

DOI:10.16136/j.joel.2023.07.0374

基于改进YOLOv5的动车组关键部件缺陷检测

徐国伟^{1,2*}, 林辉¹, 修春波^{1,3}, 杨楠¹, 刘铭阳¹

(1. 天津工业大学 控制科学与工程学院, 天津 300387; 2. 天津工业大学 机械工程学院, 天津 300387; 3. 天津工业大学 电工电能新技术天津市重点实验室, 天津 300387)

摘要:针对目前动车组 (electric multiple units, EMUs) 关键部件缺陷检测模型复杂、小目标漏检率高和检测效率低的问题, 提出一种基于改进YOLOv5的缺陷检测方法。该方法在利用生成对抗网络 (generative adversarial network, GAN) 进行数据增强的基础上, 采用轻量级网络 MobileNetV3-large 对YOLOv5m 主干网络进行替换, 同时使用深度可分离卷积优化颈部 3×3 网络结构, 以降低模型的参数量和计算量; 在改进后的主干网络中引入坐标注意力机制 (coordinate attention, CA), 以捕获小目标的位置信息和通道信息, 增强网络的特征表达能力; 对非极大值抑制 (non-max suppression, NMS) 算法进行优化, 融入重叠检测框中心点的位置信息, 以提升预测框的定位准确性。在EMUs缺陷数据集上的实验结果表明, 本文提出的检测模型相较于YOLOv5m, 参数量减少了77%, 计算量降低了80.9%, 单张图片的检测时间减少了31.7%, 平均精度均值 (mean average precision, *mAP*) 可达到0.804。另外, 在NEU-DET数据集上的实验结果表明, 改进后的模型也具有较弱的泛化能力。

关键词: 动车组 (EMs); 缺陷检测; YOLOv5; MobileNetV3; 注意力机制

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1005-0086(2023)07-0752-10

Defect detection of key components of electric multiple units based on improved YOLOv5

XU Guowei^{1,2*}, LIN Hui¹, XIU Chunbo^{1,3}, YANG Nan¹, LIU Mingyang¹

(1. School of Control Science and Engineering, Tiangong University, Tianjin 300387, China; 2. School of Mechanical Engineering, Tiangong University, Tianjin 300387, China; 3. Tianjin Key Laboratory of Advanced Electrical Engineering and Energy Technology, Tiangong University, Tianjin 300387, China)

Abstract: At present, the defect detection of key components of electric multiple units (EMUs) has the problems of complex model, high missed detection rate of small targets and low detection efficiency. To solve the existing problems, a defect detection method based on improved YOLOv5 is proposed. On the basis of using generative adversarial network (GAN) to expand the dataset, the YOLOv5m backbone extraction network is changed to the MobileNetV3-large network structure, and the neck 3×3 convolution layer is optimized by using depthwise separable convolution to further reduce the model complexity. Then, the coordinate attention (CA) is introduced into the improved backbone network to capture the location information and channel information of small targets, thereby enhancing the feature expression ability of the network. The non-max suppression (NMS) algorithm is optimized by integrating the position information of the center point of the overlapping detection box to improve the accuracy of the prediction box location. The experimental results on the EMUs defect dataset show that, compared with YOLOv5m, the improved model reduces the amount of parameters by 77%, the amount of computation by 80.9%, the detection time of a single image by 31.7%, and the mean average precision (*mAP*) can reach 0.804. In addition, the experimental results on the NEU-DET dataset show that the improved

* E-mail: xuguowei@tiangong.edu.cn

收稿日期: 2022-05-20 修订日期: 2022-07-16

基金项目: 天津市自然科学基金(18JCYBJC88300)资助项目

model also has a strong generalization ability.

Key words: electric multiple units (EMUs); defect detection; YOLOv5; MobileNetV3; attention mechanism

0 引言

近年来,我国高速铁路网规模不断增大,高铁运营里程已突破4万千米,稳居世界第一。为满足日益增长的运输需求,动车组(electric multiple units, EMUs)需全天候不间断运行,安全隐患不断增加。动车组为高速移动的复杂巨系统,含有大量关键零部件,例如齿轮箱、转向架、减振器等,易受运行环境的影响而产生缺陷。若检查处理不及时,可能会导致事故的发生。目前国内主要采用动车组运行故障图像检测系统(trouble of moving EMU detection system, TEDS)对动车组进行安全检测,但检测准确率较低,高速相机采集到的图片需要人工进行二次识别,检测效率和质量无法保证。因此,提高动车组关键部件缺陷检测效率和精度对运行安全有重要的意义。

随着基于深度学习的目标检测技术迅速发展,国内外学者逐渐将深度学习应用到动车组关键部件的缺陷检测中。利用卷积神经网络替代传统目标检测算法提取图像特征,可以有效地实现动车组关键部件的缺陷检测和分类。文献[1]针对 TEDS 系统检测精度较低的问题,提出基于卷积神经网络的缺陷检测模型和分割模型,该模型实现了动车组缺陷实例的精准分割。受电弓是动车组的关键取流部件,文献[2]基于深度学习设计出受电弓缺陷检测网络(pantograph defect detection neural network, PDDNet),用于在线检测受电弓滑板表面的缺陷,以保证动车组的供电稳定性。文献[3]提出多级缺陷检测流水线网络(multi-stage pipeline for defect detection, MPDD),该模型通过结合卷积神经网络和超分辨率策略提升检测性能,实现了动车组关键部件缺陷的自动检测。文献[4]采用生成对抗网络(generative adversarial network, GAN)对动车组部件图像进行缺陷诊断,并在定位阶段引入注意力机制,通过实验验证了该模型在复杂环境下的缺陷检测性能。上述改进虽然使算法的性能得到了提升,但对于现有动车组关键部件的缺陷检测网络,仍存在模型复杂、小目标漏检率高等弊端。

