严立成. 基于改进的 YOLOv5 轻量化织物瑕疵检测方法[J]. 智能计算机与应用,2025,15(4):210-216. DOI:10.20169/j. issn. 2095-2163. 250430

基于改进的 YOLOv5 轻量化织物瑕疵检测方法

严立成

(浙江理工大学信息科学与工程学院,杭州 310018)

摘 要:为了解决现在纺织行业织物瑕疵检测中人工检测效率低下以及存在瑕疵点分布密集、疵点尺寸小和种类繁多的问题,本文提出了一种基于 YOLOv5s 的检测方法,在骨干网络引入 CoT 上下文信息和 CBAM 通道空间注意力机制,增强网络 对特征图的空间、通道和上下文重要信息提取能力;其次,在颈部引入 BiFPN 加权双向金字塔,改善了网络对多尺寸特征图关 键信息的提取能力,并在检测头前加入 ASFF 自适应空间融合机制,提升网络的鲁棒性和特征融合能力;最后,将原有的 CloU 损失函数替换为 EloU,提升算法梯度回归的速度以及检测的精确度。基于阿里天池织物瑕疵数据集的实验结果表明:该算 法的 mAP 相比原版的 YOLOv5 的 60.2%,提升到了 65.3%,验证了该算法的有效性和鲁棒性。该研究提出的算法检测精度有 明显提高,且满足即时检测的需要,符合生产要求。

关键词:深度学习;YOLOv5;织物瑕疵检测;注意力机制;损失函数 中图分类号:TP391.4 文献标志码:A 文章编号:2095-2163(2025)04-0210-08

Lightweight fabric defect detection method based on improved YOLOv5

YAN Licheng

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: In order to solve the existing problems that in the fabric defect detection of the textile industry, the efficiency of manual detection is low, the distribution of defect points is dense, the size of defects is small and there are many kinds of defects. Based on YOLOv5s, a spatial attention mechanism of CoT context information and CBAM channel is introduced into the backbone network to enhance the ability of the network to extract important spatial, channel and context information of feature maps. After that, BiFPN weighted bidirectional pyramid is introduced in the neck to improve the ability of the network to extract key information of multi-dimensional feature maps. ASFF adaptive spatial fusion mechanism is added before the detection head to improve the robustness and feature fusion ability of the network. Finally, the original CIoU loss function is replaced by EIoU to improve the speed of gradient regression and the accuracy of detection. Experimental results based on Ali Tianchi fabric defect data set demonstrates that the mAP of the algorithm is increased to 65. 3% compared with 60. 2% of the original YOLOv5, which verifies the effectiveness and robustness of the algorithm. The detection accuracy of the proposed algorithm is obviously improved, and meets the requirements of real-time detection and production.

Key words: deep learning; YOLOv5; fabric defect detection ; attention mechanism; loss function

0 引 言

纺织行业在国内的工业领域有着举足轻重的地 位。众所周知,生产出有瑕疵的织物却会给企业造 成不必要的收益损失,因此表面疵点检测是织物质 量控制的重要环节^[1]。由于织物在颜色、材质和图 案方面都有区别,对织物瑕疵进行检测存在一定难 度。以往较为常用的织物检测方法基本为人工检 测,但传统的人工检测方法却已暴露出了一些不足, 诸如:精度低、效率低等^[2]。为了提高企业的生产 效率和产量、降低人工成本,织物表面瑕疵检测方法 的自动化已成必然趋势。随着计算机硬件性能的提 高和软件算法技术的快速发展,深度学习神经网络 算法引起了学术界的关注和兴趣。其中,计算机视 觉领域中的深度学习卷积神经网络算法已经日趋完 善,现已成为目标检测技术中的热点研发课题。

