

基于改进 YOLOv4 的变电站缺陷检测^①



陈 婷¹, 周 旻², 韩 勤², 张 湘², 茅耀斌¹

¹(南京理工大学 自动化学院, 南京 210094)

²(浙江华云信息科技有限公司, 杭州 310030)

通信作者: 茅耀斌, E-mail: myb_nust@126.com

摘 要: 为提高变电站设备缺陷的检测精度, 保障变电站运行安全, 提出一种基于改进 YOLOv4 的缺陷检测算法。不同于原始 YOLOv4, 该算法使用一维卷积替代全连接来优化 CBAM 卷积注意力模块, 然后将其嵌入主干网络中以增强特征提取能力; 同时, 在特征融合中应用空洞卷积扩大感受野, 聚合更广的语义信息。该算法在现场拍摄的样本集上进行测试, mAP 可达到 86.97%, 相比原始 YOLOv4 提高了 2.78%。实验结果表明, 本文提出的 YOLOv4 改进算法能够提升网络性能, 更好地应用于变电站设备缺陷检测任务。

关键词: YOLOv4; 缺陷检测; CBAM 注意力; 空洞卷积; 特征融合; 卷积神经网络

引用格式: 陈婷, 周旻, 韩勤, 张湘, 茅耀斌. 基于改进 YOLOv4 的变电站缺陷检测. 计算机系统应用, 2022, 31(6): 245-251. <http://www.c-s-a.org.cn/1003-3254/8520.html>

Defect Detection for Substation Based on Improved YOLOv4

CHEN Ting¹, ZHOU Min², HAN Qin², ZHANG Xiang², MAO Yao-Bin¹

¹(School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China)

²(Zhejiang Huayun Information Technology Co. Ltd., Hangzhou 310030, China)

Abstract: To increase the defect detection accuracy on substation equipment and thus ensure the operation safety of the substation, this study proposes a defect detection algorithm based on an improved YOLOv4. Unlike the original YOLOv4, the new algorithm replaces the fully connected layers with one-dimensional convolution to optimize the convolutional block attention module (CBAM), which is then embedded into the backbone network to enhance the feature extraction ability. Meanwhile, dilated convolution is used in feature fusion layers for expanding the receptive field and aggregating broader semantic information. The algorithm is tested on images captured in real substation scenes and achieves a mean average precision (mAP) of 86.97%, an increase of 2.78% on that of the original YOLOv4. Experimental results show that the proposed algorithm can improve the network performance and is thus more suitable for defect detection on substation equipment.

Key words: YOLOv4; defect detection; convolutional block attention module (CBAM) attention; dilated convolution; feature fusion; convolutional neural networks (CNN)

变电站中各类设备的可靠运行是变电站安全生产的重要基础。由于变电设备大多处于露天环境, 可能出现呼吸器破损、硅胶变色、表盘模糊等缺陷以及存在异物、鸟巢等异常, 若未及时发现此类问题, 可能造成不同等级的电力事故^[1]。因此, 变电设备的定期巡检是

变电站安全运作的重要保障。然而, 传统的人工巡检难以做到对变电站安全隐患、设备缺陷的及时、全面发现, 导致其无法为变电站的可靠运转提供安全保障。近年来, 图像视频处理以及计算机视觉技术的发展极大地提高了变电站巡检质量和效率。通过巡检机器人或

① 收稿时间: 2021-08-27; 修改时间: 2021-09-26; 采用时间: 2021-10-14; csa 在线出版时间: 2022-05-26

无人机拍摄现场设备,技术人员可利用相关技术直接对获取的图片或视频信息进行批量检测。因此,基于图像处理的缺陷检测方法研究,对实现变电站设备的全面诊断、检查,以及进一步提升巡检工作开展的效率和质量具有重要意义。如何提升设备缺陷检测性能,为智能化巡检提供更先进的技术支撑,成为现阶段电力领域的研究重点。

