

基于改进 YOLOv8n 的白酒包装缺陷检测

贺刚健 曾水玲 林方聪 向 硕 张嘉雄

(吉首大学通信与电子工程学院, 湖南 吉首 416000)

摘要: [目的] 提高白酒包装缺陷检测准确率。[方法] 提出了一种基于改进 YOLOv8n 的检测模型。在模型中引入 ADown 模块, 有效减少了模型的参数量和计算量, 同时保持了原有的特征提取能力; 通过将大型可分离核注意力 (LSKA) 融入快速空间金字塔池化 (SPPF) 结构, 进一步提升了模型对多尺度特征的捕捉和提取能力; 将原有的 CIOU 损失函数替换为 Inner-WIoU 损失函数, 将 Inner-IoU 与 Wise-IoU 相结合, 在提升检测精度的同时加快模型收敛速度。[结果] 在自制的白酒包装数据集上, 改进后的 YOLOv8n 模型平均精度达 86.4%, 较原始模型提升了 5.2%, 参数量减少了 4.9%, 计算量降低了 6.6%。[结论] 该模型能够满足对白酒包装的实时检测要求。

关键词: 缺陷检测; YOLOv8; Inner-WIoU; ADown; 大型可分离核注意力

Defect detection of liquor packaging based on improved YOLOv8n

HE Gangjian ZENG Shuiling LIN Fangcong XIANG Shuo ZHANG Jiaxiong

(School of Communication and Electronic Engineering, Jishou University, Jishou, Hunan 416000, China)

Abstract: [Objective] To improve the accuracy of liquor packaging defect detection. [Methods] A detection model based on an improved YOLOv8n is proposed. The ADown module is introduced into the model to effectively reduce parameter count and computational load while maintaining the original feature extraction capability. Large separable kernel attention (LSKA) is integrated into the spatial pyramid pooling fusion (SPPF) structure to further enhance the model's ability to capture and extract multi-scale features. In addition, the original CIOU loss function is replaced with the Inner-WIoU loss function, which combines Inner-IoU and Wise-IoU, thereby improving detection accuracy and accelerating model convergence. [Results] On a self-built liquor packaging dataset, the improved YOLOv8n model achieves an average precision of 86.4%, representing a 5.2% improvement over the original model. Moreover, the parameter count is reduced by 4.9%, and computation is reduced by 6.6%. [Conclusion] The model can meet the real-time detection requirements of liquor packaging.

Keywords: defect detection; YOLOv8; Inner-WIoU; ADown; large separable kernel attention

在白酒包装的生产过程中, 受环境、设备等多方面因素的影响, 包装往往容易出现瑕疵, 例如印刷不清晰、划痕、破损等问题。传统的白酒包装缺陷检测方法主要依赖人工检测, 虽然人工检测可以灵活应对各种复杂的情况, 但存在效率低、成本高、主观性强等问题。

随着计算机视觉技术的快速发展, 基于深度学习的目标检测算法在包装缺陷检测领域逐渐成为主流。目前, 这种方法主要分为两类: 一类是以单次多框检测器

(SSD)^[1]、YOLO 系列^[2-4]为代表的单阶段目标检测算法, 另一类是以 R-CNN^[5]、Faster R-CNN^[6]为代表的双阶段目标检测算法。在包装缺陷检测领域, 夏军勇等^[7]提出了一种基于改进 Faster R-CNN 的食品包装缺陷检测算法, 将特征提取网络替换为 ResNet50, 并融合特征金字塔网络 (FPN), 提高模型多尺度检测的能力。李志诚等^[8]提出了一种改进的 YOLOv3 算法, 将不同尺度的特征图进行多尺度特征融合, 有效提升了网络的检测能力。李建明等^[9]

基金项目: 国家自然科学基金项目 (编号: 61966014); 湖南省自然科学基金项目 (编号: 2024JJ7413); 湖南省研究生科研创新项目 (编号: LXBZZ2024310)

通信作者: 曾水玲 (1975—), 女, 吉首大学教授, 博士。E-mail: zengflsl@163.com

收稿日期: 2024-10-29 **改回日期:** 2025-06-16

引用格式: 贺刚健, 曾水玲, 林方聪, 等. 基于改进 YOLOv8n 的白酒包装缺陷检测[J]. 食品与机械, 2025, 41(11): 215-224.

Citation: HE Gangjian, ZENG Shuiling, LIN Fangcong, et al. Defect detection of liquor packaging based on improved YOLOv8n[J]. Food & Machinery, 2025, 41(11): 215-224.

提出了一种结合 Inception-V3 图像分类算法和 YOLOv3 目标检测算法的缺陷检测方法,将这两种算法结合,可以有效解决复杂背景下卷纸包装缺陷的检测问题。叶宇星等^[10]提出了一种基于 YOLOv5s 改进的方法,提高了腌制蔬菜真空包装缺陷检测的精度。曾秀云等^[11]通过改进 Faster R-CNN 中的区域候选网络(RPN)和感兴趣区域(ROI)结构,对棉布包装表面缺陷检测取得了较好的效果。Lin 等^[12]提出了一种改进的 YOLOv5 算法,该算法采用双向特征金字塔网络(BiFPN)替代原有 FPN 并新增小目标检测层,增强不同层间语义信息的融合,提高了对织物缺陷检测的效果。

尽管上述算法从模型结构、特征层融合等不同角度得到改进,但仍存在固有问题,即仅考虑提升精度,却难以避免计算开销的增加。针对这一问题,研究拟提出一种基于改进 YOLOv8n 的白酒包装缺陷检测模型,以期为白酒包装生产线提供高效、精准的质量检测方法。

1 改进的 YOLOv8n 模型

1.1 YOLOv8 整体框架

YOLOv8 是 Ultralytics 最新发布的一种先进的目标检测模型,它在之前 YOLO 版本上引入了新功能并进行了改进,以进一步提高性能。YOLOv8 一共有 5 个模型,分别是 YOLOv8-n、s、m、l 和 x。这些模型在参数量和计算复杂度方面存在差异,且随着模型深度和宽度的增加而

