

# 基于机器视觉的固体火箭发动机金属壳体表面缺陷检测方法

孙旭阳, 沈 飞, 姚怡倩, 张潇誉, 李晓豪

(西安航天动力测控技术研究所, 西安 710025)

**摘要:** 固体火箭发动机金属壳体的缺陷检测对于发动机状态的评估具有重要意义。文章提出一种基于机器视觉的检测方法, 采用基于深度卷积生成对抗网络 (DCGAN) 的数据增强技术和基于卷积注意力模块 (CBAM) 的目标检测技术实现了发动机金属壳体表面缺陷的检测。首先基于已有的少量真实壳体缺陷数据, 通过数据增强技术对数据集进行增强; 再基于增强后的数据集, 利用深度学习算法对缺陷目标进行识别和分类; 最终实现对发动机壳体缺陷目标的检测。通过对多种壳体缺陷进行检测, 验证了该方法的可行性, 且该方法在不影响检测效率的前提下, 识别准确率较传统检测方法提升 5.7%, 模型鲁棒性、泛化性较好, 在实际工程中具有良好的应用前景。

**关键词:** 固体火箭发动机; 缺陷检测; 金属壳体; 数据增强; 深度学习

中图分类号: V640; TP319.41

文献标志码: A

文章编号: 1673-1379(2025)01-0098-11

DOI: [10.12126/see.2024049](https://doi.org/10.12126/see.2024049)

## Surface defect detection method for the metal casing of solid rocket motor based on machine vision

SUN Xuyang, SHEN Fei, YAO Yiqian, ZHANG Xiaoyu, LI Xiaohao

(Xi'an Aerospace Propulsion Testing Technology Research Institute, Xi'an 710025, China)

**Abstract:** The defect detection for the metal casing of solid rocket motor is essential in the assessment of motor conditions. A machine vision-based method for detecting surface defects on metal casings of solid rocket motors was proposed in this paper. The method employed data augmentation using deep convolutional generative adversarial networks (DCGAN), and target detection technique using a convolutional block attention module (CBAM). By utilizing limited real defect data and deep learning algorithms to augment the dataset, the method identified and classified defective targets. The method was tested against various types of casing defects, showing a 5.7% improvement in identification accuracy over conventional detection methods, while maintaining detection efficiency. The model also exhibits good robustness and generalization, indicating potential for practical engineering applications.

**Keywords:** solid rocket motor; defect detection; metal casing; data enhancement; deep learning

收稿日期: 2024-04-25; 修回日期: 2024-12-04

引用格式: 孙旭阳, 沈飞, 姚怡倩, 等. 基于机器视觉的固体火箭发动机金属壳体表面缺陷检测方法[J]. 航天器环境工程, 2025, 42(1): 98-108

SUN X Y, SHEN F, YAO Y Q, et al. Surface defect detection method for the metal casing of solid rocket motor based on machine vision[J]. Spacecraft Environment Engineering, 2025, 42(1): 98-108

## 0 引言

壳体是固体火箭发动机的重要组成部分。发动机工作时,壳体内推进剂被点燃,产生巨大的压力对壳体结构造成冲击;壳体需要能够承受此压力并保持其结构的完整性,以支持发动机的正常运行。金属壳体强度高、耐压性好、耐腐蚀性强且密封性好,在部分固体火箭发动机上得到应用。发动机在工作时对金属壳体的结构完整性要求较高,但是在实际生产及运输过程中,发动机的前裙、筒段中部及后裙部位的 I 至 IV 象限处壳体表面可能存在 1~5 cm 尺寸的杂质、锈蚀或划痕等缺陷。这些缺陷不仅会影响发动机的性能,甚至可能导致发动机爆炸,造成严重损失。因此,固体火箭发动机金属壳体缺陷类别的准确、高效检测对于发动机试验状态的评估具有重要意义。

目前,针对发动机金属壳体表面缺陷的主要检测方式有目视检查和 CT 探伤<sup>[1]</sup>。这两种方法都需要工作人员全程参与,通过目视判断缺陷类型,具有一定主观性,可能导致判定不够准确,且检测效率低。其中 CT 探伤不仅所用设备价格高昂、不利于维护,而且 X 射线会对操作人员的身体健康造成损害。而用机器代替肉眼进行识别、判断的机器视觉方法具有准确、高效、安全的优点,具备良好的应用前景<sup>[1]</sup>。

基于机器视觉的金属壳体表面缺陷检测本质上是一种金属图像缺陷检测。该方法需要一定量的图像数据以及标注好的目标区域真值。但是,随着工艺制造水平的提高,壳体制造的良品率不断提升,造成真实壳体缺陷数据较少,继而导致训练时缺陷检测模型的性能无法提升,容易出现过拟合。因此需要进行数据增强,即基于已有的少量真实图像数据去获取较大量的图像数据。近年来,已有应用机器视觉对图像缺陷进行检测的相关研究。国外, Vengaloor 等<sup>[2]</sup>针对金属表面缺陷检测问题,提出一种基于 YOLOv5 的改进模型,引入特征重校准网络,增强了原模型提取全局特征的能力,提升了系统的特征辨别能力,但改进后网络运算量大,且易将干扰与实际目标混淆; Samma 等<sup>[3]</sup>针对 YOLOv7 模型网络目标检测分类准确率低的问题提出改进,添加对比损失分支功能,增强了模型的分类功能,但对于小目标、小缺陷的识别效果并不理想。国

内,高慧芳等<sup>[4]</sup>针对非金属壳体缺陷的检测,设计深度卷积生成对抗网络(deep convolutional generative adversarial networks, DCGAN)来扩充数据集,克服了缺陷样本数量少、易出现过拟合的困难,但是 DCGAN 模型训练的稳定性差、生成效果不佳,产生的无用数据较多;杨威利<sup>[5]</sup>针对金属表面缺陷数据集过小的问题,提出利用最小二乘条件生成对抗网络(least squares conditional generative adversarial networks, LS-CGAN)来增强数据集,提高了网络模型训练的稳定性,但其针对缺陷识别问题引入的 YOLOv5 模型对于金属表面小缺陷目标的识别准确率较低;崔晶楠等<sup>[6]</sup>针对钢材表面缺陷的检测,提出引入注意力机制的方法,提升了原 YOLOv5 模型的检测能力和效率,但在噪声环境干扰下检测效果并不理想。

