

DOI: 10.13382/j.jemi.B2306325

融合特征增强的轻量化罐道缺陷检测算法*

赵佰亭¹ 吴俊东¹ 贾晓芬^{1,2}

(1. 安徽理工大学电气与信息工程学院 淮南 232001;

2. 安徽理工大学省部共建深部煤矿采动响应与灾害防控国家重点实验室 淮南 232001)

摘要:为了解决矿井复杂环境下,缺陷特征提取不充分问题,融合特征增强和级联注意力机制提出一种快速智能的罐道缺陷识别算法 RDM-YOLOv5,旨在解决人工巡检效率低的现状。首先,为了提高主干网络特征图信息表征能力,设计特征增强模块 RLKM,它通过重参数化大内核卷积增强主干网络对目标特征的提取能力,并且有效降低模型参数量;然后,经过主干网络提取到高低层级特征后,由设计的级联注意力机制 DCAM 进一步挖掘缺陷目标的深层语义信息,显著增强小目标的特征信息;最后,为提升检测精度的同时保障检测网络的轻量化,在特征增强网络中引入轻量级卷积 GSConv,在保持模型检测准确性的同时降低计算成本。实验结果表明,相较于 YOLOv5s, RDM-YOLOv5 的检测精度和速度分别提高 3.7%、11.4%,模型参数量减少 15.4%。它能基本满足实际应用中精准识别和快速定位罐道表面缺陷的需求。

关键词: 罐道缺陷;特征融合;目标检测;YOLOv5s;轻量化

中图分类号: TP206;TN911.73

文献标识码: A

国家标准学科分类代码: 520.2060

Lightweight guide defect detection algorithm based on feature enhancement

Zhao Baiting¹ Wu Jundong¹ Jia Xiaofen^{1,2}

(1. China Institute of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China;

2. State Key Laboratory of Mining Response and Disaster Prevention and Control in Deep Coal Mines, Anhui University of Science and Technology, Huainan 232001, China)

Abstract: In order to solve the problem of insufficient defect feature extraction in the complex mine environment, a fast and intelligent guide defect recognition algorithm RDM-YOLOv5 is proposed by integrating feature enhancement and cascading attention mechanism, aiming to solve the current situation of low efficiency of manual inspection. Firstly, in order to improve the information representation ability of the backbone network feature map, the feature enhancement module RLKM is designed, which enhances the backbone network's ability to extract target features through reparameterized large kernel convolution, and effectively reduces the amount of model parameters. Then, after extracting the high-level and low-level features through the backbone network, under the action of the cascaded attention mechanism DCAM composed of the designed channel attention mechanism and coordinate attention mechanism, the deep semantic information of the defective target is further mined, and the feature information of the small target are significantly enhanced. Finally, in order to improve the detection accuracy while ensuring the lightweight of the detection network, a lightweight convolution GSConv is introduced into the feature enhancement network to reduce the computational cost while maintaining the accuracy of model detection. The experimental results show that compared with YOLOv5s, the detection accuracy and speed of RDM-YOLOv5 are increased by 3.7% and 11.4%, respectively, and the number of model parameters is reduced by 15.4%. It can basically meet the needs of accurate identification and rapid location of guide surface defects in practical applications.

Keywords: guide defect; feature fusion; target detection; YOLOv5s; lightweight

收稿日期: 2023-03-09 Received Date: 2023-03-09

* 基金项目: 国家自然科学基金面上项目(52174141)、安徽省自然科学基金面上项目(2108085ME158)、安徽高校协同创新项目(GXXT-2020-54)、安徽省重点研究与开发计划项目(202104a07020005)资助

0 引言

罐道是立井提升系统的重要组成部分,罐道与滑动罐耳通过摩擦接触实现导向,常年运行状态导致罐道容易出现磨损、裂痕和错位等缺陷^[1]。《煤矿安全规程》中明确规定,罐道磨损超过规定的阈值时,将会影响提升系统的安全运行,一旦发生事故将造成无法估量的损失^[2]。矿井生产环境复杂,人工检测是目前立井罐道表面缺陷的主要检测方式,此方式影响煤矿生产效率,同时受恶劣的井筒环境和人工主观因素影响,难以保证检测精度^[3]。因此,实现罐道表面缺陷高精度检测对于矿井安全生产具有重要意义。

目前,相关领域学者根据现有研究已经提出多种罐道缺陷诊断方法。文献[4]融合降噪和概率神经网络技术提出一种新的罐道缺陷检测网络,提高了罐道故障诊断性能;文献[5]基于机器视觉提出的罐道缺陷图像采集和处理方案,解决了检测信号实时性差等问题。上述方法能够实现缺陷识别,但是实际部署困难,难以满足实际工况。

随着以卷积神经网络为代表的深度学习模型在诸多计算机视觉领域成功应用,不少基于深度学习的缺陷检测方法也广泛应用在各种工业领域中^[6]。基于深度学习的方法主要有以 YOLO^[7]、SSD^[8] 为代表的一阶段网络和以 Faster R-CNN^[9] 为代表的两阶段网络。两阶段检测网络的流程是首先通过主干网络获取输入图像的特征图,利用区域生成网络计算锚框置信度,得到 Proposal 区域。然后对所获区域进行特征图 ROIpooling 操作后输进网络,经过对检测结果进行调整,得到最终的缺陷定位和分类结果。He 等^[10] 基于 Faster R-CNN 提出一种带钢表面缺陷检测网络,使用多尺度特征图替换主干网络的多级特征图,在 NEU-DET 数据集上实现了 82.3% 的 mAP;单明陶等^[11] 使用 ResNeXt50 作为主干网络,并加入改进特征金字塔结构,提出一种改进 Faster R-CNN 的刹车零件缺陷检测算法;文献[12]对铝材进行缺陷检测时,在主干网络加入 FPN 结构,并使用 ROI Align 代替 ROI Pooling,最后利用 K-means 聚类生成锚框,增强了缺陷定位效果。虽然以上算法对模型的检测性能有所提升,但是检测精度和速度之间没有达到很好的平衡,难以满足工业生产的需要。