变得更加复杂。

首先,在骨干网络(backbone)部分,YOLOv8 采用了 C2f 结构,相比于 YOLOv5 中的 CSP 结构,增加了更多的 Split 操作和跳层连接。并采用了与 YOLOv5 相同的 SPPF 操作来融合多尺度特征。其次,在检测头部分变化较大,变为了当前主流的解耦头,同时从基于锚框的方法(anchor-based)变为了无锚框的方法(anchor-free)。此外,YOLOv8 还改变了框匹配策略,将静态匹配改为了动态匹配(task-aligned assigner)方式,这有助于更准确地匹配目标框。最后,在训练策略上,YOLOv8 在最后 10 个训练周期关闭了 Mosaic 操作,这有助于模型更好地学习和收敛。

在白酒包装缺陷检测中,目标往往是细小且密集的,YOLOv8 的网络结构无法达到较高的检测要求,为解决这一问题,以 YOLOv8n 作为基础模型,从特征提取、多尺度特征融合、损失函数等方面进行优化。采用 ADown 模块替换网络中的卷积模块,降低模型复杂度。通过将大型可分离核注意力(LSKA)集成到快速空间金字塔池化(SPPF)模块中,增强各特征层之间的语义信息融合。基于 Inner-IoU 损失函数的理念,对 WIoU 进行了改进,采用 Inner-WIoU 作为模型的边框回归损失函数,以降低低质量样本的影响并加速收敛。改进后的模型网络结构如图 1 所示。

1.2 轻量化下采样 ADown 模块

YOLOv8n 模型采用卷积操作进行下采样,这种方法

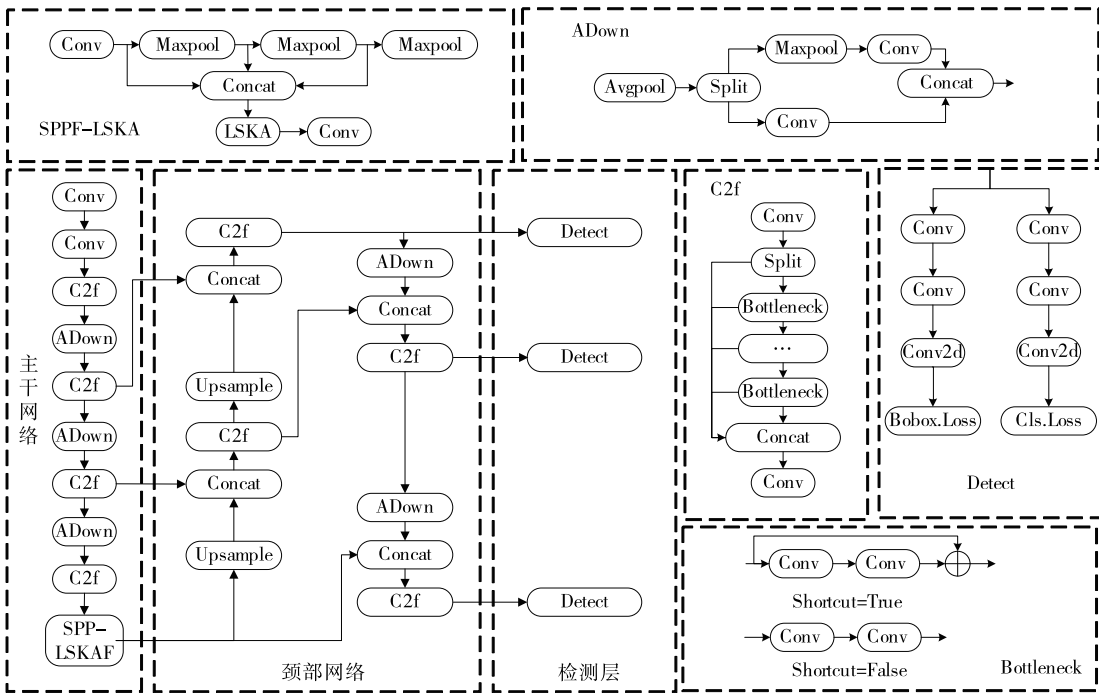


图 1 改进 YOLOv8n 网络结构图

Figure 1 Structure diagram of improved YOLOv8n network

在确保模型简洁和高效的同时,能够快速减小特征图的尺寸。然而,这也带来了一些问题。首先,卷积在下采样过程中可能导致关键信息的丢失,特别是细节部分,这对缺陷检测任务尤为重要。其次,卷积操作往往伴随着较大的参数量,增加了模型的复杂度和计算成本。为了解决这些问题,引入了YOLOv9^[13]模型中的ADown模块,替换了YOLOv8n中的卷积层。ADown模块通过将输入特征图分为两部分进行处理:一部分用于高效提取特征,另一部分用于保留更多关键细节信息。这样不仅减少了计算量,还能提取更丰富的特征,从而提升了模型的整体性能。相比传统卷积操作,ADown模块显著优化了计算效率,同时提高了检测精度。卷积模块与ADown模块结构如图2所示。

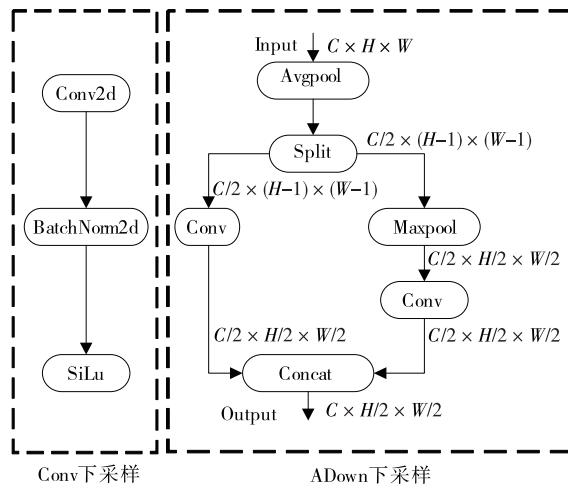


图2 卷积模块与ADown模块结构图

Figure 2 Structure diagrams of the convolutional module and ADown module

首先,输入特征图经过 2×2 的核和步长为1的平均池化操作进行初步下采样,将特征图的高度和宽度分别减小了1,接着沿着通道维度拆分为两个特征图 x_1 和 x_2 。其中, x_1 部分通过一个 3×3 卷积层进行处理,将 x_1 部分高宽减半,实现高效的特征提取和进一步的下采样,而 x_2 部分则先经过最大池化操作,将 x_2 部分高宽减半,再通过 1×1 卷积层,主要目的是保留更多的原始信息,同时压缩通道数。ADown模块采用切分特征图并分开处理的策略,能够更好地获取不同尺度的特征,在提高模型特征提取能力的同时也降低了模型的复杂度。

