实验结果的一致性,实验均在单一尺度训练下完成。

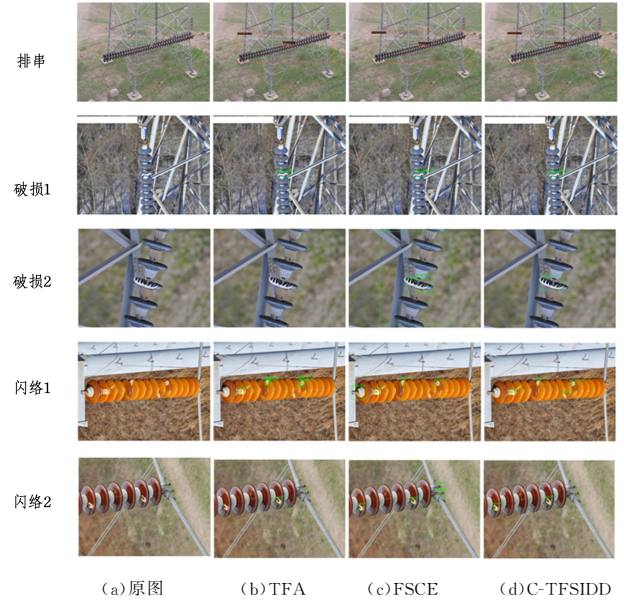


图 6 可视化结果

Fig. 6 Visualization results

本文给出了 10 shot 时基线模型 TFA,以及在基类、新类和所有类别检测精度上综合表现次优的 FSCE 模型和 C-TFSIDD 模型在绝缘子缺陷检测中的可视化结果,如图 6 所示,其直观展示出 C-TFSIDD 模型预测概率上的提升,以及在减少多个候选框生成和精确定位掉串缺陷、破损缺陷和闪络缺陷目标方面的优势。具体表现为,对于掉串缺陷和破损缺陷 1, C-TFSIDD 模型在对缺陷目标的预测概率都有明显提升,预测概率由 TFA 模型的 42% 提升到 82%;在绝缘子破损缺陷 2 检测任务中,基线模型 TFA 在样本量较少的情况下没有全面学习到破损类别的共同特征,导致出现漏检情况,而次优的 FSCE 模型在检测时出现多个目标框重叠的现象,并将图像中的目标错误预测为基类类别,在检测过程中更容易将预测结果偏向基类。而 C-TFSIDD 模型能很好地把握缺陷目标的共同特征,并准确定位出缺陷目标位置。当绝缘子目标上出现多处烧蚀缺陷时,如闪络缺陷 1 和闪络缺陷 2,由于其闪络的程度不同,因此缺陷呈现出的外观不同,具有类内差异大的特点。基线模型 TFA 和 FSCE 没有充分学习到目标的细节特征,在识别同一绝缘子目标上的多处闪络缺陷时均出现漏检的情况。此外,闪络缺陷 2 检测任务中,FSCE 模型误把输电线路上的螺栓识别为破损缺陷,没有精确定位缺陷目标。经改进后,C-TFSIDD 模型能够准确地检测出图像中的所有闪络缺陷,没有出现上述误检情况。

C-TFSIDD 模型之所以能够减少漏检和错检,主要得益于以下几个方面的创新:1)特征提取网络同时融合 CNN 的局部细节感知能力和 Transformer 的全局上下文信息的表示能力,比传统的 CNN 在特征提取时多了对图像全局信息的

把控;2)改进 PAFPN 模块优化了不同尺度的特征图融合过程,高层语义信息与低层细节信息得以充分利用;3)损失函数的优化,基于度量学习的损失函数能减小类内变异和扩大类间差异,使得模型在面对视觉上相似或变形的同类样本时,仍能够准确地进行识别和检测。在对这个具有多样化背景的数据集进行实验时,本文模型在绝缘子缺陷检测任务中表现出色。实验结果证明了该模型的有效性,尤其在样本量少的情况下,其能够显著提高缺陷检测精度。

新类中,各个缺陷类别在不同模型下的测试集检测结果如表 6 所列。

由表 6 中数据可知:

1)本文模型在绝缘子掉串缺陷的检测中在不同样本下的检测性能明显优于其他模型。其主要原因在于,绝缘子掉串缺陷图像与基类正常绝缘子图像的背景相似度高,而本文模型在训练过程中不拘泥于局部特征提取,能扩大感受野并捕获到绝缘子的全局特征,而基类图像的全局特征可以帮助模型更好地理解 and 识别新类图像,更多地关注掉串缺陷部分。

