

DOI: 10.19650/j.cnki.cjsi.J2413568

# 融合双编码与元学习的小样本输电线异物检测

陈哲煊, 高雪莲, 宋佳宇, 刘毅

(华北电力大学电气与电子工程学院 北京 102206)

**摘要:** 输电线作为电力传输的重要设施,其异物检测是电网巡检维护的重要环节。然而,受限于输电线巡检数据的可获得性和输电线所处环境的复杂性,复杂背景下的小样本数据集输电线异物检测依然是一个挑战。就此提出一种由主编码 Swin Transformer 网络和次编码卷积神经网络(CNN)构成的基于两阶段元学习训练策略的双编码目标检测网络(ML-DCTDN),该网络的创新点在于:一方面主编码 Swin Transformer 网络通过2个阶段的元学习训练获得泛化特征提取能力,即第1阶段学习输电线特征,第2阶段学习异物特征,提高其在小样本数据集的目标检测任务中的表现;另一方面该双编码网络分别采用红绿蓝图像和灰度图像的输入方式,通过分层融合模块(LFM)和特征金字塔网络(FPN)模块实现红绿蓝图像和灰度图像的多模态特征融合,既利用了红绿蓝图像丰富的色彩和纹理信息,又借鉴了灰度图像对光线和细节纹理的鲁棒特性,强化了模型在复杂背景下的抗干扰能力与检测能力。消融实验表明,元学习训练策略明显提高了模型平均准确率(mAP),灰度图像输入方法将 mAP 提高了至少4%;与 SSD、Faster RCNN、YOLOv5 以及 YOLOv8 算法的对比实验表明,小样本数据集的输电线异物检测任务中借鉴元学习策略和双编码网络结构,能明显提高复杂背景下模型的目标检测精度,mAP50 和 mAP75 值分别提高到 98.6% 和 64.7%。

**关键词:** 输电线异物检测;元学习;双编码;小样本数据集;Swin Transformer

**中图分类号:** TP391.4 TH183.3 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Small sample foreign body detection in power lines based on double coding and meta-learning

Chen Zhexuan, Gao Xuelian, Song Jiayu, Liu Yi

(College of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China)

**Abstract:** Foreign body detection is a critical component of power grid inspection and maintenance, as it plays an essential role in power transmission. However, detecting foreign objects in transmission lines with small sample data under complex environmental conditions remains a challenging task. This paper proposes a Meta-Learning-based Double Coding Target Detection Network (ML-DCTDN), combining a Swin Transformer Network and a Convolutional Neural Network (CNN). The innovation of this network lies in two key aspects: firstly, the Swin Transformer network enhances its generalization feature extraction ability through a two-stage meta-learning process. In the first stage, it learns transmission line features, while in the second stage, it focuses on foreign object features, improving performance for target detection tasks on small sample datasets. Secondly, the double coding network uses both RGB and grayscale images as inputs, and employs a Layered Fusion Module (LFM) and a Feature Pyramid Network (FPN) to achieve multi-modal feature fusion. This approach leverages the rich color and texture information of RGB images while also utilizing the robustness of grayscale images against lighting variations and fine details. The model's anti-interference and detection capabilities are thus strengthened in complex backgrounds. Ablation experiments reveal that the meta-learning strategy significantly improved the Mean Average Precision (mAP), with grayscale image input increasing the mAP by at least 4%. Comparative experiments with SSD, Faster RCNN, YOLOv5, and YOLOv8 algorithms demonstrate that the proposed meta-learning strategy and double coding network structure greatly enhance detection accuracy in foreign body detection tasks for transmission lines with small sample datasets. The mAP50 and mAP75 values achieved were 98.6% and 64.7%, respectively.

**Keywords:** foreign body detection in power lines; meta-learning; double coding; small samples database; Swin Transformer

## 0 引言

输电线是电力系统输送电能的重要设施,其安全稳定的运行对保障电网正常运转至关重要<sup>[1-2]</sup>。输电线架设环境通常很复杂,受自然环境和各种人为因素的影响,输电线常常会面临异物侵扰的问题,这些异物包括鸟巢、树枝、风筝等。这在很大程度上会影响输送电能的效率,甚至引起短路、断线等严重事故,对电网安全造成威胁。因此,输电线路异物检测技术就显得尤为重要,它可以快速准确地发现和识别输电线路上的潜在的异物风险,以便及时采取维护措施,保证输电线路的可靠性和稳定性<sup>[3]</sup>。

传统的输电线异物检测方法依赖人工巡视,这种方式不仅效率低、耗时耗力,而且在恶劣天气或复杂地形条件下极具挑战性。近年来,随着人们对计算机视觉任务的广泛关注和人工智能技术的飞速发展<sup>[4]</sup>,越来越多的研究者开始探索输电线异物检测技术。通过搭载在无人机或固定监控设备上的摄像头<sup>[5]</sup>,收集输电线路图像并通过深度学习算法进行处理和分析,以实现对外物的自动检测和识别<sup>[6]</sup>。

已发表文献中,很多目标检测模型采用卷积神经网络(convolutional neural network, CNN)提取目标特征,例如经典的目标检测模型—基于区域的卷积神经网络(regions with CNN, RCNN)<sup>[7]</sup>首先生成目标候选区域,再进行分类和边界回归。倪晨<sup>[8]</sup>将 Faster RCNN (faster region-based convolutional neural networks)应用到输电线路故障识别与检测的任务中,利用区域选取(region proposal network, RPN)和软化非极大值抑制算法改进检测模型。和上述双阶段目标检测算法相比,YOLO(you only look once)模型在提取特征后直接进行分类与边界回归,大大缩短训练时长,是单阶段目标检测网络的代表性算法<sup>[9]</sup>。Han等<sup>[10]</sup>将 YOLO-v3 主干网络改进成 Darknet53,在航拍图像的绝缘子定位任务中取得了明显突破。Liu等<sup>[11]</sup>提出单阶段多头检测算法(single shot multibox detector, SSD)算法,通过将 Faster RCNN 的 Anchor 策略融入 YOLO 算法,进一步提高了目标检测速度和准确率。Xu等<sup>[12]</sup>将 SSD 网络应用于航拍绝缘子检测任务中,并取得比 Faster RCNN 更好的检测效果。

在目标检测任务中,输入方式的选择对模型性能和应用效果的影响至关重要。大多数以 CNN 网络为底层架构的主流目标检测算法采用红绿蓝图像作为输入,例如 Faster RCNN<sup>[13-14]</sup>和 YOLO 系列<sup>[15-17]</sup>。红绿蓝图像能够提供丰富的颜色信息和细节,有助于提高对目标的识别能力及准确性。这种多通道的输入方式使得模型能够更有效地提取和学习具有显著视觉特征的目标,例如具有特定颜色或纹理的物体。然而,红绿蓝图像输入需求

的计算资源较高,在图像数据集较小的情况下,还可能会面临过拟合的风险。相对于红绿蓝图像输入而言,也有模型选择灰度图像作为输入以降低计算复杂度,这种方式在处理低光照或复杂背景时显示出良好的抗干扰性。Yao等<sup>[18]</sup>研究证明灰度图像因其降低了色彩变异性的干扰,能更有效地突出形状和运动信息,从而提升了目标检测的准确性和处理速度。

为提高模型目标检测性能,弥补单一模态输入的劣势,一些研究人员采用双编码模型,通过两个并行编码支路,分别处理不同模态的输入数据。Gupta等<sup>[19]</sup>通过将红绿蓝图像和深度图像分别输入到卷积神经网络中进行特征提取,从而提高了目标检测性能。赵轩等<sup>[20]</sup>为减小光照变化等因素的影响利用红绿蓝图像与深度图像的多模态融合进行目标检测。蒋亭亭等<sup>[21]</sup>提出了一种双分支检测网络,通过输入红绿蓝图像和红外图像进行特征融合互补,使检测结果更加准确。然而,这些方法通常在样本量和计算量上存在较大限制,并且对计算机的硬件条件要求较高,因此依然存在一定的局限性。姜国权等<sup>[22]</sup>将红绿蓝图像和灰度图像作为输入,设计了一个双编码目标检测网络,通过融合红绿蓝图像的颜色特征和灰度图像的纹理信息,提升了行人识别任务的性能。然而,该双编码分支均为 ResNet-50,在特征提取性能上稍显不足,另外模型没有实现多模态信息的多尺度有效融合,其目标检测精度较差。