单阶段网络利用整幅图作为网络的输入,直接在输出层回归边界框(bounding box)位置以及所属的类别。Yang 等^[13] 提出了一种多层次、端到端的快速钢轨表面缺陷检测方法,结合 YOLOv2 和分割算法,提升了模型的速度和准确度;刘涛等^[14] 基于 YOLOv4 提出一种印刷电路板缺陷检测网络模型,用 GhostNet 代替主干网络,提升了

模型检测效果;马燕婷等^[15] 在识别带钢表面瑕疵时,在 YOLOv5 主干网络中引入自注意力机制并结合 BiFPN 结构,有效提升带钢瑕疵检测的精度;文献[16]在对航拍绝缘子自爆缺陷定位时,使用深度可分离卷积提升检测速度,并构建四维度检测层,实现了结构的优化。尽管单阶段的检测算法兼具了速度的优势,但是检测精度依然不能满足实际需求,尤其是像立井罐道缺陷检测这样的特定场景。

综上,为了高精度、高效率完成罐道表面缺陷检测任务,必须强化检测模型对特征的表征能力,凸显模型对细小缺陷特征的提取能力,还需要降低模型参数量,保证速度与精度兼得,为此融合有效感受野和级联注意力机制等策略提出罐道小目标缺陷检测模型 RDM-YOLOv5。

1 罐道缺陷检测模型 RDM-YOLOv5

1.1 设计思想

罐道沿井筒轴线固定在罐道梁或悬挂在井架上,确保罐笼在井筒中平稳、安全运行。为了避免事故,防患于未然,对罐道上可能出现的缺陷都应该是发现在萌芽状态,换言之,应在缺陷刚出现时就能及时进行准确检测。磨损表现为块状,由表征块大小的拓扑结构和表征磨损程度的语义信息共同体现。裂痕、错位两种缺陷的形状和程度分别表现为小线条构成的浅层几何信息、深度和纹理等体现的深层语义信息。因此,无论是哪种缺陷,都必须检测出浅层特征表达的几何信息,只要能提取到缺陷的拓扑结构,就能保证发现细微的小目标缺陷。然后,再借助深层语义信息鉴别缺陷的类型及缺陷程度,通过充分融合浅层和深层特征,提高对罐道缺陷的检测精度。

YOLOv5 是目前 YOLO 系列最轻的检测算法,适用于解决实际问题。作者以 6.0 版本的 YOLOv5s^[17] 为基础,设计罐道表面缺陷的检测模型,旨在保证模型轻量化,促进智能检测算法的落地应用。

1.2 RDM-YOLOv5 网络结构

鉴于上述思考,设计图 1 的网络结构 RDM-YOLOv5,它由主干网络,特征增强网络和检测头 3 个模块组成,其中包含基础组成模块 Conv、C3 和 SPPF。Conv 为基础卷积模块,对图像提取特征;C3 模块增加网络深度,强化特征提取;SPPF 为空间金字塔池化,用来整合特征。针对 YOLOv5s 主干网络特征图信息表征能力弱,网络模型复杂度高等问题,进行如下 3 个方面改进:

1) 构建特征增强模块 RLKM,用其替换主干网络的高层和特征增强网络末端的 C3 结构,旨在提升模型的特征表征能力并减少模型参数量,达到提升检测精度和检测速度的目的。

2) 在特征增强网络中提出注意力机制 DCAM, 它级联通道注意力机制 D-ECA 与坐标注意力机制 (coordinate attention, CoordA), 来抑制背景噪声、聚焦关键特征, 通过增加对输入特征通道和位置信息的利用率, 给予重要

