



# 基于改进 YOLOv5 模型的织物疵点检测算法

郭 波<sup>1</sup>, 吕文涛<sup>1</sup>, 余序宜<sup>1</sup>, 郭 庆<sup>2</sup>, 陈亮亮<sup>3</sup>, 王成群<sup>1</sup>

(1.浙江理工大学信息学院,杭州 310018;2.浙江省技术创新服务中心,杭州 310007;  
3.浙江经贸职业技术学院应用工程系,杭州 310018)

**摘要:**针对织物疵点存在的种类多、密集度高、尺度小等检测难点,提出一种基于改进 YOLOv5 模型的织物疵点检测算法。首先,通过 K means++方法对所有真实框进行聚类,提高了模型训练时的收敛速度;其次,将 Mish 激活函数用于模型训练,提高了其非线性表达能力;再次,通过增加检测层提升了对多尺度目标的检测性能,并调整池化层位置提取多尺度的特征信息,提升了算法的鲁棒性及检测精度;最后,优化颈部网络结构,提升了算法的检测精度和速度。基于天池织物疵点数据集的实验结果表明:该算法的 mAP 达到了 76.8%,相较基于原 YOLOv5 模型的织物疵点检测算法提升了 7.7%,验证了该算法的有效性和鲁棒性。该算法在满足实时检测的要求下提高了检测精度,并优于其他主流目标检测算法,具有良好的应用前景。

**关键词:**深度学习;YOLOv5;织物疵点检测;激活函数;实时检测

中图分类号: TS107; TP202

文献标志码: A

文章编号: 1673-3851 (2022) 09-0755-09

## Fabric defect detection algorithm based on improved YOLOv5 Model

GUO Bo<sup>1</sup>, LÜ Wentao<sup>1</sup>, YU Xuyi<sup>1</sup>, GUO Qing<sup>2</sup>, CHEN Liangliang<sup>3</sup>, WANG Chengqun<sup>1</sup>

(1.School of Information Science and Technology, Zhejiang Sci-Tech University,  
Hangzhou 310018, China; 2. Zhejiang Technical Innovation Service Center,  
Hangzhou 310007, China; 3. Department of Applied Engineering, Zhejiang  
Institute of Economics and Trade, Hangzhou 310018, China)

**Abstract:** In view of the difficulty in detecting fabric defects, i.e., wide variety, high density and small scale, a fabric defect detection algorithm based on the improved YOLOv5 model was proposed. The algorithm firstly clustered all real frames with K means ++ method and increased the convergence speed during model training. Secondly, the Mish activation function was used to train the model and enhanced its nonlinear expression ability. Thirdly, the detection performance of the model for multi-scale targets was improved by adding detection layers. The position of pooling layer was adjusted to extract multi-scale feature information and improve the robustness and detection accuracy of the algorithm. Finally, the neck network structure was optimized to improve the detection accuracy and speed of the algorithm. The experimental results based on the Tianchi fabric defect data set showed that the mAP of the proposed algorithm can reach 76.8%, which was 7.7% higher than that of the original YOLOv5 model, verifying the effectiveness and robustness of the proposed algorithm. This algorithm improves detection accuracy

收稿日期: 2022-02-24 网络出版日期: 2022-06-02

基金项目: 国家自然科学基金项目(61601410); 浙江省科技厅重点研发计划项目(2021C01047, 2022C01079)

作者简介: 郭 波(1996—),男,杭州人,硕士研究生,主要从事计算机视觉方面的研究。

通信作者: 吕文涛,E-mail: alvinlwt@zstu.edu.cn

while meeting the requirements of real-time detection and outperforms other mainstream target detection algorithms, so it has a good application prospect.

**Key words:** deep learning; YOLOv5; fabric defect detection; activation function; real-time detection

## 0 引言

表面疵点检测是织物质量控制的重要环节<sup>[1]</sup>。织物疵点通常由纱线材料质量差、生产设备老旧、纺织工艺落后等原因造成,且不同的原因产生的疵点尺度、特征和纹理也不一致。传统的人工检测方法由于精度低、效率低等缺点<sup>[2]</sup>已无法满足多类别的织物疵点检测需求,开发高效准确的织物疵点检测技术已成为现代纺织行业亟须解决的问题。

