

基于改进 YOLOv5 的航空发动机表面缺陷检测模型

李鑫, 李香蓉*, 汪诚, 李秋良, 李卓越

空军工程大学基础部, 陕西 西安 710038

摘要 针对目前航空发动机表面人工缺陷检测效率低的问题, 提出一种基于改进 YOLOv5 的缺陷检测模型 YOLOv5-CE。首先, 在网络中融合数据增强策略搜索算法, 自动为当前数据集搜索最佳的数据增强策略, 实现训练效果的提升; 其次, 在 backbone 网络中引入坐标注意力机制, 在通道注意力的基础上嵌入坐标信息, 提高对小缺陷目标的检测能力; 最后, 将 YOLOv5 的定位损失函数改进为 efficient intersection over union 损失, 在加快模型收敛的同时提高预测框回归精度。实验结果表明, 所提 YOLOv5-CE 模型, 相比原 YOLOv5s 网络, 在检测速度几乎没有下降的情况下平均精度均值提高了 1.2 个百分点, 达到了 98.5%, 能够高效智能检测航空发动机 4 种常见类型缺陷。

关键词 机器视觉; 航空发动机; 表面缺陷检测; YOLOv5; 注意力机制

中图分类号 TP391.4 **文献标志码** A

DOI: 10.3788/LOP222557

Aero-Engine Surface Defect Detection Model Based on Improved YOLOv5

Li Xin, Li Xiangrong*, Wang Cheng, Li Qiuliang, Li Zhuoyue

Fundamentals Department, Air Force Engineering University, Xi'an 710038, Shaanxi, China

Abstract To improve the current low efficiency process in artificial defect detection on aero-engine surface, a YOLOv5-CE model, based on improved YOLOv5, is proposed. First, the data enhancement strategy search algorithm is integrated into the network to automatically search the best data enhancement strategy for the current dataset to improve the training effect. Second, the coordinate attention mechanism is introduced into the backbone network while the coordinate information is embedded on the basis of channel attention to improve detection of small defect targets. Finally, the location loss function of YOLOv5 is improved to efficient intersection over union loss which can accelerate the model convergence and improve the precision of prediction box regression. Experimental results show that compared with the original YOLOv5s network, the proposed YOLOv5-CE model improves the mean average precision by 1.2 percentage points to 98.5% and can efficiently, as well as intelligently, detect four common types of defects in aero-engines.

Key words machine vision; aero-engine; surface defect detection; YOLOv5; attention mechanism

1 引言

航空发动机作为飞机的心脏, 长时间工作在高温、高压、高负载的环境下, 极易产生各种缺陷和损伤^[1]。在飞行过程中, 航空发动机也非常容易受到外来物的冲击, 如飞鸟、砂石、冰雹等被吸入发动机会造成划痕、缺口和凹坑等机械损伤^[2]。这些损伤和缺陷会严重影响发动机的性能, 给飞行带来巨大安全隐患, 甚至可能造成严重事故。因此, 及时发现航空发动机上存在的缺陷或损伤, 对于保证飞机的飞行安全至关重要。

目前对于航空发动机缺陷的检测方法主要有孔探检测技术、涡流检测技术、超声波检测技术和轴承原位

检测技术等^[3], 这些技术手段对缺陷成像后还需专业技术人员通过肉眼判断缺陷类型位置等信息, 缺陷最终的定位和识别主要依赖技术人员的经验。使用人工方法进行检测不仅耗时耗力, 还容易受到疲劳、心理因素等影响, 发生漏检或者误检的情况^[4]。因此, 将图像智能识别技术应用于航空发动机缺陷检测领域, 提高检测效率, 对实现检测的智能化与自动化具有重要意义。

近年来, 得益于卷积神经网络的诞生, 基于深度学习的目标检测技术迅速发展, 并在航空航天、医疗、自动驾驶、农业等领域得到了广泛应用, 极大地方便了人们的生活^[5]。目标检测算法主要分为两类: 一类是两

收稿日期: 2022-09-15; 修回日期: 2022-10-30; 录用日期: 2022-11-23; 网络首发日期: 2023-01-04

通信作者: *lixiangrong0925@126.com

阶段(two-stage)算法,如 R-CNN^[6]、Fast R-CNN^[7]和 Faster R-CNN^[8]等,此类算法将目标检测分为两个阶段,首先进行候选区域的提取,再根据这些区域对目标进行类别和位置的回归,检测精度较高,但检测时间较长;另一类是一阶段(one-stage)算法,如 SSD^[9]和 YOLO 等,此类算法无需生成候选区域,直接输出目标的类别和位置信息,检测速度快,但在精度方面相比两阶段算法存在一定劣势。

目前许多学者已将目标检测算法应用于缺陷检测领域并加以改进,取得了不错的效果。向宽等^[10]针对传统算法对工业铝材表面缺陷识别率低和定位不准的问题,在 Faster R-CNN 算法的主干网络中加入 feature pyramid network(FPN)结构加强对小特征的提取能力,将原来的 region of interest (ROI) pooling 算法替换为更精细的 ROI align,并使用 K-means 算法对缺陷聚类,明显提升了对铝材表面缺陷的检测精度。来文豪等^[11]将多光谱成像技术与 YOLOv4 算法相结合,提出一种煤矸石智能分离方法,实现了对煤与煤矸石的准确快速识别。程松等^[12]为快速准确识别 X 射线焊缝内部小目标缺陷,提出一种基于 YOLOv5-Tiny 的轻量化焊缝缺陷识别方法,在网络中引入 squeeze-and-excitation(SE)注意力模块,将 C3 模块替换为轻量化结构,并去除了用于检测大物体的检测层,实现了模型的轻量化和检测精度的提升。孙迎春等^[13]为解决 YOLOv3 算法在检测道路交通灯时存在的漏检率高、召回率低等问题,使用 K-means 算法对数据进行了聚类分析,并对网络结构进行了精简,最后利用高斯分布特性评估边界框的准确性,实现了检测精度和速度的提升。以上研究都在实际应用中取得了较好的效果,但仍存在诸多问题:1)算法的检测速度较慢,无法满足实时检测的需求;2)航空发动机部件表面的缺陷尺寸一般较小,识别难度大;3)缺少公开的航空发动机部件表面缺陷数据集,算法训练较为困难。