在计算机视觉领域的目标检测算法中,主要分为2种框架^[3]。一种是一阶(one-stage)目标检测

作者简介:严立成(1996—),男,硕士研究生,主要研究方向:深度学习。Email:1975735205@qq.com。 收稿日期: 2023-09-26

211

算法,这种算法提供了一种端对端的目标检测和分 类框架,输入原始图像,经过框架算法处理后输出分 类结果和目标检测框,其中较为出名的有 YOLO 系 列框架^[4-8]和 SSD 算法^[9]: 另一种是二阶(twostage)目标检测算法。这种算法在检测过程中,输 入原始图片后会先产生候选框,然后对候选框进行 检测,最后再对候选框进行分类和回归,其中较为知 名的有: RCNN^[10] 及在此基础之上改进而来的 Faster - RCNN^[11]、Cascade - RCNN^[12] 和 Mask -RCNN^[13]等网络。袁华清等学者^[14]提出了一种基 于改进 YOLOv3-tiny 网络的织物瑕疵检测算法模 型。模型首先在原始框架的基础上,增添了融合浅 层和深层特征的第三预测层,通过多尺度的特征层 融合提升小目标检测能力:其次,在网络输出前使用 了基于 CBAM 的混合注意力机制和数个卷积模块, 在提升瑕疵目标权重和网络深度的同时,降低无关 背景权重,提高模型的抗干扰能力和准确度;另外, 针对织物瑕疵目标检测的特定应用,利用 K-means 方法对锚框进行了重新聚类。Yun 等学者^[15]提出 了一种针对织物特性的轻量级模块—织物模块 (Fabric Module, FM)。研究中使用 YOLOX-Nano 作为基础网络,用 FM 代替 SPP 模块,该方法提高了 检测精度和速度,更适合织物疵点的自动检测。 MAO 等学者提出了一种基于挤压和激励(SE)模块 的 YOLOv5(SE-YOLOv5)来建立一个高效的织物检 测系统^[16],并用 ActivateOrNot(ACON)激活函数取 代了 YOLOv5 cross stage partial (CSP) 中传统的 Leaky Rectified Linear Unit(ReLU)激活函数,所提方

法的准确率、泛化能力和鲁棒性均有提高。

本文提出了一种基于 YOLOv5 神经网络框架的 织物表面瑕疵检测方法。首先,在框架的骨干网络部 分引入 CoT^[17]和 CBAM 注意力集中机制^[18],提高网 络对多尺度图像的特征提取能力同时抑制特征冗余; 其次,为了提升模型的性能并减少计算量,将原有的 YOLOv5 颈部 PANet^[19]和 FPN^[20]金字塔网络替换为 BiFPN^[21],并在颈部网络后引入 ASFF^[22]自适应空间 特征融合模块,以达到更加充分利用网络高层特征的 语义信息和底层空间信息的目的。最后,将 YOLOv5 原本的损失函数由 CloU^[23] 替换为 EloU^[24]。

1 YOLOv5 的网络结构

YOLOv5 是 YOLO 系列神经网络框架中检测结 果表现较为优秀的一个,和其他目标检测神经网络 框架一样,由输入层(Input)、骨干网络(Backbone)、 特征融合层(Neck)以及预测头(Prediction Head)构 成。根据网络的不同深度和宽度,YOLOv5 被划分 成4个版本,分别是 YOLOv5s、YOLOv5m、YOLOv51 和 YOLOv5x。为了在网络的深度和宽度之间实现 相互平衡,Chen 等学者^[25]提出了 EffectiveNet 方法, 用一个常数来达到平衡网络深度和宽度的效果。本 研究的目的是对织物表面进行瑕疵检测,因此,为了 满足纺织行业中实际生产环境的需求,在保证检测 准确率以及检测实时性要求的情况下让网络规模尽 可能地小,本论文将选择 YOLOv5s 作为基础网络结 构来进行研究以及改进。YOLOv5 的网络结构如图 1 所示。



图 1 原版 YOLOv5 结构图 Fig. 1 Original YOLOv5 structure

YOLOv5 的输入端和 YOLOv4 一样,采用了 Mosaic 数据增强方式,但 CutMix 只使用了2 张图片 进行拼接,而 Mosaic 数据增强则采用了4 张图片, 通过随机缩放、随机裁剪、随机排布的方式进行拼 接。Mosaic 数据增强可以丰富数据集,同时还起到 减少 GPU 的作用。