传统的图像识别方法大多依赖于手工设计的特征,如尺度不变特征、梯度直方图特征等,根据特征点间的匹配情况锁定对象位置,实现目标检测。对于变电站场景, Cerón 等^[2]根据圆对称几何关系,利用 Canny 算子对输电线路进行检测。Pernebayeva 等^[3]使用滑窗提取绝缘子特征,利用支持向量机进行分类,实现绝缘子表面冰雪覆盖检测。侯一民等^[4]利用改进 SIFT 方法提取设备特征,实现变压器、刀闸定位。然而,传统方法存在大量漏检、错检情况,泛化能力较弱。随着深度学习的迅猛发展,使用卷积神经网络提取特征的方式逐渐取代手工特征设计。李文璞等^[5]使用 Faster R-CNN^[6]实现了对变电设备红外图像的缺陷检测。尹子会等^[7]基于改进的 Faster R-CNN 实现了变电站设备 4 类常见缺陷的检测与识别。但 Faster R-CNN 属于两阶段目标检测,模型尺寸较大,检测速度较慢。YOLO 系列^[8,9]将目标检测问题作为回归问题,在产生候选区域的同时完成分类和定位,推理速度较快。田二胜等^[10]提出基于 YOLOv4 的输电线路外破隐患识别算法,对吊车、塔吊等施工机械进行检测。

YOLOv4 是一种优秀的目标检测算法,可用于变电站异物和设备缺陷检测。然而,鸟巢、表盘模糊、呼吸器破损等缺陷检测任务存在以下难点:一方面,变电站背景信息复杂,不同类别之间的特征差距较大,同一类别部件的破损或模糊种类较多、形状多样,尤其样本数量有限且样本存在类别不均衡问题,难以提取到丰富的特征以应对不同的物体;另一方面,由于拍摄角度多变,同一类目标的尺度变化较大,也带来检测困难。原 YOLOv4 在解决这些问题上还不够强大,因此,本文从提高特征提取能力和增加感受野两点上对原始 YOLOv4 进行改进:1) 为解决背景复杂、特征差距大、形状多样且样本数量有限问题,本文在主干网络中引入优化的 CBAM^[11]注意力机制,更准确地聚合物体特征;2) 为解决拍摄角度多、尺度变化大问题,在 PAN^[12]的上、下采样模块使用空洞卷积增大感受野,

获取更广泛的特征表示,同时减少采样模块造成的特征损失。实验结果证明了两项改进的有效性,表明改进后的 YOLOv4 能够更好地实现变电站鸟巢、表盘模糊、呼吸器硅胶变色以及呼吸器硅胶筒破损 4 类缺陷的检测任务。

1 YOLOv4 目标检测算法原理

YOLOv4 作为 YOLO 系列的第 4 代检测模型,其基本思想是将整张图像作为网络输入,直接在输出层回归边界框的位置及其所属类别。YOLOv4 的网络结构如图 1 所示,图的左半部分为网络主体结构,其中 CBM、CBL、CSP、SPP、上采样和下采样的详细结构由图 1 右边的子模块部分具体描述。由图 1 可知,YOLOv4 网络主要包含以下 3 部分:用于提取特征的全卷积主干网络 CSPDarknet53、用于特征融合的 SPP^[13]和 PAN 模块以及 3 个检测头 Head。

主干部分由 5 个 CSP^[14]模块组成,每个模块的第一个卷积核步长 stride 设置为 2,根据式(1),当卷积核尺寸 F_{size} 为 3, padding 为 1 时,若步长 S 为 2,则特征图的输出宽高尺寸将缩小为输入的一半。因此,特征图在整个主干网络中将经历 5 次缩小,当输入为 416×416 时,主干最后 3 层的特征图大小分别为 52×52 、 26×26 、 13×13 。

$$out_{size} = \frac{in_{size} - F_{size} + 2P}{S} + 1 \quad (1)$$

对主干网络获取的 3 个不同尺度的特征映射, YOLOv4 进一步使用 SPP 和 PAN 做特征融合。其中, PAN 是主干层到检测层的重要纽带,它首先进行深层到浅层的上采样和特征拼接,传达强语义特征,随后进行浅层到深层的下采样和特征拼接,传达强定位特征,两者结合,实现有效特征聚合。

特征融合输出的 3 个特征层分别送往 3 个检测头 Head 进行预测,每层分配 3 个不同尺寸的先验框。假设当前特征层的宽高为 a ,则当前层将被划分为 $a \times a$ 个网格,每个网格按照 3 个先验框尺寸进行预测,则一共有 $a \times a \times 3$ 个预测框。每个预测框需要预测的内容有:当前框的中心点坐标和宽高 (x, y, w, h) 、当前框的置信度 P_{conf} 以及当前框 c 个类别的概率。因此,当前层的输出尺寸为 $a \times a \times 3 \times (c + 5)$ 。合并 3 个检测头的预测结果后,利用设定的置信度阈值以及 DIoU-NMS^[15]非极大值抑制,去除错误、冗余预测框,获得最终检测目标。

