

基于 KCC-YOLOv5 的铝型材表面缺陷检测

邓光伟¹, 尤红权², 朱志松^{1*}¹南通大学机械工程学院, 江苏 南通 226019;²南通国尚精密机械有限公司, 江苏 南通 226017

摘要 针对铝型材表面缺陷种类多、尺度差异大、小目标容易漏检等问题,提出了KCC-YOLOv5——一种基于YOLOv5s改进的铝型材表面小缺陷检测模型。首先利用IoU(intersection over union)-K-means++算法代替K-means算法聚类锚框,获得最贴合铝型材表面缺陷的锚框,提高小目标锚框的质量;其次,提出全局注意力模块C3C2F,并引入主干层,在减少参数量的同时增强小目标的语义信息和全局感知能力;最后将颈部最近邻插值上采样方式换为轻量级上采样算子CARAFE(content-aware reassembly of features),充分保留上采样特征图的小目标信息。实验结果表明,改进模型KCC-YOLOv5的均值平均精度为94.6%,相比于YOLOv5s提高了2.8个百分点,小目标漆泡和脏点的平均精度分别提高了5.2和12.4个百分点。KCC-YOLOv5模型在保持大目标检测精度小幅度提升的同时显著提升了小目标的检测精度。

关键词 机器视觉; 铝型材; 表面缺陷; KCC-YOLOv5; 小目标检测

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP230950

Defect Detection on Aluminum Profile Surface Based on KCC-YOLOv5

Deng Guangwei¹, You Hongquan², Zhu Zhisong^{1*}¹School of Mechanical Engineering, Nantong University, Nantong 226019, Jiangsu, China;²Nantong Guoshang Precision Machinery Co., Ltd, Nantong 226017, Jiangsu, China

Abstract To address the issue of various types and large-scale differences of surface defects in aluminum profiles, as well as the tendency for small targets to be missed, we suggest an improved detection model for small defects on the surface of aluminum profiles based on YOLOv5s, called KCC-YOLOv5 model. First, the IoU(intersection over union)-K-means++ algorithm is used to cluster anchor frames in place of the K-means algorithm, aiming to obtain the anchor frames that best fit the surface defects of aluminum profiles and improve the quality of small target anchor frames. Second, a global attention module C3C2F is proposed and introduced into the backbone layer to enhance the semantic information and global perception of small targets while reducing the number of parameters. Finally, the neck nearest neighbor interpolation upsampling method is replaced by a lightweight upsampling operator CARAFE(content-aware reassembly of features), which fully retains the small target information of the upsampled feature map. The experimental results show that the mean average precision of the improved KCC-YOLOv5 model is 94.6%, which represents 2.8 percentage points improvement compared to YOLOv5s. Furthermore, the average precision for small targets, such as bubbles and spots, are increased by 5.2 and 12.4 percentage points, respectively. Overall, the KCC-YOLOv5 model significantly enhances the detection accuracy of small targets while maintaining a small improvement in the detection accuracy of large targets.

Key words machine vision; aluminum profile; surface defects; KCC-YOLOv5; small targets detection

1 引言

随着全球工业化进程的快速发展,铝型材因其卓越的性能优势——低密度、高可塑性和耐腐蚀等,正在成为基础建设和工业领域的首选材料,其需求持续增

长。然而,受到加工工艺和生产设备等影响,铝型材表面容易产生擦花、漆泡、脏点等缺陷,这些缺陷可能会降低产品的质量和可靠性,甚至导致生产事故和损失。因此,需要通过表面质量检测剔除有缺陷的产品。

传统的表面缺陷人工检测效率低、成本高、精度

收稿日期: 2023-03-27; 修回日期: 2023-05-15; 录用日期: 2023-06-01; 网络首发日期: 2023-06-11

基金项目: 江苏省产学研合作项目(BY2022232)

通信作者: *zhu.zhs@ntu.edu.cn

差。随着计算机算力的提高以及各种优秀深度学习网络的不断发展,基于深度学习的铝型材表面缺陷检测得到了广泛的应用。Wei等^[1]提出了一种基于Faster R-CNN(convolutional neural network)的多尺度缺陷检测网络,将FPN(feature pyramid network)的特征融合思想添加到Faster R-CNN中,但检测速度较慢。Sun等^[2]提出了一种改进的YOLOv3网络,引入更大的预测尺度,设计注意力模块提取更多的缺陷特征,构建双塔结构加强不同层次特征的融合,但该模型检测的缺陷类型较少。Wang等^[3]提出了一种改进的MS(multi-scale fusion)-YOLOv5模型,用PE(poly-scale efficient channel attention)-Neck结构代替原算法的颈部以增强模型在不同尺度下提取和定位缺陷的能力,将多流网作为算法的第一个检测头来增强模型对随机分布缺陷的识别能力,但该模型检测的缺陷类型以大目标为主,不能很好地涵盖实际缺陷类型。