因此,为实现对航空发动机表面缺陷的定位和识别,本文以基于深度学习的目标检测算法 YOLOv5 作为研究对象,结合数据增强策略搜索算法、注意力机制和损失函数改进,对缺陷区域进行信息采集、特征提取、目标定位与分类,减少人工检测方法导致的漏检、误检情况,以此提高航空发动机表面缺陷检测的效率和精度。

2 算法原理

2.1 YOLOv5 算法介绍

YOLO 是一种典型的一阶段目标检测算法,在提取输入图像特征之后直接对目标的类别和位置信息进行回归,具有推理速度快、检测精度高等优点。YOLOv1^[14]是一阶段目标检测算法的开山之作,该算法的检测速度非常快,可轻松对物体进行实时检测,在计算机视觉领域受到了广泛关注。在此基础上改进的

YOLOv2^[15]和 YOLOv3^[16]算法相继被提出,检测精度和速度不断提升。YOLOv4^[17]算法借鉴了 CSPNet^[18]的思想,将主干网络由 Darknet-53 改进为 CSPDarknet53,在特征融合部分增加了 spatial pyramid pooling (SPP)^[19]模块,引入 path aggregation network(PANet)^[20]结构,从而在网络中增加了自底向上传递定位特征的路径,避免了特征传递过程部分信息缺失的问题。YOLOv5 相比 YOLOv4 在保持高检测精度的同时,大幅提升了检测速度,可实现高达 140 frame/s 的快速检测,具有更强的灵活性和实用性。

YOLOv5 算法引入了自适应锚框计算功能,每次训练时自动根据数据集中的目标设定锚框尺寸。YOLOv5 共包括 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 等 4 种模型,它们的网络结构基本相同,卷积核数量和网络层数依次增加,检测精度不断提高,但训练消耗的时间也逐渐变长。

2.2 YOLOv5 网络结构

YOLOv5 算法主要由 3 部分组成:backbone、neck 和 head。backbone 是主干特征提取网络,主要包括 focus、C3 和 SPP 等结构。focus 模块主要进行切片操作,对图像的宽高进行压缩,加快特征提取速度。C3 主要用于图像特征提取,基于 CSPNet 的结构设计很好地解决了计算量大的问题。SPP 模块主要通过 5×5 、 9×9 和 13×13 等 3 个尺度的最大池化来增大网络感受野,解决多尺度目标的问题。neck 是特征融合网络,采用 FPN 与 PANet 相结合的结构,对主干网络提取的特征进行多尺度融合。head 是预测网络,包含 3 个不同尺度的检测头,输出置信度得分最高的预测类别和目标的位置信息。

YOLOv5 算法不断迭代更新,目前已经有多个版本,使用 6.0 版本的 YOLOv5s 模型作为研究对象。该版本相比最初的 YOLOv5 模型更加精简:使用 6×6 大小的卷积层替代 focus 结构,使得 GPU 在运算时更加高效;将 SPP 模块改进为 spatial pyramid pooling-fast (SPPF)并置于主干网络的末端,显著提高运算速度;减少主干网络中 C3 模块的数量并在最后一个 C3 模块中引入残差结构。YOLOv5s-6.0 的网络结构如图 1 所示。

YOLOv5 算法使用 Mosaic 数据增强,将 4 张图片通过裁剪、缩放和旋转等操作随机组合在一起,增加单张图片中的目标个数,有利于提高模型的泛化能力。沿用前代算法中的先验框预测机制,网络不需要直接对物体的尺寸和位置信息进行预测,只需计算先验框与真实框之间的偏移量,再根据偏移量调整先验框,即可得到预测框的大小和位置。算法中共预设 3 组先验框,每组包含小、中、大等 3 个尺度,分别为:(10, 13)、(30, 61)、(116, 90);(16, 30)、(62, 45)、(156, 198);(33, 23)、(59, 119)、(373, 326),用于检测不同尺寸和长宽比的目标。

