

DOI:10.19651/j.cnki.emt.2108370

基于YOLOv5改进算法的印花图案疵点检测

颜学坤 楚建安

(西安工程大学电子信息学院 西安 710699)

摘要: 针对圆网印花图案疵点检测问题,采用了一种基于YOLOv5改进算法模型来检测印花图案的疵点。根据实际的情况对YOLOv5模型网络结构进行了更改,首先,对YOLOv5网络的骨干部分进行优化改进,引入了注意力机制模块,对输入图片的通道注意力和空间注意分别提取特征。其次,针对印花疵点目标较小的情况对网络的检测层结构进行了修改。实验结果显示,改进的YOLOv5检测算法精确率提升了14.4%,检测速度提升了7.6 fps,达到了43.1 fps满足实时检测要求。

关键词: 印花疵点;YOLOv5;注意力机制

中图分类号: TP38 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Printing pattern defect detection based on improved YOLOv5 algorithm

Yan Xuekun Chu Jian'an

(School of Electronics and Information, Xi'an Polytechnic University, Xi'an 710699, China)

Abstract: Aiming at the problem of defect detection of circular screen printing pattern, an improved algorithm model based on YOLOv5 is used to detect the defect of printing pattern. The network structure of YOLOv5 model is changed according to the actual situation. Firstly, the backbone of YOLOv5 network is optimized and improved, and the attention mechanism module is introduced to extract the features of channel attention and spatial attention of input pictures respectively. Secondly, aiming at the small target of printing defects, the detection layer structure of the network is modified. The experimental results show that the accuracy of the improved YOLOv5 detection algorithm is improved by 14.4%, and the detection speed is also improved, reaching 43.1 fps, which meets the requirements of real-time detection.

Keywords: printing defects; YOLOv5; attention mechanism

0 引言

随着国内外企业数字化的发展,许多传统企业需要进行数字化转型。目前,国内的许多印染企业在印花图像质量的检测上依然依靠人眼辨识,这种人工识别的方法随着观察的时间增加,人眼过度疲劳,导致准确率会降低,最终影响生产效率。文献[1-2]研究了机器视觉技术在对产品质量进行检测方面的应用,其自动、准确且能够快速识别的特点对于保证印花产品的质量具有极其重要的意义和实际价值。

近年来,有很多方法应用在织物和印花图案检测方面,文献[3]提出了基于局部片近似的织物疵点自动分割方法,该方法采用无监督方案,无需参考图像或其他先验信息;其次在对目标特征研究的基础上,文献[4]提出了一种基于主结构提取与图像签名算法的纹理织物缺陷检测方法,该方法采用总变差模型提取主结构,通过图像签名算法检测疵

点,最后使用自适应阈值分割得到疵点区域。由于传统算法的复杂性,以上的方法在检测速度和精度方面都难以满足实际生产过程的需求。随着机器学习等新技术的快速发展,文献[5]提出了基于卷积神经网络(convolutional neural networks, CNN)的深度学习在布匹缺陷检测中的应用。例如文献[6]提出的 AlexNet 网络,首先使用 GPU 加速训练,并且使用 ReLU 激活函数,相比于 Sigmoid、Tanh 等激活函数更容易求导;文献[7]提出的 VGG(visual geometry group)网络通过使用多个 3×3 的小卷积核代替大尺度卷积核,可以减少参数量;文献[8]提出的 GoogLeNet 则引入 inception 结构,融合了不同的尺度特征信息,识别到更多的细节,进一步提升了识别效果;ResNet 加深了网络结构,提出 residual 模块,并使用批量归一化加速训练;文献[9-11]提出的 YOLO 系列也因其检测速度的优势在很多领域得以应用。根据以上背景,本文对 YOLOv5 网络模型进行改

收稿日期:2021-11-17

进后对印花图案疵点进行识别。首先,引入注意力机制模块替换原 YOLOv5 主干网络中的部分特征提取模块,对输入图片的通道注意力和空间注意分别提取特征。其次,针对印花疵点目标较小的情况在保证检测精度的前提下将网络的检测头减少到两个,并通过实验调整检测头的尺寸大小。