目标检测主要分为两类:一类是两阶段算法,如R-CNN^[4]、Fast R-CNN^[5]、Faster R-CNN^[6]等;另一类是一阶段算法,如YOLO^[7]、SSD(single shot multibox detector)^[8]。其中一阶段算法YOLO因其高精度和实时性而被广泛应用于表面缺陷检测任务中,但其对于小目标的检测性能仍有待提升。Chen等^[9]提出了一种基于YOLOv3的改进算法,通过增加检测头让网络学习更多浅层的位置信息,在每个检测层之前添加SPP(spatial pyramid pooling)模块来融合不同尺度的特征,有效提高了小目标的检测精度。徐志京等^[10]提出一种基于纹理和颜色双重特征增强的舰船小目标检测方法:设计生成对抗网络增强舰船小目标的纹理特征;利用深度强化学习算法增强图像的颜色信息;设计自适应变换特征金字塔网络增强全局感受野;利用特征细化模块和圆形光滑标签完成对舰船目标边界框的特征点对齐和角度回归,实现了小目标检测精度的提升。张寅等^[11]提出了一种基于特征融合与注意力机制的遥感图像小目标检测算法FFAM(feature fusion and attention mechanism)-YOLO,利用特征增强模块(FEM)、特征融合模块(FFM)和级联注意力机制(ESM),提高了模型对小目标的检测性能。程松等^[12]提出了一种基于YOLOv5-Tiny的轻量型焊缝缺陷识别方法,在Backbone中加入注意力机制SELayer(squeeze-and-excitation layer),用GhostBottleneck模块替换Head层中的C3模块并去除大目标检测头,提高了焊缝内部小目标缺陷的识别精度。Zhu等^[13]将Transformer和CBAM(convolutional block attention module)结构引入到YOLOv5模型中,提升了无人机捕获场景中小目标的检测精度。在YOLO算法中,小目标检测精度低、检测能力不足、容易漏检的主要原因有以下几个方面:1)小目标可利用的特征有限,并且在后续网络层中特征信息容易丢失,从而导致算法难以捕捉到小目标的关键特征;2)小目标的样本数量较少

且容易受到背景干扰,使网络难以充分学习到小目标的特征;3)先验框设置不合理,导致其尺寸不适合小目标检测。

为了解决铝型材表面缺陷检测速度较慢、精度较低、缺陷类型不全面和小目标检测能力不足等问题,本文提出了一种改进的模型KCC-YOLOv5,用于铝型材表面7种不同类型尺度缺陷的检测,结合IoU-K-means++算法、C3C2F模块和轻量级上采样算子CARAFE^[14],实现锚框聚类优化和小目标语义信息增强,提升铝型材表面小目标缺陷的检测精度。

2 改进YOLOv5网络

2.1 铝型材表面缺陷检测框架

YOLOv5因其优良的检测能力和检测速度而被广泛应用于目标检测领域。YOLOv5框架主要分为3个部分:主干、颈部、头部。其中:主干部分提取图片特征;颈部融合特征;头部判别目标的类别和位置。YOLOv5s更好地平衡了精度和速度,基于YOLOv5s改进,提出了KCC-YOLOv5模型,总体结构如图1所示。KCC-YOLOv5的主干部分包括:CBS[convolution+batch normalization(BN)+SiLU]、C3(Conv3Bottleneck)、C3C2F(Conv3Conv2Former)和SPPF(spatial pyramid pooling fast)模块,其中CBS模块用于提取图像的局部特征并进行下采样;C3模块由3个标准卷积和若干瓶颈模块组成,用于学习残差特征,提高推理速度;C3C2F模块聚焦图像的重要特征,抑制背景和噪声;SPPF模块包含3个不同尺寸的最大池化层(5 pixel × 5 pixel、9 pixel × 9 pixel、13 pixel × 13 pixel),实现多尺度局部特征和全局特征的融合,有效提高图像的感受野,丰富特征图像的表达力。颈部主要由CARAFE(content-aware reassembly of features)、Concat和C3模块组成:CARAFE模块实现图像的上采样;Concat和C3模块融合不同尺度的特征图,以获取丰富的特征信息。头部的Conv模块用于调整每个检测头最终的输出通道数,以便预测目标的类别和位置。

2.2 聚类优化锚框

在YOLOv5s算法中,锚框是为COCO(common objects in context)数据集预设的。然而,本文所使用的铝型材数据集与COCO数据集存在较大差异,因此需要对锚框进行新的聚类。K-means++算法可以解决YOLOv5s中K-means算法在初始聚类中心上的依赖性问题。在K-means++的初始步骤中,仅会随机选择一个聚类中心,然后按照距离簇中心远近的概率依次确定其他聚类中心。由于K-means++算法在聚类中心的选择上具有离散性和全局性,因此可以获得更加准确和合理的聚类结果。采用IoU(intersection over union)代替欧氏距离估计两物体的相似度^[15],将该方法命名为IoU-K-means++,表示为