综上,基于机器视觉的缺陷检测方法在复杂场景的干扰下特征提取能力不足,导致算法识别准确率低、模型鲁棒性差、运算时间长;特别是在针对小样本数据(例如发动机金属壳体缺陷检测)训练时,还面临着训练难度大和过拟合的难题。而利用对抗网络进行数据增强并不能完全解决现有基于机器视觉的缺陷检测方法存在的特征提取能力弱、模型稳定性差、生成图片质量低的问题,亟需寻找一种特征提取能力强、运行速度快、鲁棒性佳的模型用于实际检测任务。最新提出的 YOLOv8 模型<sup>[7]</sup>引入了一种通过提取多层特征并进行特征融合的方法,可在不降低模型运行效率的前提下极大提升模型的特征提取能力,若合理引入注意力机制,将更有利于壳体缺陷目标的检测,提升模型的抗干扰能力及鲁棒性。

本文提出一种基于机器视觉的固体火箭发动机金属壳体表面缺陷检测方法:首先进行数据增强,通过改进 DCGAN 模型<sup>[8]</sup>、引入多尺度交叉注意力(multi-scale cross-attention, MSCA)机制<sup>[9]</sup>解决原有模型提取能力弱、模型稳定性差及生成图片质量低的问题;然后基于已有的一些金属壳体缺陷数据进行训练,生成大量的金属壳体缺陷数据,以解决样本数量少和过拟合的问题;最后针对试验现场环境复杂、特征提取难的问题,提出一种融合 YOLOv8 模型<sup>[7]</sup>和卷积注意力模块(convolutional block attention module, CBAM)<sup>[10]</sup>的改进算法模型。

## 1 研究背景

### 1.1 数据增强基础模型选择

传统的数据增强方法是对已有的图像数据进行亮度/对比度调整、旋转、平移、引入噪声及缩放尺寸等操作,从而得到新的图像数据。这虽然可以扩充图像样本的数量,但在后续缺陷检测模型的训练过程中会发现这些新数据对提升模型检测能力帮助不大。

近年来,在数据增强任务中,基于深度学习的生成式网络模型取得了比传统数据增强方法更好的效果。2014年,Goodfellow等<sup>[11]</sup>首次提出生成对抗网络(generative adversarial nets, GAN)模型,对生成器和判别器两个模型进行交替训练,通过二者之间的对抗和博弈生成真假难辨的样本数据;但GAN的训练结果不稳定且生成目标不可控,极大浪费了计算资源。2014年,Mirza等<sup>[12]</sup>提出加入条件限制的GAN,即条件生成对抗网络(conditional generative adversarial nets, CGAN),可以生成具有特定性质的样本,在提升生成图像质量的同时减少计算量;但由于CGAN缺乏特征提取机制,最终会生成较多无意义样本。2015年,Radford等<sup>[8]</sup>提出DCGAN模型,用卷积层替代GAN中的全连接层,极大提升了模型的稳定性,取得良好的生成效果;但仍存在特征提取能力弱、无用计算量较大的问题。2017年,Arjovsky等<sup>[13]</sup>提出WGAN(Wasserstein GAN)模型,进一步增强了GAN模型的稳定性,但也增加了模型训练的复杂度。2018年,Zhang等<sup>[14]</sup>提出SAGAN模型,首次将自注意力(self-attention)机制引入GAN,使得模型可以更好地捕获全局信息,但是模型复杂度大大提升,计算成本较高。

本文通过与其他生成模型的对比,认为DCGAN模型具备特征提取功能、性能稳定且参数轻量,有助于解决本文所面临的实际问题,方便工程部署,因此选择该模型作为数据增强任务的基础模型。

### 1.2 缺陷检测方法选择

目前,缺陷检测领域使用较为广泛的方法包括Faster R-CNN<sup>[15]</sup>、SSD<sup>[16]</sup>、Mask R-CNN<sup>[17]</sup>及YOLO系列模型<sup>[18-20]</sup>。其中,YOLO系列模型创新性地将缺陷检测问题转化为数学回归问题,通过模型特有

的端到端的目标检测方法,可一次性预测出目标的位置和类别;相比于其他模型,流程更加简单,训练和推理的复杂性低,可以提升识别效率,实现缺陷的实时检测。2023年,章东平等<sup>[21]</sup>针对飞机蒙皮缺陷的检测,提出一种基于YOLOv8的改进模型,通过集成残差网络及引入三支注意力机制,提升了模型的特征提取能力,取得了不错的检测效果。2023年,王耀光<sup>[22]</sup>针对手机屏幕缺陷的检测,提出一种基于YOLOv8的改进模型,通过替换主干网络及引入注意力机制,提升了算法的检测性能;但是模型的抗干扰能力稍显不足。2024年,苏佳等<sup>[23]</sup>针对工业表面缺陷检测问题,提出一种基于YOLOv8的改进策略,引入多尺度注意力模块,提升了模型的抗干扰能力。

综上,本文选择YOLOv8模型作为金属壳体缺陷检测任务的基础模型。

## 2 基于机器视觉的金属壳体缺陷检测方法

### 2.1 方案简述

在拍摄发动机金属壳体图像时,采用机械臂结合轨道的方式:将相机固定在机械臂上,随机臂沿着轨道由发动机前裙至后裙、I象限至IV象限进行顺序移动拍摄;相机每移动1次,计算机都会记录其相应的三轴坐标,从而确定所获取图像相对发动机壳体的实际位置。

在进行图像数据处理时,对相机拍摄的发动机壳体图像进行数据增强,并训练网络模型,最终实现对于壳体缺陷进行实时检测、分类以及定位。

### 2.2 数据增强方法

DCGAN的网络结构主要包括生成器和判别器。向生成器输入随机噪声向量,经过一系列转置卷积后可将该向量映射成一张“逼真的”图像,即生成“假图像”。向判别器输入生成的“假图像”或真实图像,经过一系列卷积运算,判别器得出输入图像是真实图像的概率,再通过反向传播更新网络参数,可使判别器无法辨别图像是真实数据还是生成数据,从而达到数据增强的效果。

尽管DCGAN模型在一般数据增强任务中取得了较好的成效,但在金属壳体缺陷数据增强任务中仍无法达到所需的生成效果,尤其难以捕捉任务



须关注的重要特征。因此,本文在金属壳体缺陷数据增强任务中,通过引入多尺度交叉注意力(MSCA)