目前,基于深度学习的目标检测方法主要有两种类型:基于候选区域的两阶段目标检测模型和基于回归的单阶段目标检测模型。两阶段目标检测算法的代表有 R-CNN^[5]、SPPnet^[6]、Fast R-

CNN^[7]、Faster R-CNN^[8]。由于两阶段方法需要先从输入图像中生成候选窗口,然后才能提取特征完成检测,所以检测效率较低,无法满足实时检测的要求。而单阶段目标检测方法可以直接通过网络模型回归出目标位置,推理速度较快,检测实时性高。单阶段目标检测算法的代表有 SSD^[9]、YOLO 系列算法,因检测速度快、模型较小而被广泛应用于嵌入式移动平台。文献[10]通过简化 YOLOv3 网络结构,加入密集连接的空间金字塔池化,解决了在嵌入式平台上原始网络检测速度低的问题。文献[11]使用轻量级 GhostNet 网络替换 YOLOv4 的主干网络,并优化了损失函数,使其更有利于在移动端部署。

为了降低缺陷检测模型的复杂度,提高检测实时性和小目标检测精度,本文以 YOLOv5m 网络模型为基础,将动车组关键部件常见的3类缺陷(异物、漏油、划痕)作为检测目标,对网络结构进行优化以达到检测要求。首先,将主干特征提取网络替换为轻量级网络 MobileNetV3^[12],在不损失太多精度的情况下大幅度减少计算量和参数量,方便后续在计算资源不足的终端设备上部署;其次,将主干网络中的冗余层移除,加入坐标注意力机制(coordinate attention, CA),增加主干网络的小目标特征提取能力;最后,利用深度可分离卷积替换颈部 3×3 卷积,并对非极大值抑制(non-max suppression, NMS)进行改进,提升模型的缺陷检测效果。在降低模型复杂度的同时,有效提高了动车组关键部件缺陷的检测效率和小目标检测准确率。

1 数据集构建

动车组关键部件易受环境因素的影响产生多种缺陷,根据统计数据,动车组全年产生异物、漏油和划痕缺陷的数量占比超过半数。动车组的大量缺陷数据由 TEDS 系统实时采集,经集中处理后上传至监控中心进行检查分析。TEDS 图像采集模块由侧部相机和底部相机组成,可以采集到轮轴、转向架装置等车体底部可视部件和侧裙板、轴箱等侧部可视部件。从图像采集模块的高速线阵扫描相机选取2124张缺陷图片作为基础数据集,包括异物、漏油和划痕3种缺陷,如图1所示。数据集的数量及质量在深度卷积神经网络的训练过程中起着至关重要的作用^[13],而原始数据集中的漏油和划痕缺陷图片较

少,异物图片较多,导致缺陷样本数量不均衡。为提高模型的检测性能,对漏油和划痕缺陷样本进行数据增强。

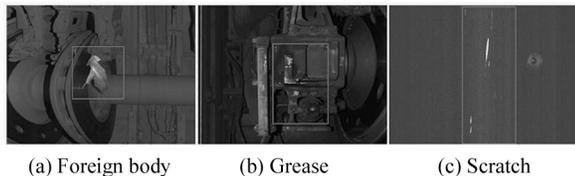


图 1 缺陷类型:(a) 异物;(b) 漏油;(c) 划痕

Fig. 1 Defect type: (a) Foreign body;
(b) Grease; (c) Scratch

1.1 有监督数据增强

有监督数据增强主要分为单样本数据增强和多样本数据增强,单样本数据增强主要分为几何变换和颜色变换。几何变换操作不改变图像的本身内容,扩增的样本过于单一,而颜色变换可以改变图片本身的信息,增加了数据样本的多样性和变化性^[14]。故本文采用添加噪声、颜色扰动和模糊处理操作对漏油和划痕缺陷数据进行扩充,以降低样本不均衡的比例,提升缺陷检测模型的泛化能力。同时采用YOLOv5模型中的Mosaic多样本数据增强方式,该方式通过随机裁剪4张图片进行拼接来丰富检测物体的背景,减少了GPU的内存使用,从而取得较好的训练效果。

1.2 无监督数据增强

考虑到有监督数据增强方式生成的图片与原图差异性较小,过量使用则会降低模型的泛化能力。故采用GAN^[15]生成质量较高的数据样本,进一步实现数据扩增,提高模型的鲁棒性。

GAN网络的设计思想来源于博弈论中的2人零和博弈^[16],由生成模型(generative model)和判别模型(discriminative model)组成。生成模型通过给定的随机噪声产生接近真实样本的图片,而判别模型用来判断输入图片为真实数据还是生成模型生成的数据,形成了一个博弈的过程。在训练过程中,循环交替地优化生成模型和判别模型的网络参数,最终达到纳什均衡点。GAN网络的目标函数如式(1)所示:

$$\min_G \max_D V(D, G) = E_{z \sim p_z(z)} [\log(1 - D(G(z)))] + E_{x \sim p_{data}(x)} [\log D(x)], \quad (1)$$

式中, x 为真实数据, p_{data} 为真实数据的分布, $D(x)$ 为真实数据分类为真的概率, z 为输入噪声, p_z 为噪声分布, $G(z)$ 为生成模型生成的数据, $D(G(z))$ 为

生成数据分类为真的概率。

原始GAN网络模型在实际训练过程中需要输入成对的数据集,且存在收敛困难和训练不稳定的缺陷,难以满足本文的数据增强需求。根据划痕缺陷样本的特点,本文采用CycleGAN^[17]进行无监督数据增强,其结构如图2所示。

