

DOI:10.19651/j.cnki.emt.2416795

基于改进 YOLOv8 的轻量化皮革缺陷检测方法^{*}

方明¹ 张娇² 徐晶^{2,3} 王绎覃²

(1. 长春理工大学人工智能学院 长春 130022; 2. 长春理工大学计算机科学技术学院 长春 130022;
3. 长春理工大学中山研究院 中山 528403)

摘要: 为了解决 YOLOv8 参数量过大影响检测速度等问题,本文以汽车座椅皮革为样本对汽车座椅表面进行缺陷检测,提出了一种轻量化的基于 YOLOv8 框架的皮革缺陷检测算法。首先,将 YOLOv8 原本的主干网络替换成轻量化网络 StarNet,StarNet 通过星型运算实现了高维和非线性特征空间的映射,从而在紧凑的网络结构和较低的能耗下展示了出色的性能和低延迟。其次,将原本的检测头替换成轻量级共享卷积检测头,通过使用共享卷积,可以大幅减少参数数量,使得模型更轻便,以便于在资源受限的设备上部署。最后,将颈部网络的 C2f 模块替换成 C2f_Star 模块,在网络更加轻量化的同时,将不同尺度的特征图进行融合,提高目标检测的准确性和鲁棒性。在自制的 HSV-Leather 数据集上对模型进行实验验证,结果表明,改进后的 YOLOv8-Leather 检测模型性能优于 YOLOv8n 模型。对比 YOLOv8n 模型,改进后的模型在参数量上降低了 57%,检测速度提升了 20%,模型权重降低了 52%,运算量降低了 53%。实验验证了改进后的模型在解决皮革表面缺陷检测问题上的可行性。

关键词: 皮革缺陷检测;YOLOv8;目标检测;轻量化;StarNet

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Lightweight leather defect detection method based on improved YOLOv8

Fang Ming¹ Zhang Jiao² Xu Jing^{2,3} Wang Yitan²

(1. School of Artificial Intelligence, Changchun University of Science and Technology, Changchun 130022, China;

2. School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China;

3. Zhongshan Research Institute, Changchun University of Science and Technology, Zhongshan 528403, China)

Abstract: In order to solve the problems such as the large amount of YOLOv8 parameters affecting the detection speed, this paper proposes a lightweight leather defect detection algorithm based on the YOLOv8 framework by using automotive seat leather as a sample for defect detection on the surface of automotive seats. Firstly, the original backbone network of YOLOv8 is replaced with the lightweight network StarNet, which achieves the mapping of high-dimensional and nonlinear feature spaces through star arithmetic, thus demonstrating impressive performance and low latency with a compact network structure and low energy consumption. Secondly, the original detection head is replaced with a lightweight shared convolutional detection head, which allows for a significant reduction in the number of parameters through the use of shared convolution, making the model lighter so that it can be easily deployed on resource-constrained devices. Finally, the C2f module of the neck network is replaced by the C2f_Star module, which fuses feature maps of different scales while the network is more lightweight to improve the accuracy and robustness of target detection. Experimental validation of the model on the home-made HSV-Leather dataset shows that the improved YOLOv8-Leather detection model outperforms the YOLOv8n model. Compared to the YOLOv8n model, the improved model reduces the number of parameters by 57%, improves the detection speed by 20%, reduces the model weights by 52%, and reduces the computation by 53%. The experiment verifies the feasibility of the improved model in solving the problem of leather surface defect detection.

Keywords: leather defect detection; YOLOv8; target detection; lightweight; StarNet

0 引言

汽车座椅是汽车内饰的重要组成部分,其外观直接影

响消费者对车辆的第一印象。精致、美观的座椅可以提升整车的档次感和品牌形象,增强市场的竞争力。因此对于实现对汽车座椅外观缺陷检测也显得尤为重要。由于人工

收稿日期:2024-09-03

* 基金项目:中山市科技局引进科研创新团队项目(CXTD2023005)资助

检测存在效率低下和检测能力有限等问题,所以自动化的外观缺陷检测在工业生产中必不可少,自动化检测可以大幅度提高检测速度和准确性,减少人工检测的主观误差和劳动强度,从而提高整体生产效率。

目前,国内外对汽车座椅外观缺陷检测的方法有很多种,大体上可以分为传统方法和深度学习的方法两大类。其中,传统方法主要包括人工目视检测、触觉检测、灯光照明辅助检测、模板对比法、使用放大镜或显微镜辅助检测等方法,但是这些方法依赖人工经验,缺乏一致性,检测速度慢,难以适应大规模生产和发现微小、复杂或隐藏的缺陷。所以,自动化检测、机器视觉和人工智能技术逐渐取代了这些传统方法,并且提高了检测的效率和准确性。基于深度学习的目标检测方法能够更好地满足汽车座椅外观缺陷检测的需求,推动汽车制造业向智能化和高质量方向发展。

