

DOI:10.19651/j.cnki.emt.2518664

基于深度学习的皮革缺陷检测方法^{*}

白先浪¹ 张群利¹ 辛志强²

(1.东北林业大学家居与艺术设计学院 哈尔滨 150040;2.东北林业大学机电工程学院 哈尔滨 150040)

摘要: 针对皮革表面缺陷纹理复杂多样,现有的缺陷检测方法存在检测精度低、漏检率和误检率高的问题,本文提出了一种基于YOLOv5s且融合了小目标检测方法与注意力机制的皮革缺陷检测算法。首先,在模型主干网络中引入多种注意力机制,使模型能够聚焦于缺陷区域,抑制背景和无关区域的干扰,增强模型的特征提取能力;其次,在颈部网络中构建加权双向特征金字塔网络,强化模型的特征融合与交互能力;最后,在头部网络中设计专门的小目标检测头,提升模型对细微缺陷特征的捕捉和定位能力。实验结果表明,所提改进方法的召回率和检测精度达到92.27%和92.16%,相较于基准模型分别提高了4.56%和3.06%。所提改进算法解决了以往缺陷检测算法在小目标检测任务中漏检率和误检率高的问题,显著增强了模型的泛化能力,使模型性能更加全面、稳健。

关键词: 皮革缺陷检测;深度学习;YOLOv5s;小目标增强;加权特征融合;注意力机制

中图分类号: TP391; TN973.1 **文献标识码:** A **国家标准学科分类代码:** 520

Leather defect detection method based on deep learning

Bai Xianlang¹ Zhang Qunli¹ Xin Zhiqiang²

(1. School of Home and Art Design, Northeast Forestry University, Harbin 150040, China;

2. School of Mechanical Engineering, Northeast Forestry University, Harbin 150040, China)

Abstract: Due to the complex and diverse textures of leather surface defects, existing detection methods often suffer from limited accuracy and elevated rates of missed and false detections. To address these challenges, this paper presents an enhanced defect detection algorithm based on YOLOv5s, incorporating small-object detection techniques and attention mechanisms. Specifically, multiple attention modules are integrated into the backbone network to guide the model's focus toward defect regions while suppressing interference from background and irrelevant features, thereby enhancing feature extraction. A weighted bidirectional feature pyramid network is introduced in the neck to strengthen feature fusion and interaction across scales. Additionally, a dedicated detection head tailored for small objects is implemented in the head network to improve the localization and recognition of subtle defect features. Experimental results show that the proposed improved method achieves a recall of 92.27% and a detection accuracy of 92.16%, representing improvements of 4.56% and 3.06%, respectively, compared to the baseline model. These enhancements effectively reduce missed and false detections in small-object scenarios and significantly improve the model's generalization capability, contributing to more robust and comprehensive performance in real-world applications.

Keywords: leather defect detection; deep learning; YOLOv5s; small object enhancement; weighted feature fusion; attention mechanism

0 引言

皮革作为一种天然材料,因其独特的质感、美观,被广泛应用于奢侈品、珠宝、手表等高附加值产品的外包装中。由于皮革在生产和运输过程中,皮革制品容易出现划伤、褶

皱、污渍等,如果让这些有缺陷的皮革包装产品流入市场,会降低产品的质量和品牌价值。因此,构建一套基于深度学习的皮革缺陷检测方法具有重大现实意义。

传统的缺陷检测方法有人工检测,这种方式对人的主观经验要求很高而且还存在漏检和误检的风险,还有一些

收稿日期:2025-04-23

^{*} 基金项目:中央高校基本科研业务费专项资金(2572019BL01)项目资助

基于数字图像处理、机器学习等传统缺陷检测方法也存在检测精度低的问题。例如张嘉宁等^[1]针对碳纤维材料在检测时存在边缘模糊等问题,提出一种基于时序信息的自适应中值滤波算法,虽然信噪比指数平均高出 6.155 dB,但速度和精度达不到工业检测要求。