1 圆网印花图像缺陷分析

圆网印花技术是利用圆网内的刮刀作用浆料,使浆料在刮刀的压力下印制到织物上去。因此圆网印花机刮刀上如果有小缺口或粘附有其他杂质会产生刮色不均,造成条状刀线;其次,印花网的网孔会被一些杂质堵塞,阻碍浆料向外渗透,造成局部无浆或少浆的情况;还会由于对花不准造成花版错位的情况。除了以上缺陷还存在例如印花水渍、皱条等其他污斑。图 1 所示为部分圆网印花缺陷示例,如图 1(a)所示为长条状疵点,图 1(b)所示为点状疵点。

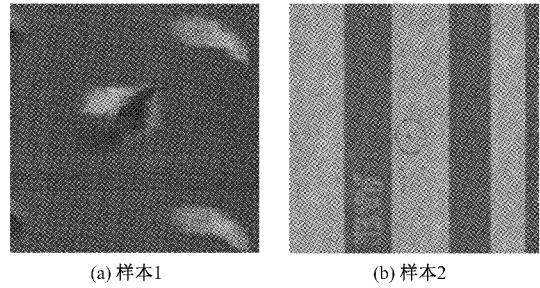


图 1 印花疵点

2 基于 YOLOv5 目标检测算法的改进

YOLOv5 是 YOLO 系列不断创新的目标检测网络,其网络结构如图 2 所示,其中 CBL 模块由卷积(convolution),批量归一化(batch normalization, BN),激活函数 Leaky ReLU 组成。Res unit 表示残差组件,残差组件的数量影响

