

基于注意力机制和多尺度特征融合的卷纸包装缺陷检测算法

黄期峰*, 李 洲

五邑大学智能制造学部, 广东 江门

收稿日期: 2023年8月25日; 录用日期: 2023年9月19日; 发布日期: 2023年9月26日

摘 要

本文针对卷纸包装缺陷检测问题, 提出了一种基于注意力机制和多尺度特征融合的卷纸包装缺陷检测算法FYOLOv5。该方法具有以下特点: 以轻量级的YOLOv5s模型为骨干网络结构, 在特征提取网络引入卷积注意力模块, 提升模型对卷纸缺陷关键特征的提取能力。在颈部网络引入自适应空间特征融合模块, 提高模型多尺度特征融合能力, 使模型对高层的语义信息以及对底层的轮廓、形状和颜色等视觉信息利用更加充分。在检测头部分利用Decouple模块对目标检测任务进行分解, 使模型更易优化。最后, 使用 α IoU损失函数优化当真实框与预测框重合时的IoU计算退化问题。本文采集并标注了卷纸包装缺陷检测数据集, 并在该数据集上进行实验。实验结果表明, 改进后的检测模型准确率达到了95.3%, 高于原始模型2.1%, 同时推理速度能够满足卷纸包装的检测需求。

关键词

卷纸包装, 缺陷检测, 深度学习, 卷积注意力模块, 多尺度特征

Defect Detection Algorithm for Roll Paper Packaging Based on Attention Mechanism and Multi-Scale Feature Fusion

Qifeng Huang*, Zhou Li

Department of Intelligent Manufacturing, Wuyi University, Jiangmen Guangdong

Received: Aug. 25th, 2023; accepted: Sep. 19th, 2023; published: Sep. 26th, 2023

*通讯作者。

文章引用: 黄期峰, 李洲. 基于注意力机制和多尺度特征融合的卷纸包装缺陷检测算法[J]. 计算机科学与应用, 2023, 13(9): 1773-1784. DOI: 10.12677/csa.2023.139176

Abstract

This thesis aims at the problem of paper roll packaging defect detection, and proposes a paper roll packaging defect detection algorithm FYOLOv5 based on attention mechanism and multi-scale feature fusion. This method has the following characteristics: The algorithm uses the lightweight YOLOv5s model as the backbone network structure, introduces the convolutional attention module in the feature extraction network, and enhances the model's ability to extract key features of paper roll defects. In the neck network, an adaptive spatial feature fusion module is introduced to improve the model's multi-scale feature fusion ability, making the model use more fully the high-level semantic information and the low-level visual information such as contour, shape and color. In the detection head part, the Decouple module is used to decompose the target detection task, making the model easier to optimize. Finally, the α IoU loss function is used to optimize the IoU calculation degradation problem when the real box and the predicted box overlap. This thesis collects and annotates a paper roll packaging defect detection dataset, and conducts experiments on this dataset. The experimental results show that the improved detection model has an accuracy of 95.3%, which is 2.1% higher than the original model, and the inference speed can meet the detection requirements of paper roll packaging.

Keywords

Roll Paper Packaging, Defect Detection, Deep Learning, Convolutional Attention Module, Multi-Scale Features

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来, 国内造纸和纸制品行业增加值增长率持续增长, 卷纸是纸制品的重要组成部分, 作为日常生活中重要而又广泛使用的产品之一, 在提升人们生活品质和满足消费需求方面具有重大作用。“十四五”是中国造纸工业进入科技赋能、转型升级、实现高质量发展, 努力攀登全球产业价值链高端的新发展时期[1]。在卷纸的生产过程中, 包装是非常重要的工序, 卷纸的包装是否完好直接影响卷纸包装质量, 但包装过程中不可避免会出现包装缺陷问题。当前对于卷纸包装质量企业主要依靠人工进行质检, 存在检测效率低、标准不统一和用人成本高等问题, 不能满足现代化生产的需求。因此使用现代化技术解决卷纸包装问题具有重要意义, 有助于提高卷纸包装的检测效率, 同时为企业节约了大量人力成本。

随着机器学习在工业领域的成功应用, 越来越多的研究者开始使用机器学习方法进行工业产品的缺陷检测, Zhang 等人[2]针对强反射金属表面常见的7种缺陷进行研究, 提出了一种基于 Ostu 阈值法和 SVM 的缺陷检测方法, 该方法利用小波平滑对图像进行预处理, 去除背景噪声, 随后对图像进行二值化处理提取有效特征并使用 SVM 分类器进行分类。Kuang 等人[3]对竹条缺陷进行研究, 首先使用阈值算法对图像进行预处理, 获得感兴趣区域, 然后使用局部二值化和灰度共生矩阵对感兴趣区域的纹理特征进行提取, 最后使用支持向量机进行分类。文献[4]使用直方图方法对食品包装缺陷进行检测, 实验也证明了该方法的有效性。文献[5]使用信息增强和模糊聚类方法, 实现对牙刷包装缺陷的检测。传统的缺陷检测方法具有一定的效果, 但是传统的方法容易受背景、光照等环境因素的影响, 造成缺陷检测准确率低。