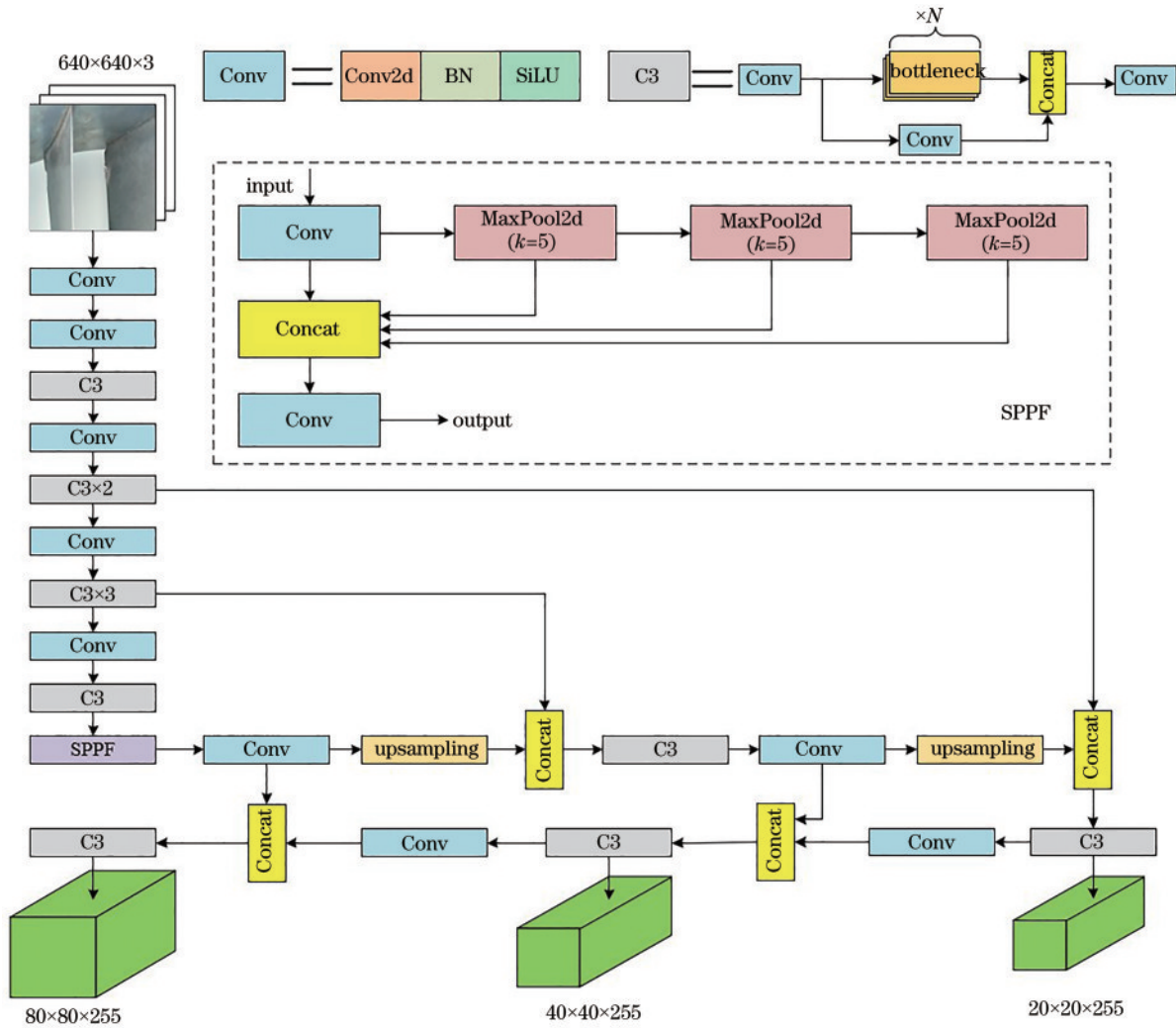


图1 YOLOv5s-6.0网络结构
Fig. 1 YOLOv5s-6.0 network structure

3 改进的YOLOv5算法

3.1 融合数据增强策略搜索算法

由于实际采集到的缺陷图像数量有限,若直接使用这些数据进行训练容易导致模型过拟合,因此需要利用数据增强技术来解决数据不足的问题,通过对现有数据进行扩充,产生充足、均衡的扩增样本来丰富数据集,从而满足深度学习算法对数据集在数量和性能等方面的要求,提升网络的泛化能力^[20]。因此在使用原始图像进行训练前,先通过数据增强技术对图像数量进行扩充。

在传统数据增强策略中,使用数据增强算法的类别和先后顺序都是人为设定的,这种依赖人工设计的数据增强算法往往不是最优的。为得到最合适的数据增强策略,采用基于神经网络架构搜索(NAS)技术的数据增强策略^[21],其核心思想是将多种数据增强策略集成到一个大的搜索空间,然后利用神经网络搜索出与当前数据集最匹配的数据增强策略,并找到它们的最佳使用顺序,使训练效果得到提升。

基于NAS的数据增强算法实现流程如图2所示。

该算法的原理是将许多数据增强子策略作为搜索空间,每个子策略由两部分组成:一部分是图像处理的函数(例如裁剪、旋转、缩放等);另一部分是使用每个函数的概率及参数的取值范围。为每个batch中的每张图像都随机选择其中一个子策略,然后采用基于强化学习的算法作为搜索算法,主要思路是使用循环神经网络(RNN)作为控制器(controller)产生子网络(child network),再对子网络进行训练和评估,得到其网络性能,最后根据网络性能不断更新控制器的参数,进而搜索出针对数据集最好的数据增强策略。搜索空间中包含的数据增强子策略如表1所示。

将数据增强策略搜索算法与YOLOv5s网络结合,在训练过程中,每次通过该算法在训练集上应用若干个数据增强子策略扩充图像数据,并进行训练,然后在验证集上对网络检测精度进行评估,并将该精度作为RNN的激励信号,生成新的若干个数据增强策略,再次对网络进行训练和验证,直到YOLOv5s网络在验证集

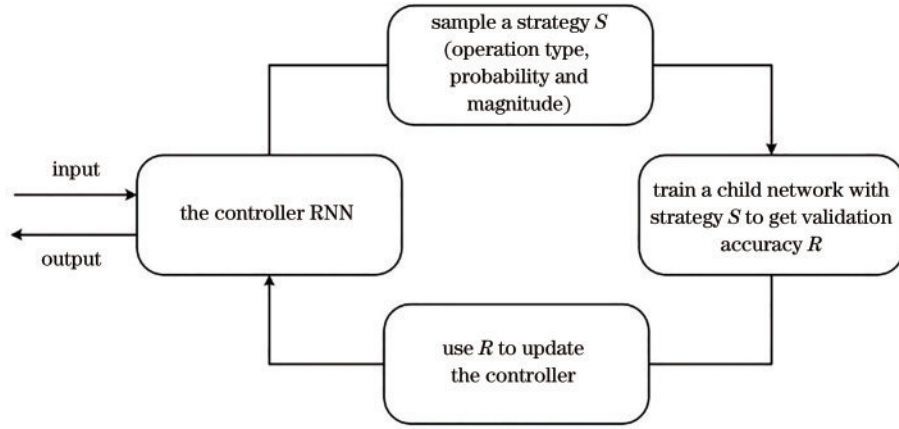


图 2 数据增强策略搜索算法流程图

Fig. 2 Flowchart of a search algorithm for data augmentation

表 1 搜索空间

Table 1 Search space

Data augmentation substrategy	Description	Range of $V_{\text{magnitude}}$
Contrast	Adjust the contrast of the image. $V_{\text{magnitude}}=0$ gives a gray image and $V_{\text{magnitude}}=1$ gives the original image	$[0, 2]$
Sharpness	Adjust the sharpness of the image. $V_{\text{magnitude}}=0$ gives a blurred image and $V_{\text{magnitude}}=1$ gives the original image	$[0, 2]$
Brightness	Adjust the brightness of the image. $V_{\text{magnitude}}=0$ gives a black image and $V_{\text{magnitude}}=1$ gives the original image	$[0, 2]$
Rotation	Rotate the image by $V_{\text{magnitude}}$ degrees	$[-90^\circ, 90^\circ]$
Scale	Enlarge or reduce the image to $V_{\text{magnitude}}$ scales	$[0.5, 2]$
Flip	Flip the image	Flip up-down/flip left-right
HSV augmentation	Adjust the H (hue), S (saturation), and V (value) of the image	$H: [0^\circ, 360^\circ]$, $S: [0, 1]$, $V: [0, 1]$
Noise	Add noise to the image	Gaussian noise and salt and pepper noise