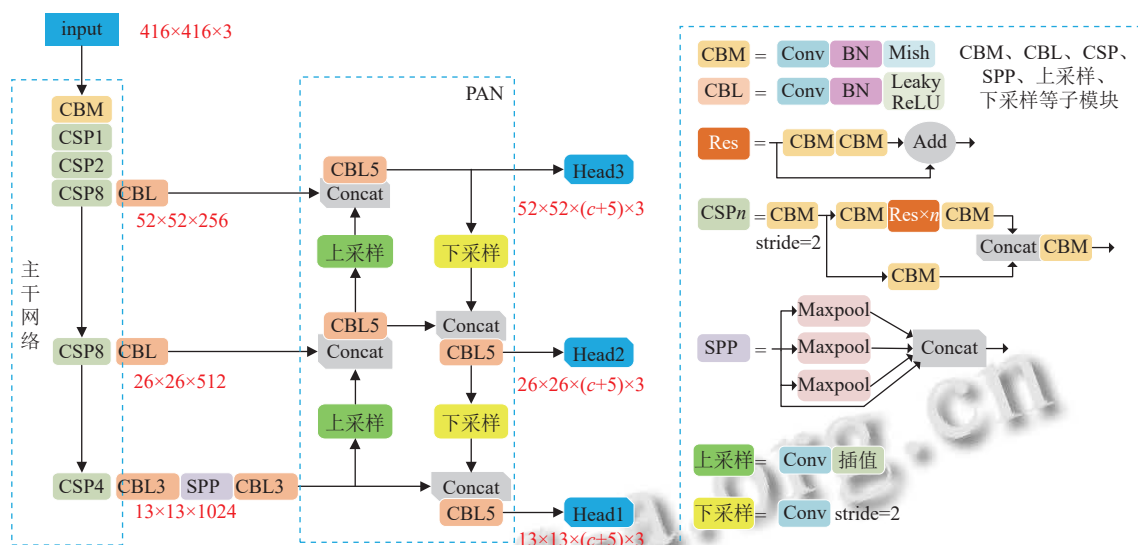


图1 YOLOv4 网络结构

2 YOLOv4 算法改进

在变电站环境中对设备进行缺陷检测时,由于背景干扰信息较多,且缺陷或异物的形状、颜色等复杂多变,尤其在电力应用中,缺陷样本的数量非常有限,直接使用 YOLOv4 难以提取足够的特征应对多样化、多尺度目标,可能出现较多误检、漏检情况。为进一步获取更丰富的多样化特征,提高检测精度,本文从主干网络和特征融合两方面对 YOLOv4 进行改进,然后基于改进的 YOLOv4 算法实现对鸟巢、表盘模糊、呼吸器硅胶变色以及呼吸器硅胶筒破损 4 类缺陷的检测识别。改进后的 YOLOv4 网络结构如图 2 所示,其中

粉色部分为改进模型增加的模块,相比图 1,主要增加了 CBAM 模块以及优化了上采样、下采样结构。根据图 2,改进后的 YOLOv4 首先使用嵌入 CBAM 注意力的全卷积网络提取输入图像特征, CBAM 依据人为引入的先验知识主动关注重要信息,有益于关键特征的提取;随后对 3 个尺度的特征图采用 PAN 进行信息融合,其中尺度的扩大缩小涉及采样模块,采样模块使用空洞卷积,使感受野在原有基础上进一步扩大,更好地包含不同形状、尺度的目标对象,同时采样模块融合了更广范围的特征,可降低采样带来的特征损失与数据破坏;最后利用融合输出的 3 个特征层实现目标预测。