随着深度学习的快速发展, 以及深度学习在工业领域中的成功应用, 针对现有缺陷检测技术的不足, 研究者利用深度学习方法进行缺陷检测。深度学习在理论方法上取得的重要进展, 加速了深度学习方法在工业领域的应用。例如 YOLO 算法的不断改进, 使得 YOLO 算法在目标检测上具有很好的准确性和检测速度, 并在一些领域获得了很好的应用[6] [7] [8]。文献[9]针对传统方法对棉布包装缺陷检测不够准确、小目标检测识别率低等问题, 提出了一种改进循环神经网络的深度学习网络, 对比于传统的方法获得了 9% 的绝对增益。文献[10]针对传统方法通用性差、在复杂环境下难以应用等问题, 提出了一种基于 YOLOv3 的缺陷检测算法。在卷纸包装缺陷检测文献中, 文献[11]与我们的工作最相关, 它针对卷纸包装缺陷检测问题, 提出了一种基于改进 YOLOv3 的缺陷检测方法, 通过引入多头注意力机制提高了模型的检测精度。

近年来, YOLOv5 在类别目标检测任务中明显优于 YOLOv3, 因此许多研究将 YOLOv5 应用到缺陷检测任务中, 并取得了不错的结果。如文献[12]针对单阶段检测网络 YOLOv5 的特征提取能力不足、模型感受野受限以及特征融合不充分等问题, 提出一种改进 YOLOv5 的钢材表面缺陷检测算法, 较 YOLOv3 提升了 8.6% 的准确率。文献[13]基于 YOLOv5 算法应用于 PCB 裸板的缺陷检测上, 使得缺陷检测精度到达了 90% 以上, 满足了工业检测需求。

基于 YOLOv5 在各种缺陷检测任务中的优异表现, 为了进一步提升模型对于卷纸包装缺陷检测的针对性, 提出了一种基于注意力机制和多尺度特征融合的卷纸包装缺陷检测算法 FYOLOv5, 本文针对 YOLOv5 检测算法提出了以下改进: 1) 为了提高模型对于不合格品关键特征的提取能力, 本文利用卷积注意力模块(CBAM)沿着通道和空间两个独立的维度捕获通道以及空间信息, 提高模型提取全局特征和局部特征的能力; 2) 在颈部网络的三个输出层添加自适应空间特征融合(ASFF)模块, 提高模型多尺度特征融合能力, 使得模型能够学习不同层级缺陷纹理特征, 使模型对高层的语义信息以及对底层的轮廓、形状和颜色等信息利用更加充分; 3) 为了使模型更易于优化, 本文使用 Decoupled 模块对目标检测任务进行分解。4) 优化了边界框损失函数, 使模型探测器获得更高的回归精度, 提升检测模型的鲁棒性。

2. 算法模型

2.1. 网络结构

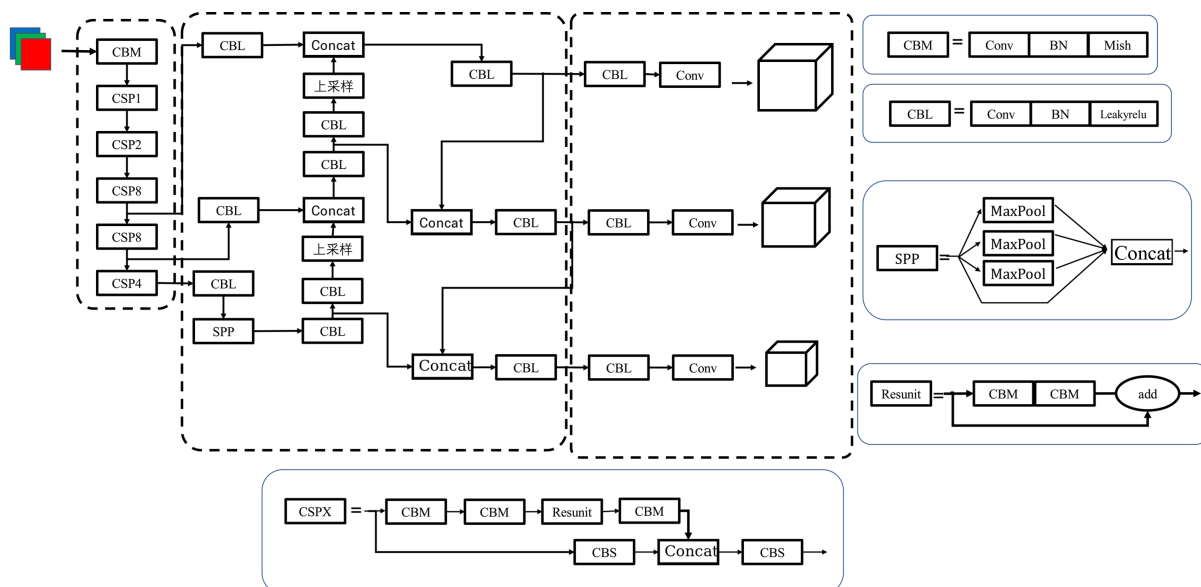


Figure 1. Roll packaging defect detection algorithm module

图 1. 卷纸包装缺陷检测算法模块

YOLOv5s 检测模型网络结构如图 1 所示, 主要由主干网络(backbone)、颈部网络(neck), 检测头(head) 三部分组成。其中主干网络主要由 Focus、CSP 和 SPP 构成用于特征提取, 颈部网络主要由特征金字塔网络和路径聚合网络组成, 将不同尺度的特征图进行融合, 以提高目标检测的精度。检测头则对每个目标的位置、类别和置信度进行回归, 从而完成目标检测任务。

