

基于双分支串行混合注意力的输电线路缺陷检测深度神经网络模型

郝然, 王红军, 李天瑞

引用本文

郝然, 王红军, 李天瑞. 基于双分支串行混合注意力的输电线路缺陷检测深度神经网络模型[J]. 计算机科学, 2024, 51(3): 135-140.

HAO Ran, WANG Hongjun, LI Tianrui. [Deep Neural Network Model for Transmission Line Defect Detection Based on Dual-branch Sequential Mixed Attention](#) [J]. Computer Science, 2024, 51(3): 135-140.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

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

[TMGAT:类型匹配约束的图注意力网络](#)

TMGAT:Graph Attention Network with Type Matching Constraint

计算机科学, 2024, 51(3): 235-243. <https://doi.org/10.11896/jsjcx.221200097>

[基于标签信息融合与多任务学习的中文命名实体识别](#)

Chinese Named Entity Recognition Based on Label Information Fusion and Multi-task Learning

计算机科学, 2024, 51(3): 198-204. <https://doi.org/10.11896/jsjcx.230200114>

[基于级联U-Net的遥感影像道路分割和轮廓提取方法](#)

Combined Road Segmentation and Contour Extraction for Remote Sensing Images Based on Cascaded U-Net

计算机科学, 2024, 51(3): 174-182. <https://doi.org/10.11896/jsjcx.221200032>

[多尺度特征融合的遥感图像目标检测方法](#)

Object Detection Method with Multi-scale Feature Fusion for Remote Sensing Images

计算机科学, 2024, 51(3): 165-173. <https://doi.org/10.11896/jsjcx.230200030>

[基于多空间属性信息融合的顺序推荐](#)

Sequential Recommendation Based on Multi-space Attribute Information Fusion

计算机科学, 2024, 51(3): 102-108. <https://doi.org/10.11896/jsjcx.230600078>

基于双分支串行混合注意力的输电线路缺陷检测深度神经网络模型

郝 然 王红军 李天瑞

西南交通大学计算机与人工智能学院 成都 611756

可持续城市交通智能化教育部工程研究中心 成都 611756

(986241422@qq.com)

摘 要 检测输电线路缺陷并及时维修可以确保电网的安全稳定,具有重大的实际意义。但输电线路图像背景复杂、元件尺寸小,导致现有的目标检测模型不能取得很好的效果,因此文中提出了基于双分支串行混合注意力的输电线路缺陷检测深度神经网络模型。该模型设计了 DBSA(Dual-branch Serial Attention)双分支串行混合注意力,从而将更多的权重放在缺陷上,并提出了 WCFPN(Well-connected Feature Pyramid Network)特征金字塔,让经 DBSA 提取的特征充分融合,从而增强模型检测小目标的能力。DBSA 将特征图沿高度和宽度两个分支压缩并用一维卷积提取注意力,WCFPN 设计了一种包含跨尺度融合和跳层连接的新型融合路径,让经 DBSA 提取的高层语义信息和低层空间信息进行更充分的交互。最后在绝缘子自爆、防振锤损坏、鸟巢异物、水泥杆破损和输电线路缺陷 5 个数据集上进行实验,结果显示所提模型取得了最佳的检测效果,在 5 个数据集上的平均 AP₅₀ 和 AP 分别为 84.3% 和 46.1%,相比目前最先进的模型 YOLOv7 分别提升了 3.7% 和 3%。

关键词: 电网缺陷;目标检测;注意力机制;特征金字塔

中图分类号 TP391

Deep Neural Network Model for Transmission Line Defect Detection Based on Dual-branch Sequential Mixed Attention

HAO Ran, WANG Hongjun and LI Tianrui

School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China

Engineering Research Center of Sustainable Urban Intelligent Transportation, Ministry of Education, Chengdu 611756, China

Abstract Detecting defects in transmission lines and repairing them in a timely manner is of great practical significance for ensuring the safety and stability of the power grid. However, due to the complex background and small component size of transmission line images, existing object detection models cannot achieve satisfactory results. Therefore, this paper proposes a deep neural network model for detecting defects in transmission lines based on dual-branch serial attention. The model designs dual-branch serial attention(DBSA) to allow the model to focus more weight on the defects, and proposes well-connected feature pyramid network(WCFPN) to enable full fusion of the features extracted by DBSA, thereby improving the model's ability to detect small targets. DBSA compresses the feature map along the height and width branches and extracts attention using one-dimensional convolution to achieve fine-grained control over the features. WCFPN designs a new fusion path that includes cross-scale fusion and skip-layer connections, allowing high-level semantic information and low-level spatial information extracted by DBSA to interact more fully. Finally, experiments are conducted on five transmission line datasets, including insulator explosion, damaged anti-vibration hammer, bird's nest debris, broken cementpole and transmission line defect, and the proposed model achieves the best detection performance. The average AP₅₀ and AP of the five datasets is 84.3% and 46.1%, respectively, which is 3.7% and 3% higher than that of the state-of-the-art model YOLOv7.