机制<sup>[9]</sup>对 DCGAN 模型进行改进和优化。改进后的 DCGAN 模型网络结构如图 1 所示。

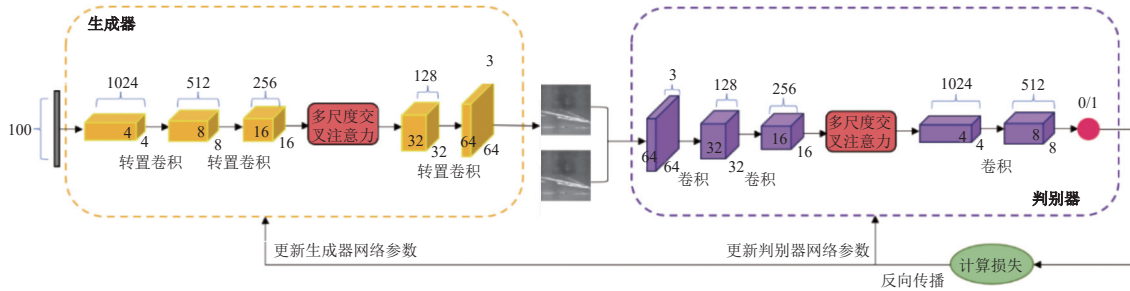


图 1 改进后的 DCGAN 模型网络结构

Fig. 1 Architectural diagram of the improved DCGAN model

本文所使用的图像数据为分辨率为  $64 \times 64$  的壳体缺陷图像。如图 1 所示,首先,将维度为 100 的随机噪声向量输入生成器,通过一个全连接层将向量映射为 1024 通道的分辨率为  $4 \times 4$  的特征图;再经过两个转置卷积层,使用批归一化和 ReLU 激活函数,将特征图上采样至分辨率  $16 \times 16$ 。此时,特征图已经包含一定空间信息,但尚未上采样至最高分辨率。

如图 2 所示,加入多尺度交叉注意力机制,即:根据前面层的输出特征图计算注意力图谱  $X$ ,以突出重要的特征区域;将计算得到的注意力图谱与原始图像数据特征图相乘,促使后续的特征图层更加关注这些重要区域,使模型在后续上采样过程中能更加精细地恢复图像细节,提高生成图像的质量,同时保持生成效率,避免过度增加计算负担;最后,通过两个转置卷积层上采样得到分辨率为  $64 \times 64$  的生成图像  $Y$ 。

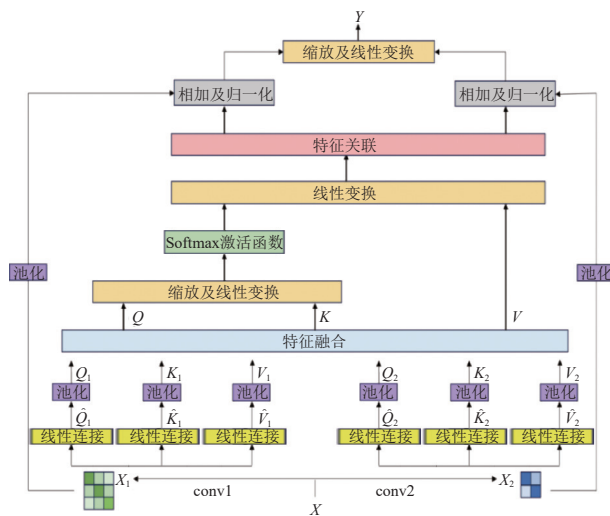


图 2 多尺度交叉注意力机制

Fig. 2 MSCA mechanism

将生成器生成的分辨率为  $64 \times 64$  的图像或真实数据图像输入判别器网络,通过两个卷积层将图像下采样至分辨率为  $16 \times 16$  的特征图;加入多尺度交叉注意力机制,使得判别器可以更有效地捕捉到生成图像的微小差异,提高判别能力;将注意力机制获取的特征图经过两个卷积层下采样至分辨率为  $4 \times 4$  的特征图,经过一个全连接层将向量映射为单一的预测值,用于判断图像是生成的还是真实的;最后,通过反向传播更新网络参数,使得判别器无法辨别图像是真实数据还是生成数据,从而达到数据增强的效果。

### 2.3 金属壳体缺陷检测方法

不同于常规的目标检测任务,发动机金属壳体缺陷检测任务的主要难点为:

1)要求准确识别缺陷的位置和种类,而壳体缺陷检测的目标小、种类多,故模型应具备对弱小目标的位置和种类进行快速、准确识别的能力。

2)在复杂场景的噪声干扰下,缺陷目标特征提取难。虽然可通过数据增强网络解决样本数量少、模型训练难度大、容易过拟合的问题,但模型的泛化性、鲁棒性以及复杂环境下的特征提取能力仍需增强。

因此,本文在基于深度学习的金属壳体缺陷检测中,结合壳体缺陷检测任务的特点进行模型的基础选型与适应性改进。在模型训练阶段,给予缺陷目标更高的关注度,以便于训练过程更多地聚焦于缺陷目标的特征提取。此外,针对复杂场景下目标提取难的问题,使用数据增强网络进行数据增强,模型训练样本量的增加使经过训练后的模型特征

提取能力得到提升。

尽管相比于其他模型, YOLOv8 模型流程简单, 训练和推理的复杂度低, 在一般的目标检测任务中取得了较好的效果, 但在金属壳体缺陷识别任务中仍无法达到较高的检测准确率和检测速度, 尤其是对于复杂环境干扰下的多类型目标识别。因此, 本文基于 YOLOv8 模型进行适应性改进, 如图 3 所示, 相比于 YOLOv8 模型的原始结构, 主要对模型进行了两部分改动(见图中红色模块)。

首先, 为提升检测的准确率和效率, 使其更适应于壳体缺陷检测任务的模型训练。在图像数据输入进行 1 次卷积后, 加入卷积注意力模块<sup>[10]</sup>。如图 4 所示, 卷积注意力模块的引入使得改进模型专注于图像中的重要特征区域, 提高了模型对于缺陷目标的关注度; 增强了特征层的输出及模型的表达能力, 使得缺陷目标特征更为明显; 提升了模型的泛化性和抗干扰能力, 使得模型可以适应不同类型、

