DOI: 10.13382/j. jemi. B2104478

基于改进 Faster R-CNN 的电厂雨排口 污染物泄漏检测*

彭道刚¹ 王永坤¹ 周 洋² 咸尔江¹ 高义民¹ (1.上海电力大学自动化工程学院 上海 200090:2. 宝山钢铁股份有限公司电厂 上海 201900)

摘 要:针对监控图像中电厂雨排口出现的废弃油污泄漏问题,提出一种基于改进 Faster 区域卷积神经网络(Faster R-CNN)的 电厂雨排口污染物泄漏检测算法。改进 Faster R-CNN 检测算法首先使用 ResNet-50 作为主干网络,在此基础上构建多尺度特 征图金字塔结构(FPN),实现高层语义和低层语义之间的信息融合,提高了检测精度;其次采用 CloU 损失和 DloU-NMS 方法, 提高 Faster R-CNN 中边框回归的准确度;最后引入 Focal Loss 损失函数,解决了区域建议网络(RPN)生成的锚点冗余导致 R-CNN 阶段出现正负样本不均衡问题。实验结果表明,此改进算法在真实样本中表现良好,平均准确率达到 90.2%,与原 Faster R-CNN 算法相比较,准确率提高,误报率和漏报率明显下降,可有效应用于实际生产环境中。

关键词:监控图像;电厂雨排口;废弃油污泄漏;改进 Faster R-CNN

中图分类号: TP183; TP391.4 文献标识码: A 国家标准学科分类代码: 520.20

Leakage detection of pollutants at rain drain outlet of power plant based on improved Faster R-CNN

Peng Daogang¹ Wang Yongkun¹ Zhou Yang² Qi Erjiang¹ Gao Yimin¹

(1. College of Automation Engineering, Shanghai University of Electric Power, Shanghai 200090, China;
2. Power Plant of Baoshan Iron & Steel Co., Ltd., Shanghai 201900, China)

Abstract: Aiming at the problem of waste oil leakage from power plant rainwater outlets in monitoring images, a pollutant leakage detection algorithm based on improved Faster R-CNN is proposed. The improved Faster R-CNN detection algorithm first uses ResNet-50 as the backbone network, and builds a multi-scale feature map pyramid structure (FPN) on this basis to achieve information fusion between high-level semantics and low-level semantics, and improve detection accuracy; Secondly, the CIoU loss and DIoU-NMS methods are used to improve the accuracy of bounding box regression in Faster R-CNN; Finally, by introducing Focal Loss function, it solves the problem of unbalanced positive and negative samples in the R-CNN training stage caused by redundant anchor generated by RPN network. The experimental results show that the improved algorithm performs well in real samples, and the accuracy rate reaches 90.2%. Compared with the original Faster R-CNN algorithm, the accuracy rate is improved, and the false positive rate and false negative rate are significantly reduced. It can be effectively used in the actual environment.

Keywords: monitoring images; power plant rainwater outlet; waste oil leakage; improved Faster R-CNN

0 引 言

电厂在运行、施工和检修过程中,会出现污染物未经 处理直接从雨排口排出的情况,这对护厂河及其沿线的 环境保护十分不利。考虑到传统人工巡检方式存在着强 度大、频次高等问题,现阶段主要通过对电厂各个雨排口 安装监控摄像头替代人工巡检,实时性效果相较人工巡 检有了很大提升,但是相应问题并未完全解决,如摄像头 回传的视频或者监控图像仍然需要人工进行查验与复 核^[1]。因此,采取图像监控和目标检测相结合的方式实 现电厂雨排口处污染物泄漏的实时监测具有重要意义。

收稿日期: 2021-06-29 Received Date: 2021-06-29

^{*}基金项目:上海市"科技创新行动计划"高新技术领域项目(21511101800)、上海市科学技术委员会工程技术研究中心项目(14DZ2251100) 资助

近年来,河道、湖泊等不同环境下的水质监测方法主要有布局物理传感器、原位监测仪、漂浮式监测基站等, 此类检测方法使用成本较高且应用于大型水域,并不适 用于雨排口场景^[2]。电厂雨排口处装有监控摄像头,可 以获得雨排口处的图像信息,而基于深度学习的目标检 测算法具有检测质量高、速度快和智能化的特点,可以很 好地提取出图像中被检目标的特征,实现对电厂雨排口 状态信息的判断,很好地契合实际使用场景。