Keywords Power defect, Object detection, Attention mechanism, Feature pyramid network

1 引言

随着我国电力需求的不断增长,我国输电线路的总长度也在迅速增加。输电线路作为电力系统重要的组成部分,其

状况关系到电力系统的安全与稳定。因此,定期检查输电线路、及时发现输电线路的缺陷以防止大面积停电具有重大的研究意义。

在电力系统缺陷检测中,无人机巡检因具有便捷的优点

到稿日期:2023-06-13 返修日期:2023-10-04

基金项目:国家自然科学基金(62176221,62276216)

This work was supported by the National Natural Science Foundation of China(62176221,62276216).

通信作者:王红军(wanghongjun@swjtu.edu.cn)

而被广泛使用。但目前无人机采集的图片仍依赖人工肉眼检查,这种方式效率低且容易造成漏检、误检的问题。近年来,深度学习与计算机视觉的快速发展,使得基于深度学习的目标检测、语义分割等算法在医疗、交通领域得到了广泛应用,将这些目标检测算法用于输电线路缺陷检测可以让电网巡检变得更高效率和智能^[1]。Zhang 等^[2]提出了基于密集 FPN 结构的 YOLOv3 算法,在提高信息利用率的同时减少了模型参数量,提高了绝缘子缺陷的检测精度。Lv 等^[3]研究了聚类算法对检测结果的影响,提出了一种基于 YOLOv4 算法的电气设备智能检测方法。Liu 等^[4]将 RepVGG 的参数重构思想引入 YOLOv5,提升了电力线关键元件的检测速度和精度。

基于深度学习的目标检测算法主要分为两类:一类是基于回归的单阶段算法,如 SSD^[5], YOLO^[6-7] 系列算法等;另一类是基于区域建议网络(RPN)的二阶段算法,如 Faster-RCNN^[8], Mask-RCNN^[9] 等。本文主要研究单阶段目标检测算法,将其应用于输电线路领域以准确高效地检测输电线路中的缺陷。相较于一般的目标检测场景,输电线路缺陷检测有以下难点:(1)缺陷尺寸小。虽然无人机航拍的图片清晰度高,分辨率高,但异物、绝缘子、防振锤等尺寸较小,在整张图片中只占很小的比例,有些缺陷肉眼难以分辨。(2)背景复杂。输电线路遍布全国各地,航拍的巡检图像背景各异,有山川、河流、树林和雪地等,大量的无关背景给检测带来了很大干扰。

现有的先进目标检测算法如 YOLOv5^[10], YOLOv7^[11] 等不能很好地克服小目标和复杂背景这两个难点。为了提高高分辨率巡检图片中复杂背景下的缺陷检测精度,本文提出了基于双分支串行混合注意力的输电线路缺陷检测深度神经网络模型。本文的主要贡献如下:

- (1)提出了串行双分支混合注意力 DBSA,让模型将更多的权重放到缺陷上,以减少无关背景信息对检测的干扰。
- (2)提出了特征金字塔 WCFPN 来增强 DBSA,让经 DBSA 提取的特征得到更充分的融合,提高模型检测小目标的能力。
- (3)在绝缘子自爆、防振锤损坏、鸟巢异物、水泥杆破损和输电线路缺陷 5 个数据集上进行实验,验证了所提模型的性能。