上的检测精度最高,此时即为最佳的数据增强策略。

3.2 坐标注意力机制

裂纹、凹坑等小缺陷目标在图像中的像素所占比例小,经过卷积层不断提取特征后,信息容易丢失,导致算法对小目标的检测效果较差^[22]。为了提高算法在雨雪、沙尘天气或光照不足等复杂场景下对小缺陷目标的检测精度,训练时重点关注目标区域,有效提取缺陷特征信息,在YOLOv5网络中引入一种新型的坐标注意力(CA)机制^[23],针对通道注意力机制容易忽略位置信息的问题,在获取通道间信息的同时将与方向相关的位置信息嵌入通道注意力中,使网络能够在更大的范围内注意到小目标^[24]。CA模块的结构如图3所示。其中,residual表示残差结构,X Avg Pool和Y Avg Pool分别表示X和Y方向的平均池化。

CA注意力机制分为坐标信息嵌入和坐标注意力生成两个阶段。通道注意力通常采用全局池化的方式编码空间信息,导致位置信息丢失,因此CA注意力机制将全局池化分解为一对一维特征编码操作。在坐标

信息嵌入过程中,对于尺寸为 $C \times H \times W$ 的输入特征图 X ,首先使用两个尺寸为 $(H, 1)$ 和 $(1, W)$ 的池化核分别沿水平和垂直方向对每个通道进行编码,则第 c 个通道在高度为 h 和宽度为 w 的输出特征图分别 $z_c^h(h)$ 和 $z_c^w(w)$,如式(1)和(2)所示。利用两个一维全局池化操作分别将两个空间方向的输入特征聚合为一对方向感知特征图,便于网络更加准确地对感兴趣的物体进行定位。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i), \quad (1)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w). \quad (2)$$

在坐标注意力生成阶段,首先对在坐标信息嵌入阶段生成的两个特征图 z^h 和 z^w 进行级联操作,再通过一个共享 1×1 卷积操作进行 F_1 变换,如式(3)所示:

$$f = \delta \left[F_1 \left(\left[z^h, z^w \right] \right) \right], \quad (3)$$

式中:在水平和垂直方向上对空间信息进行编码的中

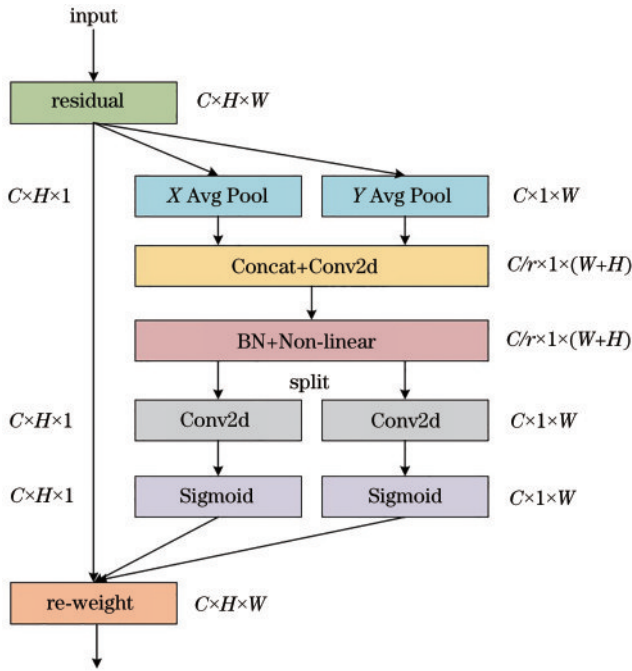


图 3 CA 模块结构

Fig. 3 Structure of the CA module

间特征映射 $f \in \mathbf{R}^{C/r \times (H+W)}$, r 为压缩比。沿空间维度将 f 分成两个单独的张量 $f^h \in \mathbf{R}^{C/r \times H}$ 和 $f^w \in \mathbf{R}^{C/r \times W}$, 然后通过两个 1×1 卷积 F_h 和 F_w 将 f^h 和 f^w 变换为与输入特征图 X 相同的通道数, 再经 Sigmoid 函数激活后可得到注意力权重 g_h 和 g_w :

$$g^h = \sigma[F_h(f^h)], \quad (4)$$

$$g^w = \sigma[F_w(f^w)]. \quad (5)$$

最后将注意力权重 g_h 和 g_w 与输入特征图 X 相乘, 即可得到坐标注意力模块输出的注意力加权特征图 Y :

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j). \quad (6)$$

3.3 定位损失函数改进

YOLOv5 的损失函数包括定位损失、置信度损失和分类损失。其中, 定位损失主要用于计算预测框与真实框之间的误差, 使用 complete intersection over union(CIoU) 损失^[25] 作为损失函数。本实验主要对定位损失函数进行改进, 解决原 CIoU 损失函数宽高比模糊的问题。

CIoU 损失的原理如图 4 和式(7)~(10)所示, 其中, A 为预测框, B 为真实框, C 为 A 、 B 两个框的最小外接矩形。

$$R_{IoU} = \frac{A \cap B}{A \cup B}, \quad (7)$$

$$L_{CIoU} = 1 - R_{IoU} + \frac{\rho^2(b^A, b^B)}{c^2} + \alpha v, \quad (8)$$

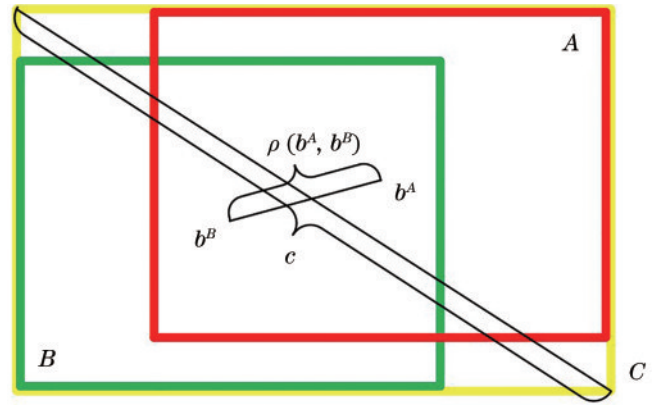


图 4 CIoU 示意图

Fig. 4 Diagram of CIoU

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^B}{h^B} - \arctan \frac{w^A}{h^A} \right)^2, \quad (9)$$

$$\alpha = \frac{v}{(1 - R_{IoU}) + v}, \quad (10)$$

式中: b^A 和 b^B 分别表示预测框和真实框的中心点; ρ 为两点之间的欧氏距离; c 表示预测框和真实框的最小外接矩形的对角线长度; α 为权重参数; v 用于衡量预测框和真实框长宽比的一致性; w^A 和 h^A 表示预测框的宽和高; w^B 和 h^B 表示真实框的宽和高。