近年来,一些研究人员开始将注意力机制和自注意力模型引入计算机视觉任务。在这一背景下,基于自注意力机制的 Transformer 模型应运而生,主要因为该模型能够处理数据之间的长距离依赖问题,并且拥有并行处理的能力。视觉 Transformer(vision transformer, ViT)作为首个引入 Transformer 结构进行图像分类任务的模型,凭借自注意力机制取得了显著的成果,证明了自注意力机制在计算机视觉中的重要性<sup>[23]</sup>。ViT 将图像块作为输入,利用自注意力机制来捕获图像内部各个部分之间的关联性。然而,在大尺寸图像和多尺度目标任务中 ViT 受限于计算复杂度和局部感知能力,SwiN Transformer 应运而生<sup>[24]</sup>。SwiN Transformer 引入了分区注意力机制,在处理大尺寸图像时,将图像分割为多个可平移窗口,成为目前在计算机视觉领域中的热门选择。

虽然 SwiN Transformer 已成功地应用于目标检测中,但 SwiN Transformer 的性能在很大程度上取决于样本的数量。SwiN Transformer 的良好表现是以大量数据的公共数据集训练为代价的。对于小样本数据,SwiN Transformer 的训练效果会大幅降低,进而降低目标检测的准确性。因此,在小样本训练中,需要针对 SwiN Transformer 进行优化,确保其训练效果。Han等<sup>[25]</sup>采用迁移学习方法冻结网络的下层结构,并使用小样本对模

型的上层结构进行微调,以提高 Transformer 在小样本中的训练效果。Zhang 等<sup>[26]</sup>将一种坐标重建注意机制嵌入多尺度卷积中,帮助双编码 Swin Transformer 提取局部特征,使其更适合于小样本训练任务。

在输电线异物目标检测任务中,输电线异物出现概率较低,收集到的数据量往往有限。通常而言,解决数据集局限性问题的主要技术是数据增强和迁移学习。在样本数量过少的情况下,数据增强可能会导致样本偏差变大。传统迁移学习的源任务与目标任务之间存在一定的相似性,但是在小样本学习场景中,源任务与目标任务可能差异较大,导致迁移的特征泛化能力较弱,从而影响模型的性能。在此背景下,元学习(meta-learning, ML)作为一种有效的小样本学习策略受到广泛关注。Thrun 等<sup>[27]</sup>将元学习概念引入深度学习网络。Kang 等<sup>[28]</sup>设计了一种基于元学习策略的目标检测模型,通过从大样本基类数据中学习元特征,在小样本新类数据上调整检测网络,证明了元学习在小样本目标检测任务中的有效性。

本研究输电线异物检测问题,就输电线检测而言,其特征非常显著<sup>[29]</sup>:输电线本身颜色单一,且具有方向性和延续性,传输架、绝缘子以及其它输电线配件有固定的结构特征。在大多数背景下,例如森林、草地、建筑物等,输电线以及其配套结构较容易与背景分割开,考虑到计算时间的优化问题,只采用灰度图像就可以较好的体现输电线与背景的特征差异。就输电线异物检测而言,传输线通常狭长且具有方向性,因此异物的颜色特征与输电线的特征差异明显,但在复杂背景下,这些异物颜色特征容易与背景中的其他元素混淆。此时红绿蓝图像更能弥补灰度图像可能失去的特征差异,便于复杂背景下的目标检测任务。

综上所述,输电线异物检测面临着数据量小、背景复杂两大问题。针对这两个问题,设计了一种基于元学习的双编码目标检测网络(double coding target detection network by meta-learning, ML-DCTDN)。对于带异物的输电线图像数据匮乏的问题,利用元学习的思想,使用大量无人机航拍图像对主编码 Swin Transformer 网络进行元训练,使其充分学习输电线以及背景的泛化特征,从而降低主编码 Swin Transformer 网络对小数据量样本的训练难度。对于背景复杂问题,首先,利用灰度图像对复杂背景和光照变化具有较好的抗干扰能力的特点,将小样本数据灰度化之后输入次编码 CNN 网络;然后,利用红绿蓝图像对目标的颜色和纹理特征进行更精确的捕捉,从而提升对小目标的识别能力和区分能力;最后,利用分层融合模块(layered fusion module, LFM)实现次编码网络对主编码网络的特征修正和特征融合,凸显主次双编码网络的结构优势。

## 1 基于元学习的训练策略

元学习就是学会学习的学习<sup>[30]</sup>。其核心原理是识别并理解任务之间的共通性和差异性,通过前一个任务的学习成果来辅助新任务的快速学习。元学习模型为了获得通用的策略或先验知识,需要在多个任务上接受训练,最终能够凭借很少的样本快速适应变化的、乃至未见过的任务。本研究采用基于元学习的两阶段训练策略,如图 1 所示。

两阶段元学习训练的目的在于解决 Swin Transformer 网络难以适用于小样本目标检测任务的问题,主要思想就是首先利用大数据集训练 Swin Transformer 网络使其可以识别输电线,再利用小样本数据集训练 Swin Transformer 网络使其可以识别异物。该训练策略利用大量无人机航拍图像训练出泛化能力较强的特征提取网络,在此基础上,可以实现对小样本数据集的微调训练,从而增强模型对新目标的适应能力。具体步骤为:

第 1 阶段将大量无人机航拍图像划分为不包含输电线的数据集(S)和包含输电线的数据集(T),将其作为基类数据集(S+T)进行分类标注。其中,不包含输电线的数据集(S)从 VisDrone2019 公共数据集中选取。为了确保基类数据集可以训练出更准确的模型,选取的图片需和包含输电线的数据集(T)图片风格相似。通过式(1)可以计算出图片间的风格相似度  $\ell$ 。

$$\ell = \frac{1}{2} \sum_{h,w} (\text{Gamm}(\mathbf{S})_{h,w} - \text{Gamm}(\mathbf{T})_{h,w})^2 \quad (1)$$

其中,  $\mathbf{S}$  表示不包含输电线的数据集中的图像数据,  $\mathbf{T}$  表示包含输电线的数据集中的图像数据,  $\text{Gamm}(\mathbf{S})_{h,w}$  表示为  $\mathbf{S}$  在  $h, w$  位置的內积,即  $\text{Gamm}(\mathbf{S})_{h,w} = \sum_c \mathbf{S}_{hc} \mathbf{S}_{wc}$ ,  $h, w, c$  对应图片的高、宽和通道,  $\text{Gamm}(\mathbf{T})_{h,w}$  同理。

利用式(1)计算出 VisDrone2019 公共数据集中的图片与包含输电线的数据集图片风格损失,选取损失最小的图片制作不包含输电线的数据集(S),作为基类数据集(S+T)的负样本。对基类数据集(S+T)执行随机采样以形成支持集 1,训练元学习模型使其可以区分两个类别,即是否包含输电线。查询集 1 是从除去支持集 1 的基类数据集(S+T)中随机选取而成,以确保类别均匀分布。在每次迭代中,主编码 Swin Transformer 网络先使用支持集 1 进行训练,通过优化模型的权重和偏置参数,从而提高模型在支持集 1 上对输电线与非输电线这两个类别的识别能力。然后在查询集 1 上测试模型性能,并基于测试结果进行参数优化。第 1 阶段的目标是利用足够大的基类数据集(S+T)学习具有代表性的输电线图像特征,从而在第 2 阶段元学习训练中能够迅速适应小样本带异物输电线数据集的训练。

