



基于多尺度融合金字塔焦点网络的接触网零部件检测

朱新宇 崔浩锐 宋洋

A detector based on a multiscale fusion pyramid focus network for catenary support components

ZHU Xinyu, CUI Haorui, SONG Yang

引用本文:

朱新宇, 崔浩锐, 宋洋. 基于多尺度融合金字塔焦点网络的接触网零部件检测[J]. 北科大: 工程科学学报, 2025, 47(2): 315–327. doi: 10.13374/j.issn2095–9389.2024.05.09.001

ZHU Xinyu, CUI Haorui, SONG Yang. A detector based on a multiscale fusion pyramid focus network for catenary support components[J]. *Chinese Journal of Engineering*, 2025, 47(2): 315–327. doi: 10.13374/j.issn2095–9389.2024.05.09.001

在线阅读 View online: <https://doi.org/10.13374/j.issn2095–9389.2024.05.09.001>

您可能感兴趣的其他文章

Articles you may be interested in

基于多尺度融合金字塔焦点网络的接触网零部件检测

朱新宇¹⁾, 崔浩锐¹⁾, 宋 洋^{2)✉}

1) 中国民用航空飞行学院航空电子电气学院, 广汉 618307 2) 西南交通大学利兹学院, 成都 611756

✉通信作者, E-mail: y.song_ac@hotmail.com

摘 要 作为高铁牵引供电系统的重要组成部分, 接触网系统承担着向动车组传输电能的重要功能. 实际工程运营表明, 受弓网交互产生的持续冲击以及外部环境的影响, 接触网支撑部件可能会出现“松、脱、断、裂”等缺陷, 导致接触网结构可靠性下降, 严重影响接触网系统稳定运行. 因此, 及时精确定位接触网支撑部件(CSCs), 对保障高铁安全运行和完善接触网检修维护策略具有重大意义. 然而, CSCs 的检测通常面临着零部件种类多、尺度差异大、部分零部件微小的问题. 针对以上问题, 本文提出一种基于多尺度融合金字塔焦点网络的接触网零部件检测算法, 将平衡模块和特征金字塔模块相结合, 提高对小目标的检测性能. 首先, 设计了可分离残差金字塔聚合模块(SRPAM), 用于优化模型多尺度特征提取能力、扩大感受野, 缓解 CSCs 检测的多尺度问题; 其次, 设计了一种基于平衡特征金字塔的路径聚合网络(PA-BFPN), 用于提升跨层特征融合效率和微小目标检测性能. 最后, 通过对比试验、可视化实验和消融实验证明了所提方法的有效性和优越性. 其中, 所提的 MFP-FCOS 在 CSCs 数据集上的检测精度(mAP)能够在达到 48.6% 的同时, 实现 30 的 FLOPs (Floating point operations per second), 表明所提方法能够在检测精度和检测速度之间保持良好的平衡.

关键词 深度学习; 目标检测; 接触网支撑组件 (CSCs); 路径聚合特征金字塔 (PA-FPN); 空洞空间卷积池化金字塔 (ASPP)

分类号 TG142.71

A detector based on a multiscale fusion pyramid focus network for catenary support components

ZHU Xinyu¹⁾, CUI Haorui¹⁾, SONG Yang^{2)✉}

1) Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China, Guanghan 618307, China

2) Southwest Jiaotong University, SWJTU-Leeds Joint School, Chengdu 611756, China

✉Corresponding author, E-mail: y.song_ac@hotmail.com

ABSTRACT As a crucial component of a high-speed rail traction power supply system, the catenary system is responsible for transmitting electrical energy to electric multiple units (EMUs). In practice, continuous impacts from pantograph-net interactions and external environmental factors can lead to defects in the catenary's supporting parts, such as looseness, detachment, fracture, and cracking. These issues compromise the reliability of the catenary structure and pose risks to its stable operation. Therefore, timely and accurate positioning of the catenary support components (CSCs) is vital for ensuring the safe operation of high-speed rails and improving the catenary maintenance strategies. In 2012, the former Ministry of Railways of China (now the China Railway Corporation) officially promulgated the General Technical Specifications for High-speed Railway Power Supply Safety Detection and Monitoring System. This study marked a shift from traditional manual inspection methods to intelligent non-contact catenary detection and maintenance using computer vision technology. This study addresses challenges in detection systems by focusing on “catenary part positioning” in the whole detection process from the perspective of the functional integrity of the detection system. Detecting CSCs is challenging because

收稿日期: 2024-05-09

基金项目: 四川省自然科学基金资助项目(24NSFSC7248); 国家自然科学基金资助项目(52477129)

of the variety of parts, scale differences, and small size of components. To overcome these challenges, this study proposes a catenary component detection algorithm that utilizes a multiscale fusion pyramid focus network. This approach integrates a balance module and a feature pyramid module to improve the detection performance of small targets. The separable residual pyramid aggregation module (SRPAM) was designed to optimize multi-scale feature extraction, expand the receptive field, and address multi-scale issues in CSC detection. Furthermore, a path aggregation network based on the equilibrium feature pyramid (PA-BFPN) was designed to improve cross-layer feature fusion efficiency and small object detection performance. Finally, the effectiveness of the proposed method is demonstrated through comparative experiments, visual analysis of the results, multi-scale feature fusion module experiments, feature pyramid network experiments, and ablation studies. The results demonstrate that the proposed multiscale feature pyramid FCOS (MFP-FCOS) algorithm offers excellent overall performance compared to many classical algorithms. Visualization experiments confirm its effectiveness in detecting targets across different scales and effectively solving small-scale and multi-scale sample detection challenges. The proposed SPRAM effectively mitigates information loss and improves feature extraction performance, whereas the proposed PA-BFPN obtains more comprehensive feature information. In summary, the proposed MFP-FCOS achieved a detection accuracy (mAP) of 48.6% on the CSC dataset with 30 floating point operations per second (FLOPs), indicating a balanced trade-off between detection accuracy and detection speed.

KEY WORDS deep learning; object detection; catenary support components (CSCs); path aggregation feature pyramids (PA-FPN); pyramid of hollow space convolutional pooling(ASPP)

在受电弓和接触网系统交互的过程中,两者之间的直接机械接触将很容易引起接触网支撑悬挂装置的剧烈震动,进而导致接触网支撑部件的机械故障,例如松动,损坏,变形.此类故障将极大地影响整个铁路系统的安全运行^[1-5].检测系统可以有效监测到接触网支撑部件异常,从而铁路管理者能及时的修复它并阻止事故的发生^[6-10]由于接触网的机械结构和拍摄条件的特殊性,接触网支撑部件的检测面临以下问题:(1)构件类型多样性;(2)不同零部件尺度差异大(例如螺钉和绝缘体);(3)部分零部件尺寸小(例如螺钉和螺母).检测系统难以有效的执行接触网支持部件检测任务.因此有必要对接触网支撑部件实现快速、准确的检测.

目前,现有的接触网零部件检测方法主要包括以下两方面:(1)传统的目标检测算法;(2)基于深度学习的图像目标检测算法^[11-12].近年,机器视觉技术已经完全代替人工完成繁琐的检测工作.起初,Ye等^[13]提出利用仿射不变矩对接触网绝缘子进行图元匹配定位,并根据绝缘子的灰度极小值分布去判断绝缘子是否存在破损故障.韩焯等^[14]提出利用 SIFT 对接触网旋转双耳进行特征提取与匹配定位,并通过检测旋转双耳的边缘线曲度判断双耳是否存在断裂故障.杨红梅等^[15]提出利用 SURF 提取接触网旋转双耳特征,实现粗定位,并利用 Hough 变化实现对接触网旋转双耳的检测和定位.张桂南等^[16]提出利用 DHOG (Difference HOG) 特征提取接触网承力索底座特

征,并使用 AdaBoost 级联分类器实现接触网底座检测.综上所述,上述传统的接触网检测方法大多采用“先特征选择,后模板匹配”两步策略.值得一提的是,这些方法虽然在单一类别的组件上取得了良好的检测效果,但受到外部环境变化和超参数设置的限制,难以具备良好的泛化能力.

基于深度学习的图像目标检测算法大致可以分为一级检测器和二级检测器.对于二级检测器而言^[17-18],Liu等^[19]设计了基于 Faster R-CNN 网络框架的特征提取网络,用来精确检测等电位线,并使用马尔可夫随机场模型精确分割等电位线区域.Zhong等^[20]利用 PAVNET 网络, Hough 变换以及 CV 模型实现对接触网紧固件开口销松脱故障检测. Kang等^[21]首先使用 Faster R-CNN 定位悬链绝缘子等部件,然后搭建一个多任务网络实现绝缘子的分类和重建.综上所述,这些方法大多数是两级检测器和传统图像检测方法的结合,这往往使得检测时间过长.一级检测器可以有效的避免上述问题. Chen等^[22]基于单阶段目标框架下,建立了一个三级网络来检测悬链线紧固件.首先通过 SSD 检测网络框架检测接触网连接件;其次,通过 YOLOV3 检测接触网紧固件;最后,通过深层卷积网络来检测紧固件的故障.总之,随着深度学习的发展,研究者开始使用深度学习检测方法来替代传统的图像检测方法.