目前,基于锚点(anchor based)的目标检测算法主要 分为单阶检测算法(one stage)和双阶检测算法(two stage)两种类型。one stage 主要有 YOLO^[3-5]、SSD^[6]等, 该类算法直接跳过区域建议网络(region proposal network, RPN),得到最终的目标类别预测概率和位置信 息,降低了网络复杂度,提高了检测速度,但是检测精度 略差,且对于小尺寸目标的检测和定位效果不佳;two stage 主要有区域卷积神经网络(R-CNN)^[7]、Fast R-CNN^[8]、Faster R-CNN^[9]以及各种变体模型^[10],该类算法 先使用合适的方法选择出建议区域,然后对选出的区域 进行预测框的类别概率预测和位置偏移,提高了检测精 度,但是检测速度略差。整体来讲,双阶目标检测算法在 多目标物体检测和小尺寸物体检测上的检测精度更高, 检测效果较为理想。

从实际场景出发,对电厂雨排口处的目标检测需要 考虑雨排口整体状态、吸油毡变化、油污泄漏等情况,被 检测目标的尺度变化较大,且存在小尺度物体、重叠度很 高和包含关系的物体类别。本文提出了一种基于改进 Faster R-CNN的电厂雨排口污染物泄漏图像检测算法, 主要改进如下。

1)基于 B/S 架构搭建电厂雨排口污染物泄漏在线 检测系统,获取雨排口图像数据,构建电厂雨排口污染物 泄漏数据集。

2)针对电厂雨排口中被检目标存在类别间特征不明 显且尺度变化大的特征,本文选择 ResNet-50^[11]作为主 干网络,并在此基础上引入多尺度特征图金字塔结构^[12] (feature pyramid networks, FPN),实现不同深度语义特征 的高度融合。

3)针对电厂雨排口被检目标存在高度重叠和包含关系的情况,引入 Zheng 等^[13]在 AAAI 2020 上提出的 CIoU 损失函数和 DIoU-NMS 方法,优化预测框和目标框之间的距离度量。

4)针对 RPN 网络生成冗余锚点导致样本不均衡的 问题,引入 Focal Loss^[14]损失函数,提高检测准确率。

1 电厂雨排口数据集构建

1.1 搭建电厂雨排口污染物泄漏在线检测系统

电厂雨排口各处高清摄像头通过内部网络将实时图 像存放在独立存储介质中。基于上述实际场景,搭建电 厂雨排口污染物泄漏在线检测系统架构,如图1所示。 首先从管理系统服务器中读取2020年度的历史图像数 据,用于数据集的构建;然后在管理系统服务器和检测系 统服务器之间建立通信连接,实现图像数据的转储,用于 实时检测;最后搭建基于 B/S 架构的可视化界面,实现图 像数据的实时查验和自动预警。





1.2 构建电厂雨排口污染物泄漏数据集

根据对电厂雨排口处的目标检测和预警要求以及电 厂相关巡检人员的指导,最终确定表1所示的电厂雨排 口污染物泄漏数据集标签。表1中以字母g开头的标签 将标注为雨排口全局状态且为大尺度目标,以单词 black 结尾的标签标注在雨排口内部且大多为中小尺度目标, 落叶、漂浮物、鸟类等无关目标不作为本数据集的检测和 标注重点。

表 1 电厂雨排口污染物泄漏数据集标签 Table 1 Label of pollutant leakage data set of power plant rainwater outlet

序号	标注名称	中文含义	备注	
1	whitelinoleumblack	白色油毡块	局部	
2	blacklinoleumblack	黑色油毡块	局部	
3	gblackwater	黑水	全局	
4	gslightlymoatgreasy	护厂河轻微油污	预警依据	
5	gseriousmoatgreasy	护厂河严重油污	预警依据	
6	gwhitelinoleum	白色油毡	全局	
7	gnormal	正常	全局	
8	gmixedwater	浑水	全局	
9	gslightlyoily	轻微油污	全局	
10	gseriousoily	严重油污	全局	

2020 年度电厂雨排口图像数据总计 2 000 000 张左 右,为了构建本文所需要的数据集,每隔 30 min 从原始 存储介质中抽取一张图像组成电厂雨排口污染物泄漏数 据集。数据集中图像数目为 6 100 张,按照 70%、20%、 10%的比例划分为训练集、验证集、测试集。

通过标注软件对电厂雨排口污染物泄漏数据集进行 矩形框标注,并填写对应标签,标注结果以 Pascal VOC 格式保存在 xml 文件中,其数据集标签的样例标注如图 2 所示。