复杂度和噪声环境下的图像数据。

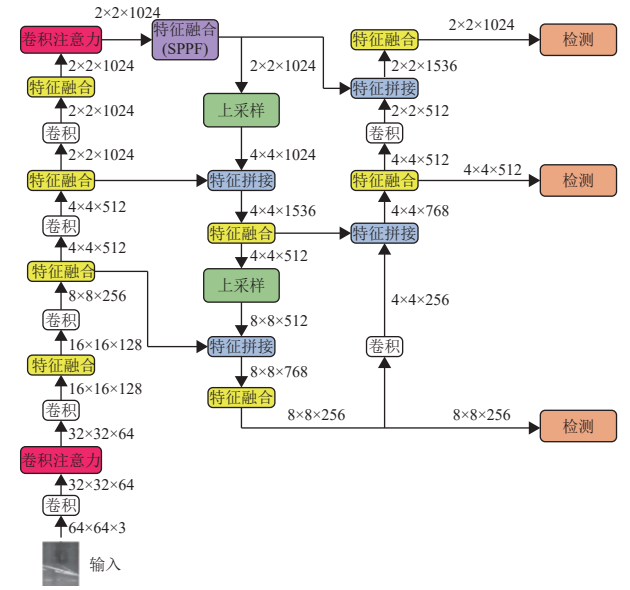


图 3 本文改进后的 YOLOv8 模型

Fig. 3 Architectural diagram of the improved YOLOv8 model

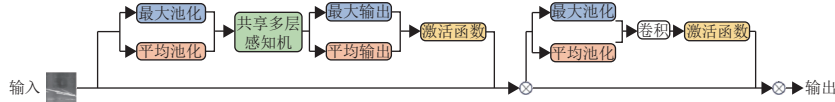


图 4 卷积注意力模块

Fig. 4 Convolutional block attention module

其次, 为提升模型的鲁棒性和复杂环境下的抗干扰能力, 使其更加适应于实际工程部署, 在进行多尺度特征融合之前加入卷积注意力模块, 使得注意力机制可以关注到整个主干网络部分的特征图, 具有多尺度的全局视野; 在复杂环境的干扰下, 模型特征提取能力更强, 抗干扰能力得到提升。

## 2.4 损失函数

基于机器视觉的壳体缺陷检测方法的性能与检测效果跟模型的训练方式紧密相关。其中, 损失函数对最终训练效果的影响很大。为提升模型最终的检测效果, 需要对训练过程中的损失函数进行改进。

### 2.4.1 数据增强网络损失函数

改进后 DCGAN 网络的损失函数为

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

式中:  $G$  代表生成器,  $G(z; \theta_g)$  中  $\theta_g$  为多层感知器的权重参数, 生成器输出假图像;  $D$  代表判别器,  $D(x;$

$\theta_d$ ) 中  $\theta_d$  为感知器权重参数(向判别器输入一张真实数据图像或生成器生成的假图像, 判别器判定图像为真时输出接近 1, 反之输出接近于 0);  $p_{\text{data}}(x)$  为真实数据  $x$  服从的概率分布;  $p_z(z)$  为随机噪声  $z$  服从的概率分布。

给定生成器  $G$  来训练判别器  $D$ , 使输入真实图像时判别器的目标函数

$$\max_D V(D, G) = E_{x \sim p_{\text{data}}(x)} [\lg D(x)] + E_{z \sim p_z(z)} [\lg (1 - D(G(z)))] \quad (2)$$

最大化。因此, 需要将  $E_{x \sim p_{\text{data}}(x)} [\lg D(x)]$  最大化, 理想状态是: 在输入真实图像时  $D(x)$  最大化, 输入生成器  $G$  生成的假图像时  $D(G(z))$  最小化。

给定判别器  $D$  来训练生成器  $G$ , 使得目标函数

$$\min_G V(D, G) = E_{x \sim p_{\text{data}}(x)} [\lg D(x)] + E_{z \sim p_z(z)} [\lg (1 - D(G(z)))] \quad (3)$$

最小化。此时,  $E_{x \sim p_{\text{data}}(x)} [\lg D(x)]$  为常数, 因此需优化  $E_{z \sim p_z(z)} [\lg (1 - D(G(z)))]$  至最小, 即最小化  $\lg (1 - D(G(z)))$ 。但在实际训练中, 由于初始生成的图像与真实图像差异大, 过于容易判别真假, 将导致无

法训练, 所以通过最大化 $\lg(D(G(z)))$ 来训练模型。

最终, 通过交替训练  $D$  和  $G$  构建自监督场景, 使得两者都达到全局最优解。

#### 2.4.2 缺陷检测模型损失函数

在壳体缺陷目标检测任务中, 缺陷目标小、种类多, 因此识别、分类的错误率高。但是, 在无缺陷壳体区域像素特征相似, 无强烈外界干扰下识别的错误率低; 在缺陷目标出现时, 与无缺陷区域的像素特征差别较大, 易识别。因此, 需要提升小目标区域在损失函数中的权重, 使得模型在训练过程中更关注小目标点。

常用的多分类交叉熵损失函数定义为

$$\text{Loss} = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \lg(p_{ic}), \quad (4)$$

式中:  $N$ 、 $M$  分别为样本数量和类别数量;  $y$  代表像素标签, 样本  $i$  的真实类别为  $c$ , 则其值为 1, 否则为 0;  $p$  表示样本  $i$  属于该标签  $c$  的概率。改进后的损失函数定义为

$$\text{Loss} = -\frac{1}{N} \sum_i \sum_{c=1}^M w_{ic} y_{ic} \lg(p_{ic}), \quad (5)$$

式中  $w_i$  为权重系数, 当像素点在无缺陷壳体区域时赋值为 1, 在缺陷目标区域时赋值大于 1, 根据每一类别  $c$  出现的概率来适应性配置权重。

### 3 实验结果与分析

#### 3.1 数据增强网络实验结果与分析

本文使用的金属壳体表面缺陷数据要求分辨率高、真实性好, 故选取工业相机+远心镜头的方案进行数据集图像的获取。具体选择了 Bummer 公司的 VCXG-124 CMOS 工业相机+OPTO ENGINEERING 公司的远心镜头的组合, 相机具备 1200 万像素分辨率和  $3.45 \mu\text{m} \times 3.45 \mu\text{m}$  的像元尺寸, 保证了所获得数据的高分辨率; 配套镜头使得成像更加真实, 保证了数据集的质量。在 1 台发动机上拍摄了 240 张金属壳体缺陷图像作为数据集; 在其他不同类型的发动机上拍摄了 24 张图像数据作为验证集。

所获得壳体缺陷的类型主要为杂质、锈蚀和划痕。图 5(a) 为典型的内含杂质图像, 说明壳体在加工过程中发生了严重的磕碰或者掺杂进了其他成分。图 5(b) 为典型的锈蚀图像, 显示壳体被化学物质腐蚀形成缺陷。图 5(c) 为典型的划痕图像, 显示

