

基于深度学习的无人机航拍输电线路侵入异物检测综述

刘传洋^{1,2}, 吴一全¹, 刘景景²

(1. 南京航空航天大学电子信息工程学院, 江苏 南京 211106;

2. 池州学院机电工程学院, 安徽 池州 247000)

摘要: 基于无人机航拍图像和深度学习算法实现输电线路侵入异物的精准有效检测, 有助于快速排查输电线路潜在安全隐患, 进而提升电网精益化运维能力。综述了基于深度学习的无人机航拍输电线路侵入异物检测方法的研究进展。首先, 从Faster R-CNN、SSD、YOLO系列算法及其他衍生算法4个维度, 系统梳理了基于深度学习的无人机航拍图像异物检测方法。其次, 阐释了无人机航拍图像下输电线路侵入异物检测的改进模型, 分别阐明了注意力机制、多尺度特征融合、轻量化网络等技术在输电线路侵入异物检测中的应用路径及其带来的性能优化效果。接着, 剖析了基于传统Transformer架构的DETR类异物检测模型, 以及面向电力行业的Transformer大模型及其工程应用。最后, 针对现有无人机航拍输电线路侵入异物智能检测面临的挑战, 对下一步研究工作进行了展望。

关键词: 输电线路; 电力巡检; 无人机航拍图像; 深度学习; 异物检测

中图分类号: TP391.41; TM930.1

文献标志码: A

doi: 10.11959/j.issn.2096-6652.ZNKX26042

Review of deep learning-based foreign object intrusion detection for transmission lines via UAV aerial photography

Liu Chuanyang^{1,2}, Wu Yiquan¹, Liu Jingjing²

1. College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

2. College of Mechanical and Electrical Engineering, Chizhou University, Chizhou 247000, China

Abstract: Efficient and accurate detection of foreign object intrusion on transmission lines based on UAV aerial images and deep learning algorithms enables rapid identification of potential safety hazards, thereby improving the lean operation and maintenance capability of power grids. The research progress of deep learning-based foreign object intrusion detection methods for transmission lines using UAV aerial images was reviewed. First, deep learning-based foreign object detection methods for UAV aerial images were systematically reviewed from four dimensions: Faster R-CNN, SSD, the YOLO series algorithms, and other derivative algorithms. Second, an improved model for UAV aerial image-based foreign object intrusion detection in transmission lines was elaborated. It illustrated the application strategies and performance optimization effects of core technologies in detection tasks, including attention mechanisms, multi-scale feature fusion, and lightweight networks. Subsequently, the DETR-based foreign object detection models built on the vanilla Transformer architecture were investigated, as well as large Transformer models tailored for the power industry and their engineering applications. Finally, in view of the current challenges in UAV aerial image-based intelligent foreign object detec-

收稿日期: 2026-04-07; 修回日期: 2026-06-05

通信作者: 吴一全, nuaaimage@163.com

基金项目: 国家自然科学基金资助项目 (No.61573183); 安徽省高校自然科学研究重点项目 (No.2025AHGXZK30826, No.KYJG002)

Foundation Items: The National Natural Science Foundation of China (No.61573183), The Key Project of Anhui Provincial Universities Natural Science Research (No.2025AHGXZK30826, No.KYJG002)

tion for transmission lines, future research prospects were proposed.

Key words: transmission line, electrical inspection, UAV aerial image, deep learning, foreign object detection

0 引言

输电线路侵入异物包括电力杆塔上的鸟巢、电力线上的悬挂物（风筝、气球、垃圾袋、覆冰等）、输电走廊内植被和工程车辆等。异物侵入输电线路势必影响其安全稳定运行，可能会引起线路跳闸，严重时甚至造成大范围停电或安全事故。因此，如何快速发现输电线路侵入异物并及时排除隐患，已经成为该领域运维工作亟待解决的问题^[1-3]。

随着视觉检测技术的快速发展，无人机巡检已成为当前输电线路巡检的主要方式。面对海量的无人机巡检图像，众多研究者利用传统目标检测和深度学习方法来检测侵入异物^[4]。传统目标检测方法大多只能进行单一类型异物检测，其检测结果易受噪声干扰影响，导致特征提取过程中存在目标信息丢失现象。深度学习方法则利用卷积神经网络（convolutional neural network, CNN）提取图像深度特征，实现侵入异物端到端的检测。与传统目标检测方法相比，深度学习方法具有较强的数据挖掘能力和较好的侵入异物检测性能。然而，复杂背景下的无人机航拍图像，目标遮挡、光照强度、多尺度及小目标等因素给准确识别输电线路中的侵入异物带来了巨大挑战^[5]。

在电力大数据与智能巡检快速发展的背景下，以无人机航拍图像为核心数据源的电力深度视觉技术，已成为输电线路状态监测、部件识别与缺陷检测的重要支撑^[6]。现有相关文献从不同角度对该领域进行了梳理，为后续研究奠定了基础。现有研究普遍关注深度学习方法在输电线路视觉检测中的应用瓶颈，文献[7]指出样本不平衡、小目标检测困难和细粒度特征难提取是制约检测精度提升的关键问题。在技术路线方面，文献[8]从智能图像处理框架、轻量化模型设计、多模态数据融合等方向展开分析。文献[9]围绕绝缘子等典型部件的缺陷检测，强调轻量化网络在小样本场景下的优势。文献[10]聚焦实时检测任务，证实多源图像融合可有效提升复杂环境下的诊断可靠性。文献[11]总结数据质量保障、小目标检测优化、嵌入式部署与统一评测基准构建是未来发展趋势。文献[12]融合深度学习方法与无人机自主巡检系统，推动视觉检

测方法从离线分析向在线感知和实时预警方向发展。

上述文献主要聚焦输电线路巡检视觉检测方法与深度学习应用的梳理及前景展望，但尚未对输电线路侵入异物检测这一细分场景开展系统性综述。迄今为止，国内外学者基于视觉检测方法开展无人机航拍输电线路侵入异物检测研究已取得一定进展，但仍缺乏对各类检测方法、改进策略与场景适应性的横向对比与归纳。鉴于此，本文系统综述了基于深度学习的无人机航拍输电线路侵入异物检测方法，以期为未来输电线路状态自动感知、智能分析及运维决策优化提供有价值的参考。

1 输电线路侵入异物类型及检测流程与挑战

1.1 输电线路侵入异物

输电线路的侵入异物是指原本不属于输电线路正常运行环境中的物体，因风力、气流、冰雪、鸟类筑巢等自然因素或人为因素进入输电线路的安全距离范围内，从而对输电线路的安全稳定运行产生潜在或实际威胁，及时发现并清除输电线路侵入异物，对于保障电网安全稳定运行至关重要。因此，国家电网公司采用无人机结合视频监控、激光雷达、红外热成像等监测技术进行数据采集，利用人工处理结合机械设备、激光技术等方法清除输电线路走廊中的侵入异物。常见的输电线路侵入异物有鸟巢、风筝、气球、线路覆冰、工程车辆等，如图1所示。

1.2 侵入异物检测流程

随着电力需求日益增加，电网覆盖范围不断扩大，输电线路走廊面临着越来越高的异物侵入风险，为了及时发现并处理输电线路侵入异物，智能检测技术应运而生。无人机航拍图像中输电线路侵入异物智能检测流程主要包括图像采集、图像预处理、特征提取、目标识别与分类等。

（1）图像采集是通过各种图像采集设备（如摄像头、扫描仪等）获取包含检测目标的图像。在输电线路侵入异物检测中，图像采集设备可以是安装在杆塔上的摄像机、无人机携带相机或激光雷达等，获取输电线路及其周边环境的图像。



图1 输电线路侵入异物

(2) 图像预处理是在对图像分析和识别之前,对图像进行灰度化、去噪、图像增强等一系列操作,其目的是通过增强图像特征或去除干扰信息等方法来改善图像质量。

(3) 特征提取是目标检测方法的关键步骤,旨在从图像中挖掘能够有效表征目标物体的信息。传统机器学习方法通过人工设计特征的方法提取异物特征(如颜色、纹理、形状等),通常会结合多种特征来提高异物检测的准确率和鲁棒性,但存在泛化能力差、计算量大、推理速度慢等缺陷。深度学习方法则利用大量标注的图像数据训练目标检测模型,通过CNN提取异物的深层特征,在应对复杂输电线路场景、开展多类异物检测时,具备较高的精度与效率。

(4) 目标识别与分类是计算机视觉(computer vision, CV)领域的重要任务,旨在从图像或视频数据中确定目标物体的类别。传统机器学习方法使用支持向量机、决策树、AdaBoost等算法对颜色、形状、纹理等特征进行分类训练,利用训练好的模型识别图像中的异物。深度学习方法通过CNN自动识别异物的类型和位置,实现目标异物端到端的检测,取得了比传统机器学习方法更好的检测

结果。

随着人工智能技术的飞速发展,基于无人机的输电线路侵入异物智能检测技术不断成熟,有效保障了电力系统的安全稳定运行。

1.3 侵入异物检测面临的挑战

随着电力行业的发展,电网规模不断扩大,输电线路的长度不断增加,传统的检测方法难以满足大规模输电线路的运维需求。智能检测技术具有高效、准确的特点,可以适应无人机巡检输电线路侵入异物检测任务,保障电网的安全可靠运行。但是输电线路侵入异物智能检测面临以下挑战。

(1) 复杂背景下异物识别。在复杂的自然环境中,受天气条件(如雾、雨、雪)、光照变化以及图像采集设备自身因素等影响,采集的图像可能存在模糊不清、目标遮挡等情况,进一步增加了后续目标检测算法的异物识别难度。

(2) 单模态图像检测。单模态图像仅包含一类图像数据,可利用的信息有限。输电线路侵入异物检测若仅采用可见光图像或红外图像,无法发挥不同图像数据的互补作用。

(3) 大规模图像数据存储与分析。智能巡检会产生海量图像和数据,对存储容量和数据处理能力

均提出了较高要求。

(4) 实时性检测。输电线路侵入异物智能检测算法需要具备高速运算能力,以快速处理各类采集数据并实现实时检测。但高精度算法计算开销大,如何在保证检测精度的同时兼顾检测实时性是一大挑战。

(5) 硬件设备限制。实现高精度的异物检测离不开高分辨率相机、精密传感器、高性能计算设备等硬件。当前电力巡检虽已应用先进传感与图像采集技术,但受恶劣天气和复杂地形影响,仍难以有效检测输电线路的侵入异物。

2 输电线路侵入异物检测的深度学习算法

基于深度学习的目标检测算法具有较强的数据挖掘能力,随着GPU计算能力的不断提升,深度学习算法在输电线路侵入异物检测方面具有广阔的应用前景。自Girshick等^[13]提出R-CNN(regions with CNN)算法以来,涌现出了很多检测性能良好的目标检测算法,如以Faster R-CNN^[14]、R-FCN^[15](region-based fully convolutional network)和Mask R-CNN^[16]为代表的双阶段目标检测算法,以及实时目标检测的SSD^[17](single shot multibox detector)、YOLO^[18-21]等单阶段目标检测算法。双阶段目标检测算法需要先生成区域建议网络并从这些区域建议网络中提取目标深层特征,然后从每个区域建议网络中获得精确的位置和类别信息。单阶段目标检测算法使用深度卷积神经网络(deep CNN, DCNN)直接从整幅图像的预设网格中获得目标的所属类别和位置信息。双阶段目标检测算法能够充分提取图像深层特征,实现准确的分类与定位,具有较高的检测精度和较低的漏检率,单阶段目标检测算法则具有更快的运行速度和较好的实时性。

针对无人机航拍图像中输电线路侵入异物检测方法,下面将从经典的Faster R-CNN、SSD、YOLO系列算法以及其他目标检测算法方面进行总结。

2.1 输电线路侵入异物检测的Faster R-CNN算法

双阶段目标检测算法为基于候选区域的检测方法,其中R-CNN、Fast R-CNN^[22]和空间金字塔池化网络(spatial pyramid pooling network, SPP-Net)^[23]采用选择性搜索算法生成候选区域,而Faster R-CNN算法提出利用区域建议网络(region proposal network, RPN)生成候选区域,与传统的选择性搜索算法和滑动窗口方法相比,其目标识别的准确率和训练速度都得到了提高,图2为基于RPN的Faster R-CNN结构。Faster R-CNN主要由4个部分组成:特征提取网络(如VGG-Net^[24]、GoogLeNet^[25]、ResNet^[26]等预训练分类模型)用于提取输入图像的特征并生成特征图;RPN用于生成感兴趣区域(region of interest, RoI),先使用Softmax确定锚框中是否包含待识别目标,再利用边界框回归确定锚框的位置以获得RoI;特征抽取部分利用RoI池化层获得用于最终分类和定位的建议特征图;分类+回归部分使用全连接层进行RoI类别计算和边界框回归以获得检测目标位置信息。