基于深度学习的目标检测方法已经取得了良好的效果, Qin等^[23]重新设计 MBCnv 中的注意力机制模块,将编码空间(ES)注意力机制作为一个

新的模块,有效解决图像空间信息被忽略的问题. Luo 等^[24]将 CBAM 注意力机制引入到骨干网络中,并增加 GSConv 模块和 SIoU 损失函数. Liu 等^[25]将骨干网络中的 C3 模块替换为 CoTNet Transformer 模块,颈部与 C3 模块集成为全局注意力机制(GAM). Ren 等^[26]利用可变形卷积提高精度,采用 SimSPPF 降低计算复杂度,最后利用 SIoU 损失函数有效解决边界回归问题.然而这些方法大多数用于检测悬链线中的局部或单个零部件.接触网零部件检测实际面临着类型多样、尺度跨度大(多尺度)、部分零部件尺寸小(小目标)的问题.对于一般的检测算法来说,会导致定位和分类之间的不匹配问题.因此,通用检测算法很难在该领域有效发挥其性能.许多研究表明多尺度融合模块^[27-30](SPP、ASPP、Inception)和特征金字塔网络^[31-34](FPN、Bi-FPN、PA-FPN)是多尺度、小目标问题的有效解决方案.融合可变形卷积的检测头可以显著提高边界框回归任务和分类任务的一致性.因此,将上述两个特征增强模块和相应的检测头引入神经网络中,可以有效地提高通用检测算法的检测性能.

因此,受上述思想启发,本文提出了一种基于多尺度融合金字塔焦点网络的无锚检测器用于检测接触网关键零部件,称为 MFP-FCOS,其算法框架如图一所示.本文的主要贡献如下:

(1)设计一种新的空洞空间金字塔池化(ASPP)模块,以提高模型提取多尺度特征和弱化背景信

息能力.

(2)设计一种新的基于平衡特征金字塔的路径聚合网络(PA-BFPN),提高对小目标的检测性能.

(3)通过对比分析等实验证明了所提 MFP-FCOS 检测器的设计合理性.

1 基于多尺度融合金字塔焦点网络的接触网零部件检测算法设计

由于接触网支持和悬挂零部件的高频振动,其接触网关键零部件发生松动、损坏和断裂等不同程度的机械故障.针对这些问题,本文在 MobileNetv2 算法的基础上提出了一种新的基于多尺度融合金字塔焦点网络(MFP-FCOS).该方法能够更好地平衡对接触网支撑部件检测的准确性和检测速度.如图 1 所示,该法由改进的 Backbone (MobileNet-P)、Neck(PA-BFPN)和 Head(FCOSX-Head)组成.其中,MobileNet-P 是由可分离残差金字塔聚合模块(SRPAM)和可变形卷积^[35](DCN)组成.

1.1 可分离残差金字塔聚合模块(SRPAM)

如下文所示,接触网支撑部件中包含大型物体和小型物体,例如绝缘子、套筒螺钉等.传统目标检测算法难以有效解决多尺度问题.文献[28]表明,空洞空间池化金字塔(ASPP)模块可以有效提高网络提取多尺度特征的能力.因此,一些研究^[29]直接在下游采样网络中加入了 ASPP 模块,解决多尺度问题.传统的 ASPP 模块主要通过“直接特征连

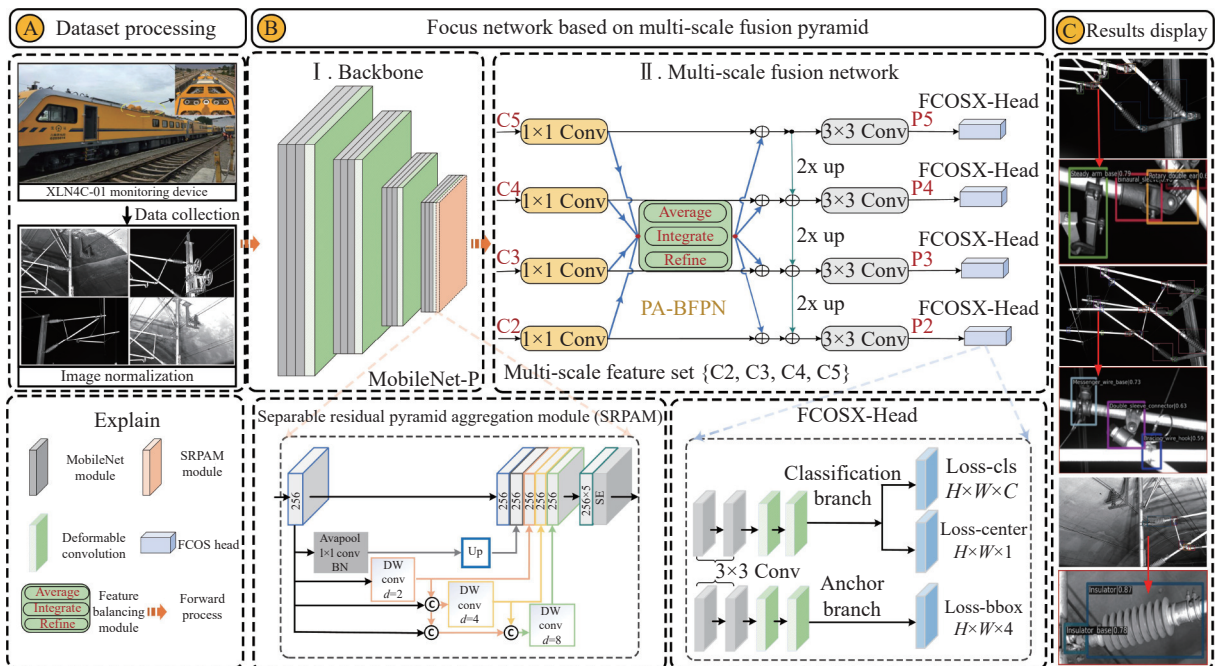


图 1 基于多尺度融合金字塔的焦点网络

Fig.1 Focus network based on a multiscale fusion pyramid

接”来提取和融合不同尺度的分支之间的特征,难以联系不同分支之间的关系.受文献 [36]的改进启发,如图 2 所示,本文在 ASPP 模块中引入了密集连接,并额外引入了注意力机制^[37]和深度可分离卷积^[38],以进一步增强模块提取关键特征的能力,减少模块的计算量.

具体来说,可分离残差金字塔聚合模块包括不同膨胀率的卷积模块、平均池化模块和卷积注意力融合模块(1×1 Conv+SE 模块^[37]).首先,SRPAM 模块通过“密集连接”的方式,将不同膨胀系数的 Block 块的输出特征进行整合,从而获得不同感受野的融合特征.然后,将融合特征和平均池化分支的输出特征整合到通道维度中,增强融合特征的全局信息;最后,将 4 个分支的融合特征导入到由 1×1 Conv 和 SE 模块组成的卷积注意力模块中,降低融合特征的通道维数,增强判别特征信息的权重.综上所述,SRPAM 模块可以高效利用各个尺度分支之间的多尺度信息.

值得一提的是,“密集连接”的引入将会导致

特征通道数目的增加,增加模块的参数和计算量.因此,进一步将深度可分离卷积(1×1 DWConv、3×1 DWConv 和 1×3 DWConv)替换 ASPP 块中的 1×1 Conv 和 3×3 Conv,并在两个 DWConv 的输出之间添加残差连接,以减少参数的数量,更有效地保留特征信息.综上所述,在 SPRAM 块中引入深度可分卷积和残差连接,可以有效缓解深度可分卷积导致的特征信息丢失问题.