的小目标特征更高的权重。

3) 引入轻量化卷积 GSConv, 旨在借助深度可分离卷积减轻模型复杂度、减少冗余信息, 保持精度的同时降低计算成本。

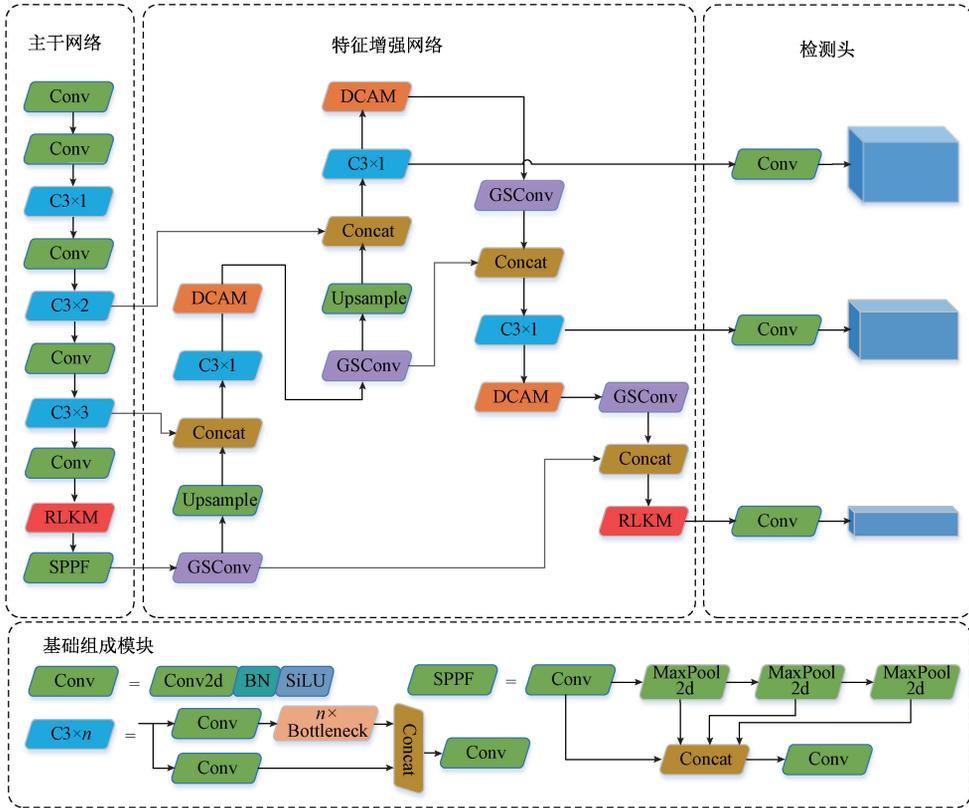


图 1 RDM-YOLOv5 网络结构
Fig. 1 RDM-YOLOv5 network structure

1.3 特征增强模块 RLKM 设计

罐道缺陷类间的形状差异较大, 如磨损缺陷的类内形状特征比较明显。随着网络层数的加深, 输出特征经过多次卷积下采样操作, 语义信息表征能力增强, 但是空间几何特征信息逐渐缺失。必须解决网络加深引起的浅层几何信息丢失问题, 增强模型对缺陷的空间几何特征提取能力, 才能提高对缺陷的检测精度。

针对多分支模型训练性能高但部署复杂, 单分支模型推理速度快但训练性能差的问题, Ding 等^[18] 提出结构重参数化方法。在训练时采用一个多分支结构模型, 然后通过等价转换将多分支转换为单分支模型, 最后部署单路模型, 这样就能同时利用多分支模型训练的高性能和单路模型推理时的优势。RepLKNet^[19] 利用结构重参数化思想, 使用超大内核的 DW 卷积, 再通过小卷积核重参数化进行模型优化, 实现模型有效感受野 (effective receptive field, ERF)^[20] 的增大, 同时降低计算量。可见, 与小核 CNNs (convolutional neural networks) 相比, 大核

CNNs 具有更大的 ERF 和更高的形状偏差。ERF 的计算公式如下:

$$\sqrt{\text{Var}[S_n]} = \sqrt{n} \sqrt{\sum_{m=0}^{k-1} \frac{m^2}{k} - \left(\sum_{m=0}^{k-1} \frac{m}{k}\right)^2} = \sqrt{\frac{n(k^2 - 1)}{12}} = O(k\sqrt{n}) \quad (1)$$

式中: k 为卷积核大小, n 为网络层数, m 为卷积核像素个数, $\sqrt{\text{Var}[S_n]}$ 为 ERF 大小。

根据 ERF 的计算公式可知, ERF 的大小随内核大小线性增长, 相比直接加深网络层数, 采用单个大卷积核可以更高效地提升 ERF。大感受野的特征提取能力可以弥补多次小卷积核下采样导致的几何信息丢失的缺陷, 受 RepLKNet 的启发, 通过重参数化 27×27 的大内核卷积块设计图 2 所示特征增强模块 RLKM, 主路分支上先经过一个标准的卷积层, 再通过 RepLK Block; 另一路分支仅经过一个基本卷积模块; 最后将两支进行 Concat 操作后, 送入标准卷积层得到输出。RepLK Block 使用了捷径分

支和 27×27 大内核的高效 DW 卷积,为模型提供足够的 ERF 和空间信息聚合能力,每个 DW 大内核卷积使用一个 5×5 的卷积核进行结构重参数化。为了提供更多的非线性和跨通道信息,DW 卷积前后均使用 1×1 卷积,因为模型的代表能力与深度密切相关,使用 1×1 卷积层可以增加深度。

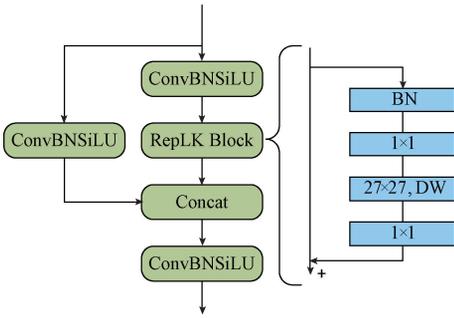


图 2 RLKM 模块结构

Fig. 2 RLKM module structure diagram

高层网络输出特征分辨率低,语义信息表征能力强,但几何细节信息缺乏,RLKM 的大核卷积与小卷积核相比,具有更大的 ERF 和更高的几何细节信息。网络末端的特征图分辨率较低,在低分辨率的特征映射上应用 RLKM 可以减少昂贵的计算和内存成本。因此,在主干网络高层特征图中加入 RLKM,可以提高模型对多特征的代表能力,有助于实现对立并罐道缺陷的更全面检测。特征增强网络末端对应的是最后一层检测头,该层检测头由深层次、低分辨率的小特征图生成,主要用来

检测大目标缺陷。罐道缺陷的大目标比小目标具有更强的形状偏差,在最后一层检测头前应用 RLKM 可以使模型更有效关注目标的几何形状特征。经过 32 倍降采样输出的特征图分辨率较低,在低分辨率的大目标检测头上应用 RLKM 可以进一步降低计算量和模型复杂度。综上,仅在主干网络高层特征图和特征增强网络末端应用 RLKM 模块。

1.4 注意力机制 DCAM 设计

罐道处于地下,背景环境比较复杂,单一的平均池化只能保留特征图整体特征,凸显背景信息的同时丧失了对细节信息的获取,而罐道裂痕和错位的纹理特征比较突出,细节信息的丢失会造成裂痕等细小目标缺陷的误检。因此,非常有必要利用最大池化提取特征图上的纹理、细节特征。

注意力机制是受人类视觉机制的启发,目的是让网络学习关注特征、忽略不相关信息。在深度学习网络中,注意力机制针对输入特征图在各维度上重要性所占的比重不同,动态赋予通道及空间维度的权重,从而重新校准特征来提高网络的表达能力。高效通道注意力机制 (efficient channel attention, ECA)^[21]通过加强通道之间的信息交互强化有效特征,CoordA^[22]通过在通道注意力中加入位置信息,使网络获取更大区域信息并避免引起大计算消耗,它可以灵活嵌入到轻量级网络中。作者沿用 ECA 的局部跨通道交互策略和自适应一维卷积结构设计通道注意力机制 D-ECA,与 CoordA 组成级联注意力机制 DCAM,整体结构如图 3 所示。

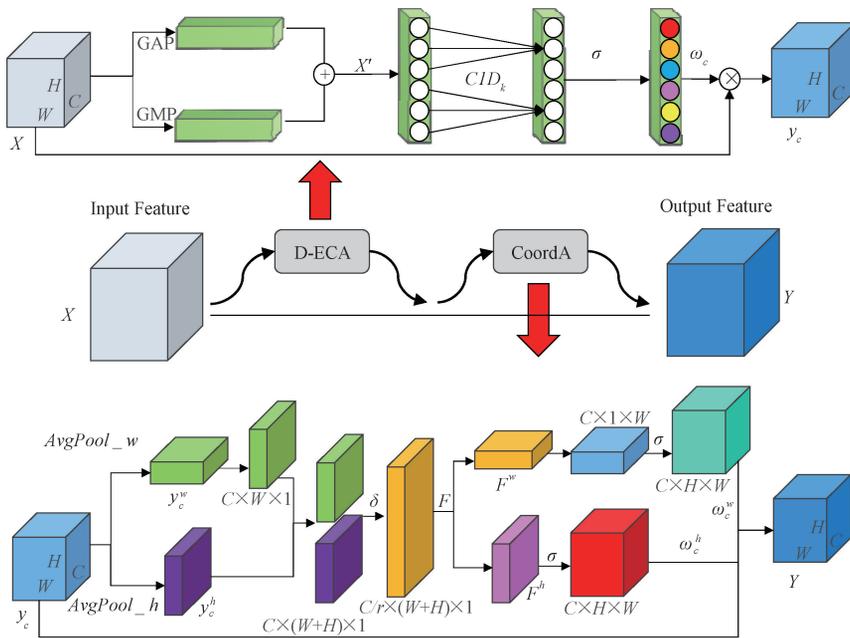


图 3 注意力机制模块 DCAM

Fig. 3 Attention mechanism module DCAM

D-ECA 对输入特征图 X 同时进行 GAP 平均池化和 GMP 最大池化操作后,按位相加两路特征图,再进行局部跨通道交互和一维动态卷积操作,得到特征图对应通道位置上的不同权重,从而更精准获得特征图在通道维度上的注意力信息。对于 D-ECA 输出的特征图,为了减轻由 2D 全局池化引起的位置信息丢失,借助 CoordA 将通道注意力分解为两个并行的一维特征编码过程,以有效地将空间坐标信息整合到生成的注意力图中。具体来说,就是先使用两个一维全局平均池化分别将沿垂直和水平方向的输入特征聚合成两个独立的方向感知特征图,然后将两个嵌入了方向特定信息的特征图通过维度拼接、降维等操作编码成两个注意力图,每个图都捕获输入特征图沿一个空间方向的依赖关系,从而使位置信息保留在生成的注意力图中,最后通过加权融合的方式将两个注意力图应用于输入特征图,使模型更好的定位感兴趣目标。

DCAM 对输入特征图 X 的特征提取过程为:

$$X' = GAP(X) \oplus GMP(X) \quad (2)$$

$$\omega_c = \sigma[C1D_k(X')] \quad (3)$$

$$y_c = \omega_c \otimes X \quad (4)$$

$$y_c^h = AvgPool_h(y_c), y_c^w = AvgPool_w(y_c) \quad (5)$$

$$F = \delta(f_{conv}^{1 \times 1} [y_c^h, y_c^w]) \quad (6)$$

$$\omega_c^h = \sigma(f_{conv}^{1 \times 1} F^h), \omega_c^w = \sigma(f_{conv}^{1 \times 1} F^w) \quad (7)$$

$$Y = y_c \otimes \omega_c^h \otimes \omega_c^w \quad (8)$$

式中: \oplus 表示特征图按位相加操作, $C1D_k$ 表示卷积核大小为 k 的一维卷积操作, k 值随训练自适应变化, σ 表示 sigmoid 激活函数, \otimes 表示特征图按位相乘操作, X' 表示池化操作后的特征图, ω_c 和 y_c 分别表示 D-ECA 学习的权重、输出的特征图 (CoordA 的输入), $AvgPool_h$ 和 $AvgPool_w$ 分别表示水平、垂直方向的平均池化操作, y_c^h 和 y_c^w 分别为平均池化得到的高度为 h 宽度为 w 的特征图, $f_{conv}^{1 \times 1}$ 表示 1×1 卷积操作, δ 表示非线性激活函数 h_swish , $[\]$ 为沿空间维度的 concat 操作, F 为对空间信息在水平方向和垂直方向上进行编码的中间特征映射, F^h 和 F^w 分别为沿水平和垂直方向维度分解后 2 个单独张量, ω_c^h, ω_c^w 分别表示水平和垂直方向坐标注意力机制学习到的权重, Y 表示 DCAM 输出的特征图。

为增加对输入特征通道和位置信息的利用率,将 DCAM 嵌入到特征增强网络中,旨在增强模型细小缺陷特征的提取能力,提升模型多尺度检测能力。

1.5 轻量级卷积 GSConv 的引入

速度和精度对于缺陷检测网络同样重要。现有的轻量级模型,使用深度可分离卷积操作大大提高了检测速度,但是牺牲了不少精确度。随着网络层数的加深,输入特征图逐步从空间信息向通道传输,如果全局都用它替

代传统卷积,会显著增加推理时间;而特征图传输到特征增强网络时,通道维度达到最大,空间维度达到最小,不需要进行变换。为了使深度可分离卷积输出效果尽可能接近传统卷积,作者仅在特征增强网络中使用 GSConv^[23] 替代传统卷积。GSConv 处理特征图可以减少冗余信息,不需要压缩,使得注意力模块效果更好,与传统卷积相比,GSConv 使用深度可分离卷积可以大大降低模型参数量,以输入 $640 \times 640 \times 3$ 和输出 $640 \times 640 \times 32$,卷积核 3×3 为例,则深度可分离卷积与传统卷积的参数量对比:

$$ratio_p = \frac{C_1 \times K_1 \times K_2 + 1 \times 1 \times C_1 \times C_2}{C_1 \times K_1 \times K_2 \times C_2} = \frac{1}{C_2} + \frac{1}{K_1 \times K_2} \approx \frac{1}{7} \quad (9)$$