在汽车座椅的制造过程中,80%以上的座椅表面选择皮革作为面料。皮革是经过一系列物理与化学加工处理所制成的一种坚固、耐用物质,因具有良好的耐热性、排湿性、耐寒性、热稳定性、透气性以及优秀的着色能力而深受人们青睐。因此,本文选取皮革作为样本对汽车座椅表面进行缺陷检测,通过对污渍、划痕、孔洞、异物、褶皱 5 种缺陷的检测来验证本实验的可行性。

随着深度学习的快速发展,通过基于深度学习的目标检测算法在目标检测中越来越重要,为实现汽车座椅外观缺陷检测提供了非常重要的基础。目前主流的目标检测算法可以分为两大类:分别是一阶段(one-stage)目标检测算法和两阶段(two-stage)目标检测算法。一阶段目标检测算法比较典型有 YOLO^[1], SSD^[2], Retina-Net^[3] 等。二阶段目标检测算法将目标检测问题划分为两个阶段,第一步先产生候选区域(region proposals),然后对候选区域分类,比较典型的算法有 Faster R-CNN^[4], Mask R-CNN^[5], SPP Net^[6]。近年来,许多研究人员针对 YOLOv8 算法的网络结构对检测算法进行了轻量化的改进。其中,吴正香等^[7]为了满足实际工业生产实时检测的要求,提出了一种轻量化 YOLOv8-SEF 的织物疵点检测算法,该算法以 YOLOv8 为基准模型,引入轻量级网络 EfficientViT 替换 YOLOv8 的主干网络并在颈部网络中融入了 SE 注意力机制,最后通过使用 Slide Loss 损失函数取代 CIoU 损失函数来促进 YOLOv8 网络结构轻量化。成顺等^[8]为了提高水下光学目标识别精度,解决计算复杂高等情况,提出了一种 ACE 图像增强的轻量化 YOLOv8 水下光学识别算法,该算法将改进的 SENetV2 网络替换了 YOLOv8 骨干网络,将 CCFM 轻量级的特征融合网络替换颈部网络,再将 DySample 替换为传统上采样器,提升图像处理的效率。改进之后的模型在数据集上的 mAP 值达到 77.3% 和 71.5%。张国鹏等^[9]针对检测模型大、运算复杂、对运算平台资源要求高的问题,通过改进主干层、优化特征模块、引入注意力机制等方法,提出了一种基于 YOLOv8 改进的轻

量化安全帽佩戴检测算法 YOLOv8-MBS,改进之后的 mAP 值达到了 87.44%。刘子洋等^[10]为了解决小目标检测方面存在漏检、错检的问题,通过引入基于 Transformer 结构的 BiFormer 注意力机制,引入 GSConv 模块等方法,提出了一种基于 YOLOv8n 改进的轻量化目标检测算法,改进后的算法精度值为 77.1%,模型参数量为 2.8 MB。李卓润等^[11]利用 YOLOv7-tiny,使用 MobileNetV3 作为新的特征提取网络,引用 PConv 减少 Neck 层的计算量,最后在 Head 层增加新的注意力机制使 YOLOv7-tiny 网络具有更高的效率和精度。任安虎等^[12]为了解决 YOLOv8 模型参数量过大影响检测精度和速度等问题,提出了一种改进 YOLOv8n 的目标检测算法,通过添加注意力机制,替换主干网络,修改损失函数等方法,减少了模型的参数量和计算量。钱尚乐等^[13]提出了一种改进 YOLOv7-tiny 的轻量化列车轮对踏面缺陷检测方法。通过替换主干网络,增加注意力机制来提升检测速度和检测精度。李忠科等^[14]通过调整网络结构,增加注意力机制等提出了一种基于轻量级 YOLOv8n 网络的 PCB 缺陷检测算法。以上方法在存在处理大规模数据时有效率瓶颈或响应时间慢,以及对计算资源或存储资源的高需求,并且在特定情况下的检测准确率不足,可能会漏检或误检导致检测精度低下等问题。

对比最新的已有的改进轻量级 YOLOv8n 网络,本文以皮革为样本对汽车座椅表面进行缺陷检测,提出了一种改进 YOLOv8 的轻量化皮革缺陷检测方法。首先,本文将 StarNet 网络替换 YOLOv8 的主干网络进行轻量化设计,这样大幅度的降低计算量,能够使模型在低性能的平台部署,保证了模型的实用性和可靠性。其次,因为缺陷位置和种类在图像中分布比较随机,随着检测头中网络层数加深,浅层特征保留的更多原始图像信息在深层网络中容易丢失,在缺陷检测时容易产生漏检、错检等问题。为了解决以上问题,本文提出了一个新的轻量级共享卷积检测头,由于轻量级共享卷积检测头中使用了共享卷积,通过共享卷积获取不同尺度的信息并进行共享,能够使网络聚焦于数据中不同类别之间的细微差异,使得检测头在参数量、计算量都减少的情况下,保证足够的检测精度。最后,本文通过改进颈部网络的设计,将 C2f 模块替换为 StarBlocks 模块,StarBlocks 模块使用了元素乘法,不需要堆叠网络深度,隐式实现高维特征映射,不依赖于激活函数来强化模型非线性表达能力,在保证模型轻量化的同时提升了网络的检测速度。