CIoU 损失中的 v 反映的是预测框与真实框宽高比的差异, 但当两个框的宽高比成线性比例时, 该惩罚项就会失效, 从而无法继续进行回归优化。因此引入 efficient intersection over union(EIoU)^[26] 将 CIoU 损失中宽高比的影响因子拆开, 分别计算预测框和真实框的宽高。EIoU 损失在 CIoU 损失的重叠面积和中心点距离两项损失的基础上将对宽高比的惩罚修正为对宽和高分别进行惩罚, 解决了宽高比模糊的问题, 并且提高了损失函数的收敛速度和预测框回归精度。EIoU 损失的原理如式(11)所示。其中, c^w 和 c^h 表示预测框和真实框的最小外接矩形的宽和高。

$$L_{EIoU} = 1 - R_{IoU} + \frac{\rho^2(b^A, b^B)}{c^2} + \frac{\rho^2(w^A, w^B)}{c_w^2} + \frac{\rho^2(h^A, h^B)}{c_h^2}. \quad (11)$$

3.4 改进后的网络结构

训练过程中使用数据增强策略搜索算法寻找最适合的数据增强方法, 提升网络训练效果。在主干网络的末端引入 CA 注意力机制, 使网络在更大的范围内注意到小缺陷和遮挡、重叠目标, 提升检测精度。将定位损失函数由 CIoU 损失改进为 EIoU 损失, 直接对宽和高的预测结果进行惩罚, 从而提高边界框的回归精度。在 YOLOv5s 的主干网络中添加 CA 注意力模块的示意图如图 5 所示, 虚线框内即为 CA 模块所在位置。

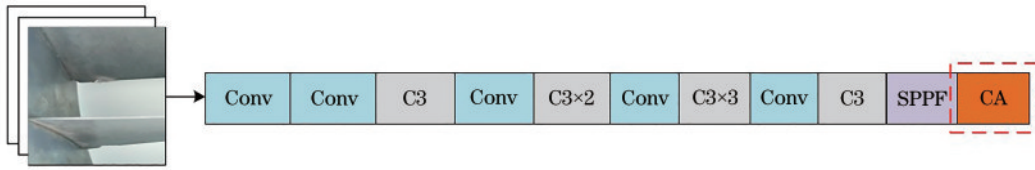


图5 CA模块添加位置示意图

Fig. 5 Diagram of the CA module adding location

4 实验方案与结果

4.1 实验环境

实验环境基于 Windows 10 操作系统, CPU 型号为 Intel (R) Xeon (R) Gold 5218 CPU@2.30 GHz, GPU 型号为 NVIDIA GeForce RTX 2080Ti, 显存为 11 GB, 运行内存为 128 GB。深度学习框架基于 PyTorch 1.9.1 搭建, Python 版本为 3.7, 用于加速 GPU 计算的 CUDA 版本为 10.0, cuDNN 版本为 7.4.1。

4.2 实验参数设定

在模型的训练过程中进行 Mosaic 数据增强, 使用 SGD 优化器对网络参数进行迭代更新, 学习率衰减策略为余弦退火。输入图像尺寸为 $640 \times 640 \times 3$, batch

size 设置为 16, 即每次输入 16 张图像, 共训练 300 个 epoch。超参数设置如下: 初始学习率为 0.001, 循环学习率为 0.2, 动量为 0.937, 权重衰减系数为 0.0005。

4.3 数据集

所使用的缺陷图像主要是在合作单位和发动机修理厂等地自行采集和网络数据爬取等途径获取的。使用的采集设备包括工业相机和孔探仪等, 在不同光照、角度以及不同背景下对缺陷进行拍摄。大部分缺陷图像通过工业相机对单个部件或直接在机体上进行拍摄, 部分缺陷由于部件在发动机内部, 因而使用孔探仪通过发动机上的孔探口进行拍摄。通过对采集到的图像进行整理、筛选、分类, 最终得到 4 种类型的缺陷: 裂纹 (crack)、缺口 (gap)、凹坑 (pit) 和划痕 (scratch), 如图 6 所示。

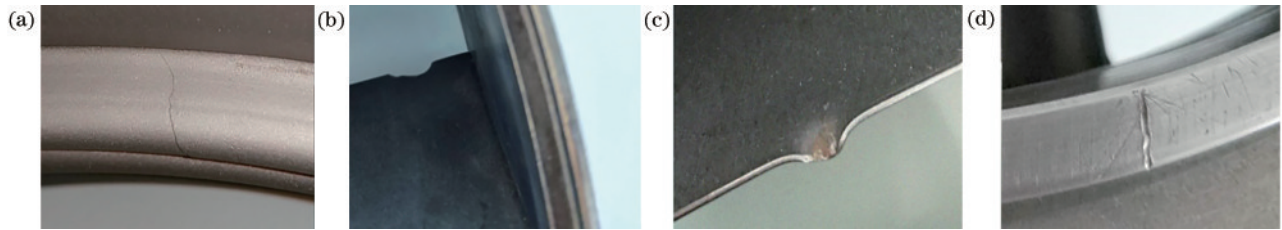


图6 缺陷类型。(a)裂纹;(b)缺口;(c)凹坑;(d)划痕

Fig. 6 Type of defect. (a) Crack; (b) gap; (c) pit; (d) scratch

共采集到 1080 张不同分辨率和不同拍摄环境下的原始图像, 其中, 裂纹、缺口、凹坑、划痕的航空发动机部件图像数量各为 543 张、192 张、135 张和 210 张。由于原始采集的图像数量有限, 若直接采用原始图像对模型进行训练, 容易出现过拟合的情况, 因此通过随机缩放、翻转、旋转和裁剪、色域变换和添加噪声等方式对图像数量进行扩充, 最终得到原始图像和增强后的图像共 3500 张, 其中, 裂纹、缺口、凹坑、划痕的图像数量分别为 1740 张、653 张、445 张和 662 张。