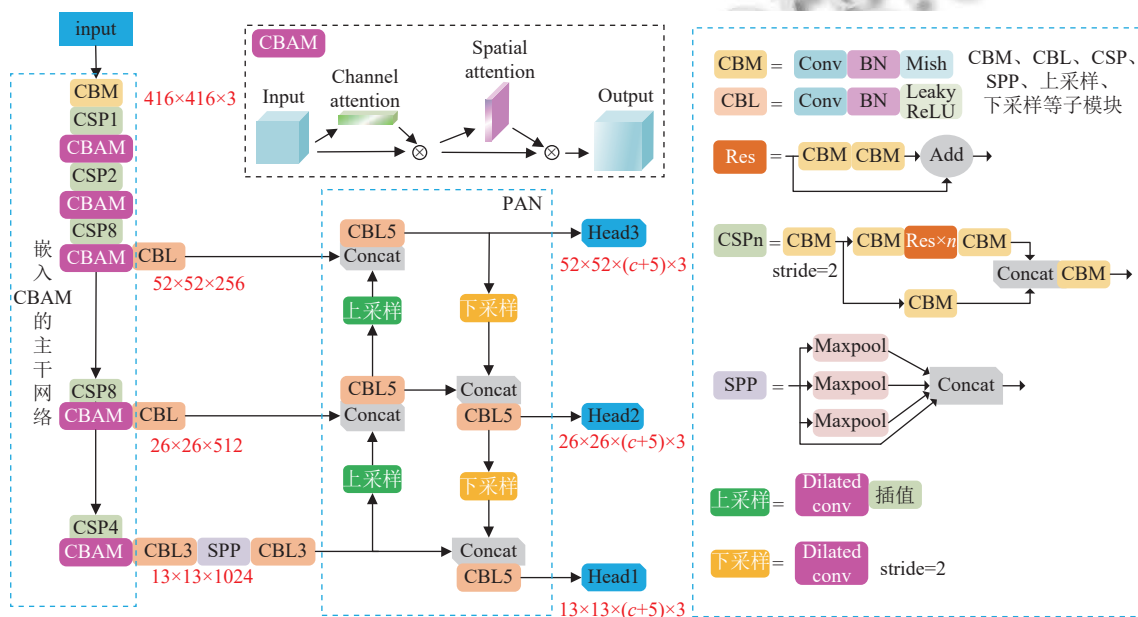


图2 YOLOv4 改进算法网络结构 (粉色部分为改进算法增加的模块)

2.1 基于注意力机制的主干网络改进

卷积注意力模块 (convolutional block attention module, CBAM) 是一种轻量、通用的注意力机制, 相较于主流的 SE (squeeze-and-excitation)^[16] 方法, CBAM 在其通道注意力的基础上叠加了空间注意力, 其中通道注意力关注输入的语义信息, 空间注意力关注位置信息, 二者结合可以获得更好的特征表示. 原始 YOLOv4 在特征融合部分使用了 SAM 注意力, CBAM 相较于 SAM 虽然实现原理上稍微复杂, 但通过引入平均池化和最大池化两种先验特征, 模型表达更为高效, 将其应用于主干网络, 可在一定程度上提升网络性能.

给定卷积神经网络生成的特征映射 $F \in \mathbb{R}^{C \times H \times W}$, CBAM 从通道和空间两个维度上计算注意力, 计算过程如图 3 所示. 首先, 为高效计算通道注意力, 使用最大池化和平均池化对空间维度进行压缩, 生成两个一维向量描述通道重要性, 将两个一维向量送入由 MLP 组成的共享网络后按元素相加获得通道注意力 $M_C \in \mathbb{R}^{C \times 1 \times 1}$, 将该注意力与输入特征按元素相乘, 获得初步调整的特征图 F' ; 其次, 在通道维度上对 F' 做最大池化和平均池化, 得到两个不同的二维特征描述, 将两个二维特征合并后再执行卷积和激活生成空间注意力 $M_S \in \mathbb{R}^{1 \times W \times H}$, 将空间注意力与 F' 按元素相乘得到最终输出 F'' . 整个过程的数学描述如下所示:

$$\begin{aligned} F' &= M_C(F) \otimes F \\ F'' &= M_S(F') \otimes F' \end{aligned} \quad (2)$$

其中, M_C 与 M_S 首先要进行广播以匹配维度, 符号 $\otimes$ 表示逐元素相乘.