目前在织物疵点检测领域,基于深度学习的检测方法凭借其特有的类人神经网络模型,可以更有效地完成复杂场景下的检测任务,已逐步取代传统检测方法。基于深度学习的检测方法可分为两种:单阶段算法和双阶段算法<sup>[3]</sup>。单阶段算法主要有SSD<sup>[4]</sup>和YOLO系列<sup>[5-9]</sup>,双阶段算法有R-CNN<sup>[10]</sup>以及在其基础上改进的Faster R-CNN<sup>[11]</sup>、Mask R-CNN<sup>[12]</sup>和Cascade R-CNN<sup>[13]</sup>等。单阶段算法因其网络中提取特征、位置信息回归和分类预测的高集成性,运算速度相较于双阶段算法更快,因此易满足工业实时检测的需求。其中,以YOLOv5为代表的单阶段轻量化模型,主要采用图像预处理技术和改良的特征信息融合方法,提高了对小目标的检测精度,同时对其他尺度的目标也有较好的检测表现。基于此,众多研究者对YOLOv5进行了多方面的改进,以使该模型具备更强的实用性和泛化性。王新等<sup>[14]</sup>在YOLOv5模型上加入自校准卷积块和通道注意力模块,增强了模型网络的特征提取能力,并改进置信度和分类损失函数,提高了模型在复杂场景下的检测精度。吕禾丰等<sup>[15]</sup>对YOLOv5模型的预测回归损失进行改进,采用EIoU(Efficient Intersection over Union)损失替代GIoU(Generalized Intersection over Union)损失,并优化加权非极大值抑制算法,显著提高了模型的检测精度。Zhu等<sup>[16]</sup>把Transformer模块和卷积注意力模块引入到颈部网络,并增加特征检测层以提高YOLOv5模型对小目标的检测性能。然而,这些算法主要面向背景简单的光学图像,当目标背景较为复杂时,模型的性能可能受到抑制。且现有大部分YOLOv5模型主要侧重于检测精度的提高,忽视了模型的检测速度。由于织物疵点种类的多样性、疵点分布的稀疏性和疵点目标的细微性,

上述基于YOLOv5模型的算法较难直接应用于织物疵点检测中。

本文提出了一种基于改进YOLOv5模型的织物疵点检测算法。首先采用K-means++<sup>[17]</sup>聚类方法对所有真实框的宽高进行聚类,以解决原始模型中聚类策略存在锚框过于密集的问题;其次,将Mish<sup>[18]</sup>激活函数用于模型训练过程,以提升模型的非线性拟合能力,并通过调整池化层的位置以提取深层网络中不同尺度的特征。此外,通过对颈部网络中的特征传递路径进行优化,从而减少特征信息丢失;同时增加特征检测层,以提升模型的检测精度。本文对织物疵点数据集进行实验,以验证该算法对织物疵点检测的有效性和鲁棒性。

## 1 YOLOv5 模型介绍

YOLOv5模型按照网络深度和维度的不同,主要分为四种模型:YOLOv5s、YOLOv5m、YOLOv5l和YOLOv5x。其中,YOLOv5s是最轻量的网络模型,具有更好的训练速度与检测速度,因此广泛应用于工业场景中。基于此,本文选用YOLOv5s模型对织物疵点进行检测(下文用“YOLOv5模型”表示)。

YOLOv5模型的网络结构如图1所示,分为输入端、骨干网络、颈部网络和检测端四个部分。输入端采用了马赛克数据增强方法,通过对输入图像进行随机缩放、裁剪、拼接、排布等方式以增加小目标的样本数量,从而提高模型对它的检测精度。骨干网络采用CSPDarknet53结构,由聚焦层(Focus)、卷积层(Conv)、CSP层与空间金字塔池化层<sup>[19]</sup>(Spatial pyramid pooling, SPP)构成。其中,Focus层对输入图像进行隔行采样与拼接,将图像平面的信息转换到通道维度上,以使其在不丢失特征信息的情况下进行下采样,同时降低了模型的计算复杂度,加快其训练速度。CSP层主要由瓶颈结构与卷积层构成,用以增强模型的学习能力,并降低参数量。SPP层采用提取与融合不同尺度的特征图的方法,增强模型的检测精度。颈部网络采用特征金字塔网络<sup>[20]</sup>和路径聚合网络<sup>[21]</sup>的融合结构,以补充深层网络中的特征信息,从而进一步提高模型的检测精度。检测端根据数据集中目标的类别数,完成对三个尺度特征图的最终检测。

