

基于改进 Yolov5 的花色布匹瑕疵检测方法

时造雄, 茅正冲

(江南大学 物联网工程学院, 江苏 无锡 214000)

摘要: 花色布匹的瑕疵检测是纺织工业中必不可少的环节, 实现快速、准确的花色布匹瑕疵检测对于提高生产效率具有重要意义; 针对花色布匹瑕疵检测中大部分瑕疵目标较小、种类分布不均、部分瑕疵长宽比较为极端以及瑕疵与背景易混淆的检测难点, 提出了一种基于 YOLOv5 网络改进的算法模型 DD-YOLOv5; 在骨干网络中采用上下文变换器网络 (CoTNet), 增强视觉表示能力; 在颈部网络中引入卷积注意力模块 (CBAM), 使网络学会关注重点信息; 在检测环节增加了一个高分辨率的检测头, 加强对小目标的检测; 并且使用 α -IoU 代替原网络中 G-IoU 方法; 经实验证明, 改进后的算法在花色布匹瑕疵数据集平均精度均值上 (mAP) 达到了较原生算法相比提升了 8.1%, 检测速度也达到了 73.6 Hz。

关键词: 瑕疵检测; 深度学习; CoTNet; 注意力机制; 交并比

Fancy Cloth Defect Detection Method Based on Improved Yolov5

SHI Zaoxiong, MAO Zhengchong

(School of Internet of Things, Jiangnan University, Wuxi 214000, China)

Abstract: The defect detection of fancy cloth is an indispensable link in the textile industry. It is of great significance to realize the rapid and accurate defect detection of the fancy cloth to improve the production efficiency. In order to solve the defect detection difficulties of small defect targets, uneven distribution of types, extreme length and width of some defects, and easy confusion with background, an improved algorithm model DD-YOLOv5 based on YOLOv5 network was proposed. Contextual transformer networks (CoTNet) are used in the backbone to enhance the visual presentation capabilities; By introducing a convolutional block attention module (CBAM) into the neck network, the key information is concerned in the network. A high resolution detector is added in the detection link to strengthen the detection of small targets. In addition, the α -IoU method is used to replace original G-IoU method. The experimental results show that compared with the original algorithm, the mean average precision (mAP) of the improved algorithm increases by 8.1%, the detection speed also reaches up to 73.6 Hz, the proposed method has a good performance in the fancy cloth defect detection.

Keywords: defect detection; deep learning; CoTNet; attention mechanism; IoU

0 引言

中国是世界上最大的布匹生产国, 布匹瑕疵的识别与检测在生产过程中是十分重要的一个环节, 传统的布匹瑕疵检测与识别是通过人工的方式进行的, 存在受主观影响较大、检测效率低下等问题。近年来, 随着计算机算力的突破, 深度学习技术发展迅猛, 在工业生产的瑕疵检测领域已经产生了越来越多的应用。

目前主流的目标检测算法分为单阶段和双阶段两种类型, 单阶段算法以 SSD (single shot multibox detector) 和 YOLO (you only look once) 为典型, 而双阶段则以 Faster R-CNN (faster region-CNN) 网络为主^[1]。

在针对花色布匹的瑕疵检测上, 文献 [2] 在 Faster R-CNN 网络模型的基础上进行改进, 提升了对较小目标瑕疵的检出能力及复杂背景下的瑕疵检出能力, 但仍存在

速度不足的缺陷^[2]; 文献 [3] 提出了双路高分辨转换网络的花色布匹瑕疵检测方法, 消除复杂花色背景的影响并且解决了少数类瑕疵样本准确率不高的问题, 该模型存在对小目标瑕疵的检测优化不足的明显短板, 整体精度较低^[3]; 文献 [4] 提出基于级联卷积神经网络的复杂花色布匹瑕疵检测算法, 该算法一定程度上改善了对瑕疵检测的速度和精度要求^[4]; 文献 [5] 提出基于深度卷积神经网络的分类不平衡纹理布匹瑕疵检测方法, 能有效减轻瑕疵种类分布不均对模型的影响, 且降低了模型对部分瑕疵较少的不敏感性, 但该模型的精度表现一般, 仍有待改善^[5]。综上所述之前国内外得研究并不能满足精度和速度的双重要求。

在以上研究的基础上, 本文主要在 YOLOv5 网络基础上进行改进, 提出了 DD-YOLOv5 (defect detection YOLOv5)

收稿日期: 2022-11-05; 修回日期: 2022-12-25。

作者简介: 时造雄(1997-), 男, 江苏盐城人, 硕士研究生, 主要从事目标检测和深度学习方向的研究。

引用格式: 时造雄, 茅正冲. 基于改进 Yolov5 的花色布匹瑕疵检测方法[J]. 计算机测量与控制, 2023, 31(4): 56-62.