2)绝缘子破损及闪络图像规模大且目标尺寸小,这使得模型在小样本条件下难以提取有效特征,导致检测精度偏低。C-TFSIDD 模型利用改进后的 PAFPN 模块,将底层特征更多地融合到高层语义特征中,提高小目标的关注度,从而提升缺陷目标检测精度。除此之外,考虑到新类缺陷目标中破损与闪络存在类内方差大而类间可分离性低的特性,加入基于度量的判别性分类损失帮助模型优化目标检测分类结果。可以看出,C-TFSIDD 模型对绝缘子破损和闪络缺陷的检测精度较其他模型仍有所提高,表明该模型对少样本的小目标绝缘子缺陷具有更好的检测性能。

表 6 不同方法在 3 类绝缘子缺陷上的 AP 比较

Table 6 Comparison of AP with different methods on three types of insulator defects

(%)

Shot	Method	Novel class		
		defect	damaged	flashover
5	FSCE ^[30]	32.5	9.6	8.0
	FsDetView ^[31]	40.2	11.4	13.3
	Meta-RCNN ^[32]	43.3	9.1	9.1
	MSPR ^[33]	18.7	4.5	5.2
	TFA ^[10]	11.5	9.1	9.1
	C-TFSIDD(Ours)	90.1	12.2	13.8
10	FSCE ^[30]	51.7	12.4	12.5
	FsDetView ^[31]	59.2	15.7	13.9
	Meta-RCNN ^[32]	58.6	12.1	11.1
	MSPR ^[33]	46.3	10.0	9.9
	TFA ^[10]	30.2	9.8	10.2
	C-TFSIDD(Ours)	97.1	30.2	29.5
20	FSCE ^[30]	85.5	24.2	26.0
	FsDetView ^[31]	75.3	14.8	15.1
	Meta-RCNN ^[32]	72.7	21.5	14.0
	MSPR ^[33]	64.5	18.5	14.3
	TFA ^[10]	29.8	10.7	15.7
	C-TFSIDD(Ours)	98.6	48.2	52.6

4.4.2 单一背景下绝缘子缺陷检测对比分析

考虑到图像背景对绝缘子缺陷的影响,对 ID 数据集中的破损和闪络缺陷按照图像背景划分为森林、农田、城市 3 类,如图 7 所示。其中,城市背景以包含人造产物和自然景观为主。

通过进行缺陷检测实验,验证所提模型在各个场景的适用性。

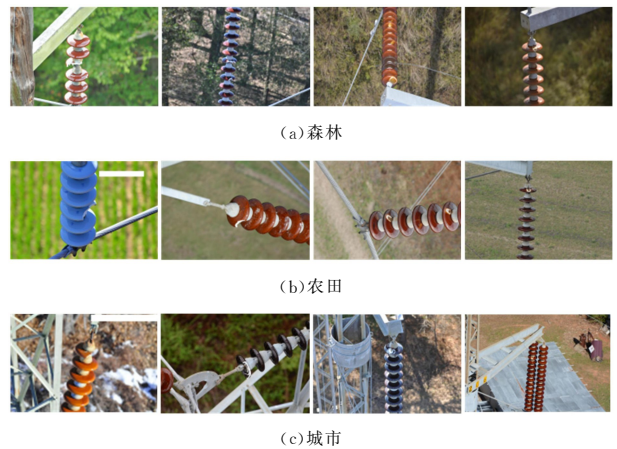


图 7 不同背景的绝缘子缺陷图像

Fig. 7 Insulator defect images in different backgrounds

为进一步验证 C-TFSIDD 模型在背景信息丰富且复杂的绝缘子缺陷数据集中的适用性,更加全面地评估模型的泛化性能,随机选取 3 类不同背景下的 $K=10$ 个绝缘子缺陷样本进行小样本微调实验。这样的实验设计,能够评估该模型在不同背景下的性能表现和对不同背景下绝缘子缺陷图像的分类准确性,以验证其对复杂绝缘子缺陷数据集的适应能力。实验结果如表 7 所列。

表 7 单一背景下的绝缘子缺陷 AP 和 nAP

Table 7 Insulator defects AP and nAP in single background

(%)