Faster R-CNN凭借较高的检测精度,在输电线路侵入异物检测中得到广泛应用。针对不同异物目标的检测难点,现有研究分别从损失函数优化、结构适配、数据增强、多模型融合等方向对Faster R-CNN进行改进,有效提升了其在复杂场景下的检测性能。针对鸟巢检测存在的适应性弱和精度偏低问题,文献[27]通过引入焦点损失函数与RoI挖掘模块强化困难样本特征学习,使鸟巢检测精度由78.29%提升至82.51%,但新增的焦点损失函数与RoI挖掘模块引入了额外计算分支,大幅增加了模

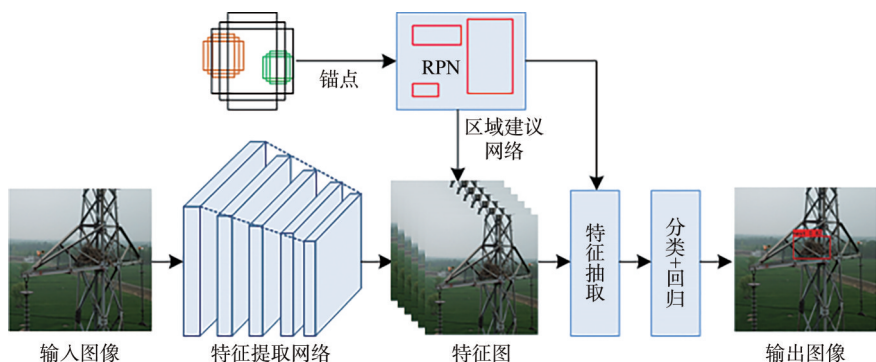


图2 基于RPN的Faster R-CNN结构

型参数量与推理计算开销,不利于无人机边缘端实时部署。面向工程车辆检测需求,文献[28]通过调整RoI池化层结构实现网络与目标特征的适配,在保证低虚警率的同时,将检测精度由原始Faster R-CNN的89.12%提升至89.93%,但该方法的检测性能提升幅度有限,在多尺度特征融合、小目标增强等深层网络架构方面还有优化空间。为同时识别多类别异物与部件缺陷,文献[29]采用镜像、旋转、噪声注入等数据增强策略扩充样本分布,大幅降低样本不均衡带来的影响,使鸟巢检测精度达到96.22%,有效降低了漏检率。在异物分类任务中,文献[30]通过融合Inception结构与Faster R-CNN(ResNet101)构建二分类模型,可在不同告警模式下实现高召回率输出,正常模式与警告模式召回率分别达到95%与100%,但该方法无法实现鸟巢、悬挂物、施工车辆等多异物细粒度分类,难以满足精细化巡检运维需求。

总体而言,基于Faster R-CNN的改进方法能够针对各类异物的检测难点,有效提升输电线路侵入异物检测精度,解决单一场景下的部分检测问题。但从无人机实时巡检的实际部署需求出发,现有改进方法普遍存在计算量大和推理速度慢的问题,难以适配无人机有限的硬件算力;同时对小目标与遮挡目标鲁棒性不足,无法满足复杂巡检场景的检测需求。

2.2 输电线路侵入异物检测的SSD算法

SSD作为一种典型的单阶段目标检测算法,以VGG-Net模型为特征提取网络,通过多尺度特征融合策略进行目标检测。其网络结构如图3所示,SSD先采用VGG-16网络生成特征图,再通过下采样卷积操作提取不同尺度特征图,最后利用边界框与类别回归预测并结合非极大值抑制(non-maximum suppression, NMS)获得最终的目标检测结果。

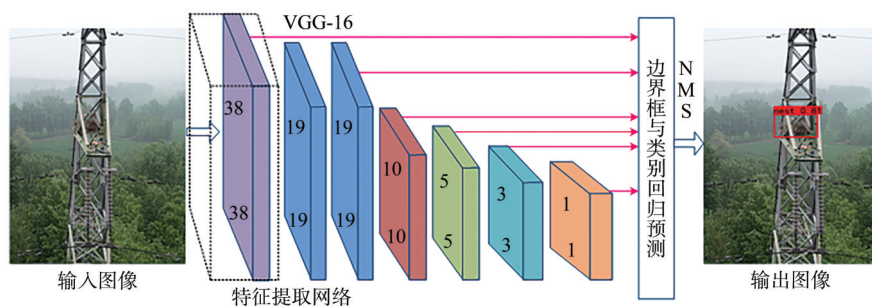


图3 SSD网络结构

SSD算法凭借结构简洁和推理快速的特点,在输电线路异物检测与部件识别任务中得到了广泛应用。为满足边缘端部署需求,现有研究多从轻量化主干网络着手,在保证检测精度的同时,大幅减少模型参数量与计算开销。针对边缘设备部署难题,文献[31]提出了一种基于轻量化SSD的输电线路侵入异物检测方法,以MobileNet作为特征提取网络,MobileNet-SSD模型仅有22 MB,与Faster R-CNN(546 MB)和VGG-SSD(94 MB)相比小得多。该方法在保证检测精度的前提下(异物识别精度达到了89%),大幅提升了网络模型的检测速度,其运行速度是Faster R-CNN的58倍,VGG-SSD的5倍,有效兼顾了轻量化与实时性需求,但该方法未对特征融合模块、锚框设计等结构进行协同优化,导致模型在复杂场景下的泛化能力不足。针对覆冰监测场景,文献[32]提出了MobileNetv3-SSD目标检测算法,在轻量化MobileNetv3特征提取网络的基础上引入多感受野模块以增强特征提取能力,MobileNetv3-SSD目标检测算法的覆冰厚度辨识准确率达到了82.1%,与YOLOv3(65.2%)和SSD(69.5%)算法相比,分别提高了近17%和12%,但受限于轻量化设计,其对小目标异物的检测能力仍不足。

SSD网络结构相对简单,计算量较小,有利于快速推理,但小目标检测效果不佳始终是制约其进一步应用的因素。为此,研究者通过优化特征提取网络、设计高效多尺度特征融合模块等技术路径,提出了DSSD(deconvolutional SSD)、FSSD(feature fusion SSD)、RSSD(rainbow SSD)等改进算法。这类算法本质上是通过强化中小目标的特征表征能力,弥补原始SSD算法在特征提取和融合中的不足。

2.3 输电线路侵入异物检测的YOLO算法

YOLO作为典型的单阶段目标检测算法,可以直接从图像中获得待检测目标的位置信息和所属类

别。YOLO 算法先将输入图像划分为 $S \times S$ 网格，如果检测目标的中心落入一个网格单元中，则该网格单元负责检测目标。输出结果包括标记边界框、目标类别和置信度分数。自 YOLOv1 开创了单阶段目标检测算法以来，短短几年间，陆续迭代出了 YOLO 系列算法（YOLOv1~YOLOv13），其发展历程如图4所示。

YOLO 系列算法兼顾检测精度与推理速度，不断优化模型实时推理性能。YOLOv3 在工程化与检测性能上实现了显著突破，成为工程化应用中首选的目标检测算法之一。YOLOv3 以 Darknet-53 作为主干特征提取网络，在每个卷积层后均加入批量归一化（batch normalization, BN）与 ReLU 激活函数，有效加快收敛速度并抑制过拟合。在颈部结构中，YOLOv3 引入特征金字塔网络（feature pyramid network, FPN），实现深层高级语义特征与浅层细节特征的高效融合，通过高分辨率浅层特征检测小尺度目标，低分辨率深层特征检测大尺度目标，显著提升多尺度目标检测能力。与前序版本相比，YOLOv3 在整体检测精度与小目标识别率上均实现大幅提升，对无人机航拍图像中的风筝、鸟巢、悬挂物等具备更优的适应性，因此被广泛作为输电线路侵入异物检测的基础算法，其网络结构如图5所示。

YOLOv5 在 YOLO 系列算法工程化应用中具有里程碑意义，其结构继承并优化了单阶段目标检测算法的核心思想，以 CSPDarknet-53 作为主干网络，并在首层引入 Focus 模块实现下采样与特征重组，在减少计算量的同时保留更多细节信息。颈部网络继续采用多尺度特征融合架构，并通过跨阶段局部（cross stage partial, CSP）模块替换部分卷积层，进一步提升特征复用效率。为适配边缘部署等不同算力场景需求，YOLOv5 提供了 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 4 种规模版本，在检测精度、运行速度与模型轻量化之间具备更优的平衡能力，模型部署灵活性显著提升，因此被广泛应用于输电线路侵入异物实时检测任务。YOLOv8 基于 YOLOv5 框架改进而来，采用 C2F 模块替换原 C3 模块，并融合 YOLOX、YOLOv6、YOLOv7 等算法的设计思想，形成了无锚框、兼容性强和易于扩展的最新技术水平（state of the art, SOTA）目标检测算法。

依托 YOLO 系列算法在推理速度与检测精度上的均衡特性，众多学者针对无人机航拍输电线路侵入异物检测的场景痛点开展了大量研究，主要围绕边界框优化、特征增强、轻量化部署、小目标与复杂背景适应等方向提升检测性能。针对异物定位

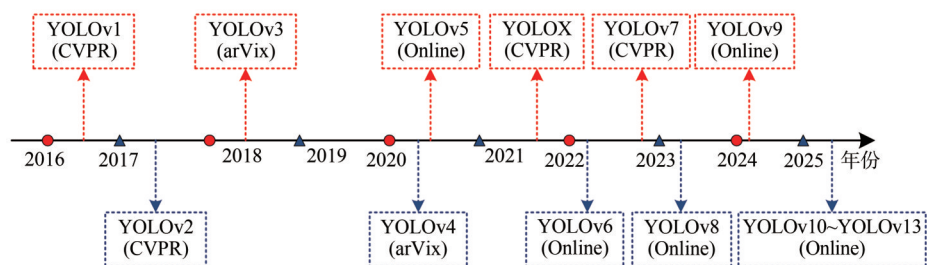


图4 YOLO 系列算法的发展历程

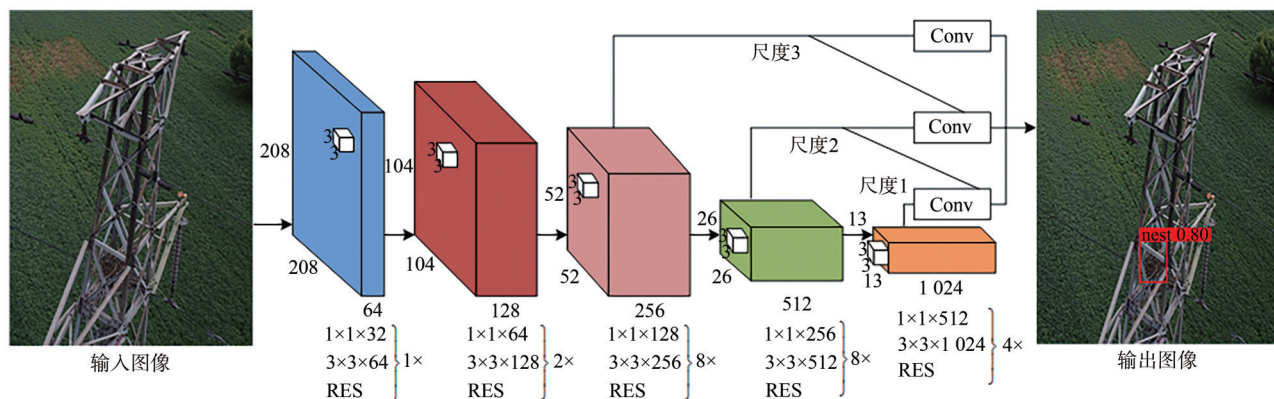


图5 YOLOv3 网络结构

不准、尺度差异大等问题,文献[33]提出了基于YOLOv3的输电线路侵入异物检测方法,通过定向边界框设计与尺度直方图匹配策略强化目标回归能力,使异物检测精度达到88.1%,较Faster R-CNN(87.4%)、SSD(77.2%)和YOLOv3(83.2%)均有明显提升,其推理速度可达45 FPS,但模型计算复杂度上升。为应对低分辨率、复杂背景、目标尺度多变等挑战,文献[34]通过空间深度卷积和多维协作注意力机制改进YOLOv7的主干网络,并结合Ghost-Conv降低计算开销,改进模型的异物检测精度进一步提升至95.98%,优于Faster R-CNN(95.47%)、SSD(85.87%)和YOLOv5(90.97%),同时异物检测速度为64 FPS,可满足实时检测需求,但模型对密集小目标和重叠异物的分辨能力仍有提升空间。在YOLOv8框架下,文献[35]提出的改进模型在检测精度(73.2%)与计算复杂度上优于Cascade R-CNN(67.7%)、YOLOv5(61.6%)、YOLOv7(67.6%)、YOLOv8(66.5%)等算法,但在复杂背景下仍存在小目标漏检和误检问题。面向无人机端侧高效巡检需求,文献[36]提出了轻量化YOLO-PowerLite模型,并成功部署于NVIDIA Jetson Xavier NX嵌入式平台,其检测精度(94.2%)优于Faster R-CNN(93.7%)、SSD(88.3%)、YOLOv5(93.7%)和YOLOv7-tiny(90.0%)。为进一步缓解遮挡、小目标及复杂背景带来的检测困难,文献[37]融合坐标注意力与Focal-EIoU损失函数提出了CFG-YOLOv8n模型,该模型将检测精度提升至97.6%,优于CenterNet(61.3%)、SSD(90.4%)、YOLOv5(91.3%)和YOLOv8(96.6%)。同时,将模型体积与参数量分别压缩至5.51 MB和 3.1×10^6 ,实现了高精度与轻量化的有效平衡。