(a) 严重油污 (a) Severe oil pollution

(b) 雨排口正常(b) Normal circumstances

图 2 标注样例 Fig. 2 Example of annotation



Girshick 在 2013 年提出了 R-CNN 算法, 2015 年提出

了 Fast R-CNN 算法,2016 年在原有研究基础上改进,又 提出了 Faster R-CNN 目标检测算法。本文在原有 Faster R-CNN 的基础上进行改进和优化,提出了电厂雨排口污 染物泄漏检测算法,算法的流程框架如图 3 所示(深色部 分为训练流程),该算法主要包含如下 5 个步骤。

步骤1)加载电厂雨排口污染物泄漏数据集,并对图 像进行标准化和缩放等预处理操作。

步骤 2)使用主干网络对批量输入的图片进行语义特征提取。网络最终生成的特征图(feature map)被后续RPN 和感兴趣区域池化层(ROI pooling layer)共享。

步骤 3)生成建议区域。RPN 作为全卷积网络,对输入特征图进行分类、边框回归和 NMS 操作,最后得到系列建议区域(映射到原图的目标矩形框)以及相对应的得分信息。

步骤 4) 感兴趣区域池化层将主干网络生成的特征 图和 RPN 网络生成的建议区域作为输入, 然后从特征图 中提取与建议区域一致的部分并且进行转换使得输出维 度相同。

步骤 5) 维度一致的特征图建议区域(proposal feature maps) 通过展平(flatten)、全连接操作获得每一个区域的 类别概率和位置偏移量,得到更加精准的目标检测框。



图 3 电厂雨排口污染物泄漏检测算法流程

Fig. 3 Flow chart of pollutant leakage detection algorithm for power plant rainwater outlet

2.1 主干网络设计

主干网络可以对输入图片进行语义特征提取。本文

分别选取 VGG16 和 ResNet-50 作为 Faster R-CNN 的主干 网络进行训练,通过性能评价实验(模型对测试集进行9

· 43 ·

次预测,准确率评估指标取均值和标准差,其余指标选取中位数),最后选择评价指标较高的网络作为主干网络。 实验结果如表 2 所示,ResNet-50 的各项评价指标均高于 VGG16^[15],且 ResNet-50 在结构上增加了残差块,增加了 信息流动,也提升了网络深度,使得网络能够提取到更加 抽象、复杂的语义特征,确保训练速度和推理准确性。因 此选择 ResNet-50 作为电厂雨排口污染物泄漏检测算法 的主干网络。

表 2 主干网络的性能评估实验 Table 2 Performance evaluation experiment of backbone network

Backbone Network	$\operatorname{Acc}_{\operatorname{av}}$	${\rm Acc}_{\rm std}$	mAP	Recall	Precision
VGG16	0.621	0.005	0.739	0.634	0.646
ResNet-50	0.801	0.014	0. 792	0.732	0.721

将 ResNet-50 作为主干网络的同时,通过 CNN 解释 器^[16]实现中间特征图可视化, ResNet-50 网络结构和不 同深度网络层输出的特征对比如表 3 所示, 网络浅层部 分输出的特征图富有细节,特征更明显, 可以看出目标物 体的位置和轮廓等局部细节信息; 随着网络结构的不断 深入, 一方面图片尺寸被不断压缩, 通道数不断增加, 另 一方面, 网络中卷积核的感受野^[17]不断加大, 网络输出 的特征图变得抽象, 方便提取出语义层次更高的全局 特征。

表 3 ResNet-50 不同深度网络输出的特征图 Table 3 Feature map of ResNet-50 network output with different depth

Layer Name	Layer	Output Size	Output feature map
Input	Normalize Resize	[1,3,576,1024]	and the second second
Conv1	7×7,64,stride 2 maxpool stride 2	[1,64,144,256]	Karangeneration (CC)
Layer1	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	[1,256,144,256]	
Lyaer2	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 4$	[1,512,72,128]	
Layer3	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 6$	[1,1024,36,64]	
Layer4	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	[1,2048,18,32]	