瑕疵检测模型, 首先将 CoTNet 与 YOLOv5 模型相结合, 用于指导动态注意力矩阵的学习, 增强视觉表示能力; 其次引入 CBAM (convolutional block attention module) 注意力机制, 让网络学会关注重点信息; 接着使用 α -IoU 损失函数代替原网络中 G-IoU 损失函数, 提高边界框的回归精度; 最后增加检测头, 提升网络对小目标的检测能力。经实验证明, 改进后的网络能同时解决上述文章中提出的花色布匹瑕疵检测的如瑕疵目标小、种类分布不均、长宽比较为极端以及易与背景混淆等检测难点, 且在速度上优于其它算法, 表明该方法在花色布匹瑕疵检测任务中具有良好的检测性能。

1 系统结构及原理

1.1 原生 YOLOv5 模型

YOLOv5 系列网络整体沿用系列网络整体布局, 主要由 Backbone (主干)、Neck (颈部)、Head (头部) 3 个部分组成, 其结构如图 1 所示。

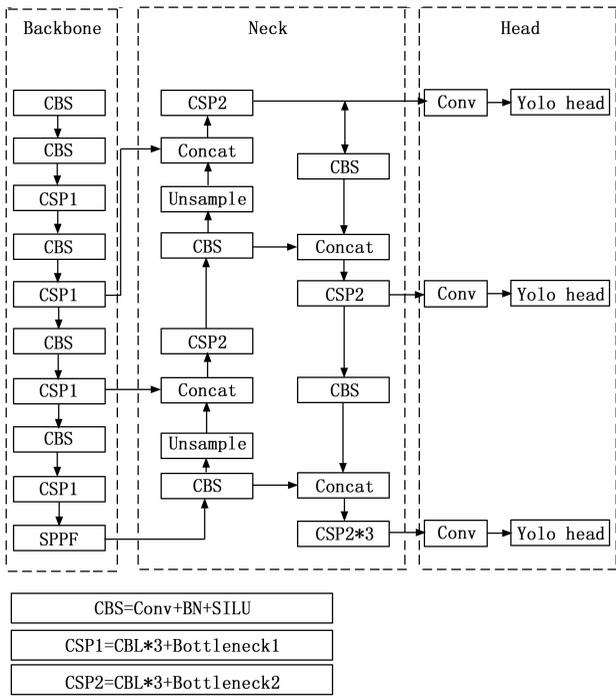


图 1 YOLOv5 网络结构图

1.2 原生算法系统结构分析

YOLOv5 目标检测网络的主干部分主要用于提取图像浅层和深层的各项特征并整合。

主干部分中 CSP1 模块主要借鉴了残差网络的思想, 由残差映射替换了以往的基础映射, 使得深层次的网络模型达到更好的训练效果, 在残差网络出现之前, 学术界往往认为网络层数越深则获取信息越多, 模型泛化效果越好。然而后续的研究表明当网络的层数加深时, 模型的准确率反而有所降低。这是由于反向传播过程中的梯度爆炸和梯

度消失所造成。也就是说, 网络越深, 模型越难优化。为了能让深层次的网络模型达到更好的训练效果, 残差网络中提出由残差映射替换了以往的基础映射。

SPPF 模块主要作用是对高层特征进行提取并融合, 在融合的过程中多次运用最大池化, 尽可能多地去提取高层次的语义特征。

颈部网络的主要作用是将主干网络提取的浅层和深层信息加强并融合。该部分采用的是特征金字塔网络 (FPN, feature pyramid network) 结合路径聚合网络 (PAN, path aggregation network) 的设计^[6-7]。金字塔结构是一种自顶向下的侧边连接, 在所有尺度上构建出高级语义特征图的结构, 该结构经过多层的网络后, 底层的目标信息已经非常模糊了, 因此网络在颈部又结合了路径聚合网络, 加入了自底向上的路线, 弥补并加强了定位信息^[8]。