综上所述,YOLO系列算法凭借端到端检测、推理速度快、检测精度较高等优势,在输电线路侵入异物检测中占据主导地位。但其检测性能仍高度依赖训练样本的数量与质量,在复杂背景、目标遮挡及极端尺度变化条件下的鲁棒性仍有待进一步提升。

2.4 输电线路侵入异物检测的其他算法

除了上述利用Faster R-CNN、SSD和YOLO算法检测输电线路侵入异物以外,还有很多基于目标检测算法的异物检测方法。文献[38]提出了多网络特征融合和随机森林分类相结合的输电线路侵入异物检测方法,通过GoogLeNet、DenseNet-201和

EfficientNet-B0级联融合进行特征提取,输电线路侵入异物检测精度达到95.88%,优于Faster R-CNN(93.09%)、SSD(85.20%)和YOLOv3(92.6%)。多网络级联融合虽提升了特征表征能力,但也导致模型参数量和计算复杂度大幅攀升,推理速度难以满足无人机实时巡检需求。文献[39]提出了一种输电线路威胁检测网络TATLNet,利用生成网络VRGNet和判别网络VTCNet预测侵入输电线路走廊的大型车辆,车辆检测准确率达到94.68%,优于MobileNet(88.35%)和ShuffleNet(83.65%)。文献[40]利用胶囊网络进行输电线路侵入异物检测,改进网络的异物检测平均准确率高达98.3%,优于AlexNet(76.3%)和GoogLeNet(85.6%)。改进网络检测精度表现突出,但是胶囊网络本身存在训练复杂度高和推理速度慢的缺陷,且未对比分析胶囊网络与主流算法的性能差异。文献[41]提出了基于RetinaNet的鸟巢检测方法,实验结果表明,RetinaNet的检测准确率达到94.1%,优于Faster R-CNN(90.2%)、SSD(89.3%)和YOLOv3(88.5%)。每幅图像的平均推理时间为68 ms,与Faster R-CNN、YOLO和SSD算法相比,RetinaNet具有良好的检测效果,能够快速准确地识别出电力塔架及鸟巢。文献[42]提出了基于分块式自纹理学习网络(patch-wise self-texture-learning network, PSTLNet)的输电线路侵入异物检测方法,结合自纹理学习模块和分块感知空间注意力模块实现输电线路上的侵入异物有效检测,鸟巢检测准确率为84.8%,优于Faster R-CNN(81.9%)、RetinaNet(80.9%)、YOLOv3(73.8%)和YOLOv7(81.4%)。文献[43]提出了基于深度弱监督和迁移学习相结合的绝缘子覆冰识别方法,该方法的绝缘子覆冰检测精度达到了90.1%,与Faster R-CNN(80.4%)、SSD(84.2%)、YOLOv5(87.6%)和RetinaNet(86.3%)相比,分别提高了9.7%、5.9%、2.5%和3.8%。文献[44]提出了基于CenterNet模型的输电线路侵入异物检测方法,CenterNet模型部署于边缘设备,异物检测精度达到了96%,优于Faster R-CNN(92.1%)、SSD(91.9%)、FCOS(95.7%)、YOLOX(95.8%)等主流检测算法,为无人机智能巡检输电线路提供了技术参考。

此外,还有研究者提出了基于语义分割算法的输电线路侵入异物检测方法,如文献[45]提出了基于U-Net算法的电力线覆冰分割方法。文献[46]提出了基于改进U-Net和遥感影像的输电线路走廊隐患检

测方法。文献[47]提出了基于实例分割算法 Mask-RCNN 的输电线路防外破检测方法。文献[48]提出了基于精细化 Mask-RCNN 的输电线路涉鸟故障检测方法。文献[49]提出了单目标深度估计融合目标检测 YOLOX 和语义分割 DeepLabv3+ 的输电线路侵入异物检测方法,具体可进一步参考相关文献。

上述典型算法在输电线路侵入异物检测中的性能对比见表1。

3 输电线路侵入异物检测的深度学习改进模型及优化策略

输电线路安全隐患排查对保障输电质量与电网安全稳定运行至关重要。为推动深度学习模型更好

地应用于输电线路侵入异物检测任务,基于无人机巡检实现复杂背景下侵入异物的快速精准识别,已成为输电线路智能巡检系统亟待解决的关键技术难题。为此,诸多研究者围绕注意力机制、多尺度特征融合、轻量化网络等方向,对深度学习模型进行改进与优化,以提升输电线路侵入异物检测性能。

3.1 基于注意力机制嵌入的异物检测模型

注意力机制本质是通过网络自主学习获得一组权重系数,以动态加权的方式在关注感兴趣区域的同时抑制不相关背景区域。一些研究者将注意力机制引入特征提取网络或特征融合网络,用于改善输电线路侵入异物检测的效果。卷积块注意力模块(convolutional block attention module,

表1 典型算法在输电线路侵入异物检测中的性能

文献	算法	检测目标	检测精度/准确率	性能对比
文献[27]	Faster R-CNN + ROI挖掘	鸟巢	82.51%	与 Faster R-CNN(78.29%)相比提高了 4.22%
文献[28]	Faster R-CNN + ROI池化层	工程车辆	89.93%	相比 Faster R-CNN(89.12%)提高了 0.81%
文献[29]	Faster R-CNN + 数据增强	鸟巢	96.22%	有效降低漏检率
文献[31]	MobileNet-SSD	鸟巢、塑料袋	89%	运行速度是 Faster R-CNN 的 58 倍、VGG-SSD 的 5 倍
文献[32]	MobileNetv3-SSD	电力线覆冰	82.1%	与 YOLOv3(65.2%)和 SSD(69.5%)算法相比,准确率分别提高了 17% 和 12%
文献[33]	YOLOv3 + 定向边界框	车辆、悬挂物	88.1%	较 Faster R-CNN(87.4%)、SSD(77.2%)和原始 YOLOv3(83.2%)分别提升 0.7%、10.9% 和 4.9%
文献[34]	YOLOv7 + SPD Conv	鸟巢、悬挂物	95.98%	检测精度优于 Faster R-CNN(95.47%)、SSD(85.87%)和 YOLOv5(90.97%)
文献[35]	TFD-YOLOv8	鸟巢、悬挂物	73.2%	检测精度优于 Cascade R-CNN(67.7%)、YOLOv5(61.6%)、YOLOv7(67.6%)和 YOLOv8(66.5%)
文献[36]	YOLO-PowerLite	鸟巢	94.2%	检测精度优于 Faster R-CNN(93.7%)、SSD(88.3%)、YOLOv5(93.7%)和 YOLOv7-tiny(90%)
文献[37]	CFG-YOLOv8n	悬挂物	97.6%	检测精度优于 CenterNet(61.3%)、SSD(90.4%)、YOLOv5(91.3%)和 YOLOv8(96.6%)
文献[38]	多特征提取网络融合+随机森林分类器	鸟巢、悬挂物	95.88%	优于 Faster R-CNN(93.09%)、SSD(85.2%)和改进 YOLOv3(92.6%)
文献[39]	轻量化 TATLNet	大型机械	94.68%	优于 MobileNet(88.35%)和 ShuffleNet(83.65%)
文献[40]	胶囊网络	悬挂物	98.3%	优于 AlexNet(76.3%)和 GoogLeNet(85.6%)
文献[41]	RetinaNet	鸟巢	94.1%	优于 Faster R-CNN(90.2%)、SSD(89.3%)和 YOLOv3(88.5%)
文献[42]	PSTL-Net	鸟巢	84.8%	优于 Faster R-CNN(81.9%)、RetinaNet(80.9%)、YOLOv3(73.8%)和 YOLOv7(81.4%)
文献[43]	弱监督+迁移学习	电力线覆冰	90.1%	与 Faster R-CNN(80.4%)、SSD(84.2%)、YOLOv5(87.6%)和 RetinaNet(86.3%)相比,分别提高了 9.7%、5.9%、2.5% 和 3.8%
文献[44]	CenterNet	鸟巢、电力部件异常目标	96%	优于 Faster R-CNN(92.1%)、SSD(91.9%)、FCOS(95.7%)、YOLOX(95.8%)等主流检测算法
文献[45]	U-Net	电力线覆冰	95.73%	分割精度优于 DeepLabV3+(93.71%)、Vit(95.87%)和 Ocrnet(94.23%)
文献[46]	改进 U-Net	施工车辆	85%	与 U-Net(79%)相比提高了 6%
文献[47]	Mask R-CNN	施工车辆	91%	优于 Faster R-CNN(89%)和 YOLOv3(81%)
文献[48]	Fine Mask R-CNN	涉鸟故障	90.9%	优于 Faster R-CNN(83%)、RetinaNet(78.2%)和 Mask R-CNN(87.8%)
文献[49]	YOLOX + DeepLabv3+	鸟巢、悬挂物	92.9%	优于 YOLOv5(91.53%)、YOLOv8(89.07%)、YOLOX(91.85%)和 RT-DETR(91.46%)

CBAM)是注意力机制的代表,被广泛嵌入各类检测主干网络以提升异物检测效果,其网络结构如图6所示。

针对Faster R-CNN算法在输电线路小目标异物检测中特征提取不足和检测精度偏低的问题,文献[50]将CBAM嵌入特征提取网络ResNet50的各卷积层中,通过整合注意力机制,使特征提取网络关注小目标的上下文信息,改进模型(ResNet50-CBAM)的鸟巢和悬挂物检测精度分别达到了87.2%和72.2%,与原始Faster R-CNN(ResNet50)相比,分别提高了4.6%和4.9%。在YOLOv5框架中,文献[51]在主干网络的SPP模块后添加CBAM以提升特征提取能力,改进模型的输电线路防外破隐患检测精度从90%提升至92.2%。针对YOLOX模型,文献[52]在多个有效特征层(80×80、40×40、20×20)引入轻量化CBAM,可在计算开销增加有限的前提下,将异物检测精度由85.67%提升至86.04%。为了解决YOLOv3模型在进行浅层与深层特征融合时引入不相关特征信息的问题,文献[53]提出了

基于注意力特征融合网络的目标检测网络模型。在YOLOv3特征融合网络的上采样过程中添加了一个注意力特征融合网络,如图7所示,通过融合低分辨率深层特征图与浅层特征图,以获得更有价值的高分辨率浅层特征。与YOLOv3网络模型相比,改进模型的鸟巢检测准确率和精度分别提高了1.15%和4.65%。CBAM有助于强化模型特征表征能力,但其引入会增加模型计算开销、降低推理速度,难以兼顾检测精度与实时性需求。

除典型CBAM外,研究者还通过更轻量化和更高效的注意力结构进一步提升检测精度,并对不同检测框架的特性实现灵活嵌入。文献[54]在YOLOv5的FPN结构中引入SimAM自注意力模块,可在无额外参数开销的情况下强化目标区域特征权重。文献[55]在YOLOv7的双向特征金字塔网络中嵌入了卷积混合注意力机制模块(mixed model of self-attention and convolution, ACmix),其能够融合自注意力与卷积的优势,兼顾全局建模与局部细节提取。文献[44]以CenterNet框架为载体,同步