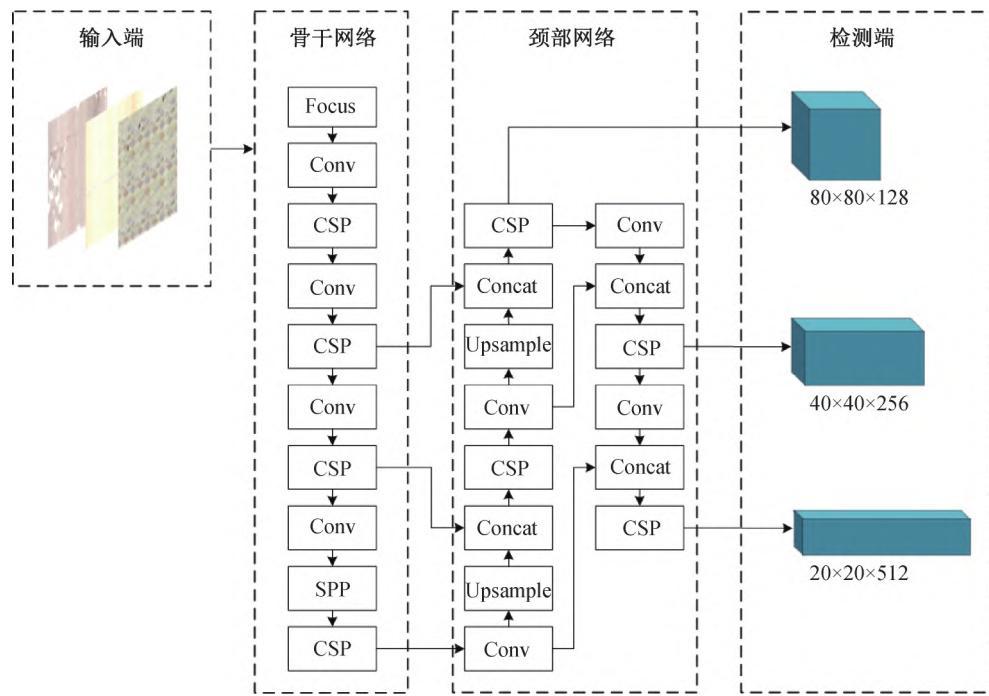


图 1 YOLOv5 模型的网络结构

## 2 YOLOv5 模型改进

### 2.1 聚类策略优化

在目标检测算法中,锚框的宽高常为人为设定,较依赖经验值,影响算法模型的检测结果。基于此,本文在 YOLOv5 模型中采用了经典的 K-means 聚类方法生成锚框尺寸,该方法的实现步骤为:

- 随机选取数据集中  $n$  个目标真实框作为初始锚框。
- 计算出所有真实框与  $n$  个初始锚框之间的交并比(Intersection over Union, IoU),将与锚框 IoU 值最大的真实框归到其所属簇中。根据每个簇中所有真实框宽高的平均值得到最适合的锚框。
- 重复执行步骤 b),直到所有锚框的宽高不再变化或变化幅度很小,即生成最优结果。

YOLOv5 模型对 K-means 聚类方法设置的迭代次数为 30 次,但由于锚框数量  $n$  的取值较大,且初始锚框的选取具有较强的随机性,因而极易导致某一个簇中包含大量的真实框或锚框大小过于接近,从而影响最终的聚类效果。

针对上述问题,本文采用 K-means++聚类方法进行锚框计算。K-means++方法主要在 K-means 聚类方法的基础上,对初始锚框的选取策略进行改进,其实现步骤为:

- 在预设初始锚框时,先随机选定一个真实框的宽高值作为第一个锚框。

b)在同一平面计算所有真实框的宽高与当前已有锚框的 IoU,然后计算每个真实框被选为下一个锚框的概率。其中,真实框与已有锚框 IoU 值越小,则有更大的概率被选为下一个锚框。

c)重复执行步骤 b),直到得出  $n$  个锚框。

d)计算出所有真实框与  $n$  个锚框之间的 IoU,将与锚框 IoU 值最大的真实框归到其所属簇中。通过计算每个簇中所有真实框宽高的平均值,得到最适合的锚框。

e)重复执行步骤 d),直到所有锚框的宽高不再变化或变化幅度很小,即生成最优结果。

因此,使用 K-means++聚类方法,可保证每个锚框的差异性尽可能大,以满足聚类要求,从而提升模型训练时边框回归的速度与精度。

### 2.2 激活函数优化

激活函数在神经网络中起着重要作用。该函数将神经元的输入通过非线性函数运算映射到输出端,从而提升神经网络的非线性建模能力,以应对复杂背景下的数据输入。目前主要的激活函数有: ReLU、Leaky ReLU、Swish、HardSwish、Mish 等。

原 YOLOv5 模型使用了 Leaky ReLU 激活函数。该函数通过保留部分负轴的值,使得负值信息输入不会全部无效,保证一定的梯度流,从而解决了 ReLU 函数在输入为负时梯度消失的问题。缺点是,该函数存在线性特性,当输入的数据信息非常复

杂时,算法模型往往难以收敛,影响模型的检测性能,且该函数的输出曲线也不够平滑,泛化能力较差。基于此,本文考虑采用非线性表达能力更好的Mish函数代替Leaky ReLU函数,其函数表达式可用式(1)表示:

$$y(x) = x \cdot \tanh(\log(1 + e^x)) \quad (1)$$

其中:tanh和log分别表示双曲正切函数和对数运算。Mish函数与Leaky ReLU的输出曲线对比如图2所示。由图2可知,Mish函数相较于Leaky ReLU函数具有更好的平滑性,尤其在输入为0处,这使得模型具备更强的泛化能力和更快的收敛速度。此外,Mish还具备无上界、有下界与非单调的特性。无上界可以防止输出封顶而导致网络饱和,即梯度消失;有下界可以保证一定的正则化效应,避免模型在训练过程中出现过拟合的情况;非单调性可使负值输入也能更新神经元的权重和偏置,从而保证信息的流动,增强模型的可解释能力和梯度流。综上可知,Mish函数在一定程度上克服了传统激活函数的缺点,虽然该函数会增加一定的计算复杂度,但其更好的非线性和平滑性可以使模型具有更强的泛化能力。