随着科学技术的发展,深度学习等先进方法在工业检测领域得到了广泛的应用。例如刘庭浔等^[2]针对传统检测模型检测仓储物体精度较低的问题,提出一种改进 YOLOv5 的仓储物体检测模型,采用 ConvNeXt 网络替换原 YOLOv5 的主干网络,检测精度达到了 89.8%,但达不到精度要求高的任务场景。赵松怀等^[3]针对目前主流烟火检测算法检测效率不高等问题,提出一种改进 YOLOv5s 的烟火轻型检测算法,引入 C3Ghost 模块和 Ghost 卷积模块,以减少网络参数数量,相较于原 YOLOv5s,检测精度提升了 1.1%,同时也实现了模型轻量化。欧晓放等^[4]引入 Transformer 结构^[5]和增加小目标检测头,在自建数据集上,改进算法检测精度达到了 89.58%,但无法满足一些精度要求特别高的任务场景,还有很大提升空间。张振^[6]受 Inception 网络^[7]和 ResNet 网络^[8]启发,通过改造 Inception-ResNetV1 中的模块,使得平均交并比和缺陷类别的交并比均达到 0.8 以上,但有一部分面积小的缺陷没有检测出来。Chen 等^[9]提出一种基于 Cycle-GAN^[10]和 YOLOv5 的石化管道缺陷检测方法,在测试集上的平均准确率和召回率分别为 93.1%和 90.96%,但仍有较多漏检发生。He 等^[11]针对绝缘子表面缺陷检测问题,提出了一种基于自适应特征融合的 YOLOv5s 改进模型,利用自适应特征融合 ASFF 模块,使得网络能够学习不同特征图之间关联性,增强语义信息传递,将主干网络替换为 EfficientNet,使网络更加高效,虽然模型大小降低了,但对微小缺陷感知能力不足,检测效果较差。Tang 等^[12]针对织物缺陷检测问题,提出了一种基于 YOLOv7-tiny 的轻量级织物缺陷检测算法,通过采用 Mish 激活函数、集成了 Ghost 卷积模块、采用轻量级上采样技术 CARAFE,虽然实现了模型轻量化,但改进模型并不适用精度要求高的检测任务。Zhang 等^[13]针对金属传感器缺陷存在误检率高和漏检率高的问题,提出了一种基于 YOLOv5s 的改进模型,通过利用 C2f 模块替换 C3 模块、引入新的改进注意力机制和损失函数,虽然精度提高了,但在实际检测时仍会有漏检发生。Xu 等^[14]针对服装缝纫缺陷检测任务,对 YOLOv8 进行改进,通过引入 Focal Loss 损失函数和 CA 注意力机制来提高模型检测框回归拟合精度与多尺度特征提取及融合能力,但在面对复杂纹理缺陷时,容易造成误检。Xie 等^[15]针对印刷电路板缺陷难以识别问题,提出一种轻量级 YOLOv5 改进模型,通过将 YOLOv5 的头部网络改为双头检测、在颈部网络引入 GSConv、还引入了 BiFPN 融合多尺度信息,虽然大大减少了计算量和参数量,但在面对小目标缺陷时,检测效果不佳。

基于此,针对目前的缺陷检测方法并不能很好的对物体表面缺陷区域进行精准的定位与识别,因此,本文选取目前性能最稳定、应用最广泛、同时兼具检测精度与速度的 YOLOv5s 作为基准模型并对其改进。具体的是在基准模型中引入注意力机制、重构颈部网络结构、设计专门的小目标检测头,理由分析如下:

1)引入注意力机制的优势:皮革表面缺陷纹理、形状复杂多变,通常表现为纹理细微变化、颜色局部差异和形状不规则、特征不明显等特点,且缺陷区域往往面积较小易被复杂纹理干扰,传统卷积操作在全局特征提取过程中容易忽略这些显著性较弱的缺陷目标。本文通过引入注意力机制,旨在增强模型对局部关键区域的感知能力,使得模型能够在通道维度与空间维度上自适应调整特征权重,从而提升对缺陷区域的关注力度,有助于模型在特征提取过程中自动聚焦图像中的缺陷区域,从而突出有效特征、抑制冗余信息,提高特征表达的判别性,进而提高模型缺陷识别的准确性与鲁棒性。

2)重构颈部网络结构的必要性:颈部结构在特征融合中起关键作用,原 YOLOv5 的特征金字塔网络和路径聚合网络在特征融合上存在一定的局限性,对不同尺度缺陷的特征融合不充分,在处理多尺度特征信息尤其是极小目标特征信息时存在上下文信息传递不充分的问题。为此,引入更高效的加权双向特征金字塔网络以强化低层高分辨率特征与高层语义特征之间的融合,使得模型对对复杂纹理和不规则形状的缺陷更加敏感,提高模型在复杂背景下对不同尺寸缺陷的感知与检测能力。

3)设计小目标检测头的有效性:皮革表面缺陷区域面积往往较小,且缺陷普遍为细小划伤、褶皱等小目标,尤其在高分辨率图像中占比很低,其纹理信息主要分布在浅层特征图中,这些纹理和细节信息在多次下采样过程中容易丢失和被淹没。常规检测头对这类小目标缺陷检测能力不足,容易导致漏检率上升。为提升模型对微小缺陷目标的检测能力,本文在原始检测架构基础上新增一个小目标检测头,设计的小目标检测头可以直接作用于高分辨率浅层特征图,能够更充分利用浅层特征信息在特征图上进行检测,在保留更多空间细节信息的同时实现小目标缺陷的高精度定位,显著缓解了传统多尺度检测框架在下采样过程中引起的小目标信息丢失问题,有效提升了模型对小尺寸缺陷的检测能力。

综上所述,上述 3 种不同改进策略协同优化提升了模型对皮革表面缺陷的检测性能,特别是在复杂纹理背景与不规则目标形态下显著提升模型的鲁棒性与检测精度。

1 基准模型

目前常见的基于深度学习的缺陷检测有两个主流方向,一个是以 YOLO 和 SSD^[16]为代表的单阶段目标检测技术,另一个是以 R-CNN、Fast R-CNN、Faster R-CNN、

Mask R-CNN 这 4 种模型^[17-20]为代表的两阶段目标检测技术。随着科学技术的不断进步,YOLO 模型已经迭代更新多种版本,其中 YOLOv5 是最经典的,在工业检测、自动驾驶、安全安防监控、医疗辅助分析等领域应用的最为广泛。根据模型的大小,YOLOv5 模型从小到大又可细分为 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 这 5 种不同的版本,这些版本在结构上相同,不同点在于网络模型的深度和宽度,图 1 为 YOLOv5 架构。