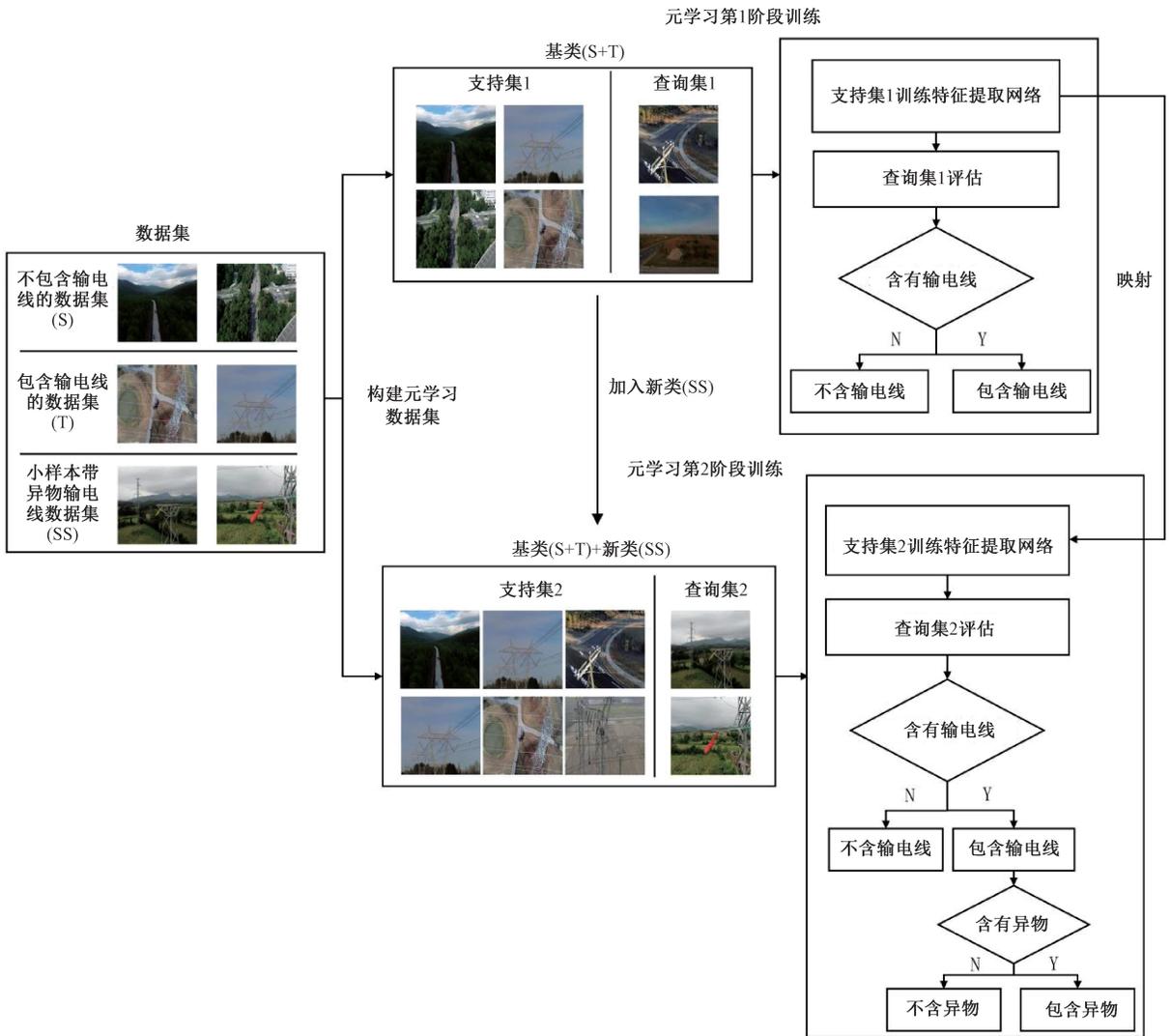


图1 基于元学习的两阶段训练策略

Fig.1 Two-stage training strategy based on meta-learning

在第1阶段基类数据集(S+T)的基础上,第2阶段将小样本带异物输电线数据集(SS)作为新类(SS)加入基类数据集(S+T)。从第2阶段基类(S+T)加新类(SS)数据集中随机选择样本构成新的支持集2,以使模型适应新类别的检测,查询集2由输电线异物数据集中未被选为支持集2的样本组成。将第1阶段训练出的最优特征提取网络作为第2阶段的元学习训练编码器,构建元学习任务。通过元学习使特征提取网络具有从少量样本中学习并泛化的能力,最终让模型能辨别输电线有无异物。

元学习阶段完成后,将训练好的特征提取网络迁移到双编码目标检测网络的主编码网络中。冻结主编码网络参数后,对双编码目标检测网络其他模块进行训练,提高模型检测性能,实现对小样本输电线异物检测数据集的目标检测。

## 2 双编码目标检测网络

ML-DCTDN 包含主编码 Swin Transformer 网络、次编码 CNN 网络、LFM、特征金字塔网络(feature pyramid network, FPN)以及检测头(head),网络结构如图2所示。元学习后特征提取网络迁移至主编码 Swin Transformer 网络,使其可以充分学习输电线以及背景特征,提高了训练效率,实现对小样本红绿蓝图像进行特征提取;次编码 CNN 网络可直接对小样本灰度图像数据进行特征提取;再由 LFM 实现主次编码网络的特征修正和融合,由 FPN 实现多尺度的特征融合,最后通过检测头输出检测结果。

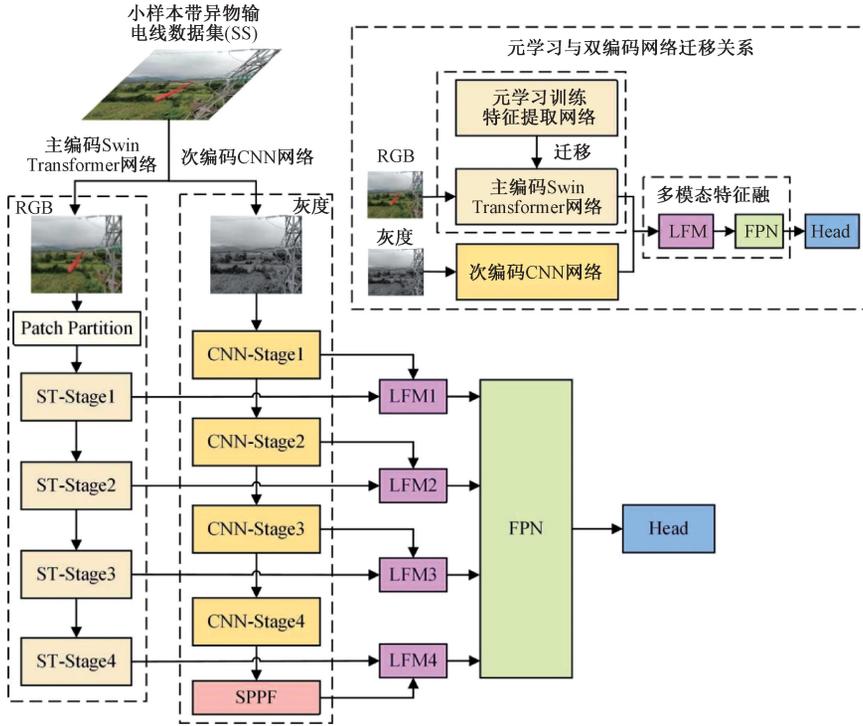


图 2 基于元学习的双编码目标检测网络

Fig. 2 Double coding target detection network by meta-learning

2.1 主编码 Swin Transformer 网络

Transformer 模型在自然语言处理领域取得了巨大成功,但针对图像处理问题,特别是处理大尺寸图像时,传统的 Transformer 结构面临着效率和性能方面的挑战。为了应对这一问题,微软亚洲研究院提出了 Swin Transformer 结构,它是一种专门处理大尺寸图像和长距离依赖关系的 Transformer 架构。本质上,Swin Transformer 模型就是一种基于滑动窗口 (shifted windows) 的分层视觉 Transformer 模型。换言之,Swin Transformer 模型对大尺寸图像进行分层处理,利用图像窗口作为模型输入,使用基于移动窗口的自注意力机制进行信息交互。

主编码 Swin Transformer 网络结构如图 3 所示。通过 Patch Partition 将输入图像划分为固定大小的图像块 (patch),例如将维度为  $h \times w \times 3$  的图片分为  $\frac{h}{4} \times \frac{w}{4} \times 48$  的 patch,再经 Linear Embedding 将每个 patch 映射为一个  $c$  维的特征向量 ( $c$  为 96),为后续的特征提取和处理提供了基础。