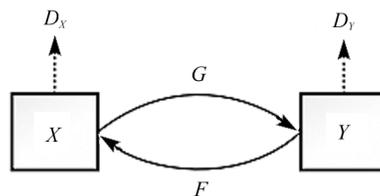


图 2 CycleGAN 网络结构

Fig. 2 CycleGAN network structure

CycleGAN由两个生成模型(G 、 F)和两个判别模型(D_X 、 D_Y)组成环形结构, X 、 Y 域中的原始图像分别由生成器 G 、 F 进行转换,判别器用来判断生成图像的真实性。同时引入了循环映射和循环一致性损失(cycle consistency loss),损失函数被定义为两个GAN网络损失和循环一致性损失的和,如式(2)所示。循环一致性原则使学习到的两种映射不会相互矛盾,从而实现了无配对的两张图片进行训练,达到少量样本也可以生成高质量图片的效果。

$$Loss_{full} = Loss_{GAN}(G, D_X) + Loss_{GAN}(F, D_Y) + \lambda Loss_{cycle} \quad (2)$$

CycleGAN在实际训练过程中没有出现模型崩溃的情况,并且生成图像的质量较高,可用于划痕数据集扩增。使用CycleGAN生成的划痕缺陷数据如图3所示,其中第一行(a)组为原始划痕缺陷图像,第二行(b)组为CycleGAN生成图像。

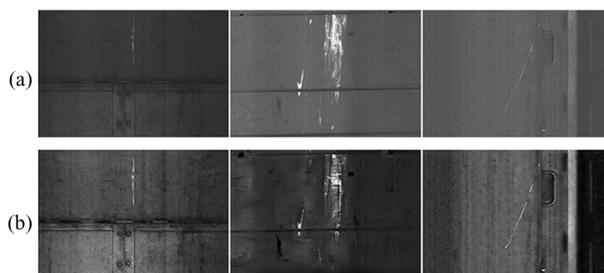


图 3 CycleGAN 生成的数据:
(a) 原始图像;(b) CycleGAN 生成图像

Fig. 3 Data generated by CycleGAN:
(a) Original image;
(b) CycleGAN generated image

2 网络结构的设计

2.1 YOLOv5m 网络结构

YOLOv5 在 YOLOv4^[18] 的基础上进行了多方面优化,根据模型深度和特征图宽度分为 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x,本文采用 YOLOv5m 作为基础模型。由于 4 种网络结构的原理相同,所以只介绍 YOLOv5m 的整体结构,结构图如图 4 所示。

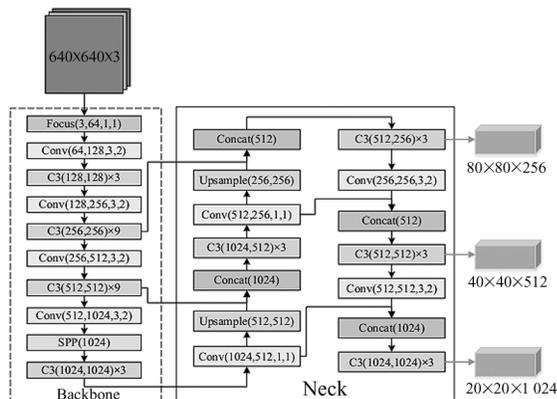


图 4 YOLOv5 网络结构图

Fig. 4 YOLOv5 network structure diagram

YOLOv5 的网络结构延续 YOLO 系列的设计,由输入端、Backbone、Neck 和预测结果输出 4 个部分组成。输入端采用 Mosaic 数据增强以加快模型的收敛速度。在主干网络中,YOLOv5 和 YOLOv4 一样借鉴了 CSPNet^[19],但 YOLOv5 中设计了两种 CSP 结构并加入 Focus 模块,Backbone 和 Neck 分别使用 CSP1_X 和 CSP2_X。Neck 通过 FPN 结构自顶向下上采样和 PAN 结构自底向上下采样进行多尺度特征融合,增加了不同尺寸特征图的语义信息和特征信息。输出端采用 CIOU_Loss 作为边界损失函数,预测框经过加权 NMS 进行筛选,输出置信度最高的边框坐标完成目标检测。虽然 YOLOv5 的检测性能在常用目标检测算法中具备一定的优势,但其主干特征提取网络和颈部网络参数量较多,计算量大,难以满足移动平台在资源受限条件下的实时缺陷检测需求。

2.2 MobileNetV3 网络

为了便于将目标检测模型在资源受限的移动设备上部署,提高缺陷检测的实时性,本文选用 MobileNetV3-large 网络结构作为特征提取网络。轻量级网络 MobileNetV3 继承了 MobileNetV1 的深度可分离卷积和 MobileNetV2 的线性瓶颈的逆残差结

构,并引入了 NetAdapt 算法和轻量级注意力模型。其激活函数在模型的第一层和后端使用 $h\text{-swish}$ 函数,如式(3)所示,其余层使用 ReLU 函数,减少了模型的计算成本。MobileNetV3 综合了前两个版本的设计优点,性能得到进一步提升,但模型的泛化能力较弱,对于复杂背景下的小目标仍存在较高的漏检率,且最后的耗时层结构冗余,容易丢失图像的多尺度特征信息。

$$h\text{-swish}(x) = x \cdot \frac{\text{ReLU6}(x+3)}{6} \quad (3)$$

为减少缺陷检测模型的参数量和计算量,移除 MobileNetV3-large 网络的最后 4 层,使用剩余的关键模块对 YOLOv5 的主干网络进行替换,得到基本的 YOLOv5-MobileNet 网络结构。在动车组关键部件的缺陷数据中,最小异物和漏油的尺度分别为 15×15 pixel 和 4×3 pixel,而最大尺度分别为 380×400 和 400×440 ,尺度不均衡会影响网络模型的检测效果。而空间金字塔池化可以进行多方面的特征提取,有效扩大感受野,丰富的特征信息可以有效解决目标尺度多样化的问题,故 YOLOv5-MobileNet 的主干网络将保留 SPP 模块,以提高异物和漏油缺陷的检测精度。