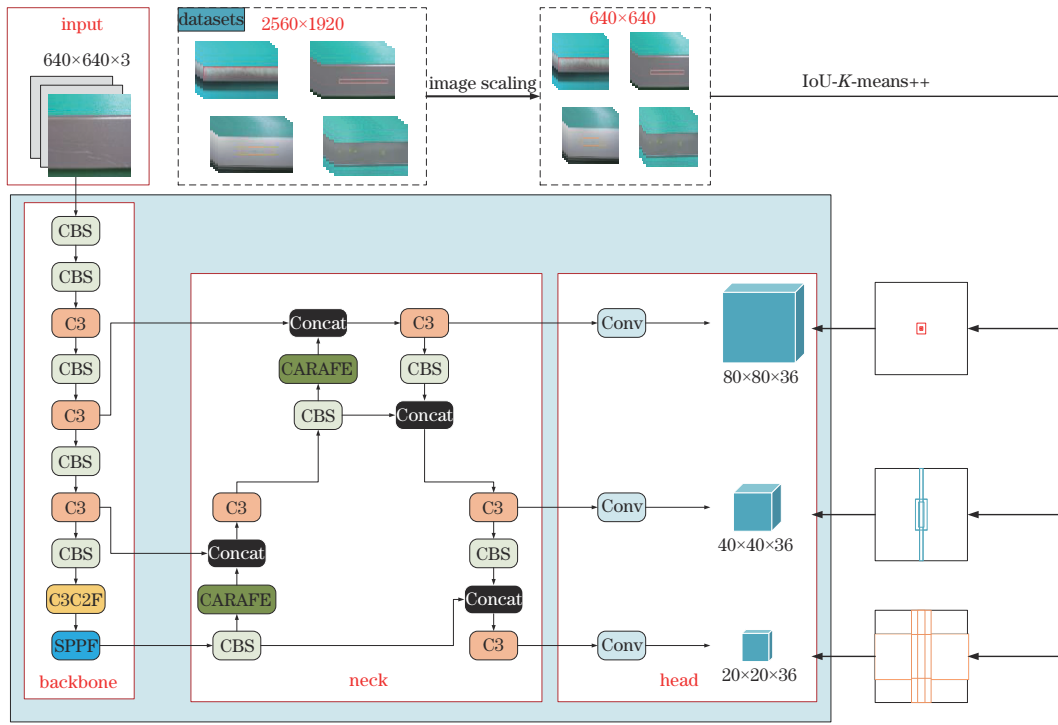


图 1 KCC-YOLOv5 缺陷检测网络结构

Fig. 1 Architecture of KCC-YOLOv5 defect detection network

$$d(b, c) = 1 - I_{IoU}(b, c), \quad (1)$$

式中： b 为真实框； c 为聚类中心； I_{IoU} 为真实框和聚类中心交集与并集的比例， I_{IoU} 越高代表真实框与聚类中心越接近，即两者相关度越高。图 2 展示了所提铝型材

数据集标签的形状和大小，以及 COCO 数据集上的 9 个锚框和 IoU-K-means++ 算法生成的 9 个新锚框，发现 IoU-K-means++ 算法聚类出的锚框更适合铝型材数据集表面的缺陷。

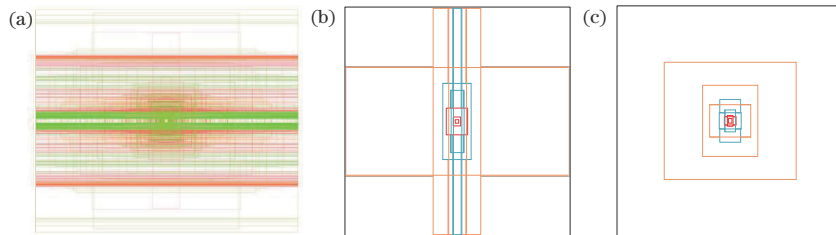


图 2 锚框对比图。(a) 铝型材数据集标签分布；(b) 新锚框；(c) 原始锚框

Fig. 2 Comparison of anchor boxes. (a) Distribution of the aluminum profile dataset labels; (b) new anchor boxes; (c) original anchor boxes

2.3 C3C2F 模块

自 2020 年以来，Vision Transformer^[16]进一步促进了视觉识别模型的发展，在图像分类、目标检测和语义分割等任务中展现出优秀的性能。与只进行局部建模的卷积操作相比，Transformer^[17]中的自注意力机制提供了一种更有效的空间信息编码方法。Conv2Former^[18]中的卷积调制模块(ConvMod)通过计算大核卷积的注意力权重 A 和输入特征 V 之间的 Hadamard 乘积来模拟自注意力机制，既减少了计算量，又有效增强了目标的语义信息和上下文感知能力。自注意力机制和卷积调制模块的对比如图 3 所示，其中： $H \times W \times C$ 是特征图的尺寸； V 为输入特征的向量； Q, K 是计算注意力权重 A 的特征向量； k 为卷积核。通过计算当前 Q 和所有 K 之间的相似度，得到一组注意力权重，然后将这组注意力权

重与对应的 V 进行乘积运算，得到基于注意力的 V 值。

令卷积调制模块 ConvMod 替换 C3 模块中的 BottleNeck 1，并将其命名为 C3C2F。C3C2F 模块能够对全局上下文信息进行建模，不仅减少了模型的参数量，而且增强了铝型材表面小目标缺陷的语义信息，同时抑制了背景和噪声的干扰。

C3C2F 模块的具体结构如图 4 所示，其中的卷积调制模块包括两个分支：第一个分支采用 1×1 卷积保留原特征图信息；第二个分支首先使用层归一化(LayerNorm)对输入特征进行归一化，以确保数据特征分布的稳定性，接着通过 1×1 卷积和激活函数 GELU(Gaussian error linear unit)，在保证特征图尺寸不变的前提下大幅增加非线性特性，最后，用一个 11×11 的深度卷积在减少参数量的同时扩大感受野。

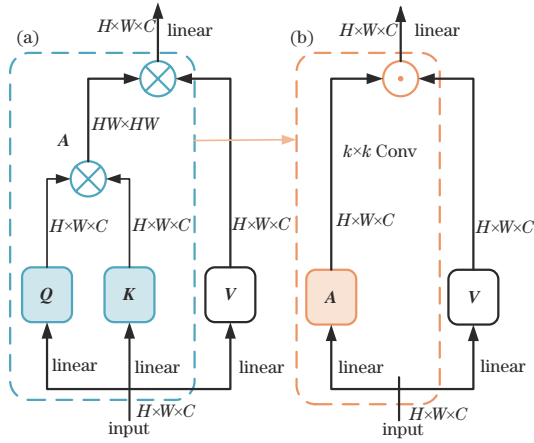


图 3 自注意力机制和卷积调制模块比较。(a)自注意力机制；(b)卷积调制模块

Fig. 3 Comparison between the self-attention mechanism and the convolutional modulation. (a) Self-attention mechanism; (b) convolutional modulation

在 ConvMod 模块中,第一个分支对应传统自注意力机制中的 V 值,第二个分支对应传统自注意力机制中的注意力权重 A 。将第二分支当作权重来调制第一分支,充分利用图像的上下文信息,使模型更集中于待检目标,减少不相关背景的干扰。