壳体在加工过程中被尖锐物体划伤损坏。

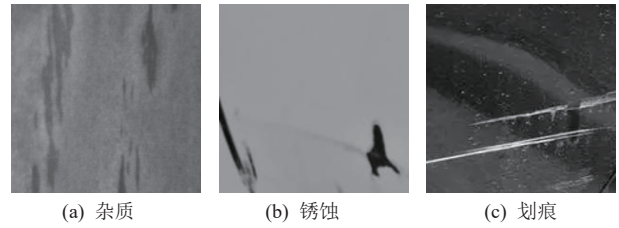


图 5 金属壳体常见的三类缺陷  
Fig. 5 Three common types of defects on metal casings

如前所述, 使用传统方法进行数据增强对模型检测性能、鲁棒性、泛化性的提升有限, 无法达到预期的效果。因此, 使用本文的数据增强方法, 通过引入多尺度交叉注意力机制对 DCGAN 网络模型进行改进, 提升原模型对于缺陷目标区域的关注度, 提高生成图像的质量。图 6 为 DCGAN 初始模型生成图像; 图 7 为改进后 DCGAN 模型生成的图像。可以看到, 后者生成图像的质量得到了提升。

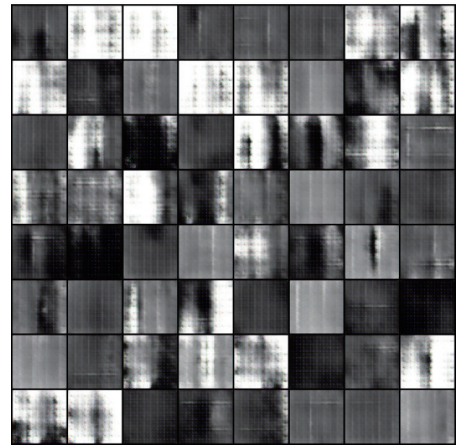


图 6 DCGAN 模型生成图像  
Fig. 6 Generated images by the DCGAN model

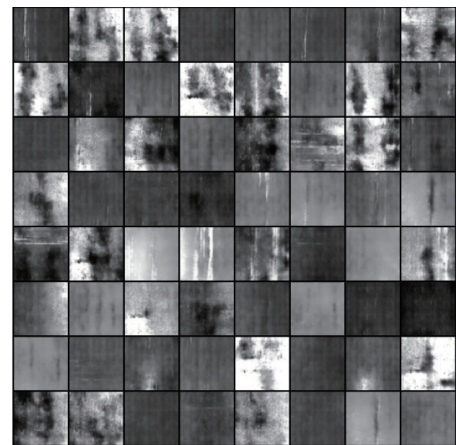


图 7 改进后 DCGAN 模型生成图像  
Fig. 7 Generated images by the improved DCGAN model

为了更加客观、定量地评估改进后 DCGAN 模型生成图像的质量以及模型的性能, 本研究使用



FID<sup>[24]</sup>(Fréchet Inception Distance)参数对模型生成图像的质量进行评价。FID 计算真实图像和生成图像之间的特征分布相似性:首先利用 Inception 网络提取真实图像和生成图像的特征,然后对于提取的特征空间进行高斯建模,最后计算两个多维高斯分布的特征距离,从而获得 FID 分数——分数越低表示生成图像质量越好。FID 的计算公式为

$$FID = \|\mu_r - \mu_g\|^2 + T_r[\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2}], \quad (6)$$

式中: $\mu_r$  和  $\mu_g$  分别为真实图像和生成图像的特征均值; $\Sigma_r$  和  $\Sigma_g$  分别为真实图像特征和生成图像特征的协方差矩阵。

作为对比实验,分别对 DCGAN<sup>[8]</sup>、WGAN<sup>[13]</sup>、CGAN<sup>[12]</sup>、SAGAN<sup>[14]</sup> 及本文数据增强模型等 5 种网络模型进行 3000 次迭代,每 100 次计算 1 次生成图像的 FID 值,最终求得 FID 平均值及最小值,如表 1 所示。可以看到,本文提出的改进模型的 FID 平均值和最小值分别为 112.31 和 79.86,均低于其他模型的结果,说明改进后的模型生成图像质量更高,生成效果更好。

表 1 对比试验 FID 值

Table 1 Comparison of FID values in different experiments

| 模型                    | FID 平均值 | FID 最小值 |
|-----------------------|---------|---------|
| DCGAN <sup>[8]</sup>  | 167.51  | 124.47  |
| WGAN <sup>[13]</sup>  | 148.41  | 110.65  |
| CGAN <sup>[12]</sup>  | 144.64  | 101.36  |
| SAGAN <sup>[14]</sup> | 131.63  | 95.63   |
| 改进后 DCGAN             | 112.31  | 79.86   |

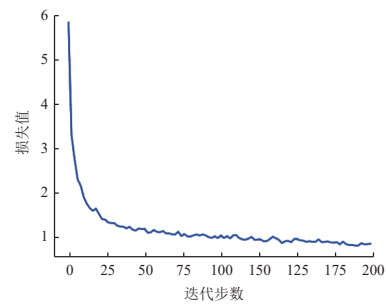
使用 LableMe<sup>[25]</sup> 标注工具手动标注出数据集中的缺陷目标位置及种类。将在 1 台发动机上拍摄的 240 张金属壳体缺陷图像数据进行划分,其中 192 张作为训练集,48 张作为测试集,验证模型的精度;使用在其他不同类型发动机上采集到的 24 张图像数据作为验证集,验证模型的泛化性。

### 3.2 壳体缺陷检测模型实验结果与分析

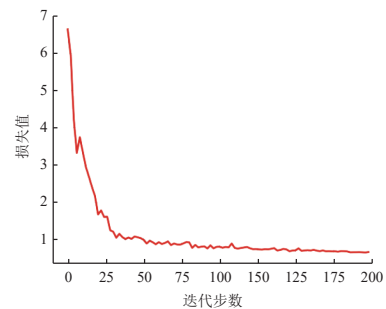
实验平台为 NVIDIA GeForce GTX 2080Ti GPU (11 GByte 显存),模型训练 1000 轮,初始学习率为 0.01,动量为 0.937。