在使用图像对算法进行训练之前, 需要对图像中的目标进行标注, 即为每张图像上的缺陷注明类别和位置, 方便算法学习。使用 LabelImg 软件对图像中的缺陷进行标注, 并生成目标的标签信息。LabelImg 是一种可视化的图像标注工具, 界面如图 7 所示。

每标注完一张图像后, 软件会自动生成一个包含图像路径、标注区域及标签类别等信息的 xml 文件, 具体内容如图 8 所示。其中: path 项表示图像的存储路径; size 项表示原始图像的宽高尺寸和通道数; object

项表示网络在训练过程中重点关注的区域; name 项为标注框的类别; bndbox 项记录了标注框左上角和右下角的横纵坐标。

将经过数据增强后得到的 3500 张图像及对应的标注文件进行整理和分类, 制成航空发动机部件表面缺陷数据集。将数据集按照 6:2:2 的比例划分为训练集、验证集和测试集。

4.4 评价指标

在目标检测领域, 常用的评价指标主要有精确率 ($R_{\text{precision}}$)、召回率 (R_{recall})、平均精度 (P_{AP}) 和平均精度均值 (P_{mAP})。 $R_{\text{precision}}$ 是模型预测正确的正样本与其检测出的所有正样本数量之比, R_{recall} 是模型预测正确的正样本与实际所有正样本的数量之比, 其计算公式分别为

$$R_{\text{precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}, \quad (12)$$

$$R_{\text{recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}, \quad (13)$$



图 7 缺陷标注示例

Fig. 7 Defect labelling example

```
<annotation>
<folder>JPEGImages</folder>
<filename>gap211.jpg</filename>
<path>D:\VOC2007\JPEGImages\gap211.jpg</path>
<source>
<database>Unknown</database>
</source>
<size>
<width>3455</width>
<height>2445</height>
<depth>3</depth>
</size>
<segmented>0</segmented>
<object>
<name>gap</name>
<pose>Unspecified</pose>
<truncated>0</truncated>
<difficult>0</difficult>
<bndbox>
<xmin>2041</xmin>
<ymin>1169</ymin>
<xmax>2192</xmax>
<ymax>1406</ymax>
</bndbox>
</object>
</annotation>
```

图 8 xml 标签文件

Fig. 8 xml label file

式中: N_{TP} 表示模型预测为正样本且实际也为正样本的数量; N_{FP} 表示模型预测为正样本但实际为负样本的数量; N_{FN} 表示模型预测为负样本但实际为正样本的数量。

以某一类别的 $R_{precision}$ 为纵轴, R_{recall} 为横轴描绘成的曲线称为 $P-R$ 曲线, 该曲线与坐标轴围成的面积表示此类别缺陷的 P_{AP} , 对所有类别缺陷的 P_{AP} 求平均值即可得到 P_{mAP} , 其计算公式分别为

$$P_{AP} = \sum_{i=1}^N R_{precision}(i) \Delta R_{recall}(i) = \int_0^1 R_{precision}(R_{recall}) dR_{recall}, \quad (14)$$

$$P_{mAP} = \frac{\sum_{i=1}^N P_{APi}}{N}, \quad (15)$$

式中: N 表示总类别数。

4.5 不同目标检测算法对比

为了验证所改进的 YOLOv5-CE 算法的综合性能, 选择几种主流的目标检测算法, 包括两阶段算法 Faster R-CNN, 一阶段算法 YOLOv3、YOLOv4、YOLOv5s 和 YOLOXs, 在相同的数据集、实验参数和训练策略下, 对以上目标检测算法进行训练和测试, 得到每种算法检测各类别缺陷的 P_{AP} 、 P_{mAP} 、检测速度和模型容量, 具体如表 2 所示。

由实验结果可知, 所改进的 YOLOv5-CE 模型相比几种主流目标检测算法, 在检测精度、速度和模型容量等方面都具有明显优势。相比两阶段算法 Faster R-CNN, 不仅在 P_{mAP} 上领先 24.7 个百分点, 而且检测速度也有大幅提升; 相比一阶段算法 YOLOv3 和 YOLOv4, 不仅在 P_{mAP} 上分别领先 15.2 个百分点和 11.3 个百分点, 并且在检测速度上也具有明显优势, 特别是模型容量远小于这 3 种目标检测算法, 使得 YOLOv5-CE 模型在移动端的部署和低成本的工业应用上更加方便。相比改进前的 YOLOv5s 网络, YOLOv5-CE 在几乎不增加模型容量的情况下实现了 1.2 个百分点的 P_{mAP} 的提升, 并且仍能保持较高的检测速度, 满足实时检测的任务需求。YOLOv5-CE 虽然在检测速度方面低于 YOLOXs 算法, 但 P_{mAP} 相比后

表 2 各算法检测性能对比

Table 2 Comparison of the detection performance of algorithms

Model	$P_{AP} / \%$				$P_{mAP} / \%$	Speed / (frame/s)	Capacity / MB
	crack	gap	pit	scratch			
Faster R-CNN	75.4	78.1	58.7	83.1	73.8	14.29	109
YOLOv3	82.1	89.8	85.1	76.1	83.3	22.63	235
YOLOv4	90.0	93.2	86.6	79.0	87.2	24.72	244
YOLOv5s	94.9	99.5	98.4	96.5	97.3	41.32	14.4
YOLOXs	90.8	90.7	90.9	90.8	90.8	110.62	68.7
YOLOv5-CE	97.3	99.4	99.2	98.2	98.5	40.65	14.5