2.4 上采样模块 CARAFE

为了进一步增强模型对铝型材表面小目标缺陷的识别能力,在 YOLOv5s 模型的颈部,用轻量级上采样算子 CARAFE 替换最近邻插值上采样方式,以扩大感受野并自适应输入内容。CARAFE 结构如图 5 所示,包括两部分:上采样核预测模块和特征重组模块。

假设上采样倍率为 σ , CARAFE 模块的上一层输出特征图 X 的尺寸为 $H \times W \times C$ 。在上采样核预测模块中:为了降低后续步骤的计算量,首先用一个 1×1 卷积将通道数压缩到 C_m ;然后用一个 $k_{encoder} \times k_{encoder}$ 的卷积层预测上采样核,得到尺寸为 $\sigma H \times \sigma W \times k_{up}^2$ 的上采样核;最后用 Softmax 函数对上采样核进行归一化,使得卷积核权重和为 1。在特征重组模块中,取出输入特征图 X 上以目标位置 l 为中心的 $k_{up} \times k_{up}$ 区域 $[N(x_l, k_{up})]$ 和以 l 为目标位置的重组内核 W_l 作点积,最终得到输出特征图 X' 。

增大 $k_{encoder}$ 可以扩大感受野,利用更大区域的上下文信息;增大重组内核 k_{up} ,可以充分利用特征图的语义信息。实验发现保证 $k_{encoder} = k_{up} - 2$ 可以在性能和效率之间获得一个更好的平衡,因此取 $k_{encoder} = 3, k_{up} = 5$ 。

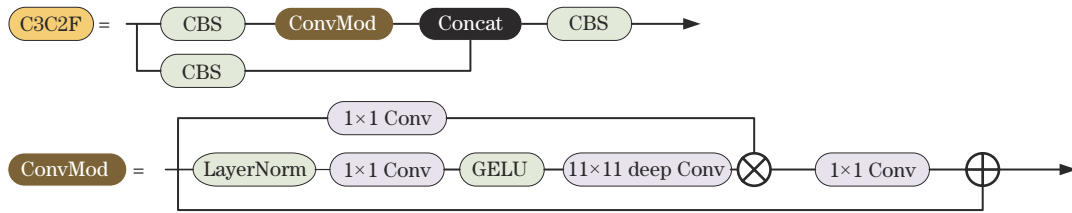


图 4 C3C2F 模块结构

Fig. 4 Structure of the C3C2F module

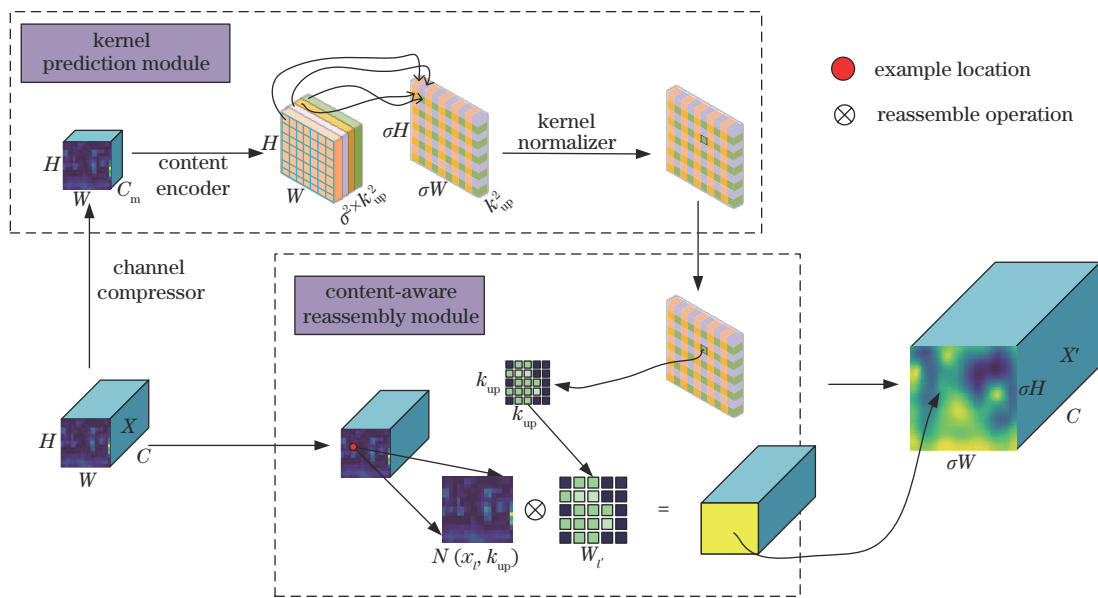


图 5 CARAFE 模块结构

Fig. 5 Structure of the CARAFE module

KCC-YOLOv5的浅层网络更关注细节信息以定位目标,深层网络更关注语义信息以推理目标。将CARAFE运用于特征金字塔FPN^[19],可以加强深层特征信息,从而使融合后的特征具有更丰富的表达能力。

3 实验分析

3.1 数据集。

铝型材数据集来自2018广东工业制造大数据创新大赛——铝型材表面瑕疵识别。铝型材表面有7种类型的缺陷:不导电、桔皮、起坑、擦花、喷流、漆泡和脏点,如图6所示,图片尺寸为2560 pixel×1920 pixel。由于原始数据样本不均衡,因此采用水平翻转、竖直翻转、平移等方法的随机组合来执行数据增强,数据增强结果如表1所示。为了加快模型训练,实验中图像被自适应地缩放至640 pixel×640 pixel。图7为数据集信息的可视化结果,图7(a)为关于数据集的类型和标签数量的信息,图7(b)为缺陷中心点的位置分布信息,图7(c)为缺陷大小分布信息。由图7可见,该数据集中缺陷的位置分布均匀,缺陷尺度变化大,小目标缺陷相对较多。