#### 1) 测试本文缺陷检测模型的检测性能

模型训练损失函数曲线如图 8 所示,可以看到曲线收敛,且训练集与测试集曲线相差很小,表明模型学习效果良好、性能优秀。



(a) 训练集损失函数曲线



(b) 测试集损失函数曲线

图 8 模型训练损失函数曲线对比

Fig. 8 Comparison of loss function curves during model training

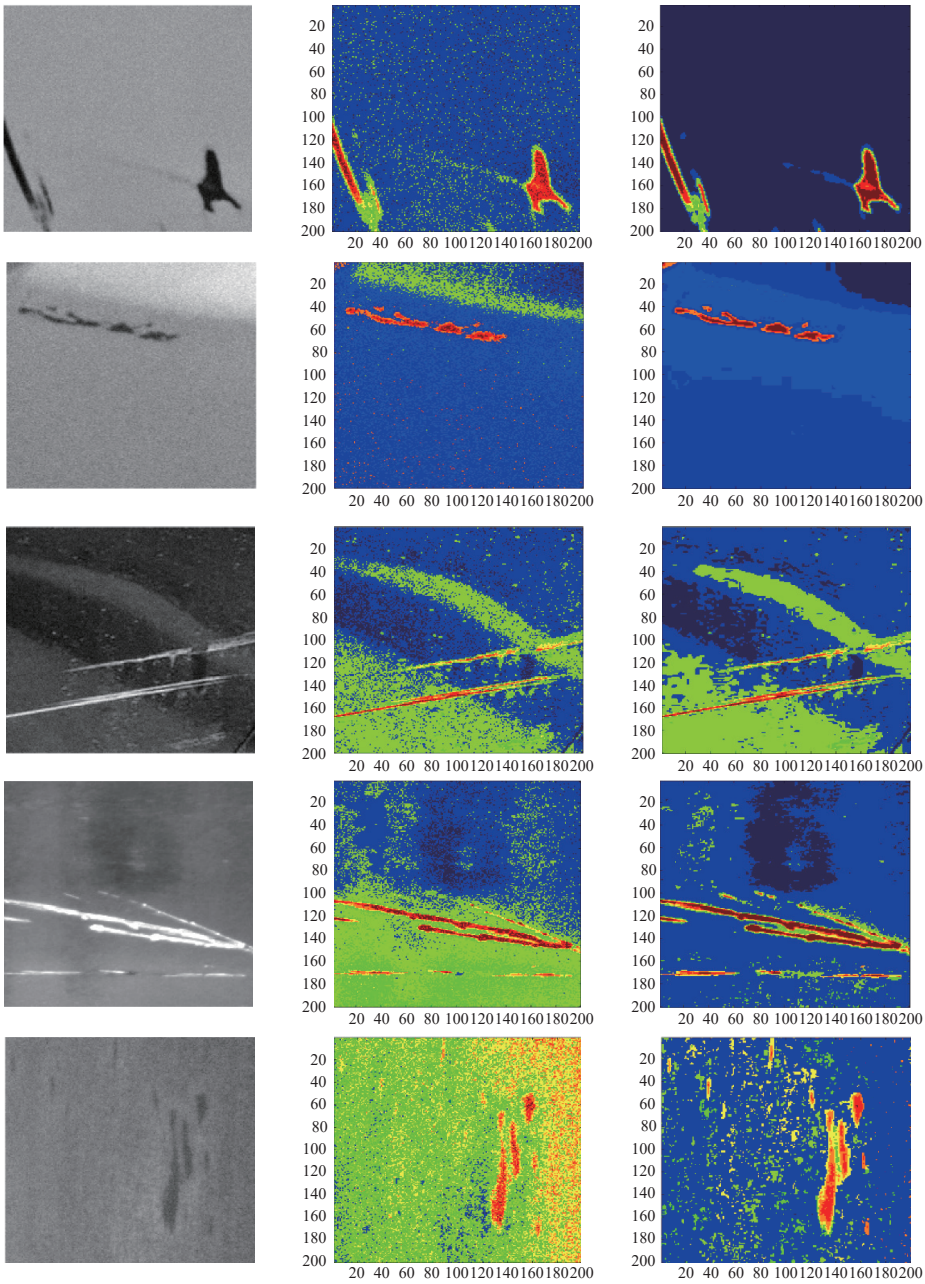
基于本文提出的改进后 DCGAN 模型增强过的数据集,使用 YOLO 系列模型经典评价指标对壳体缺陷检测模型的性能进行评估。指标包括:精确度(Precision)、召回率(Recall)、精确度-召回率综合指标(F1-Score)以及多分类平均精确度(mAP)——mAP50 和 mAP50-95等;另外,对于模型的检测效率用 FPS(即 1 s 内可处理的图像样本数量)指标来评估。对比本文所用的改进后 YOLOv8 模型与相关研究文献中模型的评价指标,结果如表 2 所示。可以看到,本文模型相比于原始 YOLOv8 模型检测精度提升较大,与其他目标检测模型相比,在不影响检测效率的前提下,检测准确率更高,更适应于发动机金属壳体缺陷检测任务。

同时,为了直观展示加入卷积注意力模块后模型特征提取能力的提升,绘制改进模型特征注意力热力图(图 9)。图中,左列为待检测原图,中间列为原始 YOLOv8 模型检测特征注意力热力图,右列为本文方法检测特征注意力热力图。可以看到,通过引入注意力机制能够提升缺陷目标区域的权重(红色为模型特征提取时需要关注的缺陷目标特征),使得模型不易受复杂环境干扰的影响,更加突出主题目标,鲁棒性佳。

此外, 将本文方法在由其他类型发动机上摄取的 24 张真实缺陷图像数据构成的验证集上进行了测试验证, 结果表明, 取得了 0.865 7 的检测精度, 证明了模型的泛化性佳。

表 2 模型性能对比  
Table 2 Comparison of model performance

| 模型                         | Precision | Recall  | F1      | mAP50   | mAP50-95 | FPS |
|----------------------------|-----------|---------|---------|---------|----------|-----|
| Fast R-CNN <sup>[15]</sup> | 0.846 5   | 0.667 5 | 0.746 4 | 0.775 9 | 0.301 1  | 10  |
| YOLOv8 <sup>[7]</sup>      | 0.824 7   | 0.703 9 | 0.759 5 | 0.784 6 | 0.285 6  | 27  |
| 章东平等的方法 <sup>[21]</sup>    | 0.860 1   | 0.675 1 | 0.756 5 | 0.786 2 | 0.312 3  | 27  |
| 王耀光的方法 <sup>[22]</sup>     | 0.859 8   | 0.661 3 | 0.747 6 | 0.778 5 | 0.305 1  | 26  |
| 苏佳等的方法 <sup>[23]</sup>     | 0.863 7   | 0.673 2 | 0.756 6 | 0.787 6 | 0.312 6  | 29  |
| 改进后YOLOv8                  | 0.876 4   | 0.736 8 | 0.801 6 | 0.823 7 | 0.320 1  | 29  |