1.3 SPPF-LSKA模块的设计

SPPF模块通过3个 5×5 大小的最大池化操作来捕获多尺度特征,从而提高模型对不同尺度目标的检测能力。然而,在白酒包装缺陷检测任务中,许多如划痕、破损等细小缺陷,这些小目标在最大池化过程中容易被忽

略或压缩,导致模型对细小缺陷的检测效果下降。为了解决这一问题,在SPPF模块中引入了LSKA^[14]模块,以增强模型对小目标缺陷的特征提取能力。SPPF模块、改进后的SPPF-LSKA模块和LSKA模块如图3所示。

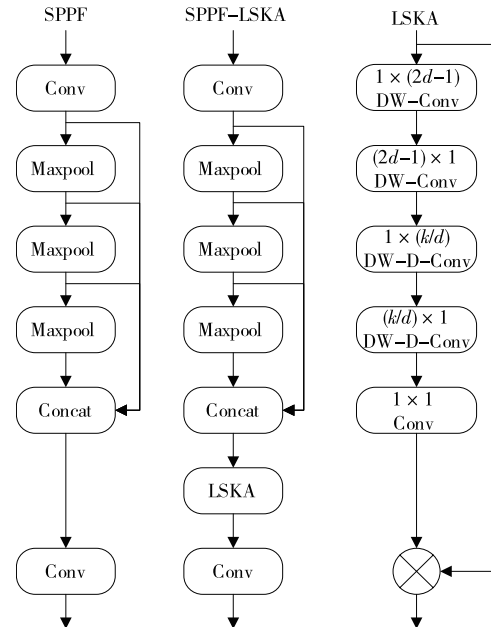


图3 SPPF、SPPF-LSKA和LSKA模块结构图

Figure 3 Structure diagram of SPPF, SPPF-LSKA, and LSKA modules

LSKA是在大核注意力机制(LKA)的基础上改进的,LSKA首先将深度卷积和深度扩展卷积的二维卷积核分解为两个级联的1D卷积核,再将拆分后的卷积进行串联。这样大大减少了所需参数数量和计算量,同时这些卷积能够分别在水平和垂直方向上,以及不同尺度和范围上提取特征,扩大了模型的感受野,从而增强了网络的特征表达能力。LSKA输出:

$$\bar{Z}^C = \sum_{h,w} W_{(2d-1) \times 1}^C * (\sum_{h,w} W_{1 \times (2d-1)}^C * F^C), \quad (1)$$

$$Z^C = \sum_{h,w} W_{\lfloor \frac{k}{d} \rfloor \times 1}^C * (\sum_{h,w} W_{1 \times \lfloor \frac{k}{d} \rfloor}^C * \bar{Z}^C), \quad (2)$$

$$A^C = W_{1 \times 1} * Z^C, \quad (3)$$

$$\bar{F}^C = A^C \otimes F^C, \quad (4)$$

式中:

$*$ 、 \otimes ——卷积、哈达玛乘积;

k, d ——卷积核大小、扩展率;

F^C ——特征图;

\bar{Z}^C, Z^C ——深度卷积的输出、深度扩展卷积的输出;

$W_{(2d-1) \times 1}^C, W_{1 \times (2d-1)}^C$ ——大小为 $(2d-1) \times 1, 1 \times (2d-1)$ 的卷积核;

$W_{\lfloor \frac{k}{d} \rfloor \times 1}^C$ 、 $W_{1 \times \lfloor \frac{k}{d} \rfloor}^C$ ——大小为 $\lfloor \frac{k}{d} \rfloor \times 1$ 、 $1 \times \lfloor \frac{k}{d} \rfloor$ 的卷积核；

$W_{1 \times 1}$ ——大小为 1×1 的卷积核；

h, w ——特征图的高度和宽度；

A^C ——注意力权重图；

F^C ——注意力权重图 A^C 与特征图 F^C 的哈达玛乘积；

C ——输入通道数。

1.4 改进的损失函数

YOLOv8n 采用 CIOU^[15] 作为边界框损失函数。CIOU 在考虑 IoU 值的基础上,还综合考虑了边界框的中心点距离和宽高比,从而提升了检测框的定位精度。但由于训练数据集往往含有低质量样本,距离、纵横比等几何度量都会加大对低质量样本的惩罚,从而降低模型的泛化性^[16]。Tong 等^[17]提出的 WIoU 能有效地解决低质量样本对模型的影响。基于距离注意力,构建了具有两层注意力机制的 WIoUv1 如式(5)和式(6)所示。

$$L_{\text{WIoUv1}} = R_{\text{WIoU}} L_{\text{IoU}}, \quad (5)$$

$$R_{\text{WIoU}} = \exp \left[\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*} \right], \quad (6)$$

式中:

R_{WIoU} ——距离度量函数；

L_{WIoUv1} ——WIoUv1 的边界框损失；

W_g, H_g ——最小围框的宽、高；

$*$ ——从计算图中分离 W_g, H_g ；

(x, y) ——预测框中心点坐标；

(x_{gt}, y_{gt}) ——目标框中心点坐标；

L_{IoU} ——基本边界框损失。

WIoUv3 通过定义一个离群度 β 来描述锚框的质量,较小的离群度意味着锚框质量较高,因此为其分配较小的梯度增益,以便使边界框回归更专注于质量一般的锚框。而对离群度较大的锚框分配较小的梯度增益,可以有效防止低质量样本产生较大的有害梯度。WIoUv3 公式如式(7)~式(9)所示。

$$\beta = \frac{L_{\text{IoU}}^*}{L_{\text{IoU}}} \in [0, +\infty), \quad (7)$$

$$r = \frac{\beta}{\delta \alpha^{\beta - \delta}}, \quad (8)$$

$$L_{\text{WIoUv3}} = r L_{\text{WIoUv1}}, \quad (9)$$

式中:

β, r ——离群度、梯度增益；

$L_{\text{IoU}}^*, \overline{L_{\text{IoU}}}$ ——单个样本 IoU 与平均 IoU；

α, δ ——可以调整的超参数；

L_{WIoUv3} ——WIoUv3 的边界框损失。

由于 $\overline{L_{\text{IoU}}}$ 是动态的,锚框的质量划分标准也是动态的,因此 WIoUv3 在每一时刻都能做出最符合当前情况的梯度增益分配策略。目前现有的基于交并比(IoU)的边框回归方法通过引入新的损失项来尝试加速模型的收敛过程,但这些方法往往忽视了 IoU 损失项本身所固有的局限性。为此,Zhang 等^[18]提出了 Inner-IoU。

Inner-IoU 引入了尺度因子(ratio)来控制辅助边框的尺度大小用于计算损失,同时加速边界框回归。对于高 IoU 样本,使用较小的辅助边框计算损失能够加速收敛,而对于低 IoU 样本,使用较大的辅助边框计算损失则更为合适。Inner-IoU 示意图如图 4 所示,Inner-IoU 关于辅助边框的尺度变换方法如式(10)~式(20)所示。

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} * r_{\text{atio}}}{2}, \quad (10)$$

$$b_r^{gt} = x_c^{gt} + \frac{w^{gt} * r_{\text{atio}}}{2}, \quad (11)$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} * r_{\text{atio}}}{2}, \quad (12)$$

$$b_b^{gt} = y_c^{gt} + \frac{h^{gt} * r_{\text{atio}}}{2}, \quad (13)$$

$$b_l = x_c - \frac{w * r_{\text{atio}}}{2}, \quad (14)$$

$$b_r = x_c + \frac{w * r_{\text{atio}}}{2}, \quad (15)$$

$$b_t = y_c - \frac{h * r_{\text{atio}}}{2}, \quad (16)$$

$$b_b = y_c + \frac{h * r_{\text{atio}}}{2}, \quad (17)$$

$$i_{\text{inter}} = [\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)] * [\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)], \quad (18)$$

$$u_{\text{union}} = (w^{gt} * h^{gt}) * (r_{\text{atio}})^2 + (w * h) * (r_{\text{atio}})^2 - i_{\text{inter}}, \quad (19)$$

$$I_{\text{IoU}^{\text{inner}}} = \frac{i_{\text{inter}}}{u_{\text{union}}}, \quad (20)$$

式中:

w, h ——锚框的宽度、高度；

w^{gt}, h^{gt} ——真实框的宽度、高度；

r_{atio} ——尺度因子,且 $r_{\text{atio}} \in [0.5, 1.5]$ ；

x_c^{gt}, y_c^{gt} ——真实框、真实框内部的中心点；

x_c, y_c ——锚框和锚框内部的中心点；

$b_l^{gt}, b_r^{gt}, b_t^{gt}, b_b^{gt}$ ——真实框的左边界、右边界、顶部边界和底部边界；

b_l, b_r, b_t, b_b ——锚框的左边界、右边界、顶部边界和底部边界；

$i_{\text{inter}}, u_{\text{union}}$ ——锚框与真实框的交集、并集；

$I_{\text{IoU}^{\text{inner}}}$ ——交并比。

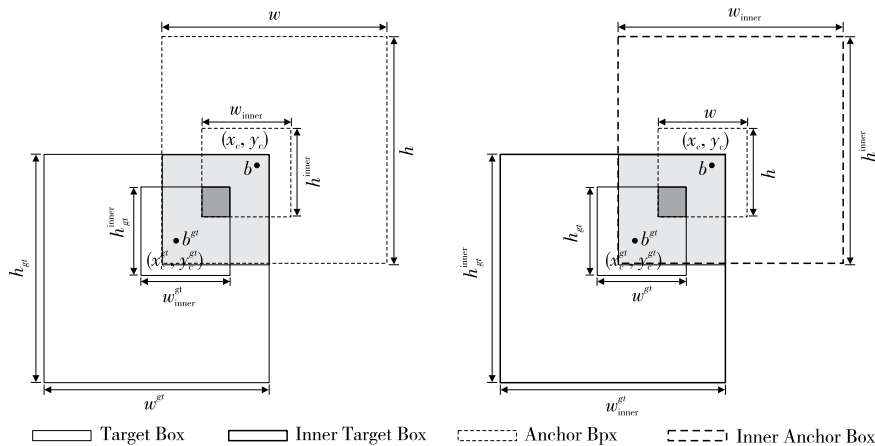


图4 Inner-IoU示意图

Figure 4 Schematic diagram of Inner-IoU

虽然 Inner-IoU 克服了传统 IoU 无法根据不同检测器与检测任务进行自我调整,泛化性差的问题,但也容易面临低质量样本对模型损害的问题。因此,将 Inner-IoU 与 WIoU 相结合形成 Inner-WIoU,使用尺度因子(ratio)控制生成不同尺度的辅助边框用于计算损失,加快边界框回归,同时也能够减小低质量样本产生的有害梯度。Inner-WIoU 公式如式(21)和式(22)所示。

$$L_{\text{Inner-IoU}} = 1 - I_{\text{IoU}^{\text{inner}}}, \quad (21)$$

$$L_{\text{Inner-WIoU}} = L_{\text{WIoUv3}} + L_{\text{IoU}} - I_{\text{IoU}^{\text{inner}}}, \quad (22)$$

式中:

$L_{\text{Inner-IoU}}$ ——Inner-IoU 的边界框损失;