表1 数据增强结果

Table 1 Results of data augmentation

Class	Original image /frames	Enhanced image /frames
Non-conduction	390	500
Scratch	128	500
Orange Peel	173	500
Jet	86	500
Crater	407	500
Bubble	82	500
Spot	261	500
Total	1527	3500

3.2 评价指标和实验设置

实验使用Windows 10操作系统,中央处理器(CPU)为AMD EPYC 7543,图形处理器(GPU)为NVIDIA GeForce RTX3090,显存24 G,基于PyTorch深度学习框架构建网络模型,开发环境为PyTorch1.11.0、CUDA11.5、Python3.8。

通过指标mAP@0.5、FPS、GFLOPs和参数量评估性能,其中:mAP@0.5表示总类别的平均精度;FPS为每秒处理帧数,用于评估模型执行速度;GFLOPs为浮点计算量,用于衡量模型复杂度。

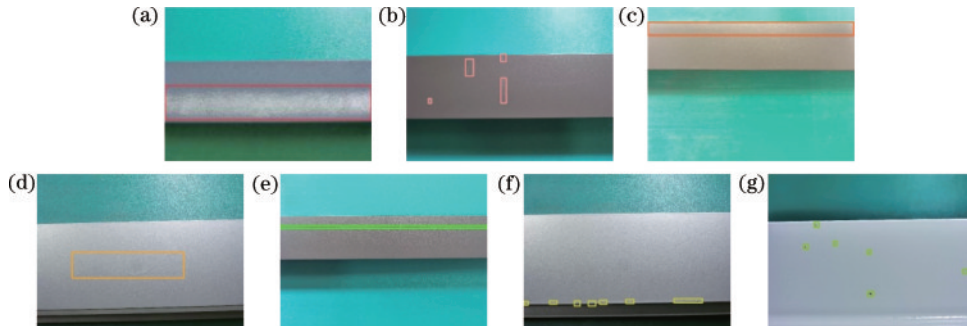


图6 铝型材表面缺陷图片。(a)不导电;(b)擦花;(c)桔皮;(d)喷流;(e)起坑;(f)漆泡;(g)脏点

Fig. 6 Pictures of surface defect of aluminum profile. (a) Non-conduction; (b) scratch; (c) orange peel; (d) jet; (e) crater; (f) bubble; (g) spot

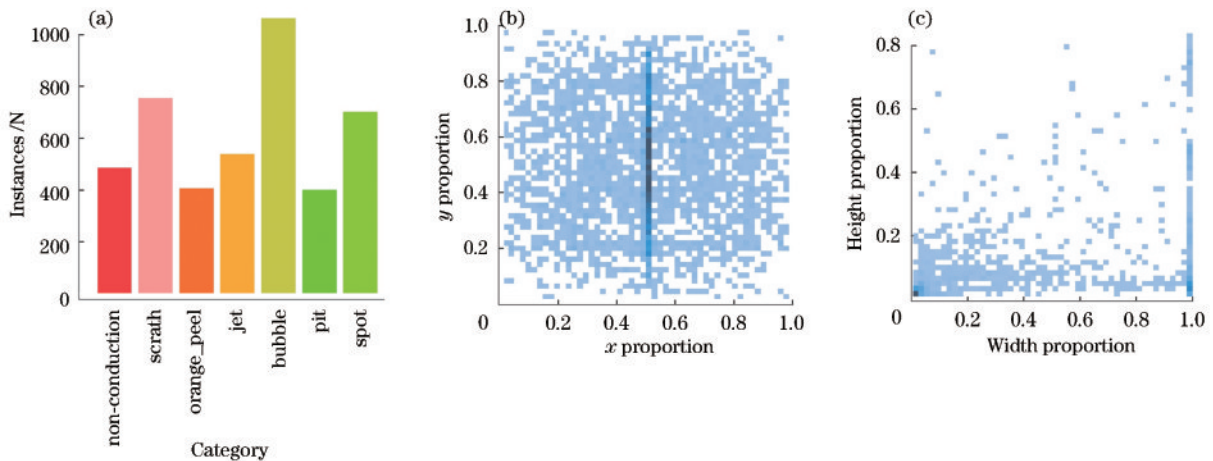


图7 铝型材数据集信息可视化结果。(a)种类和数量;(b)中心点分布;(c)缺陷大小分布

Fig. 7 Visualization results of aluminum profile dataset information. (a) Category and quantity; (b) central point coordinate distribution; (c) defect size distribution

3.3 KCC-YOLOv5 模型结果分析对比

基于损失函数和均值平均精度曲线对原始模型和改进模型进行评估,评估结果如图 8 所示。图 8(a)显示 KCC-YOLOv5 模型的损失函数收敛更快,并且最

终的损失函数值更小。这表明 KCC-YOLOv5 模型的预测结果与真实结果之间的误差更小,具有更强的泛化能力。图 8(b)显示 KCC-YOLOv5 模型相较于原始模型具有更高的检测精度。

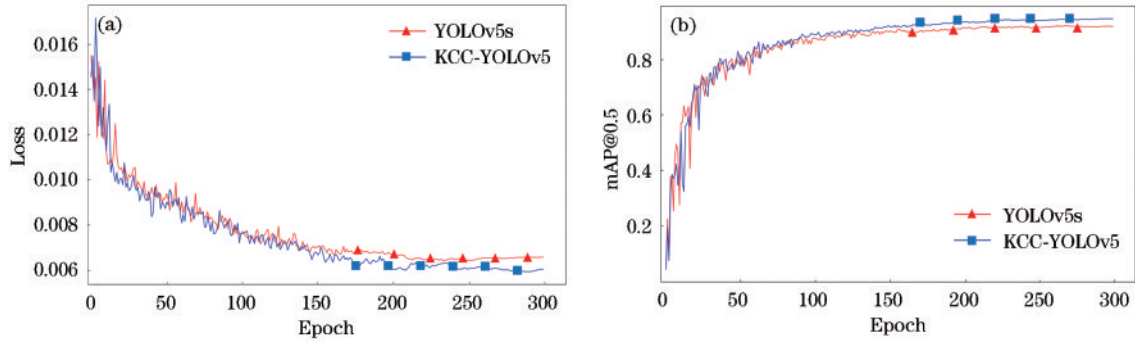


图 8 KCC-YOLOv5 模型和 YOLOv5s 模型的损失函数和均值平均精度曲线对比。(a) 损失函数的变化曲线; (b) 均值平均精度的变化曲线

Fig. 8 Comparison of loss and mAP@0.5 curves between KCC-YOLOv5 model and YOLOv5s model. (a) Curves of the loss function; (b) curves of the mAP@0.5