1.1 YOLOv5 的输入端

输入端(Input)包括 Mosaic 数据增强、自适应锚 框计算以及自适应图片缩放。其中, Mosaic 数据增 强可以随机使用4张图片并将其随机缩放, 最后再 将图片随机分布并进行图片的拼接, 可以丰富图片 集的数量, 还可以随机增加小目标的数量, 提高网络 的鲁棒性并且减少 GPU 使用, 解放资源使用。自适 应锚框计算将针对不同数据集, 设定初始长度和宽 度的锚框, 可以使用代码自动化设计锚框的尺寸, 不 需要人工手动计算相对应数据集图片大小的锚框。 YOLOv5 的自适应图片缩放将原始图片尺寸统一缩 放至设定的标准尺寸, 并自适应填充最少的黑边, 因 此 YOLOv5 推理速度得到了有效的提高, 同时减少 图片冗余信息。

1.2 YOLOv5 Backbone

YOLOv5 的骨干(Backbone)网络部分包含 Focus 结构、CSP^[26]结构和 SPPF^[27]结构。其中, Focus 对原始图像进行切片操作,缩小图像尺寸的 同时增加了图像的通道数,丰富了图像语义信息。 YOLOv5 包含 2 种 CSP 结构,可以起到下采样的作 用,CSP 模块首先将基础层特征映射为 2 部分,之后 通过跨阶段层次结构将这 2 部分合并,起到减少网 络优化阶段的梯度信息重复问题,在保证准确率的 前提下简化了网络的计算量。SPPF 主要用于提升 特征图的表达能力。

1.3 YOLOv5 Neck

YOLOv5的颈部(Neck)采用 FPN+PAN 结构, 其中,FPN 是自顶向下,通过上采样的方式将高层 的特征信息传递并融合,保证多尺度的特征具有丰 富的语义信息。PAN 的作用是通过加入自底向上 的线路,对特中层进行下采样操作,弥补因经过 FPN 的多层网络后丢失的低层细节信息,同时增强定位 信息。

1.4 YOLOv5 Prediction

Prediction 使用了二维普通卷积,采用了 CloU_loss 计算位置损失,以 BCEWithLogits_loss 计 算置信度损失和分类损失。CloU_Loss 可由如下公 式求得:

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

其中,α表示权重;v表示衡量长宽比的相似度; b,bst分别表示预测框和目标框的中心点,两者之间 距离采用欧式距离ρ;c表示能同时包含预测框和目 标框的最小包围框的斜距。

进一步,推得 BCEWithLogits_Loss 的数学公式 如下:

BCEWithLogits_loss =
$$-w_n [y_n \cdot \log \sigma(x_n) +$$

$$(1 - y_n) \cdot \log(1 - \sigma(x_n))]$$
 (2)

其中, n 表示样本总数; w_n 表示权重调节系数; y_n 表示数据标签值; x_n 表示数据预测值; σ 表示 Sigmoid 函数。

2 改进 YOLOv5 的网络结构

2.1 改进的 YOLOv5 Backbone

2.1.1 引入 CBAM 注意力机制

引入注意机制模块的目的是为了关注图像的关键特征,同时抑制非关键以及不必要的特征。卷积模块的注意力机制(Convolutional Block Attention Module,CBAM),是一种结合了空间(Spatial)和通道(Channel)的注意力机制模块。CBAM 注意模块通过卷积运算并混合跨通道(Channel)和空间(Spatial)信息来提取特征信息,该模块将一个通道注意力模块(Channel Attention Module)和一个空间注意模块(Spatial Attention Module)相连接,从通道和空间两个维度来让神经网络学习关键特征。其整体结构如图 2 所示。



图 2 CBAM 模块结构图 Fig. 2 CBAM module structure

2.1.2 设计原理解析

通道注意力模块的输入特征为 F/inR^{C×H×W},通 道注意力模块中的一维卷积 M_e/inR^{C×1×1},卷积得到 的结果和原图相乘,将通道注意力模块的输出结构 作为空间注意力模块的输入特征,和空间注意力模 块的二维卷积模块进行卷积操作: M_s/inR^{1×H×W},最 后将空间注意力模块的输出结果与原图相乘。通道 注意力模块结构如图 3 所示。