$L_{\text{Inner-WIoU}}$ ——Inner-WIoU 的边界框损失。

2 试验结果与分析

2.1 试验配置

试验环境为 Pytorch 深度学习框架,CPU 为 Intel(R) Xeon(R) CPU E5-2696 v3,GPU 为 NVIDIA RTX 3060Ti,操作系统为 Linux,输入图像尺寸为 640×640 ,批处理大小为 16,训练周期(epoch)为 300,初始学习率为 0.010 0,权重衰减系数为 0.000 5,动量为 0.937,编程语言为 Python3.11,Cuda 版本为 12.2。

2.2 数据集介绍

采用自制的白酒包装数据集,在湖南吉首某公司实地采集,然后利用 Labelling 进行手工数据标注。数据集共计 918 张图片,每张图片中包含一种或多种缺陷,其中缺陷类型分为 6 种,包括印刷不良、脏迹、褶皱、划痕、折痕、破损。数据集部分图片如图 5 所示。

为了增强白酒包装缺陷数据集的多样性和提高模型的鲁棒性与泛化能力,对原始数据集采用了一系列的数据增强方法进行处理。包括加噪声、调整亮度、旋转、裁剪和平移等操作,通过这些增强技术,模型能够更好地适应在不同光照、角度和噪声条件下的检测任务。这种处理不仅有效增加了样本的多样性,还模拟了各种可能的实际场景,使得模型在应对白酒工业质检复杂工况时表现更加稳健。经过数据增强处理后,最终生成了一个包含 2 754 张图片的数据集。数据集按照 7:1:2 的比例进行划分,其中 70% 用于训练,10% 用于验证,20% 用于测试。这种划分策略确保了模型在训练阶段能够充分学习,在验证阶段能够及时调整,在测试阶段能够准确评估其在真实场景中的性能。

2.3 评价指标

在评估网络性能时,通常需要同时关注精确率和召回率,其计算式分别见式(23)和式(24)。采用平均精度

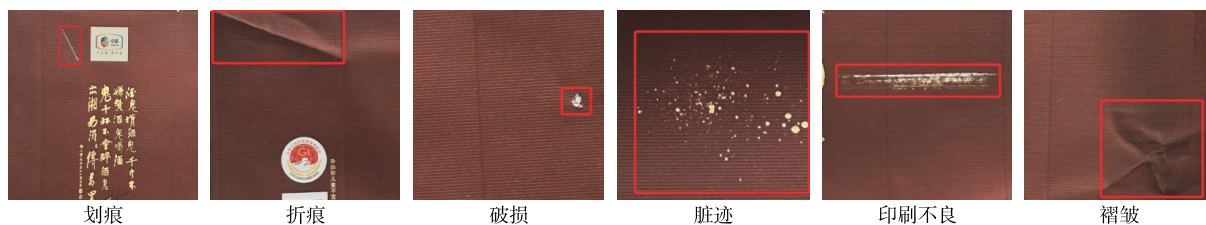


图5 白酒包装缺陷分类

Figure 5 Classification of liquor packaging defects

(AP)和平均精度均值(mAP)作为主要评价指标,以全面反映模型的检测性能。

$$R_{\text{recall}} = \frac{T_p}{T_p + F_N}, \tag{23}$$

$$P_{\text{recision}} = \frac{T_p}{T_p + F_p}, \tag{24}$$

式中:

R_{recall} ——召回率,%;

P_{recision} ——精确率,%;

T_p ——实际为正类且模型也预测为正类;

F_N ——实际为正类但模型预测为负类;

F_p ——实际为负类但模型预测为正类。

平均精度是指在不同召回率下计算出的精度的平均值,公式如式(25)所示。在此基础上,平均精度均值则是所有不同召回率下计算出的平均精度(AP)的均值,其公式如式(26)所示。

$$A_p = \int_0^1 P(t) dt, \tag{25}$$

式中:

A_p ——平均精度,%;

$P(t)$ ——不同召回率下的精确度。

$$m_{\text{AP}} = \frac{\sum_{i=1}^k (A_p)_i}{k}, \tag{26}$$

式中:

k ——类别的总数;

$(A_p)_i$ ——第*i*类的平均精度;

m_{AP} ——平均精度均值,%。

此外,在目标检测任务中通常使用 mAP@0.5 表示在 IOU 阈值为 0.50 时计算的平均精度均值,用 mAP@0.5:0.95 表示在多个 IOU 阈值下(从 0.50 到 0.95,步长为 0.05)计算出的平均精度均值。

在评估模型性能时还需兼顾计算效率。计算量(FLOPS)衡量模型的计算复杂度,参数量代表可训练参数总数,影响模型的复杂性与资源消耗;而检测速度(FPS)则表示每秒处理的图像帧数,是衡量实时性能的重要指标。

2.4 试验分析

2.4.1 损失函数对比 在相同的试验条件下,为了找到最合适的损失函数,进行了多种损失函数的对比和评估,结果如表 1 所示。第 1 组试验使用的是 YOLOv8n 原始损失函数 CIoU,作为基准试验,第 2~5 组试验选择了 WIoU 的 3 个版本进行对比,第 6~8 组试验选择了 WIoU 的 3 个版本,第 5~8 组试验结合 Inner-IOU、Inner-WIoU 等改进版本。

试验结果表明,CIoU 作为基准损失函数,在 mAP@0.5 上达到 81.2%,但其在更严格的 mAP@0.5:0.95 上表现一般,仅为 42.6%。WIoU 系列通过引入权重机制,使得不同版本在 mAP@0.5 上有所优化,其中 WIoU-v1 表现最佳,mAP@0.5 为 81.5%,mAP@0.5:0.95 达到 44.5%,说明其权重机制在复杂场景中对边界框优化有显著效果。相比之下,WIoU-v2 表现稍差,mAP@0.5 降至 79.9%。

在引入内框优化机制的 Inner-WIoU 系列损失函数中,Inner-WIoU-v3 在多项指标上均优于其他损失函数。在 mAP@0.5 指标上,其表现达到了 82.1%,而在 mAP@0.5:0.95 上也达到了 44%。同时,其参数量也保持不变,表明其能在保持模型轻量化的前提下,实现精度的提升。