针对于电厂雨排口污染物泄漏数据集尺度变化较大

和浅层网络比深层网络输出特征图细节更加丰富,特征 更加明显的特点,引入了多尺度图像特征金字塔结构,进 一步提升主干网络特征提取的能力。

由表 3 可知, ResNet-50 主干网络可划分为 Conv1、 Layer1、Layer2、Layer3、Layer4, 且生成5个不同尺度的特 征图(生成顺序自下而上,图4(a)),基于 FPN 原理,本 文采用后面 4 层(Layer1~Layer4) 为基础构建多尺度特 征图金字塔(生成顺序自上而下,横向连接,图4(b))。 首先,Layer4 层网络输出通过池化层(MaxPooling2D)生 成新特征图 P1,特征图尺寸为[1,256,9,16];其次,分别 对主干网络当前层(Laveri, i=4,3,2,1)使用 1×1 卷积进 行降维处理,对生成的新特征图 Pi(i=1,2,3,4)采用上 采样:最后,将新特征图 Pi 和经过 1×1 卷积处理过的 Layeri 层相加在一起再经过一个 3x3 卷积,生成和当前层 尺度一致的新特征图 Pi(i=2,3,4,5),实现效果如图 4 所示,可以看出两种特征图存在较大的差别。通过这种 方式使得当前层在保持尺度不变的基础上,得到了高层 语义的补充,进而增强了主干网络对多尺度特征的表达 和提取能力。从性能评估实验中得出的实验数据(表4) 表明 FPN 网络对算法的评价指标提升很明显。



图 4 ResNet-50 与 FPN 相结合效果



表 4 加入 FPN 结构的性能评估实验 Table 4 Performance evaluation experiment with FPN structure

	1						
Backbone	FDN	Acc	Acc	mAP	Booall	Provision	
Network	FFN	necav	rec _{std}	шлі	necan	1 recision	
ResNet-50		0.850	0.008	0.842	0.770	0.783	
		0.801	0.014	0.792	0.732	0.721	

2.2 基于 RPN 网络的分析与改进

由图 3 可以看出, RPN 网络生成多个建议区域, 主干 网络会输出 5 个不同尺度的增强语义特征图作为 RPN 网络的输入, 由于 RPN 网络采用了基于锚点的多尺度、 多比例设计机制实现不同尺寸建议区域的选取, 所以 RPN 网络需要进行调整以适应输入的多特征图。

原始 Faster R-CNN 算法对 Anchor 的尺寸和宽高比 进行了预设,3个尺度{128,256,512},3种比例{1:1, 1:2,2:1},特征图中的每一个点都会生成9(3×3=9) 个锚点框(anchor boxes)。本文在应用中增加尺度到5 个,分别为{32,64,128,256,512},比例保持不变,特 征图与锚点尺度进行单一对应,所以基于 FPN 地 RPN 网 络特征图中的任意一点共生成 15(3×5=15)个锚点。

RPN 网络的损失函数为.

$$L_{oss}(\{pre_{i}\},\{o_{i}\}) = \frac{1}{N_{clss}} \sum_{i} L_{cls}(pre_{i},pre_{i}^{*}) + \frac{1}{N} \sum_{i} pre_{i}^{*}(pre_{i},pre_{i}^{*}) L_{regress}(o_{i},o_{i}^{*})$$
(1)

其中预测框回归损失采用了 smooth L1 函数, 如式(2) 和(3)所示。

$$L_{regress}(o_i, o_i^*) = smooth_{L1}(o_i - o_i^*)$$
(2)

$$smooth_{L1}(z) = \begin{cases} 0.5z^2, |z| < 1\\ |z| - 0.5, \\ \ddagger (0, 1) \end{cases}$$
(3)

实际上,上述预测框回归损失在计算的时候,存在如 下3个方面的问题。

1)提前假设目标框4个点坐标相互独立,无相关性。 上述公式先求出4个点的损失,然后在进行相加得到最 终的损失输出,实际上目标框4个点坐标因为长度和宽 度的存在具有坐标相关性。

2) 真实目标标注框和目标预测框之间采用的评价指 标为交并比(IoU),这与上述回归损失计算并不是等价 的,多个检测框可能存在相同大小的 smooth "损失,但是 IoU 可能差别很大,如图 5 所示。





IoI = 0.27

 $\|\cdot\|_1 = 9.07$ IoU = 0.27

 $\|\cdot\|_1 = 9.07$ $I_0 U = 0.27$

实线:真实标注框 虚线:物体预测框

图 5 相同 smooth₁₁ 损失,不同 IoU Fig. 5 Same $smooth_{I1}$ Loss, different IoU

3)本文制作的电厂雨排口污染物泄漏数据集中,真 实标注目标框之间存在重叠和包含关系,加大了回归的 难度。

针对回归损失缺点和数据集特征,本文将使用 IoU 作为回归损失函数并且引入 Zheng 等^[13]提出的 CloU (complete IoU loss)损失函数和基于 DIoU(distance-IoU) 的非极大值抑制(NMS)预测框选择方法。