2 相关工作

本文在所提模型中引入了 DBSA 注意力机制和 WCFPN 特征金字塔,下面介绍注意力机制和特征金字塔的相关知识。

2.1 注意力机制

注意力机制^[12]是一种增强卷积网络特征提取能力的方法,它让模型聚焦于重要的目标区域,并将更多的权重投入到该部分,从而更准确地捕捉输入中的重要信息。

对输入的特征图 $U \in R^{H \times W \times C}$, SENet^[13] 对每个通道上的空间信息 $H \times W$ 进行全局平均池化操作,生成通道维度上的一个向量 $z \in R^{1 \times 1 \times C}$,经过两个全连接层的 ω_1, ω_2 和激活函数 δ, g 生成通道注意力权重 s ,如式(1)所示。再对原始的特征图 U 每个通道上的 $H \times W$ 个值乘上对应通道权重 s ,得到权重加持后的特征图,完成特征图校准,如式(2)所示。

$$s = F_{ex}(z, \omega) = \sigma(g(z, \omega)) = \sigma(\omega_2 \delta(\omega_1 z)) \quad (1)$$

$$X = F_{scale}(U, s) = s \cdot U \quad (2)$$

CBAM^[14]在 SE 通道注意力基础上加入空间注意力,如式(3)所示,对输入 $U \in R^{H \times W \times C}$ 沿通道维度进行平均池化和最大池化操作,得到两个 $z \in R^{H \times W \times 1}$ 的空间特征,在通道维度上进行拼接和卷积操作后,生成空间注意力。

$$\begin{aligned} M_s(F) &= \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \\ &= \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \end{aligned} \quad (3)$$

ECA^[15]使用自适应一维卷积替代 SE 的全连接层,在产生更高效的通道注意力的同时减少了模型的参数量。

2.2 特征金字塔

特征金字塔(Feature Pyramid Network, FPN)^[16]从主干网络中提取不同分辨率的特征图,用多尺度特征融合的方式,如图 1(a)所示,通过自顶向下的路径和横向连接将低分辨率的空间信息和高分辨率的语义信息融合,聚合不同分辨率下的特征,有效地提升了目标检测的检测精度。

FPN 这一开创性工作证明了多尺度特征融合在目标检测任务中的有效性。PANet^[17]指出 FPN 只有一条自顶向下的路径存在单向路径的信息局限性,因此增加了一个额外的自底向上的路径聚合网络,如图 1(b)所示,保证了信息的可靠传递。BiFPN^[18]引入可学习的权重来学习不同输入特征的重要性。

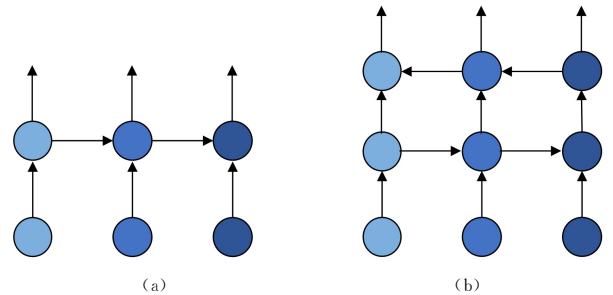


图 1 不同结构的特征金字塔

Fig. 1 Feature pyramids with different structures

3 本文模型

本章首先在 3.1 和 3.2 小节提出了 DBSA 双分支串行混合注意力和 WCFPN 特征金字塔,并在 3.3 小节提出了基于双分支串行混合注意力的输电线路缺陷检测深度神经网络模型。

3.1 DBSA 双分支串行混合注意力

无人机拍摄的巡检图片有复杂的背景信息,在模型网络中添加注意力机制能让缺陷得到更多的关注,减少背景等冗余特征对检测的干扰,从而增强网络的表征能力并提升检测精度。通道注意力关注特征的内容是什么,而空间注意力更关注特征的位置在何处,两者是互补关系,因此混合注意力往往会取得更好的效果。为此,本文提出了 DBSA 双分支串行混合注意力,其结构如图 2 所示。

空间特征图包含了特征的空间分布信息,提取空间信息可以更准确地定位到缺陷的位置。DBSA 对输入 F 沿高度和宽度两个分支分别进行全局平均池化压缩,生成 $1 \times W \times C$ 的

宽度特征 f_1 和 $H \times 1 \times C$ 的高度特征 f_2 , 如式(4)和式(5)所示:

$$f_1(w_j, c_z) = \frac{1}{H} \cdot \sum_{i=1}^H x(h_i, w_j, c_z), j=0, 1, 2, \dots, W; \\ z=0, 1, 2, \dots, C \quad (4)$$

$$f_2(h_i, c_z) = \frac{1}{W} \cdot \sum_{j=1}^W x(h_i, w_j, c_z), i=0, 1, \dots, H; \\ z=0, 1, 2, \dots, C \quad (5)$$

其中, i, j, z 分别是高度 h 、宽度 w 和通道 c 的下标, 其最大值分别为 H, W, C 。

特征 f_1 和 f_2 分别再沿通道维度取平均值, 如式(6)和式(7)所示, 生成 $1 \times W \times 1$ 和 $H \times 1 \times 1$ 的一维宽度信息 g_1 和一维高度信息 g_2 , 再用一维卷积运算得到一维宽度注意力因子和一维高度注意力因子, 两者相乘得到 $H \times W \times 1$ 的空间注意力。

$$g_1(w_j) = \frac{1}{C} \cdot \sum_{z=1}^C f_1(w_j, c_z), j=0, 1, 2, \dots, W \quad (6)$$

$$g_2(h_i) = \frac{1}{C} \cdot \sum_{z=1}^C f_2(h_i, c_z), i=0, 1, 2, \dots, H \quad (7)$$

经空间注意力完善的特征图 $F' \in R^{H \times W \times C}$ 有 C 个通道, 但每个通道的重要程度不同。通过通道注意力为每个通道分配不同的权重, 模型能更好地识别输入特征的内容, 从而更好地识别巡检图像中缺陷的种类。DBSA 的通道注意力首先对输入 F' 进行全局平均池化, 将每个通道上的 $H \times W$ 个特征信息加和求平均值, 得到每个通道对应的特征 f_3 , 再采用一维卷积核代替 SENet 的全连接层进行跨通道交互, 能够在显著降低模型复杂度的同时获得更准确的通道注意力。最后, F' 再经过通道注意力的计算得到输出 F'' 。经 DBSA 重校准的特征 F'' 相比原输入 F 具有更突出的缺陷信息和更少的背景信息, 从而增强了模型的特征提取能力, 进而提升了对缺陷的检测精度。

$$f_3(c_z) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x(h_i, w_j, c_z), z=0, 1, 2, \dots, Z \quad (8)$$

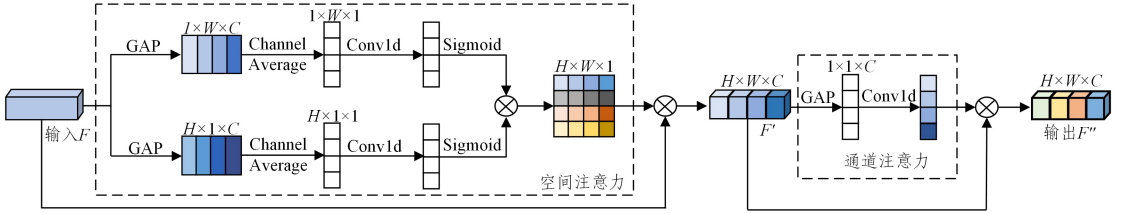


图2 DBSA 注意力机制

Fig. 2 DBSA attention mechanism

DBSA 的整体流程如算法 1 所示。

算法 1 DBSA 算法流程

输入: 特征图 F

输出: 经 DBSA 校正计算的特征图 F''

1. $x = F$
2. 求特征图各维度大小 $(n, h, w, c) = \text{Size}(x)$
3. 获取特征图的空间注意力
4. 在高度方向全局平均池化 $f_1 = \text{Gap}(x)$
5. 在通道方向求平均值 $g_1 = \text{Mean}(f_1, \text{dim}=c)$
6. 求宽度注意力因子 $xw = \text{Conv1d}(g_1)$
7. 在宽度方向上全局平均池化 $f_2 = \text{Gap}(x)$
8. 在通道维度求平均值 $g_2 = \text{Mean}(f_2, \text{dim}=c)$
9. 求高度注意力因子 $xh = \text{Conv1d}(g_2)$
10. $F' = F * xh * xw$
11. $y = F'$
12. 获取通道注意力
13. 在通道维度全局平均池化 $f_3 = \text{Gap}(y)$
14. 一维卷积求通道注意力: $xc = \text{Conv1d}(f_3)$
15. $F'' = F' * xc$
16. 返回 F''

3.2 WCFPN 特征金字塔

本文提出 WCFPN 特征金字塔对经 DBSA 提取的特征进行融合, 其结构如图 3 所示。金字塔内部共有 M, N, L, P 4 个特征层, 主干网络提取的特征 $C3, C4, C5$ 经过跨尺度特征融合和跳层连接最终生成特征信息更丰富的 $P3, P4, P5$ 。

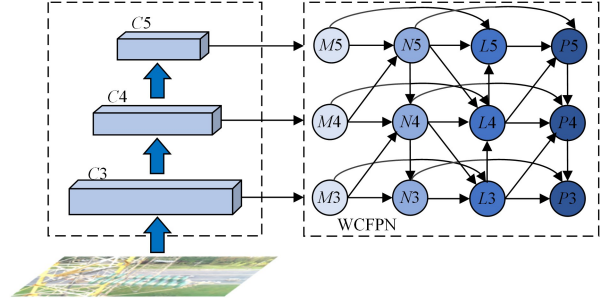


图3 WCFPN 结构图

Fig. 3 Structure of WCFPN

WCFPN 的跨尺度特征融合加强了金字塔内部特征间的连接, 以 N_4 节点为例介绍具体的融合细节。如图 4 所示, 当前层 N_4 融合的特征来源有先前层 M_3 、先前层 M_4 和当前层 N_5 这 3 个不同大小的特征图。在融合的过程中, 对 N_5 和 M_3 分别进行上采样和下采样操作, 得到和 M_4 大小相同的 40×40 特征图, 然后在通道维度上将 3 个特征图拼接, 再经过 L-ELAN 层级聚合结构完成融合。L-ELAN 将拼接后的特征图在通道维度上分组, 下分支经过 3×3 卷积核不断提取特征, 最后将 4 个相同大小的特征图在通道维度上拼接, 生成下一层特征 N_4 。

为了让金字塔不同层的特征之间进行信息交互, WCFPN 还引入了跳层连接, 如图 3 所示, 在具有相同分辨率的同一特征图维度上, 第一层 M 的特征信息参与了第三层 L 的特征融合, 第二层的特征信息 N 参与了第四层 P 的特征融合, 这让特征融合操作考虑了相同分辨率的隔层特征, 使得模型能够

得到更深层次的广义信息。

WCFPN 的 L-ELAN 层级聚合结构受 CSP, ELAN 和

YOLOv7 的 E-ELAN 网络启发,通过控制梯度路径让深层网络更加有效地学习和收敛。

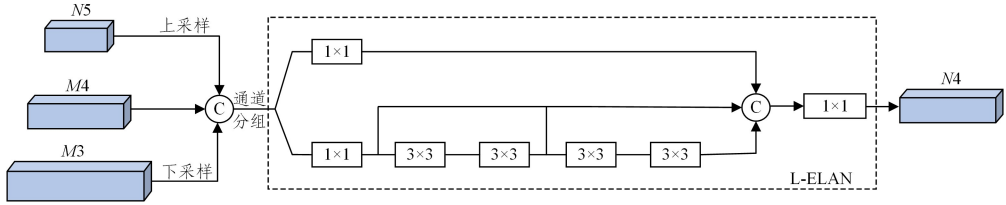


图 4 L-ELAN 结构

Fig. 4 Structure of L-ELAN

L-ELAN 层级聚合的整体流程如算法 2 所示。

算法 2 L-ELAN 层级聚合

输入: 特征图 M_3, M_4, N_5

输出: 经 L-ELAN 结构特征融合后的特征图 N_4

1. 对 N_5 上采样 $N = \text{Upsample}(N_5, \text{scale_factor}=2)$
2. 对 M_3 下采样 $M = \text{Downsample}(M_3, \text{scale_factor}=0.5)$
3. 将 N, M, M_4 通道维度拼接 $L = \text{Concat}(N, M, M_4)$
4. 将 L 沿通道维度拆分 $L_1, L_2 = \text{Split}(L)$
5. 对 L_1 卷积操作 $P_1 = \text{Conv2d}(L_1, \text{kernel_size}=1)$
6. 对 L_2 卷积操作 $P_2 = \text{Conv2d}(L_2, \text{kernel_size}=1)$
7. $T_2 = P_2$
8. 对 T_2 卷积操作 $P_3 = \text{Conv2d}(\text{Conv2d}(T_2, \text{kernel_size}=3), \text{kernel_size}=3)$
9. $T_3 = P_3$
10. 对 T_3 卷积操作 $P_4 = \text{Conv2d}(\text{Conv2d}(T_3, \text{kernel_size}=3), \text{kernel_size}=3)$
11. 将 P_1, P_2, P_3 和 P_4 在通道维度拼接: $P = \text{Concat}(P_1, P_2, P_3, P_4)$
12. 对 P 卷积操作 $N_4 = \text{Conv2d}(P, \text{kernel_size}=1)$
13. 返回 N_4

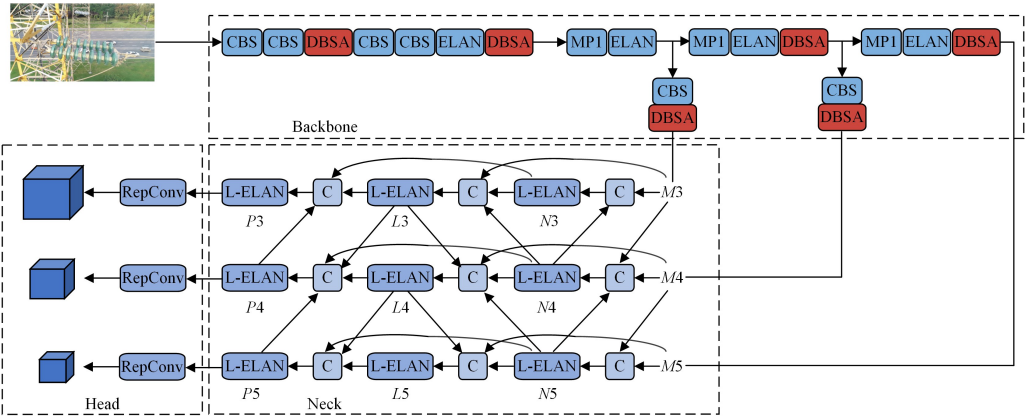


图 5 基于双分支串行混合注意力机制的输电线路缺陷检测深度神经网络模型

Fig. 5 Deep neural network model for transmission line defect detection based on dual-branch sequential mixed attention

4 实验结果与分析

4.1 实验环境

实验的硬件配置与参数如表 1 所列,实验运行的环境为:操作系统 Ubuntu22.04、Python3.8、Pytorch1.12.1。

表 1 实验配置和参数

Table 1 Experiment configuration and parameters

名称	型号	参数	参数值
GPU	NVIDIA Titan X	优化器	SGD
CPU	Ryzen5 5600x	学习率	0.01
CUDA	11.6	分辨率	640×640

4.2 数据集

本文在 5 个输电线路缺陷数据集上进行了实验,分别是绝缘子自爆、防振锤损坏、鸟巢异物、水泥杆破损和输电线路缺陷。其中输电线路缺陷是前 4 个数据集的集合,用于检验模型检测多种类缺陷的能力。表 2 列出了 5 种数据集的详细信息。

表 2 输电线路缺陷数据集

Table 2 Transmission line defect datasets

名称	简称	数量
绝缘子自爆	insulator	376
鸟巢异物	nest	385
防振锤损坏	hammer	188
水泥杆破损	pole	130
输电线路缺陷	transmission	1079

4.3 评价指标

本文采用的4个评价指标为精准率 P 、召回率 R 、 AP_{50} 和 AP 。 P 和 R 的计算式如式(9)和式(10)所示,其中 TP 、 FP 和 FN 分别表示混淆矩阵的真阳性、假阳性和假阴性。

平均精准度 AP 是 P - R 曲线的面积,计算式如式(11)所示,其中 AP_{50} 为IOU取0.5时的 AP 值, AP 为IOU在 $[0.5, 0.95]$ 区间上每隔0.05取一次值计算得到的所有 AP 结果的平均值。

$$P = \frac{TP}{TP + FP} \quad (9)$$

$$R = \frac{TP}{TP + FN} \quad (10)$$

$$AP = \int_0^1 P(R) dR \quad (11)$$

4.4 实验结果

4.4.1 注意力对比实验

为了验证DBSA的有效性,本文在主干网络中分别加入CBAM、ECA and DBSA注意力机制。由表3可知,注意力机制可以增强Backbone的特征提取能力,从而提升检测准确率,相比CBAM和ECA,本文的DBSA取得了更显著的效果。因为DBSA从宽度和高度两个分支获取空间信息,挖掘了更多空间各像素的相关性,获得了更多的缺陷位置信息,从而更好地定位到缺陷的位置。