检测头部分主要包括不同的损失函数。其主要目标是通过降低预测框和真实框的损失函数来获得最佳的目标检测框^[9]。

2 改进措施

2.1 改进后系统原理图

本文提出的 DD-YOLOv5 模型结构如图 2 所示。

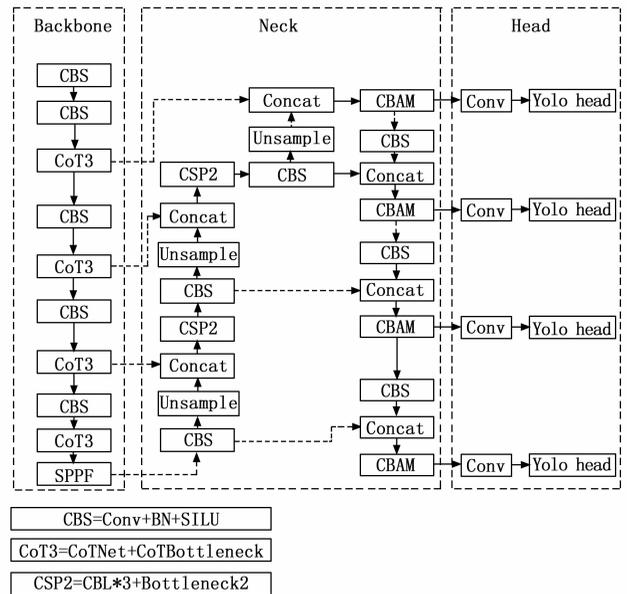


图 2 DD-YOLOv5 网络结构图

通过图 2 可知, 与原生网络相比, 本文网络主要改进点为: 首先用 CoT3 模块代替原有 CSP1 模块, 主要考虑 CoT 模块能够充分利用输入信息上下文的特点, 以克服布匹瑕疵检测中大部分瑕疵为较小目标的难点^[10]; 其次在颈部网络中引入 CBAM 注意力机制, 在骨干网络后进一步加强针对特征的信息筛选; 最后在网络的检测层处, 算法增加了一个检测头, 使检测头覆盖更大的检测范围, 网络能够更加全面的关注到想要的信息。

2.2 CoTNet

Transformer 的出现极大地促进了自然语言处理领域的发展, 该风格的网络架构近期在计算机视觉领域也掀起了一波浪潮并取得了极具竞争力的结果^[11]。尽管如此, 现有的架构直接在二维特征图上执行自注意力并得到注意力矩阵, 但是关于近邻丰富的上下文信息却并未得到充分探索。

CoT 模块是一种能够充分利用二维特征图上下文信息的注意力机构, 其结构如图 3 所示。

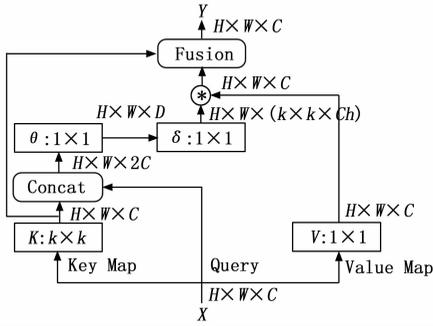


图 3 CoT 模块

本文中所提到的 CoTNet 结构由 CoT 模块代替主干网络中的 CSP1 模块而组成。

由图 3 可知假设输入二维特征 $X \in R^{H \times W \times C}$, 而键向量 K , 输入特征 Q 以及值向量 V 分别定义为:

$$K = X, Q = X, V = XW_v \quad (1)$$

其中: W_v 是指将 X 进行了特征的映射得到新的 V 。CoT 模块首先采用 $k \times k$ 卷积提取上下文信息, 所得 $K^1 \in R^{H \times W \times C}$ 自然地可以反应近邻间的上下文信息, 我们将其视