IoU 损失函数的定义如下:

$$Loss = 1 - IoU + \omega(B, G_i)$$
⁽⁴⁾

其中 $\omega(B,G_t)$ 作为预测框和真实目标框之间的惩罚 项。DIoU 的定义如下:

$$Loss_{DloU} = 1 - IoU + \frac{\rho^2(b, g_t)}{c^2}$$
 (5)

其中, $\frac{\rho^2(b,g_t)}{c^2}$ 作为惩罚项可以最小化距离 d,其中

b和 t 表示为 B 和 G, 的中心点, $\rho^2(b,g_t)$ 表示为欧氏距 离,c为最小B和G,之间的对角线长度,示意图如图6 所示。



Fig. 6 DIoU diagram

从定义和图例可以看出, DIoU 考虑了两大因素: 重 叠面积和中心点距离,但是没有考虑到长宽比。针对这 个问题,本文在 DIoU 的基础上增加一个影响因子 $\beta\gamma$,形 成 CIoU 回归损失函数,相应地在训练过程中也缓解了发 散问题。

CloU 的定义如下所示:

$$Loss_{CloU} = 1 - IoU + \frac{\rho^2(b, g_t)}{c^2} + \beta\gamma$$
(6)

其中β作为平衡参数,定义为:

$$\beta = \frac{\gamma}{(1 - IoU) + \gamma} \tag{7}$$

γ用来衡量长宽比一致性,定义为:

$$\gamma = \frac{4}{\pi^2} \left(\tan^{-1} \frac{w_{gt}}{h_{gt}} - \tan^{-1} \frac{w_b}{h_b} \right)$$
(8)

从式(6)可以看出, CloU回归损失能够解决 G, 和 B 重叠或不重叠问题,且两框处于重叠关系时,将使用基于 边界中心点距离和边界框长宽比的尺度信息来进行综合 度量。

RPN 网络会出现同一目标物体存在很多预测框的 问题,而非极大值抑制算法(NMS)采用直接设定阈值的 方法剔除认为是同一目标物体的预测框,存在造成重叠 度较高的物体漏检的可能。

利用 DIoU 可以衡量两框中心距离的特性,采用

DIoU 加 NMS 方法来对预测框进行过滤,对比表 5 的实验数据,本文所使用的方法对网络性能有所提升。

表 5 CloU 损失加 DloU-NMS 和原始 RPN 网络性能评估实验

 Table 5
 CIoU loss plus DIoU-NMS and original RPN network performance evaluation experiment

Backbone Network	CIoU DIoU+NMS	$\mathrm{Acc}_{\mathrm{av}}$	${\rm Acc}_{\rm std}$	mAP	Recall	Precision
ResNet-50		0.864	0.011	0.878	0.784	0.826
+FPN		0.850	0.008	0.842	0.770	0.783

2.3 ROI 检测网络分析与改进

图 3 中, ROI 中 Pooling 层主要用于建议区域和主干 网络输出特征图的匹配。通过分析发现, ROI Pooling 算 法在特征图映射和特征图输出这两个步骤使用了量化操 作,造成了像素偏差且对后续分类和回归定位造成较大 的影响。

为了提高性能,减少误差,本文使用 ROI Align 算法 替代 ROI Pooling。ROI Align 算法在特征图映射时保留 浮点数,在特征图输出时采用双线性差值的方法获得区 域最大像素值。性能评估实验(表 6)证明采用 ROI-Align 能够实现算法性能的提升。

表 6 采用 ROI Align 的性能评估实验 Table 6 Performance evaluation experiment using ROI Align

Backbone	ROI	$\mathrm{Acc}_{\mathrm{av}}$	Acc. Acc.		Becall	Precision	
Network	Align		recostd	man	necan	1 ICCISION	
ResNet-50		0.856	0.013	0.852	0.763	0.811	
+FPN		0.850	0.008	0.842	0.770	0.783	

2.4 R-CNN 的损失函数改进

考虑到 R-CNN 部分所采用正负样本不均衡而导致 模型检测准确率降低的问题,本文引入具有正负样本均 衡能力的 Focal Loss^[18],降低模型训练时负样本损失在 损失函数中的权重,纠正模型梯度的更新方向,使网络能 够学习到更加有用的语义信息。