图 3 通道注意力模块结构图

Fig. 3 Channel Attention module structure

其中,通道注意力模块的主要任务是提取图片 的有用信息,将输入特征图分别通过并行的最大池 化层和平均池化层,压缩特征图的尺寸,将特征图 由 *C*×*H*×*W*压缩成*C*×1×1(这里,*C*为通道数, *H*为图像的高度,*W*为图像宽度),再将压缩后的特 征图经过共享权重模块,将特征图的通道数压缩为 原有尺寸的1/r(这里r为减少率),然后再将通道数 扩展回原通道数*C*,经过*ReLU*激活函数得到2个结 果。将2个激活结果相加,最后经过一个 Sigmoid 激活函数得到最终的通道注意输出结果,将此输出 结果和输入的特征图进行相乘操作,使输出特征尺 寸变回 *C*×*H*×*W*。

通道注意模块的公式为:

 $M_{c}(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F)))$

(3)

其中, σ 表示 Sigmoid 激活函数; *MLP* 表示共享 权重层; F 表示输入特征图; *MaxPool* 表示通道全局 最大池化; *AvgPool* 表示通道全局平均池化。

研究可知,通道注意模块能提取到更丰富和全 面的语义特征。

空间注意模块结构如图 4 所示。空间注意模块 是用于是更好地提取目标的位置信息,此模块将通 道注意的输出结果分别通过并行的最大池化层和平 均池化层,将特征图的通道数 C 压缩为 1,然后经过 张量拼接(Concat)操作将经过上述 2 个池化层的输 出结果进行拼接,经过一个大小为 7×7 的卷积核变 为特征通道数为 1 的特征图,然后使用一个 Sigmoid 激活函数得到空间注意的特征图,最后将输出结果 和原特征图相乘,得到一个尺寸为 C×H×W的最终 输出结果。





空间注意模块的公式为:

 $M_{s}(F) = \sigma(f^{7\times7}([AvgPool(F); MaxPool(F)])) (4)$

其中, σ 表示 Sigmoid 激活函数; F 表示输入特征图; *MaxPool* 表示空间全局最大池化; *AvgPool* 表示空间全局平均池化; $f^{7\times7}$ 表示 7×7 大小的卷积核; Concat 表示张量连接操作。

本研究将在 CSP 后增加一个 C3+CBAM 注意力 集中模块。为提高主干网络对目标的特征提取能 力,在 Backbone 阶段引入软注意力机制(Softattention)中的卷积注意力机制模块,该模块可以在 空间维和通道维分别推断出注意力的权重,使网络 能够关注到图像中关于目标的重点信息。

2.2 CoT

CoT(Contextual Transformer),可以重复利用输入键图(Key Map)之间的上下文(Context)信息来指导动态注意力矩阵的学习,增强了视觉表示的能力。 CoT首先通过上下文对输入进行卷积得到输入的静态上下文表示,然后通过2个连续的卷积,将得到的注意力矩阵和输入相乘,实现输入的动态上下文表示。最后,将静态和动态的上下文表示进行特征融





Fig. 5 CoT structure

2.3 改进的 YOLOv5Neck

2.3.1 双向特征金字塔网络

双向特征金字塔网络(Bidirectional Feature Pyramid Network, BiFPN)引入可学习权重来学习不同 输入特征的重要性,同时重复使用自顶向下和自底 向上的多尺度特征融合作用,起到高效的多尺度特 征融合作用。BiFPN 网络结构如图 6 所示。