者高 7.7 个百分点,并且在模型容量方面具有明显优势,约为 YOLOXs 的 21.1%。

4.6 消融实验

为了检验数据增强策略搜索算法、CA 注意力机制以及 EIou 损失函数的有效性,设计了消融实验来验证每部分改进的效果,结果如表 3 所示。其中:YOLOv5_A 表示在原 YOLOv5s 网络中融合数据增强

搜索算法;YOLOv5-C 表示在原 YOLOv5s 网络中添加 CA 注意力机制;YOLOv5-E 表示将原 YOLOv5s 网络中的定位损失函数由 CIoU 损失改进为 EIou 损失;YOLOv5-CE 表示所提模型,在 YOLOv5s 网络的基础上同时使用了数据增强搜索算法、CA 注意力机制和定位损失函数等改进。

表 3 消融实验

Table 3 Ablation experiments

Model	$R_{precision} / \%$	$R_{recall} / \%$	$P_{AP} / \%$				$P_{mAP} / \%$	Speed / (frame/s)
			crack	gap	pit	scratch		
YOLOv5s	97.7	94.7	94.9	99.5	98.4	96.5	97.3	41.32
YOLOv5_A	96.8	95.7	95.2	99.5	98.9	97.8	97.9	40.49
YOLOv5-C	97.8	95.7	96.9	99.5	99.3	96.9	98.2	41.15
YOLOv5-E	97.3	96.6	96.4	99.5	99.3	98.2	98.4	46.73
YOLOv5-CE	98.1	95.4	97.3	99.4	99.2	98.2	98.5	40.65

由表 3 可知:YOLOv5_A 模型相比原网络的精确率和检测速度出现下降,但召回率有所提升,并且 P_{mAP} 提高了 0.6 个百分点,这表明数据增强策略对于提高算法的检测精度具有一定作用,但会造成检测速度的损失;YOLOv5-C 模型相比原网络的精确率、召回率均有提升, P_{mAP} 增加了 0.9 个百分点,而检测速度几乎没有区别,并且对于裂纹(crack)和凹坑(pit)这两类小缺陷的检测精度提升尤为明显,分别为 2.0 个百分点和 0.9 个百分点,这验证了 CA 注意力机制能够在更大的范围内关注小目标,提高对小缺陷的检测精度,并且不影响检测速度;YOLOv5-E 模型相比原网络的精确率略微下降,召回率提升了 1.9 个百分点, P_{mAP} 值提高了 1.1 个百分点,尤其是对划痕(scratch)类缺陷的检测精度提升了 1.7 个百分点,检测速度也有了大幅提升,约为 13.1%,这表明 EIou 损失提高了算法的收敛速度,并且带来了回归精度的提升。

所提 YOLOv5-CE 模型,相比原 YOLOv5s 网络,精确率和召回率均有所提升,分别为 0.4 个百分点和 0.7 个百分点,在检测速度几乎没有下降的情况下 P_{mAP} 值提高了 1.2 个百分点,特别是对 crack 和 scratch 类缺

陷的检测精度有了明显提升,分别为 2.4 个百分点和 1.7 个百分点。

4.7 检测效果对比

为了验证 YOLOv5-CE 模型实际检测效果的提升,分别使用原始 YOLOv5s 模型和 YOLOv5-CE 模型对实际航空发动机部件表面的缺陷进行了检测,部分检测效果对比如图 9 所示。对比图 9(a)、(d)可以看出,YOLOv5-CE 模型对于 gap 类缺陷的检测的置信度有所提高;对比图 9(b)、(e)可以发现,YOLOv5-CE 模型提高了对于 pit 和 scratch 类缺陷的检测置信度;对比图 9(c)、(f)可以看出,原 YOLOv5s 模型对于 pit 类缺陷存在漏检和误检的情况,而改进后的 YOLOv5-CE 模型很好地解决了这一问题,同时对该类别缺陷的检测置信度也普遍高于原 YOLOv5s 模型。

5 结 论

为实现航空发动机表面缺陷的智能化与高精度检测,提出一种基于 YOLOv5s 算法改进的 YOLOv5-CE 模型。不同于传统数据增强策略的随机选取,YOLOv5-CE 融合数据增强搜索算法自动搜索与数据

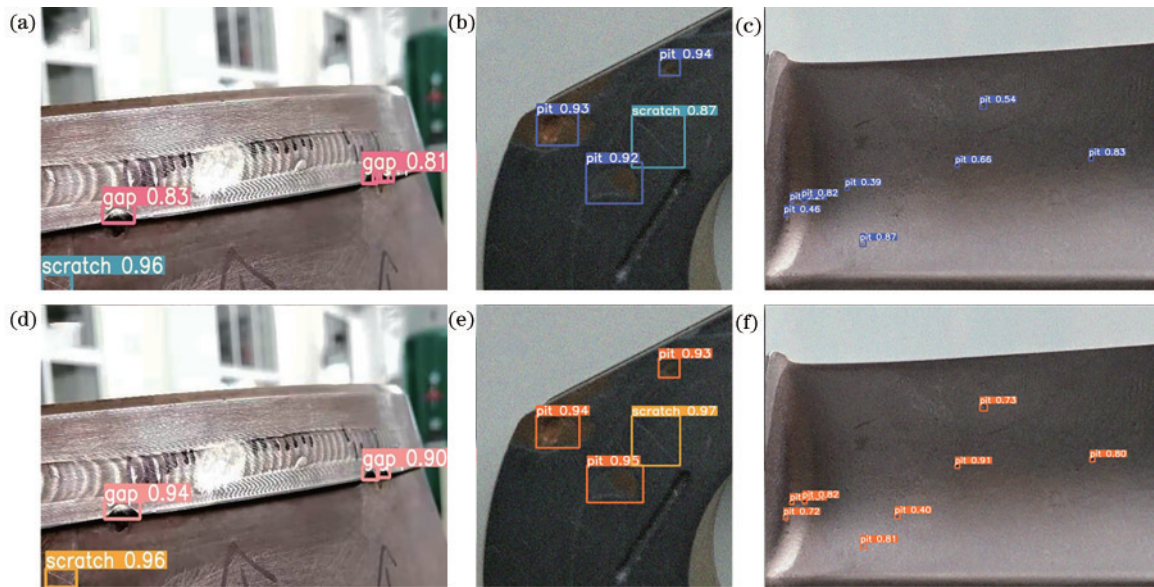


图9 两种模型检测效果对比。(a)(b)(c) YOLOv5s 检测效果;(d)(e)(f) YOLOv5-CE 检测效果
Fig. 9 Comparison of the detection effect of two models. (a) (b) (c) Detection effect of YOLOv5s;
(d) (e) (f) Detection effect of YOLOv5-CE