Focal Loss 的定义为:

$$L_{loc}(p_i, t_i) = \begin{cases} -\alpha (1 - p_i)^{\gamma} \log(p_i), & t_i = 1\\ -(1 - \alpha) p_i^{\gamma} \log(1 - p_i), & t_i = 0 \end{cases}$$
(9)

式中:*p_i* 表示样本预测为正样本的概率;*t_i* 表示真实样本 对应的标签(0或者1);α可以控制正负样本均衡,本文 中取值为 0.33;γ用于使用网络专注于困难样本学习,本 文取值为 2。

性能评估实验数据如表 7 所示, R-CNN 部分引入

Focal Loss 以后,算法的整体性能较改进前有所提升,位置回归和类别预测分数也相应提高。

表 7 引入 Focal Loss 的性能评估实验 Table 7 Performance evaluation experiment with Focal Loss

Backbone	Focal	$\mathrm{Acc}_{\mathrm{av}}$	Acc	A D	D 11	Precision	
Network	Loss		Accestd	MAP	necali		
ResNet-50		0.853	0.007	0.841	0.790	0.801	
+FPN		0.850	0.008	0.842	0.770	0.783	

3 实验结果与分析

本文使用 Java 作为后端接口, VUE 框架作为前端界 面显示,以此为基础构建电厂雨排口污染物泄漏在线检 测系统并进行数据采集、数据集制作和目标实时检测。 在 Windows 10 操作系统上采用 PyTorch^[19]作为深度学习 框架;使用 NVIDIA GeForce RTX2080 的 GPU 进行电厂 雨排口污染物泄漏检测算法的并行加速训练。

3.1 数据集训练过程

由于本文所做数据集拍摄角度单一,所以需要进行 数据增强^[20-21]来减少数据集的不相关特征。本文采用线 上增强方法,进行小批量数据的图像增强操作。

本文采用的数据增强方式有改变尺寸,随机裁剪,水 平或竖直翻转,亮度、对比度、颜色改变等,具体实现效果 如图 7 所示。



Fig. 7 Data enhancement rendering

本文采用迁移学习的方法来训练网络,训练过程中 设置 batch size 为4,epoch 为15;采用带动量参数的随机 梯度下降算法^[22](stochastic gradient descent,SGD),动量 参数为0.9,学习率为0.005,在开始第1轮训练阶段启 用学习率预热(warmup)策略,并设置等间隔(5个epoch)调整学习率。

电厂雨排口污染物泄漏检测算法的学习率和损失函数如图8所示,可以看出在进行16000次迭代之后,损失函数曲线已经趋近于收敛;当迭代到5000次和11000次前,曲线一直趋于震荡不下降的状态,在经过学习率调整之后,损失函数曲线继续下降直至收敛。



Fig. 8 Loss function curve and learning rate curve

3.2 算法测试

通过对不同电厂雨排口状态进行测试,测试效果如 图 9 所示,检测结果和巡检人员所观察到的异常状态保 持一致,检测效果优良;检测结果包含目标类别和置信度 分数,且模型可以准确检测出雨排口的状态信息和处于 雨排口之内的吸油毡块。图 9(d)中由于雨排口处于严 重油污状态且内部含有大量黑色油毡块,所以产生了很 多的检测框,经查验,图 9 中油毡块数和检测框数目等 同,检测结果无误。



Fig. 9 Test results of rain outlet status in power plant

3.3 结果分析

使用 Faster R-CNN 算法对数据集进行同参数训练, 并且采用平均精度均值(mAP)作为评价指标,每一个 epoch 后计算一次 mAP 值,训练结束后,对比改进前后两 种算法的检测效果。如图 10 所示,以第 15 个 epoch 作为 标准,原始 Faster R-CNN 算法全类平均准确率为 80%,改 进后算法的全类平均准确率为 88.3%。相较之下,改进 后算法的评估指标有了 8.3%的性能提升。



Fig. 10 mAP values of pollutants at the outlet of power plant before and after improvement