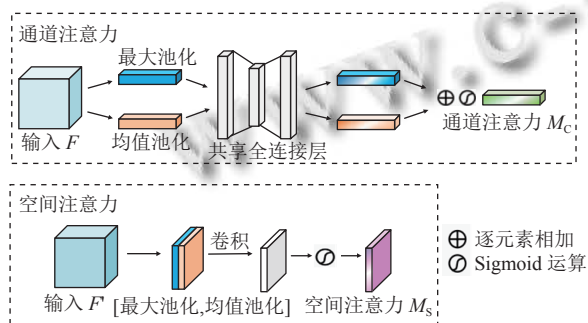


图3 CBAM通道注意力与空间注意力计算示意图

观察通道注意力模块可以发现, 对于初步提取的最大值池化特征和均值池化特征, CBAM 采用全连接层分别对二者做通道特征聚合. 考虑到全连接层的参数量较大, 且由于相邻通道间往往相关性更大, 使用全

连接会产生较多冗余计算, 因此, 本文采用一维卷积代替全连接^[17]. 一维卷积只在宽或高单个方向上进行滑动窗操作并相乘求和, 由于同一个窗口内是权值共享的, 因此能有效降低参数量与计算量, 轻松实现跨通道的信息交互与整合. 通过改变卷积核的数量, 可以实现降维和升维, 本文将通道数先缩小为原来的 1/8, 随后升至初始维度. 两个池化特征经过两次卷积后按元素相加, 最后通过 Sigmoid 函数生成通道注意力.

对于 YOLOv4 的特征提取主干网络, 本文在其每个 CSP 模块后引入优化后的 CBAM 注意力模块, 如图 2 的主干部分所示, 对每个 CSP 的输出特征层先后进行通道和空间两个维度上的加权处理, 使网络更关注重要的有效信息而抑制无效信息, 提高检测性能.

2.2 基于多尺度特征融合的采样模块改进

特征融合是对不同尺度的特征映射在上采样和下采样的过程中进行信息交互, 实现浅层网络的语义信息学习和深层网络的定位信息学习, 然而在一系列上采样和下采样过程中, 信息的传递不可避免地存在损失. 具体来说, 上采样是通过插值的方式将特征映射的尺寸扩大, 插值会导致一些小目标的数据被破坏, 无法重建回原始空间信息; 下采样是通过将步长 stride 设为 2 使特征映射的尺寸减小, 等价于做特征压缩, 这一过程会导致内部数据结构和空间层级化信息丢失. 另外, 由于部件本身以及拍摄角度的原因, 目标对象的尺度和形状差异较大, 需要包含更广的尺度以融入更多、更丰富的特征. 为改善以上问题, 本文使用空洞卷积代替普通卷积来增大感受野, 所得更广尺度的特征与原始尺度特征进行融合, 层次化信息更丰富, 在采样时, 能够保留更多信息, 减少特征丢失.

空洞卷积又称扩张卷积, 相比于标准卷积, 增加了一个扩张率参数, 用来表示卷积核各数据间的间隔数量. 图 4 展示了两种不同扩张率的卷积, 其中, 图 4(a) 表示 3×3 且扩张率为 1 的卷积核, 等价于标准卷积; 图 4(b) 表示 3×3 且扩张率为 2 的空洞卷积, 在执行卷积运算前, 空洞处先进行补零操作, 形式上类似于 5×5 卷积. 可见, 一个扩张率为 2 的 3×3 空洞卷积, 感受野等价于一个 5×5 的标准卷积, 而参数只需要 9 个. 表明在相同计算条件下, 空洞卷积可以提供更大的感受野. 考虑到空洞卷积本身的实现效率低于普通卷积, 因此仅在关键的、易造成特征丢失的上采样、下采样模块中引入空洞卷积, 如图 2 中 PAN 的采样部分所示, 采用空洞

卷积有效增大感受野,融合更广尺度的特征,从而更好地保留图像空间特征,减少信息损失和数据破坏。

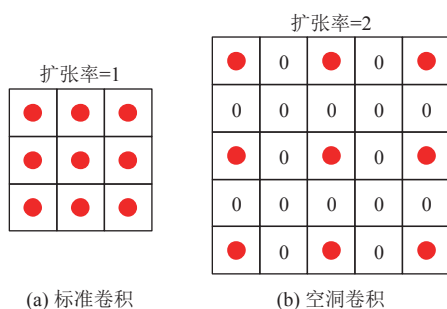


图4 标准卷积与空洞卷积示意图

3 实验与结果分析

3.1 实验数据集及评价指标

通过变电站现场拍摄和网络抓取两种方式,本文获取了包含表盘模糊、呼吸器硅胶变色、呼吸器破损和鸟巢4种缺陷的6571张图片,其中70%数据为现场采集图片。各类样本图像分布为:表盘模糊2259张,硅胶变色1549张,呼吸器破损405张,鸟巢2358张。该数据集涵盖了不同天气状况、光照强度下的图片,且图片中同时包含了一定数量的清晰表盘和正常呼吸器设备作为负样本。选用LabelImg工具标注图片,生成相应标签集。然后将整个数据集按5:1随机分为训练集和测试集,其中训练集再以10%的比例随机划分出验证集。在训练前期,将使用Mosaic数据增强方法对训练集进一步扩充。