ST-block (Swin Transformer block) 是主编码 Swin Transformer 网络的核心部分。主编码 Swin Transformer 网络的每一层计算过程要经过 2 次 ST-block。第 1 个 ST-block 中采用窗口多头自注意力机制 (window based multi-headed self attention, WMSA),其 ST-block 计算过程如式(2)、(3)所示。

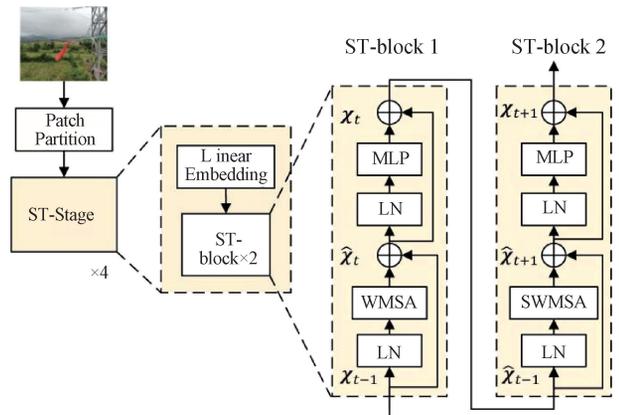


图 3 主编码 Swin Transformer 网络结构

Fig. 3 The main network structure with Swin Transformer

$$\hat{\chi}_t = WMSA(LN(\chi_{t-1})) + \chi_{t-1} \tag{2}$$

$$\chi_t = MLP(LN(\hat{\chi}_t)) + \hat{\chi}_t \tag{3}$$

其中,  $\hat{\chi}_t$  和  $\chi_t$  分别表示 WMSA 模块和多层感知机 (multi-layer perceptron, MLP) 模块的输出特征。

WMSA 计算过程为:

$$Attention = SoftMax(\mathbf{Q}\mathbf{K}^T / \sqrt{d} + B) \mathbf{V} \tag{4}$$

其中,  $\mathbf{Q}, \mathbf{K}$  与  $\mathbf{V}$  为分别为注意力机制的 Query、Key 与 Value 矩阵,  $B$  为位置编码,  $\sqrt{d}$  为缩放因子,  $\mathbf{W}$  是权重矩阵。

WMSA 结构如图 4 所示。

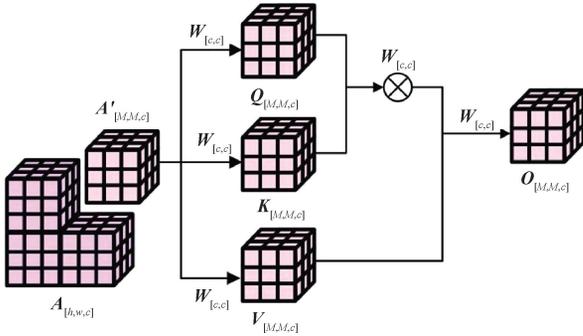


图 4 窗口多头自注意力机制 WMSA

Fig. 4 Window based multi-headed self attention of WMSA

图 4 中, WMSA 在标准多头自注意力机制 (multi-headed self attention, MSA) 基础上引入  $M \times M \times c$  大小的子窗口, 即图 4 中的  $A'$ , 共将  $A$  切割成  $\frac{h}{M} \times \frac{w}{M}$  个子窗口。通过对子窗口的  $Q, K, V$  矩阵加权计算, 使得自注意力机制计算量大大减少, 计算量的计算过程为:

$$\rho(MSA) = 4hwc^2 + 2(hw)^2c \quad (5)$$

$$\rho(WMSA) = (4M^2c^2 + 2M^4c) \times \left(\frac{h}{M} \times \frac{w}{M}\right) = 4hwc^2 + 2M^2hwc \quad (6)$$

将  $h, w$  和  $M$  代入式(5)和(6)后, MSA 与 WMSA 的计算量可差几十甚至上百倍, 本研究中  $h, w$  都为 640, 当  $M$  为 7 时, 2 种计算量相差 1 700 倍。WMSA 很好的解决了内存和计算量的问题, 但由于各个窗口之间无法通信, 限制了它的全局建模能力。因此在第 2 个 ST-block 中使用移动窗口多头自注意力机制 (shifted window based multi-headed self attention, SWMSA), 计算过程如式(7)、(8)所示。

$$\hat{\chi}_{t+1} = SWMSA(LN(\chi_t)) + \chi_t \quad (7)$$

$$\chi_{t+1} = MLP(LN(\hat{\chi}_{t+1})) + \hat{\chi}_{t+1} \quad (8)$$

其中,  $\hat{\chi}_{t+1}$  和  $\chi_{t+1}$  分别表示 SWMSA 模块和 MLP 模块的输出特征。SWMSA 对图像各个窗口进行位移并组合, 再利用 Mask 掩码屏蔽掉不相连部位的注意力计算, 即将图 5 中灰色部分 Mask 掩码为负无穷, 最后将图像窗口还原。

Patch Merging 阶段将经过  $2 \times 2$  采样后的图像块进行合并, 维度由  $\frac{h}{4} \times \frac{w}{4} \times c$  采样合并为  $\frac{h}{8} \times \frac{w}{8} \times 4c$ , 利用卷积操作将  $4c$  维度降为  $2c$  维度, 保持与次编码 CNN 网络每层通道数一致。

Swin Transformer 作为主干网络提取特征, 解决了传统 Transformer 模型面对大尺寸图像时的计算复杂性和内

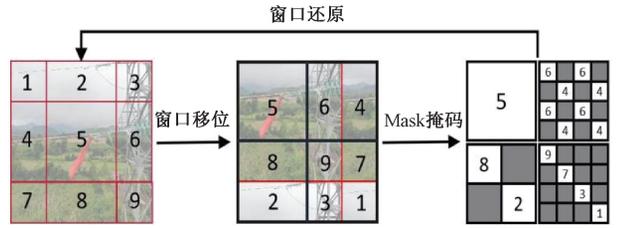


图 5 SWMSA 窗口移位

Fig. 5 SWMSA window shift

存消耗问题。同时, WMSA 提高了模型的感知能力, 从而提高捕获图像特征信息的能力。SWMSA 有效解决了窗口间无法建立联系的问题, 使得局部特征和全局特征得以交互, 从而有效地捕获完整图像中的信息。

### 2.2 次编码 CNN 网络

输电线路所处的自然环境多变, 从光照、天气到背景干扰, 都可能对算法的检测性能带来影响。针对这一问题, 灰度图像因其对复杂背景和光照变化具有较好的抗干扰能力成为一种可行的解决方案。灰度图像相较于红绿蓝图像的三通道颜色而言, 前者通道更加简化, 减少了需要处理的数据量, 提高了算法的效率。另外, 灰度图像更侧重于目标的纹理和结构信息, 而非颜色信息, 这使得灰度图像可以去除颜色信息的干扰, 突出图像的轮廓和细节, 有助于在复杂背景中区分目标。进一步而言, 灰度图像对光照变化的鲁棒性较强, 即使背景光照变化较大, 也不会对目标的亮度信息产生很大影响, 有助于减少因光照变化引起的检测误差。

次编码 CNN 网络结构如图 6 所示。

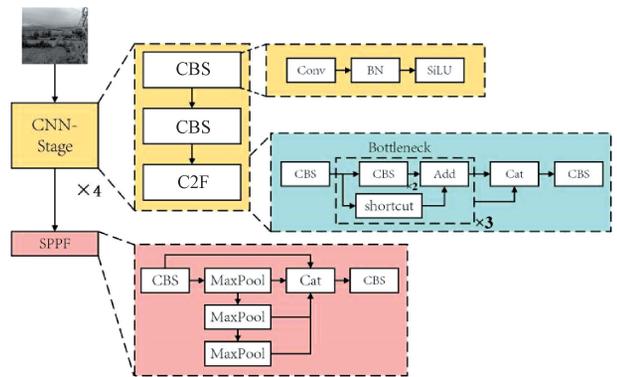


图 6 次编码 CNN 网络结构

Fig. 6 Sub-network with conventional neural network

次编码 CNN 网络的灰度图像输入由彩色图像的 RGB 值经加权平均得到, 如式(9)所示。

$$gray = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (9)$$

次编码 CNN 网络与主编码 Swin Transformer 网络一样, 都采用 4 层结构和多尺度输出。每层结构由 CBS

(convolution、batch normalization、sigmoid linear unit) 模块和 C2F(cross stage partial bottleneck with 2 convolutions) 模块构成。CBS 模块由卷积层(convolution, Conv)、批量归一化层(batch normalization, BN)和 SiLU(sigmoid linear unit) 激活函数组成。卷积层提取特征,批量归一化层用于减少内部协变量偏移,SiLU 激活函数在保持非线性的同时,可以让网络学习更复杂的特征表示。C2F 模块中的 Bottleneck 模块是一种残差学习模块,在卷积层提取特征的同时,通过跳跃连接将输入直接添加到 Bottleneck 模块的输出,避免了深度网络中常见的梯度消失问题,提升了网络的训练稳定性和性能。经过 4 层 CNN-Stage 模块后,加入 SPPF(spatial pyramid pooling-