1.2 PA-BFPN

特征金字塔融合网络通常添加在主干网络之后,目的是融合主干网络中不同分辨率的特征图,有效增强模型对小目标和多尺度目标的检测效果.虽然典型的特征金字塔融合网络如 FPN^[31]、NAS-FPN^[34] 和 Bi-FPN^[33] 在许多检测框架中发挥了良好的作用,但仍然面临以下问题:(1)许多方法集中于相邻特征层的信息,而忽略跨层信息;(2)采用固定的拓扑结构,使得特征提取的性能陷入局部最优.PA-FPN^[32] 通过将路径聚合网络(PANet)技术引入 FPN 来实现其结构和性能优化.如图 3 所

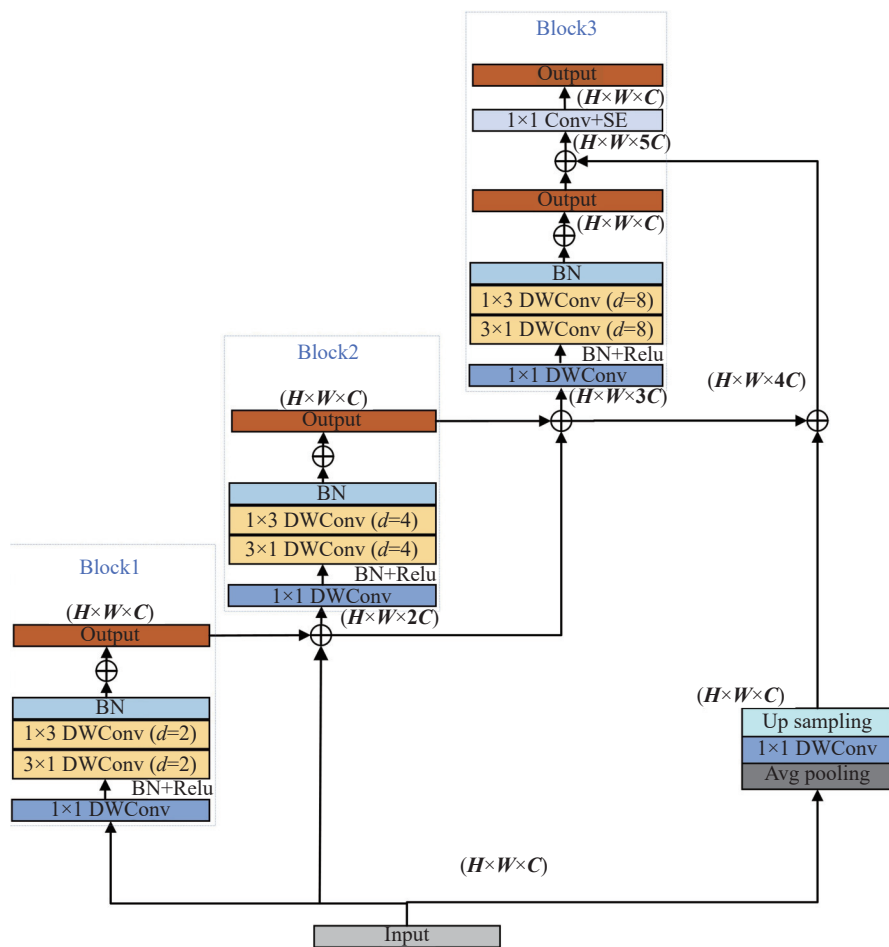


图 2 SRPAM 模块

Fig.2 SRPAM module

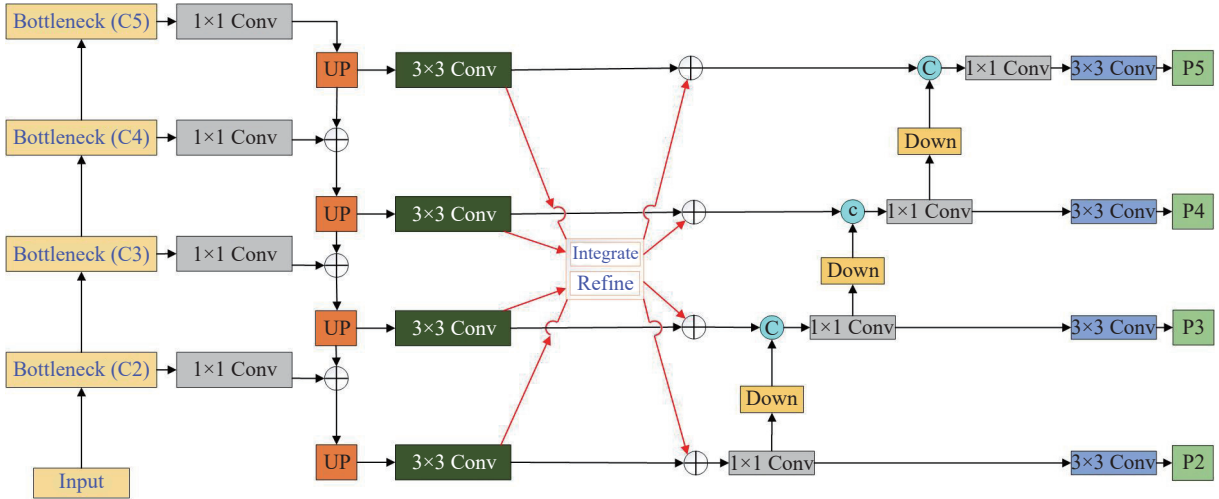


图3 PA-BFPN 的结构图

Fig.3 Structural diagram of PA-BFPN Block

示, 将设计的跨层特征平衡模块引入到 PA-FPN 的横向连接中, 以构建新型层平衡特征金字塔网络 (PA-BFPN). 具体来说, 该模块通过跨层特征平衡模块将跨层信息引入 PA-FPN 各层的融合特征中, 从而能够提取具有层级特异性和层级平衡性的特征, 提高特征融合的性能.

首先, 为了从跨层特征中获得不同尺度的特征信息, 通过跨级特征平衡模块集成主干网络的多级特征 $\{C_5, C_4, C_3, C_2\}$. 设主干网络第 i 层的横向连接输出定义为 L_i , 跨层特征平衡模块的输出定义为 L_i^{out} , 多层特征的数量定义为 N , 底层特征表示和顶层特征表示分别定义为 i_{bom} 和 i_{top} . 首先, 选择中间层的特征分辨率作为目标分辨率, 例如 L_3 . 其次, 使用插值上采样操作和自适应池化操作将剩余的特征分辨率 $\{L_2, L_4, L_5\}$ 调整为目标分辨率.

$$B = \frac{1}{N} \sum_{i_{\text{bom}}}^{i_{\text{top}}} \text{Rescale}(L_i), \text{Rescale} \begin{cases} \text{adapool}, & i < 3 \\ \text{none}, & i = 3 \\ \text{unsample}, & i \end{cases} \quad (1)$$

其中, B 表示将自适应池化后的跨层特征图 $\{C_5, C_4, C_3, C_2\}$ 相加并求平均值; adapool 代表自适应池化操作; unsample 表示插值下采样操作; none 表示不进行任何操作. 接下来, 为了进一步细化和增强跨级信息的平均输出, 将非局部注意力模块 (Non-local, NL)^[39] 引入到跨级特征平衡模块中, 强化特征图中的重要特征. 对应的公式表示如下.

$$\text{Att} = \text{Refine}_{\text{NL}}(B) \quad (2)$$

其中, Att 表示多级特征输出的精细表示; $\text{Refine}_{\text{NL}}$ 代表非局部注意力模块. 为了有效地将跨层信息引入 PA-FPN 中每个特征层的横向连接中, 采用逆

运算公式 (式 (1)) 表示 B , 通过逐元素相加法表示 Att, 并恢复到对应的横向连接卷积特征 $\{L_2, L_3, L_4, L_5\}$ 的大小.

$$R_i = \text{Restore}(\text{Att}), \text{Restore} \begin{cases} \text{unsample}, & i < 3 \\ \text{none}, & i = 3 \\ \text{adapool}, & i > 3 \end{cases} \quad (3)$$

$$L_i^{\text{out}} = L_i + R_i, \quad i \in \{2, 3, 4, 5\} \quad (4)$$

其中, R_i 表示第 i 层横向卷积精细表示 Att 的恢复结果; L_i^{out} 表示跨级平衡特征的最终输出 (具有层级特异性和层级平衡性的增强横向特征). 最后, 将增强的横向特征表示 L_i^{out} 引入到 PA-FPN 的剩余部分, 进行“从上到下”的特征融合.

$$P_i = \text{PABFPN}(L_i^{\text{out}}), \quad i = \{2, 3, 4, 5\} \quad (5)$$

其中, P_i 为 PA-BFPN 模块第 i 级增强融合特征. 因此, 模块从主干网络中获得各层次增强的融合特征.