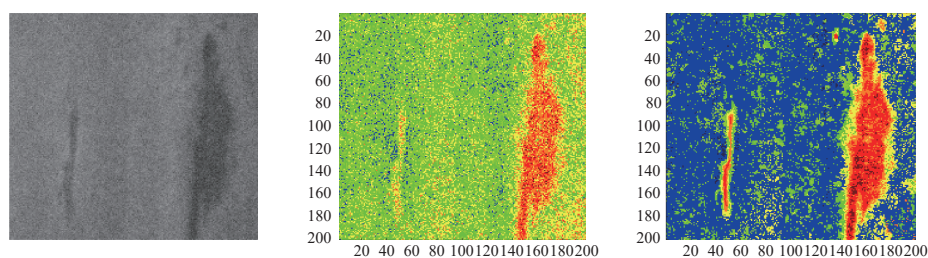


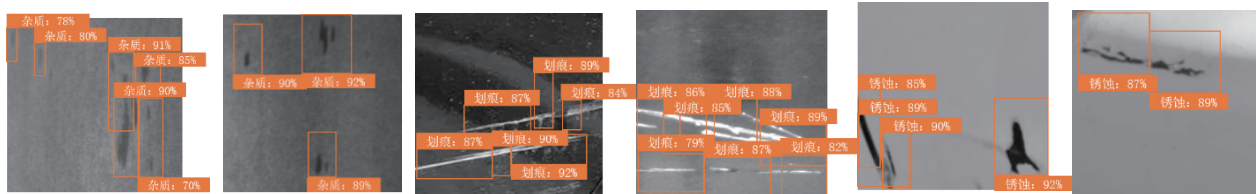
图9 改进模型特征注意力热力图

Fig. 9 Feature attention heat maps of the improved model

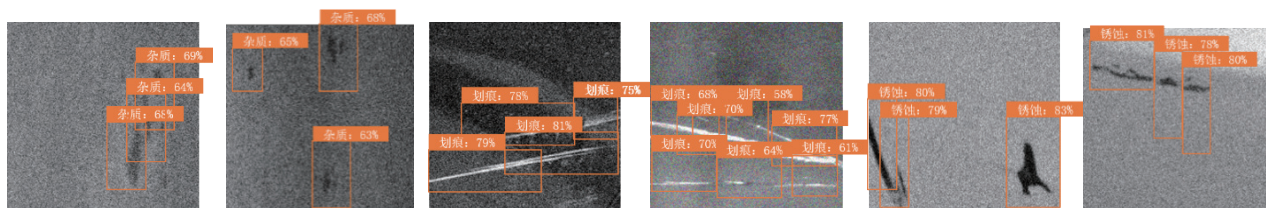
## 2) 测试本文缺陷检测模型的抗干扰能力

向本文数据增强方法扩充过的数据集中引入光照噪声干扰——用加入高斯噪声的方法来模拟光照干扰,再对比本文模型与原始YOLOv8模型进

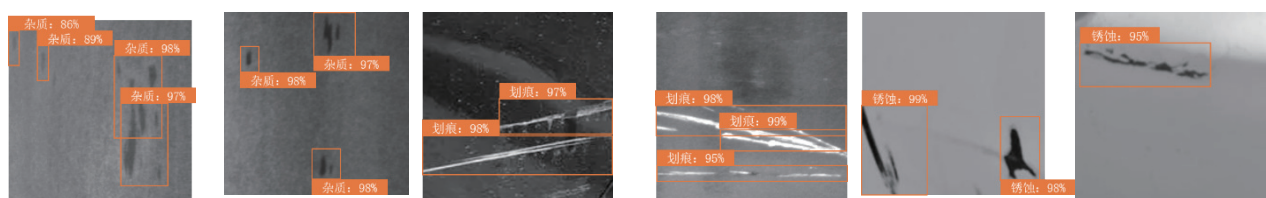
行缺陷目标检测的结果,如图10所示:在复杂环境干扰、目标弱小的情况下,改进后模型仍可准确地对壳体缺陷目标进行检测和分类,虽类型识别概率有所下降但不影响最终判读,效果优于改进前。



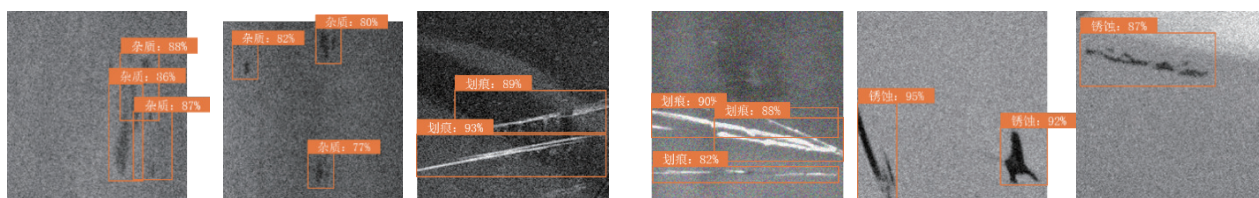
(a) 无干扰数据YOLOv8模型检测效果



(b) 有干扰数据YOLOv8模型检测效果



(c) 无干扰数据本文方法检测效果



(d) 有干扰数据本文方法检测效果

图10 改进后模型的金属壳体缺陷检测效果

Fig. 10 Effectiveness of the improved model for detecting defects on metal casings

## 3.3 消融实验

为了验证本文提出的引入多尺度交叉注意力机制的DCGAN数据增强模型以及引入YOLOv8模型的卷积注意力模块对金属壳体缺陷检测准确

率的提升作用,进行消融实验,结果如表3所示。可以看出:改进DCGAN模型的引入可以提升数据集的质量、数量,在避免过拟合的同时增强检测模型的识别准确性、鲁棒性和泛化性,对最终结果有

较大改善，mAP50 值提升较大；卷积注意力模块的引入使得模型更加专注于图像中的重要特征区域，进一步提升了模型对于缺陷目标检测的效果，mAP50 值再次得到提升。

表 3 基于改进模型的消融实验结果  
Table 3 Ablation study results with the improved model

| 方法                 | Precision | Recall  | F1      | mAP50   | mAP50-95 | FPS |
|--------------------|-----------|---------|---------|---------|----------|-----|
| YOLOv8+DCGAN       | 0.806 5   | 0.665 7 | 0.729 4 | 0.758 4 | 0.285 4  | 28  |
| YOLOv8+改进后DCGAN    | 0.824 7   | 0.703 9 | 0.759 5 | 0.784 6 | 0.285 6  | 27  |
| 改进后YOLOv8+DCGAN    | 0.845 8   | 0.692 5 | 0.761 5 | 0.803 7 | 0.296 4  | 28  |
| 改进后YOLOv8+改进后DCGAN | 0.876 4   | 0.738 6 | 0.801 6 | 0.823 7 | 0.320 1  | 29  |