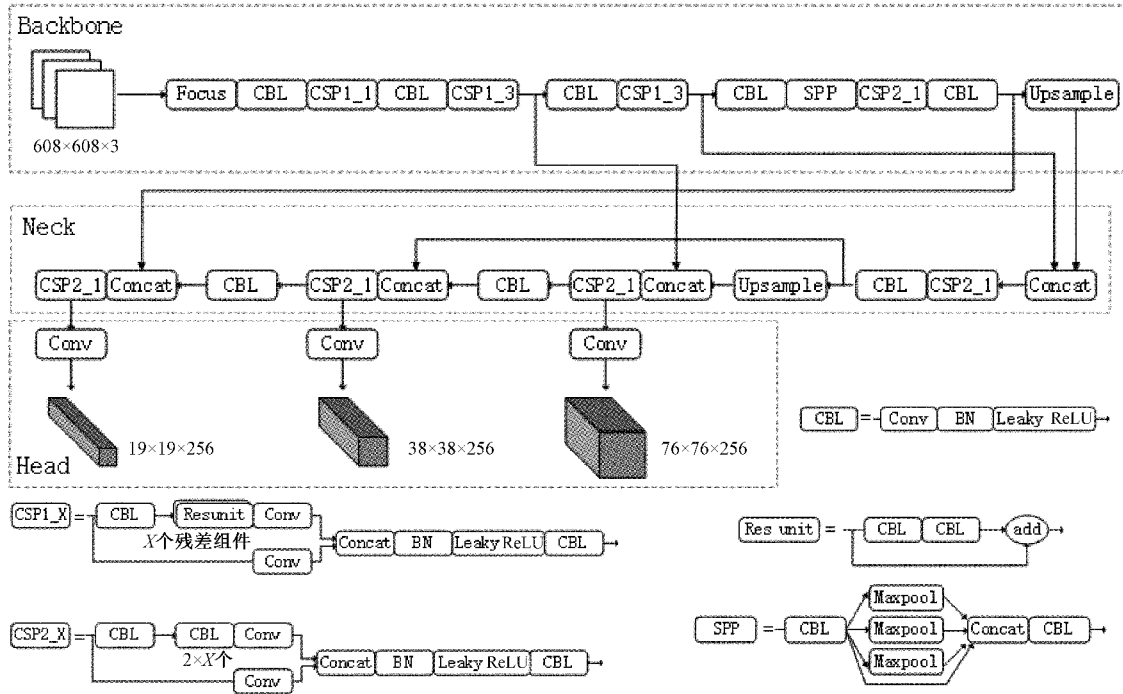


图 2 YOLOv5 网络结构

着网络模型的大小。相比于 YOLOv4,网络灵活性和检测速度有所提高。目前,YOLOv5 根据网络的宽度和深度的不同,分为 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 共 4 个版本。为了平衡检测速度和精度,本文选取 YOLOv5s 版本进行作为基础网络进行改进。

2.1 YOLOv5 网络架构与组件

Backbone:在骨干网络部分,首先对输入样本进行数据增强操作,包括 Mosaic 增强、随机缩放、裁剪等方式,随后样本进入 Focus 结构进行切片操作,如图 3 所示,在横向和纵向上每隔一个像素点去一个值,则输出的图片长宽

为输入的 1/2,通道数为原来的 4 倍,之后经过一系列卷积、池化、跨阶段局部网络(cross stage partial network, CSPNet)^[12]和特征金字塔池化(spatial pyramid pooling, SPP)^[13]结构,得到特征图。

Neck:在特征融合部分主要通过特征金字塔网络(feature pyramid networks, FPN)^[14]和路径聚合网络(perceptual adversarial network, PAN)^[15]结构完成,高层特征和低层特征的融合,使得网络可以获得更多的原图信息,对目标识别的更加精确。

Head:在检测层部分,针对检测目标的大小不同,设计

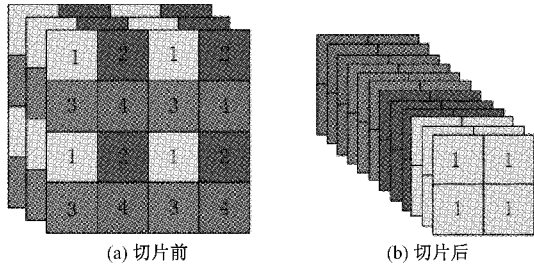


图3 切片操作

了3个不同尺寸的特征图,分别为 19×19 、 38×38 、和 76×76 来检测大、中、小目标。最后通过非极大抑制(non-maximum suppression, NMS)方法保留置信度最高的预测框信息,完成目标检测过程。

2.2 改进的YOLOv5算法模型

印花图案缺陷往往目标较小,YOLOv5模型对于小目标检测精度无法满足现实需求,存在检测精度低和漏检等情况。本文针对印花缺陷较小的特点,引入卷积块注意力模块(convolutional block attention module, CBAM)机制^[16],提高对小目标的检测精度。其次,根据实际检测情况,对多尺度检测层的特征尺度进行了修改,使得模型更好地适应小目标检测。

1)改进的Backbone网络结构

本文提出了使用CBAM模块替换骨干网络中的部分CSP模块,提高网络的特征提取能力,对通道特征具有强化重要信息,抑制无关信息的作用,可以使网络提取更加关键的特征信息。

CBAM模块如图4所示,主要由两个模块组成,分别为通道注意力模块(channel attention module, CAM)和空间注意力模块(spatial attention module, SAM),它们分别在通道和空间上进行特征的提取。

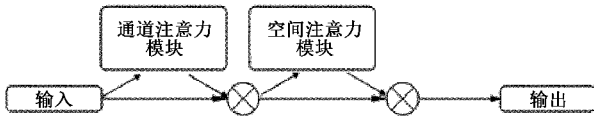


图4 CBAM网络结构

CAM结构如图5所示,输入特征图分别经过全局最大池化和全局平均池化得到特征向量,分别经过全连接层后相加得到尺寸为 $1 \times 1 \times C$ 的特征向量,将输入特征图与该向量相乘即为该模块的输出特征图。