fast) 模块。SPPF 模块通过多层池化操作提取特征图中不同区域的信息,并将这些信息合并为一个特征向量,增强了特征在不同尺度下的表达能力。

### 2.3 分层融合模块 LFM

LFM 模块实现了主编码 Swin Transformer 网络和次编码 CNN 网络的多层次特征互补。LFM 在不同层次,整合主次 2 个编码网络的特征,并将次编码 CNN 网络捕捉到的纹理细节特征嵌入主编码 Swin Transformer 网络获得的特征中,最终使模型获得既包含高层语义也包含丰富细节的全面特征描述,这对于识别不同尺寸和复杂背景下的目标而言至关重要。LFM 结构如图 7 所示。

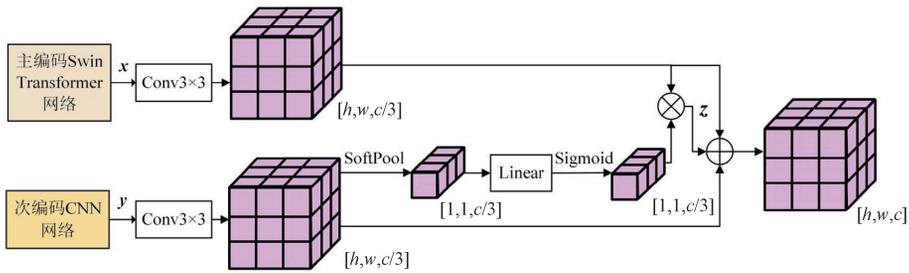


图 7 LFM 结构

Fig. 7 LFM structure

在图 7、式(10)和(11)中,  $x$  与  $y$  分别代表主编码 Swin Transformer 和次编码 CNN 网络的输出,通过  $3 \times 3$  卷积核将其通道数变为原来的  $1/3$ 。对卷积后的  $y$  使用 SoftPool 池化策略来获得次编码网络的通道权重数,通过线性层和激活函数后,与卷积后的  $x$  相乘,最后将 3 个特征相连接,形成 LFM 的输出,通道数恢复为  $c$ 。

$$z = Conv(x) \otimes$$

$$Sigmoid(Linear(SoftPool(Conv(y)))) \quad (10)$$

$$LFM = Conv(x) \oplus z \oplus Conv(y) \quad (11)$$

通过 LFM 的方式,克服了单一模型在处理多尺度、多样性目标时的缺点,相互补充了红绿蓝图像和灰度图像在复杂背景、光线、日照等条件下的图像信息,修正并融合了主编码网络和次编码网络的特征,增强了模型在复杂背景下的特征表达和检测能力,进而提高了目标检测任务的整体性能。

### 2.4 特征金字塔网络 FPN 模块

双编码网络通过分层结构提取不同尺度的特征,其语义信息表征能力和空间几何细节特征抓取能力成反比,也就是说小尺度特征图的语义信息表征能力强,感受野比较大,但是特征图的分辨率低,会造成几何信息细节的缺乏,大尺度特征图与此相反,这就造成了使用单一尺度的特征图对不同尺度目标进行检

测,其表征能力会明显弱化。所以,将多尺度特征进行融合可以进一步提升模型在目标检测任务中的表现和效果。

FPN 是一种自上而下的网络结构<sup>[31]</sup>,可以将小尺度特征图与大尺度特征图横向连接并通过卷积融合在一起,从而实现多尺度特征的融合。如图 8 所示,FPN 的输入为双编码网络经过 LFM 后输出的 4 个特征图,通道数分别为 96、192、384、768。首先,利用  $1 \times 1$  卷积将通道维度均调整为 256。然后,构建自顶向下的通道,经过上采样操作,将高层特征图与相对应的底层特征图连接融合。其次,在每个尺度的特征融合后,添加一层  $3 \times 3$  卷积,用于减少上采样产生的混叠效应,同时平滑特征。

### 2.5 检测头 Head 模块

检测头 Head 模块结构如图 9 所示。首先,对 FPN 输出的各尺度特征图进行采样并按通道维度进行相加。然后,得到的特征图经过 CBS 模块提取重要特征,模块中的批归一化和激活函数使得特征表示更为丰富和有效。然后,处理后的特征图输入 2 条路径,一条路径经过一个  $3 \times 3$  卷积和一个  $1 \times 1$  卷积操作生成物体的分类结果,另一条路径同样经过卷积操作生成边界框回归的

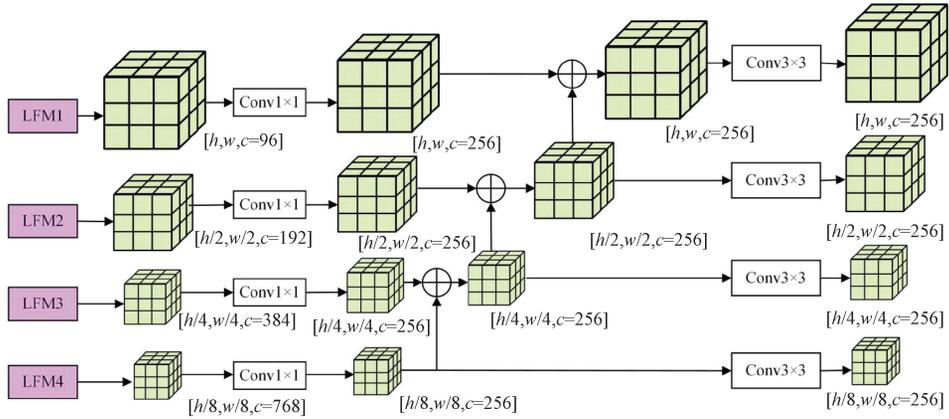


图8 FPN结构

Fig. 8 FPN structure

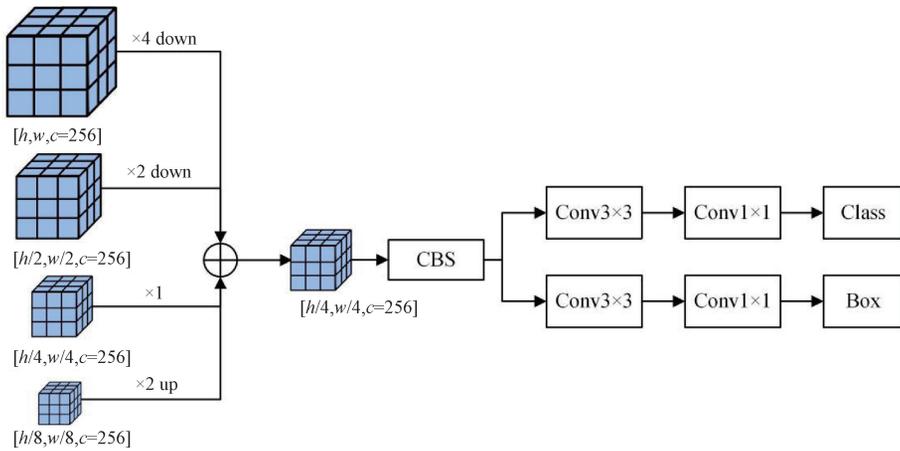


图9 检测头 Head 模块

Fig. 9 Head module structure

结果。通过这种设计,模型能够同时获得分类得分和边界框参数,并分别计算分类损失函数和回归损失函数,以便提升目标检测的整体性能。

### 3 实验与分析

本章通过算法对比和消融实验,验证模型在复杂背景下小样本输电线异物目标检测的可行性和有效性。

#### 3.1 实验数据集

本研究选取 VisDrone2019 数据集(S)和输电线巡检数据集(T)作为元学习基类数据集(S+T),输电线异物巡检数据集(SS)作为元学习新类(SS)和目标检测训练数据集。数据集分布如表1所示。VisDrone2019 数据集由天津大学机器学习和数据挖掘实验室 AISKYEYE 团队收集。该数据集由无人机摄像头采集得到,包含多种场景和不同的天气、光照条件下的图片。为了让主编码网