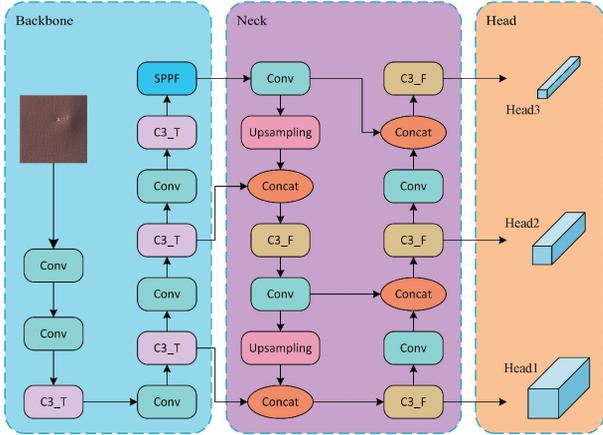


图 1 YOLOv5 架构

Fig. 1 YOLOv5 architecture

在 YOLOv5 架构中,主干网络由 5 个 Conv 模块、4 个 C3_T 模块、1 个空间金字塔池化快速模块(spatial pyramid pooling-fast, SPPF)组成。Conv 模块采用标准卷积、批归一化、激活函数,用来提取图像局部特征。C3_T 模块采用跨阶段局部连接能够减少计算冗余并且增强梯度流动。SPPF 是在空间金字塔池化网络(spatial pyramid pooling, SPP)^[21]基础上改进而来,具有多尺度特征融合功能,能够适应不同尺寸目标,加快网络推理速度。颈部网络是由 4 个 Conv 模块、4 个 C3_F 模块、2 个上采样模块组成。采用特征金字塔网络(feature pyramid network, FPN)^[22]将网络深层特征图中的高级语义信息传递到网络浅层特征图,利用路径聚合网络(path aggregation network, PANet)^[23]将网络浅层特征图中的位置细节信息传递到网络深层特征图。头部网络则是由 3 个检测头组成,其中大型特征图擅长检测小目标物体,中型特征图负责检测中型目标,小型特征图擅长检测大目标物体。

2 改进方法

2.1 主干网络改进

由于基准模型在检测小目标缺陷时,预测框有时会出现定位偏差和误检等问题,因此可以通过在模型的主干网络中引入注意力机制,能够根据数据的重要程度不同,动态的赋予其不同的权重,从而使得模型能够更加关注数据的重要特征部位。本文通过在主干网络中引入压缩与激励模

块(squeeze-and-excitation, SE)、高效通道注意力机制(efficient channel attention, ECA)、坐标注意力机制(coordinate attention, CA)、卷积块注意力模块(convolutional block attention module, CBAM)注意力机制增强模型特征提取能力、上下文感知能力,以及对缺陷区域的关注度,提升了模型在复杂场景下的检测精度。

SE 注意力机制^[24]如图 2 所示,能够自主学习特征图各个通道的重要性并且能够动态赋予每个通道不同的权重。

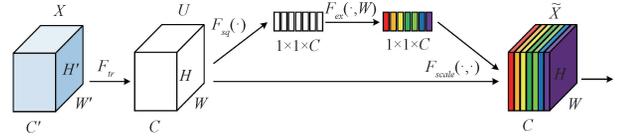


图 2 SE 结构

Fig. 2 SE structure

SE 注意力机制首先采用全局平均池化操作将特征图的每个通道按空间维度的方向进行压缩。然后,通过两个全连接层生成注意力权重。最后,将得到的注意力权重去乘以原始特征图的每个通道,得到带权值的特征图。原理计算如式(1)~(3)所示。

$$\mathbf{z}_c = F_{sq}(\mathbf{u}_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W \mathbf{u}_c(i, j) \quad (1)$$

$$\mathbf{s} = F_{ex}(\mathbf{Z}, \mathbf{W}) = \sigma(\mathbf{W}_2 \delta(\mathbf{W}_1 \mathbf{Z})) \quad (2)$$

$$\tilde{\mathbf{x}}_c = F_{scale}(\mathbf{u}_c, \mathbf{s}_c) = \mathbf{s}_c \cdot \mathbf{u}_c \quad (3)$$

式中: $\mathbf{u}_c(i, j)$ 为特征图 \mathbf{u} 的第 c 个通道位于 (i, j) 处的元素值; \mathbf{z}_c 为特征图 \mathbf{z} 第 c 个元素值; δ 为 ReLU 激活函数; σ 为 Sigmoid 激活函数; $\mathbf{W}_1, \mathbf{W}_2$ 分别对应的是第 1 个全连接层和第 2 个全连接层变换; \mathbf{s}_c 为第 c 个通道的注意力权重; $\tilde{\mathbf{x}}_c$ 为最终的输出特征图。

ECA 注意力机制^[25]如图 3 所示,是 SE 注意力机制的一种改进版,属于通道注意力机制。ECA 相比于 SE,其最大的优点就是用一维卷积替换了 SE 模块中的两个全连接层,这样不仅能够实现跨通道信息共享交融,还能极大地减少模型的参数量,有效提高模型的检测速度。