图 9 为原始模型和 KCC-YOLOv5 模型的 PR (precision-recall) 曲线图,曲线图的面积代表缺陷的检测精度,发现不导电、桔皮、起坑、擦花、喷流、漆泡和脏点的平均精度分别提高了 3.4, 0.1, 0.2, 4, 0.1, 5.2,

12.4 个百分点。其中,小目标缺陷漆泡和脏点的检测精度提升显著,验证了改进模型对铝型材表面小目标缺陷的有效性。

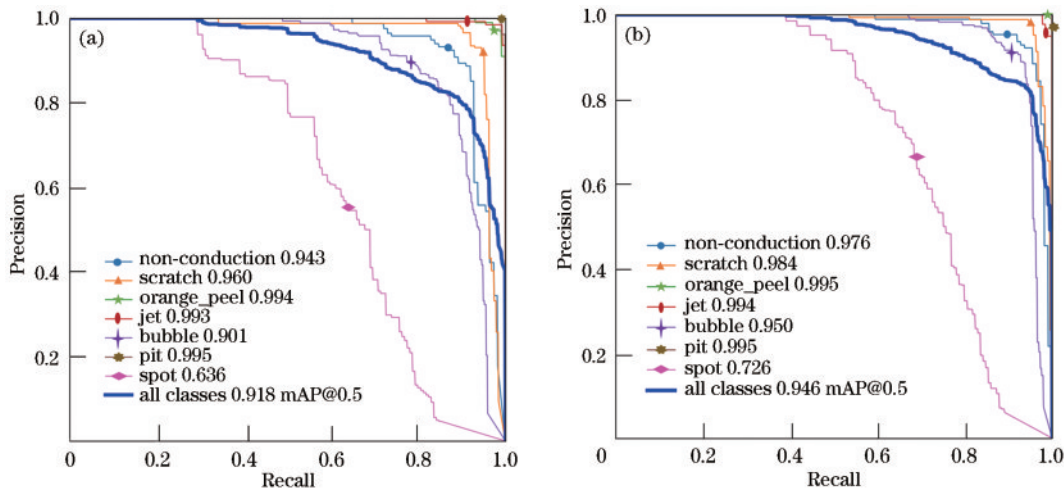


图 9 PR 曲线对比。(a) YOLOv5s; (b) KCC-YOLOv5

Fig. 9 Comparison of PR curves. (a) YOLOv5s; (b) KCC-YOLOv5

3.4 消融实验

为了验证不同的改进部分对原始 YOLOv5s 算法性能的影响,进行了消融实验,结果如表 2 所示。原始 YOLOv5s 的 mAP 为 91.8%;通过 IoU-K-means++

算法聚类优化锚框, mAP 提高了 1.3 个百分点,且 FPS 小幅度提升;在主干部分引入 C3C2F 模块增强铝型材表面缺陷的语义信息,抑制无关信息的干扰,降低了模型的参数量和模型复杂度, mAP 获得了 0.9 个百

表 2 不同改进部分对原始 YOLOv5 算法的影响

Table 2 Impact of different improvement parts on the original YOLOv5 algorithm

IoU-K-means++	C3C2F	CARAFE	Params. /M	GFLOPs	mAP@0.5 /%	FPS / (frame·s ⁻¹)
			7.03	15.8	91.8	137.0
✓			7.03	15.8	93.1	138.9
✓	✓		6.60	15.5	94.0	131.6
✓	✓	✓	6.79	16.1	94.6	113.6

分点的提升;通过替换颈部的上采样方式进一步提升了模型对铝型材表面缺陷的识别能力,mAP提升了0.6个百分点。KCC-YOLOv5模型最终的mAP值为94.6%,比原始YOLOv5s模型的mAP提升了2.8个百分点,FPS为113.6 frame/s,满足实时检测要求,充分证明了改进的有效性。

对锚框的聚类方法进行消融实验,实验评价指标为BRP(best possible recall)和Fitness。其中:BRP表示锚框宽高与真实框宽高的符合程度;Fitness表示锚框与真实框的重合程度。实验结果如表3所示,表3中SSE(sum of square error)为欧氏距离,可以看出IoU-K-means++聚类方法生成的锚框宽高更符合铝型材数据集表面缺陷的宽高,并且锚框与真实框的重合度更高。

针对小目标缺陷漆泡和脏点,将包括C3C2F模块和不包括C3C2F模块两种情况下的热力图可视化,结果如图10所示。由图10可见,通过添加C3C2F模块使模型的铝型材表面两种小目标缺陷语义信息得到了

表3 不同聚类算法生成的锚框与铝型材数据集真实框架的符合度

Table 3 Degree of conformity of anchor frames generated by different clustering algorithms to the true frames of the aluminum profile dataset

K-means	K-means++	SSE	IoU	BRP / %	Fitness / %
✓		✓		97.8	68.4
	✓	✓		98.8	68.7
✓			✓	99.4	70.8
	✓		✓	99.5	71.2