本文以 ResNet-50 作为主干网络,在此基础上引入多 尺寸 FPN 结构、修改 RPN 回归损失为 CloU 损失、替换 NMS为 DIoU 加 NMS、使用 ROI Align 算法替代 ROI Pooling、引入 Focal Loss 平衡 R-CNN 阶段的正负样本这 几种方法,最后为了验证上述几种方法对模型的影响,进 行消融实验,数据汇总如表8所示。实验数据表明使用 ResNet 作为主干网络能够极大提升算法性能:使用 ResNet 加 FPN 结构的算法在进行某一方面的改进时,会 出现指标下降的情况,但是整体性能稳步提升。后续消 融实验的数据表明,改进 Faster R-CNN 检测算法在引 入 FPN 结构以后能够极大提高整体算法的性能,可以 证明 FPN 结构强化了特征图的特征表达能力;而引入 CIoU与 DIoU+NMS 方法和 POI Align 方法能够有效提 升物体的检测精度,从而提升算法性能;Focal Loss 方法 通过平衡 R-CNN 阶段的正负样本来提升算法的性能且 提升有限。整体实验表明基于改进 Faster R-CNN 的电 厂雨排口污染物泄漏检测算法与改进前相比,综合性 能有了明显提升。

图 11 所示为 Faster R-CNN 算法改进前后的检测效 果,图 11(a)为原算法的检测效果图,图 11(b)为改进后 算法的检测效果图。整体来看,检测目标的类别置信度 提高,这是由于改进后的主干网络添加了 FPN 结构使得 网络可以提取到更加丰富的特征语义;对比第2组和第3 组图片,可以看到改进后算法对于重叠物体的检测效果 有了明显的提升,这是因为采用了 CloU 函数作为 RPN 部分的损失函数、使用 DloU 加 NMS 对原始 NMS 进行了 代替,对重叠物体检测能力大大提高;对比第4组的图 片,改进后算法的误检测概率大大降低,这是由于在 R-CNN 部分使用 Focal Loss 均衡了正负样本,使得物体检 测变得更加的精确,训练也更快速和稳定。综上所述,改 进后算法在预测框位置信息和类别信息上的预测能力上 大大提升,减少了误检率。 实验粉据汇首

±ο

秋 6 关键数据汇芯											
Table 8 Summary of experimental data											
Backhone			改进力	方法		评价指标					
Detection networ	Detection network	FDN	CIoU	ROI	Food Loog	$\mathrm{Acc}_{\mathrm{av}}$	$\mathrm{Acc}_{\mathrm{std}}$	mAD	Recall	Precision	
Inetwork		FFIN	DIoU+NMS	Align	Focal Loss			mAP			
VGG16	Faster R-CNN					0.621	0.005	0.739	0.634	0.646	
ResNet-50	Faster R-CNN					0.801	0.014	0.792	0.732	0.721	
ResNet-50	Faster R-CNN	\checkmark				0.850	0.008	0.842	0.77	0.783	
ResNet-50	Faster R-CNN	\checkmark	\checkmark			0.864	0.011	0.878	0.784	0.826	
ResNet-50	Faster R-CNN	\checkmark				0.856	0.013	0.852	0.763	0.811	
ResNet-50	Faster R-CNN	\checkmark				0.853	0.007	0.841	0.79	0.801	
ResNet-50	Faster R-CNN		\checkmark	\checkmark		0.862	0.017	0.829	0.779	0.731	
ResNet-50	Faster R-CNN	\checkmark		\checkmark	\checkmark	0.871	0.009	0.848	0.781	0.829	
ResNet-50	Faster R-CNN	\checkmark	\checkmark		\checkmark	0.882	0.021	0.867	0.801	0. 786	
ResNet-50	Faster R-CNN	\checkmark	\checkmark	\checkmark		0.889	0.005	0. 855	0.796	0.821	
ResNet-50	Faster R-CNN					0.902	0.003	0.883	0.826	0.848	









(a) 改进前算法检测效果 (a) Algorithm detection effect before improvement



(b) 改进后算法检测效果 (b) Algorithm detection effect before improvement

图 11 Faster R-CNN 改进前后算法检测效果

Fig. 11 Effect picture of Faster R-CNN algorithm before and after improvement

4 结 论

本文提出了一种基于改进 Faster R-CNN 的电厂雨排 口污染物泄漏图像检测算法,通过线上增强的方式对输 入图片进行处理,解决数据集拍摄角度单一而引入不相 关特征的问题;主干网络采用 ResNet-50,并在此基础上 构建多尺度特征图金字塔,实现了数据集特征的融合; RPN 网络部分使用 CloU 作为回归损失并且使用 DloU 加 NMS 对生成的建议区域进行后续处理,解决了重叠区 域目标检测效果差的问题;最后使用 ROI Align 提升预测 框的定位精度,使用 Focal Loss 均衡样本提高训练效果。 本文所研究方法可以实现电厂雨排口在线实时检测,检 测效果和巡检人员判读结果一致,未来可以通过调整雨 排口处监控摄像头位置进而实现数据集的扩充,对其他 电厂雨排口实现在线检测。

参考文献

杨洋.视频监控系统智能分析与图像质量诊断[D]. [1] 兰州:兰州理工大学,2020.