作输入 X 的静态上下文表达。将前述所得 K^1 与 Q 进行拼接并通过两个连续 1×1 卷积 W_θ 和 W_δ 来计算并得到注意力矩阵 A :

$$A = [K^1, Q] W_\theta W_\delta \quad (2)$$

接下来, 基于上述所得注意力矩阵 A 和值向量 V , 我们生成增强特征 K^2 :

$$K^2 = V * A \quad (3)$$

上述所得增强特征 K^2 可以捕获关于输入的动态特征交互, 我们将其称之为输入 X 的动态上下文表达。最后将上述两种上下文表达通过注意力机制融合即可得到 CoT 模块的输出。

CoT 模块能将上下文信息挖掘与自注意力学习集成到统一架构中, 充分探索近邻间的上下文信息以一种高效方式提升自注意力学习, 进而提升输出特征的表达能, 提升了视觉表达能力。本文在 DD-YOLOv5 模型的主干网络中多次应用该模块, 有利于模型对图像深浅特征的全局提取, 重点信息可以被加强, 从而使花色布匹瑕疵中占比较多的瑕疵的全局特征能够得到加强。

2.3 CBAM 注意力机制

在花色布匹瑕疵检测的过程中, 部分瑕疵目标仅占几个像素点, 部分瑕疵与背景极为相似, 这些瑕疵的特征信息极易在深层网络中丢失, 很容易出现漏检、错检。而 CBAM 注意力机制可有效增大目标在整张特征图中的权重, 使得各种瑕疵信息更容易在网络中学习到^[12]。其结构如图 4 所示。

由图可知, CBAM 模块包括通道注意力模块和空间注意力模块两个子模块。

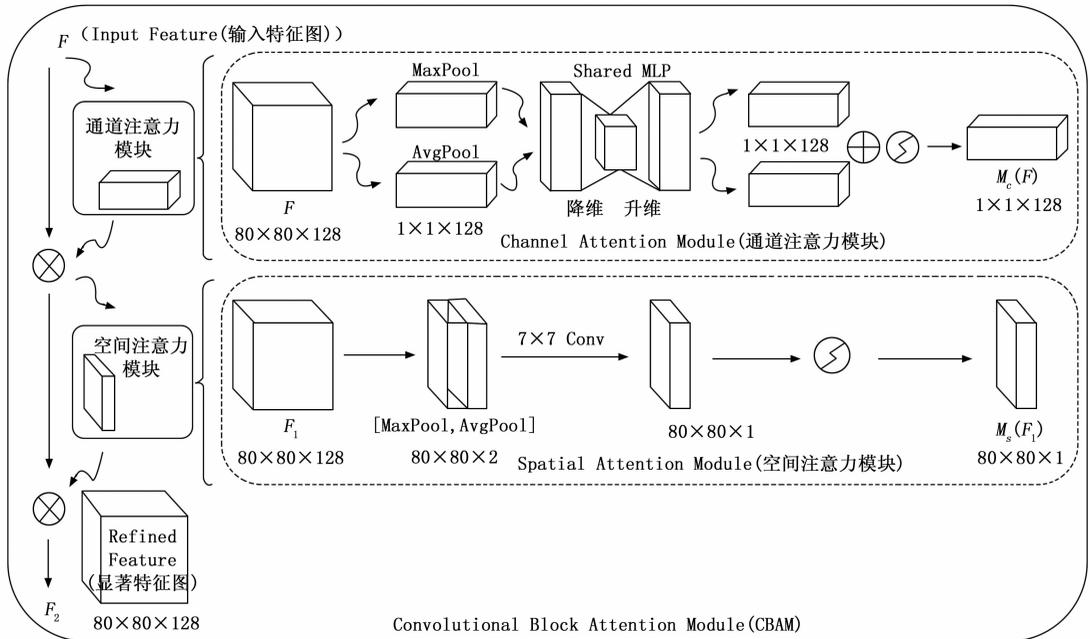


图 4 CBAM 模块

首先利用通道注意力模块来生成通道注意力图, 特征图的每个通道都被认为是一个特征检测器, 因此通道的注意力能够集中在给定输入图像中有意义的部分; 在该模块中, 为了有效地计算通道注意力, 采用压缩输入特征映射的空间维度的方法, 即同时使用平均池化和最大池化的方法, 这种做法比单独使用一种池化方法更具有表征力。