SAM结构如图6所示,对经过CAM后得到的特征图进行基于通道的全局最大池化和全局平均池化后在通道方向上进行拼接,然后经过卷积层并使用Sigmoid函数激活,将得到的结果和经过CAM得到的结果相乘,即为CBAM网络结构的输出。

2)改进的检测层结构

如图7所示,为改进后的网络结构,本文保留了YOLOv5中使用的FPN和PAN结构,但是改进了检测层

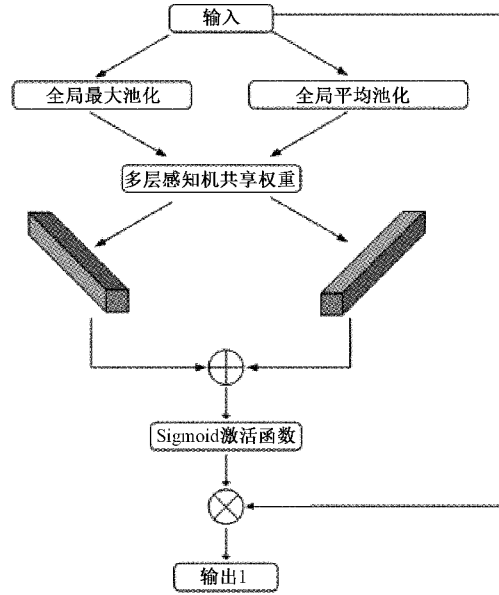


图5 CAM网络结构

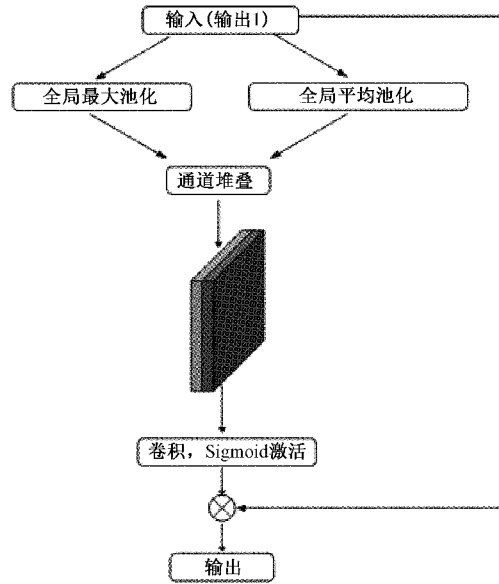


图6 SAM网络结构

部分,输入图片经过特征提取网络和上采样以及拼接操作最终得到 120×120 尺度和 60×60 尺度的检测头,提升网络对小目标识别的准确率。

2.3 损失函数

本实验所使用的损失函数与YOLOv5相同,分为分类损失(classification loss),位置损失(localization loss)和置信度损失(confidence loss),且使用二元交叉熵损失函数计算方法计算 classification loss 和 confidence loss,使用GIoU (generalized intersection over union) Loss 作为 bounding box 回归的损失,即总损失为:

$$Loss = \lambda_1 L_{cla} + \lambda_2 L_{loc} + \lambda_3 L_{conf} \quad (1)$$

其中, $\lambda_1, \lambda_2, \lambda_3$ 为平衡系数。

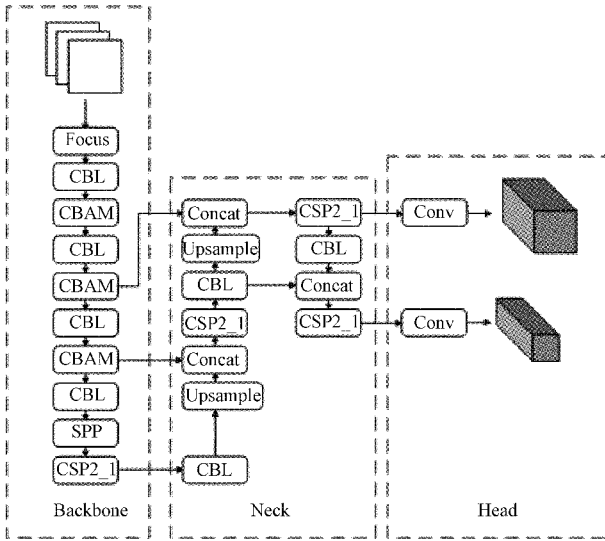


图 7 改进的 YOLOv5 模型简图

1) 二元交叉熵损失

$$y_i = \text{Sigmoid}(x_i) = \frac{1}{1 + e^{-x_i}} \quad (2)$$

$$L = - \sum_{i=1}^N y_i^* \log(y_i) + (1 - y_i^*) \log(1 - y_i) \quad (3)$$

式中： x_i 表示当前类别预测值， y_i 表示经过 Sigmoid 函数后得到的当前类别概率， N 表示类别总数， y_i^* 表示当前类别的真实值(0 或 1)。

2) GIoU Loss

GIoU 是通过计算预测框与实际标注边框的交并比来计算损失。如图 8 所示，假设实际框的面积为 A ，预测框的面积为 B ，包围实际框和预测框的矩形的最小面积为 C ，则：

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (4)$$

$$GIoU = IoU - \frac{|C \setminus (A \cup B)|}{|C|} \quad (5)$$

其中，“ $C \setminus (A \cup B)$ ”表示从面积 C 中去掉 A 与 B 的并积部分。考虑两个框重叠的极端情况可知 GIoU 的取值范围为 $[-1, 1]$ 。GIoU 损失为：

$$GIoU_Loss = 1 - GIoU \quad (6)$$

即两个框重叠区域越大，损失越小。

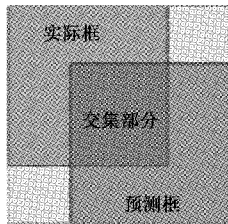


图 8 交并比计算图

2.4 目标框回归

在 YOLOv5 网络中目标框回归是属于偏移量的回归，

使得计算量大大降低，提升了网络的训练速度。如图 9 所示，实线框为预测框，虚线为先验框，预先设定不同尺度的先验框即 anchor box，网络对输入进行预测生成预测框，预测框的计算公式如下：

$$b_x = \sigma(t_x) \times 2 - 0.5 + c_x \quad (7)$$

$$b_y = \sigma(t_y) \times 2 - 0.5 + c_y \quad (8)$$

$$b_w = p_w (\sigma(t_w) \times 2)^2 \quad (9)$$

$$b_h = p_h (\sigma(t_h) \times 2)^2 \quad (10)$$

式中： t_x, t_y, t_w, t_h 为网络输出的预测框中心点坐标以及宽高的偏移值，函数 $\sigma(x)$ 选取的是 Sigmoid 函数，使得预测偏移值在 $(0, 1)$ 之间； p_w, p_h 为 anchor 给出的目标宽高的初始值； b_x, b_y, b_w, b_h 为预测框的中心坐标以及宽度和高度。

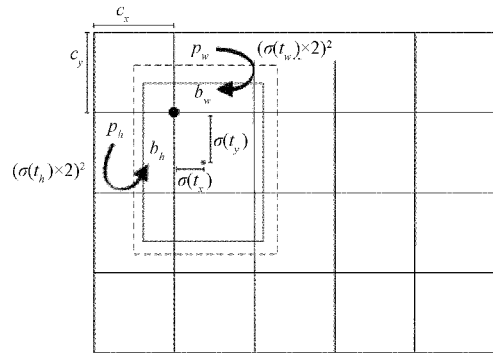


图 9 预测框计算示意图

3 实验结果与分析

3.1 实验环境

本文的实验环境搭建在服务器工作站上，硬件配置为：Intel(R)Core(TM)i7 10750H CPU@2.60 GHz, Nvidia Geforce RTX 2060 显卡, 16 GB 内存, 1 TB 硬盘。软件配置情况为 Windows10 版本, CUDNN 7.6.5 版本, CUDA 10.2 版本。深度学习框架平台为 Pytorch, Python 版本 3.7。