表 1 损失函数对比试验

Table 1 Comparison experiment of loss functions					
模型	精确率/ %	召回率/ %	mAP@ 0.5/%	mAP@0.5: 0.95/%	参数量/ M
CIoU	81.5	76.9	81.2	42.6	3.01
WIoU-v1	81.1	71.9	81.5	44.5	3.01
WIoU-v2	81.6	69.5	79.9	42.3	3.01
WIoU-v3	82.3	75.8	80.9	41.2	3.01
Inner-IOU	80.3	71.7	79.1	43.1	3.01
Inner-WIoU-v1	78.0	72.6	79.3	43.0	3.01
Inner-WIoU-v2	75.9	73.2	78.4	42.8	3.01
Inner-WIoU-v3	84.9	74.3	82.1	44.0	3.01

2.4.2 SPPF-LSKA 热力图分析 为验证 SPPF-LSKA 模块的引入对特征提取的增强效果,对比了引入该模块前后模型的热力图可视化结果,如图 6 所示。通过对比发现,加入 SPPF-LSKA 模块后,模型对缺陷区域的关注度显著提升,尤其是在处理细微和多尺度特征时,表现出更强的捕捉能力。这表明 SPPF-LSKA 模块能够有效增强模型对关键特征的感知,减少漏检和误检现象,使检测结果更加精准。

2.4.3 消融试验 为了验证改进模型的有效性,共设计 6 组消融试验,在 YOLOv8n 网络基础上引入改进模块,结果如表 2 所示。以第 1 组试验为基准,该模型的 mAP@0.5 为 81.2%。第 2 组试验采用轻量化下采样模块 ADown 后,较基准网络 mAP@0.5 提高了 0.8%,参数量和计算量均有不同程度的下降,检测速度提升了 7 ms,验证了切分特征图处理策略的有效性。第 3 组试验采用 SPPF-LSKA 模块后,mAP@0.5 较基准网络提高了 1.5%,参数量与计算量略有增加,检测速度保持不变。第 4 组试验采用 Inner-WIoU 损失函数,mAP@0.5 较基准网络提高了 0.9%,参数

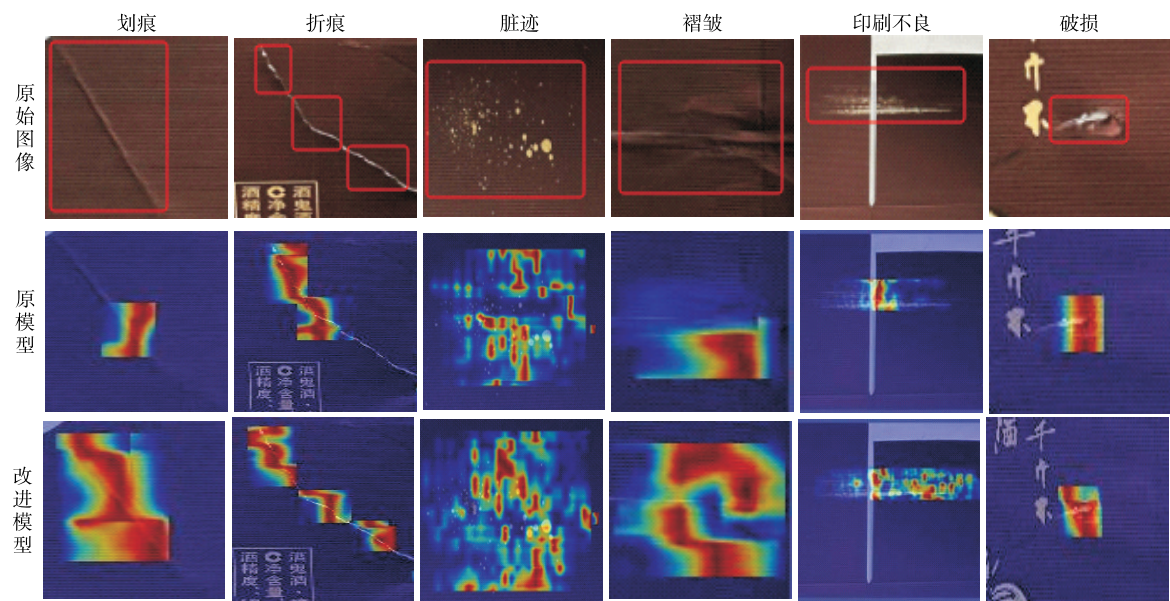


图6 引入SPPF-LSKA模块前后的热力图

Figure 6 Heatmaps before and after the introduction of the SPPF-LSKA module

量与计算量保持不变,检测速度增加了3 ms。第5组试验在采用ADown模块的基础上加入SPPF-LSKA模块,mAP@0.5较基准网络提升了3.6%,同时参数量和计算量略有增加,检测速度降低了4 ms,主要归功于级联的深度卷积和深度扩展卷积增强了特征提取能力,而分解卷积核的思想减少了参数量的提高。第6组试验最后加入Inner-WIoU损失函数后,较基准网络mAP@0.5提高了5.2%,参数量和计算量保持不变,检测速度提高了7 ms,表明了Inner-WIoU损失函数的有效性。

消融试验结果表明,改进模型具有更小的参数量、更高的精度、更低的计算量和更高的检测速度,可应用于工业设备部署。

图7展示了改进的YOLOv8n模型与YOLOv8n模型的性能指标。由图7可知,改进的YOLOv8n模型在各个阶段的平均精度均值曲线均优于YOLOv8n模型。早期的训练阶段改进的YOLOv8n模型就展现出了较高的性

能,随着训练轮数的增加,曲线稳步上升,并在后期稳定在一个更高的值。这表明改进模型在多个尺度上都有较好的适应性和泛化能力。

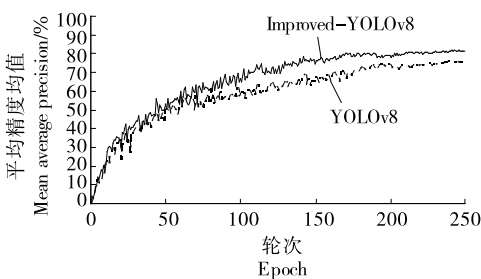


图7 性能指标对比

Figure 7 Comparison of performance indicators

2.5 不同模型性能对比试验

为了进一步验证改进模型的性能,在相同的试验条件下,将其与当前先进的目标检测模型进行对比,结果如

表2 消融试验

Table 2 Ablation experiment

试验号	ADown	SPPF-LSKA	Inner-WIoU	参数量/M	mAP@0.5/%	mAP@0.5:0.95/%	计算量/G	速度/ms
1				3.01	81.2	42.6	8.1	89
2	✓			2.59	82.0	46.6	7.4	80
3		✓		3.28	82.7	45.6	8.3	89
4			✓	3.01	82.1	44.0	8.1	92
5	✓	✓		2.87	84.8	49.4	7.6	85
6	✓	✓	✓	2.87	86.4	47.9	7.6	96