1.3 MobileNet-P 主干网络和 FCOSX-Head 检测头

MobileNetv2^[40] 作为轻量化网络, 具有参数量小、复杂度低、推理速度快等优点. 原始的 MobileNetv2 有五个阶段, 第一阶段由 Conv-2D、BN 批处理标准化和 Relu 激活函数组成. 后四个阶段中的第一模块为下采样模块 (Down-sampling block), 如图 4(b) 所示, 其余模块为残差块 (Identity block) 组成, 如图 4(a) 所示. 其中, 残差块的主流分支是由 1×1 Conv、 3×3 DWConv、 1×1 Conv 和一个残差边组成. 下采样模块结构由 1×1 Conv、 3×3 DWConv 和 1×1 Conv 组成. 从 MobileNetv2 的网络结构可以得出, 该网络主要是由二维卷积 (Conv-2D) 和深度可分离卷积 (DWConv) 组成. 虽然此网络具有良好的特征提取能力和相对较低的计算要求, 但并不

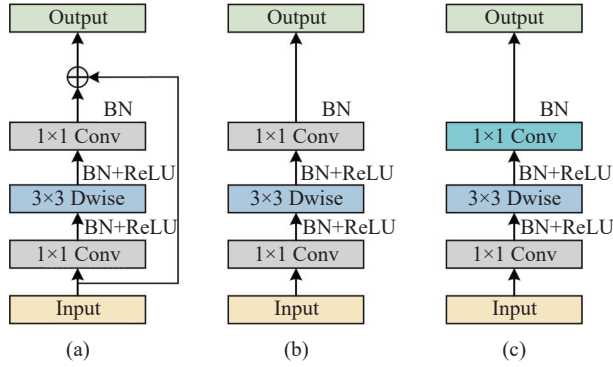


图 4 MobileNetv2 的基本模块. (a) 基本模块, 步长=1; (b) 下采样模块, 步长=2; (c) 使用可变形卷积的下采样模块

Fig.4 Basic modules of the MobileNetv2 block: (a) identity block, $s=1$; (b) down-sampling block, $s=2$; (c) down-sampling block with DCN

能解决接触网零部件尺度差异大和背景复杂的问题. 针对上述问题, 本节对原有的 MobileNet2 进行了以下优化: (1) 在 MobileNet2 的最后一阶段的末尾增加了所提出的 SRPAM 模块; (2) 每个下采样模块中第二个 1×1 Conv 替换为可变形卷积 (DCN), 如图 4(c) 所示.

如图 5(b) 所示, 通过上述方法改进了原有的 MobileNet2 模块, 设计了新的主干网络 (MobileNet-P). MobileNet-P 能有效地增强网络特征提取能力, 缓解小目标漏检问题, 提高背景复杂环境下的适应能力.

此外, 受文献 [41] 的启发, 本文设计了一种可

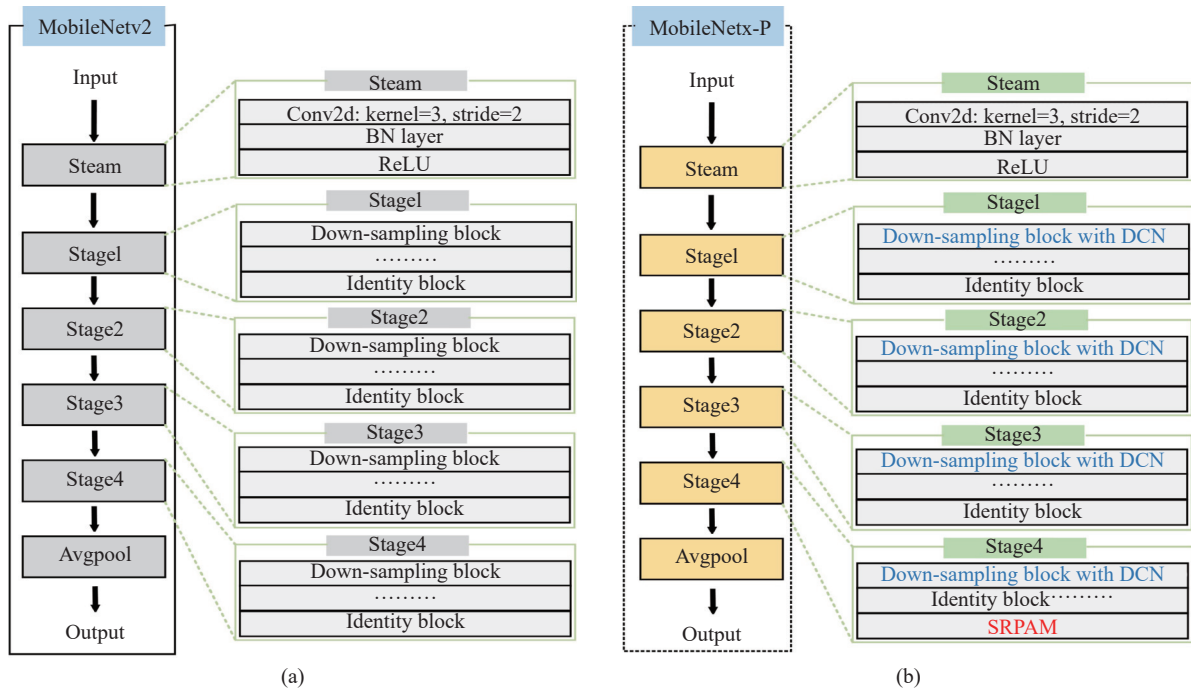


图 5 主干网络的结构图. (a) MobileNetv2; (b) MobileNet-P

Fig.5 Structure diagram of the backbone network: (a) MobileNetv2; (b) MobileNet-P

变形双支路检测头 (FCOSX-Head). 如图 6 所示, 将可变形卷积引入双分支结构中, 模型能够通过学习偏移量自适应地选择对应的空间特征区域. 其中, 第一个分支是分类-中心点分支, 主要回归图像中目标的类型特征图和中心点偏置. 第二个是回归分支, 主要用于回归、修正检测框边界. 进一步, 为了提升检测头对于多尺度信息的捕获, 将两分支的后两层卷积替换为可变形卷积. 综上所述, FCOSX-Head 可以自适应地提取相应的特征信息, 提高检测缺陷的准确性.

2 学习策略

如下所示, FCOS 检测器的总损失函数由分类损失、中心点损失和边界损失组成.

$$L_{FCOS} = L_{cls} + L_{reg} + L_{cen} \quad (6)$$

其中, L_{cls} 、 L_{reg} 和 L_{cen} 分别代表分类损失, 回归损失和中心损失. 本文采用焦点损失 (Focal loss) 作为检测器的分类损失函数, 以减少正样本和负样本之间的不平衡带来的负面影响.

$$L_{cls} = \frac{1}{N_{pos}} \sum_{x,y} L_{focal}(p_{x,y}^s, p_{x,y}^*) \quad (7)$$

其中, L_{focal} 表示焦点损失, N_{pos} 表示正样本的数量, $p_{x,y}^s$ 像素点 (x,y) 的分类分数, $p_{x,y}^*$ 表示像素点 (x,y) 的真值类标签, 中心度损失通过交叉熵损失 (CE) 计算.

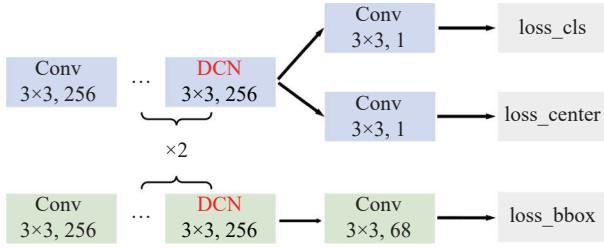


图6 FCOSX-Head的结构图

Fig.6 Structure diagram of FCOSX-Head

$$L_{\text{cen}} = \frac{1}{N_{\text{pos}}} \sum_{x,y} I_{c_{x,y}^* > 0} L_{\text{cen}}(c_{x,y}^s, c_{x,y}^*) \quad (8)$$

其中, L_{cen} 表示交叉熵损失. $c_{x,y}^s$ 表示预测中心度分数, $c_{x,y}^*$ 表示像素点 (x,y) 的中心度真值. 中心度损失仅使用正样本计算. 回归损失通过计算 GIoU 损失. 公式如下.

$$L_{\text{reg}} = \frac{1}{N_{\text{pos}}} \sum_{x,y} I_{c_{x,y}^* > 0} L_{\text{reg}}(t_{x,y}^s, t_{x,y}^*) \quad (9)$$

其中, L_{reg} 表示 GIoU 损失. $t_{x,y}^s$ 表示回归像素点 (x,y) 的结果. $t_{x,y}^*$ 表示回归像素点 (x,y) 的真值. I 表示指标功能, 当 $c_{x,y}^* > 0$, 则为 1, 相反则为 0. 只有正样本才能用于计算边界回归损失.