集最匹配的数据增强策略,实现数据增强策略选择的最优化。其次,为进一步提高小缺陷的检测效果,在网络中引入坐标注意力机制,在通道注意力的基础上嵌入坐标位置信息,使网络在更大的范围内注意到小目标。最后,将定位损失函数由 CIoU 损失改为 EIoU 损失,加快模型收敛速度并提高预测框的回归精度。实验结果表明,在自制的航空发动机缺陷数据集上,YOLOv5-CE 模型相比原 YOLOv5s 网络,精确率和召回率均有所提升,分别为 0.4 个百分点和 0.7 个百分点,在保持原有检测速度的情况下 P_{mAP} 提高了 1.2 个百分点,并且几乎不增加模型容量,相比主流的目标检测算法在检测精度、检测速度和模型容量方面均有明显优势。后续工作将进一步研究如何提升模型的检测速度以及在复杂场景下的检测性能,并将模型在移动端进行部署以实现表面缺陷检测的便捷化。

参 考 文 献

- [1] 孟娇茹. 航空发动机孔探损伤识别方法[J]. 黑龙江科技学院学报, 2009, 19(1): 50-53.
Meng J R. Aero-engine interior damage recognition based on support vector machine[J]. Journal of Heilongjiang Institute of Science and Technology, 2009, 19(1): 50-53.
- [2] 马超, 王玉娜, 武耀罡, 等. 航空发动机风扇叶片硬物冲击损伤特征[J]. 航空动力学报, 2017, 32(5): 1105-1111.
Ma C, Wang Y N, Wu Y G, et al. Hard object impact damage characteristics of aero engine fan blade[J]. Journal of Aerospace Power, 2017, 32(5): 1105-1111.
- [3] 郑丽, 罗泽明, 张波. 无损检测在航空发动机维修中的应用和发展[J]. 现代制造技术与装备, 2014(4): 62-63.
Zheng L, Luo Z M, Zhang B. The application and development of non-destructive testing in aircraft engine maintenance[J]. Modern Manufacturing Technology and Equipment, 2014(4): 62-63.
- [4] 何嘉辉, 张栋善, 赵成, 等. 航空发动机叶片裂纹检测技术及应用分析[J]. 内燃机与配件, 2020(15): 151-152.
He J H, Zhang D S, Zhao C, et al. Detection technology and application analysis of aero-engine blade crack[J]. Internal Combustion Engine & Parts, 2020(15): 151-152.
- [5] 邱天衡, 王玲, 王鹏, 等. 基于改进 YOLOv5 的目标检测算法研究[J]. 计算机工程与应用, 2022, 58(13): 63-73.
Qiu T H, Wang L, Wang P, et al. Research on object detection algorithm based on improved YOLOv5[J]. Computer Engineering and Applications, 2022, 58(13): 63-73.
- [6] 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, OH, USA. New York: IEEE Press, 2014: 580-587.
- [7] Girshick R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1440-1448.
- [8] Ren S Q, He K M, 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.
- [9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [10] 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 product surface defect detection method based on improved Faster RCNN[J]. Chinese Journal of Scientific Instrument, 2021, 42(1): 191-198.
- [11] 来文豪, 周孟然, 胡锋, 等. 基于多光谱成像和改进 YOLO v4 的煤矸石检测[J]. 光学学报, 2020, 40(24): 2411001.
- Lai W H, Zhou M R, Hu F, et al. Coal gangue detection based on multi-spectral imaging and improved YOLO v4 [J]. Acta Optica Sinica, 2020, 40(24): 2411001.
- [12] 程松, 杨洪刚, 徐学谦, 等. 基于 YOLOv5 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): 2104005.
- Cheng S, Yang H G, Xu X Q, et al. Improved lightweight X-ray aluminum alloy weld defect detection algorithm based on YOLOv5[J]. Chinese Journal of Lasers, 2022, 49(21): 2104005.
- [13] 孙迎春, 潘树国, 赵涛, 等. 基于优化 YOLOv3 算法的交通灯检测[J]. 光学学报, 2020, 40(12): 1215001.
- Sun Y C, Pan S G, Zhao T, et al. Traffic light detection based on optimized YOLOv3 algorithm[J]. Acta Optica Sinica, 2020, 40(12): 1215001.
- [14] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [15] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, July 21-26, 2017, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [16] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2022-08-05]. <https://arxiv.org/abs/1804.02767>.
- [17] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2022-08-09]. <https://arxiv.org/abs/2004.10934>.
- [18] Wang C Y, Mark Liao H Y, 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 (CVPRW), June 14-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1571-1580.
- [19] He K M, Zhang X Y, Ren S Q, 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.
- [20] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 8759-8768.
- [21] 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.
- Zhu X H, Qian L P, Fu W. Overview of research on image data enhancement technology[J]. Software Guide, 2021, 20(5): 230-236.
- [22] 王玲敏, 段军, 辛立伟. 引入注意力机制的 YOLOv5 安全帽佩戴检测方法[J]. 计算机工程与应用, 2022, 58(9): 303-312.
- Wang L M, Duan J, Xin L W. YOLOv5 helmet wear detection method with introduction of attention mechanism[J]. Computer Engineering and Applications, 2022, 58(9): 303-312.
- [23] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 13708-13717.
- [24] 宋甜, 李颖, 王静. 改进 YOLOv5s 的车载红外图像目标检测[J]. 现代计算机, 2022, 28(2): 21-28.
- Song T, Li Y, Wang J. Improved YOLOv5s vehicle infrared image target detection[J]. Modern Computer, 2022, 28(2): 21-28.
- [25] Zheng Z H, Wang P, Ren D W, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2022, 52(8): 8574-8586.
- [26] Zhang Y F, Ren W, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.