背景类别	Shot	Method	Novel class			
			defect	damaged	flashover	
森林	10	FSCE ^[30]	—	33.3	29.9	31.6
		FsDetView ^[31]	—	26.2	25.8	26.0
		Meta-RCNN ^[32]	—	23.5	22.7	23.1
		MSPR ^[33]	—	27.0	19.3	23.2
		TFA ^[10]	—	17.4	22.1	19.6
		C-TFSIDD(Ours)	—	50.1	52.1	51.1
农田	10	FSCE ^[30]	—	26.0	21.2	21.4
		FsDetView ^[31]	—	22.0	9.1	15.6
		Meta-RCNN ^[32]	—	25.5	9.1	17.3
		MSPR ^[33]	—	16.2	11.7	14.0
		TFA ^[10]	—	10.5	17.1	13.8
		C-TFSIDD(Ours)	—	48.3	52.6	50.5
城市	10	FSCE ^[30]	79.7	10.5	11.7	34.0
		FsDetView ^[31]	31.1	11.0	9.1	17.0
		Meta-RCNN ^[32]	66.6	6.1	9.1	27.2
		MSPR ^[33]	27.9	9.5	9.8	15.7
		TFA ^[10]	24.2	10.2	11.6	15.3
		C-TFSIDD(Ours)	90.9	22.2	35.2	49.4

注: defect 缺陷图像均为城市背景。

从表 7 中可以看出,与其他小样本目标检测模型相比,无论是在森林、农田还是城市的背景环境中,C-TFSIDD 模型在 10 个样本量训练后,都取得了较高的准确率。小样本目标检测模型存在一定的抗干扰能力,因此,在单一背景下的森林、农田环境中对绝缘子破损、烧蚀缺陷检测的精度高于多以包含人造产物和自然景观为主的背景中的检测精度。表中数据反映出城市背景下的绝缘子这两类缺陷的检测精度低于其他背景下的缺陷检测结果。但仍可以发现,C-TFSIDD 模型在城市背景中对破损、烧蚀缺陷的检测精度达到 22.2% 和

35.2%，高于次优模型 FSCE 的 11.7% 和 23.5%。绝缘子缺陷检测算法往往会应用于各种不同的场景和背景环境中，这些实验结果进一步验证了 C-TFSIDD 模型在小样本绝缘子缺陷检测任务上的优越性，表明该模型在特定场景中具有适应性、可靠性和有效性。

综上所述，C-TFSIDD 模型无论在复杂背景还是在单一背景的数据集中，都能够精准识别出大部分缺陷绝缘子。通过

局部和全局特征提取，捕捉缺陷区域的局部信息，更好地理解其与周围环境之间的关联，模型能够学习到更全面的图像特征表示，并且在缺陷检测任务中展现出良好的检测效果。

图 8 给出了单一场景下模型微调过程的准确率。由于不同模型的迭代次数不同，因此在归一化迭代次数后进行绘制，使得迭代次数从一个统一的起点开始，到一个统一的终点结束，曲线之间的比较更加直观。

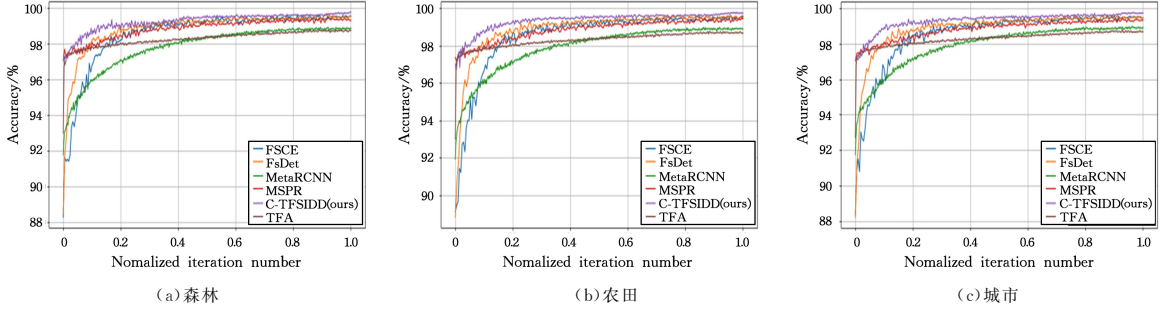


图 8 不同背景下微调模型的准确率

Fig. 8 Accuracy of fine-tuning models in different backgrounds

4.5 消融实验

为更具体地证明本文方法中每个模块的可行性和有效性，在包含丰富信息的复杂背景下的绝缘子缺陷数据集上进行验证， K 取 5 shot, 10 shot。对模型主要组成部分进行 4 组消融实验：1) 更换特征提取网络模块 (Next-ViT)；2) 使用改进