其中, C_1 和 C_2 为输入和输出特征图的通道数, $K_1 \times K_2$ 为卷积核大小, $C_1 \times K_1 \times K_2 \times C_2$ 为传统卷积层的参数量, $C_1 \times K_1 \times K_2 + 1 \times 1 \times C_1 \times C_2$ 为深度可分离卷积层的参数量, $ratio_p$ 为深度可分离卷积层与标准卷积层参数量的比值。

由式(9)可见,相同条件下深度可分离卷积的参数量仅为传统卷积的 $1/7$ 。因此,本文应用 GSConv 在保证精度的情况下大大降低模型参数量,提升网络模型在罐道缺陷检测中的检测速度。

2 实验结果与分析

2.1 实验环境与参数设定

实验在 CPU 型号为 Intel Core i9-10900k CPU @ 3.70 GHz, GPU 型号为 NVIDIA GeForce RTX 3080Ti, 12 G 显存, 32 G 内存的 PC 上进行,深度学习环境搭载于 CUDA11.1 和 Pytorch1.9.0 框架,Python 版本为 3.8。实验相关参数设置如表 1 所示。

表 1 实验相关参数

Table 1 Relevant parameters of the experiments

名称	参数	名称	参数
批处理	16	动量值	0.937
图片尺寸	640	学习率	0.01
训练周期	150	衰减系数	0.0005

2.2 数据集构建与评价指标

由于立井罐道还没有公开的标准缺陷数据集,本文依托深部煤矿采动响应与灾害防控国家重点实验室搭建如图 4 所示立井罐道模拟平台,构建罐道表面缺陷数据集。

使用工业相机采集不同光照强度下 900 张罐道缺陷数据,模拟罐道常见的磨损 (abrasion)、裂痕 (crack)、错位 (dislocation) 3 种缺陷,对每一类别缺陷,划分 3 个光照



图 4 立井罐道模拟平台

Fig. 4 Shaft guide simulation platform

强度 (10 lux、20 lux 和 50 lux) 进行采集,每个光照梯度采集 100 张图片数据。利用 labeling 标注软件对缺陷部位进行标注。为避免类别数量不平衡对算法产生不必要的影响,通过旋转、镜像、裁剪等方法扩充数据集,最终将数据集扩充至 3 000 张,并按照 8 : 2 的比例划分训练集和验证集。图 5 为正常罐道表面和常见的 3 种罐道缺陷类型。

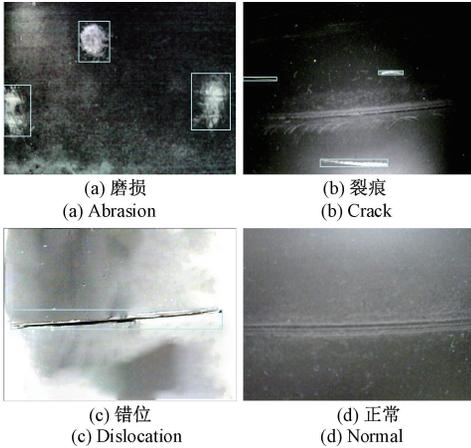


图 5 正常罐道和常见 3 种缺陷

Fig. 5 Normal guide and three common defects

区别于缺陷分类任务,仅通过精确率作为缺陷检测的评价指标是不够的,还需要考虑模型的检测速度。因此,采用平均精度均值 (mean average precision, mAP) 来评价整个网络模型的性能,文中选取 mAP@ 0.5 (即 IoU 为 0.5 时的 mAP) 作为衡量指标, IoU 计算的是“预测边框”和“真实边框”的交叠率,即两者交集与并集的比值;同时引入模型检测速度 (frames per second, FPS), FPS 为每秒检测图片的数量,用来评估模型的检测速度;采用参数量 (parameters, Params) 作为反应模型大小的评价指标。计算公式如下:

$$P = \frac{TP}{TP + FP} \tag{10}$$

$$R = \frac{TP}{TP + FN} \tag{11}$$

$$AP = \int_0^1 P(R) dR \tag{12}$$

$$mAP = \frac{1}{n} \sum_{i=0}^n AP(i) \tag{13}$$

其中, P 为精确率, R 为召回率和平均精度 AP , 真正例记为 TP , 假正例记为 FP , 假反例记为 FN 。 $P(R)$ 为精确率-召回率曲线, AP 为 P-R 曲线下的面积。

2.3 改进效果验证

为了论证 RDM-YOLOv5 的缺陷检测效果, 利用上述数据集在表 1 的实验参数限定下对它和 YOLOv5s 进行训练和测试, 结果如图 6 所示。由图 6(a) 的准确率曲线可知, RDM-YOLOv5 的准确率在迭代 25 次时趋于稳定, 且准确率相比原 YOLOv5s 网络有明显提升。损失函数采用原 YOLOv5 损失函数, 包含分类损失、定位损失和置信度损失, 分类损失采用的是二元交叉熵损失函数 BCE_loss。图 6(b) 的分类损失曲线显示, RDM-YOLOv5 网络的收敛效果更好, 模型的拟合效果也具有优势, 可以更精准的对罐道缺陷进行分类。上述结果表明, RDM-YOLOv5 针对 YOLOv5s 做出的改进策略可以优化网络性能, 提升网络检测的准确率。