此时将通道注意力模块的输出作为空间注意力模块的输入, 然后利用空间注意力模块去生成空间注意力图; 与通道注意模块不同的是, 空间注意模块更加关注的是信息的位置去作为通道注意模块的补充; 为了计算空间注意力, 需要在通道中使用平均池化和最大池化操作, 将它们连接起来以生成一个有效的特征描述符, 然后使用两个池化操作聚合一个特征图的通道信息, 生成两个二维的特征图, 这两个特征图分别表示了通道的平均池化特性和最大池化特性, 然后利用一个标准的卷积层进行连接和卷积操作, 得到最终需要的二维空间注意力图。

这样一来 CBAM 模块可同时关注空间和通道信息, 使得提取到的高层特征更全面更丰富^[13-14]。

2.4 α -IoU 损失函数

边界框 (Bbox, bounding box) 是目标检测中用于框选出待检测目标的一个方框, 一个目标检测器通常由边界框回归和一个分类回归组成^[15]。

早期的目标检测工作使用 n 范数损失函数进行边界框回归, 这种方法对边界框的大小极为敏感, 交并比 (IoU, intersection over union) 损失函数的出现弥补了这一缺陷。与 n 范数损失函数相比, IoU 损失函数对边界框的尺度是不变的。然而, 当预测框与边界框的真实值不重叠时, IoU 损失会出现梯度消失问题, 导致收敛速度减慢, 进而导致检测器不准确。这激发了几种改进的基于 IoU 的损失函数设计, 包括 GIoU (generalized IoU)、DIoU (distance IoU) 和 CIoU (complete IoU)。YOLOv5 原本采用的损失函数为 Loss_GIoU, 当预测的边界框在真实的边界框内部时, GIoU 无法判断出正确的最小边界框^[16]。

在花色布匹瑕疵检测中有部分瑕疵长宽比较为极端, 处理不当极容易出现漏检情况, 采用更有效的边界框损失函数此时显得尤为重要。为解决该问题, 本文引入 α -IoU 予以改进, 在原有的基础上更加关注高 IoU 目标, 能够优化所有层次的目标, 实现不同水平的检测框回归精度^[16]。Loss $_{\alpha$ -IoU 的公式如下所示:

$$Loss_{\alpha$$
-IoU} = \frac{1-IoU^\alpha}{\alpha}, \alpha > 0 \quad (4)

通过对 Loss $_{\alpha$ -IoU 中的参数 α 进行调制, 可以推导出现有的大多数 IoU。本文中采用目前常用效果最好的 Loss $_{\alpha$ -IoU 改进为 Loss $_{\alpha$ -IoU, 其定义如下所示:

$$Loss_{\alpha$$
-IoU} = 1-IoU^\alpha + \frac{\rho^2 \alpha (b, b^{gt})}{c^{2\alpha}} + (\beta v)^\alpha \quad (5)

其中: c 代表的是能够同时包含预测框和真实框的最小

闭包区域的对角线距离, 而 v 衡量长宽比的一致性, 其定义如公式 (6); β 是一个正的权衡参数, 其数值如式 (7), $\rho^{2\alpha} (b, b^{gt})$ 分别代表了预测框和真实框的中心点的欧式距离。

$$v = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{gt}}{h^{gt}} \arctan \frac{\omega}{h} \right)^2 \quad (6)$$

$$\beta = \frac{v}{(1-IoU) + v} \quad (7)$$

式 (6) 中, $\frac{\omega^{gt}}{h^{gt}}$ 和 ω/h 分别代表目标框和预测框各自的宽高比, 至此便完成了对边界框损失函数的优化, 能有效地改善部分瑕疵长宽比悬殊的难题^[17]。

2.5 多尺度检测头

本文改进的算法在检测头部分增加了一个低维且高分辨率的检测头, 该检测头主要是用来处理难以检测的小目标部分^[18]。

增加的检测头由低层高分辨率的特征图生成, 当网络的输入图像尺寸为 960×960 时, 原算法中的 3 个检测头对应的特征图大小分别为 120×120 , 60×60 , 30×30 , 而增加的检测头特征图大小为 240×240 , 该检测头能充分利用到浅层特征信息, 明显降低卷积带来的小目标区域特征信息丢失的影响, 虽然在计算量上有所增加, 但能明显改善小目标的检测精度^[19-20]。