4 结束语

本文面向固体火箭发动机金属壳体表面缺陷检测，针对其中图像数据缺乏的问题，基于 DCGAN 引入多尺度交叉注意力机制，提出了一种改进型图像生成模型；为提升模型的检测性能，基于 YOLOv8 模型引入了卷积注意力模块，提出了一种改进型目标检测模型。通过在同一发动机壳体缺陷测试数据集上进行测试表明，本文方法相比于其他模型及改进前的方法，在不影响检测效率的前提下，检测精度提升了 5.7%，更适应于发动机金属壳体表面缺陷检测任务，在复杂试验环境下识别准确率更高、抗干扰能力更强；并且通过在由其他不同种类发动机壳体上采集到的验证集数据上进行试验，证明本方法的泛化性较好，具有较高的实际工程应用价值。

参考文献（References）

[1] 孙旭阳, 沈飞, 谢俊彦. 基于机器视觉的固体火箭发动机喷管喉径测量方法[J]. 固体火箭技术, 2023, 46(1): 119-127.  
SUN X Y, SHEN F, XIE J Y. Measurement method for nozzle throat diameter of solid rocket motor based on machine vision[J]. Journal of Solid Rocket Technology, 2023, 46(1): 119-127.

[2] VENGALLOOR R, MURALIDHAR R. Deep learning based feature discriminability boosted concurrent metal surface defect detection system using YOLOv-5s-FRN[J]. The International Arab Journal of Information Technology, 2024, 21(1): 94-106.

[3] SAMMA H, AL-AZANI S, LUQMAN H, et al. Contrastive-based YOLOv7 for personal protective equipment detection[J]. Neural Computing & Applications, 2024, 36(5): 2445-2457.

[4] 高慧芳, 金永, 柴国强, 等. 基于 DCGAN 与 YOLOv5s 的火箭弹非金属壳体缺陷识别方法研究[J]. 固体火箭技术, 2022, 45(6): 949-955.

GAO H F, JIN Y, CHAI G Q, et al. Defect identification method of rocket non-metallic shell based on DCGAN and YOLOv5s[J]. Journal of Solid Rocket Technology, 2022, 45(6): 949-955.

[5] 杨威利. 基于深度学习的复杂金属表面缺陷检测技术研究[D]. 西安: 西安工业大学, 2023. doi: 10.27391/d.cnki.gxagu.2023.000490.

[6] 崔晶楠, 黄春艳, 李艳玲. 改进 YOLOv5s 算法的钢材表面缺陷检测[J/OL]. 电子科技. [2024-02-25]. <https://doi.org/10.16180/j.cnki.issn1007-7820.2024.12.008>.

CUI J N, HUANG C Y, LI Y L. Improvement of YOLOv5s algorithm for steel surface defect detection[J/OL]. Electronic Science and Technology. [2024-02-25]. <https://doi.org/10.16180/j.cnki.issn1007-7820.2024.12.008>.

[7] JOCHER G, CHAURASIA A, QIU J. Ultralytics YOLO (Version 8.0. 0)[CP/OL]. [2024-04-24]. <https://github.com/ultralytics/ultralytics>.

[8] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. Computer Science, 2015. doi: 10.48550/arXiv.1511.06434.

[9] ZHU S, DU B, ZHANG L, et al. Attention-based multiscale residual adaptation network for cross-scene classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-15.

[10] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//15th European Conference on Computer Vision (ECCV-2018). Munich, 2018: 3-19.

[11] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, 2014: 2672-2680.

[12] MIRZA M, OSINDERO S. Conditional generative adversarial nets[J]. Computer Science, 2014. doi: 10.48550/arXiv.1411.1784.

[13] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[C]//Proceedings of the

- 34th International Conference on Machine Learning. Sydney, 2017: 214-223.
- [14] ZHANG H, GOODFELLOW I, METAXAS D, et al. Self-attention generative adversarial networks[C]//Proceedings of the 36th International Conference on Machine Learning. Long Beach, CA, USA, 2019: 7354-7363.
- [15] 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 and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//14th European Conference on Computer Vision. Amsterdam, 2016: 21-37.
- [17] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice, 2017: 2961-2969.
- [18] 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, NV, USA, 2016: 779-788.
- [19] JOCHER G, STOKEN A, BOROVEC J, et al. Ultralytics/yolov5: v5.0-YOLOv5-P6 1280 models, AWS, Supervise.ly and YouTube integrations[Z], 2021. doi: [10.5281/zenodo.4679653](https://doi.org/10.5281/zenodo.4679653).
- [20] JOCHER G, CHAURASIA A, STOKEN A, et al. Ultralytics/yolov5: v7.0-YOLOv5 SOTA realtime instance segmentation[Z], 2022. doi: [10.5281/zenodo.7347926](https://doi.org/10.5281/zenodo.7347926).
- [21] 章东平, 王杼涛, 夏岳键, 等. 基于改进 YOLOv8 的飞机蒙皮缺陷检测算法[J]. 北京航空航天大学学报. doi: [10.13700/j.bh.1001-5965.2023.0744](https://doi.org/10.13700/j.bh.1001-5965.2023.0744).
- ZHANG D P, WANG Z T, XIA Y J, et al. Aircraft skin defect detection algorithm based on enhanced YOLOv8[J]. *Journal of Beijing University of Aeronautics and Astronautics*. doi: [10.13700/j.bh.1001-5965.2023.0744](https://doi.org/10.13700/j.bh.1001-5965.2023.0744).
- [22] 王耀光. 基于 YOLOv8 改进模型的手机屏幕缺陷检测算法研究[D]. 上海: 华东师范大学, 2023.
- [23] 苏佳, 贾泽, 秦一畅, 等. 面向工业表面缺陷检测的改进 YOLOv8 算法[J/OL]. 计算机工程与应用, 2024(14): 187-196 [2024-04-24]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20240422.1827.012.html>.
- SU J, JIA Z, QIN Y C, et al. Improved YOLOv8 algorithm for industrial surface defect detection[J/OL]. *Computer Engineering and Applications*, 2024(14): 187-196 [2024-04-24]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20240422.1827.012.html>.
- [24] HEUSEL M, RAMSAUER H, UNTERTHINER T, et al. GANs trained by a two time-scale update rule converge to a local Nash equilibrium[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, CA, USA, 2017: 6629-6640.
- [25] RUSSELL B C, TORRALBA A, MURPHY K P, et al. LabelMe: a database and web-based tool for image annotation[J]. *International Journal of Computer Vision*, 2008, 77: 157-173.

(编辑: 张艳艳)