本文采用查准率 P 、查全率 R 、精度 AP 以及各类的平均精度 mAP 作为性能评价指标。查准率与查全率的计算方法如下:

$$P = \frac{TP}{TP+FP}, R = \frac{TP}{TP+FN} \quad (3)$$

其中, TP 表示实际为某类正样本,预测为该类别且预测框与真实框IoU大于0.5的框的数量; FP 表示预测类别错误或IoU小于0.5的情况; FN 表示实际为负样本但预测出正样本的情况。根据以上3项即可求出查准率 P 和查全率 R 。对不同置信度阈值求取不同的 P 、 R 组合,即可得到 P - R 曲线,该曲线与坐标轴围成的面积表示为 AP ,用来衡量某个类别的检测精度, AP 越大,精度越高。对每类 AP 求平均值,得到各类的平均检测精度 mAP ,衡量模型整体检测精度。

另外,采用每秒处理帧数(frames per second, FPS)来描述模型的推理速度,其大小等于模型检测时间的倒数。

3.2 实验过程

基于改进YOLOv4算法的变电站缺陷检测实验均在同一实验环境下完成。实验平台为Ubuntu 18.04.5操作系统,显卡为NVIDIA GeForce RTX 3090,运行内存为24 GB。深度学习软件环境为CUDA 11.1、Python 3.7.10、PyTorch 1.8.1。

本文设定网络的输入大小为 416×416 ,训练时首先加载YOLOv4在COCO数据集上的预训练模型作为特征提取器,当前网络的主干层先冻结,仅训练剩余层的参数,迭代50轮后解冻主干层,训练全局网络参数。前50轮训练中,初始学习率设置为0.001,解冻后再次训练50轮,学习率设置为0.0001,均选择Adam优化器对参数进行优化,批次大小设为64。

3.3 结果与分析

(1) 基于改进YOLOv4的缺陷检测结果与分析

将基于改进YOLOv4算法训练所得模型对缺陷进行检测,典型检测结果如图5所示。进一步将模型在测试集上进行性能评估,结果如表1所示,各类别中,表盘模糊的查准率最高,为96.43%,呼吸器变色的查全率最高,为97.06%,该类别精度也达到了最高的98.42%,模型整体的平均精度为86.97%。同时由表1可知,呼吸器破损的查准率较低,主要原因是该类别的数据量相对不足,对于各种破损情况,无法精确识别;鸟巢的查全率最低,这与鸟巢处于高空、大小不一、存在遮挡等因素有重要关联。根据结果分析可以发现,对于表盘模糊、呼吸器变色两类缺陷,由于目标的基本形状较为统一,主要通过区分颜色和纹理特征判断是否存在缺陷,因此精度较高;而鸟巢和呼吸器硅胶筒破损,会出现结构残缺的情况,目标形状复杂多变,因而检测难度较大,精度相对较低。

(2) 消融实验

对比本文算法、初始YOLOv4以及单独应用第2.1节、第2.2节改进方法在测试集上的性能,其结果如表2所示。从表中可知,单独应用CBAM注意力对各类别的精度以及平均精度均有一定提升,证明了网络特征提取能力的加强能够有效提升模型精度;使用空洞卷积除了鸟巢精度没有提升外,其余指标尤其是 mAP 均都得到优化,证明扩大感受野对模型性能提升有一

定的帮助. 同时运用两种优化方法时, 最终的 mAP 达到了 86.97%, 相比原始 YOLOv4 提升了 2.78%, 表明了两项优化结合的有效性. 尽管鸟巢 AP 略有降低, 但

模型更加关注数据量较少的呼吸器破损类别, 并将其 AP 提升了 7.77%, 极大地弥补了样本量不足的弊端, 同时这也对模型整体性能优化发挥了重要作用.