3.2 实验数据

本次实验使用的数据来自于阿里云公开的布匹瑕疵检测数据集以及陕西兴洲纺织科技有限公司采集的部分数据，共 1 260 张图片，其中 883 张用于训练，377 张用于测试。图片尺寸统一裁剪为 608×608 大小且每张图片至少存在一处瑕疵。由于训练数据集较小，网络中使用的 Mosaic 数据增强，每次给网络输入图片时先随机选取的 4 张图片进行随机的缩放、裁剪后进行拼接，然后开始训练，新的图片相比于原来的图片可识别信息更加丰富，特别是随机缩放增加了很多小目标，在一定程度上提升了网络的鲁棒性。

3.3 网络训练

在网络训练过程中,将输入图片前向传播,得到预测值和真实值的误差,通过随机梯度下降(stochastic gradient descent, SGD)不断对误差进行调整,使得整个网络模型得到优化。本实验通过引入CBAM机制并根据实际检测对象调整检测层的尺寸,最终在印花图案质量检测上取得了显著的效果。

本实验训练时使用的主要超参数如表1所示。

表1 主要超参数

Learning rate	Momentum	Weight decay	Optimizer	Epoch
0.01	0.937	0.0005	SGD	300

在训练时记录下模型的边界框损失、置信度损失和分类损失,如图10(a)所示是模型改进前的损失曲线,可以看出改进前损失曲线都波动较大,且存在过拟合现象;图10(b)是模型改进后的损失,边界框回归损失和分类损失明显收敛更快,且置信度损失相比改进前效果也更加稳定。由此可知,改进后的模型在对疵点的定位和分类方面效果更加显著,这对于实际应用具有重要意义。

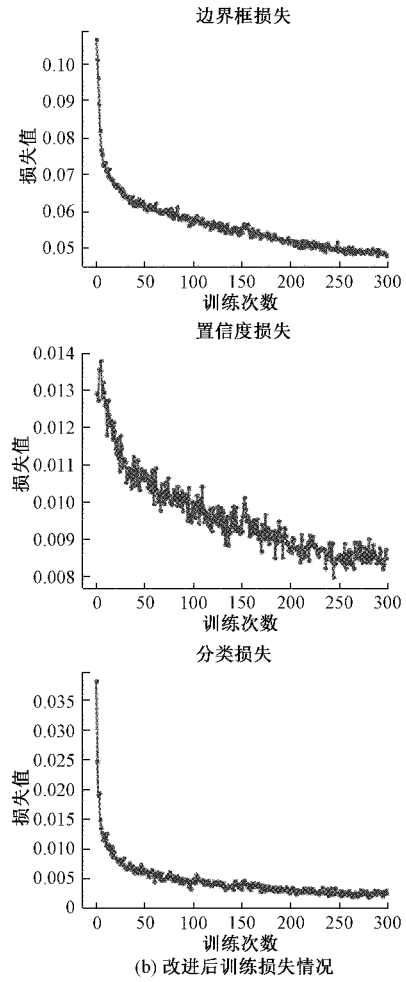
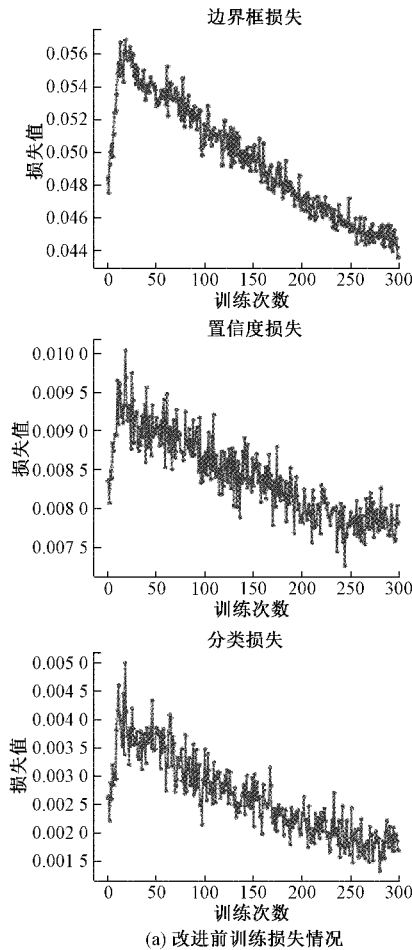