3 改进的 YOLOv5-MobileNet 网络

3.1 Backbone 网络的改进

本文采用的数据集中含有大量分辨率低、通道特征信息有限的小目标异物缺陷图像,YOLOv5-MobileNet 网络模型在训练过程中易丢失相关通道间的特征信息,导致小目标缺陷检测效果较差。而注意力机制可以优化深层网络的特征信息,依据权重去衡量不同信息的重要性,从而加强关键信息、弱化无用信息,提高小目标的检测性能^[20,21]。常用的通道注意力模块 SENet^[22]会赋予小目标通道特征信息更高的权重,但忽略了特征空间的位置信息,而位置信息会影响小目标的最终检测效果。CA^[23]将特征图的位置信息嵌入到通道注意力中,在增加少量计算开销的同时,获取了更大区域的信息,并且可以灵活地添加到轻量级主干网络,提升网络的特征表达能力。CA 模块的结构如图 5 所示。

图 5 中,CA 模块通过坐标信息嵌入和坐标注意力生成两个步骤来编码通道关系和远程依赖关系:1) 对于输入的特征图分别进行水平方向和垂直方向的平均池化,得到一对方向感知特征图 z^h 、 z^w ,沿着两个方向聚合特征有助于网络准确地定位感兴趣区

域;2) 级联生成的两个特征图,经过共享 1×1 卷积变换函数 F_1 得到中间特征图 $f = \delta(F_1([z^h, z^w]))$, 其中 δ 为非线性激活函数;然后沿着空间维度将 f 拆分为两个单独的张量 f^h, f^w , 利用两个 1×1 卷积变换 f^h, f^w 的通道数,得到通道数和输入相同的 g^h, g^w ;拓展后的 g^h, g^w 分别作为注意力权重。CA 模块的输出如式(4)所示:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (4)$$

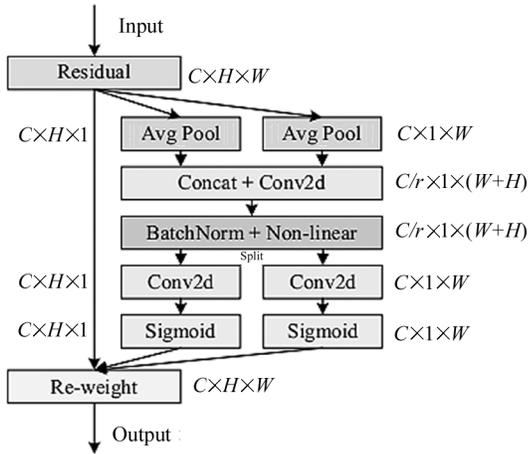


图 5 坐标注意力机制

Fig. 5 Coordinate attention mechanism

基于现有缺陷数据集中小目标异物数量多的特点和轻量级网络的设计需求,本文在 YOLOv5-MobileNet 主干网络的最后一层加入轻量级 CA 模块,在不增加主干网络复杂度的基础上,提升模型检测小目标缺陷的性能。

3.2 Neck 结构的优化

YOLOv5 的 Neck 网络采用了 FPN+PAN 的结构,其中 PANet 采用自底向上的传递路径进行特征融合,使顶层特征图也具备丰富的底层位置信息,增强了多尺度目标的定位能力。而 FPN+PAN 结构中采用多个标准卷积层进行特征提取,不仅增加了模型的参数量和计算量,而且影响特征提取的速率和多样性。

深度可分离卷积(depthwise separable convolution)^[24]是一种由深度卷积(depthwise convolution)和逐点卷积(pointwise convolution)组成的高效特征提取模块,其参数量和计算成本远小于基本卷积,并可以提取到更加丰富的特征表示。深度可分离卷积操作分为两个过程:对输入数据每个通道分别进行卷积提取特征,然后采用 1×1 卷积对第一步产生的特征进行加权组合,产生新的特征图。若输入数据

为 $P \times P \times S$,输出特征图大小为 $P \times P \times T$,卷积核大小为 $K \times K$,则标准卷积的计算量为: $K \times K \times S \times T \times P \times P$;深度可分离卷积的计算量为深度卷积和逐点卷积的计算量之和,其值为: $K \times K \times S \times P \times P + P \times P \times S \times T$,两者计算量之比为 $1/T + 1/K^2$,即深度可分离卷积的计算量为标准卷积的 $1/T + 1/K^2$ 。

在保证缺陷检测精度的前提下,为了进一步减少模型复杂度,本文借鉴了 MobileNet 网络的设计思想,采用深度可分离卷积替代 Neck 结构中的标准卷积操作。改进后的 YOLOv5-MobileNet 网络模型结构如图 6 所示。