表 3 所示。YOLOv8n、YOLOv5n、YOLOv10n、Mamba-YOLOv8 4 个模型具备参数量少、计算量低等优势,但其检测精度仍需进一步提升;YOLOv8s、YOLOv7-tiny、TOOD、DINO、ATSS、RetinaNet 6 个模型虽然取得了不错的检测精度,但模型过于复杂,计算量较大,这使得在资源有限的设备上部署变得困难;RT-DETR、YOLOv6n、YOLOv9t、Faster-RCNN 4 个模型检测精度较

低,难以胜任复杂情况下白酒包装缺陷检测任务。相比以上模型,改进的 YOLOv8n 模型在保证计算量与参数量低的前提下,明显提升了检测精度,参数量仅为 2.87 M,mAP@0.5 达到了 86.4%,计算量仅为 7.6 G,可满足实际白酒包装生产线的检测需求。

2.6 检测结果可视化分析

如图 8 所示,YOLOv8n 模型在识别较小尺寸目标缺

表 3 不同模型性能对比试验
Table 3 Comparative experiment on the performance of different models

模型	参数量/M	mAP@0.5/%	mAP@0.5:0.95/%	计算量/G	精确率/%	召回率/%
YOLOv5n	1.77	79.6	38.8	4.2	81.0	74.8
YOLOv5s	7.03	85.7	48.6	15.8	86.3	81.1
YOLOv7-tiny	6.02	84.0	40.8	13.1	82.4	80.2
YOLOv8n	3.01	81.2	42.6	8.1	81.5	76.9
YOLOv8s	11.13	86.3	50.7	28.4	86.0	82.5
RT-DETR ^[19]	28.46	67.2	34.3	100.6	70.2	60.6
YOLOv9t	1.97	70.6	35.0	7.6	75.8	62.4
YOLOv6n	4.23	63.0	30.2	11.8	72.6	55.3
Mamba-YOLOv8	5.99	80.9	45.0	13.6	82.6	75.4
YOLOv10n ^[20]	2.27	74.1	39.5	6.5	76.4	64.9
Faster-RCNN	41.37	55.6	—	164.0	—	—
TOOD ^[21]	32.03	82.6	—	154.0	—	—
DINO ^[22]	47.55	85.5	—	217.0	—	—
ATSS ^[23]	38.90	85.7	—	85.1	—	—
RetinaNet ^[24]	36.43	78.3	—	161.0	—	—
Improved-YOLOv8n	2.87	86.4	47.9	7.6	81.7	83.0

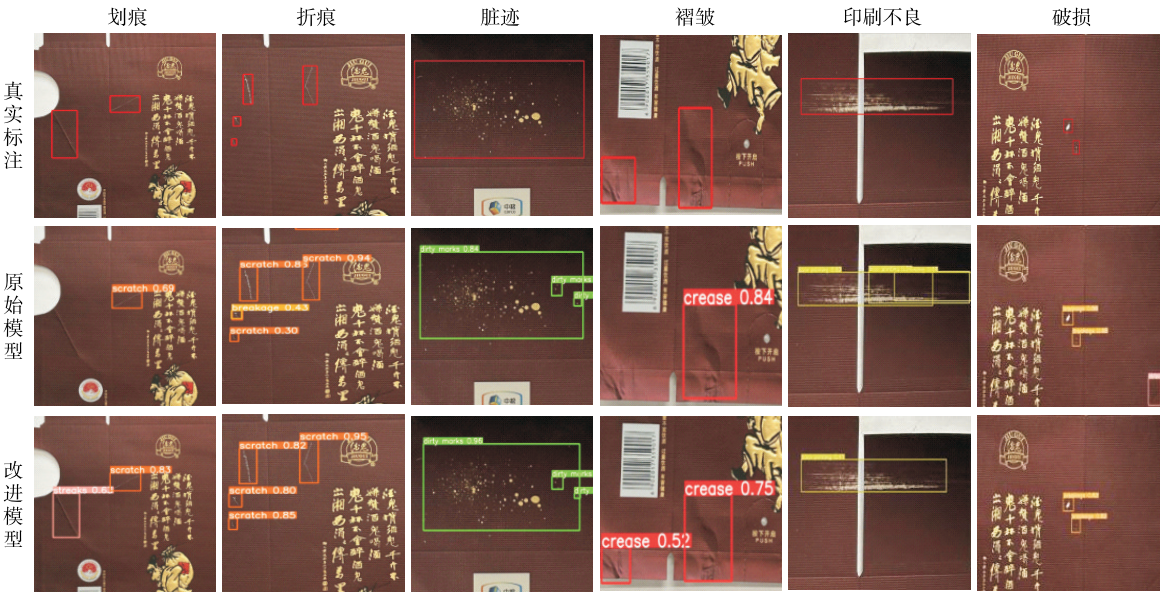


图 8 改进模型检测结果的对比

Figure 8 Comparison of detection results from improved model

陷(如划痕、折痕、褶皱)以及较不明显的缺陷时,存在较明显的漏检问题。此外,YOLOv8n模型在缺陷检测过程中还经常伴随误检现象,尤其是在背景复杂或缺陷边界模糊的情况下,容易将某一类缺陷误识为另一类缺陷。相比之下,改进的YOLOv8n模型在面临这些情况时展现出更强的鲁棒性。通过优化损失函数和模型架构,改进的YOLOv8n模型在检测小目标和难以辨别的缺陷方面表现出更高的准确性,显著减少了漏检和误检的发生。