3 实验与分析

本文使用高速铁路采集的接触网关键零部件数据集, 并对所提的(MFP-FCOS)进行了性能评估. 在装载 Intel(R) Core(TM) i9-13900HX、NVIDIA RTX 4060 GPU、32 G RAM 的 PC 端 Pytorch 深度学习平台上进行实验. 其中, 具体的软件环境为

CUDA11.3、cuDNN8.7.0、Torch1.9、Python3.9 和MM-detection2.28. 在训练细节方面, 实验将非极大抑制值设置为 0.45. 使用随机梯度下降法(SGD)作为优化器(权重衰减和动量分别为 0.0001 和 0.9), 并在训练前的 500 个 Step 中逐渐将学习率从 1×10^{-5} 线性增加到 1×10^{-3} , 而在 14 和 22 个 Epoch 之后依次将学习率降低 10 倍. 此外, 在使用接触网零部件数据集之前, 使用 COCO 2017 数据集对模型进行预训练以提高模型训练效率. 最后, 模型训练的 Batch-size 和 Epoch 分别被设置为 4 和 30.

3.1 数据集

实验数据集图像主要是通过接触网检测车的图像采集系统进行采集. 数据集的总样本数为 5226 张, 分辨率为 6600×4400 . 其中训练数据集 3500 张, 验证数据集 700 张, 测试数据集 1026 张. 如图 7 所示, 本数据集中的关键部件包括以下 12 类: 绝缘子、绝缘子底座、双耳套筒、腕臂支持套筒、顶紧螺栓、斜拉线定位钩、旋转双耳、定位器支座、双套筒连接器、等电位线、承力索底座、防风拉线环.

3.2 评估指标

本文选取准确率(AP)、平均 AP(mAP)、Precision(P)、Recall(R) 等常用检测指标来评估检测模型的整体性能, 这些指标定义如下.

$$P = \frac{TP}{TP + FP} \times 100\% \quad (10)$$

$$P = \frac{TP}{TP + FN} \times 100\% \quad (11)$$

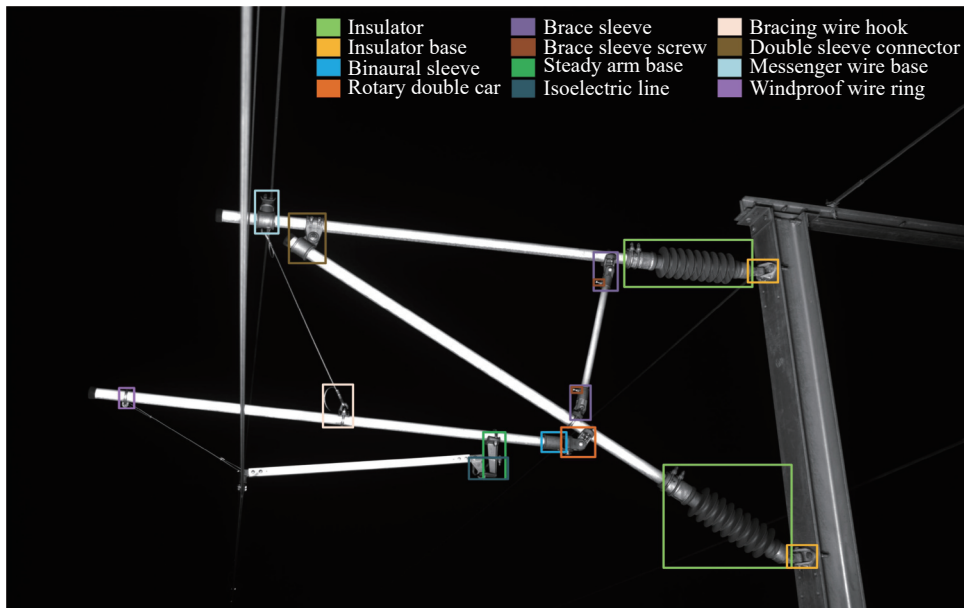


图7 接触网关键部件示意图

Fig.7 Schematic diagram of key catenary components

其中, TP 是正确正样本, FN 是错误负样本, FP 是错误正样本. 此外, AP 表示 $P-R$ 曲线的关于坐标轴的面积, 与检测性能呈正相关.

$$AP = \int_0^1 P(R)dR \quad (12)$$

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (13)$$

其中, P 和 R 分别表示精确率和召回率, mAP 表示每个类别 AP 的平均得分. 此外, 本文还引入了 FPS (Frames per second)、FLOPs (Floating point operations) 和模型参数量 (P_a) 作为模型计算量和计算速度的评价指标.

3.3 模型算法

本文所提出的模型算法步骤如下:

输入: 接触网支撑部件数据集的训练集和测试集.

步骤 1: 输入的 RGB 图像按照以下参数设置进行归一化: (1) 均值为 [123.675, 116.28, 103.53]; (2) 标准差为 [58.395, 57.12, 57.375].

步骤 2: 先将数据集图像的短边随机缩放到 800 ~ 1333, 长边设为短边的 1.5 倍. 然后, 对处理后的图像进行归一化.

步骤 3: 将处理后的图像放入改进后的主干网

络 (MobileNet-P) 中. 然后从主干中提取不同特征层计算多尺度特征 {C2, C3, C4, C5}. 最后, 将多尺度特征引入到 PA-BFPN 中计算融合特征 {P2, P3, P4, P5}, 帮助检测头预测结果.

步骤 4: 将 PA-BFPN 的输出特征输入到 FCO-SX-Head 中, 计算分类分数、中心点偏置分数和定位分数, 用于预测最终的边界框和对象的类别.

步骤 5: 根据检测头得到的分类分数和回归分数计算分类损失 (Focal loss) 和回归损失 (GIoU loss).

步骤 6: 通过最小化损失函数来训练 MFP-FCOS, 然后保存最优模型参数.

3.4 对比实验分析

为了验证所提方法 (MFP-FCOS) 的性能, 本文对比了一些经典的目标检测算法. 结果如表 1 所示. 实验配置条件如下: (1) 表中所有实验的预测结果为 5 次随机训练和测试的平均值; (2) 随机训练不同尺寸大小的图像, 使用 800×800 的随机图像进行测试. 进行了以下对比实验: (1) 对于一阶检测器, 本文以 YOLOv3 和 YOLOX 为基线; (2) 对于二阶检测器, 以 Faster R-CNN 和 Cascade R-CNN 为基线; (3) 对于无锚检测器, 将引入 Center-Net、CornerNet 和 ASTT 对比试验.

如表 1 所示, 得出以下结论: (1) 除了 YOLO 系列检测器之外, 在检测方法相同的条件下, 检测

表 1 经典检测方法对比实验分析

Table 1 Comparative experiments for classic detection methods

Methods	Backbone	mAP	AP ₅₀	AP ₇₅	P_a/M	FLOPs	FPS
YOLOv3	Darknet53	46.23	90.56	45.38	62.4	122.2	32
YOLOX-m	CSP-Darknet53	48.21	91.23	47.16	25.43	58.7	34
Faster-RCNN	ResNet50	45.98	90.32	45.68	41.24	122.4	25
	ResNet101	47.16	91.34	46.54	60.34	171.2	15
Cascade-RCNN	ResNet50	46.14	90.96	46.23	69.54	125.1	21
	ResNet101	47.86	91.84	47.01	89.12	173.2	13
FCOS	ResNet50	46.23	90.87	46.98	33.23	118.3	28
	ResNet101	47.51	91.52	47.12	51.34	161.8	19
CornerNet	ResNet50	45.67	90.08	45.45	31.34	319.7	11
	ResNet101	46.71	90.67	46.19	50.14	367.1	6
ATSS	ResNet50	46.35	90.78	45.29	31.91	126.28	22
	ResNet101	47.24	91.46	46.95	50.91	173.83	16
MFP-FCOS	ResNet50	47.98	91.12	46.47	37.34	123.8	23
	ResNet101	48.43	92.56	47.24	55.45	165.4	12
	MobileNetv2	47.35	90.41	45.67	12.21	35.1	32
	MobileNet-P	48.62	92.76	46.82	14.43	38.6	30

Note: AP₅₀ represents the average accuracy at IoU=0.5; AP₇₅ represents the average accuracy at IoU=0.75.