PAFPN 替换传统 FPN 模块；3) 使用基于度量的判别性分类损失函数 (DL) 以辅助参数更新；4) 使用 Focal Loss 损失函数替换交叉熵损失函数。此外，本文还给出不同模块下的模型参数量及计算量。

结果如表 8 所列。

表 8 绝缘子缺陷数据集上消融实验结果的对比

Table 8 Comparison of ablation experimental results on insulator defect dataset

Method	nAP50		mAP		Param	FLOPs	Image-size
	5 shot	10 shot	5 shot	10 shot			
Base(ResNet101)	9.9	16.8	43.0	55.7	60.4×10^6	291.6×10^6	1333×800
Base(NextViT)	36.1	45.7	51.4	60.5	47.8×10^6	247.7×10^6	1333×800
Base(NextViT)+改进 PAFP	37.4	46.8	42.1	60.9	51.4×10^6	268.1×10^6	1333×800
Base(NextViT)+改进 PAFP+DL	37.5	50.8	49.3	61.5	—	—	—
Base(NextViT)+改进 PAFP+DL+Focal Loss	38.6	52.3	45.6	62.4	—	—	—

第一组实验为基线模型采用 ResNet101 作为特征提取网络的算法，为后续实验提供对比。

第二组实验将基线模型特征提取网络 ResNet101 更换成 Next-ViT 网络。该网络结合 CNN 的局部特征捕捉能力和 Transformer 的全局上下文建模，有效整合局部与全局特征，让模型更好地理解图像内容。利用多头卷积注意力和多头自注意力模块并行地在不同的表示子空间中工作，使得模型能够同时捕捉到缺陷区域的多种特征。可视化出测试过程中特征网络模块最后一层特征图，如图 9 所示，红色表示关注区域。可以发现，在传统 ResNet101 特征提取网络中，模型更加关注绝缘子本身而忽略缺陷部分，而融合局部和全局特征的 Next-ViT 网络对缺陷区域关注度较高。这反映出模型对目标区域有着更多的关注度，充分发挥了其在图像中提取有效信息方面的优势。实验结果支持了本文的可视化发现，相较于基线模型，使用 Next-ViT 后，新类检测结果在 5shot, 10shot 下分别提升 26.2%，28.9%，明显优于基线模型。多头分组的设计，使得模型的参数和计算量也大幅降低，表明 Next-ViT 的引入能有效捕捉目标检测任务中目标的局部

细节信息和全局上下文信息，提高绝缘子缺陷检测的精度。

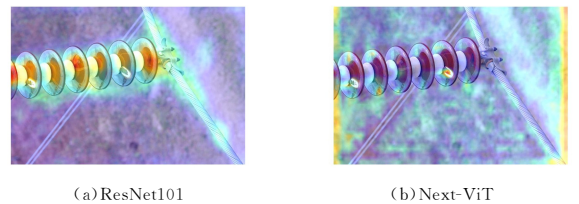


图 9 特征图可视化(电子版为彩图)

Fig. 9 Visualization of feature map

第三组实验为在第二组实验的基础上将 FPN 替换成改进 PAFP 模块。由表 8 中实验结果可知，绝缘子缺陷目标检测结果在 5 shot, 10 shot 下分别提升 1.3%，1.1%。这一性能的提升得益于：1) 改进 PAFP 模块深层特征图融合了浅层特征细节信息，从而提供了更准确的目标位置信息；2) 浅层特征图融合了深层特征丰富的语义信息，在提供细节信息的基础上，有助于绝缘子缺陷分类的完成。

第四组实验为使用基于度量的判别性分类损失函数 (DL) 优化分类器。实验结果表明，利用 DL 辅助参数更新，

新类精度在 10shot 下提升 4.0%。这一精度大幅提升的主要原因在于,当样本量增多时,该损失函数能不断学习不同类目标之间的关系,以及同类目标的共同特征,帮助模型做出更好的分类决策,在检测精度提升上发挥的作用更明显。

第五组实验为在基于度量学习的判别性损失函数(DL)的基础上,引入 Focal Loss 来替代传统的交叉熵损失函数进行分类损失的优化。实验结果表明,通过这种方式,5 shot 和 10shot 下的新类精度分别提升 1.1% 和 1.5%。这一提升凸显了 Focal Loss 在处理类别不平衡问题时的优势,特别是在样本数量受限的情况下,它通过减少对易分类样本的关注,增加对难分类样本的关注,促进了精确的分类预测。