络可以有效的从复杂背景中提取输电线路及配套设施的特征,本研究从 VisDrone2019 数据集中选取了 2 484 张与输电线巡检数据集场景和风格类似的图片,两者一起作为元学习第1阶段的基类数据集(S+T)。输电线异物巡检数据集(SS)包含2类异物图片,即230张悬挂物图片和200张鸟巢图片。所有图片像素调整为640×640,扩充后按8:1:1的比例划分训练集、测试集和验证集。

表1 数据集分布

Table 1 Data set distribution

数据集	数量	背景	输电线	悬挂物	鸟巢
VisDrone2019 数据集(S)	2 484	√			
输电线巡检数据集(T)	2 484	√	√		
输电线异物巡检数据集 (悬挂物,SS-1)	230	√	√	√	
输电线异物巡检数据集 (鸟巢,SS-2)	200	√	√		√

### 3.2 实验环境与参数设置

实验代码在 Pytorch 开发框架下实现。硬件环境为: A800 显卡, 32 G 内存, 80 G 显存, 32 核 CPU。实验环境如表 2 所示。另外, 本研究使用随机梯度下降法 (stochastic gradient descent, SGD) 完成模型训练, 初始学习率设置为 0.01, 权重衰减为 0.0001。

表 2 实验环境配置

Table 2 Experimental environment configuration

运行环境	详细配置
显卡	A800
内存	32 GB
显存	80 GB
CUDA 版本	11.1
开发语言	Python 3.6
深度学习框架	PyTorch 1.8.1

### 3.3 模型评价指标

采用平均准确率 (mean average precision, mAP) 来评判模型的检测性能。mAP 由各类目标识别准确率 (average precision, AP) 取均值得到。AP 能够体现模型对某一类别目标的检测性能, 该指标由精确度  $P$ 、召回率  $R$  计算得到。计算公式为式 (12)~(15)。

$$P = \frac{TP}{TP + FP} \quad (12)$$

$$R = \frac{TP}{TP + FN} \quad (13)$$

$$AP = \int_0^1 P(R) dr \quad (14)$$

$$mAP = \frac{\sum AP}{N} \quad (15)$$

式 (12)~(15) 中,  $TP$  表示预测框与真实框匹配的数量;  $FP$  表示预测框与真实框不匹配的数量;  $FN$  表示未检测出的真实框数量。判断成功匹配的标准为交并比  $IOU$  (intersection over union) 阈值, 当检测框与真值框的  $IOU$  高于阈值时视为成功检测到异物。将交并比  $IOU$  阈值设为 0.5 和 0.75, 对应平均准确率为  $mAP50$ 、 $mAP75$ 。

### 3.4 消融实验结果与分析

为验证本研究所提出的基于元学习的小样本训练策略和双编码目标检测网络在输电线异物检测任务中的有效性, 本节对模型以及训练策略进行消融实验。

在对 Swin Transformer 网络进行小样本输电线异物检测训练时, 发现其损失函数无明显下降趋势, 为使 Swin Transformer 网络在小样本输电线异物检测任务中发挥优势, 设计了基于元学习的小样本训练策略。对基于元学

习的小样本训练策略进行消融实验, 其结果如表 3 所示。

表 3 元学习消融实验结果

Table 3 Ablation experiment results of meta-learning

(%)

模型	$mAP50$	$mAP75$
DCTDN	39.3	13.2
DCTDN+ML1	74.6	44.1
ML-DCTDN; DCTDN+ML1+ML2	<b>98.6</b>	64.7

表 3 中, DCTDN 是将未经元学习训练的双编码目标检测网络直接进行小样本的目标检测训练。由于训练样本较少, 主编码 Swin Transformer 网络在训练时损失函数收敛较慢, 经过多轮次迭代后  $mAP50$  也仅有 39.3%, 验证了 Swin Transformer 网络自身对小样本学习的不适应性。DCTDN+ML1 为主编码网络经过第 1 阶段元学习训练后迁移到目标检测网络中, 并将其冻结。相较于 DCTDN, 其  $mAP50$  与  $mAP75$  均有较大提升, 分别为 74.6% 和 44.1%。本研究模型 ML-DCTDN (即 DCTDN+ML1+ML2) 为主编码 Swin Transformer 网络经过两阶段元学习训练后迁移到目标检测网络中, 其  $mAP50$  与  $mAP75$  相较于 DCTDN+ML1 均有约 20% 的提升, 分别达到了 98.6% 和 64.7%。可以看出主编码 Swin Transformer 网络结构复杂、参数量大, 应用于小样本数据集时训练困难, 而经过两阶段的元学习训练后, 主编码网络可以提取到较为泛化的特征, 这可以明显减小训练压力, 显著提升后续的小样本学习任务的准确率。

为验证双编码目标检测网络各模块对目标检测效果的影响, 对模型各模块做消融实验, 结果如表 4 所示。

表 4 中, 单编码网络 1 只使用次编码 CNN 网络, 输入为红绿蓝图像; 单编码网络 2 只使用次编码 CNN 网络, 输入为灰度图像; 单编码网络 3 只使用主编码 Swin Transformer 网络, 输入为红绿蓝图像。由单编码网络 1 与单编码网络 2 可知, 次编码 CNN 网络中, 输入灰度图像比输入红绿蓝图像的  $mAP50$  提高了 4%,  $mAP75$  提高了 4.9%。双编码网络 1 与双编码网络 2 使用主编码 Swin Transformer 网络和次编码 CNN 网络, 即使用 ML-DCTDN 模型, 前者双编码网络的次编码网络采用红绿蓝图像作为输入, 后者次编码网络采用灰度图作为输入, 前者比后者的  $mAP$  值均有 4% 左右的提高。可见灰度图像排除了光线、噪声等影响, 强化了边缘细节, 有效提高了仿真效果。表 4 中使用双编码网络的模型比仅使用单编码网络的模型的  $mAP$  均有较大提升, 本研究方法的  $mAP50$  高达 98.6%,  $mAP75$  达到 64.7%。由此可见, 具有异物的输电线图像采用双编码的模型进行检测, 利用红绿蓝图像和灰度图像两种输入方式, 对复杂背景

表4 ML-DCTDN 各模块消融实验结果  
Table 4 Ablation experiment results of each module in ML-DCTDN

模型	主编码 Swin Transformer 网络	次编码 CNN 网络 (红绿蓝图像)	次编码 CNN 网络 (灰度图像)	LFM	$mAP50/\%$	$mAP75/\%$
单编码网络 1		√			89.4	53.9
单编码网络 2			√		93.4	58.8
单编码网络 3	√				91.2	60.4
双编码网络 1	√	√		√	94.1	60.9
双编码网络 2(本研究模型)	√		√	√	98.6	64.7

具有较好的抗干扰能力,既确保模型的多模态信息捕获,又丰富了模型的特征表示,提升了目标检测的性能。

### 3.5 主流模型对比

为了验证本研究模型优势,将其与目前主流的几种目标检测模型进行性能比较,即 SSD、Faster RCNN、YOLOv5、YOLOv8,仿真结果如表 5 所示。

表5 主流模型对比实验结果

Table 5 Comparison of experimental results of mainstream models (%)

模型	$mAP50$	$mAP75$
SSD	87.2	52.5
Faster RCNN	91.1	62.7
YOLOv5	93.6	60.2
YOLOv8	94.1	62.5
YOLOv8+两阶段元学习策略	93.9	62.8
ML-DCTDN	<b>98.6</b>	64.7