表3 YOLOv7在不同注意力下的结果

Table 3 Results of YOLOv7 with different attention

模型	数据集	P	R	AP_{50}	AP
YOLOv7	insulator	0.925	0.876	0.903	0.592
	nest	0.876	0.789	0.770	0.358
	hammer	0.863	0.826	0.852	0.487
	pole	0.750	0.857	0.78	0.335
	transmission	0.791	0.681	0.727	0.382
YOLOv7+ CBAM	insulator	0.904	0.892	0.895	0.595
	nest	0.862	0.781	0.763	0.349
	hammer	0.859	0.847	0.854	0.494
	pole	0.753	0.844	0.769	0.332
	transmission	0.787	0.687	0.705	0.398
YOLOv7+ ECA	insulator	0.903	0.910	0.907	0.604
	nest	0.892	0.792	0.791	0.357
	hammer	0.872	0.833	0.862	0.491
	pole	0.779	0.862	0.812	0.345
	transmission	0.776	0.692	0.732	0.385
YOLOv7+ DBSA	insulator	0.935	0.892	0.911	0.613
	nest	0.903	0.772	0.787	0.363
	hammer	0.883	0.839	0.876	0.492
	pole	0.763	0.865	0.801	0.347
	transmission	0.789	0.712	0.739	0.389

4.4.2 FPN对比实验

为了验证WCFPN的有效性,本文将YOLOv7的Neck依次转换为BiFPN和WCFPN,从表4可以看出,WCFPN取得了比PAFPN和BiFPN更好的效果。因为WCFPN的跨尺度特征融合和跳层连接让特征之间发生了更充分的融合,进而使低层空间信息和高层语义信息发生了更有效的交换。低层空间信息包含了整个图像的空间整体布局,高层语义信息包含了缺陷的位置信息,两者有效的信息交换让模型更加关注全局与局部的关系,从而更容易识别缺陷等小目标。

表4 YOLOv7在不同Neck下的实验结果

Table 4 Results of YOLOv7 with different Necks

模型	数据集	P	R	AP_{50}	AP
YOLOv7 (PAFPN)	insulator	0.925	0.876	0.903	0.592
	nest	0.876	0.789	0.770	0.358
	hammer	0.863	0.826	0.852	0.487
	pole	0.750	0.857	0.780	0.335
	transmission	0.791	0.681	0.727	0.382
YOLOv7+ BiFPN ^[18]	insulator	0.954	0.872	0.914	0.610
	nest	0.892	0.783	0.781	0.349
	hammer	0.875	0.855	0.881	0.492
	pole	0.760	0.862	0.791	0.344
	transmission	0.776	0.693	0.717	0.376
YOLOv7+ WCFPN	insulator	0.945	0.892	0.923	0.621
	nest	0.882	0.795	0.783	0.342
	hammer	0.879	0.852	0.879	0.513
	pole	0.782	0.883	0.811	0.369
	transmission	0.803	0.688	0.745	0.397

4.4.3 消融实验

为了评估模型DBSA和WCFPN的效果,本文设计了消融实验,实验结果如表5所列。表中average AP_{50} , average AP 指模型在5个数据集上 AP_{50} 和 AP 的均值。结果表明DBSA和WCFPN两个模块均使得原模型的检测效果有所提升,因为DBSA抑制了图像大量无关背景的干扰,让模型更关注缺陷的信息,而WCFPN更充分的特征融合保证了信息的有效交换,让模型获取了更广义的信息。两者结合使用可以取得最高的检测精度,说明经DBSA提取的特征被WCFPN金字塔融合后缺陷的信息更加突出,进而使检测效果得到提升。

表5 消融实验

Table 5 Ablation experiment

算法	DBSA	WCFPN	Average AP_{50}	Average AP
YOLOv7			0.806	0.431
模型1	✓		0.823	0.441
模型2		✓	0.828	0.448
本文模型	✓	✓	0.843	0.461

4.4.4 对比实验

为了验证所提模型的有效性,本文采用不同的模型在5个数据集上进行训练,并将 AP_{50} 、 AP 作为评价指标进行验证。实验结果如表6和表7所列。相较于原模型YOLOv7,本文模型在5个数据集上的平均 AP_{50} 和 AP 分别提升了3.7%和3%。取得了最好的检测效果。因为模型主干网络中的DBSA增强了特征提取能力,WCFPN实现了信息的有效传输,两者的结合使模型提取到了更多的缺陷种类信息和位置信息,进而增强了模型在缺陷检测中的表现。