器模型越大越深, 检测准确性越高; (2) 当本文的改进模块分别加载到不同的主干网络中时, MFP-FCOS 检测网络实现了最佳性能, 具体表现在 mAP、AP₅₀、AP₇₅、P_a 和 FLOPs 中, 其最优的性能指标值为 mAP=48.62%、AP₅₀=92.76%、AP₇₅=47.24%、P_a=12.21 和 FLOPs=34; (3) YOLOX-m 达到了最佳的 FPS 指数 (FPS=34), 在 mAP 指数上优于 Faster RCNN、Cascade RCNN、FCOS、CornerNet、ATSS 等检测方法约 0.45%~2.5%, 但仍比 MobileNet-P 低约 1.62%; (4) 当 MFP-FCOS 和 ATSS 加载同一主干网络时, MFP-FCOS 的 mAP、AP₅₀ 和 AP₇₅ 指标比 ATSS 高 0.4%~1.6%, 证明 MFP-FCOS 的有效性; (5) 当 MFP-FCOS 加载所提出的主干网络 MobileNet-P 中时, 模型的检测性能指标比原始的 MobileNetv2 骨干网络提高约 1.15%~2.8%; (6) 在 FPS 指数上, 本文提出的 MobileNet-P 的每秒传输帧数 (FPS) 达到 30 帧, 仅次于最优参数 FPS=34, 但仍优于大多数目标检测算法 (Faster-RCNN、Cascade-RCNN、FCOS 和 ATSS) 40% 左右. 表明 MobileNet-P 仍有高效的检测速率. 总之, 本文提出的 MFP-FCOS 网络在检测速度上比 YOLO 算法略低, 但其综合性能良好, 性能优于其他经典算法.

3.5 实验结果可视化分析

本节对提出的方法进行了可视化实验, 直观地验证模型的检测效果. 其中, 图 8(a)~(c) 是在室外夜晚拍摄的结果, 图 8(d)~(f) 表示夜晚隧道环境下的检测结果图, 图 8(g)~(h) 表示复杂环境下的检测结果图.

从检测结果可以看出, MFP-FCOS 网络不存在漏检和重复框的问题. 对于背景复杂的隧道环境下, 仍能准确检测不同尺度的零部件. 其次, 网络可以精确定位小尺度部件 (腕臂支持套筒和等电位线) 和小部件与大部件之间的中尺度目标 (绝缘子底座). 由此可见, MFP-FCOS 网络能够很好的检测各种尺寸的接触网支撑零部件. 这直观地表明本文所提出的方法可以解决小目标, 多尺度样本问题.

3.6 对比实验结果可视化分析

本节将 YOLOX-m、CornerNet 和 ATSS 与本文提出的 MFP-FCOS 进行可视化对比实验分析, 验证本文所提出的 MFP-FCOS 的有效性.

如图 9 所示, 本文采用四种不同的目标检测算法的可视化结果图, 论证本文所提方法的可行性. 字母 (a)~(d) 分别代表 YOLOX-m、CornerNet、ATSS、MFP-FCOS 检测器对同一张图像的检测效果. 本节

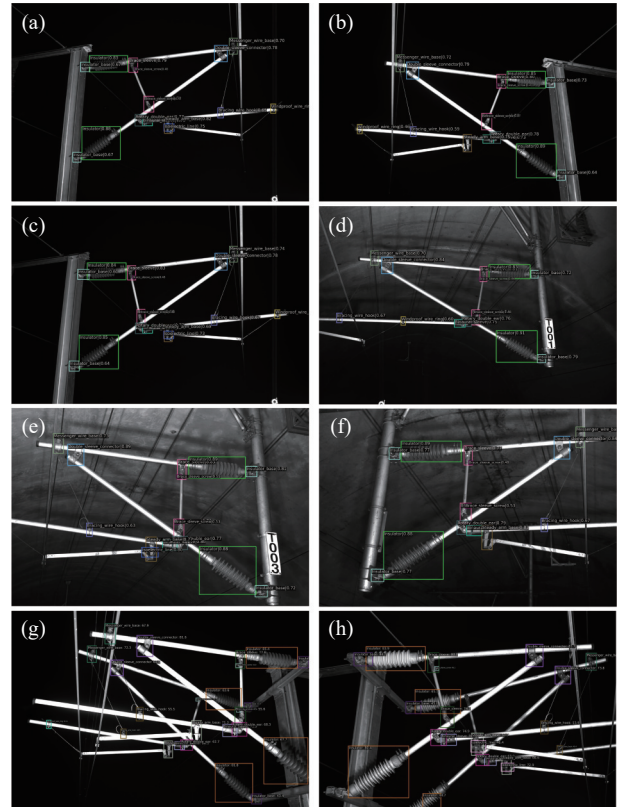


图 8 检测结果可视化图

Fig.8 Visualization of detection results

通过红色箭头表示目标组件的遗漏区域. 通过观察图 9, 我们可以得出以下结论: (1) YOLOX-m 和 MFP-FCOS 检测器可以有效检测出所有接触网支撑组件; (2) 对于 CornerNet 检测器, 图像中有 4 个缺失组件和 8 个重复框. 4 个缺失区域包括 2 个顶紧螺栓、1 个等电位线和 1 个绝缘子; (3) 对于 ATSS 检测器, 缺失区域包括 2 个顶紧螺栓; (4) 通过对比目标交并比 (IoU) 大小, 得出 MFP-FCOS 整体性能优于 YOLOX-m 检测器. 综上所述, 许多检测方法很难检测到尺寸较小的部件 (顶紧螺栓) 和尺寸较大的部件 (绝缘子), 尤其是尺寸差异大的目标. 但是, 本文提出的 MFP-FCOS 检测器在检测不同尺度目标方面表现非常好. 这直观地表明本文提出的方法可以有效解决小目标和多尺度样本检测问题.

3.7 多尺度特征融合模块分析

本节将 SPP、ASPP、Dense-ASPP、Inception 等经典方法与本文改进的模块 SRPAM 进行对比实验, 验证本文所提出的多尺度特征融合模块 (SRPAM) 的有效性和优越性, 实验结果如表 2 所示.

从表 2 的结果可以得出, 与其他经典方法相比, SRPAM 模块的综合检测性能是最优的, 最优参数为 mAP=48.62%、AP₅₀=92.76% 和 AP_S=8.89%.

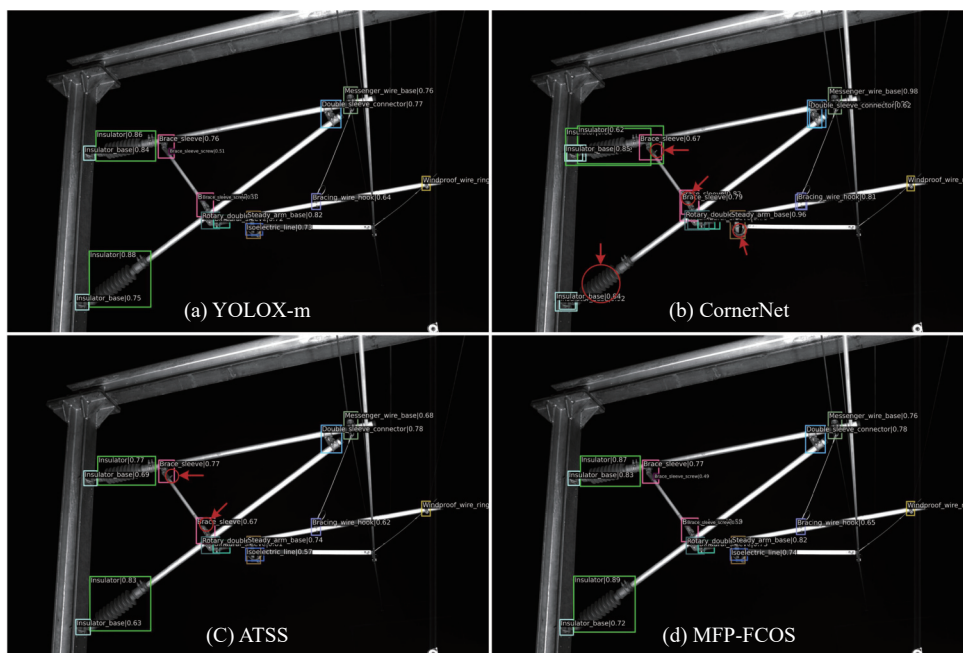


图 9 对比试验检测结果

Fig.9 Visualization of comparative test results

表 2 多尺度特征融合模块实验分析

Table 2 Analysis of the multi-scale feature fusion module

Methods	mAP	AP ₅₀	AP ₇₅	AP _S	AP _m	AP _L
SPP	45.51	88.63	43.07	5.23	44.53	49.27
ASPP	46.65	89.43	44.17	6.43	45.86	48.8
Dense-ASPP	47.32	89.78	44.86	8.16	47.41	50.28
Inception	45.77	88.91	43.54	6.94	44.91	48.00
SRPAM	48.62	92.76	46.82	8.89	46.17	49.43

Note: AP_S represents the performance of the model for small-sized objects; AP_m represents the performance of the model for medium-sized objects; AP_L represents the performance of the model for large-scale objects.