由表 5 可知,SSD 的  $mAP50$  和  $mAP75$  仿真效果都低于其他模型,主要原因在于 SSD 固定尺寸的锚框设计无法

完全适应目标的实际大小,且特征图之间的融合不足,导致对小物体的检测性能相对较弱,从而影响了整体检测效果。Faster RCNN、YOLOv5、YOLOv8 和本研究方法 ML-DCTDN 的  $mAP50$  均达到 90% 以上,其中 ML-DCTDN 的  $mAP50$  最高,达到了 98.6%;它们的  $mAP75$  也达到 60% 以上,其中 ML-DCTDN 的  $mAP75$  高达 64.7%。ML-DCTDN 相较于 Faster RCNN 和 YOLOv5 而言, $mAP50$  分别提升了 7.5% 和 5%, $mAP75$  分别提升了 2% 和 4.5%。YOLOv8 经过两阶段元学习策略后再进行目标检测训练,其  $mAP$  与未经过元学习的 YOLOv8 相比,并无明显差距,说明 YOLOv8 在小样本目标检测训练中,已充分提取了特征,因此元学习增强 YOLOv8 特征提取网络的效果并不明显。对比 YOLOv8,ML-DCTDN 的  $mAP50$  提高了 4.5%, $mAP75$  提高 2.2%,分别达到了 98.6% 和 64.7%。进一步证明了 ML-DCTDN 中的主编码分支经过元学习策略后,充分发挥了 Swin Transformer 在目标检测上的优势,达到了比 YOLOv8 模型更高的准确率。

图 10 与图 11 为主流算法检测效果图,其中虚线框为真值框,实线框为模型检测框。

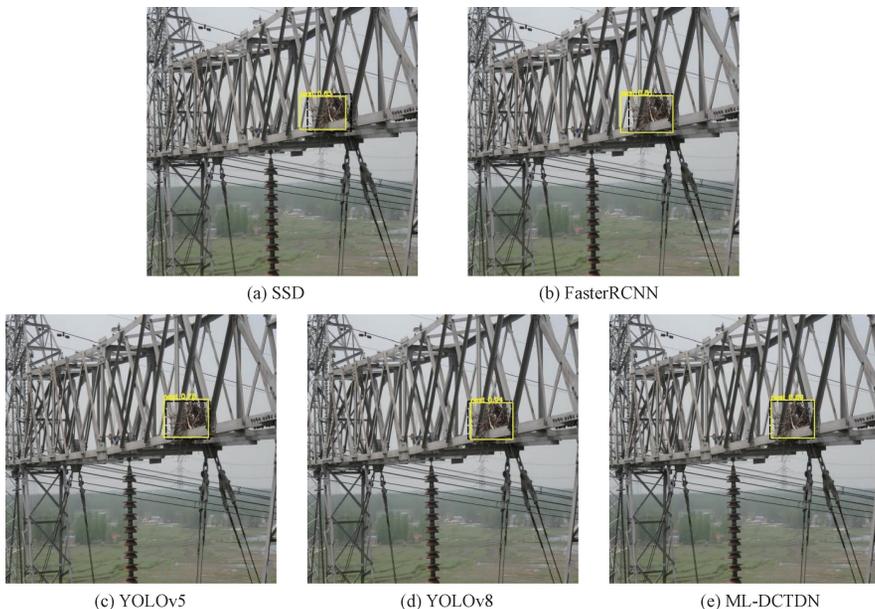


图 10 主流算法检测效果图(鸟巢)

Fig. 10 Detection effect of mainstream algorithm detection(nest)

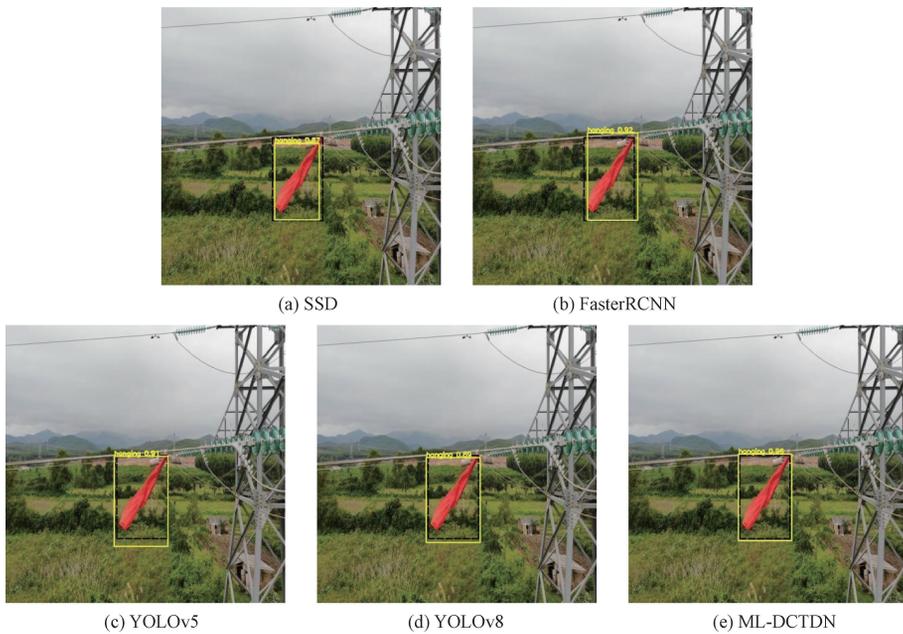


图 11 主流算法检测效果图(悬挂物)

Fig. 11 Detection results of mainstream algorithm detection(hanging object)

由于 YOLOv8 是否增加元学习策略,算法效果影响不大,因此本研究选择了两者中效果最好的进行对比。由图 10(a) 和 11(a) 可知,SSD 是通过预设的多尺度锚框进行检测,其锚框设计存在一定局限性,在捕捉小目标和复杂场景目标的能力上表现稍弱;图 10(b) 和 11(b) Faster RCNN 的检测准确性高度依赖于候选区域质量,而且在小样本数据集训练方面也弱于元学习训练策略,进而影响最后的检测效果;从图 10、11(c) 和 (d) 可看出, YOLOv5 与 YOLOv8 作为单阶段检测器,在检测速率上有天然优势,但在处理复杂背景或小尺度目标时效果欠佳;由图 10(e) 和 11(e) 的检测效果可看出 ML-DCTDN 的检测效果最好。ML-DCTDN 一方面利用元学习策略改善 Swin Transformer 的小样本学习能力,另一方面利用双编码网络结构充分发挥红绿蓝图像和灰度图像的互补信息优势,提高模型在复杂背景下的检测性能,使其在小目标与复杂背景情况下的检测效果均好过其它模型。

## 4 结 论

本研究提出的基于两阶段元学习策略的双编码目标检测网络,旨在解决小样本、复杂背景下的输电线异物检测任务。

针对数据库样本量不足的实际问题,模型采用两阶段元学习的训练策略,结合 Swin Transformer 网络,利用大样本数据库,先识别是否包含输电线,再利用小样本输电线异物数据库,检测是否有异物。提升了模型处理新

类别小样本时的快速适应性和泛化能力,在小样本数据集上有优秀表现;消融实验也表明,元学习的引入使模型的  $mAP$  均有明显提升,证明了两阶段元学习策略在小样本目标检测中的有效性。

针对复杂背景下的目标检测问题,设计双编码网络结构,弥补了单一模型在处理多尺度和复杂目标时的缺陷;利用红绿蓝图像丰富的色彩信息和纹理细节信息以及灰度图像对光线和细节纹理的抗干扰性与鲁棒性,让主编码 Swin Transformer 网络采用红绿蓝图像作为输入,让次编码 CNN 网络采用灰度图像作为输入,通过 LFM 对这两种网络提取的不同特征表达进行互补和融合,增强了双编码模型在复杂背景下的特征表达和检测能力,进而提高了目标检测任务的整体性能。灰度图像输入方法的引入可以使  $mAP50$  和  $mAP75$  至少提高了 4% 左右。本研究方法与 SSD、Faster-RCNN、YOLOv5 与 YOLOv8 算法相比, $mAP$  均有明显提升, $mAP50$  和  $mAP75$  值分别提高到 98.6% 和 64.7%。

本研究为解决小样本数据集应用于复杂背景下的输电线异物检测提供了一种可行的解决方案,未来工作可以考虑模型轻量化以及提高检测速率,探索其在实时检测方面的应用。

## 参考文献

- [1] 和敬涵,罗国敏,程梦晓,等. 新一代人工智能在电力系统故障分析及定位中的研究综述[J]. 中国电机工程学报, 2020, 40(17): 5506-5516.