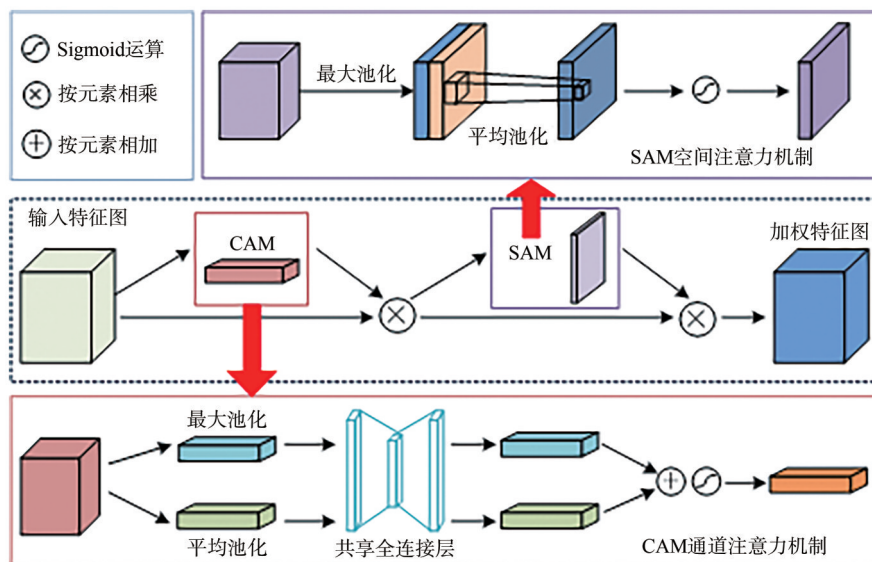


图6 CBAM网络结构

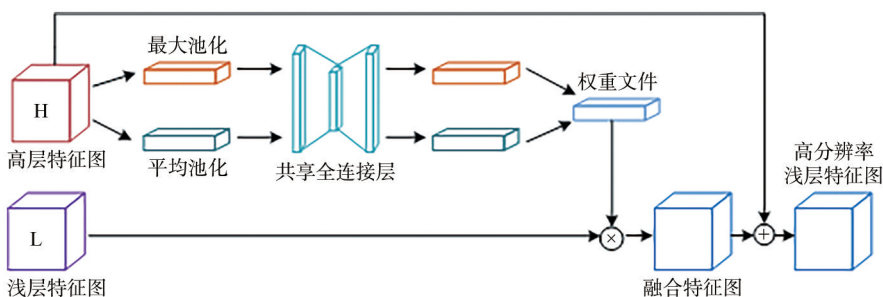


图7 注意力特征融合网络

引入高效通道注意力（efficient channel attention, ECA）机制和同尺度残差注意力特征融合（same-scale residual attention feature fusion, SRAFF）模块，其中ECA机制以一维通道交互方式实现重要特征通道的自适应增强，参数量与计算量远低于传统注意力模块，SRAFF模块则在同尺度特征维度引入残差关联与注意力加权，有效缓解特征融合过程中的信息衰减问题。文献[56-58]在YOLO网络模型中引入窗口自注意力（Swin Transformer）加强特征提取网络，凭借窗口划分与层级自注意力建模的优势，有效捕获输电线路场景中远距离目标关联与大范围背景语义信息，弥补了传统卷积网络长距离依赖建模能力的不足。

注意力机制能够让模型自动聚焦于输入数据中比较重要的部分，在处理复杂数据时，能够有效地筛选出对当前任务最具价值的信息，从而提高模型的目标检测性能。但是注意力机制需要计算输入数据中各个元素之间的相关性，以确定注意力权重，使处理大规模数据或复杂模型时消耗的计算成本较高。结合输电线路无人机巡检的实际场景，注意力机制可针对性捕捉小目标和弱特征，过滤无关信息，解决小目标异物检测精度不足的问题。但该机制在运行过程中需计算大量相关性参数，计算开销偏高。尤其在处理大规模图像数据、复杂背景及多类别异物检测任务时，计算成本会大幅上升，难以满足无人机巡检对模型轻量化与实时推理的要求。

3.2 基于多尺度特征融合的异物检测模型

多尺度特征融合是计算机视觉与图像处理领域的常用技术，旨在充分融合不同层级的特征信息以提升模型性能。深层特征图具备丰富的语义信息，而浅层特征图保留更多细节信息，通过特征金字塔网络结构可实现二者的有效融合。无人机巡检图像受拍摄角度、距离与高度影响，异物目标尺度差异显著，小目标占比高，直接导致检测精度下降。为此，研究者基于多尺度特征融合思想，分别从特征提取增强、FPN结构优化、双尺度级联检测等路径对模型进行改进，有效提升了复杂场景下的适应性。

在Faster R-CNN框架中，文献[59]采用多深度特征图与滑动窗口协同提取区域建议，可强化对不同尺度鸟巢的特征捕获能力，使检测精度从77.83%提升至88.68%，尺度适配性得到明显改善，但是仅通过多深度特征图与滑动窗口实现尺度适

配，未对特征金字塔融合逻辑进行结构性优化。针对输电线路异常目标检测需求，文献[60]在Mask R-CNN的基础上，利用FPN提取多尺度特征，并引入全局网络、横向连接增强多尺度特征表达，显著提升小目标检测能力，使烟火、工程车辆等目标检测精度提升1.4%。为实现检测精度与速度的均衡，文献[61]提出了基于双尺度YOLOv3网络模型的鸟巢检测方法，其检测框架如图8所示，采用两个YOLOv3网络模型对双尺度图像进行级联检测，相比单尺度模型，检测准确率提升5.7%，检测速度提升2.1 FPS，在检测实时性与检测性能间取得更优平衡。双网络级联架构虽提升了检测精度与速度，但模型整体参数量与结构冗余度有所增加。

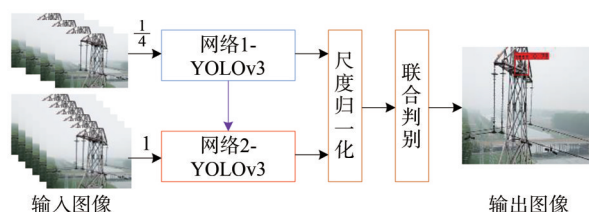


图8 双尺度检测框架

针对无人机航拍图像中输电线路小目标异物识别准确率偏低的核心痛点，研究者围绕特征复用、网络结构优化、感受野扩充、超参数调优等方向开展研究，有效提升了小目标异物的检测能力。在YOLOv3框架下，研究者通过引入Dense-Net实现特征的高效复用与融合，形成两种典型改进路径。文献[62]将Dense-Net嵌入特征提取网络倒数第二层，强化全流程特征传递，使风筝、鸟巢、施工机械等异物识别准确率从87.5%提升至96%。文献[63]将Dense-Net集成于三尺度检测头，聚焦目标识别关键环节，使异物检测精度提升至95.54%，较原始YOLOv3提高1.72%。密集连接结构会小幅增加模型参数量与计算量，对无人机边缘端实时推理存在一定影响。针对RetinaNet模型对输电线路小目标异物特征挖掘不充分的问题，文献[64]通过在FPN结构中新增P2特征层，有效扩充网络感受野，使鸟巢检测精度提升至90.2%，较基准ResNet50提高0.6%。新增P2特征层虽能保留更多浅层细节，但会增加特征融合分支与运算量。在YOLOv7架构中，文献[65]通过将空间到深度卷积（spatial-to-depth convolution, SPD-Conv）融入ELAN模块，并结合遗传算法优化超参数加速收敛，进一步提升

低分辨率图像与小目标异物的检测性能，其异物检测精度达到92.2%，相较于YOLOv7、YOLOv5、Faster R-CNN、CenterNet等主流模型均有显著优势，提升幅度为1.1%~8.28%。SPD-Conv与ELAN模块相结合虽能提升小目标特征提取效果，但模块堆叠会带来一定计算冗余。

多尺度特征融合可聚合网络深层语义特征与浅层细节特征，全面捕捉输电线路异物多维度信息，有效提升无人机复杂航拍场景下的检测精度与鲁棒性。该方法存在明显局限：多尺度结构会增大模型体量与运算开销，不符合轻量化部署要求；算法对图像质量敏感，易受环境因素干扰，加之小目标异物样本少、标注困难，模型易出现过拟合，泛化能力随之降低。

3.3 基于轻量化网络的异物检测模型

为实现输电线路巡检图像中侵入异物的智能识别，众多研究者致力于轻量化网络研究。深度学习模型轻量化可依据实现方式划分为模型压缩与轻量化网络设计两类。模型压缩以现有成熟网络为基础，通过模型剪枝、知识蒸馏、低秩分解、网络量化等手段精简模型结构。轻量化网络设计则从网络架构源头出发，采用MobileNet、SqueezeNet、ShuffleNet、GhostNet等轻量化主干模型，结合深度可分离卷积（depthwise separable convolution，DSC）、分组卷积等算子重构网络结构。目前，针对输电线路侵入异物检测的轻量化研究仍相对有限，现有研究多聚焦于优化特征提取网络，通过减少模型参数量与计算量，在维持检测精度的前提下提升推理效率。

为了快速识别并精确定位输电线路巡检图像中的侵入异物和电力部件缺陷，文献[66]提出了基于轻量化Faster R-CNN的输电线路巡检图像检测方法，利用轻量化ZFNet作为训练模型，增强目标特征提取能力，对鸟巢、绝缘子、均压环等目标的识别准确率和检测速度明显优于Faster R-CNN、SSD和YOLO模型，但是模型泛化能力不足。文献[67]提出了基于边缘智能（YOLOv3）的输电线路侵入异物检测方法，以轻量化ECA注意力机制改进的MobileNetv3作为主干网络，MobileNetv3-YOLOv3网络模型实现了塔吊、吊车、导线异物、山火等输电线路侵入异物的智能识别，该模型大小为23.6 MB，其异常目标识别精度和速度要优于MobileNet-SSD和YOLOv5s，但是对密集分布的小目标检测效果

欠佳。文献[68]提出了一种改进YOLOv3的输电线路侵入异物识别方法，MobileNetv2取代Darknet-53作为主干网络，并采用DSC代替检测头网络中的3×3卷积，与YOLOv3、YOLOv4-tiny和YOLOv5s相比，改进YOLOv3模型具有较小的模型尺寸（27 MB）和较快的检测速度，实现了检测速度和精度的有效平衡，但是对复杂背景下的小目标检测效果有限。为了精确实时检测输电线路悬挂的异物，文献[69]提出了基于轻量化YOLOv4-EDAM（YOLOv4 with embedded dual attention mechanism）模型的异物检测方法，使用嵌入挤压和激励网络（squeeze-and-excitation network，SENet）的MobileNetv2代替CSPDarkNet53特征提取网络，通过DSC代替SPP和PANet模块中的标准卷积，将CBAM嵌入SPP和PANet模块中，YOLOv4-EDAM网络结构如图9所示。该模型能够有效地检测输电线路上的鸟巢、风筝、气球等异物，其异物检测精度高达96.71%，检测速度为45 FPS，能够满足实时检测要求。但该模型未充分考虑边缘设备的算力限制，且对密集型小目标的检测效果仍有提升空间。

针对大多数深度学习网络模型复杂、检测速度慢，难以部署于嵌入式边缘设备的问题，文献[70]提出了用于输电线路安全隐患实时检测的轻量化YOLOv5s网络模型，该模型大小仅有14.4 MB，其检测速度高于200 FPS，能够实现工程车辆、烟火、异物等高精度实时检测，但是模型牺牲了部分小目标检测精度，对复杂光照和遮挡场景的适应性较弱，新场景迁移能力不足。文献[71]将YOLOv5s网络模型部署在无人机上，实现了输电线路鸟巢自动检测，其中鸟巢检测精度达到了92.1%，其检测速度为33.9 FPS，与YOLOv3（23.2 FPS）和YOLOXs（31.1 FPS）相比分别快了10.7 FPS和2.8 FPS，但是在不同光照条件、拍摄角度、目标遮挡等实际应用场景下，模型的检测性能会明显下降。文献[72]提出了一种改进YOLOv5的输电线路侵入异物检测方法，通过引入RepConv结构来减少计算量和内存消耗，在主干网络和特征融合网络中嵌入C2F结构，进一步提升模型的检测精度和速度。实验结果表明，改进模型的检测速度高达258 FPS，其异物检测精度达到了96.9%，相较于原始YOLOv5提升了2.2%。但是模型对密集型、小尺寸异物的检测效果仍有提升空间。为了快速准确地识别无人机