3 实验结果与分析

3.1 数据集及瑕疵介绍

本文选用的数据采集自佛山南海纺织车间, 共计 2 739 张含有瑕疵的图片, 其中 2 122 张作为训练集, 617 张作为测试集。图片中包括沾污、错花、水印、花毛、缝头、缝头印、虫粘、破洞、褶子、织疵、漏印、蜡斑、色差、网折共计 14 种瑕疵类型, 图 5 为瑕疵数量图, 沾污这种瑕疵的占比超过 80%, 种类分布极其不均衡。

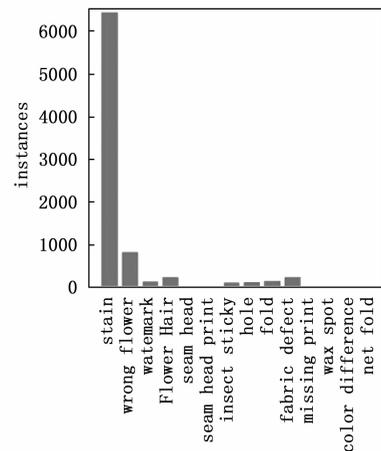


图 5 瑕疵数量

图 6 分别为小目标瑕疵、长宽比极端瑕疵、与背景及其相似的瑕疵展示。

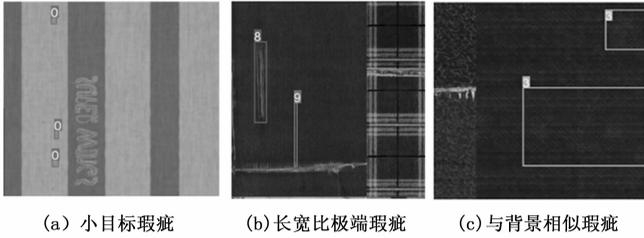


图 6 布匹瑕疵检测中的 3 种瑕疵检测难点

本文主要针对图 6 中 3 种瑕疵检测中的难点问题搭建检测平台, 并使用改进后的算法进行检测。

3.2 实验环境搭建

本文硬件配置环境如表 1 所示。

表 1 实验环境

系统	Windows 11 (64 位)
CPU	Intel Core i7-12700F
RAM	32.0 GB
GPU	NVIDIA RTX 3070Ti
深度学习框架	PyTorch
GPU 加速库	CUDA 11.0

3.3 实验结果展示

图 7 为多种瑕疵实际检测效果图的典型瑕疵检测细节图的展示。

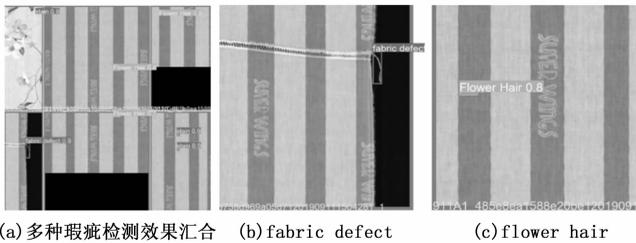


图 7 部分瑕疵检测实际效果

通过以上实际检测效果图可以发现, 在本文的算法下, 对布匹瑕疵的检测效果已经可以达到工业现场级别的要求, 且大部分的瑕疵检测识别率更是在 0.9 以上。可以充分说明本文算法在针对花色布匹的瑕疵检测上效果卓然。

3.4 评估标准

本文中引入以下评估标准评估模型性能:

精确率 (P, precision) 又称查准率, 指正确预测为正占全部预测为正的的比例, 如式 (8) 所示:

$$P = \frac{TP}{TP + FP} \tag{8}$$

召回率 (R, recall) 又称查全率, 指正确预测为正占全部实际为正的的比例, 如式 (9) 所示:

$$R = \frac{TP}{TP + FN} \tag{9}$$

平均精度 (AP, average precision,) 由精确率和召回率计算得出, 根据召回率值绘出精确率的折线图, 计算出该

折线下的面积, 如式 (10) 所示:

$$AP = \int_0^1 p(r) dr = \sum_{k=1}^N P(k) \Delta r(k) \tag{10}$$

平均精度均值 (mAP, mean average precision) 是指 C 个不同缺陷的平均精度 AP 的平均值, 如式 (11) 所示:

$$mAP = \frac{\sum_{i=1}^C AP_i}{C} \tag{11}$$

3.5 消融实验

为验证本研究提出的关于 YOLOv5 的 4 种改进策略, 在数据集上展开消融试验, 以判断每个改进点的有效性, 依次在原有模型基础上加上改进措施, 使用改进用“√”表示, 未使用改进则表格为空。训练过程使用相同的参数配置, 结果如表 2 所示。