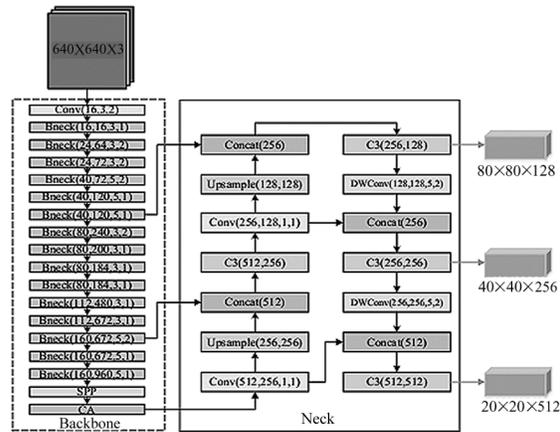


图 6 YOLOv5-CATiny 网络结构图

Fig. 6 YOLOv5-CATiny network structure diagram

3.3 NMS 的改进

在缺陷检测过程中,同一目标的周围区域会产生多个预测框,通常会采用 NMS 筛选出最佳边界框。交并比(injection over union, IOU)是 NMS 的判别依据,用于评价真实框和预测框的差距,如式(5)所示。YOLOv5 采用 NMS 算法进行预测框筛选,但 NMS 算法存在以下几个缺点:1) 根据边框的重叠程度和置信度高低来剔除冗余框时,阈值的选择会对召回率和检测精度有很大的影响;2) 当冗余框较多时,GPU 难以并行处理导致筛选速度下降。

$$IOU = \frac{(A \cap B)}{(A \cup B)}, IOULoss = 1 - IOU \quad (5)$$

DIOU-NMS^[25]在 DIOULoss 的基础上被提出,采用 DIOU 公式进行计算,增加了新的惩罚项以抑制冗余框,如式(6)所示:

$$s_i = \begin{cases} s_i, & IOU - R_{DIOU}(M, B_i) < \epsilon \\ 0, & IOU - R_{DIOU}(M, B_i) \geq \epsilon \end{cases}$$

$$R_{DIOU} = \frac{\rho^2(b, b^{gt})}{c^2} \quad (6)$$

式中, M 为当前得分最高的预测框, B_i 为待处理框, b, b^{gt} 分别为预测框和真实框的中心点位置, ρ 表示计算两点的欧氏距离, c 表示能同时覆盖 b, b^{gt} 两框的最小矩形的对角线长度。

YOLOv5 选用不同算法的预测结果如图 7 所示, NMS 算法的筛选机制导致(a)组测试图片出现了漏检的情况, 而(b)组准确检测出了左上方的划痕, 且异物缺陷的预测框相对于(a)组也更为准确。由于 DIOU 可以最小化预测框和真实框中心点的距离, 使得 DIOU-NMS 收敛速度较快, 预测效果更好。综上所述, 本文选用 DIOU-NMS 来抑制冗余检测框, 以提升模型的缺陷检测效果。

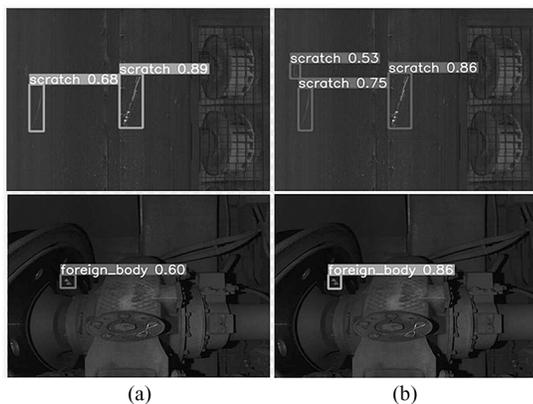


图 7 不同算法预测结果对比:
(a) NMS 算法; (b) DIOU-NMS 算法

Fig. 7 Comparison of prediction results of different algorithms: (a) NMS algorithm;
(b) DIOU-NMS algorithm

4 结果与分析

本文采用 TEDS 系统采集到的动车组关键部件缺陷图片作为基础数据集, 数据增强后的图片总数为 2919 张, 包含 478 张颜色变换扩充图片, 317 张 CycleGAN 生成图片。本文的模型性能对比实验均采用扩充后的数据集进行训练, 缺陷图片使用 LabelImg 工具进行标注, 标签文件为 YOLO 格式。将数据集按 8:1:1 的比例随机划分为训练集、验证集和测试集, 分配结果如表 1 所示, 其中 F、G、S 分别代表异物、漏油和划痕。

表 1 数据集中各类别的数量

Tab. 1 The number of categories in the dataset

Category	Train	Val	Test	Sum
F	997	112	108	1217
G	652	76	78	806
S	725	81	90	896

4.1 实验环境和模型评估指标

实验平台的操作系统为 Windows10, 软件框架为 Pytorch, 硬件环境为 Intel®Core™i7-7820X CPU 和 NVIDIA RTX 2080Ti GPU。相关参数设置: 输入图片尺寸为 640×640 , 训练次数为 300, batch-size 设置为 16, 初始学习率设置为 0.01, 权重衰减系数设置为 0.0005, mosaic 数据增强系数设置为 1.0。

实验采用平均精度 (average precision, AP) 和平均精度均值 (mean average precision, mAP) 作为模型缺陷检测性能的评价指标。AP 为平均精度, 其值为不同阈值下精确率 (Precision) 和召回率 (Recall) 曲线与坐标轴围成的面积, 用于评估单个类别的检测性能; mAP 为所有类别 AP 的平均值, 可以衡量模型的多目标检测性能。采用单张图片的处理延迟 (latency) 作为模型检测速度的评估指标, 延迟时间为推理时间和 NMS 时间之和, latency 越小, 缺陷检测速度越快。计算量和参数量 (params) 分别用来衡量模型的时间复杂度和空间复杂度。相关公式如下所示:

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN}, \quad (7)$$

$$AP = \int_0^1 p(r) dr, mAP = \frac{AP}{Classes_num}, \quad (8)$$

式中, TP 为正确识别的正样本, TN 为正确识别的负样本, FP 为错误识别的正样本, FN 为错误识别的负样本, $p(r)$ 为 Precision-Recall 曲线, $Classes_num$ 为类别数。

4.2 数据增强实验

为了验证数据增强对模型缺陷检测性能的影响, 使用 YOLOv5m 和 YOLOv5-CAtiny 模型分别进行数据增强前后的对比实验, 测试结果如表 2 所示。其中, “√”表示使用扩充后的数据集训练模型; 无标记则表示使用基础数据集训练模型。可以看出, 由于数据增强扩充了漏油和划痕缺陷的图片数量, 降

表 2 不同模型在数据增强前后的检测性能对比

Tab. 2 Comparison of detection performance of different models before and after data enhancement

Model	Data augmentation	mAP	AP		
			F	G	S
YOLOv5m	√	0.78	0.851	0.723	0.766
		0.821	0.868	0.786	0.810
YOLOv5-CAtiny	√	0.771	0.847	0.741	0.726
		0.804	0.877	0.764	0.769

低了样本不平衡比例,使得训练后模型的检测效果较好,各类别的AP值均有所提高,验证了本文数据增强的有效性。

4.3 模型性能对比

为了验证改进网络结构的缺陷检测效果,本文对 MobileNetV3 替换 YOLOv5m 主干网络(YOLOv5-MNtv3)、YOLOv5-MNtv3 + CA 模块(YOLOv5-MCA)、改进 YOLOv5-MCA 的 Neck 结构(YOLOv5-CAtiny)以及原始 YOLOv5m 4 种网络结构分别训练两次,每种模型选择最后一轮 epoch 的训练权重进行测试,缺陷检测结果取两次测试的平均值,不同网络结构的性能对比如表 3 所示。

表 3 4 种网络结构的性能对比

Tab. 3 Performance comparison of four network structures

Model	Params /M	GFL OPs	Latency /ms	mAP	AP		
					F	G	S
YOLOv5m	21.1	50.3	12.0	0.821	0.868	0.786	0.81
YOLOv5-MNtv3	5.47	10.3	8.6	0.778	0.835	0.727	0.774
YOLOv5-MCA	5.49	10.3	8.7	0.807	0.889	0.755	0.776
YOLOv5-CAtiny	4.85	9.6	8.2	0.804	0.877	0.764	0.769

由表 3 可知,对 YOLOv5m 主干网络使用 MobileNetV3 进行替换后,网络模型参数量下降了 74.1%,计算量下降了 79.5%,延迟时间减少了 28.3%,mAP 降低了 4.3%;在此基础上添加轻量级 CA 模块后,提高了模型对于小目标缺陷的检测能力,异物和漏油的 AP 值分别提升了 5.5% 和 2.8%,而且没有增加额外的计算开销;最后对 Neck 结构进行优化,深度可分离卷积替换标准卷积后并没有对检测精度造成很大的影响,并且延迟时间进一步减小。

不同模型的缺陷检测结果对比如图 8 所示,3 种关键部件缺陷的检测结果相似,YOLOv5-CAtiny 并没有出现漏检情况,且异物、漏油缺陷的置信度均高于 YOLOv5m。YOLOv5-CAtiny 相较于原始 YOLOv5m 网络结构,在 mAP 仅降低 1.7% 的情况下,网络模型的参数量减少了 77%,计算量减少了 80.9%,延迟时间减少 31.7%,在保持较高缺陷检测精度的同时,极大降低了模型复杂度,提高了缺陷检测实时性。因此,改进后的网络结构在模型复杂度和速度上均有很大的优势,为海量动车组关键部件的缺陷图片检测和后续移动平台的部署提供了有利条件。

为进一步验证本文改进模型的有效性,在相同数据集和实验环境的条件下,将本文算法同常见的

两阶段目标检测算法 Faster R-CNN、单阶段目标检测算法 SSD、YOLOv3Tiny、YOLOv4Tiny、YOLOv5s 进行对比实验,测试结果如表 4 所示。由表 4 可知,YOLOv5-CAtiny 的检测精度仅次于 YOLOv5m,明显优于其他目标检测算法;在模型复杂度上,计算量略高于 YOLOv4Tiny,但本文算法的 mAP 提升了 18.4%;在检测速度上,YOLOv3Tiny、YOLOv4Tiny 的延迟时间均低于本文算法,但其检测精度均无法满足缺陷检测要求。综上所述,本文算法的缺陷检测性能优于其他算法。

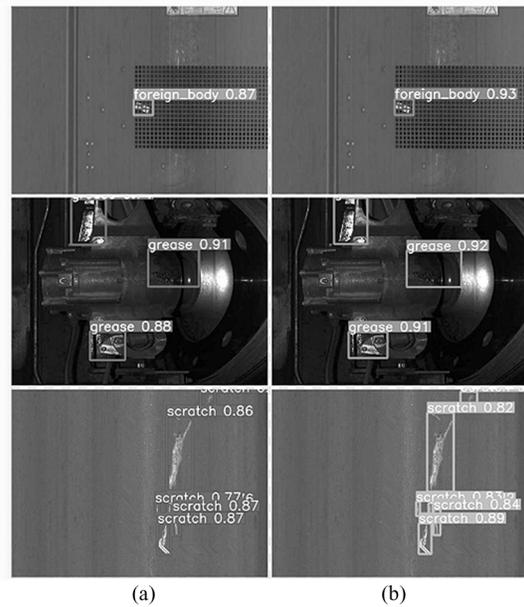


图 8 缺陷检测结果对比图:

(a) YOLOv5m 算法; (b) YOLOv5-CAtiny 算法

Fig. 8 Comparison of defect detection results:

(a) YOLOv5m algorithm;

(b) YOLOv5-CAtiny algorithm

表 4 不同模型的性能对比

Tab. 4 Performance comparison of different models

Model	Params /M	GFL OPs	Latency /ms	mAP	AP		
					F	G	S
Faster R-CNN	137.1	185.2	53.9	0.549	0.80	0.43	0.42
SSD	26.2	31.4	22.8	0.71	0.81	0.67	0.63
YOLOv5s	7.2	16.4	7.9	0.681	0.753	0.676	0.616
YOLOv5m	21.1	50.3	12.0	0.821	0.868	0.786	0.810
YOLOv3Tiny	8.9	13.2	4.5	0.575	0.638	0.468	0.619
YOLOv4Tiny	6.1	8.2	7.0	0.62	0.69	0.59	0.56
YOLOv5-CAtiny	4.85	9.6	8.2	0.804	0.877	0.764	0.769

4.4 NEU-DET 测试实验

NEU-DET 表面缺陷数据集是缺陷检测领域常用的公开数据集,包含 1800 张热轧带钢的 6 种典型

表面缺陷图像^[26]。如图 9 所示, NEU-DET 数据集和动车组关键部件缺陷图像特性和背景具有一定的相似性^[3], 使用该数据集在一定程度上可以反映出模型的泛化能力。鉴于动车组关键部件缺陷样本相对稀缺, 本文选用 NEU-DET 数据集对不同模型进行测试实验, 实验结果如表 5 所示。可以看出, 在检测速度上, YOLOv5-CAtiny 的单张图片检测时间为 13.7 ms, 远低于 Faster R-CNN、MPDD、DDN; 在检测精度上, *mAP* 仅比 MPDD 模型低 1%, 相比其他模型具有一定的优势, 验证了 YOLOv5-CAtiny 的优越性能和实际应用价值。

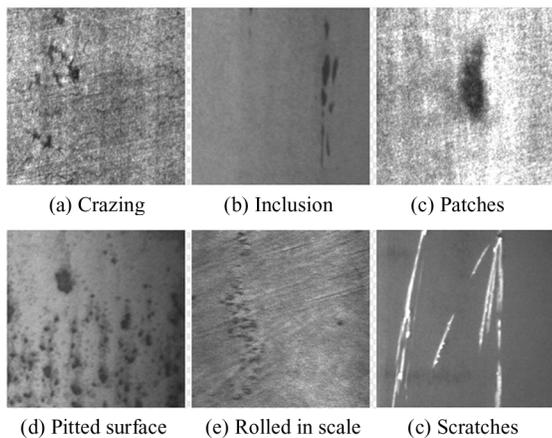


图 9 NEU-DET 表面缺陷数据集

Fig. 9 NEU-DET surface defect database

表 5 不同模型在 NEU-DET 上的检测性能对比

Tab. 5 Comparison of detection performance of different models on NEU-DET

Model	Latency/ms	<i>mAP</i>
Faster R-CNN	57.9	0.601
SSD	14.2	0.615
MPDD ^[3]	203.0	0.765
DDN(ResNet34) ^[26]	50.0	0.748
YOLOv5-CAtiny	13.7	0.764

5 结 论

为解决当前动车组关键部件缺陷检测实时性低、小目标检测效果差和模型复杂而难以在移动端部署的问题, 本文提出一种基于改进 YOLOv5 的缺陷检测方法, 并通过对比实验验证了该方法的有效性。

1) 该方法通过轻量级网络 MobileNetV3-large 对 YOLOv5m 的主干网络进行替换, 在检测精度降低 1.7% 的情况下, 网络模型的参数量和计算量分别减少了 77%、80.9%, 延迟时间减少了 28.3%, 有利

于在资源受限的移动设备上部署。

2) 在主干网络添加了 CA 模块, 训练过程可以融合特征图的位置信息, 减少了异物、漏油小目标缺陷的漏检率, 两者的 *AP* 值分别提升了 5.5%、2.8%。

3) Neck 结构选用深度可分离卷积进行卷积操作, 在保证检测精度的同时, 进一步降低模型计算量。DIOU-NMS 增加了预测框定位的准确性, 同时加快了收敛速度, 提升了模型的缺陷检测性能。

实验结果表明, YOLOv5-CAtiny 在保持较高检测精度的情况下, 大幅度降低了模型的参数量和计算量, 提高了小目标检测准确率和缺陷检测实时性, 检测性能与目前常见的目标检测算法相比也具备一定的优势。在 NEU-DET 数据集上的结果表明, YOLOv5-CAtiny 的 *mAP* 可达到 0.764, 检测速度优于其他方法, 验证了改进算法具有较强的泛化能力和鲁棒性。下一步将对异物缺陷遮挡问题进行研究, 对 YOLOv5 网络结构进一步优化, 提升缺陷检测模型的实用性。

参考文献:

- [1] ZHOU W, SHI T Y, LI P, et al. Defects detection and segmentation of operation safety image of EMU based on convolutional neural network[J]. Journal of The China Railway Society, 2019, 41(10): 76-83.
周雯, 史天运, 李平, 等. 基于卷积神经网络的动车组行车安全图像缺陷检测与分割[J]. 铁道学报, 2019, 41(10): 76-83.
- [2] WEI X, JIANG S, LI Y, et al. Defect detection of pantograph slide based on deep learning and image processing technology[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(3): 947-958.
- [3] ZHAO B, DAI M R, LI P, et al. Defect detection method for electric multiple units key components based on deep learning[J]. IEEE Access, 2020, 8: 136808-136818.
- [4] ZHANG H D, YUAN X, LI D Y, et al. An effective framework using identification and image reconstruction algorithm for train component defect detection[J]. Applied Intelligence, 2022.
- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, USA. New York: IEEE, 2014, 580-587.