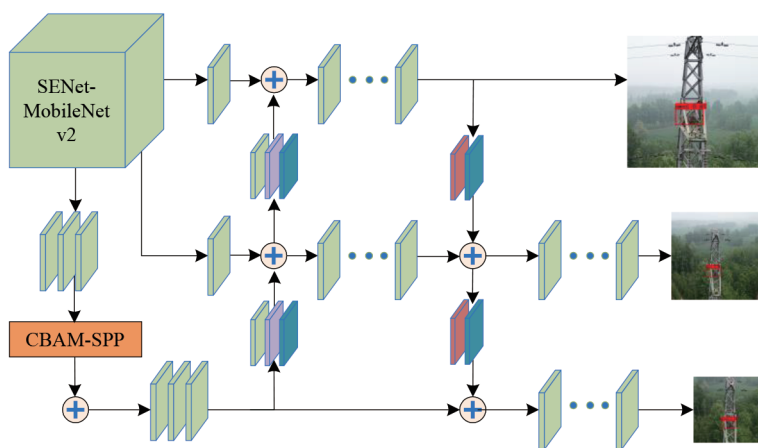


图9 YOLOv4-EDAM网络结构

巡检图像中的鸟巢，文献[73]提出了基于改进YOLOv4的输电线路中鸟巢识别方法，通过在CSP-DarkNet53中引入DSC提高模型检测速度，并在检测头网络中添加SPP模块以提高小目标检测能力，改进模型的鸟巢检测准确率达到91.38%，检测速度为39.24 FPS。综合检测准确率和速度，改进YOLOv4有较高的鸟巢检测准确率和较低的运算开销，其整体性能优于Faster R-CNN、SSD和YOLOv5。但模型轻量化程度不足，难以适配低算力边缘设备。

轻量化网络通过简化结构、删减冗余参数和优化卷积逻辑，有效降低模型计算量。在硬件受限条件下，这类模型推理与训练效率更高，资源占用少，适配嵌入式平台，十分契合无人机巡检的应用需求。但网络轻量化会削弱特征提取能力，压缩模型容量后，其在小样本、恶劣环境、目标遮挡等场景下的泛化性能变差，易出现漏检和误检，难以兼顾轻量化与检测鲁棒性。

综上所述，注意力机制、多尺度特征融合与轻量化网络3类主流改进思路，从不同维度针对性解决输电线路侵入异物检测难题：注意力机制侧重于自适应抑制背景干扰、强化关键目标区域特征响应；多尺度特征融合能够统筹深浅层特征信息，适配巡检图像目标尺度跨度大和小目标占比高的特点；轻量化网络则以降低计算开销为核心，保障模型在边

缘设备的可部署性与实时推理能力。不同改进策略的优缺点及适用场景见表2。表3展示了基于深度学习算法的输电线路侵入异物检测研究成果。

4 Transformer及大模型在智能巡检中的应用

Transformer模型于2017年由Google团队在“Attention Is All You Need”一文中首次提出。该模型以自注意力机制为核心，摒弃了传统循环神经网络（recurrent neural network, RNN）与CNN依赖序列迭代或局部卷积核提取特征的局限，能够实现输入数据不同位置特征信息的自适应聚合与高效建模。这一突破不仅大幅提升了特征提取效率与模型训练速度，还能有效扩大目标感受野、精准捕捉长距离特征依赖。该模型在自然语言处理领域取得突破性进展后，迅速成为跨领域研究热点。

4.1 基于传统Transformer架构的DETR类异物检测模型

随着研究的不断深入，Transformer模型核心架构被快速迁移至CV领域，研究人员通过改进Transformer模型的输入层与注意力机制，将图像数据转化为序列特征，成功将该模型应用于图像分类、目标检测、语义分割等任务，并提出了ViT、Swin-Transformer、DETR等经典改进模型^[74-77]。其中，ViT首次将纯Transformer架构应用于图像分类任务，通过将图像分割为固定尺寸的图像块，采

表2 不同改进策略性能对比分析

改进模型	优势	局限性	适用场景
注意力机制	聚焦目标区域,抑制背景干扰	计算成本增加	复杂背景、小目标检测
多尺度特征融合	覆盖多尺度异物(如风筝/工程车辆)	特征融合难度大,易过拟合	异物尺度差异大
轻量化网络	低资源消耗,适配边缘设备	复杂特征表达能力下降	无人机实时巡检、边缘部署

表3 基于深度学习算法的输电线路侵入异物检测研究成果

异物类型	算法	文献	改进策略	自建数据集	开源工具	GPU	检测精度	推理速度/FPS
鸟巢	RetinaNet	文献[41]	—	5 000	Keras	GTX 2080	94.1%	14.7
	YOLOv3	文献[53]	注意力特征融合(AFF)	3 328	PyTorch	RTX TITAN	93.16%	41.9
	Faster R-CNN	文献[59]	多尺度融合+SE 模块	12 000	—	—	88.68%	12
	YOLOv3	文献[63]	双尺度 YOLOv3 网络	3 000	—	GTX 1080Ti	76.2%	7.14
	RetinaNet	文献[64]	ResNet50+多尺度 P2 层	1 629	—	GTX 1080Ti	90.4%	—
	YOLOv5s	文献[71]	—	2 000	PaddlePaddle	Volta	92.1%	33.9
悬挂物	YOLOv4	文献[73]	检测头增加两个 SPP+DSC	3 000	—	—	97.54%	39.24
	Faster R-CNN	文献[50]	ResNet-CBAM	2 701	—	GTX 2080	79.7%	—
	YOLOX	文献[52]	ASPP+CBAM+GIoU	4 517	PyTorch	RTX TITAN	86.57%	30.08
	YOLOv3	文献[62]	Dense-Net 替换预测网络中的尺度 2 和尺度 3	1 000	TensorFlow	GTX 1080Ti	96%	—
	SSD	文献[31]	轻量化 MobileNet+NMS	4 882	Caffe	GTX 1080Ti	89%	—
	YOLOv4	文献[69]	SENet-MobileNetv2+CBAM-(SPP+PANet)	4 517	TensorFlow+Keras	GTX 1660	96.71%	45
工程车辆	YOLOv5s	文献[72]	RepConv+C2F+Meta-ACON loss	4 517	PyTorch	RTX TITAN	96.9%	258.36
	Mask R-CNN	文献[47]	基于特征映射的迁移学习方法	3 571	—	RTX 2080Ti	91%	5.06
	YOLOv5	文献[51]	CBAM+DIoU+NMS+DANN	3 750	PyTorch	RTX 2080Ti	92.2%	121.9
	Mask R-CNN	文献[60]	PFN+全局特征融合+特征指导候选框生成	17 290	PyTorch	GTX 1080Ti	93.8%	1.22
	YOLOv7	文献[65]	遗传算法+SDP 卷积	7 923	PyTorch	RTX 3070Ti	92.2%	19
	YOLOv3	文献[67]	轻量化 MobileNetv3+模型剪枝	20 000	PyTorch	GTX 2080Ti	69.9%	6.06
覆冰	YOLOv3	文献[68]	轻量化 MobileNetv2	35 000	TensorFlow	GTX 1080	78.8%	60
	SSD	文献[32]	MobileNetv3+多感受野模块	2 924	PyTorch	GTX1080Ti	82.1%	24.5
	YOLOv5	文献[43]	弱监督学习+迁移学习	1 981	PyTorch	GTX 2080	84.3%	—
	U-Net	文献[45]	Swin-Transformer 模块+CBAM+GAN	235	—	—	95.73%	—

用图像分块序列化编码方式完成全局特征建模，但受限于全局注意力机制，存在小目标特征表征能力弱、计算开销偏大等缺陷。在此基础上，Swin-Transformer 引入窗口受限自注意力与层级多尺度特征融合机制，通过分层下采样构建多尺度特征金字塔网络，有效弥补了 ViT 在小目标定位与细粒度特征提取上的不足，显著增强了模型应对密集检测任务的能力。不同于前两者以通用视觉分类和基础特征建模为核心，DETR 实现了 Transformer

与端到端目标检测的深度融合^[78-79]，其网络结构与检测流程如图 10 所示。该算法将目标检测重构为集合预测问题，利用编码器挖掘全局上下文特征，再通过解码器并行预测目标类别与边界框。凭借强大的全局建模能力，DETR 抗干扰与泛化性能表现突出，不仅拓展了 Transformer 的视觉应用范围，也为线路巡检、工业缺陷检测等实际场景提供了新的技术思路。

尽管标准 DETR 及其衍生 RT-DETR 模型具备

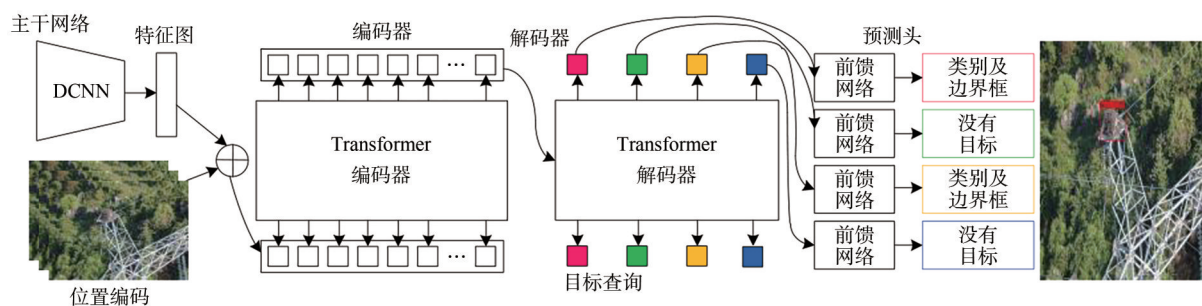


图10 DETR 的网络结构与检测流程

端到端检测与全局建模的优势,但应用于输电线路侵入异物检测任务时,仍面临模型参数量冗余、推理时延高、难以适配无人机边缘设备部署等问题。针对原始DETR算力开销大、边缘部署难的不足,文献[80]从模型结构轻量化角度出发,提出了LMFC-DETR模型,通过替换轻量化主干网络、精简自注意力计算模块,在大幅压缩模型参数量与计算量的前提下,最大程度保留了Transformer全局上下文建模能力。该模型在精度损失可控的前提下,实现了检测性能与部署成本的均衡,适用于算力资源受限的无人机巡检终端,但该方法在复杂背景与遮挡工况下的泛化能力仍有提升空间。为弥补单一轻量化策略的局限性,文献[81]以RT-DETR为基线模型,采用多维度协同优化策略,同时引入轻量化主干网络重构、注意力机制自适应改进与多尺度特征融合增强优化措施。一方面,该模型通过轻量化主干削减冗余计算,满足边缘实时推理需求;另一方面,依托注意力机制强化小目标异物的特征权重,结合多尺度特征融合策略缓解了Transformer对细粒度小目标建模不足的问题。但多模块叠加优化也在一定程度上增加了模型训练难度与超参数调优成本。针对巡检图像中小目标异物特征微弱、检测难度大的问题,文献[82]提出了TLI-DETR模型,在主干网络中嵌入GhostConv残差模块与MobileNetV4轻量化结构,借助卷积模块的局部特征提取优势,增强模型对局部细粒度特征的提取能力,强化小目标异物的浅层细节表征。同时,新增输电线路场景先验引导模块,通过线路区域分割技术预筛选RoI,有效降低了背景噪声对小目标特征的干扰。该模型融合卷积局部特征提取与Transformer全局上下文建模的优势,弥补了Transformer模型对浅层细节特征提取不足的缺陷,但场景先验模块会额外增加前处理环节的推理开销。

综合来看,Transformer架构凭借自注意力机制拥有强大的全局特征建模能力,能够突破传统CNN局部感受野的局限,有效捕捉输电线路异物与复杂背景间的长距离依赖,对光照多变、背景杂乱等巡检场景适应性较强。但自注意力机制需要遍历所有特征位置,存在运算复杂度高、参数冗余、推理时延大等问题,与无人机机载设备算力不足的现状相矛盾。此外,该架构侧重于全局语义建模,浅层细粒度特征表达能力较弱,面对线路中小尺寸和弱特征异物时易受背景噪声影响,检测精度难以

保障。现有研究多通过轻量化降参、强化小目标特征、简化网络结构等方式,平衡模型精度与运行效率。后续可围绕轻重量协同优化、先验知识引入、异构特征融合等方向深入研究,推动Transformer模型在电力巡检场景落地。

4.2 基于Transformer架构的电力行业大模型及工程应用

大模型是基于Transformer架构构建的机器学习模型,其核心架构依托自注意力机制实现对输入数据的全局依赖建模,通过多头注意力并行捕捉不同维度的特征关联,结合前馈神经网络完成特征的非线性转换与强化,最终实现对复杂任务的精准建模。作为大模型的核心基础,十亿级至万亿级的海量参数为特征学习提供充足空间,能够精准拟合电力巡检等复杂场景下的非线性关系。TB级训练数据则带来丰富场景样本,有效提升模型泛化能力与环境适配性。这类模型的典型特征可概括为参数量大、数据规模大和计算开销大。典型代表包括生成式预训练Transformer(generative pre-trained Transformer, GPT)系列与基于Transformer的双向编码器表征(bidirectional encoder representations from Transformer, BERT)系列,其中GPT系列依托Decoder结构实现生成式任务,BERT系列则通过Encoder结构强化双向特征提取,可为电力巡检中的多模态识别、语义理解等任务提供核心架构支撑。