图 6 BiFPN 结构图 Fig. 6 BiFPN structure

图 6 中, 橙色部分为自顶向下的通路, 传递的是 高层特征的语义信息:蓝色部分是自底向上的通路, 传递的是低层特征的位置信息:紫色部分是同一层在 输入节点和输出节点间新加的一条边。为了减少网 络参数量,删除那些只有一条输入边的节点,因为单 个输入边无法进行特征融合。为了融合更多特征信 息、且不增加计算成本,BiFPN 将在融合处的节点和 同一级别的、且输出特征图分辨率相同的输入节点之 间增加额外的连接。图 6 中,5,6 节点和 8,9 节点为 不同级别节点的聚合,需要一个转换函数 F 将不同的 特征有效地聚合并输出新的特征信息。传统的特征 融合主要有简单的特征图相加,如 contcat 或者 shorcut 拼接,然而,不同输入特征图有不同的分辨率,简单拼 接无法区分不同分辨率特征图对融合输入的贡献。因 此,BiFPN 提出了一种快速归一化融合方法,针对不同 级别的节点特征融合,其融合方法公式如下:

$$O = \sum_{i} \frac{\omega_{i}}{\varepsilon + \sum_{i} \omega_{j}} \cdot I_{i}$$
(5)

其中, ω_i 表示特征权重, ε 表示学习率。使用 Relu 激活函数,确保每个特征权重 $\omega_i \ge 0$,学习率 ε 设 为 0.000 1,在 GPU 上,这种融合方法可以更快地运行。

本研究中改进的 YOLOv5 结构如图 7 所示。图 7 中,将 YOLOv5s Backbone 中的第二个 C3 结构和 Neck 中的第 19 层进行 add 连接,在不改变特征图 参数的前提下,提升了网络中对尺寸目标的特征融 合能力。其融合过程公式如下:

$$P_{19}^{\text{out}} = Conv \left(\frac{\omega_1 \cdot P_6^{\text{in}} + \omega_2 \cdot P_{14}^{\text{in}} + \omega_3 \cdot P_{18}^{\text{in}}}{\omega_1 + \omega_2 + \omega_3 + \varepsilon} \right) \quad (6)$$

其中, P_i^{out} 表示网络中第 i 层的输出特征; P_i^{in} 表示网络中第 i 层的输入特征; Conv 表示卷积操作。



图 7 改进后的 YOLOv5 结构图

Fig. 7 Improved YOLOv5 structure

2.3.2 自适应空间特征融合

自适应空间特征融合(Adaptively Spatial Feature Fusion, ASFF)模块分别对 PANet 输出的 3 个水平特 征图进行加权融合,通过增加可学习的参数抑制梯 度反向传播过程中的不一致,充分利用不同尺度的 特征,从而达到更充分利用网络中高层特征的语义 信息和底层特征中的空间信息的目的。

2.4 改进的 YOLOv5 Loss Function

原生的 YOLOv5 的 loss-function 为 *CloU*,但是 *CloU* 损失虽然在 *DloU* 损失的基础上添加了衡量预 测框和真实框纵横比 v,在一定程度上可以加快预 测框的回归速度,但是存在的问题是在预测框回归 过程中,一旦预测框和真实框的宽高纵横比呈现线 性比例时,预测框 w 和 h 就不能同时增加或者减少, 也就不能继续进行回归优化了。本文将 YOLOv5 的 loss-function 改进为 *EloU*,将纵横比的损失项拆分 成预测的宽高分别与最小外接框宽高的差值,加速 了预测框的收敛、提高了预测框的回归精度。*CloU* 函数公式如下:

$$R_{\text{CloU}} = \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha v \tag{7}$$

其中,α表示正向权衡参数,其公式如下:

$$\alpha = \frac{v}{(1 - IoU) + v} \tag{8}$$

这里,用于衡量恒定的预测框长宽比,其公式如下:

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w}{h} \right)^2$$
(9)

由此可得, CloU的损失函数定义如下:

$$L_{\rm CloU} = 1 - IoU + \frac{\rho^2(b, b^{\rm gt})}{c^2} + \alpha v \qquad (10)$$