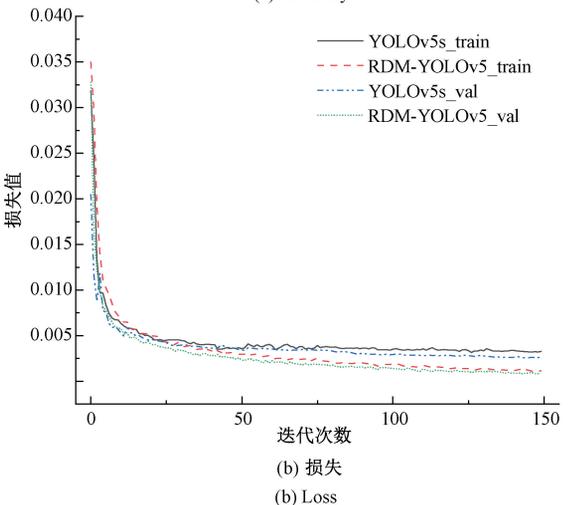
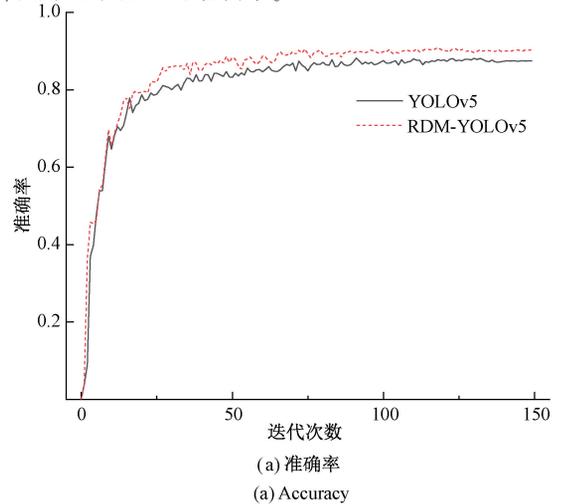


图 6 检测准确率及损失曲线

Fig. 6 Detection accuracy and loss curve

2.4 消融实验

RDM-YOLOv5 是在 YOLOv5s 的基础上,依次加入 RLKM、DCAM 以及 GSConv 这 3 种改进策略,为了衡量改进效果,在罐道缺陷数据集上开展 4 组实验,结果如表 2 所示,其中“√”表示采用对应策略。

表 2 消融实验结果

Table 2 Ablation results

实验	RLKM	GSConv	DCAM	mAP/%	FPS	Params/M
1				87.4	70	7.02
2	√			89.2	74	6.36
3	√	√		89.6	78	5.92
4	√	√	√	90.6	78	5.94

由表 2 可以看出,实验 1 是未加入改进策略的 YOLOv5s 网络,其 mAP 值为 87.4%,FPS 为 70,模型参数量为 7.02 M,实验 4 是采用了上述 3 种改进策略后的 RDM-YOLOv5,缺陷检测的 mAP 为 90.6%,比 YOLOv5s 上升了 3.7%,检测速度提升了 11.4%,模型参数量相比 YOLOv5s 下降了 15.4%。实验 1 和实验 4 证明 RDM-YOLOv5 不仅降低了模型参数,还在提高模型检测精度的前提下,提升了模型的检测速度。

实验 2 是在 YOLOv5s 主干网络末端和增强网络末端使用 RLKM 代替原有的 C3 模块,替代后网络的 mAP 值比 YOLOv5s 上升了 1.7%,检测速度提升了 5.7%,并且模型参数量下降了 9.4%。实验 2 在提升精度的同时,有效降低参数量,提高了模型检测速度,主要得益于 RLKM 的重参数大卷积核,它能有效提升模型的特征表征能力并减少模型参数量。实验 3 在实验 2 的基础上引入 GSConv,虽然 mAP 涨幅不大,但检测速度较实验 2 提升了 5.4%,模型参数量降低了 6.9%,充分证明深度可分离卷积减轻模型复杂度,保持精度的同时降低计算成本。实验 4 在实验 3 的基础上加入 DCAM 模块,提升网络对小目标的检测精度,可以看出 mAP 值与实验 3 相比提升了 1.1%,该效果主要源于 DCAM 级联通道注意力机制 D-ECA 与坐标注意力机制 CoordA,有效来抑制背景噪声,给予细小目标特征更高的权重。表 2 的实验结果表明,RDM-YOLOv5 中提出的 3 种改进策略设计合理,均有助于提升模型的检测性能,均具有较好的检测效果。

2.5 对比实验

消融实验充分证明了 RDM-YOLOv5 模型的合理性。为了进一步验证其检测效果,将它和当前主流的目标检测网络 Faster R-CNN、SSD、YOLOv3^[24]、YOLOv4^[25] 以及 YOLOv7-tiny^[26] 在罐道表面缺陷数据集上进行对比实验,实验结果的客观指标如表 3 所示。

表 3 不同目标检测网络对比实验

Table 3 Comparative experiment of different target detection networks

实验	mAP/%	FPS	Params/M
Faster R-CNN	81.2	6	137
SSD	73.1	64	26.29
YOLOv3	75.7	60	61.95
YOLOv4	70.2	58	64.36
YOLOv7-tiny	71.0	76	6.23
YOLOv5s	87.4	70	7.02
RDM-YOLOv5	90.6	78	5.94

由表 3 可知,与上述网络相比,RDM-YOLOv5 的检测速度、平均精度均值均为最高,模型参数量最少。以 Faster R-CNN 为主的两阶段检测网络相比于传统的单阶段目标检测网络 SSD、YOLOv3、v4,其 mAP 值达到最高 81.2%,但其模型参数量大,检测速度低,难以满足实际需求。在单阶段目标检测网络中,YOLOv7-tiny 的检测速度 FPS 值最高为 76,但检测精度仅为 71.0%,相较 YOLOv3、YOLOv5s 检测网络检测精度较低。可见,RDM-YOLOv5 在提升网络检测精度的同时,提高了网络检测速度,为在工业生产中的落地应用奠定了基础。