大模型凭借多模态融合、高精度识别等核心优势,已在输电线路智能巡检领域形成一批可复制和可推广的典型示例,不仅大幅提升了巡检作业效率与缺陷识别准确率,更为电力行业数字化与智能化转型提供了坚实的技术支撑^[83-88]。南方电网公司发布了电力生产应用场景首个行业大模型——“大瓦特”。该模型采用基础预训练和行业微调的技术路径,基础预训练阶段依托通用多模态数据构建模型通用特征提取能力,行业微调阶段融入电力行业基础知识与业务规范,实现模型与输电线路巡检等电力场景的深度适配,是跨自然语言处理与计算机视觉双模态的全栈自主可控大模型。其参数量达百亿级,通过优化多头注意力机制的特征分配策略,强化对输电设备缺陷特征和线路环境特征的捕捉能力,实现了算力、算法和应用全流程的自主可控。基于“大瓦特”模型底座,广西电网自主研发了专属CV大模型,该模型针对输电设备图像识别场景

优化了卷积注意力模块，强化缺陷区域的特征提取精度。相较于传统百万级参数量的小模型，该大模型参数量提升至亿级，充足的参数空间使其可精准识别输电线路细微缺陷，在输电设备缺陷检测和隐患识别场景中，平均识别准确率达91.24%，较传统小模型提升15%，缺陷隐患识别率提升近5倍。依托“大瓦特”模型，贵州电网通过CV模型与语义模型的跨模态融合，利用跨模态注意力机制实现图像缺陷特征与语义描述的精准映射，构建输电侧专属大模型。该模型将输电线路缺陷检出率提升至90%以上，识别效率较传统模型提升12倍，人工复核量由日均20人次缩减至1人次，显著降低了输电线路巡检人工成本投入。

在百度的技术支持下，国家电网成功推出“光明电力大模型”，该模型基于文心大模型的Transformer改进架构，采用多模态数据协同预训练和行业场景精调模式，训练数据涵盖亿级图像、文本等多源多模态数据，其中电力行业专属数据占比超40%，该模型通过引入电力设备缺陷标注数据、巡检规范文本数据，具备强大的多模态融合分析能力。依托该模型，国家电网多家分公司将无人机巡检等17类百余个场景规模化落地，年巡检杆塔超500万基，显著提升了电网运维的智能化水平。国网湖南电科院自主研发的10亿节点配网视觉大模型，优化了Transformer架构中的特征金字塔模块，提升了对不同距离和角度配网设备图像的特征提取能力，已完成超30万公里配网巡检任务，识别效率较主流视觉模型提升10%，并有效支撑数十个配网场景的缺陷检测工作。国网安徽合肥供电公司依托“玄视电力视觉大模型”，构建了一套电力巡检图像智能分析体系。针对日常巡检涉及的18类运行环境与杆塔信息、14类电力设备、23类电力元件的状态识别任务，以及26类缺陷隐患的检测场景，该模型通过注意力机制聚焦缺陷关键区域，将隐患识别准确率提升至94%以上。深圳供电局推出的“祝融2.0”，基于Transformer Encoder-Decoder架构强化逻辑推理与自然语言交互能力，通过Decoder模块实现缺陷识别结果的自然语言生成，同时优化Encoder模块的特征提取精度，显著提升了电网安全隐患的告警精度。此外，该模型在输电线路山火烟雾、外力破坏等隐患识别场景中，通过引入场景自适应注意力机制，精准区分干扰因素与真实隐患，准确率高达98%，已在深圳、云南两地实

现跨区域推广应用。

此外，百度与国家电网联合攻关，研发出“国网-百度·文心行业大模型”，该模型基于文心大模型Transformer架构，融入电力行业知识图谱，强化语义理解与场景推理能力，适配输电线路巡检、电网调度等多场景需求。华为云依托盘古电力大模型，针对无人机电力巡检细分场景，优化Transformer架构的轻量化设计，推出“盘古电力巡检大模型”，降低无人机端的计算资源消耗，提升巡检实时性。商汤科技以自研通用多模态大模型为基础，深度融合电力行业实际业务需求与应用场景，依托行业领域数据开展专项微调，对Transformer架构的注意力机制进行优化重构，实现电力目标特征权重的自适应精准分配，为电力系统智能化感知、分析与决策提供技术支撑，有效推动能源电力行业的数字化、智能化转型升级。

大模型凭借Transformer架构与海量参数训练的技术优势，有效突破了传统深度学习算法在输电线路智能巡检中的应用瓶颈，展现出了突出的技术价值与落地优势。首先，大模型依托自监督、半监督学习及AIGC样本生成能力，大幅降低对人工标注数据的依赖，借助涌现特性实现少样本快速微调，有效解决电力缺陷样本稀缺和场景适配性差的行业难题，跨场景泛化能力显著优于YOLO等传统目标检测算法。其次，大模型具备强大的多模态融合能力，可统一整合可见光、红外、点云、声纹及文本台账等异构信息，实现全局特征感知与深层特征挖掘，跳出单模态识别局限，降低缺陷漏检率与误检率。同时，通过模型量化、知识蒸馏等轻量化改造，大模型能够适配无人机、巡检机器人等边缘终端，依托云边端协同架构平衡算力开销与实时检测需求。此外，视觉语言大模型融合行业专业知识与语义分析能力，可完成缺陷智能定级、故障溯源与运维决策推理，实现从目标检测到语义理解、从被动巡检到风险预判的升级，为输电线路智能化和数字化巡检提供坚实的技术支撑。

大模型的应用仍存在明显局限，受算力成本与可解释性的双重约束，其短期内难以完全替代传统目标检测模型。大模型参数量通常达亿级乃至百亿级，训练与推理过程需依托高性能GPU集群，部署成本高昂，难以适配电力巡检中大量边缘终端的轻量化需求。大模型的缺陷识别过程依赖复杂的特征关联计算，无法清晰地输出识别依据，一旦出现

误判,难以快速定位问题根源,这与工业场景对可靠性和可追溯性的要求存在差距。相比之下,传统模型凭借轻量化架构、高实时性及低成本部署等特性,能够快速响应巡检现场的实时检测需求,在中小规模缺陷检测、简单场景巡检中仍具备不可替代的优势。未来,大模型与传统模型的深度融合,将成为解决电力巡检等实际工程问题的关键路径:一方面,以大模型驱动传统模型优化,利用大模型在复杂特征提取和少样本学习中的优势,为传统模型提供预训练特征迁移,提升传统模型在复杂场景中的缺陷识别精度,同时保留其轻量化和高实时性的特点;另一方面,采用混合架构部署模式,在云端利用大模型完成复杂场景分析、缺陷趋势预测等计算任务,在边缘终端部署优化后的传统模型,实现实时缺陷检测、现场告警,兼顾检测精度与部署成本。

5 未来研究展望

在无人机航拍输电线路侵入异物检测领域,基于深度学习的目标检测算法利用DCNN自动提取图像特征,具有检测速度快、效率高等优点,在异物检测任务中取得了显著效果。但同时也面临着诸多挑战,如复杂背景下异物识别、单模态图像检测、大规模图像数据存储与分析、实时性检测、硬件设备限制等。基于目前的研究进展,将未来的输电线路侵入异物智能检测研究归纳为以下几个方面。

(1) 高质量数据集建设。异物检测性能良好的深度学习网络模型离不开数据集的支持,用于输电线路侵入异物检测的自建小样本数据集参差不齐,且数据集没有统一的质量评估标准。建立更大规模和更高质量的输电线路侵入异物数据集,有助于训练出检测性能良好的异物检测网络模型,便于不同的深度学习网络模型之间进行横向比较。

(2) 多模态信息融合^[89-90]。可见光图像是目前无人机航拍输电线路侵入异物检测的主要研究对象,除了传统的可见光图像,结合其他模态的数据,如红外图像、紫外图像、激光雷达点云数据等,将不同模态的数据进行有效融合,可以提供更丰富的信息,充分发挥各种模态的优势,提升异物检测性能。

(3) 云-边-端协同融合^[91-92]。基于云-边-端协同融合的输电线路智能巡检,采用云端全局协同、

边缘节点本地处理和终端实时响应的分层协作模式,针对性解决传统巡检中数据传输冗余、实时响应性能差、算力分配不均衡等核心痛点,推动巡检从被动缺陷检测向主动风险预警演进,为电网运维全流程自动化提供技术支撑。

(4) 传统模型与大模型协同应用^[93-94]。大模型与传统模型深度融合:一方面,利用大模型对传统模型的检测结果进行二次验证与修正,通过语义推理缓解传统模型在复杂场景下易出现的误检和漏检问题;另一方面,通过大模型引导传统模型优化检测策略,根据不同巡检场景与任务需求动态调整模型参数,提升检测效率与适配性。此外,可结合两者优势构建端到端智能巡检系统,实现从数据采集、处理到故障诊断、报告生成的全流程端到端智能化,进一步提升输电线路巡检的可靠性与智能化水平。

(5) 无监督深度学习。输电线路上的悬浮物多种多样、形态各异,且有限的训练样本无法覆盖输电线路上的所有真实场景,网络模型对实际场景中的遮挡、模糊、变形等异物,以及数据集中没有合适种类样本的目标预测效果不佳。无监督学习不依赖特定样本标注,能适应新的检测任务和环境变化。如何利用无监督深度学习方法提升网络模型的检测性能,此研究具有一定的参考价值。

(6) 强化学习与深度学习融合。强化学习可以使网络模型在与环境的交互过程中不断学习和优化检测策略。将强化学习与深度学习深度融合用于输电线路侵入异物智能检测,网络模型可以根据检测结果自动调整检测策略,提高检测效率和准确性。

6 结束语

传统人工巡检方式与早期目标检测算法在无人机巡检输电线路异物识别任务中,存在检测准确率低和作业效率不高的缺陷,已无法满足智能电网高速发展的建设需求。以无人机巡检图像为数据源,依托深度学习目标检测算法实现输电线路侵入异物的高效精准检测,可有效实现输电线路安全隐患的快速排查,进而提升电网精益化运维能力。

本文系统综述了近年来基于深度学习的无人机航拍图像中输电线路侵入异物检测方法的研究进展。首先,按照经典Faster R-CNN、SSD、YOLO及其他检测算法的分类,对各类方法进行了系统介绍。随后,针对深度学习模型在该任务中的优化方

向, 从注意力机制、多尺度特征融合、轻量化网络等技术角度, 归纳总结了各类改进模型的技术路径与应用成效, 这不仅有利于不同深度学习目标检测算法之间的相互比较, 更有助于从事异物检测的研究人员针对不同的异物特点选取合适的目标检测方法。接着, 剖析了基于传统Transformer架构的DETR类异物检测模型, 以及面向电力行业的Transformer大模型及其工程应用。最后, 基于目前深度学习技术的发展现状及近期异物检测的研究思路, 指明高质量数据集建设、多模态信息融合、云-边-端协同融合、传统模型与大模型协同应用、无监督深度学习、强化学习与深度学习融合等是该领域未来的重点研究方向。通过以上论述, 希望为未来输电线路异物检测场景的深度学习目标检测算法研究提供有价值的参考。

参考文献:

- [1] 黄新波. 基于图像感知的输电线路智能巡检综述[J]. 高电压技术, 2024, 50(5): 1826-1841.
Huang X B. Survey of intelligent inspection based on image perception[J]. High Voltage Engineering, 2024, 50(5): 1826-1841.
- [2] 田庆, 胡蓉, 李佐勇, 等. 基于SE-YOLOv5s的绝缘子检测[J]. 智能科学与技术学报, 2021, 3(3): 312-321.
Tian Q, Hu R, Li Z Y, et al. Insulator detection based on SE-YOLOv5s[J]. Chinese Journal of Intelligent Science and Technology, 2021, 3(3): 312-321.
- [3] 张焕龙, 齐企业, 张杰, 等. 基于改进YOLOv5的输电线路鸟巢检测方法研究[J]. 电力系统保护与控制, 2023, 51(2): 151-159.
Zhang H L, Qi Q Y, Zhang J, et al. Bird nest detection method for transmission lines based on improved YOLOv5[J]. Power System Protection and Control, 2023, 51(2): 151-159.
- [4] 何植仟, 曹立杰. UAVAI-YOLO: 无人机航拍图像的小目标检测模型[J]. 智能科学与技术学报, 2024, 6(2): 262-271.
He Z Q, Cao L J. UAVAI-YOLO: dense small target detection algorithm based on UAV aerial images[J]. Chinese Journal of Intelligent Science and Technology, 2024, 6(2): 262-271.
- [5] 付晶, 谈家英, 周立玮, 等. 输电线路无人机巡检图像故障智能识别方法分析[J]. 高电压技术, 2023, 49(S1): 103-110.
Fu J, Tan J Y, Zhou L W, et al. Analysis of intelligent defect recognition method in UAV inspection images of transmission lines[J]. High Voltage Engineering, 2023, 49(S1): 103-110.
- [6] 刘传洋, 吴一全, 刘景景. 基于视觉的输电线路金具锈蚀缺陷检测方法研究进展[J]. 仪器仪表学报, 2024, 45(3): 286-305.
Liu C Y, Wu Y Q, Liu J J. Research progress of vision-based rust defect detection methods for metal fittings in transmission lines[J]. Chinese Journal of Scientific Instrument, 2024, 45(3): 286-305.
- [7] 赵振兵, 蒋志钢, 李延旭, 等. 输电线路部件视觉缺陷检测综述[J]. 中国图象图形学报, 2021, 26(11): 2545-2560.
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.
- [8] 刘志颖, 缪希仁, 陈静, 等. 电力架空线路巡检可见光图像智能处理研究综述[J]. 电网技术, 2020, 44(3): 1057-1069.
- [9] Liu Z Y, Miao X R, Chen J, et al. Review of visible image intelligent processing for transmission line inspection[J]. Power System Technology, 2020, 44(3): 1057-1069.
- [9] 刘开培, 李博强, 秦亮, 等. 深度学习目标检测算法在架空输电线路绝缘子缺陷检测中的应用研究综述[J]. 高电压技术, 2023, 49(9): 3584-3595.
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.
- [10] Luo Y H, Yu X, Yang D S, et al. A survey of intelligent transmission line inspection based on unmanned aerial vehicle[J]. Artificial Intelligence Review, 2023, 56(1): 173-201.
- [11] Liu X Y, Miao X R, Jiang H, et al. Data analysis in visual power line inspection: an in-depth review of deep learning for component detection and fault diagnosis[J]. Annual Reviews in Control, 2020, 50: 253-277.
- [12] Nguyen V N, Jenssen R, Roverso D. Automatic autonomous vision-based power line inspection: a review of current status and the potential role of deep learning[J]. International Journal of Electrical Power & Energy Systems, 2018, 99: 107-120.
- [13] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 580-587.
- [14] Dai J F, Li Y, He K M. R-FCN: object detection via region-based fully convolutional networks[C]//Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS). Massachusetts: MIT Press, 2016: 379-387.
- [15] 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.
- [16] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN[PP]. V3. (2018-01-24)[2026-04-07]. arXiv: arXiv.1703.06870.
- [17] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[C]//Computer Vision-ECCV. Berlin: Springer, 2016: 21-37.
- [18] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2017: 6517-6525.
- [19] Redmon J, Farhadi A. YOLOv3: an incremental improvement[PP]. V1. (2018-04-08)[2026-04-07]. arXiv: arXiv.1804.02767.
- [20] Bochkovskiy A, Wang C Y, Liao H M. YOLOv4: optimal speed and accuracy of object detection[PP]. V1. (2020-04-23)[2026-04-07]. arXiv: arXiv.2004.10934.
- [21] Wang C Y, Bochkovskiy A, Liao H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[PP]. V1. (2022-07-06)[2026-04-07]. arXiv: arXiv.2207.02696.
- [22] Girshick R. Fast R-CNN[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2015: 1440-1448.
- [23] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[C]//Computer Vision-ECCV. Berlin: Springer, 2014: 346-361.
- [24] Chatfield K, Simonyan K, Vedaldi A, et al. Return of the devil in the details: delving deep into convolutional nets[C]//Proceedings of the British Machine Vision Conference. Berlin: Springer, 2014: 1-12.

- [25] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2015: 1-9.
- [26] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 770-778.
- [27] Li J, Yan D F, Luan K, et al. Deep learning-based bird's nest detection on transmission lines using UAV imagery[J]. Applied Sciences, 2020, 10(18): 6147.
- [28] Xiang X Z, Lv N, Guo X L, et al. Engineering vehicles detection based on modified faster R-CNN for power grid surveillance[J]. Sensors, 2018, 18(7): 2258.
- [29] Hao J, Lin S J, Wang L, et al. Research on detection methods of foreign objects and component defects in transmission lines based on deep learning[C]//Proceedings of the 2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2). Piscataway: IEEE Press, 2021: 3070-3075.
- [30] Zhang F, Fan Y L, Cai T, et al. OTL-classifier: towards imaging processing for future unmanned overhead transmission line maintenance[J]. Electronics, 2019, 8(11): 1270.
- [31] 路艳巧, 孙翠英, 曹红卫, 等. 基于边缘计算与深度学习的输电设备异物检测方法[J]. 中国电力, 2020, 53(6): 27-33.
Lu Y Q, Sun C Y, Cao H W, et al. Foreign body detection method for transmission equipment based on edge computing and deep learning[J]. Electric Power, 2020, 53(6): 27-33.
- [32] 马富齐, 王波, 董旭柱, 等. 面向输电线路覆冰厚度辨识的多感受野视觉边缘智能识别方法研究[J]. 电网技术, 2021, 45(6): 2161-2169.
Ma F Q, Wang B, Dong X Z, et al. Receptive field vision edge intelligent recognition for ice thickness identification of transmission line[J]. Power System Technology, 2021, 45(6): 2161-2169.
- [33] Zhu J G, Guo Y, Yue F D, et al. A deep learning method to detect foreign objects for inspecting power transmission lines[J]. IEEE Access, 2020, 8: 94065-94075.
- [34] 严宇平, 杨秋勇, 谢翰阳, 等. 基于改进YOLOv7的输电线路异物检测模型[J]. 南方电网技术, 2024, 18(9): 47-58.
Yan Y P, Yang Q Y, Xie H Y, et al. Foreign object detection model of transmission line based on improved YOLOv7[J]. Southern Power System Technology, 2024, 18(9): 47-58.
- [35] 王亚茹, 冯利龙, 宋晓珂, 等. TFD-YOLOv8: 一种用于输电线路的异物检测方法[J]. 图学学报, 2024, 45(5): 901-912.
Wang Y R, Feng L L, Song X K, et al. TFD-YOLOv8: a transmission line foreign object detection method[J]. Journal of Graphics, 2024, 45(5): 901-912.
- [36] Liu C Y, Wei S F, Zhong S B, et al. YOLO-PowerLite: a lightweight YOLO model for transmission line abnormal target detection[J]. IEEE Access, 2024, 12: 105004-105015.
- [37] Li S, Du K, Li Z Y, et al. An improved YOLOv8n network for transmission line intrusion foreign object detection[C]//Proceedings of the 2024 9th Asia Conference on Power and Electrical Engineering (ACPEE). Piscataway: IEEE Press, 2024: 1559-1565.
- [38] Yu Y Z, Qiu Z B, Liao H S, et al. A method based on multi-network feature fusion and random forest for foreign objects detection on transmission lines[J]. Applied Sciences, 2022, 12(10): 4982.
- [39] 李梅, 郭飞, 张立中, 等. 基于TATLNet的输电场景威胁检测[J]. 工程科学学报, 2020, 42(4): 509-515.
Li M, Guo F, Zhang L Z, et al. Threat detection in transmission scenario based on TATLNet[J]. Chinese Journal of Engineering, 2020, 42(4): 509-515.
- [40] 宋立业, 王诗翱, 刘昕明. 基于改进胶囊网络的电力线巡线异物检测[J]. 电子测量与仪器学报, 2020, 34(12): 49-56.
Song L Y, Wang S A, Liu X M. Detection of foreign objects in power line patrol based on improved capsule network[J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(12): 49-56.
- [41] Zhang H, Zhao J, Chen Y R, et al. Intelligent bird's nest hazard detection of transmission line based on RetinaNet model[J]. Journal of Physics: Conference Series, 2021, 2025(1): 012235.
- [42] Yi J F, Mao J X, Zhang H, et al. PSTL-net: a patchwise self-texture-learning network for transmission line inspection[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 5005714.
- [43] Hao Y P, Liang W, Yang L, et al. Methods of image recognition of overhead power line insulators and ice types based on deep weakly-supervised and transfer learning[J]. IET Generation, Transmission & Distribution, 2022, 16(11): 2140-2153.
- [44] 李利荣, 陈鹏, 张云良, 等. 基于改进CenterNet的输电线路电力器件及异常目标检测[J]. 高电压技术, 2023, 49(11): 4757-4765.
Li L R, Chen P, Zhang Y L, et al. Detection of power devices and abnormal objects in transmission lines based on improved CenterNet[J]. High Voltage Engineering, 2023, 49(11): 4757-4765.
- [45] Hu T, Shen L Q, Wu D, et al. Research on transmission line ice-cover segmentation based on improved U-Net and GAN[J]. Electric Power Systems Research, 2023, 221: 109405.
- [46] 王彤, 朱凌, 范亚洲, 等. 基于改进U-net语义分割遥感影像的线路走廊隐患检测方法[J]. 南方电网技术, 2019, 13(8): 67-73.
Wang T, Zhu L, Fan Y Z, et al. Hidden danger detection method of line corridor in remote sensing image based on improved U-net semantics segmentation[J]. Southern Power System Technology, 2019, 13(8): 67-73.
- [47] 魏贤哲, 卢武, 赵文彬, 等. 基于改进Mask R-CNN的输电线路防外破目标检测方法研究[J]. 电力系统保护与控制, 2021, 49(23): 155-162.
Wei X Z, Lu W, Zhao W B, et al. Target detection method for external damage of a transmission line based on an improved Mask R-CNN algorithm[J]. Power System Protection and Control, 2021, 49(23): 155-162.
- [48] 张焯, 高玉蕊, 黄新波, 等. 基于Fine Mask RCNN的110~220 kV输电铁塔涉鸟故障识别与评估[J]. 电网技术, 2022, 46(6): 2132-2140.
Zhang Y, Gao Y H, Huang X B, et al. Fault identification and evaluation of tower wading in 110~220 kV transmission line with Fine Mask RCNN[J]. Power System Technology, 2022, 46(6): 2132-2140.
- [49] 胡广怡, 韩军, 倪源松, 等. 联合单目深度估计的输电导线异物检测方法[J]. 计算机工程与应用, 2025, 61(23): 305-315.
Hu G Y, Han J, Ni Y S, et al. Method for foreign object detection on transmission lines using monocular depth estimation[J]. Computer Engineering and Applications, 2025, 61(23): 305-315.
- [50] Wang W K, Meng P, Huang W J, et al. Improved faster R-CNN algorithm for transmission line small target detection[C]//International Conference on Neural Computing for Advanced Applications. Singapore: Springer, 2022: 427-441.
- [51] 龙乐云, 周腊吾, 刘淑琴, 等. 改进YOLOv5算法下的输电线路外破隐患目标检测研究[J]. 电子测量与仪器学报, 2022, 36(11): 245-253.
Long L Y, Zhou L W, Liu S Q, et al. Identification of hidden damage targets by external forces based on domain adaptation and attention mechanism[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(11): 245-253.
- [52] Wu M H, Guo L M, Chen R, et al. Improved YOLOX foreign object detection algorithm for transmission lines[J]. Wireless Communications