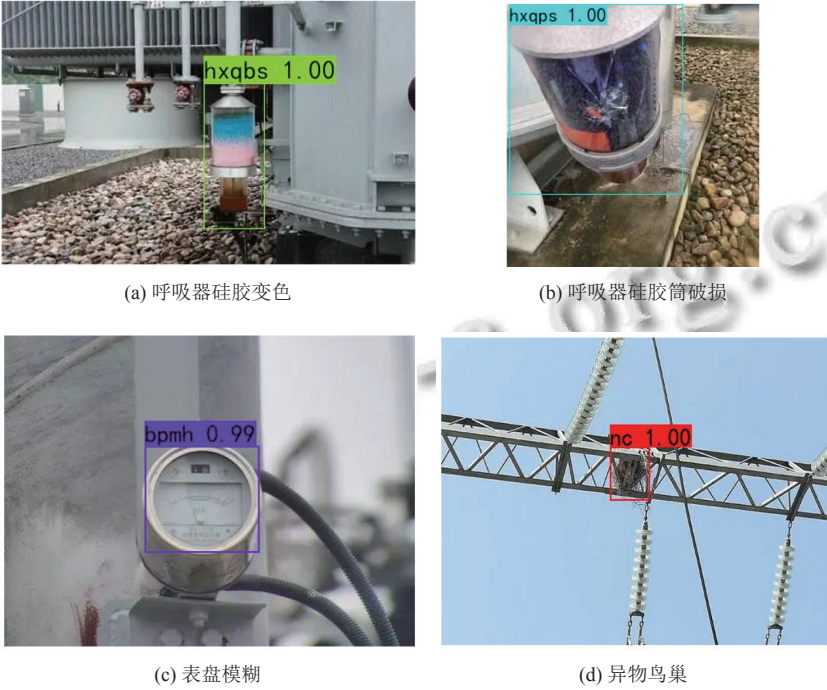


图5 缺陷检测效果图

表1 改进 YOLOv4 算法性能评估 (%)

类别	P	R	AP	mAP
表盘模糊	96.43	86.75	93.06	86.97
呼吸器变色	95.93	97.06	98.42	
呼吸器破损	80.00	85.71	84.54	
鸟巢	81.28	72.06	71.87	

表2 消融实验性能对照 (%)

算法	表盘模糊AP	呼吸器变色AP	呼吸器破损AP	鸟巢AP	mAP
YOLOv4	89.26	98.25	76.77	72.46	84.19
YOLOv4+CBAM	92.13	98.78	77.16	73.44	85.38
YOLOv4+空洞	92.08	98.59	78.05	72.43	85.29
本文算法	93.06	98.42	84.54	71.87	86.97

(3) 不同模型输入尺寸性能对比

为进一步探索模型输入尺寸对性能的影响, 本文将模型的输入尺寸扩大到 608×608, 然后将其与 416×416 的模型做性能对比, 结果如表3所示. 可以看出, 增大模型输入尺寸对鸟巢的检测精度有了较大提升, 模型 mAP 也因此而提高到 87.61%; 然而, 基于 GPU 的推理速度从 57.9 fps 降至 39.3 fps, 减少了 32%, 这是由

于增大输入尺寸, 模型的计算量也随之增加, 对推理平台的算力要求也更高. 因此, 当模型的应用场景对实时性或推理时间要求较高时, 适合采用 416×416 的输入; 相反, 若对模型精度要求更高, 且设备算力允许时, 可增大输入尺寸.

表3 不同模型输入尺寸性能对照

模型输入尺寸	表盘模糊AP (%)	呼吸器变色AP (%)	呼吸器破损AP (%)	鸟巢AP (%)	mAP (%)	速度 (fps)
416×416	93.06	98.42	84.54	71.87	86.97	57.9
608×608	91.97	97.92	84.84	75.71	87.61	39.3

4 结论

本文为提高变电站缺陷检测精度, 提出基于改进 YOLOv4 的变电站缺陷检测算法, 首先, 利用优化的 CBAM 注意力提升主干网络的特征提取能力; 其次, 将空洞卷积应用于特征融合层中的上采样与下采样中, 通过增大感受野获得较全的语义信息; 最后, 将改进算