1 相关工作

YOLOv8 算法:

YOLOv8^[15]是一个全新的 SOTA 模型,支持图像分类、物体检测和实例分割任务,是一种基于图像全局信息进行预测的目标检测系统。YOLOv8 的网络结构主要可分为输入端(Input)、主干特征提取模块(Backbone)、特征加

强模块(Neck)和检测模块(Head)共4个部分。

YOLOv8网络在训练过程中自带数据增强,在训练时会数据集实时进行数据增强操作,这样确保了模型数据集的多样性与泛化能力。YOLOv8的主干特征提取模块仍然沿用了CSPDarkNet^[16]结构,并采用梯度流更加丰富的C2f模块代替C3模块,并对不同尺度模型调整了不同的通道数,降低了模型计算量,提升了收敛速度与收敛效果。特征加强模块采用PANet^[17](path aggregation network)结构,不仅会对特征进行上采样实现特征融合,还会对特征进行下采样实现特征,最终将PANet的3个输出分支送入检测模块。YOLOv8的检测头是一个解耦头,解耦头采用两条并行的分支分别提取类别特征和位置特征,然后每一条分支各用一层 1×1 卷积完成分类和定位任务。

对比YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x,本文采用的是轻量级的YOLOv8n网络,对比同YOLOv8系列,YOLOv8n的模型体积更小,检测速度更快。

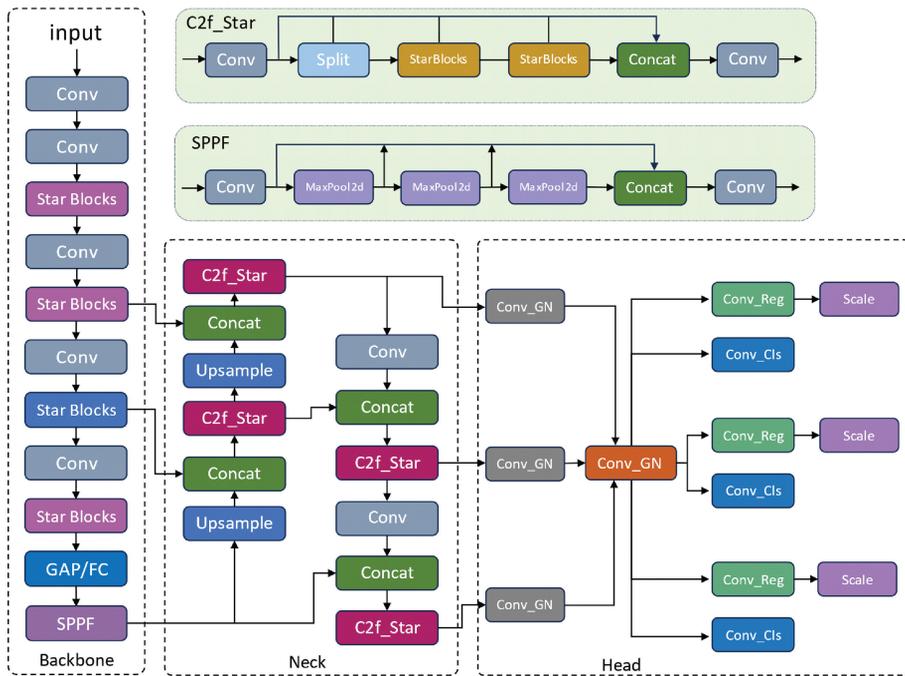


图1 YOLOv8n-Leather网络结构图

Fig.1 Network structure of YOLOv8n-Leather

YOLOv8模型改进策略:

1) YOLOv8-StarNet模型

StarNet是2024年3月微软发布的最新网络模型。本文使用的StarNet^[18]遵循传统的分层网络,直接使用卷积层降低分辨率,并在每个阶段将通道数增加一倍。通过重复多个星形块来提取特征,没有任何复杂的结构和精心选择的超参数,StarNet展示了在紧凑网络结构和高效预算下出色的性能和低延迟。StarNet结构如图2所示。StarNet采用4级分层架构,利用卷积层进行下采样,并使用StarBlocks模块进行特征提取。为了满足效率的要求,

2 本文方法

为了解决现有的皮革缺陷检测方法模型参数量大,导致计算资源需求增加等问题,本文对YOLOv8的轻量化做出以下3点改进:

针对皮革缺陷检测模型速度慢、模型大,难以部署在嵌入式模块的问题,首先,将YOLOv8原本的主干网络替换成轻量级的StarNet网络结构,StarNet通过星操作实现高维和非线性特征空间的映射,从而提升了网络的性能,降低了延迟。其次,将原本的检测头替换成了轻量级共享卷积检测头,通过使用两个 3×3 的共享卷积,大幅减少参数数量的同时,在检测头的后面添加了一个Scale层对特征进行缩放,让检测头做到在参数数量更少、计算量更少的情况下,减少精度的损失。最后,将颈部网络的C2f模块替换成C2f_Star,以便将不同尺度的特征图进行融合,提高目标检测的准确性和鲁棒性。改进之后的模型结构图如图1所示。

将Layer Normalization^[19]替换为Batch Normalization^[20],并将其放置在深度卷积之后(可以在推理时融合)。在这个过程中,将通道扩展因子始终设置为4,并且在每个阶段将网络宽度翻倍。遵循MobileNetv2^[21]设计,将GELU激活函数替换为ReLU6,相比GELU,ReLU6在输出值的范围上进行了限制,使得输出值不会超过6,这样做具有一定的防止梯度爆炸的作用。

对比其他深度学习网络,将StarNet用于替换YOLOv8的主干网络进行轻量化改进可以显著减少模型的参数量和计算量,从而提高模型的部署效率和推理速

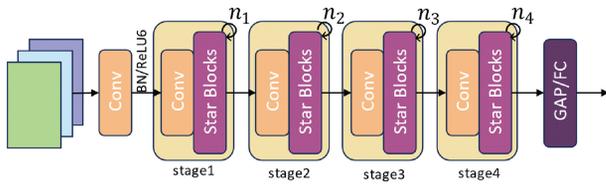


图 2 StarNet 网络结构图

Fig. 2 StarNet network structure diagram

度。使得模型更具有扩展性和适应性,能够更好地应对不同的硬件平台和应用场景,使得模型在不同的部署环境中都能够保持高效和可靠。

2) 轻量级共享卷积检测头

检测头通常负责将网络的特征映射转换为目标检测框及其相关的类别置信度和位置信息。在 YOLOv8 中,检测头参数量大幅增加的原因是 YOLOv8 采用了解耦头结构。因为 YOLOv8 的检测头有 6 个 1×1 的卷积和 12 个 3×3 的卷积,所以造成 YOLOv8 检测头的参数量过大。为了解决 YOLOv8 解耦头参数量过大的问题,本文通过优化 YOLOv8 检测头的设计,来改善目标检测的准确性,使模型在复杂场景中更准确地定位和识别目标。改进的检测头还可以带来检测速度的提升和模型运行效率的改善。改进后的检测头会减少模型的计算量和提高模型在硬件上的并行处理能力,从而加快目标检测的速度,使其能够在实时或高吞吐量的应用中表现更好。

轻量级共享卷积检测头 (lightweight shared convolutional detection head, LSCD), GroupNorm (GN)^[22] 把通道分为组,并计算每一组之内的均值和方差,以进行归一化处理可以提升检测头定位和分类的性能。GN 的计算与批量大小无关,其精度也在各种批量大小下保持相对稳定。将 GN 用于卷积网络中,归一化做特征融合,获得更多的感受野,可以提升检测头定位和分类的性能。LSCD 网络结构如图 3 所示。本文改进的轻量级共享卷积检测头,首先进行一个 Detect 的输入,经过一个 1×1 的 Conv_GN 的卷积之后,进行一个合并操作,通过使用两个 3×3 的共享卷积进行参数共享,再转换到输出尺度的卷积网络。最后,运用 Scale 层增加网络的表征能力,允许网络学习对输入数据进行缩放和平移的操作,从而提高模型的灵活性和拟合能力。

通过使用共享卷积,大幅减少了参数的数量,当模型在资源受到限制的设备上使用,可以提升模型的检测速度,使模型更轻便。在使用共享卷积的时,为了解决模型在检测中每个检测头所检测的目标尺度不一致问题,LSCD 使用 Scale 层对特征进行缩放。在缩放的同时,让检测头做到在参数量更少、计算量更少的情况下,尽可能减少精度的损失。

3) YOLOv8-Star

YOLOv8 在颈部网络中使用了 C2f 模块来进行特征

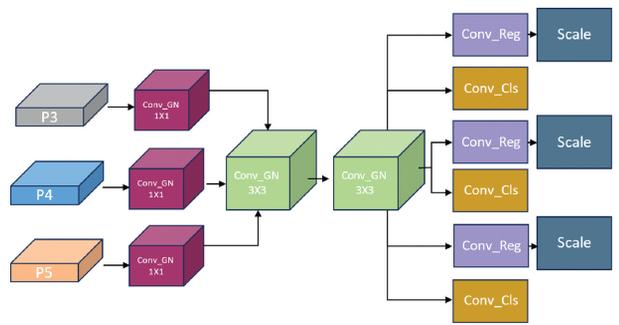


图 3 轻量级共享卷积检测头网络结构图

Fig. 3 Structure of lightweight shared convolutional detection head network

融合,来提高目标检测的性能。颈部网络通常负责将来自主干网络不同层级的特征图进行融合和增强,通过利用低级特征图和高级特征图的优势,提高目标检测的精度和鲁棒性。

由于需要进行特征融合操作,会增加模型计算的复杂度,导致模型训练和推理的时间成本增加,C2f 模块的引入大幅增加了模型的参数量,这样就需要更多的存储空间和计算资源。为了解决以上问题,本文通过改进颈部网络的设计,将 C2f 模块替换为更加轻量化的 C2f_Star,在降低模型参数量和计算量的同时,能够更有效地整合多层次、多尺度的特征信息,提升模型对目标的表征能力,有助于改进检测精度和模型的泛化能力。受到 MobileNeXt^[23] 的启发,在每个块的末尾加入了一个深度可分离卷积 DW Conv, DW Conv 是由可分离卷积拓展而来, DW Conv 每个卷积核深度为 1,相较于其他常规卷积, DW Conv 是一种轻量级的卷积操作,能够减少参数量和运算量。 DW Conv 结构图如图 4 所示。

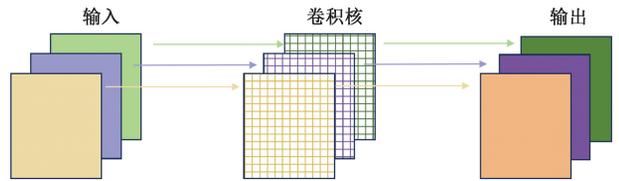


图 4 DW Conv 结构图

Fig. 4 DW Conv structure diagram

C2f_Star 结构使用了 StarNet 中的 StarBlocks 模块替换了 YOLOv8 颈部网络中的 C2f 模块, StarBlocks 模块网络结构图如图 5 所示,在 StarBlocks 模块中使用了 Star Operation (元素乘法),通过逐元素乘法融合两个线性变换的特征。揭示了 Star Operation 在无需加宽网络下,将输入映射到高维非线性特征空间的能力。与增加网络宽度的传统神经网络不同, Star Operation 类似于在不同通道上进行成对特征乘法的核函数,特别是多项式核函数。当应用到神经网络中并通过多层堆叠时,每一层都会带来隐式维度复杂性的指数增长。只需几层, Star Operation 就

可以在紧凑的特征空间内实现几乎无限的维度。与此同时,Star Operation 不引入额外的参数,只是利用输入特征之间的乘积或加权乘积。这种设计使得网络结构变得轻量化,减少了训练和推理时的计算成本和存储需求。

Star Operation 在性能上始终优于求和,使用 C2f_Star 轻量化颈部网络可以降低计算复杂性,进行轻量化改进,并且 C2f_Star 能够使颈部网络保持足够的特征表达能力。通过堆叠多个层,即使只是几层,Star Operation 可以以指数方式显著放大隐式维度,以维持 YOLOv8 模型的高检测精度和泛化能力。C2f_Star 结构图如图 6 所示。

3 实验验证

3.1 实验环境与参数设置

本文训练的实验环境为 64 位的 win11 操作系统,实验环境为 python = 3.9,框架为 TorchVision = 0.14.1,

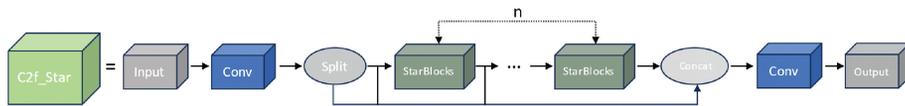


图 6 C2f_Star 结构图

Fig. 6 C2f_Star structure diagram

TorchAudio=0.13.1,PyTorch=1.13.1,相关硬件配置和模型参数如表 1 所示。

表 1 实验相关硬件配置和模型参数设置

Table 1 Experimental hardware configuration and model parameter settings

名称	配置	参数	参数值
GPU	NVIDIA GeForce RTX 4060	图像大小	640×640
CPU	Intel(R) i7-12700H	学习率	0.000 1
CUDA	11.6	批量大小	16
Workers	8	Auto_augment	SGD
Batchsize	16	Close_mosaic	10

3.2 数据集准备

为了进行实证分析,本文采用自行搭建的皮革缺陷样本采集系统获取常见的汽车座椅皮革缺陷样本,HSV-Leather 数据集包括常见的污渍、划痕、孔洞、异物、褶皱 5 种缺陷,每种缺陷各拍摄 1 900 张缺陷样本,总计 9 500 张缺陷样本。本数据集通过 Labelimg 对各种缺陷进行标注,总共标注了 9 500 个标记边界框。其中,数据集按照 7:2:1 划分为训练集、验证集、测试集,训练集为 6 500 张,验证集为 1 900 张,测试集为 950 张。此外,为了提升检测模型对目标任务的泛化能力和鲁棒性,数据集的拍摄使用的是 HSV 光源,主要通过 HSV 光源对不同缺陷的显示程度,增强泛化能力来提升目标检测的检测精度。在第 3.4 节

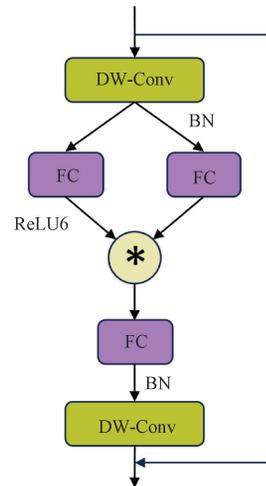


图 5 StarBlocks 模块图

Fig. 5 StarBlocks block diagram

中报告了在 HSV-Leather 数据集上的实验结果。

3.3 评价指标

本文的主要目标是在保证精度的前提下降低模型的计算量、参数量、模型权重和提升检测速度。因此,本实验将多类别平均精度(mean average precision, mAP)、模型权重(Weights)、模型参数量(Parameters)、模型计算量(GFLOPs)、检测速度(frame per second, FPS)作为模型检测效果的评估指标。这些评估指标综合考虑了模型的检测精度、模型大小与复杂度、模型的泛化能力这 3 个方面,有助于在汽车座椅的表面缺陷检测中提供更高效的目标检测方案。

3.4 HSV-Leather 数据集性能评估

为了提升 YOLOv8 的表现性能,有效解决模型资源占用过大等问题,本文对 YOLOv8 模型进行了轻量化改进,包括融合了轻量化 StarNet 网络、LSCD、使用 StarBlocks 模块对 C2f 模块进行轻量化。本文把融合了 StarNet 模块的 YOLOv8 称为 YOLOv8-StarNet 模型;把添加了 LSCD 的 YOLOv8 模型称为 YOLOv8-LSCD 模型;把 C2f 替换为 StarBlocks 模块的 YOLOv8 模型称为 YOLOv8-Star 模型。将以上 3 种模型融合在一起的称为 YOLOv8-Leather。为了验证以上改进带来的性能提升,本文在 HSV-Leather 数据集上进行了消融实验和对比实验,并用实验验证了模型的性能表现。

1) 消融实验

为了证明改进损失函数后模型的优势,本文在 HSV-Leather 数据集上进行了消融实验,通过 Parameters、

Weights、FPS、GFLOPs 等 4 个指标进行评估模型的效果。结果如表 2 所示,从实验结果上来看 YOLOv8-Leather 模型相较于 YOLOv8n 模型 Params 降低了 56.7%;Weights 降低了 51.6%;GFLOPs 降低了 52.9%;FPS 提升了 19.8%。本研究验证了轻量化 YOLOv8 设计的有效性。

同时,为了更好地展示本文改进后各缺陷的精度,对污渍、划痕、孔洞、异物、褶皱 5 种缺陷的检测结果如图 7 所示。从图 7 中可以看出,对比基线模型,YOLOv8-Leather 精度的保持相对稳定,这更加说明了本文改进之后的模型在皮革缺陷检测方面的可行性。

表 2 消融实验结果

Table 2 Results of ablation experiment

模型	mAP	Params/M	FLOPs/G	FPS/fps	Weights/M
YOLOv8n	93.34	3.0	8.7	121.7	6.2
YOLOv8-StarNet	91.96	2.2	6.5	142.7	4.7
YOLOv8-LSCD	92.72	2.4	6.5	147.9	4.9
YOLOv8-Star	91.99	2.5	6.9	135.9	5.3
YOLOv8-Leather	93.48	1.3	4.1	145.8	3.0

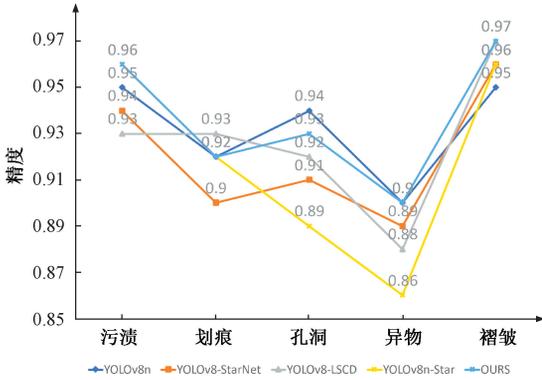


图 7 各类别缺陷精度

Fig. 7 Accuracy of defects of various types

2)对比实验

为了进一步验证所提出的模型的检测性能,本文选择了与 8 个优秀的目标检测模型进行比较,包括 YOLOv5n、YOLOv6-N、YOLOv7-tiny、YOLOv8n、YOLOv8s、RTDETR-YOLO、YOLO-SOP、YOLOv8-FNBF。YOLOv8-Leather 与目前主流的目标检测模型相比,在检测精度、检测速度、模型参数量、模型权重、模型计算量等方面都优于其他的模型,对比实验结果如表 3 所示。

3.5 实验结果及展示

为了更具体的展示本文算法与其他算法的比较结果,分别对 YOLOv8n 和改进之后的 YOLOv8-Leather 在测试集上进行了测试,图 8 展示了在 HSV-Leather 数据集上的

表 3 对比实验结果

Table 3 Comparative experimental results

模型	mAP	Params/M	FLOPs/G	FPS/fps	Weights/M
YOLOv5n	90.14	2.5	4.5	133.2	5.2
YOLOv6-N	88.98	4.2	11.4	111.6	8.7
YOLOv7-tiny	82.58	6.1	13.2	104.5	12.3
YOLOv8n	93.34	3.0	8.7	121.7	6.2
YOLOv8s	93.42	11.2	28.6	110.3	7.4
RTDETR-YOLO	95.36	19.9	56.9	78.2	40.4
YOLO-SOP ^[24]	93.07	3.2	15.8	139.1	4.3
YOLOv8-FNBF ^[25]	92.15	5.4	18.7	142	5.1
YOLOv8-EfficientViT	90.2	4.1	10.1	106.1	8.8
YOLOv8-Leather	93.48	1.3	4.1	145.8	3.0

一些检测结果。从图 8 中可以看出,每个框代表检测到的缺陷区域,每个缺陷区域上方标注了缺陷类别、缺陷的置信度,并且改进之后的模型能够更精确地定位缺陷目标的

位置,从而对边界框的位置和大小预测更自信,实现对缺陷的定位和分类,更接近实际目标的位置和大小,对图像的分析可以展示出改进之后模型的高效性和精确性。

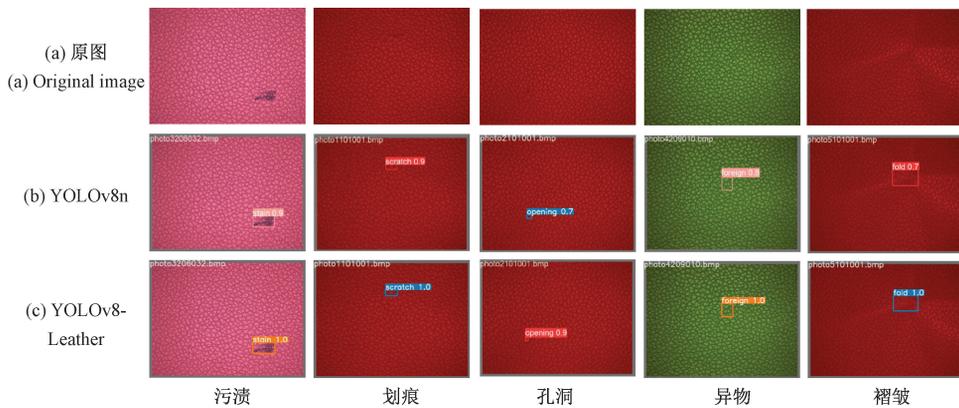


图8 各类型缺陷检测结果对比

Fig. 8 Comparison of detection results of different types of defects

4 结论

为本文针对皮革缺陷检测问题,提出了一种基于改进YOLOv8的轻量化皮革缺陷检测方法。该方法首先是将YOLOv8的主干网络替换为轻量级StarNet网络,然后将YOLOv8的检测头替换为轻量级共享卷积检测头,最后,将颈部网络的C2f模块替换为StarNet中的StarBlocks模块。通过以上3个方面的轻量化改进,显著减少了模型参改进之后的YOLOv8-Leather检测模型能够有效的降低基线模型的参数量和计算量,本文的方法不仅在准确性上与传统方法相媲美,还大大降低了在资源受限设备上的部署成本。在未来的研究中可以增加单张数据集中缺陷种类的数量,以促进模型应用的广泛性。未来,计划进一步探索知识蒸馏、权值量化、模型剪枝、通道剪枝、注意力迁移这方面的研究,以进一步提升模型的性能和适用范围,推动轻量化模型在实际应用中的广泛应用。

参考文献

- [1] REDMON, 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(CVPR), 2016:779-788.
- [2] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision-ECCV 2016: 14th European Conference, 2016: 21-37.
- [3] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. 2017 IEEE International Conference on Computer Vision(ICCV), 2017: 2999-3007.
- [4] 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, 2017, 39(6): 1137-1149.
- [5] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask

R-CNN[C]. 2017 IEEE International Conference on Computer Vision(ICCV), 2017: 2980-2988.

- [6] 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 & Machine Intelligence, 2015, 37(9): 1904-1916.
- [7] 吴正香,宋荣,刘娟,等.基于轻量化YOLOv8-SEF的色织物疵点检测算法[J/OL].棉纺织技术,1-8[2024-11-13]. <http://kns.cnki.net/kcms/detail/61.1132.TS.20240815.0815.002.html>.
- WU ZH X, SONG R, LIU X, et al. Color fabric defect detection algorithm based on lightweight YOLOv8-SEF[J/OL]. Cotton Textile Technology, 1-8 [2024-11-13]. <http://kns.cnki.net/kcms/detail/61.1132.TS.20240815.0815.002.html>.
- [8] 成顺,李建荣,王志乾,等.基于YOLOv8轻量化水下光学图像识别算法[J/OL].激光与光电子学进展,1-20[2024-11-13]. <http://kns.cnki.net/kcms/detail/31.1690.tn.20240809.1638.010.html>.
- CHEN SH, LI J R, WANG ZH Q, et al. Lightweight underwater optical image recognition algorithm based on YOLOv8 [J/OL]. Advances in Laser and Optoelectronics, 1-20[2024-11-13]. <http://kns.cnki.net/kcms/detail/31.1690.tn.20240809.1638.010.html>.
- [9] 张国鹏,周金治,马光岑,等.改进YOLOv8的轻量化安全帽佩戴检测算法[J].电子测量技术, 2024, 47(17):147-154.
- ZHANG G P, ZHOU J ZH, MA G C, et al. Improved YOLOv8 lightweight helmet wear detection algorithm[J]. Electronic Measurement Technology, 2024,47(17):147-154.
- [10] 刘子洋,徐慧英,朱信忠,等. Bi-YOLO: 一种基于YOLOv8n改进的轻量化目标检测算法[J]. 计算机工

- 程与科学, 2024, 46(8): 1444-1454.
- LIU Z Y, XU H Y, ZHU X ZH, et al. Bi-YOLO: An improved lightweight object detection algorithm based on YOLOv8n[J]. Computer Engineering and Science, 2024, 46(8): 1444-1454.
- [11] 李卓润, 李波, 邱鹏程, 等. 轻量化 YOLOv7-tiny 的水下压印字符识别[J]. 国外电子测量技术, 2024, 43(4): 162-169.
- LI ZH R, LI B, QIU P CH, et al. Underwater imprinting character recognition of lightweight YOLOv7-tiny[J]. Foreign Electronic Measurement Technology, 2024, 43(4): 162-169.
- [12] 任安虎, 李宇飞, 陈洋. 改进 YOLOv8 的高速公路交通异常事件检测[J/OL]. 激光杂志, 1-9[2024-11-13]. <http://kns.cnki.net/kcms/detail/50.1085.TN.20240628.0926.002.html>.
- REN AN H, LI Y F, CHEN Y. Improved YOLOv8 motorway traffic anomaly detection[J/OL]. Laser Journal, 1-9 [2024-11-13]. <http://kns.cnki.net/kcms/detail/50.1085.TN.20240628.0926.002.html>.
- [13] 钱尚乐, 曹伟, 高军伟. 基于改进 YOLOv7-tiny 的轻量化列车轮对踏面缺陷检测方法[J/OL]. 光电子·激光, 1-12[2024-11-13]. <http://kns.cnki.net/kcms/detail/12.1182.O4.20240711.1323.002.html>.
- QIAN SH L, CAO W, GAO J W. Defect detection method of lightweight train wheelset tread based on improved YOLOv7-tiny[J/OL]. Optoelectronics and Lasers, 1-12 [2024-11-13]. <http://kns.cnki.net/kcms/detail/12.1182.O4.20240711.1323.002.html>.
- [14] 李忠科, 刘小芳. 基于轻量级 YOLOv8n 网络的 PCB 缺陷检测算法[J]. 电子测量技术, 2024, 47(4): 120-126.
- LI ZH K, LIU X F. PCB defect detection algorithm based on lightweight YOLOv8n network [J]. Electronic Measurement Technology, 2024, 47(4): 120-126.
- [15] VARGHESE R, SAMBATH M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems(ADICS), 2024: 1-6.
- [16] WANG C Y, LIAO H Y M, YEH I 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), 2020: 1571-1580.
- [17] LIU SH, LU Q, QIN H F, et al. Path aggregation network for instance segmentation[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [18] MA X, DAI X Y, BAI Y, et al. Rewrite the stars[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024.
- [19] BA J L, KIROS J R, HINTON G E, et al. Layer normalization [J]. ArXiv preprint arXiv: 1607.06450, 2016.
- [20] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [J]. Journal of Machine Learning Research(JMLR), 2015(37): 448-456.
- [21] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [22] WU Y X, HE K M. Group normalization [C]. European Conference on Computer Vision, 2018.
- [23] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3 [C]. 2019 IEEE/CVF International Conference on Computer Vision(ICCV), 2019: 1314-1324.
- [24] 彭鸿瑞, 杨桂华. 基于改进 YOLOv8 的 SOP 芯片缺陷检测研究[J]. 电子测量技术, 2024, 47(12): 71-82.
- PENG H R, YANG G H. Research on defect detection of SOP chip based on improved YOLOv8[J]. Electronic Measurement Technology, 2024, 47(12): 71-82.
- [25] 陈梓延, 王晓龙, 何迪, 等. 基于改进 YOLOv8 的轻量化车辆检测网络[J/OL]. 计算机工程, 1-13[2024-11-12]. <https://doi.org/10.19678/j.issn.1000-3428.0069122>.
- CHEN Z Y, WANG X L, HE D, et al. Lightweight vehicle detection network based on improved YOLOv8 [J/OL]. Computer Engineering, 1-13[2024-11-12]. <https://doi.org/10.19678/j.issn.1000-3428.0069122>.

作者简介

方明, 教授, 博士生导师, 主要研究方向为图像处理、计算机视觉。

E-mail: fangming@cust.edu.cn

张娇(通信作者), 硕士研究生, 主要研究方向为图像处理、缺陷检测。

E-mail: 2022101007@mails.cust.edu.cn

徐晶, 教授, 博士生导师, 主要研究方向为图像处理、计算机视觉。

E-mail: xujing@cust.edu.cn

王绎覃, 硕士研究生, 主要研究方向为图像处理、计算机视觉。

E-mail: 2021101076@mails.cust.edu.cn