此外,在 10shot 学习设置下重新训练模型,通过消融实验讨论不同分类和回归损失函数与 DL 组合对目标检测性能的影响。特别地,本文关注回归损失函数 L1 Loss 的一种改进形式(Balanced L1 Loss)^[34],它引入额外参数,通过动态调整梯度上来优化模型的学习过程,实验结果如表 9 所列。可以发现,10shot 时,单独使用 Focal Loss 代替交叉熵损失对分类损失进行优化可以提升新类精度 1.5%;而单独使用 Balanced L1 Loss 代替 L1 Loss 回归损失时精度略有下降(0.4%);当 Balanced L1 Loss 与 Focal Loss 损失联合使用时,新类精度略下降(0.6%)。经过分析认为,在训练数据有限的情况下利用该回归损失调整梯度的大小,可能会对模型的学习能力造成负担,这可能是回归损失精度没有显著提升的原因之一。因此,在综合考虑实验结果后,选择基于度量学习的损失函数与 Focal Loss 分类损失和 L1 Loss 回归损失的组合。该组合策略旨在帮助模型更好地学习和预测,在提高检测精度的同时保持模型的泛化能力。

表 9 损失函数组合消融实验

Table 9 Ablation experiments with combined loss functions

Classification Loss			Regression Loss		nAP50
Cross Entropy Loss	Focal Loss	DL	L1Loss	Balanced L1Loss	10 shot
✓			✓		46.8
✓		✓	✓		50.8
✓		✓		✓	50.4
	✓	✓	✓		52.3
	✓	✓		✓	51.7

本研究的综合性消融实验结果显示,使用所提模块的目标检测性能最佳,证明本文方法在少量样本的情况下能有效提升绝缘子缺陷检测性能。

结束语 针对绝缘子缺陷图像样本量不足且背景复杂、目标尺寸小导致绝缘子缺陷检测精度不高的问题,提出一种融合局部和全局特征的小样本目标检测模型。首先,采用融合局部和全局特征信息的 Next-ViT 模型成功提取出更为丰富、全面的特征,为后续的缺陷检测提供坚实的基础;其次,使用改进 PAFPN 模块进行多尺度特征的深度融合,增强小目标细节信息的表示能力;最后,优化损失函数,使得分类器在有限样本下依然能够展现出卓越的判别性能。此外,由于绝缘子检测场景与自然场景下类别差距大,因此采用了一种更适合本文模型的训练策略。由实验可知,所提小样本绝缘子

缺陷检测模型优于其他模型,表现出很好的检测效果。但其也存在不足,对比实验显示,在提升新类绝缘子缺陷检测精度的同时,基类的平均检测精度略有下降。未来研究会聚焦于在不遗忘先验知识的前提下解决小样本绝缘子缺陷检测问题,旨在提升新类检测效果的同时保持基类的检测性能。

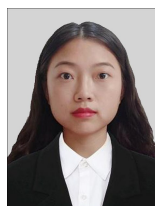
参 考 文 献

- [1] LIU C Y, WU Y Q. Research Progress of Vision Detection Methods Based on Deep Learning for Transmission Lines[J]. Proceedings of the CSEE, 2023, 43(19): 7423-7446.
- [2] ZHAO Z B, JIANG Z G, LI Y X, et al. Overview of visual defect detection of transmission line components[J]. Journal of Image and Graphics, 2021, 26(11): 2545-2560.
- [3] LIU Y, HUANG X. Efficient Cross-Modality Insulator Augmentation for Multi-Domain Insulator Defect Detection in UAV Images[J]. Sensors, 2024, 24(2): 428.
- [4] MA B, FU Y K, WANG C P, et al. High Performance Insulators Location Scheme Based on YOLOv4 with GDloU Loss Function [J]. Computer Science, 2022, 49(S1): 412-417.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 779-788.
- [6] ZHAO W Q, CHENG X F, ZHAO Z B, et al. Insulator recognition based on attention mechanism and Faster RCNN[J]. CAAI Transactions on Intelligent Systems, 2020, 15(1): 92-98.
- [7] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]// Proceedings of the 29th International Conference on Neural Information Processing Systems. Cambridge, MA: MIT, 2015: 91-99.
- [8] LIU J H, ZHAO Z, FU J R, et al. Active small sample learning based the pipe weld defect detection method[J]. Chinese Journal of Scientific Instrument, 2022, 43(11): 252-261.
- [9] ZHAO Z F, HUANG J H, LUO H J, et al. Simulation Research on Ultrasonic Total Focus Method Detection of Internal Defects of Composite Insulators[J]. Piezoelectrics and Acousto-optics, 2024, 46(1): 136-142.
- [10] WANG X, HUANG T E, DARRELL T, et al. Frustratingly simple few-shot object detection[J]. arXiv: 2003. 06957, 2020.
- [11] LI J, XIA X, LI W, et al. Next-vit: Next generation vision transformer for efficient deployment in realistic industrial scenarios [J]. arXiv: 2207. 05501, 2022.
- [12] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 8759-8768.
- [13] ZHANG T, ZHANG X, ZHU P, et al. Generalized few-shot object detection in remote sensing images[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2023, 195: 353-364.
- [14] KÖHLER M, EISENBACH M, GROSS H M. Few-shot object detection: A comprehensive survey [J]. arXiv: 2112. 11699, 2021.
- [15] SHI Y D, WANG H M, JING C, et al. A Few-Shot Defect De-