- review of new generation artificial intelligence in power system fault analysis and location[J]. Proceedings of the CSEE, 2020, 40(17):5506-5516.
- [2] 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 CH 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.
- [3] 徐业东, 蔡亚恒, 李严, 等. 基于YOLOv5s的轻量化架空输电线路鸟巢检测网络[J]. 电子测量技术, 2024, 47(7):138-148.  
XU Y D, CAI Y H, LI Y, et al. Nest detection network of lightweight overhead transmission line based on YOLOv5s [J]. Electronic Measurement Technology, 2024, 47(7):138-148.
- [4] 赵永强, 饶元, 董世鹏, 等. 深度学习目标检测方法综述[J]. 中国图象图形学报, 2020, 25(4):629-654.  
ZHAO Y Q, RAO Y, DONG SH P, et al. Review of deep learning object detection methods [J]. Journal of Image and Graphics, 2020, 25(4):629-654.
- [5] 冷佳旭, 莫梦竟成, 周应华, 等. 无人机视角下的目标检测研究进展[J]. 中国图象图形学报, 2023, 28(9):2563-2586.  
LENG J X, MO M J CH, ZHOU Y H, et al. Research progress of object detection from UAV perspective [J]. Journal of Image and Graphics, 2023, 28(9):2563-2586.
- [6] 陈驰, 彭向阳, 宋爽, 等. 大型无人机电力巡检 LiDAR 点云安全距离诊断方法[J]. 电网技术, 2017, 41(8):2723-2730.  
CHEN CH, PENG X Y, SONG SH, et al. LiDAR point cloud safety distance diagnosis method for large UAV power inspection [J]. Grid Technology, 2017, 41(8):2723-2730.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [J]. IEEE Conference on Computer Vision and Pattern Recognition, 2014:580-587.
- [8] 倪晨. 基于卷积神经网络的输电线路故障快速检测与识别[D]. 呼和浩特: 内蒙古工业大学, 2019.  
NI CH. Rapid fault detection and identification of transmission lines based on convolutional neural networks [D]. Hohhot: Inner Mongolia University of Technology, 2019.
- [9] 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.
- [10] HAN J M, YANG ZH, ZHANG Q Y, et al. A method of insulator faults detection in aerial images for high-voltage transmission lines inspection [J]. Applied Sciences, 2019, 9(10):2009.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [J]. Computer Vision-ECCV 2016, 2016:21-37.
- [12] XU CH F, BO B, LIU Y, et al. Detection method of insulator based on single shot multibox detector [J]. Journal of Physics: Conference Series, 2018, 1069(1):012183.
- [13] REN SH 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 & Machine Intelligence, 2017, 39(6):1137-1149.
- [14] 叶飞, 骆星智, 宋永春, 等. 基于双特征融合的改进 R-CNN 电力小金具缺陷检测方法研究 [J]. 电子测量与仪器学报, 2023, 37(7):213-220.  
YE F, LUO X ZH, SONG Y CH, et al. Research on improved R-CNN defect detection method based on dual feature fusion [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(7):213-220.
- [15] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. ArXiv preprint arXiv:1804.02767, 2018.
- [16] 彭继慎, 孙礼鑫, 王凯, 等. 基于模型压缩的 ED-YOLO 电力巡检无人机避障目标检测算法 [J]. 仪器仪表学报, 2021, 42(10):161-170.  
PENG J SH, SUN L X, WANG K, et al. ED-YOLO power inspection UAV obstacle avoidance target detection algorithm based on model compression [J]. Chinese Journal of Scientific Instrument, 2021, 42(10):161-170.
- [17] 刘熹, 陈晨, 双丰. 基于改进 YOLOv7-tiny 的多种类绝缘子检测算法 [J]. 仪器仪表学报, 2024, 45(9):101-110.  
LIU X, CHEN CH, SUANG F. Multi-type insulator detection algorithm based on improved YOLOv7-tiny [J]. Chinese Journal of Scientific Instrument, 2024, 45(9):101-110.
- [18] YAO H, ZHANG Y J, JIAN H CH, et al. Nighttime pedestrian detection based on Fore-Background contrast learning [J]. Knowledge-Based Systems, 2023, 275:110719.
- [19] GUPTA S, GIRSHICK R, ARBELÁEZ P, et al.

- Learning rich features from RGB-D images for object detection and segmentation [C]. Computer Vision-ECCV 2014, 2014: 345-360.
- [20] 赵轩,郭蔚,刘京. RGB-D 图像中的分步超像素聚合和多模态融合目标检测[J]. 中国图象图形学报, 2018, 23(8):1231-1241.
- ZHAO X, GUO W, LIU J. Step superpixel aggregation and multimodal fusion target detection in RGB-D images[J]. Journal of Image and Graphics, 2018, 23(8):1231-1241.
- [21] 蒋亭亭,刘昱,马欣,等. 多支路协同的 RGB-T 图像显著性目标检测[J]. 中国图象图形学报, 2021, 26(10):2388-2399.
- JIANG T T, LIU Y, MA X, et al. Multi-path collaborative salient object detection based on RGB-T images [J]. Journal of Image and Graphics, 2021, 26(10): 2388-2399.
- [22] 姜国权,肖祺祺,霍占强. 融合 RGB 与灰度图像特征的行人再识别方法[J]. 计算机工程, 2021, 47(4): 226-233, 240.
- JIANG G Q, XIAO ZH ZH, HUO ZH Q. Pedestrian re-identification method combining RGB and grayscale image features[J]. Computer Engineering, 2021, 47(4): 226-233, 240.
- [23] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[J]. ArXiv preprint arXiv: 2010.11929, 2020.
- [24] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: Hierarchical vision Transformer using shifted windows[J]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10012-10022.
- [25] HAN S Y, SHAO H D, CHENG J SH, et al. Convformer-NSE: A novel end-to-end gearbox fault diagnosis framework under heavy noise using joint global and local information [J]. IEEE/ASME Transactions on Mechatronics, 2023:340-349.
- [26] ZHANG Y ZH, ZHAO X Q, LIANG H P, et al. Multiscale dilated convolution and swin-transformer for small sample gearbox fault diagnosis [J]. Applied Intelligence, 2024: 7716-7732.
- [27] THRUN S, PRATT L. Learning to learn: Introduction and overview [M]. Learning to learn. Boston, MA: Springer US, 1998: 3-17.
- [28] KANG B Y, LIU ZH, WANG X, et al. Few-shot object detection via feature reweighting [C]. 2019 IEEE/CVF International Conference on Computer Vision, 2019: 8420-8429.
- [29] 廖圣龙. 航拍输电线图像中部件检测关键技术研究[D]. 大连:大连海事大学,2017.
- LIAO SH L. Research on key technology of component detection in aerial transmission line image [D]. Dalian: Dalian Maritime University, 2017.
- [30] 李凡长,刘洋,吴鹏翔,等. 元学习研究综述[J]. 计算机学报, 2021, 44(2): 422-446.
- LI F ZH, LIU Y, WU P X, et al. Meta-learning research review [J]. Chinese Journal of Computers, 2021, 44(2): 422-446.
- [31] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.

## 作者简介



陈哲煊, 2022 年于华北电力大学获得学士学位, 现为华北电力大学硕士研究生, 主要研究方向为深度学习在电力领域的应用。

E-mail: 2397930220@qq.com

**Chen Zhexuan** received his B.Sc. degree from North China Electric Power University in 2022. He is currently a master student of North China Electric Power University. His main research direction is the application of deep learning technology in the field of electric power.



高雪莲(通信作者), 2005 年获得中科院半导体所博士学位, 现为华北电力大学副教授、硕士生导师, 主要研究方向为人工智能在电力领域的应用。

E-mail: xuelian\_gao@ncepu.edu.cn

**Gao Xuelian** (Corresponding author) received her Ph.D. degree from Institute of Semiconductors, Chinese Academy of Sciences in 2005. She is currently an associate professor and master's supervisor of North China Electric Power University. Her research interest is the application of artificial intelligence in the field of electric power.