法在四类缺陷样本集上进行测试,4类缺陷的 mAP 达到 86.97%,相较于原始 YOLOv4 提升了 2.78%,证明了算法的优越性。

实验结果表明,改进的 YOLOv4 在缺陷样本数量有限、类别数量不平衡且目标形状和尺度多变的条件下,提升了变电站设备缺陷检测性能,为变电站运维智能化提供了有效手段,具有一定的实际应用价值。后续工作将进一步对所改进的网络进行压缩和约简,在保证精度的同时降低参数量与计算量,提高推理速度,实现实时检测。

参考文献

- 1 应樱. 基于改进 Faster R-CNN 的变电站设备缺陷检测算法研究[硕士学位论文]. 杭州: 浙江大学, 2021. [doi: [10.27461/d.cnki.gzjdx.2021.000392](https://doi.org/10.27461/d.cnki.gzjdx.2021.000392)]
- 2 Cerón A, Mondragon IF, Prieto F. Power line detection using a circle based search with UAV images. Proceedings of 2014 International Conference on Unmanned Aircraft Systems. Orlando: IEEE, 2014. 632–639 [doi: [10.1109/ICUAS.2014.6842307](https://doi.org/10.1109/ICUAS.2014.6842307)]
- 3 Pernebayeva D, Bagheri M, James A. High voltage insulator surface evaluation using image processing. Proceedings of 2017 International Symposium on Electrical Insulating Materials. Toyohashi: IEEE, 2017. 520–523.
- 4 侯一民, 邸建铭. 改进的尺度不变特征转换精确图像匹配在电力设备目标定位中的应用. 中国电机工程学报, 2012, 32(19): 134–139. [doi: [10.13334/j.0258-8013.pcsee.2012.19.003](https://doi.org/10.13334/j.0258-8013.pcsee.2012.19.003)]
- 5 李文璞, 谢可, 廖道, 等. 基于 Faster RCNN 变电设备红外图像缺陷识别方法. 南方电网技术, 2019, 13(12): 79–84. [doi: [10.13648/j.cnki.issn1674-0629.2019.12.012](https://doi.org/10.13648/j.cnki.issn1674-0629.2019.12.012)]
- 6 Ren SQ, He KM, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137–1149. [doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)]
- 7 尹子会, 孟荣, 范晓丹, 等. 融合边缘计算和改进 Faster R-CNN 的变电站设备典型视觉缺陷检测系统. 中国科技论文, 2021, 16(3): 343–348. [doi: [10.3969/j.issn.2095-2783.2021.03.018](https://doi.org/10.3969/j.issn.2095-2783.2021.03.018)]
- 8 Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv: 1804.02767, 2018.
- 9 Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: Optimal speed and accuracy of object detection. arXiv: 2004.10934, 2020.
- 10 田二胜, 李春蕾, 朱国栋, 等. 基于 YOLOv4 的输电线路外破隐患识别算法. 计算机系统应用, 2021, 30(7): 190–196. [doi: [10.15888/j.cnki.csa.008082](https://doi.org/10.15888/j.cnki.csa.008082)]
- 11 Woo S, Park J, Lee JY, et al. CBAM: Convolutional block attention module. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 3–19. [doi: [10.1007/978-3-030-01234-2_1](https://doi.org/10.1007/978-3-030-01234-2_1)]
- 12 Liu S, Qi L, Qin H F, et al. Path aggregation network for Instance segmentation. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 8759–8768. [doi: [10.1109/CVPR.2018.00913](https://doi.org/10.1109/CVPR.2018.00913)]
- 13 He KM, Zhang XY, Ren SQ, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. Proceedings of the 13th European Conference on Computer Vision. Zurich: Springer, 2014. 346–361. [doi: [10.1007/978-3-319-10578-9_23](https://doi.org/10.1007/978-3-319-10578-9_23)]
- 14 Wang CY, Liao HYM, Wu YH, et al. CSPNet: A new backbone that can enhance learning capability of CNN. Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Seattle: IEEE, 2020. 1571–1580. [doi: [10.1109/CVPRW50498.2020.00203](https://doi.org/10.1109/CVPRW50498.2020.00203)]
- 15 Zheng ZH, Wang P, Liu W, et al. Distance-IoU loss: Faster and better learning for bounding box regression. Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI, 2020. 12993–13000.
- 16 Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011–2023. [doi: [10.1109/TPAMI.2019.2913372](https://doi.org/10.1109/TPAMI.2019.2913372)]
- 17 付国栋, 黄进, 杨涛, 等. 改进 CBAM 的轻量级注意力模型. 计算机工程与应用, 2021, 57(20): 150–156.