YANG Y. Intelligent analysis and image quality diagnosis of video monitoring system [D]. Lanzhou: College of Electrical and Information Engineering Lanzhou University of Technology, 2020.

- [2] ALILOU H, NIA A M, SARAVI M M, et al. A novel approach for selecting sampling points locations to river water quality monitoring in data-scarce regions [J]. Journal of Hydrology, 2019, 573: 109-122.
- REDMON J, DIVVALA S, GIRSHICK R, et al. You [3] only look once: Unified, real-time object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- 郭璠,张泳祥,唐琎,等. YOLOv3-A:基于注意力机制 [4] 的交通标志检测网络[J]. 通信学报, 2021, 42(1): 87-99.

GUO F, ZHANG Y X, TANG J, et al. YOLOv3-A: A traffic sign detection network based on attention mechanism [J]. Journal on Communications, 2021, 42(1):87-99.

[5] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. ScaledYOLOv4: Scaling cross stage partial network [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13029-13038.

- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision, Springer, 2016: 21-37.
- GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [8] GIRSHICK R. Fast R-CNN [C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [9] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2015, 28: 91-99.
- [10] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [12] 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.
- ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C].
 Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [15] QASSIM H, VERMA A, FEINZIMER D. Compressed residual-VGG16 CNN model for big data places image recognition [C]. 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), 2018: 169-175.
- [16] WANG Z J, TURKO R, SHAIKH O, et al. CNN explainer: Learning convolutional neural networks with interactive visualization [J]. IEEE Transactions on Visualization and Computer Graphics, 2020.
- [17] LUO W, LI Y, URTASUN R, et al. Understanding the effective receptive field in deep convolutional neural networks [C]. Proceedings of the 30th International Conference on Neural Information Processing Systems,

2016: 4905-4913.

[18] 陶晓天,何博侠,张鹏辉,等. 基于深度学习的航天密 封圈表面缺陷检测[J]. 仪器仪表学报,2021,42(1): 199-206.
TAO X T, HE B X, ZHANG P H, et al, Surface defect detection of aerospace sealing rings based on deep learning [L] Chinese Journal of Scientific Instrument

detection of aerospace sealing rings based on deep learning [J]. Chinese Journal of Scientific Instrument, 2021,42(1):199-206.

- PASZKE A, GROSS S, MASSA F, et al. Pytorch: An imperative style, high-performance deep learning library [J]. Advances in Neural Information Processing Systems, 2019, 32: 8026-8037.
- [20] BUSLAEV A, IGLOVIKOV V I, KHVEDCHENYA E, et al. Albumentations: Fast and flexible image augmentations [J]. Information, 2020, 11(2): 125.
- [21] 张晓丽,董昱. 面向铁路货车车号定位的 Faster R-CNN 卷积神经网络[J]. 电子测量与仪器学报, 2020, 34(10):65-73.
 ZHANG X L, DONG Y. Faster R-CNN convolutional neural network for the location of freight train number[J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(10):65-73.
- [22] GOWER R M, LOIZOU N, QIAN X, et al. SGD: General analysis and improved rates [C]. International Conference on Machine Learning, 2019: 5200-5209.





彭道刚(通信作者),分别 2001 年和 2004 年于华北电力大学获得学士学位和硕 士学位,2009 年于同济大学获得博士学位, 现为上海电力大学教授,硕士生导师,主要 研究方向为智能发电、综合智慧能源与能源 互联网等。

E-mail:pengdaogang@126.com

Peng Daogang (Corresponding author) received his B. Sc. degree in 2001 from North China Electric Power University, received his M. Sc. degree in 2004 from North China Electric Power University, received his Ph. D. degree in 2009 from Tongji University. Now he is a professor and master's supervisor in Shanghai University of Electric Power. His main research interests include intelligent power generation, integrated intelligent energy and energy Internet.



王永坤,2018年于太原工业学院获得 学士学位,现为上海电力大学硕士研究生, 主要研究方向为深度学习和目标检测。 E-mail:wang_yongkun@yeah.net

Wang Yongkun received his B. Sc. degree in 2018 from Taiyuan Institute of

Technology. Now he is a M. Sc. candidate in Shanghai University of Electric Power. His main research interests include deep learning and object detecting.