增强,并抑制了背景和噪声。为了更清晰地展示CARAFE模块对铝型材表面小目标缺陷的影响,将输入图像为漆泡和脏点时,第11层和第15层的原始上采样模块和CARAFE上采样模块输出的特征图可视化,如图11所示,发现在CARAFE上采样后的特征图中小目标特征更明显,位置信息更清晰,检测效果得到有效提升。

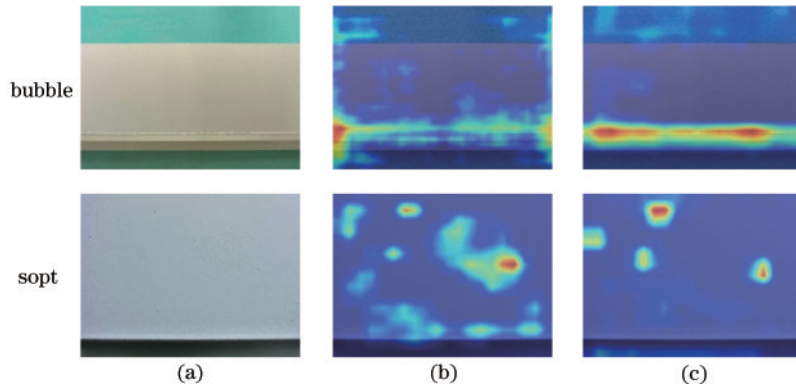


图10 热力图可视化结果。(a)输入图像;(b)不包括C3C2F;(c)包括C3C2F

Fig. 10 Visualization results of the heat maps. (a) Input images; (b) without C3C2F; (c) with C3C2F

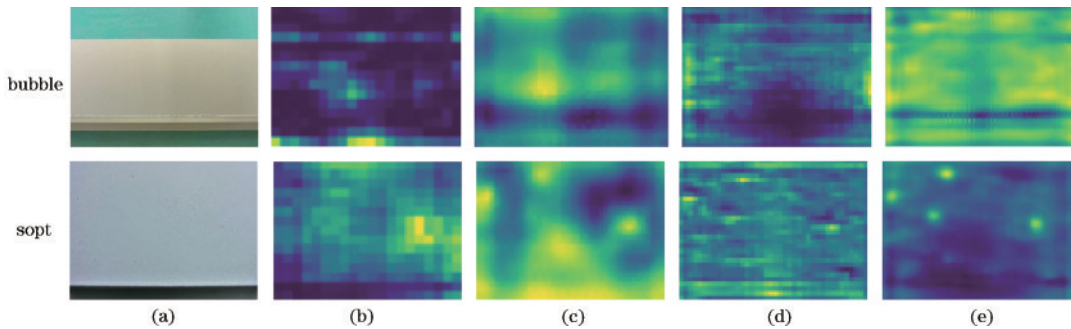


图11 特征图可视化结果。(a)输入图像;(b)原始上采样模块(stage 11)输出的特征图;(c) CARAFE上采样模块(stage 11)输出的特征图;(d)原始上采样模块(stage 15)输出的特征图;(e) CARAFE上采样模块(stage 15)输出的特征图

Fig. 11 Visualization results of the feature maps. (a) Input images; (b) feature maps of the original upsampling (stage 11) module; (c) feature maps of the CARAFE upsampling (stage 11) module; (d) feature maps of the original upsampling (stage 15) module; (e) feature maps of the CARAFE upsampling (stage 15) module

最后,为了进一步证明改进模型KCC-YOLOv5对铝型材数据集中小目标缺陷的有效性,对原始模型和KCC-YOLOv5模型进行了可视化比较。由图12可见,KCC-YOLOv5模型有效降低了原始YOLOv5s模型对小目标缺陷的漏检问题,并且提高

了置信度。

3.5 与其他算法模型的性能比较

为了验证改进模型KCC-YOLOv5的性能,将其与主流目标检测模型Faster R-CNN、YOLOv3、YOLOv4、YOLOv7和文献[1-3]中的改进方法进行对

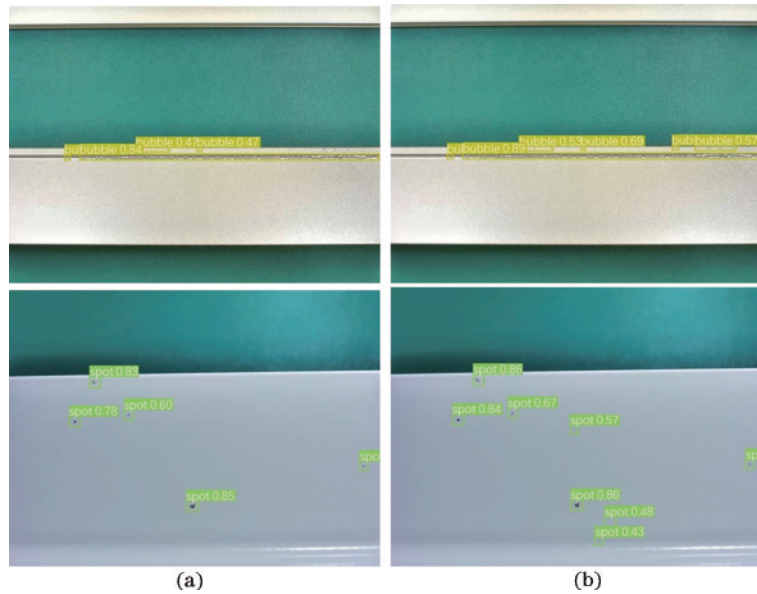


图 12 原始模型和 KCC-YOLOv5 模型对小目标缺陷检测的结果对比。(a)YOLOv5s;(b)KCC-YOLOv5

Fig.12 Comparison of defect detection results for small targets between the original model and the KCC-YOLOv5 model. (a) YOLOv5s; (b) KCC-YOLOv5