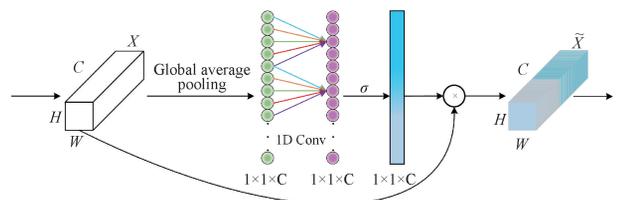


图 3 ECA 结构

Fig. 3 ECA structure

CA 注意力机制^[26]如图 4 所示,其核心思想是通过引入坐标信息,使模型能够更好地理解特征图不同空间位置信息,不仅如此,还考虑了长距离依赖关系。

CA 注意力机制首先对输入特征图按 X 方向和 Y 方向

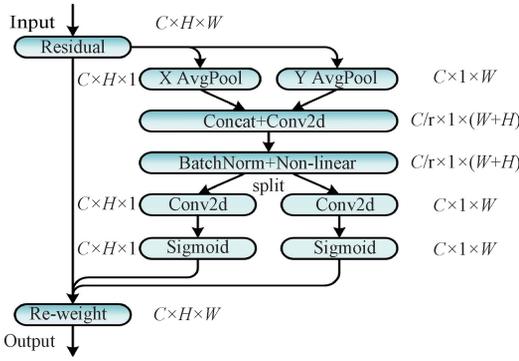


图 4 CA 结构
Fig. 4 CA structure

上进行平均池化，得到两个特征向量。然后通过拼接、卷积、批归一化、非线性化得到注意力权重。接着将注意力权重按空间维度进行拆分，再用卷积提取特征和非线性化得到两个方向上注意力权重特征向量。最后将得到的两个注意力权重向量乘以原始特征图进行特征重标定，原理计算如式(4)~(9)所示。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (4)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (5)$$

$$f = \text{ReLU}(\text{Batchsize}(\text{Conv2d}(\text{Concat}[z^h; z^w]))) \quad (6)$$

$$g^h = \text{Sigmoid}(\text{Conv2d}(f^h)) \quad (7)$$

$$g^w = \text{Sigmoid}(\text{Conv2d}(f^w)) \quad (8)$$

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (9)$$

式中： $x_c(h, i)$ 为输入特征图 x 在第 c 通道上位于 (h, i) 处的元素值； $z_c^h(h), z_c^w(w)$ 为 X、Y 方向上平均池化得到的特征向量； f 是 $z_c^h(h), z_c^w(w)$ 特征向量经过拼接、卷积、批归一化、非线性化后得到的新的特征向量； f^h, f^w 为拆分后的特征向量； g^h, g^w 为不同方向上的注意力权重； y_c 为经过特征重新标定的输出特征图。

CBAM 注意力机制^[27]如图 5 所示，其核心思想是同时考虑通道注意力和空间注意力，使得模型既能关注到重要的通道信息，也能关注到重要的空间位置信息。

CBAM 注意力机制原理是先将基于通道注意力机制得到的通道注意力权重用来对原始输入特征图进行通道加权表示得到一张新的特征图，接着再用空间注意力机制得到的空间注意力权重对这张新的特征图进行空间位置加权表示得到最终的输出特征图，原理计算方法如式(10)和式(11)所示。

$$F' = M_c(F) \otimes F \quad (10)$$

$$F'' = M_s(F') \otimes F' \quad (11)$$

式中： F 为输入特征图； M_c 为通道注意力权重； F' 为经过通道注意力机制作用后输出的特征图； M_s 为空间注意力权重； F'' 为最终的输出特征图； \otimes 表示元素逐点相乘。

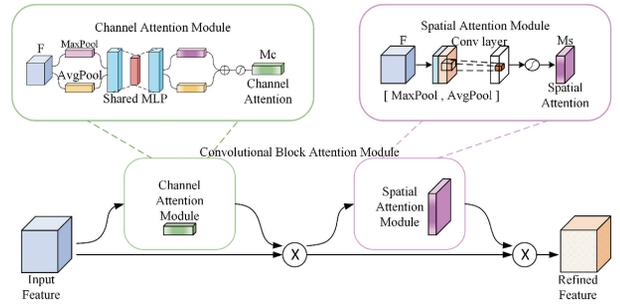


图 5 CBAM 结构
Fig. 5 CBAM structure

鉴于皮革缺陷通常具有小尺寸、低对比度等特点，模型对关键缺陷区域的感知能力不强。为了增强模型在特征提取阶段对缺陷区域的关注能力，本文考虑在 YOLOv5s 主干网络中引入注意力机制以提升模型特征表达能力。具体而言，考虑到不同注意力机制在通道建模、空间建模及定位表达能力方面存在差异，本文在实验结果与分析章节设计了对比实验，选取了 4 种典型的注意力机制，即 SE 注意力机制、ECA 注意力机制、CA 注意力机制和 CBAM 注意力机制，分别嵌入模型主干网络中的 C3 模块内，并对上述 4 种注意力机制进行横向对比，以量化评估不同注意力机制在皮革缺陷检测任务中的性能差异。对比实验结果表明，CA 注意力机制在检测精度、召回率等关键指标上表现优异，尤其在小目标缺陷检测任务中具有更强的局部感知能力与空间、通道特征建模能力。因此，本文最终采用 CA 注意力机制作为改进方法。

2.2 颈部网络改进

原始的 YOLOv5 颈部网络采用 FPN 和 PANet，这种传统的结构存在冗余连接、网络计算量大、特征交互不充分等问题，为此本文对颈部网络进行重构，设计了基于加权双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)^[28]，如图 6 所示。BiFPN 采用双向跨尺度连接，能够减少计算量、加强跨尺度特征交互。同时 BiFPN 采用加权特征融合方式，使模型能够关注到更加有价值的特征，提升融合效果。