- and Mobile Computing, 2022(10): 1-10.
- [53] 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.
- [54] 亢洁, 王勃, 刘文波, 等. 融合 CAT-BiFPN 与注意力机制的航拍绝缘子多缺陷检测网络[J]. 高电压技术, 2023, 49(8): 3361-3372.
- Kang J, Wang Q, Liu W B, et al. Detection model of multi-defect of aerial photo insulator by integrating CAT-BiFPN and attention mechanism[J]. High Voltage Engineering, 2023, 49(8): 3361-3372.
- [55] Zhang J, Qi Q, Zhang H, et al. Detection of bird's nest of transmission lines from aerial images based on deep learning model[J]. International Journal of Innovative Computing, 2022, 18(6): 1755-1768.
- [56] 薛昂, 姜恩宇, 张文涛, 等. 基于窗口自注意力网络与 YOLOv5 融合的输电线路通道异物检测[J]. 上海交通大学学报, 2025, 59(3): 413-423.
- Xue A, Jiang E Y, Zhang W T, et al. Detection of foreign bodies in transmission line channels based on the fusion of swin transformer and YOLOv5[J]. Journal of Shanghai Jiaotong University, 2025, 59(3): 413-423.
- [57] Tang C L, Dong H Y, Huang Y R, et al. Foreign object detection for transmission lines based on swin Transformer V2 and YOLOX[J]. The Visual Computer, 2024, 40(5): 3003-3021.
- [58] Sun H B, Shen Q C, Ke H C, et al. Power transmission lines foreign object intrusion detection method for drone aerial images based on improved YOLOv8 network[J]. Drones, 2024, 8(8): 346.
- [59] 王纪武, 罗海保, 鱼鹏飞, 等. 基于 Faster R-CNN 的多尺度高压塔鸟巢检测[J]. 北京交通大学学报, 2019, 43(5): 37-43.
- Wang J W, Luo H B, Yu P F, et al. Bird's nest detection in multi-scale of high-voltage tower based on Faster R-CNN[J]. Journal of Beijing Jiaotong University, 2019, 43(5): 37-43.
- [60] 李辉, 董燕, 刘祥, 等. 基于两阶段深度网络的输电线路异常目标检测方法[J]. 控制与决策, 2022, 37(7): 1873-1882.
- Li H, Dong Y, Liu X, et al. Transmission line abnormal object detection method based on deep network of two-stage[J]. Control and Decision, 2022, 37(7): 1873-1882.
- [61] 丁建, 黄陆明, 朱迪锋, 等. 面向高空塔架鸟巢检测的双尺度 YOLOv3 网络学习[J]. 西安理工大学学报, 2021, 37(2): 253-260.
- Ding J, Huang L M, Zhu D F, et al. Learning to detect high altitude tower nest using dual-scale YOLOv3 network[J]. Journal of Xi'an University of Technology, 2021, 37(2): 253-260.
- [62] 杨剑锋, 秦钟, 庞小龙, 等. 基于深度学习网络的输电线路异物入侵监测和识别方法[J]. 电力系统保护与控制, 2021, 49(4): 37-44.
- Yang J F, Qin Z, Pang X L, et al. Foreign body intrusion monitoring and recognition method based on dense-YOLOv3 deep learning network[J]. Power System Protection and Control, 2021, 49(4): 37-44.
- [63] 张焕坤, 李军毅, 张斌. 基于改进型 YOLO v3 的绝缘子异物检测方法[J]. 中国电力, 2020, 53(2): 49-55.
- Zhang H K, Li J Y, Zhang B. Foreign object detection on insulators based on improved YOLO v3[J]. Electric Power, 2020, 53(2): 49-55.
- [64] 刘国文, 张彩霞, 李斌, 等. 基于改进 RetinaNet 模型的接触网鸟巢检测[J]. 数据采集与处理, 2020, 35(3): 563-571.
- Liu G W, Zhang C X, Li B, et al. Detection of Birds' nest in catenary based on improved RetinaNet model[J]. Journal of Data Acquisition & Processing, 2020, 35(3): 563-571.
- [65] Yu C H, Liu Y K, Zhang W R, et al. Foreign objects identification of transmission line based on improved YOLOv7[J]. IEEE Access, 2023, 11: 51997-52008.
- [66] 魏业文, 李梅, 解园琳, 等. 基于改进 Faster-RCNN 的输电线路巡检图像检测[J]. 电力工程技术, 2022, 41(2): 171-178.
- Wei Y W, Li M, Xie Y L, et al. Transmission line inspection image detection based on improved Faster-RCNN[J]. Jiangsu Electrical Engineering, 2022, 41(2): 171-178.
- [67] 张盛, 王继业, 宋睿, 等. 基于边缘智能的输电线路异常目标高效检测方法研究[J]. 电网技术, 2022, 46(5): 1652-1661.
- Zhang J, Wang J Y, Song R, et al. Research on efficient detection technology of transmission line abnormal target based on edge intelligence[J]. Power System Technology, 2022, 46(5): 1652-1661.
- [68] Li H, Liu L Z, Du J, et al. An improved YOLOv3 for foreign objects detection of transmission lines[J]. IEEE Access, 2022, 10: 45620-45628.
- [69] Qiu Z B, Zhu X, Liao C B, et al. A lightweight YOLOv4-EDAM model for accurate and real-time detection of foreign objects suspended on power lines[J]. IEEE Transactions on Power Delivery, 2023, 38(2): 1329-1340.
- [70] Li H, Jiang F, Guo F, et al. A real-time detection method of safety hazards in transmission lines based on YOLOv5s[C]//Proceedings of SPIE Conference on International Conference on Artificial Intelligence and Intelligent Information Processing (AIIP 2022). Bellingham: SPIE Press, 2022: 283.
- [71] Li H, Dong Y Q, Liu Y X, et al. Design and implementation of UAVs for bird's nest inspection on transmission lines based on deep learning[J]. Drones, 2022, 6(9): 252.
- [72] Zhang H, Zhou X J, Shi Y K, et al. Object detection algorithm of transmission lines based on improved YOLOv5 framework[J]. Journal of Sensors, 2024(9): 1-14.
- [73] 王杨杨, 曹晖, 莫文昊. 基于深度学习的改进型 YOLOv4 输电线路鸟巢检测与识别[J]. 智慧电力, 2023, 51(1): 101-107.
- Wang Y Y, Cao H, Mo W H. Bird's nest detection and identification on improved YOLOv4 transmission line based on deep learning[J]. Smart Power, 2023, 51(1): 101-107.
- [74] 姚云, 胡振斌, 邓涛, 等. 基于自适应池化注意力 Transformer 的唇语识别方法[J]. 智能科学与技术学报, 2025, 7(2): 211-220.
- Yao Y, Hu Z X, Deng T, et al. A lip reading method based on adaptive pooling attention Transformer[J]. Chinese Journal of Intelligent Science and Technology, 2025, 7(2): 211-220.
- [75] 杨澳男, 莫红, 赵世礼, 等. HCANet: 基于分层 Transformer 架构的微表情识别模型[J]. 智能科学与技术学报, 2025, 7(2): 277-286.
- Yang A N, Mo H, Zhao S L, et al. HCANet: a micro expression recognition model based on hierarchical Transformer architecture[J]. Chinese Journal of Intelligent Science and Technology, 2025, 7(2): 277-286.
- [76] 张志成, 王静, 张阳, 等. OrthoGPT: 面向精准诊疗的多模态骨科大模型[J]. 智能科学与技术学报, 2024, 6(3): 338-346.
- Zhang Z C, Wang J, Zhang Y, et al. OrthoGPT: multimodal generative pre-trained transformer models for precise diagnosis and treatment of orthopedics[J]. Chinese Journal of Intelligent Science and Technology, 2024, 6(3): 338-346.
- [77] 丛晓峰, 桂杰, 章军. 基于视觉 Transformer 的多损失融合水下图像增强网络[J]. 智能科学与技术学报, 2022, 4(4): 522-532.
- Cong X F, Gui J, Zhang J. Underwater image enhancement network based on visual Transformer with multiple loss functions fusion[J]. Chinese Journal of Intelligent Science and Technology, 2022, 4(4): 522-532.
- [78] Xu W F, Zhong X H, Luo M, et al. End-to-end insulator string defect detection in a complex background based on a deep learning model[J]. Frontiers in Energy Research, 2022, 10: 928162.
- [79] Cheng Y, Liu D M. An image-based deep learning approach with im-

- proved DETR for power line insulator defect detection[J]. Journal of Sensors, 2022(9): 1-22.
- [80] Li T Y, Zhu C S, Wang Y X, et al. LMFC-DETR: a lightweight model for real-time detection of suspended foreign objects on power lines[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 2539319.
- [81] 王震洲, 孙冬冬, 王建超, 等. 改进 RT-DETR 的输电线路异物检测算法研究[J]. 计算机工程与应用, 2026, 62(2): 116-125.
Wang Z Z, Sun D D, Wang J C, et al. Research on improved RT-DETR algorithm for foreign object detection in transmission lines[J]. Computer Engineering and Applications, 2026, 62(2): 116-125.
- [82] 徐程, 郑春厚, 胡祥, 等. 基于 TLI-DETR 的输电线路巡检图像小目标检测方法[J]. 智慧电力, 2025, 53(9): 117-125.
Xu C, Zheng C H, Hu X, et al. A small object detection method for transmission line inspection images based on TLI-DETR[J]. Smart Power, 2025, 53(9): 117-125.
- [83] 闫玮丹, 齐冬莲, 闫云凤, 等. 面向电力领域的知识图谱与大模型融合关键技术及其典型应用[J]. 高电压技术, 2025, 51(4): 1747-1762.
Yan W D, Qi D L, Yan Y F, et al. Key technologies and typical applications of knowledge graph and large language model fusion in the power sector[J]. High Voltage Engineering, 2025, 51(4): 1747-1762.
- [84] 蒲天骄, 赵琦, 王新迎. 电力人工智能技术研究框架、应用现状及展望[J]. 电网技术, 2025, 49(5): 1751-1770.
Pu T J, Zhao Q, Wang X Y. Technology framework, application status and prospects on electric power artificial intelligence[J]. Power System Technology, 2025, 49(5): 1751-1770.
- [85] 赵振兵, 冯烁, 席悦, 等. 大模型时代: 电力视觉技术新起点[J]. 高电压技术, 2024, 50(5): 1813-1825.
Zhao Z B, Feng S, Xi Y, et al. The era of large models: a new starting point for electric power vision technology[J]. High Voltage Engineering, 2024, 50(5): 1813-1825.
- [86] Guo X L, Bao Y F, Jiang H, et al. RDBL-net: detection of foreign objects on transmission lines based on positional encoding multiscale feature fusion[J]. International Journal of Sensor Networks, 2025, 47(2): 61-71.
- [87] Zhong Y Y, Luo P X, Yan Y F, et al. PowerGPT: a multimodal foundation model for power inspection[J]. Applied Soft Computing, 2026, 186: 456-462.
- [88] Wang J H, Li M X, Luo H C, et al. Power-Llava: large language and vision assistant for power transmission line inspection[C]//Proceedings of the 2024 IEEE International Conference on Image Processing (ICIP). Piscataway: IEEE Press, 2024: 963-969.
- [89] 曹现刚, 刘航, 刘家辉, 等. 面向原煤分选场景的多模态融合异物开集检测方法[J]. 煤炭科学技术, 2026, 54(1): 464-474.
Cao X G, Liu H, Liu J H, et al. Multimodal fusion method for foreign object open-set detection in raw coal sorting scenario[J]. Coal Science and Technology, 2026, 54(1): 464-474.
- [90] 张辉, 杜瑞, 钟杭, 等. 电力设施多模态精细化机器人巡检关键技术及应用[J]. 自动化学报, 2025, 51(1): 20-42.
Zhang H, Du R, Zhong H, et al. The key technology and application of multi-modal fine robot inspection for power facilities[J]. Acta Automatica Sinica, 2025, 51(1): 20-42.
- [91] 王涓, 赵大伟, 朱琪, 等. 大模型驱动的云边协同多智能体具身学习[J]. 自动化学报, 2026, doi: 10.16383/j.aas.c260096.
Wang J, Zhao D W, Zhu Q, et al. Cloud-edge collaborative multi-agent embodied learning driven by large models[J]. Acta Automatica Sinica, 2026, doi: 10.16383/j.aas.c260096.
- [92] 吴杏平, 周飞, 吴春鹏, 等. 电力人工智能算力网络原型构建与调度验证[J]. 电力系统自动化, 2026, doi: cnki.net/urlid/32.1180.TP.20260513.2151.008.
Wu X P, Zhou F, Wu C P, et al. Prototype construction and scheduling verification of computing network for electric power artificial intelligence[J]. Automation of Electric Power Systems, 2026, doi: cnki.net/urlid/32.1180.TP.20260513.2151.008.
- [93] Waga A, Benhlila S, Bekri A, et al. A survey on autonomous navigation for mobile robots: from traditional techniques to deep learning and large language models[J]. Journal of King Saud University Computer and Information Sciences, 2025, 37(7): 198.
- [94] 李柏, 郝金第, 孙跃硕, 等. 平行智能范式视角下的视觉-语言-动作模型发展现状与展望[J]. 智能科学与技术学报, 2025, 7(3): 290-303.
Li B, Hao J D, Sun Y S, et al. Vision-language-action models under parallel intelligence paradigm: the state of the art and future perspectives[J]. Chinese Journal of Intelligent Science and Technology, 2025, 7(3): 290-303.

[作者简介]



刘传洋 (1982-), 男, 南京航空航天大学电子信息工程学院博士生, 池州学院机电工程学院副教授, 主要研究方向为图像处理与机器视觉检测。



吴一全 (1963-), 男, 南京航空航天大学电子信息工程学院教授、博士生导师, 主要研究方向为视觉检测与图像测量、遥感图像处理与理解、红外目标检测与识别、视频处理与智能分析。



刘景景 (1981-), 女, 池州学院机电工程学院副教授, 主要研究方向为图像处理与机器视觉检测。