图10 YOLOv5改进前后训练结果对比

3.4 模型评估对比

模型的评估指标对于反应模型的真实性能有着很关键的作用,本实验选取的评估指标是精确率(Precision)和召回率(Recall),因为这两个指标可以清晰地反应出网络对检测对象的误检和漏检情况,其公式如下:

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

其中,TP表示正样本预测为正样本的数量,FP表示负样本预测为正样本的数量,FN表示正样本预测为负样本的数量。

如表2所示为改进前后的模型性能对比,本实验改进的模型参数降低了 2.19×10^6 个参数量,检测精度和召回率分别提升了14.4%和6.2%,检测速度提高了7.6 fps,能够满足印花实时检测的要求。因此,本实验提出的改进YOLOv5网络模型对于圆网印花疵点检测是有效的。

图11为YOLOv5改进前后的检测效果对比,由图11(a)和(b)可以看出,改进后的YOLOv5模型对于小

表 2 改进前后模型性能对比

模型	参数 个数	检测精度/ %	召回率/ %	检测速度/ fps
改进前	9.28×10^6	72.1	85.2	35.5
改进后	7.09×10^6	86.5	91.4	43.1

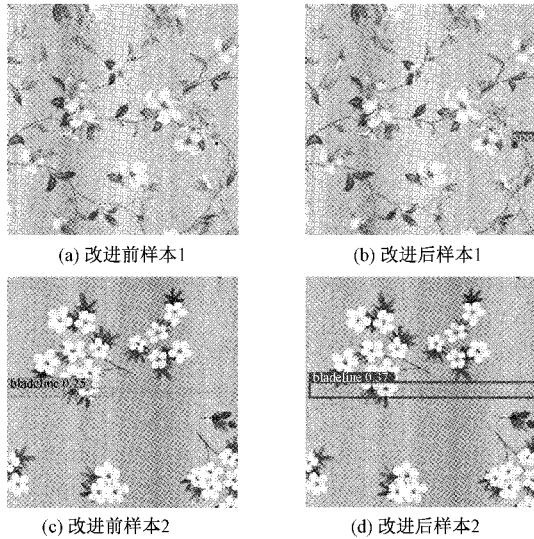


图 11 YOLOv5 改进前后检测效果对比

目标漏检情况有所改善;同时,如图 11(c)和(d)所示,改进后的 YOLOv5 模型对于印花斑点的定位更加准确,且置信度也有所提高。

3.5 不同检测算法对比

本文选取了 3 种不同的算法和本实验的算法进行对比实验,对本文改进后的 YOLOv5 模型的检测效果进行充分评估。为了满足实际生产过程中对识别精度和速度的要求,本文选取平均精确度和检测速度作为该实验的评估指标,实验结果如表 3 所示。

表 3 不同检测算法对比

算法名称	平均精确度/%	检测速度/fps
Resnet50-SSD	61.2	52.4
Resnet50-Faster R-CNN	74.5	21.8
YOLOv5	73.8	35.5
改进的 YOLOv5	89.2	43.1

由表 3 可知,基于 Resnet50 的 SSD 检测算法的检测速度高达 52 fps,但平均精确度最低。而 Faster R-CNN 和 YOLOv5 检测算法在平均精确度方面相差不大,但由于 YOLO 系列使用的是 one-stage 方法直接回归物体的类别概率和位置坐标,比 two-stage 方法先产生候选区域后进行训练要快。本文提出的改进 YOLOv5 检测算法比 Faster R-CNN 的平均精确度高出 14.7%,检测速度虽然比 SSD 检测算法慢 9 fps 左右,但仍然可以满足圆网印花