此外, 其他模块的综合性能排名为 Dense-ASPP>ASPP>Inception>SPP. 本文所提出的多尺度模块性能优越的原因如下: (1)SPP 模块是多尺度池化操作的集合, 检测效果与 ASPP 模块相比, 其存在严重的特征信息丢失问题; (2)与 Dense-ASPP 相比, ASPP 仅通过特征联合的方法融合多尺度特征, 忽视了不同分支间的交互信息. 综上所述, ASPP 模块中引入深度可分离卷积、残差连接和密集连接可以有效地提高其特征提取性能.

3.8 特征金字塔网络分析 (FPN)

在本节中, 对比了经典的特征金字塔网络 (例如 FPN、PA-FPN、NAS-FPN 和 Bi-FPN), 以验证所提出的 PA-BFPN 网络的有效性, 表 3 展示相关实验结果.

如表 3 所示, 得到以下结论: (1)本文所提的 PA-BFPN 综合性能最好, 其最优参数分别为 mAP=

48.62%、AP₅₀=92.76%、AP₇₅=46.82%、AP_S=8.89%、AP_m=46.17%; (2)PA-FPN 综合检测性能达到次优, 在 AP_L 上达到最优值为 50.06%; (3)PA-FPN 和 Bi-FPN 检测性能相近, 各指标均比 FPN 高 0.5 ~ 1.5% 左右. 其原因如下: (1)相比于 FPN、NAS-FPN 和 Bi-FPN 模块, PA-FPN 和 PA-BFPN 通过增加自底向上路径, 用底层中的精确信号增强整个特征金字塔; (2)与 FPN 相比, PA-FPN 和 PA-BFPN 可以通过平衡特征金字塔模块获得跨层特征信息和多尺度信息; (3)与 FPN 相比, PA-FPN 和 Bi-FPN 包括“自上向下”和“自下向上”的特征融合过程, 因此可以获得更全面的特征信息. 由此可见, PA-BFPN 具有一定的优势和设计的合理性. 如表 2 和表 3 所示, SRPAM 模块和 PA-BFPN 模块在悬链线部件检测任务的数值分析中表现良好. 为了更加直观地展示两个模块在网络前向过程中对语义信

息的影响,对主干层和颈部层的层输出特征进行了热力图可视化分析。

如图 10 所示,热力图可视化结果可以得出, MFP-FCOS 网络首先对大尺度目标进行关注,其次随着网路的加深,会额外激活小尺度目标区域,并关注接触网关键零部件区域。这进一步表明 MFP-FCOS 对不同尺度目标的有效性,证明了 MFP-FCOS 设计的合理性。

3.9 消融实验分析

为了进一步验证本文所提出方法的有效性和合理性,本节对 SRPAM、DCN、PA-BFPN、FCOSX-Head 模块进行了消融实验分析,结果如表 4 所示。(-)DCN 表示可变形卷积模块的消融,替换后的网

络精度提升了 2.25%; (-)SRPAM 表示 SRPAM 模块的消融; (-)PA-BFPN 表示 PA-BFPN 模块的消融; (-)FCOSX-Head 表示 FCOSX-Head 的消融; (-)表示不进行消融实验。从结果可以得出,各个模块的消融都会对模型带来不同程度的负面影响。通过替换 DCN、SRPAM、PA-BFPN 和 FCOSX-Head 模块,在 AP_{50} 指标方面分别提高 2.25%、3.05%、3.23% 和 2.78%。(-)表示加入四种改进策略,对于五个指标均有不同程度的提升,在 mAP、 AP_{50} 、 AP_{75} 、 AP_S 和 AP_I 分别最高提升 3.07%、3.23%、2.96%、2.95% 和 1.72%。由此可见,消融对模型的影响顺序依次为 PA-BFPN、SRPAM、FCOSX-Head、DCN。进一步验证了所提模型的有效性。

表 3 特征金字塔网络实验分析

Table 3 Analysis of feature pyramid networks

Methods	mAP	AP_{50}	AP_{75}	AP_S	AP_m	AP_I
FPN	45.55	89.48	43.86	5.94	44.47	47.7
NAS-FPN	46.95	90.01	44.4	6.47	45.01	48.85
Bi-FPN	47.05	89.89	44.75	5.83	45.41	48.68
PA-FPN	48.28	90.51	45.65	6.96	46.11	50.06
PA-BFPN	48.62	92.76	46.82	8.89	46.17	49.43

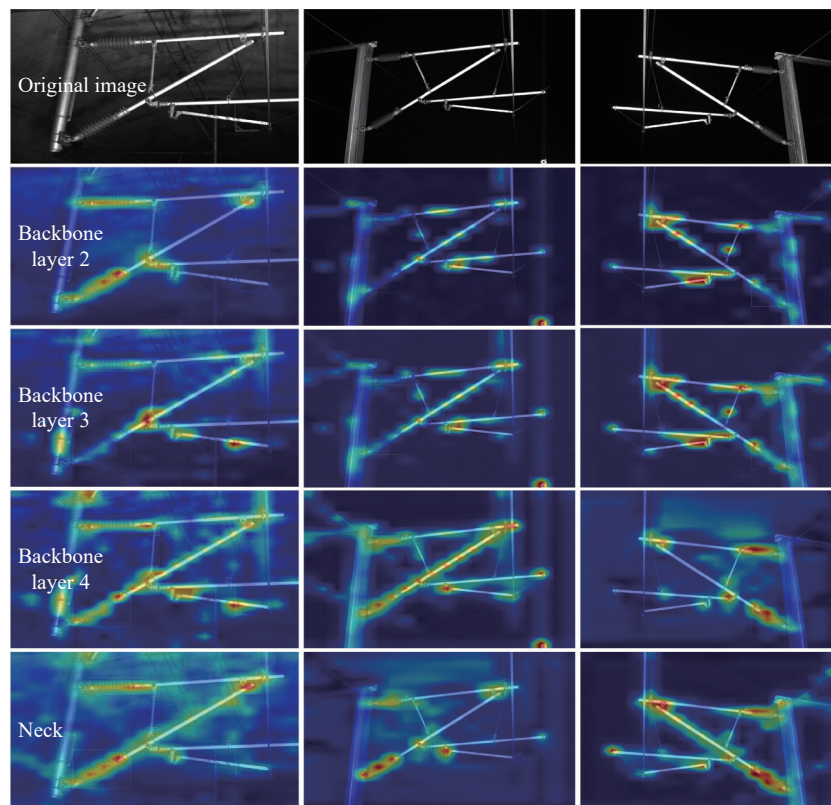


图 10 热力图可视化

Fig.10 Visualization of heat maps

表 4 消融实验分析

Table 4 Experiments analysis of model ablation

Ablation methods	mAP	AP ₅₀	AP ₇₅	AP _S	AP _m	AP _l	Latency/ms
(-) DCN	48.17	90.51	45.01	8.41	45.83	48.66	28.66
(-) SRPAM	46.51	89.71	44.57	6.92	45.2	48.16	25.33
(-) PA-BFPN	45.55	89.53	43.86	5.94	44.47	47.71	26.03
(-) FCOSX-Head	47.55	89.98	45.07	6.54	47.41	48.24	29.71
(-)	48.62	92.76	46.82	8.89	46.17	49.43	31.25

Note: Latency represents the test time of an image.

4 结论

针对接触网零部件多尺度问题和复杂背景问题, 本文提出了一种 MFP-FCOS 检测器. 该方法包括以下贡献: (1) 通过在 ASPP 中引入密集连接和深度残差模块, 增强主干网络特征提取能力; (2) 在 MobileNet-P 中引入可变形卷积和 SRPAM 模块, 优化其特征提取能力; (3) 在 PA-BFPN 中引入跨层特征平衡模块, 构建新的 PA-BFPN 模块, 有效地融合跨层特征信息, 提高小尺度目标检测性能; (4) 检测头引入可变形卷积, 优化分类和定位精度. 最后通过对比试验和可视化分析得出以下结论: (1) 与经典检测器相比, 例如 Faster-RCNN、YOLOX-m、FCOS 和 ATSS, MFP-FCOS 网络对小尺度零部件有更好的检测能力, 其综合性能更优秀; (2) 通过对比一些经典的多尺度融合模块和特征金字塔网络, SRPAM 和 PA-BFPN 在接触网检测性能上具有显著优势. 综上所述, MFP-FCOS 网络设计合理, 具有一定的优势. 在未来研究中, 将结合小样本学习和无监督检测方法解决接触网零部件检测问题.