2.7 通用性试验

试验选自百度飞桨平台公开发布的瓦楞纸板缺陷数据集。该数据集包含多种缺陷类型,如纸板破损、划痕等。数据集共计1 056张照片,每张照片包含多种缺陷。在保证相同参数的情况下,将改进的YOLOv8n模型和YOLOv8n模型在相同的环境下训练200个周期,结果如表4所示。相较于原模型,改进后的模型在mAP@0.5上高出1.1%,mAP@0.5:0.95上高出1.1%,精确率上高出3.3%,同时参数量与计算量更小。该试验结果表明,改进后的模型不仅在特定数据集上表现出良好的检测效果,同时在其他缺陷检测数据集上也展现了较好的性能。

表4 对比试验结果

Table 4 Comparative experimental results

模型	参数 量/M	mAP@ 0.5/%	mAP@0.5: 0.95/%	计算 量/G	精确 率/%	召回 率/%
YOLOv8n	3.01	91.5	63.9	8.1	89.1	84.8
改进的YOLOv8n	2.87	92.6	65.0	7.6	92.4	84.0

3 结论

针对白酒包装缺陷的问题,提出了一种基于改进YOLOv8n的模型。①在网络结构中引入ADown模块,既保持了网络的特征提取能力,又显著降低了参数量和计算量,使模型更适合工业生产和资源受限的应用场景;②提出了SPPF-LSKA模块,将大尺寸可分离核注意力机制引入SPPF模块,提升了模型对不同尺度特征的提取能力;③设计了一种融合Inner-IoU和WIoU的损失函数,有效加速了边框回归,同时减少了低质量样本的负面影响。与YOLOv5、YOLOv6、RT-DETR、Faster-RCNN等经典目标检测模型相比较,改进的模型在白酒包装缺陷检测精度方面有所提升,同时计算量与参数量也有所下降,这表明了改进的模型在白酒包装缺陷检测中具有更好的适应性。未来工作将进一步研究如何在不增加计算量和参数量的前提下,进一步提升对细小缺陷的检测能力,以更好

地适应复杂的白酒包装生产工况环境。

参考文献

- [1] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]// Computer Vision-ECCV 2016: 14th European Conference. Amsterdam: Springer International Publishing, 2016: 21-37.
- [2] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]// 2016 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE Press, 2016: 779-788.
- [3] 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[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 7 464-7 475.
- [4] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Hawaii: IEEE, 2017: 7 263-7 271.
- [5] GIRSHICK R. Fast R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision. Santiago: IEEE Press, 2015: 1 440-1 448.
- [6] 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): 106680.
- [7] 夏军勇, 王康宇, 周宏娣. 基于改进Faster R-CNN的食品包装缺陷检测[J]. 食品与机械, 2023, 39(11): 131-136, 151.
- [8] 李志诚, 曾志强. 基于改进YOLOv3的卷纸包装缺陷实时检测算法[J]. 中国造纸学报, 2022, 37(2): 87-93.
- [9] 李建明, 杨挺, 王惠栋. 基于深度学习的工业自动化包装缺陷检测方法[J]. 包装工程, 2020, 41(7): 175-184.
- [10] 叶宇星, 孙志锋, 马风力, 等. 基于改进YOLOv5s的腌制蔬菜真空包装缺陷检测[J]. 包装工程, 2023, 44(9): 45-53.
- [11] 曾秀云, 陆华才, 吕禾丰. 基于改进Faster R-CNN的棉布包

- 装缺陷检测的方法研究[J]. 电子测量与仪器学报, 2022, 36(4): 179-186.
- ZENG X Y, LU H C, LYU H F. Research on cotton packaging defect detection method based on improved Faster R-CNN[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(4): 179-186.
- [12] LIN G J, LIU K Y, XIA X K, et al. An efficient and intelligent detection method for fabric defects based on improved YOLOv5[J]. Sensors, 2022, 23(1): 97.
- [13] WANG C Y, YE H I, MARK LIAO H Y. Yolov9: learning what you want to learn using programmable gradient information[C]// European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2025: 1-21.
- [14] LAU K W, PO L M, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN [J]. Expert Systems with Applications, 2024, 236: 121352.
- [15] ZHENG Z, WANG P, REN D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2021, 52(8): 8 574-8 586.
- [16] 张嘉雄, 曾水玲, 李昭贤. 基于改进 YOLOv7 的钢板表面缺陷检测的方法[J/OL]. 控制工程. (2024-04-01) [2024-09-15]. <https://doi.org/10.14107/j.cnki.kzgc.20231034>.
- ZHANG J X, ZENG S L, LI Z X. Method for surface defect detection in steel plates based on improved YOLOv7 [J/OL]. Control Engineering of China. (2024-04-01) [2024-09-15]. <https://doi.org/10.14107/j.cnki.kzgc.20231034>.
- [17] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[J/OL]. ArXiv. (2023-01-24) [2024-09-28]. <https://arxiv.org/abs/2301.10051>.
- [18] ZHANG H, XU C, ZHANG S. Inner-IoU: more effective intersection over union loss with auxiliary bounding box[J/OL]. ArXiv. (2023-11-06) [2024-09-28]. <https://arxiv.org/abs/2311.02877>.
- [19] LV W Y, XU S L, ZHAO Y A, et al. DETRs beat YOLOs on real-time object detection[C]// proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2024: 16 965-16 974.
- [20] WANG A, CHEN H, LIU L, et al. Yolov10: real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107 984-108 011.
- [21] FENG C J, ZHONG Y J, GAO Y, et al. TOOD: task-aligned one-stage object detection[C]// 2021 IEEE/CVF International Conference on Computer Vision. New York: IEEE Press, 2022: 3 490-3 499.
- [22] ZHANG H, LI F, LIU S, et al. Dino: detr with improved denoising anchor boxes for end-to-end object detection[J/OL]. ArXiv. (2022-03-07) [2024-09-28]. <https://arxiv.org/abs/2203.03605>.
- [23] ZHANG S F, CHI C, YAO Y Q, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2020: 9 756-9 765.
- [24] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.