CloU Loss 虽然考虑了边界框回归的重叠面积、 中心点距离、纵横比。但是通过式(7)中的v反映的 纵横比的差异,而不是宽高分别与其置信度的真实 差异,所以有时会影响模型的有效优化相似性。针 对这一问题,在 CloU 的基础上将纵横比拆开,提出 了 EloU Loss,其损失函数定义如下:

$$L_{\rm EIoU} = L_{\rm IoU} + L_{\rm dis} + L_{\rm asp} = 1 - IoU + \frac{\rho^2(b, b^{\rm gt})}{c^2} + \frac{\rho^2(b, b^{\rm gt})}{C_w^2} + \frac{\rho^2(b, b^{\rm gt})}{C_h^2}$$
(11)

其中, C_w 和 C_h 分别表示覆盖2个Box的最小外接框的宽度和高度。

该损失函数包含3个部分:重叠损失、中心距离

损失和宽高损失。前两部分延续 CloU 中的方法, 但是宽高损失直接使目标盒与锚盒的宽度和高度之 差最小,这就使得收敛速度会更快。

3 实验结果及分析

3.1 数据集

本实验使用公开的百度飞桨织物瑕疵数据集,数 据集为百度飞桨的4371张有图案带瑕疵织物布匹 图片,分辨率分别为4096×1080,4096×1086和 4096×1920,图像包含有复杂图案的织物和纯色的织 物,缺陷种类有缝头、沾污、织疵、虫粘、缝头印、破洞、 花毛、遮字和水印共9种瑕疵。数据集选取3749张图 像用于训练,622张图像用于测试。

3.2 实验环境和参数

本研究的实验操作系统为 ubuntu20.04,显卡为 NVIDIA GeForce RTX 3090,其显存为 24 GB。CPU 为 Intel(R) Xeon(R) Platinum 8358P,系统内存为 80 GB。集成开发环境(IDE)选择使用 Pycharm 2021.3.3,编程语言使用 Python,深度学习框架选择 使用 Pytorch。在上述实验条件下对有瑕疵的织物 图像数据集进行训练,设置初始学习率(learning rate)为0.01,权重衰减设置为0.000 5,动量设置为 0.937,优化器选择使用 SGD。

3.3 实验结果和对比

本研究将数据集的训练迭代次数(epoches)设置 为 300 次,检测结果见图 8。此后将原版 YOLOv5 模 型和本研究改进后的 YOLOv5 模型训练得出的结果 进行对比,对比主要参数有均值平均精度(mean Average Precision, mAP)和损失值 Loss。结果见表 1。

由结果可知,在添加了各个模块后,模型的精度、召回率、mAP以及检测速度均有所提升,在添加 CBAM后,mAP提升1.4%;添加BiFPN后,mAP提 升了1.8%;添加ASFF后,mAP提升了1.4%;将损 失函数由 CIoU 改为 EIoU后,mAP提升了1.8%。

训练后模型的 box、目标损失函数和分类损失 函数如图 9 所示。



图 8 检测结果示意图 Fig. 8 Detection result samples

表 1 算法改进前后比较 Table 1 Comparison of the algorithm before and after improvement

算法	mAP/%	精度 P/%	召回率 R/%	参数量	检测速度/(帧・s ⁻¹)
YOLOv5	60.2	59.8	57.8	7 235 389	112.3
YOLOv5+CBAM	61.6	60.3	58.6	7 268 255	110.5
YOLOv5+CBAM+BiFPN	63.4	63.0	61.9	12 599 813	104.8
YOLOv5+CBAM+BiFPN+ASFF YOLOv5+CBAM+BiFPN+ASFF+ <i>EIoU</i>	64.8	62.8	62.4	12 862 567	102.6
	65.3	63.7	62.8	12 883 943	98.8



由图 9 可知, 原版 YOLOv5 模型的目标损失函数在 200 个 *Epoch* 后依旧在大幅度跳动, 而改进后的模型 3 项损失函数回归曲线更为平滑, 收敛速度更快, 说明改进后的算法训练效果优于原版模型。