表 2 消融实验结果

改进策略	第 1 组	第 2 组	第 3 组	第 4 组	第 5 组
增加检测层		√	√	√	√
CoTNet			√	√	√
CBAM				√	√
α-IoU					√
mAP/%	71.1	75.6	77.4	79.3	80.0

消融实验以原生 YOLOv5 网络模型为基准, 输入图像分辨率均为 960 * 960。

通过表 2 中第 1 组和第 2 组的实验对比, 可以发现当检测层由三层增加到四层, mAP 由 71.1% 升到了 75.6%, 结果表明增加的针对小目标的提取浅层信息的检测头起到了显著效果, 很明显的降低了漏检率。

通过第 2 组和第 3 组对比, 当网络中增加了 CoT 模块形成 CoTNet 后, mAP 由 75.6% 提升到 77.4%, 主要原因是因为 CoT 模块能多挖掘网络中的上下文信息, 进而对图像深浅特征进行全局提取, 这一举动虽然损失了较小的检测速率, 但能显著提高对目标的检测准确率。

通过第 3 组与第 4 组实验对比显示, 在颈部网络增加了 CBAM 模块的网络 mAP 值从 77.4% 提升到 79.3%, 表明 CBAM 模块能够更多地关注有用的信息, 提高对重点瑕疵如沾污的识别率。

通过第 4 组与第 5 组的实验数据对比, 瑕疵识别的 mAP 值由 79.3% 提升到了 80.0%, 说明在引入了新的边界框损失函数后, 检测器能更好地进行边界框的回归, 从而使得长宽比极端的瑕疵的漏检率下降。

综合上述可知, 在结合了多种针对 YOLOv5 算法的改进措施后, 本文的算法能够在布匹瑕疵的专项检测中取得较好的效果, 充分说明上述改进措施在实际检测中是有实际成效的。

3.6 损失函数收敛曲线

图 8 和图 9 为原生算法和改进后算法的损失函数收敛情况。

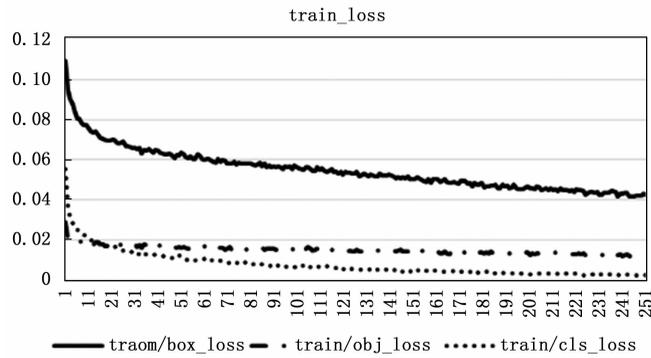


图 8 原生网络损失函数收敛曲线

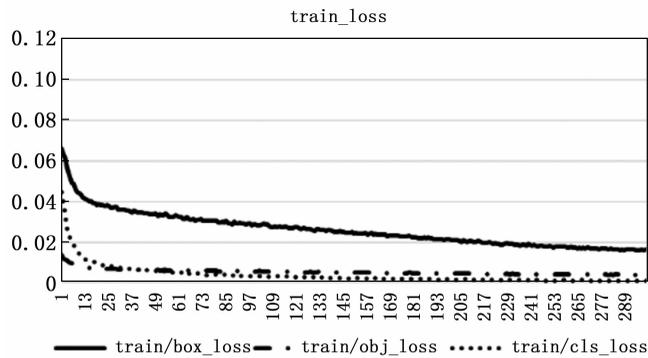


图 9 改进后网络损失函数收敛曲线

其中实线为定位损失 box_loss , 指预测框与标定框之间的误差, 点划线为置信度损失 obj_loss , 用于计算网络的置信度, 虚线为分类损失 cls_loss , 用于计算锚框与对应的标定分类是否正确。