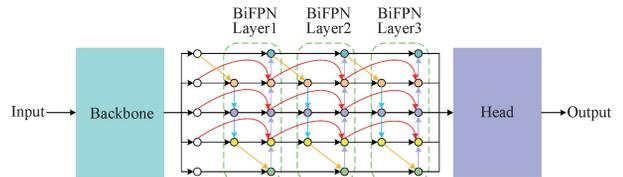


图 6 改进后的颈部网络
Fig. 6 Improved neck network

2.3 头部网络改进

大型特征图通常位于模型的浅层，由于其尺寸大、分辨率高、含有丰富的位置信息，可以精确地描述小目标物体的位置和形状，更有利于捕捉小目标物体的细节，因此大型特征图更擅长检测小目标物体。鉴于上述原理，本文通过设

表 2 不同注意力机制对比

Table 2 Comparison of different attention mechanisms

%

SE	ECA	CA	CBAM	P	R	mAP@0.5	mAP@0.5:0.95
×	×	×	×	87.61	87.71	89.10	82.40
√	×	×	×	85.54	86.26	88.61	81.29
×	√	×	×	86.22	86.18	88.07	81.17
×	×	√	×	89.62	90.58	91.67	84.78
×	×	×	√	89.19	89.90	90.89	83.91

可视化,如图 8 所示。图 8(a)~(e)分别是基准模型和引入 SE 注意力机制、ECA 注意力机制、CBAM 注意力机制、CA 注意力机制的梯度加权类激活图。从图 8(a)~(c)可以看出基准模型和引入 SE 注意力机制、ECA 注意力机制的改进模型激活区域较为分散,且部分缺陷区域未能被有

效关注。从图 8(d)和(e)可以看出引入 CA 注意力机制、CBAM 注意力机制的改进模型激活区域更为紧凑,最为精确,均能全面覆盖缺陷区域。实验表明,引入 CA 注意力机制、CBAM 注意力机制能够有效提升模型对稀疏分布小目标的上下文感知能力,强化小目标的高频细节信息。

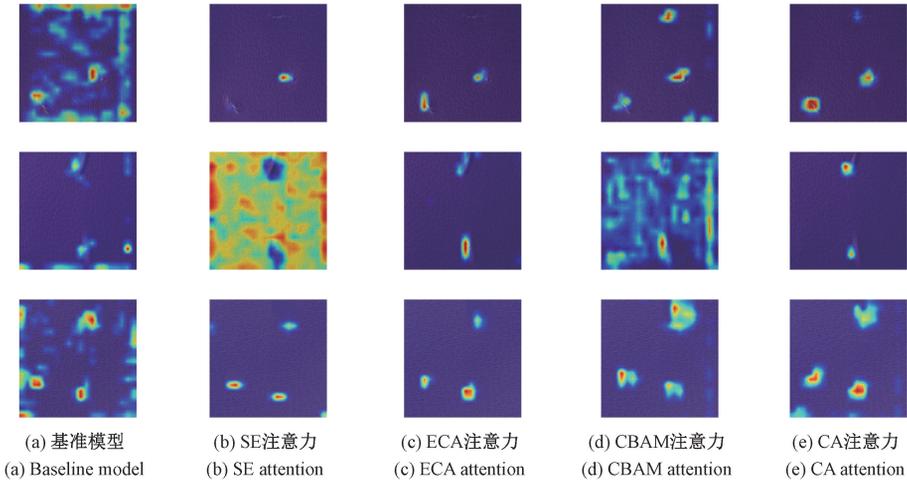


图 8 基于 Grad-CAM 技术可视化

Fig. 8 Visualization based on Grad-CAM technology

3.5 颈部网络重构实验

为了探究颈部网络重构对模型性能的影响,将原有的颈部网络 FPN+PAN 设计成加权双向特征金字塔网络,实验结果如表 3 所示。由表 3 可知颈部网络重构后,模型各项评价指标均有显著提升,其中召回率达到了 89.01%,相较于基准模型提高了 1.3%,检测精度达到了 90.75%,相较于基准模型提高了 1.65%。实验表明,颈部网络重构后,逐步细化多尺度特征表示,提升了模型对尺度变化的鲁棒性,实现了特征在不同尺度之间的高效双向流动,增强了对目标的多尺度建模能力,显著增强了模型多尺度特征融合能力和小目标检测性能。

表 3 颈部网络重构

Table 3 Neck network reconstruction

%

模型	P	R	mAP@0.5	mAP@0.5:0.95
Baseline	87.61	87.71	89.10	82.40
Baseline+BiFPN	89.42	89.01	90.75	84.24

为了直观展示颈部网络重构与基准模型在物体感知方面的差异,下面对模型检测过程中的特征图进行可视化,如图 9 所示。图 9(a)和(b)依次展示的是基准模型和颈部网络重构模型在划伤、褶皱、污渍皮革图像上的特征图可视化,且每张子图又含有 9 张小子图,从左到右、从上到下分别为主干网络第一个 Conv 模块层,4 个 C3 模块层, SPPF 模块层,以及头部网络 3 个检测头层的特征图可视化结果。从图 9(a)可以看出基准模型的特征图可视化提取到的有用信息较为有限,可视化效果较差。从图 9(b)可以看出颈部网络重构模型在网络浅层中提取的位置细节信息更加丰富;在网络深层中提取的语义特征更加高级抽象。实验表明,改进模型相较于基准模型,对数据信息的特征提取更加成功,模型能更好地学习到输入数据的模式特征。

3.6 小目标检测头实验

为了验证在模型头部网络中设计一个小目标检测头对模型检测性能的提升效果,本文将改进模型与基准模型

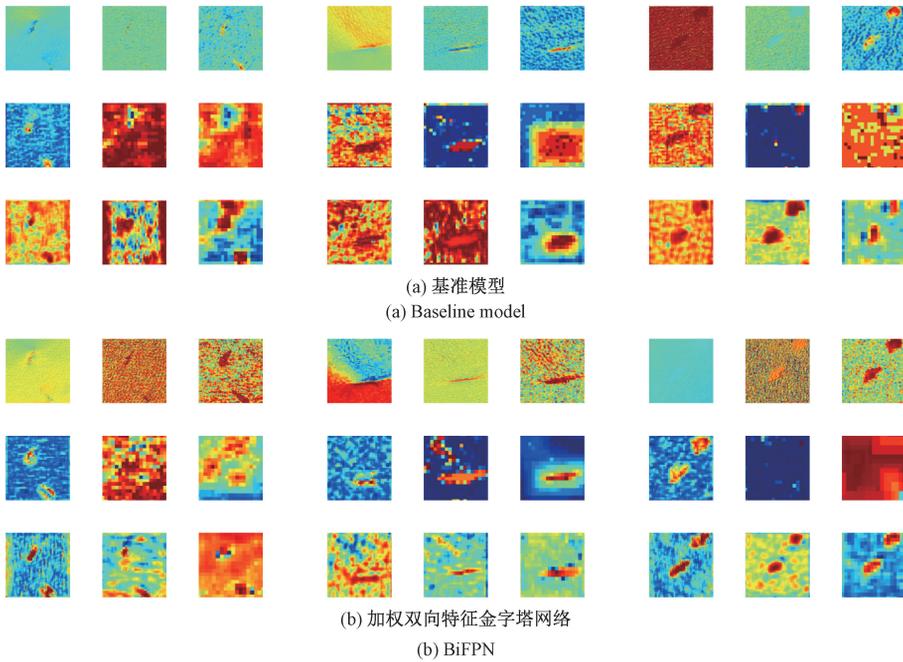


图9 特征图可视化

Fig.9 Feature map visualization

在训练过程中的关键性能指标进行可视化对比,绘制如图10所示的4组评价指标曲线。图10的4张子图分别从精准率、召回率、 $mAP@0.5$ 和 $mAP@0.5:0.95$ 这4个维度上反映模型在训练过程中的性能演变趋势。图10(a)展示了精准率随训练轮数变化的曲线,从图中可以看出,改进模型在训练初期精准率迅速提升,并在中后期持续维持在高位,收敛后高于基准模型。这表明小目标检测头的引入有效提升了模型对缺陷的判别能力,降低了误检率,尤其在复杂背景或纹理干扰显著的场景下仍保持了良好的判别性能。图10(b)为召回率变化曲线,对比可见,基准模型在训练后期召回率趋于平稳但略显不足,而改进模型的召回率在多个训练阶段均保持较高水平,尤其在后期均超过基准模型,说明小目标检测头增强了模型对小尺寸缺陷目标的捕获能力,缓解了漏检现象,提升了检测模型的鲁棒性。图10(c)所示为 $mAP@0.5$ 曲线,表示在IoU阈值为0.5时的平均检测精度,从曲线可见,改进模型在训练中期即开始与基准模型拉开差距,并逐步稳定在更高的收敛值。说明其在位置预测和类别判定方面整体准确率更优,该提升来源于模型在特征融合和小目标信息保持方面的能力增强。图10(d)展示了 $mAP@0.5:0.95$ 曲线,衡量模型在多尺度IoU阈值下的综合检测性能,该指标对模型的泛化能力和定位精度要求更高。实验结果显示,改进模型在该指标均高于基准模型,收敛速度更快且波动更小,进一步印证了小目标检测头在提升模型稳定性、多尺度适应能力以及整体检测精度方面的有效性。综上分析,本文提出的改进方法不仅提升了模型在小目标检测任务

中的准确性和鲁棒性,还在保持高性能的同时,确保了训练过程的稳定收敛性与泛化能力。

3.7 消融实验

为了验证各个改进方法的有效性以及改进方法组合集成对模型检测性能的影响,在基准模型基础上,将改进方法排列组合进行消融实验,结果如表4所示。由表4可知,本文所提到的各项单一的改进方法对模型检测性能的提升均有促进作用。尤其是融合CA注意力机制、加权双向特征金字塔网络以及小目标检测头的集成方法对模型性能提升最多,检测精度达到了92.16%,相较于基准模型提高了3.06%。实验表明,本文所提融合的改进方法比单一采取一种改进策略检测效果要更加显著,此时模型的综合性能能达到最佳。

为了检验所提改进方法和基准模型在皮革缺陷检测过程时的效果,下面在测试集上进行检测验证,如图11所示。图11(a)和(b)分别展示的是基准模型和本文所提改进模型在划伤、褶皱、污渍皮革图像上的检测效果图,从图11(a)可以看出基准模型在检测时会出现冗余的预测框,且预测框定位的范围会超出缺陷区域、出现定位偏差。从图11(b)可以看出相较于基准模型,所提融合改进方法在检测时预测框内的置信度是最高的,且当图像中出现多个缺陷目标时,预测框也能够精准定位且均能覆盖到缺陷区域,并且能够准确识别缺陷类型。实验表明,所提融合改进方法能够有效解决基准模型存在的预测框定位偏差、漏检误检等问题,说明改进方法是成功的。

3.8 其他模型对比实验

经过上述实验结果与分析,为了验证本文所提改进方

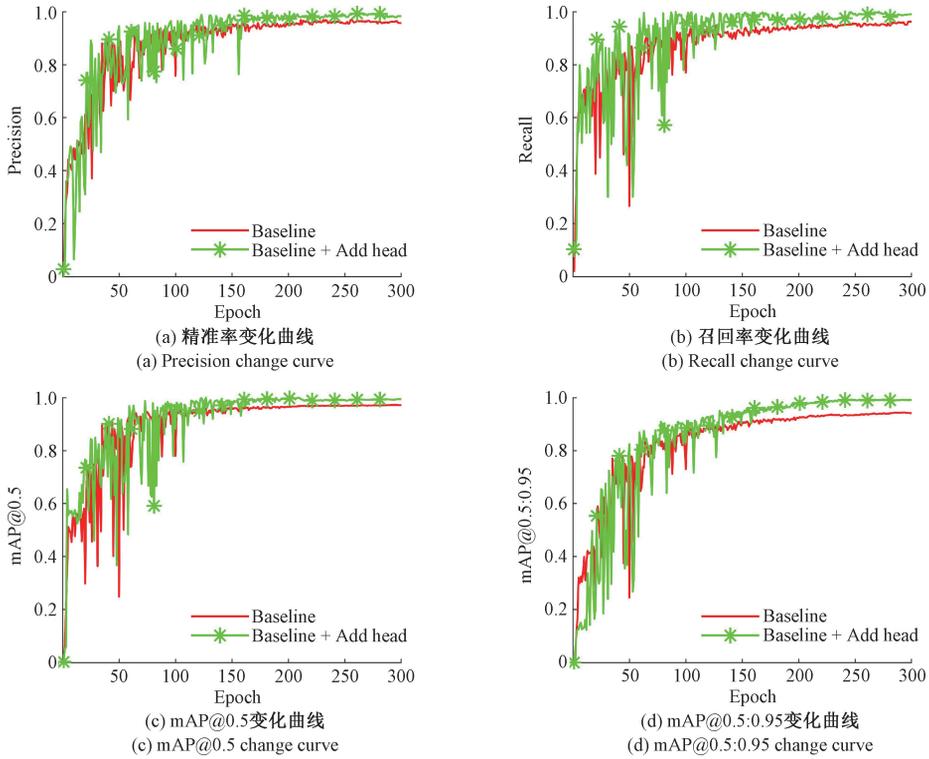


图 10 性能评价指标曲线
Fig. 10 Performance evaluation index curve

表 4 消融实验结果

Table 4 Results of ablation experiment

方法	P	R	mAP@0.5	mAP@0.5:0.95
Baseline	87.61	87.71	89.10	82.40
Baseline + Add head	88.92	90.03	90.90	84.09
Baseline + BiFPN	89.42	89.01	90.75	84.24
Baseline + CA	89.62	90.58	91.67	84.78
Baseline + Add head + CA	91.43	90.65	91.80	85.04
Baseline + BiFPN + CA	91.93	90.92	91.92	85.17
Baseline + Add head + BiFPN + CA	92.25	92.27	92.16	85.36

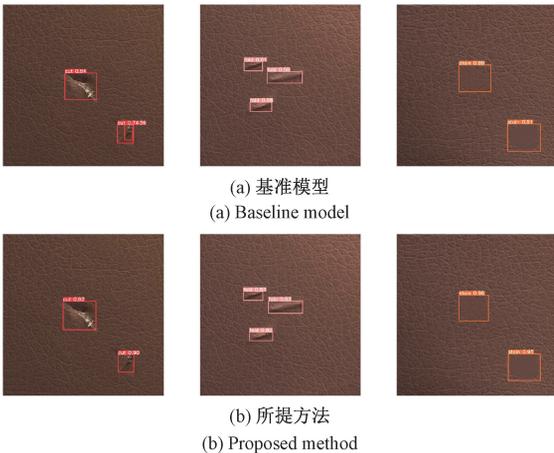


图 11 检测效果图

Fig. 11 Detection effect diagram

法的有效性,将其与当前主流的几种目标检测模型进行对比,结果如表 5 所示。从表 5 可知,所提方法相较于其他主流的单阶段目标检测模型和两阶段目标检测模型,在各项评价指标上均是最优的。其中召回率达到了 92.27%,相较于基准模型提高了 4.56%,模型检测精度达到了 92.16%,相较于基准模型提高了 3.06%。实验表明,本文所提改进方法通过在主干网络引入 CA 注意力机制、颈部网络重构,增加小目标检测头,显著增强了网络特征提取能力,使网络能够更有效地学习层次化的语义信息,提升融合的精度和语义表达能力,增强了模型对小目标细粒度特征的建模能力,解决了小目标因低像素占比和语义模糊导致的检测瓶颈,显著提高了小目标检测精度与召回率。

表5 不同算法对比实验结果

Table 5 Comparison of experimental results of different algorithms

算法	P	R	mAP@0.5	%
				mAP@0.5:0.95
Faster R-CNN	88.12	88.27	90.06	83.23
SSD	84.96	84.98	86.79	79.83
YOLOv5n	84.66	85.31	86.55	80.53
YOLOv5m	81.33	85.40	86.61	77.81
YOLOv5l	74.82	83.86	80.47	71.99
YOLOv6s	87.23	86.95	88.25	81.97
YOLOv7-Tiny	87.02	87.12	88.76	82.39
YOLOv8n	88.27	87.96	89.51	82.56
YOLOv9s	88.73	88.21	89.87	82.74
Baseline	87.61	87.71	89.10	82.40
Proposed method	92.25	92.27	92.16	85.36

4 结 论

针对当前主流的目标检测模型在检测缺陷皮革时存在漏检率和误检率高、检测精度低的问题,本文提出了一种基于改进YOLOv5s的皮革缺陷检测方法。首先为了增强模型特征提取能力,使网络更有效地学习层次化的语义信息,增强上下文感知能力,在主干网络中引入CA注意力机制。其次为了增强模型特征融合与交互能力,细化多尺度特征表示,提升融合的精度和语义表达能力,优化梯度流动,将颈部网络设计成加权双向特征金字塔网络。最后为了扩大小目标在特征图上的感受野,提高模型对小目标细粒度特征的建模能力和感知能力,在头部网络中专门设计一个小目标检测头。在缺陷皮革数据集上进行实验验证,结果显示,基于改进YOLOv5s的皮革缺陷检测方法的检测精度达到了92.16%,相较于基准模型提高了3.06%。不仅如此,本文所提改进方法相较于其他主流目标检测方法,在各项评价指标上是最优的,检测效果也是最佳的,相比于其他主流的目标检测模型拥有更显著的优势。实验表明,基于改进YOLOv5s的皮革缺陷检测方法为皮革包装产品外表面缺陷检测研究提供了新的思路和技术支持,具有重要的应用价值。

参考文献

- [1] 张嘉宁,韩凤霞,王红军,等. 3D打印碳纤维复合材料件缺陷识别方法研究[J]. 电子测量技术, 2025, 48(5): 137-146.
ZHANG J N, HAN F X, WANG H J, et al. Research on defect identification method for 3D printed carbon fiber composite structural components [J]. Electronic Measurement Technology, 2025, 48(5): 137-146.
- [2] 刘庭浔,王飞,刘旭,等. 基于改进YOLOv5的仓储物体检测研究[J]. 电子测量技术, 2024, 47(14): 149-158.
LIU T X, WANG F, LIU X, et al. Research on warehouse object detection based on improved YOLOv5 [J]. Electronic Measurement Technology, 2024, 47(14): 149-158.
- [3] 赵松怀,周敏,申飞,等. 基于改进YOLOv5s的烟火轻型检测算法[J]. 电子测量技术, 2024, 47(17): 140-146.
ZHAO S H, ZHOU M, SHEN F, et al. Light detection algorithm of pyrotechnics based on improved YOLOv5s [J]. Electronic Measurement Technology, 2024, 47(17): 140-146.
- [4] 欧晓放,韩凤春,田竞,等. 基于改进YOLOv5s模型的电动三轮车检测[J]. 激光与光电子学进展, 2025, 62(2): 214-222.
OU X F, HAN F CH, TIAN J, et al. Electric tricycle detection based on improved YOLOv5s model [J]. Laser & Optoelectronics Progress, 2025, 62(2): 214-222.
- [5] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]. 31st International Conference on Neural Information Processing Systems, 2017: 5998-6008.
- [6] 张振. 基于深度学习的电池壳表面缺陷检测技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2021.
ZHANG ZH. Research on surface defect detection technology of battery shell based on deep learning [D]. Harbin: Harbin Institute of Technology, 2021.
- [7] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2818-2826.
- [8] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [9] CHEN K, LI H T, LI CH SH, et al. An automatic defect detection system for petrochemical pipeline based on Cycle-GAN and YOLOv5 [J]. Sensors, 2022, 22(20): 7907.
- [10] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]. IEEE International Conference on Computer Vision, 2017: 2223-2232.
- [11] HE ZH D, WANG Y M, ZHENG AN P, et al. Insulator defect detection algorithm based on adaptive feature fusion and lightweight YOLOv5s [J]. Journal

- of Real-Time Image Processing, 2024, 22(1): 12.
- [12] TANG L, MEI SH Q, SHI Y SH, et al. Research on fabric defect detection algorithm based on lightweight YOLOv7-Tiny[J]. Journal of Natural Fibers, 2024, 21(1): 2352753.
- [13] ZHANG B F, ZHU X F, YU J H, et al. Metal sensor base defects detection using deep learning based YOLO network [J]. Signal, Image and Video Processing, 2024, 19(1): 47.
- [14] XU Z B, BAO Y CH, TIAN B Q. Improved YOLOv8 garment sewing defect detection method based on attention mechanism [J]. Journal of Measurements in Engineering, 2024, 12(4): 706-721.
- [15] XIE Y G, ZHAO Y W. Lightweight improved YOLOv5 algorithm for PCB defect detection[J]. The Journal of Supercomputing, 2024, 81(1): 261.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision-ECCV 2016; 14th European Conference, 2016: 21-37.
- [17] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [18] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [19] REN SH 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, 2016, 39(6): 1137-1149.
- [20] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [C]. IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [21] HE K M, ZHANG X Y, REN SH 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.
- [22] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [23] LIU SH, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [24] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [25] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534-11542.
- [26] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [27] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]. European Conference on Computer Vision(ECCV), 2018: 3-19.
- [28] TAN M X, PANG R, LE Q V. EfficientDet: Scalable and efficient object detection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.
- [29] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]. IEEE International Conference on Computer Vision, 2017: 618-626.

作者简介

白先浪, 硕士研究生, 主要研究方向为数字图像处理与识别、深度学习与目标检测、智能包装技术等。

E-mail: 810339624@qq.com

张群利(通信作者), 博士, 副教授, 主要研究方向为智能包装技术、包装测试技术等。

E-mail: zhangqunli@nefu.edu.cn

辛志强, 硕士研究生, 主要研究方向为高光谱技术、深度学习与目标检测等。

E-mail: 1256745@qq.com