图 10 展示了改进前后的模型检测结果。



Fig. 10 Comparison of detection results

从图 10 中可知,改进后的模型解决了原有模型 漏检和错检等问题,对小目标的检测更加精确。

4 结束语

本研究针对织物瑕疵存在的种类繁多、发布密 集且不规律与尺寸小等问题,提出了一种基于改进 的 YOLOv5 模型的织物表面瑕疵检测算法。该算法 通过在骨干网络部分加入 CoTNet 上下文注意力机 制和 CBAM 注意力机制,提高模型对特征图重要信 息的提取能力;在颈部引入了 BiFPN 双向金字塔网 络,提高模型对多尺寸特征融合的效率;加入 ASFF 剔除了模型在多尺度特征融合阶段时产生的冗余信 息,提高了模型的检测精度和鲁棒性;将损失函数由 *CloU* 替换为 *EloU*,提升了模型的检测精度以及在 学习时的损失函数回归速率。实验结果表明:改进 后的 YOLOv5 模型具有更好的检测精度、鲁棒性和 可部署性,能对纺织产业的自动化和智能化产生发 挥重要推动作用。

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 Huosheng, ZHANG Yafeng, WU Zesen. Fabric defect detection method combing 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]. 现代纺织技术,2022,30(5):1-11.
- [4] 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. Piscataway, NJ:IEEE, 2016: 779-788.
- [5] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2017: 7263–7271.
- [6] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv, 1804. 02767, 2018.
- [7] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv,2004. 10934, 2020.
- [8] ZHU Xingkui, LYU S, WANG Xu, et al. TPH YOLOV5: Improved YOLOv5 based on transformer prediction head for object detection on drone – captured scenarios [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ; IEEE, 2021; 2778–2788.
- [9] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]//Proceedings of the 14th European Conference on Computer Vision(ECCV 2016). Cham: Springer, 2016: 21–37.

- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2014: 580-587.
- [11] REN Shaoqing, HE Kaiming, 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.
- [12] CAI Zhaowei, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2018: 6154-6162.
- [13] HE Kaiming, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2017: 2961–2969.
- [14] 袁华清, 刘桂华, 王娅琼. 基于改进 YOLOv3-tiny 的织物表面 小目标瑕疵检测[J]. 制造业自动化, 2022, 44(12):172-176.
- [15] YUN Y Y, LI L Q. A lightweight algorithm for woven fabric defect detection [C]//2022 China Automation Congress (CAC).
 Xiamen, China: Chinese Association of Automation, 2022:1025–1029.
- [16] ZHENG L, WANG X, WANG Q, et al. A fabric defect detection method based on improved YOLOv5 [C]//Proceedings of the IEEE 7th International Conference on Computer and Communications(ICCC). Picataway,NJ:IEEE, 2021;620–624.
- [17] LI Yehao, YAO Ting, PAN Yingwei, et al. Contextual transformer networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45 (2): 1489–1500.
- [18] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham:Springer,2018: 3-19.
- [19] WANG Wenhai, XIE Enze, SONG Xiaoge, et al. Efficient and accurate arbitrary-shaped text detection with pixel aggregation net-

work [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2019: 8440-8449.

- [20] 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. Piscataway, NJ:IEEE, 2017: 2117–2125.
- [21] TAN Mingxing, PANG Ruoming, LE Q V. Efficientdet: Scalable and efficient object detection [C]//Proceedings of the IEEE/ CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2020; 10781–10790.
- [22] LIU Songtao, HUANG Di, WANG Yunhong. Learning spatial fusion for single-shot object detection [J]. arXiv preprint arXiv, 1911.09516, 2019.
- [23] ZHENG Zhaohui, WANG Ping, LIU Wei, et al. Distance-IoU loss: Faster and better learning for bounding box regression [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [24] ZHANG Yifan, REN Weiqiang, ZHANG Zhang, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146–157.
- [25] CHEN Wei, YIN Hongzhi, WANG Weiqing, et al. Effective and efficient user account linkage across location based social networks [C]//Proceedings of the IEEE 34th International Conference on Data Engineering (ICDE). Piscataway, NJ:IEEE, 2018: 1085– 1096.
- [26] LIU Wei, HASAN I, LIAO Shengcai. Center and scale prediction: Anchor-free approach for pedestrian and face detection[J]. arXiv preprint arXiv, 1904. 02948, 2019.
- [27] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, 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.