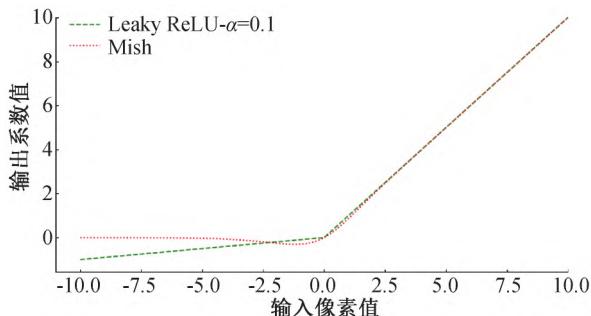


图2 Mish函数与Leaky ReLU函数输出曲线

## 2.3 网络结构优化

### 2.3.1 增加特征检测层

织物疵点数据集中含有大量的小疵点,且部分种类的疵点具高相似性,对模型的检测性能带来较大的影响。基于此,本文在检测端增加特征检测层,以提升模型对不同尺度目标的检测能力,并抑制检测目标的尺度不一致性所带来的不利影响。四个特征检测层对应的特征图分辨率和通道数分别是 $80 \times 80 \times 128$ 、 $40 \times 40 \times 256$ 、 $20 \times 20 \times 384$ 和 $10 \times 10 \times 512$ 。其中,高分辨率特征图没有进行过多下采样,小目标特征信息的整体性较好,因而能显著提高模型对小目标的检测精度。同时,由于低分辨率特征图进行了充分下采样,特征图感受野较大,使得特征信息被有效提取,在此情

况下,大尺度目标的特征更加凸显,利于提高模型对大目标的检测精度。

### 2.3.2 SPP层位置优化

通常情况下,卷积神经网络受其全连接层特性的影响,需统一输入图像的分辨率。然而,该操作通常会对图像进行放缩或裁剪,极易导致目标信息的缺损甚至丢失,从而影响模型的检测性能。大部分算法使用SPP层将特征图分成多个固定尺度的网格,并对所有网格进行特征提取与连接,以实现固定分辨率的输出,满足全连接层的输入要求。

本文改进的YOLOv5模型去除了全连接层,极大降低了模型大小,同时保留了SPP层。因为在模型中,SPP层不仅能输出固定分辨率的特征图,还能提取图像中不同尺度的特征信息,模拟多尺度的特征图输入,防止过拟合的发生。

在原YOLOv5模型中,SPP层设置在骨干网络中,有利于提取有效的目标语义信息,提升模型对物体变形的鲁棒性及检测精度。为了提取更加丰富的语义信息,本文将SPP层设置在颈部网络。在不影响颈部网络对特征图进行上下采样及特征融合的前提下,经过CSP层对目标特征的强化学习,SPP层能最大限度地提取出不同尺度的特征图所包含的目标语义信息,对模型的检测性能有较好的提升。

### 2.3.3 颈部网络结构优化

颈部网络对最终检测结果具有较大影响,主要表现在:a)颈部网络采用特征金字塔网络和路径聚合网络的融合结构,可增强特征图上的有效信息;b)颈部网络将其中的CSP层作为特征输出层,把对应的特征图传递给检测端。随着颈部网络的加深,特征图经过数个卷积层的下采样与上采样,丢失的特征信息增多,从而影响检测端对目标的检测性能。同时,CSP层包含大量堆叠的卷积层加剧了特征信息的丢失。基于此,本文对颈部网络的整体结构进行优化,将其中四个CSP层移出,使其不参与网络对特征图信息的传递,只作为特征检测层,仍与Concat层相连。这种改进既保留了CSP层在深层网络中对特征信息的强化学习,也不会使网络丢失过多的特征信息,同时大幅提升了网络中特征信息的传递,提升了模型的检测精度。此外,CSP层含有的拆分与融合结构,能使梯度路径的数量增加一倍,缓解Concat层直接拼接特征图所带来的梯度信息缺少的问题。