- tection Method for Transmission Lines Based on Meta-Attention and Feature Reconstruction[J]. *Applied Sciences*, 2023, 13(10): 5896.
- [16] ZHAI Y J, YANG K, WANG Q M, et al. Disc Insulator Defect Detection Based on Mixed Sample Transfer Learning[J]. *Proceedings of the CSEE*, 2023, 43(7): 2867-2877.
- [17] CUI K B, PAN F. A CycleGAN small sample library amplification method for faulty insulator detection[J]. *Computer Engineering & Science*, 2022, 44(3): 509-515.
- [18] KARLINSKY L, SHTOK J, HARARY S, et al. Repmet: Representative-based metric learning for classification and few-shot object detection[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019: 5197-5206.
- [19] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The pascal visual object classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88: 303-338.
- [20] LIU K P, LI B Q, QIN L, et al. Review of Application Research of Deep Learning Object Detection Algorithms in Insulator Defect Detection of Overhead Transmission Lines[J]. *High Voltage Engineering*, 2023, 49(9): 3584-3595.
- [21] ZHOU L J, MAO J N. Vision Transformer-based recognition tasks: a critical review[J]. *Journal of Image and Graphics*, 2023, 28(10): 2969-3003.
- [22] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook, NY: Curran Associates Inc., 2017: 6000-6010.
- [23] XI Y, ZHOU K, MENG L W, et al. Transmission Line Insulator Defect Detection Based on Swin Transformer and Context[J]. *Machine Intelligence Research*, 2023, 20: 729-740.
- [24] GUO J, LI T, DU B. Segmentation Head Networks with Harnessing Self-Attention and Transformer for Insulator Surface Defect Detection[J]. *Applied Sciences*, 2023, 13(16): 9109.
- [25] DU Z W, ZHOU H, LI C Y, et al. Small Object Detection Based on Deep Convolutional Neural Networks: A Review[J]. *Computer Science*, 2022, 49(12): 205-218.
- [26] ZHAO Y, YANG L. Distance metric learning based on the class center and nearest neighbor relationship[J]. *Neural Networks*, 2023, 164: 631-644.
- [27] 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*. 2017: 2980-2988.
- [28] RAIMUNDO A. Insulator data set-Chinese power line insulator dataset(CPLID)[DB/OL]. <https://github.com/InsulatorData/InsulatorDataSet>.
- [29] DENG J H, GUO W Q, CHEN H J, et al. Few-shot diatom detection combining multi-scale multi-head self-attention and on-line hard example mining[J]. *Journal of Computer Applications*, 2022, 42(8): 2593-2600.
- [30] SUN B, LI B, CAI S, et al. FSCE: Few-shot object detection via contrastive proposal encoding[C]//*Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 2021: 7352-7362.
- [31] XIAO Y, LEPETIT V, MARLET R. Few-shot object detection and viewpoint estimation for objects in the wild[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 45(3): 3090-3106.
- [32] YAN X, CHEN Z, XU A, et al. Meta R-CNN: Towards general solver for instance-level low-shot learning[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019: 9577-9586.
- [33] WU J, LIU S, HUANG D, et al. Multi-scale positive sample refinement for few-shot object detection[M]//*Lecture Notes in Computer Science*. Cham: Springer, 2020: 456-472.
- [34] PANG J, CHEN K, SHI J, et al. Libra R-CNN: Towards balanced learning for object detection[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019: 821-830.



CUI Kebin, born in 1979, Ph.D, professor. His main research interests include image processing, pattern recognition and machine learning.



HU Zhenzhen, born in 2000, postgraduate. Her main research interests include object detection and few-shot learning.

(责任编辑:柯颖)