表6 不同模型的 AP_{50}

Table 6 AP_{50} of different models

模型	insulator	nest	hammer	pole	transmission	average
Faster-RCNN	0.818	0.765	0.834	0.717	0.748	0.776
Grid-RCNN ^[19]	0.795	0.724	0.815	0.636	0.729	0.740
CascadeRCNN ^[20]	0.836	0.742	0.846	0.725	0.773	0.784
Ddod ^[21]	0.828	0.761	0.849	0.741	0.762	0.788
YOLOv7	0.903	0.77	0.852	0.78	0.727	0.806
本文模型	0.937	0.797	0.887	0.823	0.769	0.843

表7 不同模型的AP

Table 7 AP of different models

模型	insulator	nest	hammer	pole	transmission	average
Faster-RCNN	0.536	0.315	0.442	0.288	0.384	0.393
Grid-RCNN ^[19]	0.525	0.297	0.417	0.308	0.348	0.379
CascadeR-CNN ^[20]	0.482	0.322	0.435	0.342	0.413	0.392
Ddod ^[21]	0.493	0.323	0.451	0.331	0.375	0.395
YOLOv7	0.592	0.358	0.487	0.335	0.382	0.431
本文模型	0.625	0.366	0.509	0.388	0.415	0.461

结束语 为了解决电网巡检中图像背景复杂、缺陷尺寸小的难题,本文提出了基于双分支串行混合注意力机制的输电线路缺陷检测深度神经网络模型。DBSA 双分支串行混合注意力机制用一维卷积核先后提取特征图的空间注意力和通道注意力,使模型提取到的特征包含更多的缺陷信息而不是无关背景噪音;WCFPN 特征金字塔加强了金字塔内部特征间的连接,让特征更充分地融合,提升了模型发现小目标的能力,最后在5个数据集上测试了模型的性能,实验结果表明本文模型取得了最好的检测结果。

本文模型虽然提升了输电线路缺陷的检测精度,但也略微增加了模型的复杂度,下一步的研究工作还需从轻量化的角度出发,在基本不影响检测精度的前提下,降低模型的复杂度从而轻量化模型。

参考文献

- [1] ZHAO Z, QI H, QI Y, et al. Detection method based on automatic visual shape clustering for pin-missing defect in transmission lines[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(9): 6080-6091.
- [2] ZHANG X, ZHANG Y, LIU J, et al. InsuDet: A fault detection method for insulators of overhead transmission lines using convolutional neural networks[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-12.
- [3] LV F C, NIU L L, WANG S H, et al. Intelligent detection and parameter adjustment strategy of major electrical equipment based on optimized YOLOv4[J]. Transactions of China Electrotechnical Society, 2021, 36(22): 4837-4848.
- [4] LIU M, LI Z, LI Y, et al. A fast and accurate method of power line intelligent inspection based on edge computing[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-12.
- [5] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single Shot Multibox Detector[C]// Computer Vision — ECCV 2016: 14th European Conference. Springer International Publishing, 2016: 21-37.
- [6] 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.
- [7] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv: 2004. 10934, 2020.
- [8] REN S, HE K, 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.

- [9] HE K, GKIOXARI G, DOLLÁR P, et al. Mask r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision. 2017: 2961-2969.
- [10] GLENN J. YOLOv5 release v6.1. [OL]https://github.com/ultralytics/yolov5/releases/tag/v6.1, 2022. 2, 7, 10.
- [11] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. arXiv: 2207. 02696, 2022.
- [12] NIU Z, ZHONG G, YU H. A review on the attention mechanism of deep learning[J]. Neurocomputing, 2021, 452: 48-62.
- [13] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 7132-7141.
- [14] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.
- [15] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 11534-11542.
- [16] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 2117-2125.
- [17] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 8759-8768.
- [18] ZHAO Q, SHENG T, WANG Y, et al. M2det: A single-shot object detector based on multi-level feature pyramid network[C]// Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence. 2019: 9259-9266.
- [19] LU X, LI B Y, YUE Y X, et al. Grid R-CNN plus: faster and better [J]. arXiv: 1906. 05688, 2019.
- [20] CAI Z, VASCONCELOS N. Cascade r-cnn: Delving into high quality object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 6154-6162.
- [21] CHEN Z, YANG C, LI Q, et al. Disentangle your dense object detector[C]// Proceedings of the 29th ACM International Conference on Multimedia. 2021: 4939-4948.



HAO Ran, born in 1999, postgraduate. His main research interests include deep learning and defect detection.



WANG Hongjun, born in 1977, Ph.D., associate researcher, Ph.D supervisor, is a senior member of CCF(No. 19036S). His main research interests include artificial intelligence, machine learning, and data mining.