参 考 文 献

- [1] Song Y, Liu Z G, Duan F C, et al. Study on wind-induced vibration behavior of railway catenary in spatial stochastic wind field based on nonlinear finite element procedure. *J Vib Acoust*, 2018, 140(1): 011010
- [2] Han Y, Liu Z G, Lyu Y, et al. Deep learning-based visual ensemble method for high-speed railway catenary clevis fracture detection. *Neurocomputing*, 2020, 396: 556
- [3] Gao S B, Liu Z G, Yu L. Detection and monitoring system of the pantograph-catenary in high-speed railway (6C) // 2017 7th International Conference on Power Electronics Systems and Applications—Smart Mobility, Power Transfer & Security (PESA). Hong Kong, 2017: 1
- [4] Song Y, Liu Z G, Ouyang H J, et al. Sliding mode control with PD sliding surface for high-speed railway pantograph-catenary contact force under strong stochastic wind field. *Shock Vib*, 2017, 2017(1): 4895321
- [5] Liu W Q, Liu Z G, Wang H, et al. An automated defect detection approach for catenary rod-insulator textured surfaces using unsupervised learning. *IEEE Trans Instrum Meas*, 2020, 69(10): 8411
- [6] Song Y, Liu Z G, Rönquist A, et al. Contact wire irregularity stochastics and effect on high-speed railway pantograph-catenary interactions. *IEEE Trans Instrum Meas*, 2020, 69(10): 8196
- [7] Liu Z G, Xiang C, Wang Y Q, et al. A model-based predictive direct power control for traction line-side converter in high-speed railway. *IEEE Trans Ind Appl*, 2017, 53(5): 4934
- [8] Song Y, Wang Z W, Liu Z G, et al. A spatial coupling model to study dynamic performance of pantograph-catenary with vehicle-track excitation. *Mech Syst Signal Process*, 2021, 151: 107336
- [9] Duan F C, Song Y, Gao S B, et al. Study on aerodynamic instability and galloping response of rail overhead contact line based on wind tunnel tests. *IEEE Trans Veh Technol*, 2023, 72(6): 7211
- [10] Song Y, Rönquist A, Jiang T J, et al. Railway pantograph-catenary interaction performance in an overlap section: Modelling, validation and analysis. *J Sound Vib*, 2023, 548: 117506
- [11] Hou G Y, Chen Q H, Yang Z H, et al. Helmet detection method based on improved YOLOv5. *Chin J Eng*, 2024, 46(2): 329 (侯公羽, 陈钦煌, 杨振华, 等. 基于改进 YOLOv5 的安全帽检测算法. 工程科学学报, 2024, 46(2): 329)
- [12] Xue S, Wang Y B, Lü Q Y, et al. Anti-occlusion target detection algorithm for anti-UAV system based on YOLOX-drone. *Chin J Eng*, 2023, 45(9): 1539 (薛珊, 王亚博, 吕琼莹, 等. 基于 YOLOX-drone 的反无人机系统抗遮挡目标检测算法. 工程科学学报, 2023, 45(9): 1539)
- [13] Ye Han, Liu Z, Han Z, et al. Foreign body detection between insulator pieces in electrified railway based on affine moment invariant. *J Chi Rail Soc*, 2013, 35(4): 30
- [14] Han Y, Liu Z G, Han Z W, et al. Fracture detection of ear pieces of catenary support devices of high-speed railway based on SIFT feature matching. *J China Railw Soc*, 2014, 36(2): 31 (韩烨, 刘志刚, 韩志伟, 等. 基于 SIFT 特征匹配的高速铁路接触网支撑装置耳片断裂检测研究. 铁道学报, 2014, 36(2): 31)

- [15] Yang H M, Liu Z G, Han Y, et al. Defective condition detection of insulators in electrified railway based on feature matching of speeded-up robust features. *Power Syst Technol*, 2013, 37(8): 2297
(杨红梅, 刘志刚, 韩焯, 等. 基于快速鲁棒性特征匹配的电气化铁路绝缘子不良状态检测. *电网技术*, 2013, 37(8): 2297)
- [16] Zhang G N, Liu Z G, Han Y, et al. Loss fault detection for auxiliary catenary wire of high-speed railway catenary wire holder. *J China Railw Soc*, 2017, 39(5): 40
(张桂南, 刘志刚, 韩焯, 等. 高速铁路接触网承力索座识别与辅助承力索缺失故障检测方法研究. *铁道学报*, 2017, 39(5): 40)
- [17] Yang L, Zhong J H, Zhang Y, et al. An improving faster-RCNN with multi-attention ResNet for small target detection in intelligent autonomous transport with 6G. *IEEE Trans Intell Transp Syst*, 2023, 24(7): 7717
- [18] Cai Z W, Vasconcelos N. Cascade R-CNN: High quality object detection and instance segmentation. *IEEE Trans Pattern Anal Mach Intell*, 2021, 43(5): 1483
- [19] Liu Z G, Wang L Y, Li C J, et al. A high-precision loose strands diagnosis approach for isoelectric line in high-speed railway. *IEEE Trans Ind Inform*, 2018, 14(3): 1067
- [20] Zhong J P, Liu Z G, Han Z W, et al. A CNN-based defect inspection method for catenary split pins in high-speed railway. *IEEE Trans Instrum Meas*, 2019, 68(8): 2849
- [21] Kang G Q, Gao S B, Yu L, et al. Deep architecture for high-speed railway insulator surface defect detection: Denoising autoencoder with multitask learning. *IEEE Trans Instrum Meas*, 2019, 68(8): 2679
- [22] Chen J W, Liu Z G, Wang H R, et al. Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network. *IEEE Trans Instrum Meas*, 2018, 67(2): 257
- [23] Qin R B, Chen N J, Huang Y H. EDDNet: An efficient and accurate defect detection network for the industrial edge environment // 2022 *IEEE 22nd International Conference on Software Quality, Reliability and Security (QRS)*. Guangzhou, 2022: 854
- [24] Luo S Y, Jia B, Zhao J, et al. Surface defect detection method for steel materials based on improved YOLOv7 // 2023 *8th International Conference on Image, Vision and Computing (ICIVC)*. Dalian, 2023: 194
- [25] Liu L, Feng X F, Li F, et al. Surface defect detection of industrial components based on improved YOLOv5s. *IEEE Sens J*, 2024, 24(15): 23940
- [26] Ren F, Fei J J, Li H S, et al. Steel surface defect detection using improved deep learning algorithm: ECA-SimSPPF-SIoU-Yolov5. *IEEE Access*, 2024, 12: 32545
- [27] Yan X C, Zhang Y, Jin Q B. Chemical process fault diagnosis based on improved ResNet fusing CBAM and SPP. *IEEE Access*, 2023, 11: 46678
- [28] Fu D Y, Zeng X F, Han S C, et al. Nested densely atrous spatial pyramid pooling and deep dense short connection for skeleton detection. *IEEE Trans Hum Mach Syst*, 2023, 53(1): 75
- [29] Yang M K, Yu K, Zhang C, et al. DenseASPP for semantic segmentation in street scenes // 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 3684
- [30] Chen J D, Chen W R, Zeb A, et al. Lightweight inception networks for the recognition and detection of rice plant diseases. *IEEE Sens J*, 2022, 22(14): 14628
- [31] Chen L, An S M, Zhao S, et al. MS-FPN-based pavement defect identification algorithm. *IEEE Access*, 2023, 11: 124797
- [32] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation // 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 8759
- [33] Syazwany N S, Nam J H, Lee S C. MM-BiFPN: Multi-modality fusion network with Bi-FPN for MRI brain tumor segmentation. *IEEE Access*, 2021, 9: 160708
- [34] Ghiasi G, Lin T Y, Le Q V. NAS-FPN: Learning scalable feature pyramid architecture for object detection // 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, 2019: 7029
- [35] Dai J F, Qi H Z, Xiong Y W, et al. Deformable convolutional networks // 2017 *IEEE International Conference on Computer Vision (ICCV)*. Venice, 2017: 764
- [36] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks // 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, 2017: 2261
- [37] Hu J, Shen L, Sun G. Squeeze-and-excitation networks // 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 7132
- [38] Yang H N, He J, Liu Z G, et al. LLD-MFCOS: A multiscale anchor-free detector based on label localization distillation for wheelset tread defect detection. *IEEE Trans Instrum Meas*, 2024, 73: 5003815
- [39] Wang X L, Girshick R, Gupta A, et al. Non-local Neural Networks // 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 7794
- [40] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks // 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, 2018: 4510
- [41] Chen Z H, Yang C, Li Q F, et al. Disentangle Your Dense Object Detector // *Proceedings of the 29th ACM International Conference on Multimedia*. Seattle, 2021: 4939