- [6] HE K,ZHANG X,REN S,et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,37(9):1904-1916.
- [7] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision,December. 13-16,2015, Santiago,Chile. New York:IEEE,2015,1440-1448.
- [8] XU G J,ZHANG Y,LAI X Y. Recognition approaches of tea bud image based on faster R-CNN depth network[J]. Journal of Optoelectronics • Laser,2020,31(11):1131-1139.
许高建,张蕴,赖小焱.基于Faster R-CNN深度网络的茶叶嫩芽图像识别方法[J].光电子·激光,2020,31(11):1131-1139.
- [9] XU G Y,YI M Y. Improved SSD object detection algorithm based on space-channel attention[J]. Journal of Optoelectronics • Laser,2021,32(9):970-978.
许光宇,尹孟园.基于空间-通道注意力的改进SSD目标检测算法[J].光电子·激光,2021,32(9):970-978.
- [10] LI C Y,YAO J M,LIN Z X,et al. Object detection method based on improved YOLO lightweight network[J]. Laser & Optoelectronics Progress,2020,57(14):141003.
李成跃,姚剑敏,林志贤,等.基于改进YOLO轻量化网络的目标检测方法[J].激光与光电子学进展,2020,57(14):141003.
- [11] LI W J,XU G W,KONG W G,et al. Research on target detection of plant leaf-stem intersection based on improved YOLOv4[J]. Computer Engineering and Applications,2022,58(4):221-228.
李文婧,徐国伟,孔维刚,等.基于改进YOLOv4的植物叶茎交点目标检测研究[J].计算机工程与应用,2022,58(4):221-228.
- [12] HOWARD A,SANDLER M,CHEN B,et al. Searching for mobileNetV3[C]//2019 IEEE International Conference on Computer Vision,October 27-November. 2,2019,Seoul, Korea. New York:IEEE,2019,1314-1324.
- [13] RAN R,XU X H,QIU S H,et al. Review of crack detection methods based on deep convolutional neural networks [J]. Computer Engineering and Applications,2021,57(9):23-35.
冉蓉,徐兴华,邱少华,等.基于深度卷积神经网络的裂纹检测方法综述[J].计算机工程与应用,2021,57(9):23-35.
- [14] ZHU X H,QIAN L P,FU W. Overview of research on image data enhancement technology[J]. Software Guide,2021,20(5):230-236.
朱晓慧,钱丽萍,傅伟.图像数据增强技术研究综述[J].软件导刊,2021,20(5):230-236.
- [15] GOODFELLOW I J,POUGET-ABADIE J,MIRZA M,et al. Generative adversarial nets [C]//International Conference on Neural Information Processing Systems,December 8-13,2014, Montreal, Canada. New York:IEEE,2014:2672-2680.
- [16] LIANG J J,WEI J J,JIANG Z F. Generative adversarial networks GAN overview[J]. Journal of Frontiers of Computer Science and Technology,2020,14(1):1-17.
梁俊杰,韦舰晶,蒋正锋.生成对抗网络GAN综述[J].计算机科学与探索,2020,14(1):1-17.
- [17] ZHU J Y,PARK T,LSOLA P,et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision, October. 27-29,2017, Venice, Italy. New York:IEEE,2017:2242-2251.
- [18] WANG Z Y,ZHANG G,YANG Q,et al. Research on surface defect recognition of copper strip based on YOLOv4 [J]. Journal of Optoelectronics • Laser,2022,33(2):163-170.
王紫玉,张果,杨奇,等.基于YOLOv4的铜带表面缺陷识别研究[J].光电子·激光,2022,33(2):163-170.
- [19] WANG C Y,LIAO H,WU Y H,et al. CSPNet:A new backbone that can enhance learning capability of CNN [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, June 13-19,2020, Seattle, USA. New York:IEEE,2020:1571-1580.
- [20] LI W T,PENG L. Small objects detection algorithm with multi-scale channel attention fusion network[J]. Journal of Frontiers of Computer Science and Technology,2021,15(12):2390-2400.
李文涛,彭力.多尺度通道注意力融合网络的小目标检测算法[J].计算机科学与探索,2021,15(12):2390-2400.
- [21] REN H,WANG X G. Review of attention mechanism[J]. Journal of Computer Applications,2021,41(S1):1-6.
任欢,王旭光.注意力机制综述[J].计算机应用,2021,41(S1):1-6.
- [22] JIE H,LI S,GANG S. Squeeze-and-excitation networks

- [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, USA. New York: IEEE, 2018: 7132-7141.
- [23] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 19-25, 2021, Online. New York: IEEE, 2021: 13713-13722.
- [24] CHOLLET F. Xception: Deep learning with depth-wise separable convolutions[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Hawaii, USA. New York: IEEE, 2017: 1251-1258.
- [25] GONG M, WANG D, ZHAO X, et al. A review of non-maximum suppression algorithms for deep learning target detection[C]//Seventh Symposium on Novel Photoelectric Detection Technology and Application 2020, March 12, 2021, Kunming, China. Bellingham, WA, USA: SPIE, 2021, 11763: 1176332.
- [26] HE Y, SONG K C, MENG Q G, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(4): 1493-1504.

作者简介:

徐国伟 (1972—),男,工学博士,教授,硕士生导师,研究方向为智能控制理论与人工智能。