表 4 为 YOLOv5s 网络和 RDM-YOLOv5 网络在罐道表面缺陷数据集的检测性能对比结果。从表中可知对比 YOLOv5s 网络,RDM-YOLOv5 具有更高的检测精度,并且对于 3 种罐道缺陷均具有明显的 AP 值提升,尤其针对 Crack 这类细小缺陷目标精度提升最大,相较于 YOLOv5s 原始网络 AP 值提高了 6.2%。结果表明,RDM-YOLOv5 具有更优的检测性能。

表 4 网络检测性能对比

Table 4 Comparison of network detection results

网络模型	mAP/%	AP/%		
		Abrasion	Crack	Dislocation
YOLOv5s	87.4	89.7	84.3	88.1
RDM-YOLOv5	90.6	91.5	89.6	90.8

为了直观地展示检测效果,图 7 给出了 RDM-YOLOv5 与各主流单阶段目标检测算法的检测结果,其中虚线方框处为漏检情况。可见,各算法均能有效检测出罐道缺陷小目标,但 YOLOv4 和 YOLOv7-tiny 受背景噪声的干扰,对小目标特征提取不充分,出现如图 7(a) 和 (b) 中的漏检情况;YOLOv3 检测效果略优于以上两种算法,虽然没有出现漏检情况,但是 YOLOv3 和 SSD 对细小目标缺陷检测结果不理想,出现图 7(b) 中缺陷目标检测的置信度低于 0.6 的情况;原始 YOLOv5s 网络对小目标的检测效果较以上算法具有较大提升,但是与 RDM-YOLOv5 相比,YOLOv5s 对 3 种缺陷目标检测的置信度均小于 0.9,对图 7(b) 缺陷目标的检测置信度均值仅为 0.74;RDM-YOLOv5 对图 7(a) 和 (c) 缺陷检测的置信度

均达到 0.9 以上,对图 7(b)细小目标缺陷的检测置信度均值达到 0.86,相比 YOLOv5s, RDM-YOLOv5 对 3 种缺陷目标检测的置信度均值分别提高了 11.98%、15.70%

和 5.88%。综上, RDM-YOLOv5 不仅能高精度地检测出多尺度的罐道缺陷,还能显著提升对细小缺陷目标的检测效果。

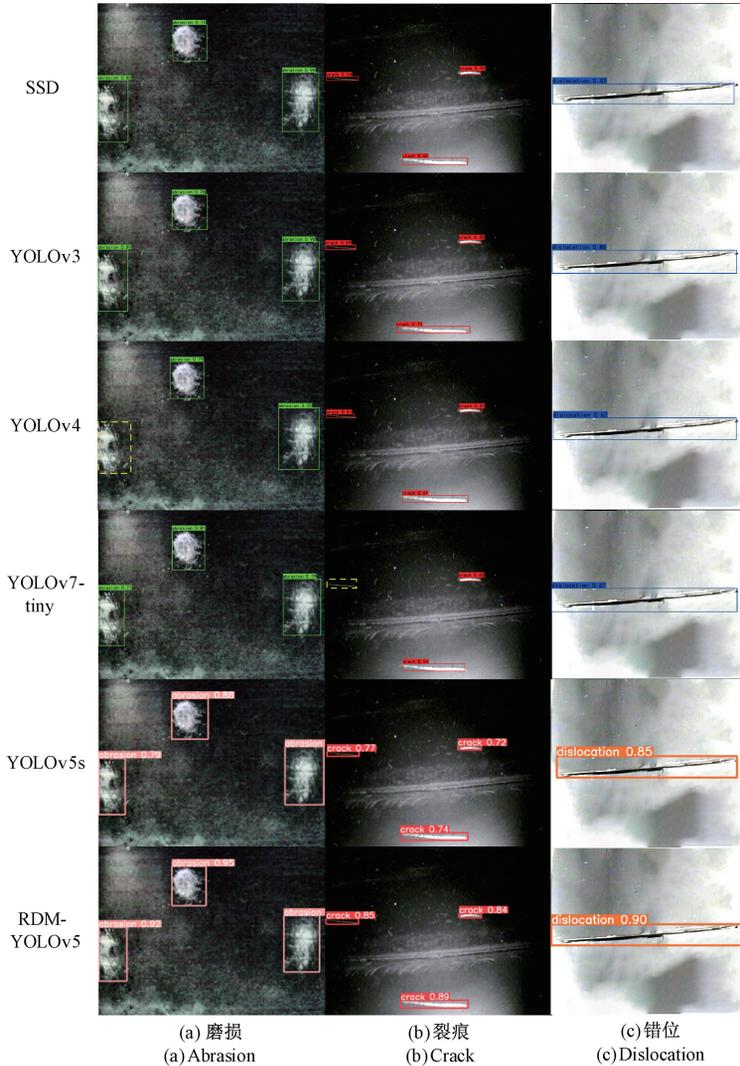


图 7 RDM-YOLOv5 对比实验检测结果

Fig. 7 RDM-YOLOv5 comparative detection results

3 结 论

针对立井罐道表面缺陷检测问题,提出了融合特征增强和级联注意力的缺陷检测算法 RDM-YOLOv5。该算法中构建的特征增强模块 RLKM,通过重参数化大内核卷积,能够增大特征提取层的 ERF,并增强模型的特征提取能力,该效果保证了检测精度的提高和模型参数数量的降低。设计的 D-ECA 在级联 CoordA 后,能够增加对输入特征通道和位置信息的利用率,凸显对小目标检测有利的融合特征。轻量级卷积的引入能减少冗余信息,促进模型检测速度的进一步提升。在罐道表面缺陷数据集

上, RDM-YOLOv5 对缺陷检测的 mAP 值达到 90.6%,模型参数量降低至 5.94 M,检测速度达到 78,算法的准确性和实时性均得到保障。

RDM-YOLOv5 为煤矿井下罐道缺陷识别任务提供了新的研究思路和理论模型,后续需进一步构建井下实际罐道缺陷数据信息样本库,充分训练增强算法模型的鲁棒性。

参 考 文 献

[1] WU B, LI W, JIANG F. Fault diagnosis of mine shaft guide rails using vibration signal analysis based on dynamic time warping [J]. Symmetry, 2018, 10(10): 500.

- [2] 封华,李强,陈朋朋. 磁轮驱动钢罐道巡检机器人研究[J]. 煤炭科学技术,2022,50(11):183-193.
FENG H, LI Q, CHEN P P. Research on inspection robot of steel tank track driven by magnetic wheel [J]. Coal Science and Technology,2022,50(11):183-193.
- [3] 丁鑫. 基于机器视觉的提升系统钢罐道摩擦副磨损检测研究[D]. 徐州:中国矿业大学,2021.
DING X. Research on wear detection of rail tank track friction pair in lifting system based on machine vision [D]. Xuzhou: China University of Mining and Technology,2021.
- [4] 马天兵,王鑫泉,王孝东. 基于 EMD-PNN 网络的刚性罐道故障诊断方法[J]. 电子测量与仪器学报,2019,33(3):58-64.
MA T B, WANG X Q, WANG X D. Fault diagnosis method of rigid tank channel based on EMD-PNN network [J]. Journal of Electronic Measurement and Instrumentation, 2019,33(3):58-64.
- [5] 肖飞. 基于 ZYNQ 的刚性罐道图像采集处理系统研究[D]. 徐州:中国矿业大学,2021.
XIAO F. Research on rigid tank channel image acquisition and processing system based on ZYNQ [D]. Xuzhou: China University of Mining and Technology,2021.
- [6] 陶显,侯伟,徐德. 基于深度学习的表面缺陷检测方法综述[J]. 自动化学报,2021,47(5):1017-1034.
TAO X, HOU W, XU D. A review of surface defect detection methods based on deep learning [J]. Acta Automatica Sinica,2021,47(5):1017-1034.
- [7] 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.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C]. European Conference on Computer Vision. Springer, Cham, 2016: 21-37.
- [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.
- [10] HE Y, SONG K, MENG Q, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features [J]. IEEE Transactions on Instrumentation and Measurement, 2019, 69 (4): 1493-1504.
- [11] 单明陶,高玮玮,杨亦乐,等. 基于改进 Faster R-CNN 的刹车零件缺陷检测算法[J]. 国外电子测量技术, 2022,41(4):22-28.
SHAN M T, GAO W W, YANG Y L, et al. Defect detection algorithm for brake parts based on improved Faster R-CNN [J]. Foreign Electronic Measurement Technology,2022,41(4):22-28.
- [12] 向宽,李松松,栾明慧,等. 基于改进 Faster RCNN 的铝材表面缺陷检测方法[J]. 仪器仪表学报, 2021, 42(1): 191-198.
XIANG K, LI S S, LUAN M H, et al. Aluminum surface defect detection method based on improved Faster RCNN [J]. Chinese Journal of Scientific Instrument, 2021, 42(1): 191-198.
- [13] YANG H, WANG Y, HU J, et al. Deep learning and machine vision-based inspection of rail surface defects[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 71: 1-14.
- [14] 刘涛,张涛. 基于 GhostNet-YOLOv4 算法的印刷电路板缺陷检测[J]. 电子测量技术,2022,45(16): 61-70.
LIU T, ZHANG T. Defect detection of printed circuit board based on GhostNet-YOLOv4 algorithm [J]. Electronic Measurement Technology, 2022, 45 (16): 61-70.
- [15] 马燕婷,赵红东,阎超,等. 改进 YOLOv5 网络的带钢表面缺陷检测方法[J]. 电子测量与仪器学报,2022, 36(8):150-157.
MA Y T, ZHAO H D, YAN CH, et al. Surface defect detection method of strip steel based on improved YOLOv5 network [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36 (8): 150-157.
- [16] 贾晓芬,于业齐,郭永存,等. 航拍绝缘子自爆缺陷的轻量化检测方法[J]. 高电压技术,2023,49(1): 294-300.
JIA X F, YU Y Q, GUO Y C, et al. Lightweight detection method for self-detonation defect of aerial photograph insulators [J]. High Voltage Technology, 2023,49(1):294-300.
- [17] GLENN J. YOLOv5 release v6.0. <https://github.com/ultralytics/yolov5/tree/v6.0>,2021.
- [18] DING X, ZHANG X, MA N, et al. Repvgg: Making vgg-style convnets great again [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13733-13742.
- [19] DING X, ZHANG X, HAN J, et al. Scaling up your kernels to 31x31: Revisiting large kernel design in CNNs [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11963-11975.
- [20] LUO W, LI Y, URTASUN R, et al. Understanding the

effective receptive field in deep convolutional neural networks[J]. *Advances in Neural Information Processing Systems*, 2016, 29.

- [21] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2020: 11531-11539.
- [22] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021: 13713-13722.
- [23] LI H, LI J, WEI H, et al. Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles [J]. *arXiv preprint arXiv: 2206.02424*, 2022.
- [24] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. *arXiv preprint arXiv: 1804.02767*, 2018.
- [25] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. *arXiv preprint arXiv:2004.10934*, 2020.
- [26] 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 preprint arXiv:2207.02696*, 2022.

作者简介



赵佰亭, 2002 年于青岛科技大学获得学士学位, 2005 年于青岛科技大学获得硕士学位, 2010 年于哈尔滨工业大学获得博士学位, 现为安徽理工大学教授, 主要研究方向为深度学习、图像处理。

E-mail: btzhao@ aust. edu. cn

Zhao Baiting received his B. Sc. degree from Qingdao University of Science and Technology in 2002, M. Sc. degree from Qingdao University of Science and Technology in 2005 and Ph. D. degree from Harbin Institute of Technology in 2010. Now he is a professor at Anhui University of Science and Technology. His main research interests include deep learning and image processing.



吴俊东 (通信作者), 现为安徽理工大学硕士研究生, 主要研究方向为图像处理、缺陷检测。

E-mail: 18502190958@163.com

Wu Jundong (Corresponding author) is now a M. Sc. candidate at Anhui University of Science and Technology. His main research interests include image processing and defect detection.