机检测的需求。

此外,为了更加直观地体现不同算法之间的检测效果,对比了不同算法对不同尺寸斑点检测的情况,如图 12 所示,从左到右分别为改进的 YOLOv5、Resnet50-Faster R-CNN、和 Resnet50-SSD 的检测结果,可以看出后两种算法在对不同尺度目标检测时存在漏检、错检的情况,改进后的检测模型在这方面有了很大的改善,检测精度有了很大地提高。

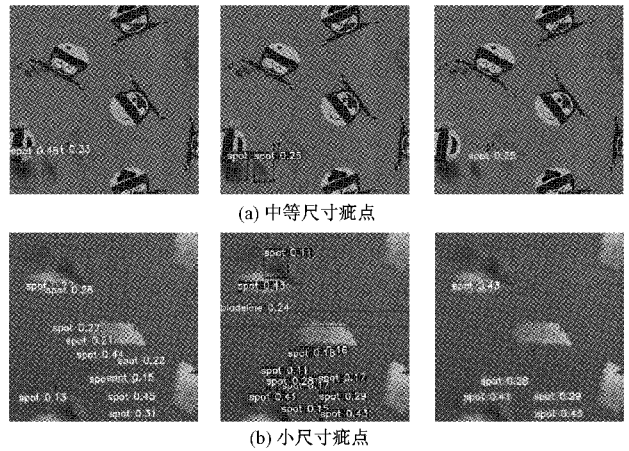


图 12 不同尺度斑点模型检测结果对比

4 结 论

本实验针对圆网印花机的印花图案斑点检测任务,提出了一种基于 YOLOv5 改进的网络模型,引入了 CBAM 模块改进骨干网络,并且针对印花图案的小目标斑点检测问题,优化了检测层,将原有的 3 种不同的尺度改为 120×120 和 60×60 尺度。实验结果显示,改进后的 YOLOv5 模型的参数量更少,同时检测精度和检测速度也有明显提升。由于现有数据集的有限性,以后的研究计划需要加强数据集的采集,并且进一步提高网络的检测精度,在实际圆网印花机上部署时,还需要提高网络的抗干扰能力。

参考文献

- [1] 刘泽,王崑,王平. 钢轨表面缺陷检测机器视觉系统的设计[J]. 电子测量与仪器学报, 2010, 119(11): 1012-1017.
- [2] 蓝金辉,王迪,申小盼. 卷积神经网络在视觉图像检测的研究进展[J]. 仪器仪表学报, 2020, 41(4): 167-182.
- [3] ZHOU J, WANG J. Unsupervised fabric defect segmentation using local patch approximation [J]. Journal of the Textile Institute, 2016, 107(6): 1-10.
- [4] 王震,景军锋,张缓缓,等. 基于主结构提取与签名算法的织物缺陷检测[J]. 电子测量与仪器学报, 2019, 220(4): 36-41.
- [5] 赵志勇,叶林,桑红石,等. 深度学习在布匹缺陷检测中

- 的应用[J]. 国外电子测量技术, 2019, 297(8): 110-116.
- [6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25: 1097-1105.
- [7] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. International Conference on Learning Representations, 2015, 1: 1-14.
- [8] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, 1: 1-9.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 1: 779-788.
- [10] REDMON J, FARHADI A. YOLO 9000: Better, faster, stronger [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, 1: 7263-7271.
- [11] FARHADI A, REDMON J. Yolov3: An incremental improvement [C]. Computer Vision and Pattern Recognition, 2018, 1: 1804.
- [12] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, 1: 390-391.
- [13] HE K, ZHANG X, REN S, 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.
- [14] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, 1: 2117-2125.
- [15] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 1: 8759-8768.
- [16] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision(ECCV), 2018, 1: 3-19.

作者简介

颜学坤, 硕士研究生, 主要研究方向为人工智能与机器视觉技术。

E-mail: yxk0529@163.com

楚建安, 副教授, 高级工程师, 主要研究方向为印染机械、纺织行业。

E-mail: 13809153570@163.com