比。在划分相同的训练集和验证集的铝型材数据集上进行实验,结果如表 4 所示,由于铝型材数据集上有大量的小目标缺陷漆泡和脏点,大多数主流模型对漆泡和脏点的检测效果并不理想,而 KCC-YOLOv5 具有

更好的检测效果,并且在大目标缺陷检测中也具有良好的表现。此外,在所提小型铝型材数据集上,KCC-YOLOv5 相对于文献[1-3]中的 3 种方法也展现出更好的性能。

表 4 所提模型与其他算法模型的对比结果

Table 4 Comparison results between the proposed model and other models

Model	Average precision / %							Param. /M	GFLOPs	mAP@ 0.5 / %	FPS / (frame·s ⁻¹)
	Non-conduction	Scratch	Orange Peel	Jet	Pit	Bubble	Spot				
FasterR-CNN	90.5	95.9	96.3	93.7	89.2	79.7	60.6	41.2	91.3	86.6	16.0
FasterR-CNN(FPN) ^[1]	96.0	98.1	96.5	98.8	97.2	85.4	63.3	46.0	118.1	90.8	9.3
YOLOv3	93.2	91.8	98.3	98.3	97.0	88.6	57.1	61.5	154.6	89.2	45.3
MSFF-YOLOv3 ^[2]	93.6	97.1	99.5	99.0	99.5	91.1	61.9	62.8	175.9	91.7	41.7
YOLOv4	93.4	95.7	98.4	98.8	99.5	91.2	63.7	52.5	119.8	91.5	54.7
YOLOv5s	94.3	96.0	99.4	99.3	99.5	90.1	63.6	7.0	15.8	91.8	137.0
MS-YOLOv5 ^[3]	93.7	97.1	99.5	99.5	99.5	90.0	67.9	8.1	17.9	92.5	105.6
YOLOv7	95.7	97.1	99.5	99.5	99.5	90.5	65.4	37.2	105.2	92.4	73.5
KCC-YOLOv5	97.6	98.4	99.5	99.4	99.5	95.0	72.6	6.8	16.1	94.6	113.6

4 总 结

本文提出了一种用于铝型材表面缺陷检测的改进算法模型 KCC-YOLOv5,以解决传统铝型材表面缺陷检测精度低、检测速度慢和小目标缺陷检测能力不足等问题。利用 IoU-K-means++ 算法聚类优化锚框;用 C3C2F 模块替换 C3 模块,突出小目标的特征信息;用轻量级上采样算子 CARAFE 替换最近邻插值上采样方式,有效保留小目标特征信息。实验结果表明,改进模型检测效果提升明显,有效改善了铝型材表面小目标缺陷的漏检问题。

参 考 文 献

- [1] Wei R F, Bi Y B. Research on recognition technology of aluminum profile surface defects based on deep learning [J]. Materials, 2019, 12(10): 1681.
- [2] Sun L S, Wei J X, Du H C, et al. MSFF: a multi-scale feature fusion network for surface defect detection of aluminum profiles[J]. IEICE Transactions on Information and Systems, 2022, E105.D(9): 1652-1655.
- [3] Wang T, Su J H, Xu C A, et al. An intelligent method for detecting surface defects in aluminium profiles based on the improved YOLOv5 algorithm[J]. Electronics,

- 2022, 11(15): 2304.
- [4] 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.
- [5] Girshick R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2016: 1440-1448.
- [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): 1137-1149.
- [7] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2023-03-25]. <https://arxiv.org/abs/1804.02767>.
- [8] 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.
- [9] Chen B Y, Liu Y T, Sun K. Research on object detection method based on FF-YOLO for complex scenes [J]. IEEE Access, 2021, 9: 127950-127960.
- [10] 徐志京, 柏雪. 基于双重特征增强的遥感舰船小目标检测[J]. 光学学报, 2022, 42(18): 1828002.
Xu Z J, Bai X. Small ship target detection method for remote sensing images based on dual feature enhancement [J]. Acta Optica Sinica, 2022, 42(18): 1828002.
- [11] 张寅, 朱桂熠, 施天俊, 等. 基于特征融合与注意力的遥感图像小目标检测 [J]. 光学学报, 2022, 42(24): 2415001.
Zhang Y, Zhu G Y, Shi T J, et al. Small object detection in remote sensing images based on feature fusion and attention[J]. Acta Optica Sinica, 2022, 42(24): 2415001.
- [12] 程松, 杨洪刚, 徐学谦, 等. 基于YOLOv5的改进轻量型X射线铝合金焊缝缺陷检测算法[J]. 中国激光, 2022, 49(21): 2104005.
- Cheng S, Yang H G, Xu X Q, et al. Improved lightweight X-ray aluminum alloy weld defects detection algorithm based on YOLOv5[J]. Chinese Journal of Lasers, 2022, 49(21): 2104005.
- [13] Zhu X K, Lyu S C, Wang X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), October 11-17, 2021, Montreal, BC, Canada. New York: IEEE Press, 2021: 2778-2788.
- [14] Wang J Q, Chen K, Xu R, et al. CARAFE: content-aware ReAssembly of FEatures[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2020: 3007-3016.
- [15] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [16] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at scale[EB/OL]. (2020-10-22) [2023-03-25]. <https://arxiv.org/abs/2010.11929>.
- [17] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Advances in Neural Information Processing Systems, December 4-9, 2017, Long Beach, California, USA. Massachusetts: MIT Press, 2017: 5998-6008.
- [18] Hou Q B, Lu C Z, Cheng M M, et al. Conv2Former: a simple transformer-style ConvNet for visual recognition [EB/OL]. (2022-11-22)[2023-03-25]. <https://arxiv.org/abs/2211.11943>.
- [19] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.