通过两图对比可发现定位损失由明显下降, 这是由于算法中使用 α -IoU 代替原网络中的 G-IoU 交并比计算方法, 进而使得另外两种损失也跟着有所下降, 网络整体收敛更快, 效果更好。

3.7 对比实验

为了证明 DD-YOLOv5 网络模型在花色布匹表面瑕疵检测方面的优势, 我们将其与目前几种主流的目标检测算法进行了对比, 其对比结果如表 3 所示。

表 3 对比实验结果

算法	$mAP/\%$	$Precision/\%$	$Recall/\%$	FPS/Hz
Faster-RCNN	55.8	45.8	81.4	22.4
SSD	49.3	62.9	52.6	54.6
YOLOv4	68.4	74.5	63.0	38.7
YOLOv5	71.1	76.1	65.5	75.2
DD-YOLOv5	80.0	85.2	77.4	73.6

从表 3 中可以得知, Faster-RCNN 网络在精确率方面不高, 但召回率却是最高, 这是因为双阶段网络结构采用候选框机制遍历特征图, 故能将图中的目标较齐全的检测

到, 但这也同样会使得二阶段网络计算量较大故而速度较慢, 仅能达到 22.4 帧的速度。而同样作为单阶段目标检测的 SSD 网络虽然在检测速率上与 YOLO 系列没有过大的差距, 但是在检测精度上, 无论是精确率还是召回率, 比起 YOLO 系列都差距较大, 原因在于 SSD 的主干网络为 VGG 系列网络, 网络层数较低, 深层特征提取不全, 针对花色布匹瑕疵的检测效果较差。

同为 YOLO 系列的 YOLOv4 算法, 比起 YOLOv5 系列算法在检测精度上看来差距不大, 仅小幅度落后于 YOLOv5 系列算法, 但 YOLOv4 算法在检测速度上比起 YOLOv5 系列网络存在一定差距。

改进后的 DD-YOLOv5 由于增加了针对小目标检测的检测头以及其他的改进措施, 使得整个网络的复杂程度有所提高, 相应的在速度上有所损失, 检测速率为 73.6 Hz, 比原生 YOLOv5 算法的 75.2 Hz 降低了 2%, 但在精度综合评估标准的 mAP 值上, 本文算法比原生 YOLOv5 算法有着 12.5% 的检测精度提升, 故综合精度和速度两方面, 本文算法在众多算法中取得了较大的优势。

总之, 本文提出的 DD-YOLOv5 算法在速度和精度的综合考虑下获得了最佳效果, 比起经典算法性能更好, 相比同系列 YOLO 其他算法表现也最为优异。

4 结束语

为了解决花色布匹瑕疵检测问题, 进一步提高花色布匹表面瑕疵检测的准确率和识别速度, 本文提出了一种基于 YOLOv5 的 DD-YOLOv5 模型, 使用 COT 模块替代 CSP1 模块形成 CoTNet 并作为主干网络, 利用输入的上下文信息来指导自注意力机制的学习, 能够对图像深浅特征进行全局提取, 提升目标检测的准确率, 同时在颈部网络中引入 CBAM 注意力机制, 能同时关注空间和通道信息, 促使专注于有用的目标对象, 在检测头部分还增加了一个关注于小目标的检测头, 明显改善了网络的精度, 降低了漏检率, 最后用 α -IoU 替代原 G-IoU 交并比计算方法, 使得网络整体收敛更快, 检测效果更好。

通过实验结果证明, DD-YOLOv5 网络模型在花色布匹瑕疵的数据集上 mAP 达到了 80.0%, 检测速度也达到 73.6fps, 同时满足了检测精度和检测效率的要求, 对花色布匹瑕疵检测具有重要的现实意义。今后将进一步尝试不损失精度的前提下继续减小网络模型的参数量, 使模型得到进一步的轻量化。

参考文献:

[1] 周天宇, 朱启兵, 黄敏, 等. 基于轻量级卷积神经网络的载波芯片缺陷检测 [J]. 计算机工程与应用, 2022, 58 (7): 213-219.

[2] 李俊杰. 基于深度学习的布匹缺陷检测算法研究 [D]. 成都: 电子科技大学, 2020.

[3] 吕祥聪. 基于深度学习的布匹缺陷检测方法研究 [D]. 青岛:

