

# 基于 YOLOv8 的发动机缸内异物 检测算法开发与应用

房运涛,李爽,韩晓琴,翟强,庄顺胥,齐伟,宋丽娟

潍柴动力股份有限公司,山东 潍坊 261061

**摘要:**为解决人工检测发动机缸内异物时的漏检和误检等问题,设计基于改进目标检测算法 YOLOv8 的发动机缸内异物检测算法并进行试验验证。基于 CoFNet 中的注意力机制,设计 Contextual Attention 模块,重构 C2f 中的 Bottleneck 结构为 CoA\_C2f,替换 YOLOv8 骨干网络中的 C2f 模块;将模型 Neck 部分连续上采样后的特征图 Concat 模块替换为上下文聚合模块 CAM;在 Neck 和 Head 之间嵌入 Triplet Attention 模块。试验结果表明:设计的发动机缸内异物检测模型可有效识别缸内异物,在原始 YOLOv8 基础上引入 CoA\_C2f、CAM 和 Triplet Attention 3 个模块后的平均检测精度提高 21.65%。

**关键词:**改进 YOLOv8 算法;目标检测;机器视觉;异物检测

**中图分类号:**TP391.41

**文献标志码:**A

**文章编号:**1673-6397(2024)04-0033-08

**引用格式:**房运涛,李爽,韩晓琴,等. 基于 YOLOv8 的发动机缸内异物检测算法开发与应用[J]. 内燃机与动力装置,2024,41(4):33-40.

FANG Yuntao, LI Shuang, HAN Xiaoqin, et al. Design and application of the foreign object detection algorithm for engine cylinder based on YOLOv8[J]. Internal Combustion Engine & Powerplant, 2024, 41(4):33-40.

## 0 引言

在发动机装配过程中,活塞缸内异物检测是一项非常重要的任务。目前,发动机缸内异物检测仍然依靠人眼进行观测,但由于人眼容易疲劳以及个体认知有限,极易出现漏检和误检等问题。近年来,基于深度学习的视觉检测系统在工业生产中的应用越来越广泛<sup>[1-2]</sup>。基于深度学习的目标检测方法通过学习大量的图像数据自动识别目标,不再依赖于传统的手工特征提取方法,使发动机缸内异物的自动检测成为可能<sup>[3-4]</sup>。

基于深度学习的目标检测算法主要分为两阶段和单阶段 2 类。两阶段算法需要先提取候选区域,再对候选区域进行分类和回归处理,例如 R-CNN<sup>[5]</sup>、Fast R-CNN<sup>[6]</sup>、Faster R-CNN<sup>[7]</sup>等。单阶段算法直接对输入的图像应用算法并输出类别和相应的定位,例如 SSD 系列算法<sup>[8-10]</sup>和 YOLO 系列算法<sup>[11-14]</sup>,其中 YOLO 系列算法是工业自动检测中广泛使用的目标检测框架。YOLO 算法将整个图像划分为网格单元,并在每个单元中预测多个边界框和类别概率,从而实现目标检测。相比其他算法,YOLO 算法具有更高的检测速度和准确性<sup>[15-16]</sup>。但是,作为一种单阶段算法,YOLO 算法在目标位置的精度上比两阶段算法低,尤其是在对一些目标的检测方面。因此,国内外学者都致力于如何提高 YOLO 系列算法的检测精度。Hu 等<sup>[17]</sup>在 YOLOv5 中的骨干网络中引入压缩和激励(squeeze and excitation, SE)模块,以突出特征图中的重要特征,提高了网络检测的性能;Zhang 等<sup>[18]</sup>提出了一种跨阶段局部 Neck 模块,并结合 YOLOv5 模型,在多尺度目标上获得了较好的精度;Liu 等<sup>[19]</sup>提出了一种动态检测头,融合 3 种注意力机制来提高

收稿日期:2024-03-01

第一作者简介:房运涛(1989—),男,山东潍坊人,工学博士,工程师,主要研究方向为机器视觉、人工智能,E-mail:fangyun.tao@163.com。

网络的检测性能;Lou 等<sup>[20]</sup>针对小目标检测,在 YOLOv8 的基础上提出了一种新的下采样方法和特征融合网络,在保留了背景信息的前提下,能够兼顾浅层和深层信息,有效提高了模型的检测精度;程松等<sup>[21]</sup>通过在 YOLOv5 中引入不同的损失函数,提高了算法的检测精度。

综上所述,引入注意机制、完善特征融合网络、优化损失函数,是提高 YOLO 系列算法检测精度的重要手段。但是,上述研究对于算法的优化均是针对特定的应用场景,而且引入注意力机制、优化特征融合网络等都可能网络复杂度增加。本文中构建更有针对性的注意力机制,提出一种基于改进 YOLOv8 算法的发动机缸内异物检测模型,实现发动机缸内螺栓、螺母和垫片等异物的自动检测,解决由于依靠人工观察导致的漏检和误检等问题,并有效平衡计算代价及性能增益。

## 1 基于 YOLOv8 的异物检测算法

本文中在 YOLOv8 算法的基础上,有针对性地对其输入端、网络结构进行改进,构建发动机缸内异物检测模型,旨在增强模型对发动机缸内异物的识别能力,以及模型在不同环境下的泛化能力,实现对发动机缸内异物的准确、稳定和高效的检测。

### 1.1 YOLOv8 算法整体架构

YOLOv8 算法由 Glenn-Jocher 提出,与 YOLOv3 和 YOLOv5 算法一脉相承。YOLOv8 算法网络结构示意图如图 1 所示。YOLOv8 算法的总体结构包括输入端(Input)、骨干网络(Backbone)、特征融合网络(Neck)和输出端(Head) 4 个部分。Backbone 主要用于特征提取,YOLOv8 将 YOLOv5 中的跨阶段局部网络(cross stage partial network, CSPNet)模块替换为轻量级 C2f 模块,通过密集残差结构增强特征表达的能力,在尾部使用快速空间金字塔池化层(spatial pyramid pooling fast, SPPF)增加感受并捕获场景中不同层次的特征信息;Neck 部分主要用于特征融合,采用路径聚合网络与 C2f 模块,针对 Backbone 三个阶段输出的不同尺度的特征图进行融合,帮助浅层信息向深层特征聚合;Head 部分使用解耦头结构,分为分类和定位预测端,缓解了分类和定位任务之间存在的冲突。

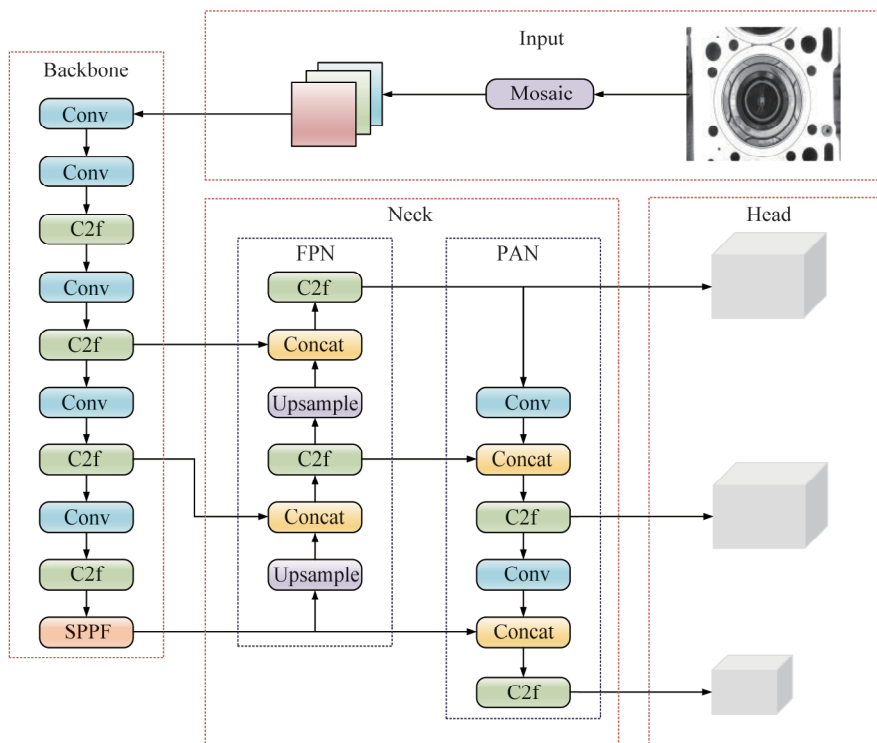


图1 YOLOv8 算法网络结构示意图

### 1.2 YOLOv8 算法改进

为解决人工检测发动机缸内异物时的漏检和误检等问题,实现发动机缸内异物的自动检测,本文中以 YOLOv8s 模型为基础,基于 CoTNet 中的注意力机制,设计了 Contextual Attention 模块,基于该模块重构了 C2f 中的 Bottleneck 结构,命名为 CoA\_C2f,替换 YOLOv8 骨干网络中的 C2f 模块,改善骨干网络的多尺度细节特征的提取能力。在模型的 Neck 部分,将连续上采样后的特征图 Concat 模块替换为上下文聚合模块 CAM,促进不同层次特征图的交互融合。充分发挥注意力机制的优势,在 Neck 和 Head 之间嵌入三重注意力(triplet attention, TA)模块,提高模型提取特征的能力。最后,采用 WIoU 作为边界框损失函数。改进的 YOLOv8 算法网络结构示意图如图 2 所示。

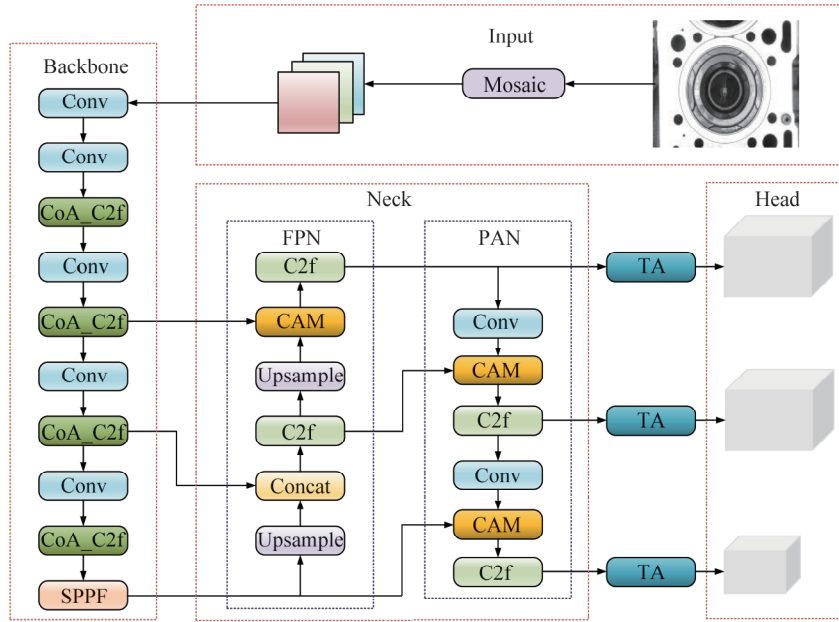


图 2 改进的 YOLOv8 算法网络结构示意图

#### 1.2.1 重构骨干网络

YOLOv8 借鉴 CSPDarkNet 网络结构设计了特征提取骨干网络,与 YOLOv5 最大的区别是用 C2f 模块替换了 C3 模块,其特征提取能力仍有欠缺。本文中基于 CoTNet 中的注意力机制,设计了 Contextual Attention 模块,其结构如图 3 所示,图中  $h, w$  分别为空间的高度、宽度维度,  $c$  为通道维度。首先,通过 3 个不同的卷积操作提取特征层 keys、query 和 value,将有上下文特征的 keys 和特征层 query 拼接,然后通过两个  $1 \times 1$  卷积块并经 Softmax 操作将结果转化为注意力权重矩阵  $A$ ;其次,将权重矩阵  $A$  与特征层 value 相乘后得到全局动态特征层;最后,将全局动态特征层和 keys 特征层进行融合,并通过一个  $1 \times 1$  卷积得到对突出点着重关注、同时不缺失对普通点关注的注意力输出特征层 output。本文中结合 Contextual Attention 模块,对 YOLOv8 的骨干网络进行了重建,将 Contextual Attention 模块嵌入到 YOLOv8 骨干网络的残差块中,替换原有的普通  $3 \times 3$  卷积,形成融合注意力机制的 C2f 模块——CoA\_C2f 模块,可有效改善骨干网络的特征提取能力。

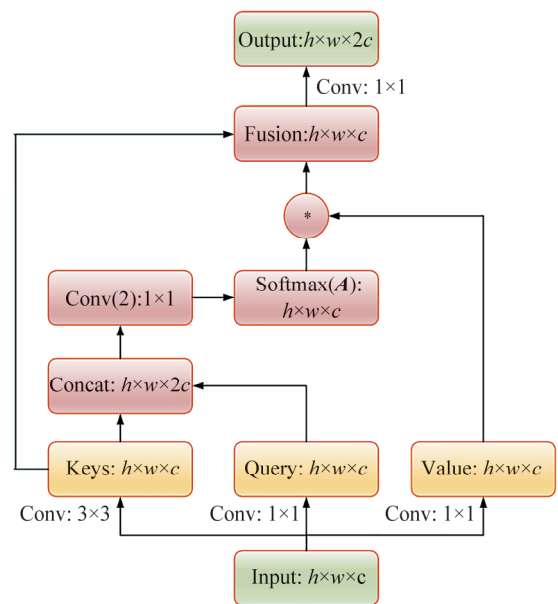


图 3 Contextual Attention 模块

### 1.2.2 改进特征融合方式

在处理复杂背景下异物图像时,由于不同层级的特征图含有信息差异,直接拼接这些特征图导致信息冗余。此外,由于图像经过了上采样处理,使浅层细节特征被深层语义特征遮盖,丢失包含丰富空间细节的低级特征,最终导致模型的特征提取能力降低,影响异物检测精度。为了解决上述问题,参考文献[22],引入上下文聚合模块CAM,即在模型的Neck部分,将连续上采样后的特征图Concat模块替换为CAM模块,可以自适应地调整局部特征之间的空间叠加关系,促进不同层次特征图的交互融合。CAM模块的结构如图4所示。它能有效学习不同层级特征图之间的差异,通过使用可变形卷积学习像素间的偏移关系,并采用注意力机制智能分配融合不同特征图的比例,从而提高特征信息的利用率。

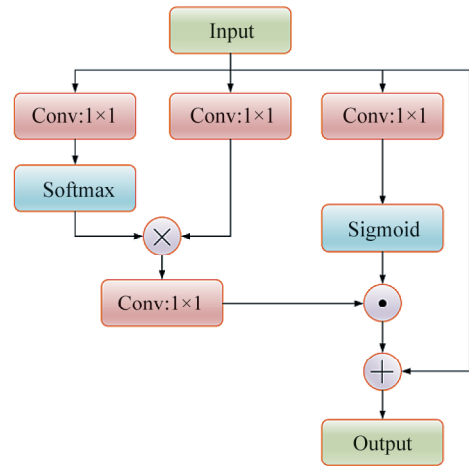


图4 CAM模块结构示意图

### 1.2.3 引入TA模块

传统注意力机制在计算通道注意力时在输入张量上进行池化操作,导致大量空间信息丢失。TA模块是利用3个分支结构捕获输入数据的跨维度交互,从而计算注意力权重,在尽量减小空间信息损失的前提下,构建输入通道或空间位置之间的相互依赖性。TA模块结构示意图如图5所示。TA模块由3个分支组成,前两个分支负责捕获空间维度 $h$ 或 $w$ 与通道维度 $c$ 之间的交互特征,第3个分支——identity用于构建空间注意力,所有3个分支的输出通过简单平均进行聚合<sup>[23]</sup>。计算时,对每个分支中的输入张量进行排列变换,通过池化层、卷积层和Sigmoid激活层生成注意力权重,然后经排列变换转换为输入张量的形状。为了能够使发动机缸内异物检测模型的性能得到进一步的提高,本文中在特征融合部分之后添加注意力模块以提高模型提取特征的能力。

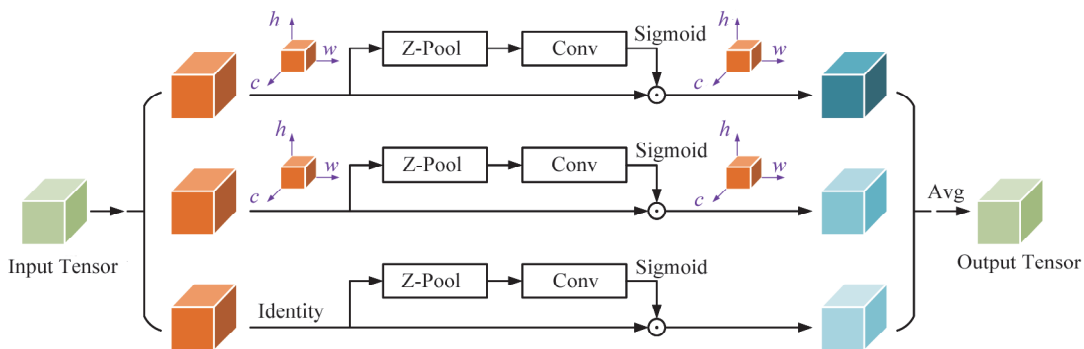


图5 TA模块结构示意图

## 2 异物检测模型部署与测试

### 2.1 数据集及试验环境

#### 2.1.1 缸内异物检测数据集

目前还没有为发动机缸内异物检测量身定制的公共数据集。为捕捉发动机缸内存在异物的图像,将螺栓、螺母和垫片等常见异物放置发动机活塞缸内,通过人工拍照采集在不同光线和角度下的异物照片。为了扩充数据集,采用泊松融合算法,将不同结构、尺寸的螺栓等异物无缝缝合到正常的发动机活塞图像上,得到了一个包含5 000张图像、专门针对发动机缸内异物检测的数据集,如图6所示。数据集标签包括螺栓、螺母、垫片等。

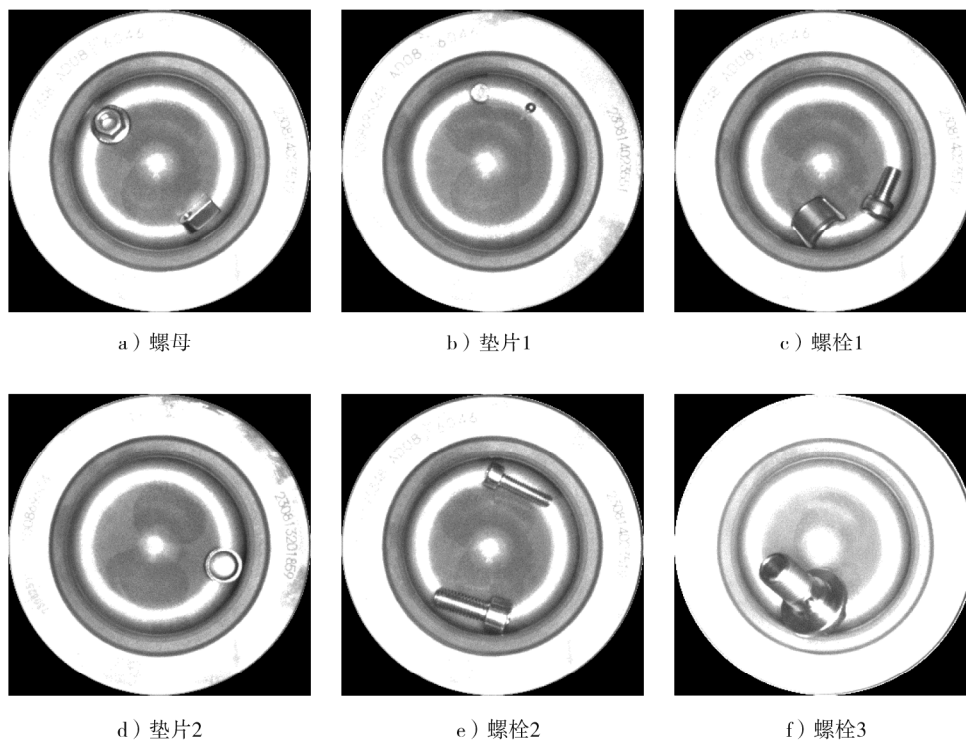


图 6 发动机缸内异物检测数据集示例

### 2.1.2 试验环境

硬件配置包括 1 个 NVIDIA GeForce RTX 3060 图形处理器 (graphics processing unit, GPU) 和一个 Intel(R) Core i9-13900 中央处理器 (central processing unit, CPU), 其中 GPU 的内存为 6 Gbyte; 软件配置包括 Python3.9、PyTorch1.13.1 和 Cuda11.6 深度学习框架。试验中, 将缸内异物检测数据集按 8 : 1 : 1 的比例分为训练集、验证集和测试集, 训练迭代次数设置为 300, 使用 SGD 优化器, 设置初始学习率为 0.01、批处理大小为 8、动量参数为 0.8、置信度阈值为 0.5、输入图像像素为 640×640。

### 2.2 异物检测模型部署

OpenVINO 是优化神经网络模型和加速推理计算的软件工具套件, 可以优化绝大多数框架 (Tensorflow、Pytorch 等) 的深度学习模型, 可在英特尔处理器和其他硬件平台上以最佳性能进行计算。OpenVINO Runtime 可以自动使用激进的图形融合, 内存重用, 负载平衡和跨 CPU、GPU、视频处理单元 (video processing unit, VPU) 等进行集成并行处理, 可以优化深度学习流水线。此外, OpenVINO 工具包还包含了 OpenCV 图片处理工具包, 用于前处理和推理结果后处理等。鉴于上述优势, 本文中基于 OpenVINO 工具套件, 将基于 YOLOv8 的异物检测模型部署在边缘设备上, 其部署流程如图 7 所示。

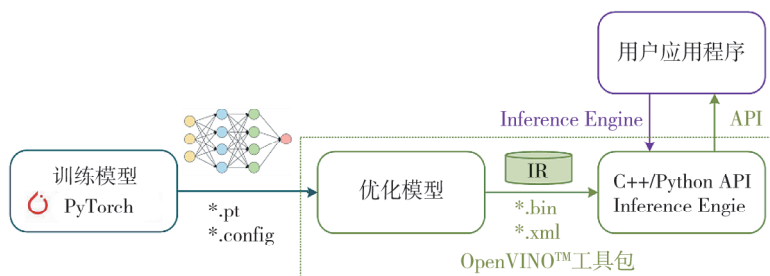


图 7 基于 OpenVINO 的异物检测模型部署流程

### 2.3 异物检测实验及结果分析

选取交并比 (intersection over union, IOU) 阈值为 0.5 时的平均精度 mAP0.5 作为异物检测模型的衡量指标。为了验证 CoA\_C2f、CAM 和 TA 模块对异物检测模型整体性能的影响, 对异物检测模型进行拆



分,并在相同的数据集上测试,设计的异物检测模型在自制发动机缸内异物检测数据集上测试结果如图8所示,模型检测精度如表1所示。由图8可知:设计的模型可以有效地检测出发动机活塞缸内的螺栓、螺母等异物。由表1可知:只采用原始YOLOv8s模型的检测精度为67.11%;在原始YOLOv8s的基础上至引入一个模块时,平均精度mAP0.5分别提高了3.78%、2.50%、3.20%,当引入2个模块时,平均精度分别提高了12.37%、11.28%、13.74%;当引入3个模块时,平均精度提高了21.65%,表明设计的检测模型相对于原始YOLOv8s模型具有较好的异物检测性能。

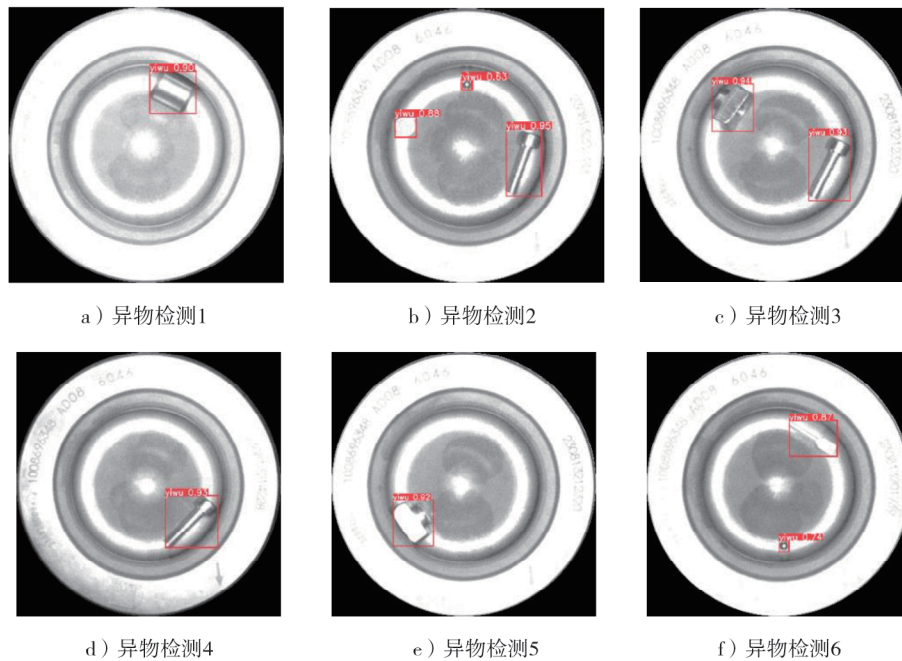


图8 缸内异物检测数据集测试结果

表1 不同模型的检测精度

YOLOv8s	CoA_C2f	CAM	Triplet Attention	mAP0.5/%
√				67.11
√	√			69.65
√		√		68.79
√			√	69.26
√	√	√		75.41
√	√		√	74.68
√		√	√	76.33
√	√	√	√	81.64

注:“√”表示采用该模型检测。

综上所述,在原始YOLOv8s基础上引入CoA\_C2f、CAM和TA模块后均能有效提高模型的检测精度,且CoA\_C2f、CAM和TA模块能够相互耦合,进一步提高模型的检测精度。

### 3 结论

为解决依靠人眼检测发动机缸内异物时的漏检和误检等问题,设计基于改进YOLOv8的发动机缸内异物检测算法。在原始YOLOv8的基础上,重构骨干网络,改进特征融合方式,引入Triplet Attention模

块;通过人工拍照建立缸内异物检测数据集。在发动机缸内异物检测数据集上的试验结果表明:设计的发动机缸内异物检测模型可有效识别缸内异物,且引入 CoA\_C2f、CAM 和 Triplet Attention 3 个模块后的缸内异物平均检测精度提高了 21.65%。

#### 参考文献:

- [1] JIAN C, GAO J, AO Y. Automatic surface defect detection for mobile phone screen glass based on machine vision[J]. *Applied Soft Computing*, 2017, 52: 348–358.
- [2] HE Z, LIU Q. Deep regression neural network for industrial surface defect detection[J]. *IEEE Access*, 2020, 8: 35583–35591.
- [3] LIU Y, XIAO H, XU J, et al. A rail surface defect detection method based on pyramid feature and lightweight convolutional neural network[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1–10.
- [4] SINGH S A, DESAI K A. Automated surface defect detection framework using machine vision and convolutional neural networks[J]. *Journal of Intelligent Manufacturing*, 2023, 34(4): 1995–2011.
- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// *Proceedings of 2014 27th IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, USA: IEEE 2014: 580–587.
- [6] ALRUWAILI M, SIDDIQI M H, ATTA M N, et al. Deep learning and ubiquitous systems for disabled people detection using YOLO models[J]. *Computers in human behavior*, 2024, 154: 1–13.
- [7] 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 & Machine Intelligence*, 2017, 39(6): 1137–1149.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]// *Proceedings of Computer Vision-ECCV 2016*. Berlin, Germany: Springer, 2016: 21–37.
- [9] FU C Y, LIU W, RANGA A, et al. DSSD: Deconvolutional single shot detector[EB/OL]. (2017-06-23) [2023-08-22]. <https://arxiv.org/abs/1701.06659> 2017: 1701.06659.
- [10] LI Z, ZHOU F. FSSD: feature fusion single shot multibox detector[EB/OL]. (2017-11-04) [2023-09-18]. <https://arxiv.org/abs/1712.00960v1>.
- [11] 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*. Las Vegas, USA: IEEE, 2016: 779–788.
- [12] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[J]. *IEEE*, 2017: 6517–6525.
- [13] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-05) [2018-09-30]. <https://arxiv.org/pdf/1804.02767.pdf>.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOV4: optimal speed and accuracy of object detection [EB/OL]. (2020-04-23) [2023-01-26]. <https://zhuanlan.zhihu.com/p/520922132>.
- [15] LI C, LI L, JIANG H, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. (2022-10-23) [2023-03-05]. <https://arxiv.labs.arxiv.org/html/2209.02976>.
- [16] WANG C Y, BOCHKOVSKIY A, LIAO H. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real time object detectors[EB/OL]. (2022-06-06) [2023-03-05]. <https://arxiv.org/abs/2207.02696>.
- [17] HU J, SHEN L, ALBANIE S. Squeeze-and-excitation networks[J]. *IEEE CVF Conference on Computer Vision and Pattern Recognition*, 2017: 7132–7141.
- [18] ZHANG T. GBH-YOLOv5: Ghost convolution with Bottleneck CSP and tiny target prediction head incorporating YOLOv5 for PV panel defect detection[J]. *Electronics*, 2023, 12(3): 561–576.
- [19] LIU D, LIANG J, GENG T, et al. Tripartite feature enhanced pyramid network for dense prediction[J]. *IEEE Transactions on Image Processing*, 2023, 32: 3272826.
- [20] LOU H, DUAN X, GUO J, et al. DC-YOLOv8: small size object detection algorithm based on camera sensor[J]. *Electronics*, 2023, 12(10): 172–173.
- [21] 程松, 杨洪刚, 徐学谦, 等. 基于 YOLOv5 的改进轻量型 X 射线铝合金焊缝缺陷检测算法[J]. *中国激光*, 2022, 49(21): 136–144.

- [22] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: convolutional triplet attention module[C]// Proceedings of the IEEE/CVF winter conference on applications of computer vision. New York, USA: IEEE, 2021:3139-3148.
- [23] 付锦燧,张自嘉,孙伟,等.改进YOLOv8的航拍图像小目标检测算法[J].计算机工程与应用,2024,60(6):100-109.

## Design and application of the foreign object detection algorithm for engine cylinder based on YOLOv8

FANG Yuntao, LI Shuang, HAN Xiaoqin, ZHAI Qiang, ZHUANG Shunxu,  
QI Wei, SONG Lijuan

Weichai Power Co., Ltd., Weifang 261061, China

**Abstract:** To solve the problems of missed and false detection when detecting foreign objects in engine cylinders by manual testing, an engine cylinder foreign object detection algorithm based on the improved object detection algorithm YOLOv8 is designed and experimentally verified. Based on the attention mechanism in CoTNet, a Contextual Attention module and reconstruct the Bottleneck structure in C2f, named CoA\_C2f, are designed to replace the C2f module in the YOLOv8 backbone network. In the Neck section of the model, the continuously upsampled feature map Concat module is replaced with the context aggregation module CAM. Triplet Attention module is embed between Neck and Head. The experimental results show that the designed engine cylinder foreign object detection model can effectively identify foreign objects in the cylinder, and the average detection accuracy is improved by 21.65% after introducing CoA\_C2f, CAM, and Triplet Attention modules on the basis of the original YOLOv8s.

**Keywords:** improved YOLOv8 algorithm; target detection; machine vision; foreign object detection

(责任编辑:臧发业)

.....  
(上接第32页)

optimized. The analysis results show that the presence of corrosive elements such as P and K in the fuel used by the unit created an acidic environment, the installation process leads to asymmetric tensile stress at the shoulder of the nozzle, and the small radius of the shoulder corner causees stress concentration, these factors result in coarse martensitic microstructure and high hardness on the nozzle surface, making it sensitive to stress, which is the primary reason for the nozzle cracking. By improving the nozzle material, increasing the shoulder corner radius from 0.45 mm to 1.20 mm, and adding a pre-tightening step before the final fastening of the injector and high-pressure connector, after 4 000 h of durability testing, the optimized nozzle does not crack, the cracking fault of the oil nozzle has been effectively solved.

**Keywords:** common rail system; injector nozzle; intergranular cracking; hydrogen stress corrosion

(责任编辑:臧发业)