改进后的YOLOv5模型的网络结构如图3所示。

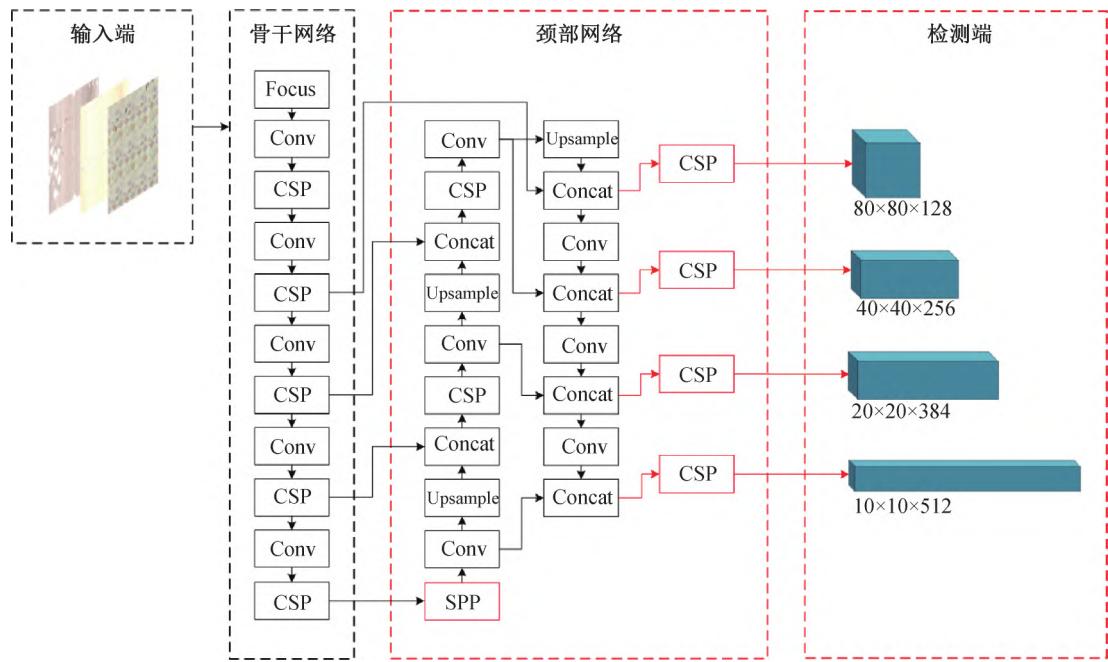


图 3 改进后 YOLOv5 算法网络结构

### 3 模型训练及结果分析

#### 3.1 数据集

本文实验采用公开的阿里云天池织物疵点数据集<sup>[22]</sup>。该数据集包含疵点图像 4371 张, 正常图像 4371 张, 图像分辨率为  $4096 \times 1696$  像素, 共有 15 种疵点类别。由于硬件条件有限, 本文对数据集进行切分处理, 将原图像垂直切分为两张  $2048 \times 1696$  像素的图像, 并对图像进行适当的缩放; 然后, 去掉不包含疵点的图像, 从中选取网折(Fold)、破洞(Hole)、漏印(Miss Print)、色差(Color Shade)、织疵(Flaw)、虫粘(Bug)、褶皱(Scrimp)、缝头印(Sewing Print)、缝头(Sewing)共 9 种疵点类别组成新的疵点数据集, 每个类别的示例图像和图像数量分布分别如图 4、图 5 所示。数据集包含 4972 张图像, 实验过程中随机选取 4337 张图像用于训练, 634 张图像用于测试。

#### 3.2 实验环境与训练参数

本文的实验操作系统为 Ubuntu18.04.1, 显卡为 NVIDIA GeForce RTX 3060, 其中显存为 12 GiB。CPU 为 AMD Ryzen 5 5600X, 系统内存为 32 GiB。IDE 工具采用 PyCharm, 并使用 Python 编程语言和 Pytorch 深度学习框架。在上述实验环境下对图像数据集进行训练, 初始学习率为 0.01, 动量和权重衰减分别设为 0.937 和 0.0005。

#### 3.3 结果和分析

将织物疵点数据集分别通过 YOLOv5 模型和

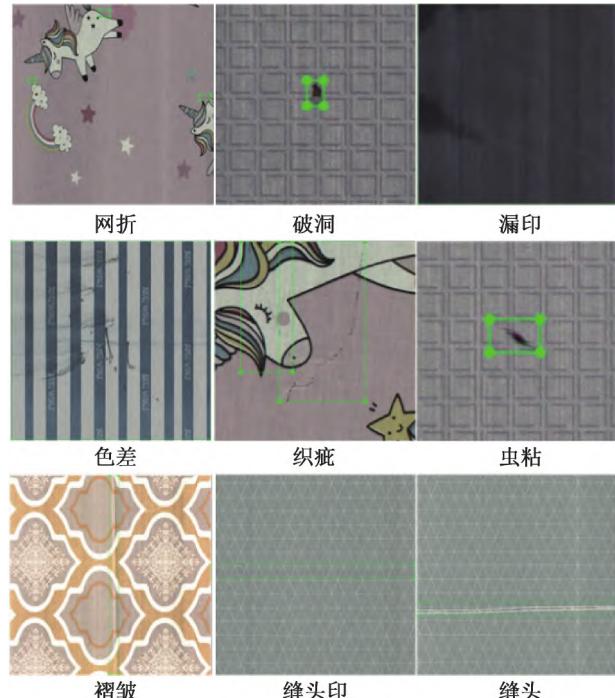


图 4 各类别织物疵点的示例图像

改进后的模型进行训练, 迭代次数设置为 500 次, 训练过程中对应模型的各项损失值对比如图 6 所示。

根据图 6 所知, 改进的 YOLOv5 模型的各项损失拥有更快、更平滑及更稳定的收敛速度。基于 YOLOv5 的改进模型的网络结构及其检测结果如表 1 所示, 检测性能评价指标为 mAP、模型尺寸与检测速度。

由表 1 可得, 改进后的 YOLOv5 模型获得了 76.8% 的 mAP, 比原 YOLOv5 模型提高了 7.7%。

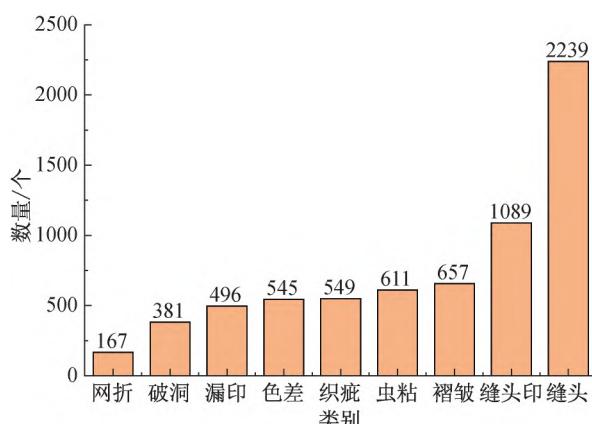


图 5 各类别织物疵点的数量

同时,由于改进算法使用较多的检测层以获取更高的检测精度,导致模型计算量和存储成本的些许上升。从表 1 可知,额外的检测层增加了 10.9 MiB, 相应检测速度下降 20.1 Hz, 即每张图像的检测速度降低了 1.9 ms。

为验证本文所提算法的有效性和鲁棒性,基于同一数据集,本文与其他主流模型在检测性能上进行了对比,结果如表 2 所示。由表 2 可知,相比其他模型,改进后的 YOLOv5 模型在 mAP、模型尺寸和检测速度上均获得了最优结果。其中,mAP 比 YOLOv4 模型提高 11.7%,模型尺寸约为 YOLOv3 模型的 1/9, 检测速度比 YOLOv4 模型提高约 45 Hz。

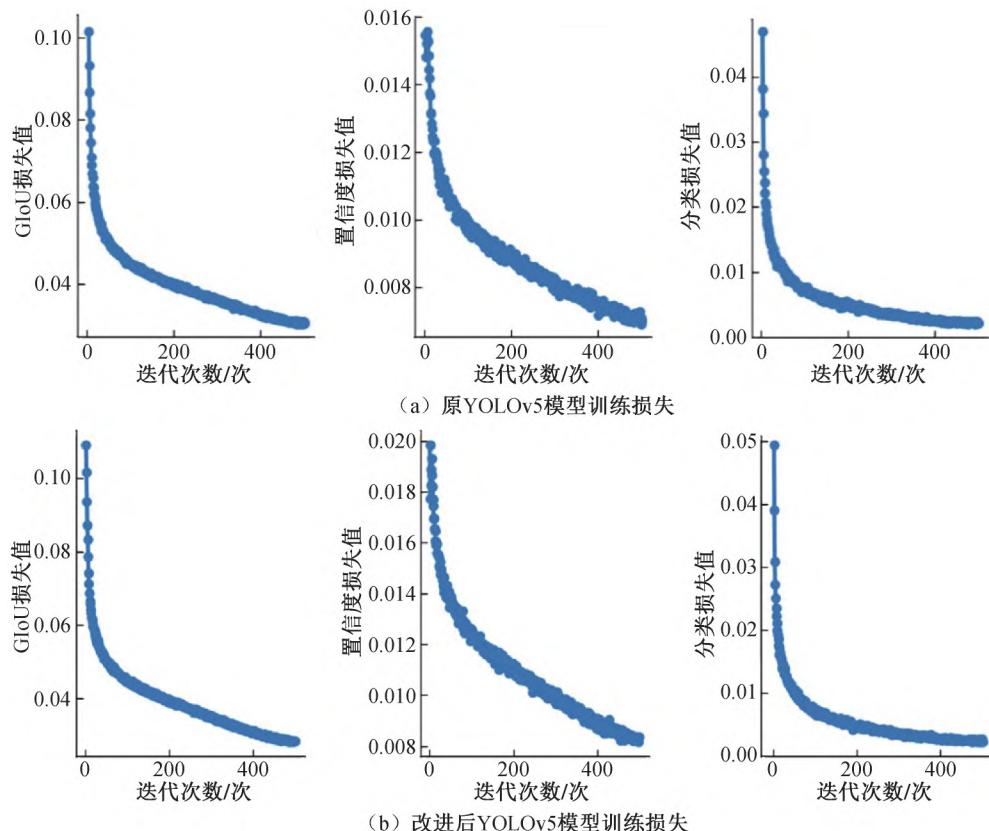


图 6 改进前后 YOLOv5 模型各项训练损失曲线

表 1 改进 YOLOv5 模型的检测指标

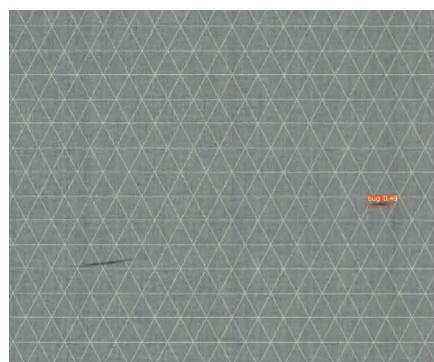
算法	mAP/%	模型尺寸/MiB	检测速度/Hz
YOLOv5	69.1	14.9	113.6
YOLOv5+K means++	69.8(+0.7)	14.9	137.0
YOLOv5+Mish 激活函数+K means++	71.6(+1.8)	14.9	123.5
YOLOv5+Mish 激活函数+4 检测层+K means++	74.4(+2.8)	25.8	94.3
YOLOv5+Mish 激活函数+4 检测层+SPP 调整+K means++	75.8(+1.4)	25.8	96.6
YOLOv5+Mish 激活函数+4 检测层+SPP 调整+ 颈部网络结构优化+K means++	76.8(+1.0)	25.7	93.5

表 2 主流模型检测性能

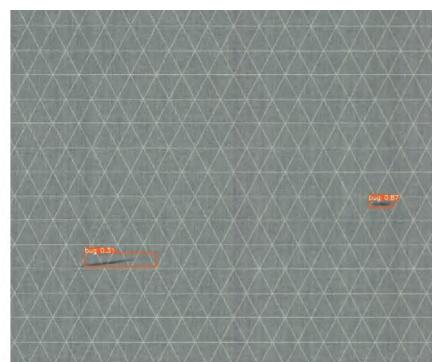
算法	mAP/%	模型尺寸/MiB	检测速度/Hz
Faster R-CNN(ResNet50+FPN)	60.8	315.4	15.4
RetinaNet(ResNet50+FPN)	59.1	246.5	15.1
YOLOv3	57.4	235.2	36.2
YOLOv4	65.1	244.2	48.1
改进的 YOLOv5(本文)	76.8	25.7	93.5

改进的 YOLOv5 模型在织物疵点数据集上的检测结果对比如图 7 所示,其中检测结果由类别名与置信度组成,置信度越高,表示预测框与真实框越接近。通过对比结果可知,原 YOLOv5 模型对包含目标的预测框不具有高置信度,模型的检测性能较差,同时也存在大量漏检。相比之下,改进后的

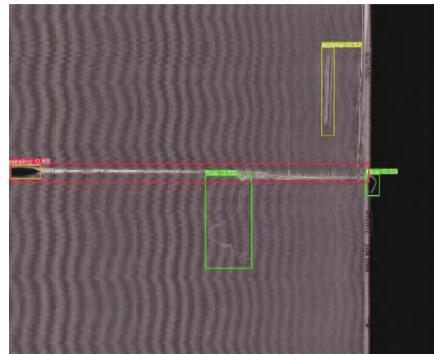
YOLOv5 模型可以很好地改善上述缺陷,提高了检测精度。值得注意的是,本文使用的天池数据集主要以机织物图像为主,而文中所提模型也可扩展到其他类型织物的瑕疵图像检测中,如针织物图像、纱线产品图像等。



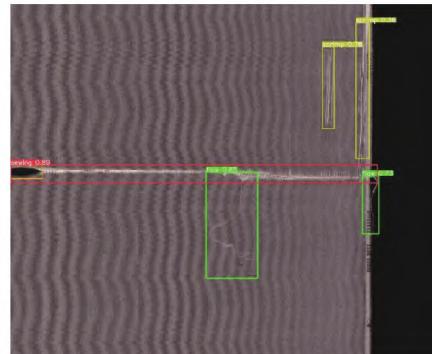
(a) 原 YOLOv5



(b) 改进后 YOLOv5



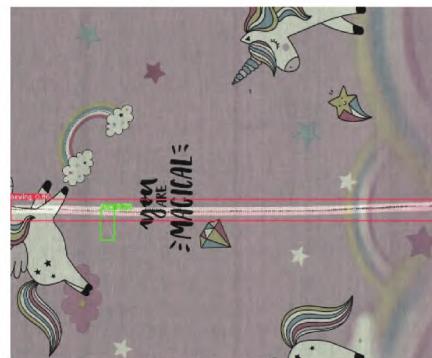
(c) 原 YOLOv5



(d) 改进后 YOLOv5



(e) 原 YOLOv5



(f) 改进后 YOLOv5

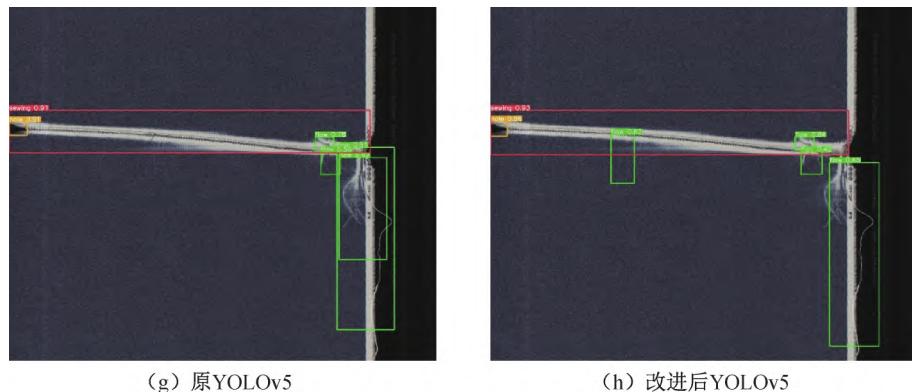


图7 改进前后YOLOv5模型的检测结果示例

## 4 结语

针对织物疵点存在的种类多、密集度高、尺度小等检测难点,本文提出了一种基于改进YOLOv5模型的织物疵点检测算法。该算法通过K-means++方法对所有真实框进行聚类,提高模型训练时的收敛速度;引入Mish激活函数,提高模型非线性表达能力;增加检测层提升模型对多尺度目标的检测性能;调整池化层位置以提取多尺度的特征信息,提升模型的鲁棒性及检测精度;优化颈部网络结构方式以缩短信息的传递路径,进一步提高了模型的检测精度和速度。实验结果表明改进后的YOLOv5模型具有鲁棒性强、检测率高、部署方便等优势,对纺织产业的智能化水平能起到一定的促进作用。

## 参考文献:

- [1] 吕文涛,林琪琪,钟佳莹,等.面向织物疵点检测的图像处理技术研究进展[J].纺织学报,2021,42(11):197-206.
- [2] Xie H S, Zhang Y F, Wu Z S. Fabric defect detection method combining image pyramid and direction template [J]. IEEE Access, 2019, 7: 182320-182334.
- [3] 郑雨婷,王成群,陈亮亮,等.基于卷积神经网络的织物图像识别方法研究进展[J/OL].现代纺织技术.(2022-02-24)[2022-04-28].<https://kns.cnki.net/kcms/detail/33.1249.TS.20220307.1913.010.html>.
- [4] Zhai S P, Shang D R, Wang S H, et al. DF-SSD: An improved SSD object detection algorithm based on DenseNet and feature fusion[J]. IEEE Access, 2020, 8: 24344-24357.
- [5] Fang W, Wang L, Ren P M. Tinier-YOLO: A real-time object detection method for constrained environments [J]. IEEE Access, 2019, 8: 1935-1944.
- [6] Kim C, Kim H M, Lyuh C G, et al. Implementation of yolo-v2 image recognition and other testbenches for a CNN accelerator [C]//2019 IEEE 9th International Conference on Consumer Electronics. Berlin, Germany: IEEE, 2019: 242-247.
- [7] Redmon J, Farhadi A. YOLOv3: An incremental improvement [EB/OL]. (2018-04-08) [2022-02-24]. <https://arxiv.org/abs/1804.02767v1>.
- [8] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal speed and accuracy of object detection [EB/OL]. (2020-4-23)[2022-02-24]. <https://arxiv.org/abs/2004.10934>.
- [9] Jocher G. Yolov5[EB/OL]. (2020-7-24)[2022-02-24]. <https://github.com/ultralytics/yolov5/tree/v2.0>.
- [10] 张静,农昌瑞,杨智勇.基于卷积神经网络的目标检测算法综述[J/OL].兵器装备工程学报.(2022-04-28)[2022-04-29].<http://kns.cnki.net/kcms/detail/50.1213.TJ.20220427.1656.035.html>.
- [11] He Z W, Zhang L. Domain adaptive object detection via asymmetric tri-way faster-RCNN [C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2020: 309-324.
- [12] Cheng T H, Wang X G, Huang L C, et al. Boundary-preserving mask R-CNN[C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2020: 660-676.
- [13] Zhao W D, Huang H C, Li D, et al. Pointer defect detection based on transfer learning and improved cascade-RCNN[J]. Sensors, 2020, 20(17): 4939.
- [14] 王新,王赛.基于改进YOLOv5算法的交警手势识别[J].电子测量技术,2022,45(2):129-134.
- [15] 吕禾丰,陆华才.基于YOLOv5算法的交通标志识别技术研究[J].电子测量与仪器学报,2021,35(10):137-144.
- [16] 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). Montreal, BC, Canada: IEEE, 2021: 2778-2788.
- [17] Lattanzi S, Sohler C. A better k-means++ algorithm via local search [C]//International Conference on Machine Learning. Long Beach, California, USA: PMLR, 2019: 3662-3671.
- [18] Misra D. Mish: A self regularized non-monotonic activation function [EB/OL]. (2019-08-23) [2022-02-24]. <https://arxiv.org/abs/1908.08681>.
- [19] Huang Z C, Wang J L, Fu X S, et al. DC-SPP-YOLO: Dense connection and spatial pyramid pooling based YOLO for object detection[J]. Information Sciences, 2020, 522: 241-258.
- [20] Ghiasi G, Lin T Y, Le Q V. NAS-FPN: Learning scalable feature pyramid architecture for object detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019: 7029-7038.
- [21] Zhou L M, Rao X H, Li Y H, et al. A Lightweight object detection method in aerial images based on dense feature fusion path aggregation network [J]. ISPRS International Journal of Geo-Information, 2022, 11 (3): 189.
- [22] Yu X Y, Lyu W T, Zhou D, et al. ES-Net: Efficient scale-aware network for tiny defect detection[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-14.

(责任编辑:康 锋)