在卷纸包装缺陷检测中, 外包装种类繁多, 同时不合格产品与合格产品之间可能存在较小的差异性, 为了进一步提高检测精度, 使模型能够更加适应当前检测任务。对检测模型进行优化, 优化主要包括四个部分: 1) 在主干提取网络融入卷积注意力模块(CBAM)。2) 在颈部网络中融入自适应空间特征融合模块。3) 在检测头部分引入 Decouple 模块对目标检测任务进行分解。4) 引入损失函数 α IoU 对边界框损失进行优化。下面将对各个部分进行阐述。

2.2. 卷积注意力模块

卷纸包膜种类繁多且样式多变, 合格产品与不合格产品之间差异性较小, 此外部分不合格产品突出的透明薄膜受光照影响在图像中表现不明显, 因此在特征提取网络中添加卷积注意力模块[14]以提高模型的特征提取能力。从不同样式的包装中提取出缺陷的共性, 并为其分配更大的权重, 提升模型的学习能力, 在特征提取网络添加的位置如图 2 所示。

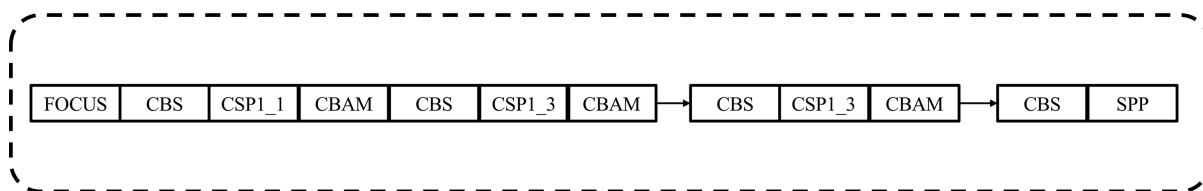


Figure 2. Adding CBAM structure diagram

图 2. 添加 CBAM 结构图

卷积注意力模块是一种用于增强卷积神经网络特征提取能力的模块, 它可以结合通道和空间两个维度的注意力机制, 自适应地调整输入特征图的权重, 从而突出有用的特征, 抑制无关的特征。它可以嵌入任何卷积神经网络中, 在不增加太多计算开销和参数数量的情况下, 提高网络在各种视觉任务上的性能[15]。卷积注意力模块包括两个子模块: 通道注意力模块和空间注意力模块。整体 CBAM 模块如图 3 所示, 首先通过一个通道注意力模块得到通道加权结果, 再经过一个空间注意力模块, 最终得到输出结果。

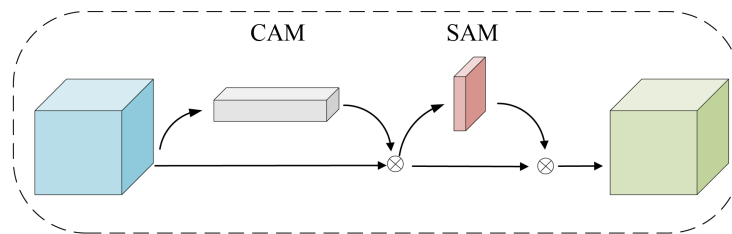


Figure 3. Spatial channel attention map

图 3. 空间通道注意力图

通道注意力模块的原理是利用全局池化操作来获取输入特征图中每个通道的全局信息, 然后通过一个共享的全连接层来学习每个通道的重要性权重, 从而提升特征的表达能力。对输入特征图进行全局平均池化和全局最大池化, 得到两个向量, 分别表示每个通道的平均值和最大值。这两种池化操作可以捕获不同的信息, 平均池化可以反映每个通道在整个空间上的程度, 最大池化可以反映每个通道在某些区域上的显著性。然后, 将这两个向量分别通过一个共享的全连接层得到两个向量。随后, 将这两个向量

相加, 并通过一个 sigmoid 激活函数, 得到通道注意力权重。最后将该向量与输入特征图按元素相乘, 得到经过通道注意力调整后的输出特征图。这个过程是自适应的, 可以根据输入图像的不同部分进行调整, 从而获取更具有区分性的特征。通道注意力结构图如图 4 所示。其数学表达式如式 1 所示。

$$M_C(F) = \sigma(MLP(Avgpool(F)) + MLP(\max Pool(F))) \quad (1)$$

其中 $M_C(F)$ 表示通道注意力权重, σ 表示 sigmoid 激活函数, MLP 表示多层感知机用于共享参数, $AvgPool$ 和 $\max Pool$ 分别表示平均池化和最大池化, F 表示输入特征图。

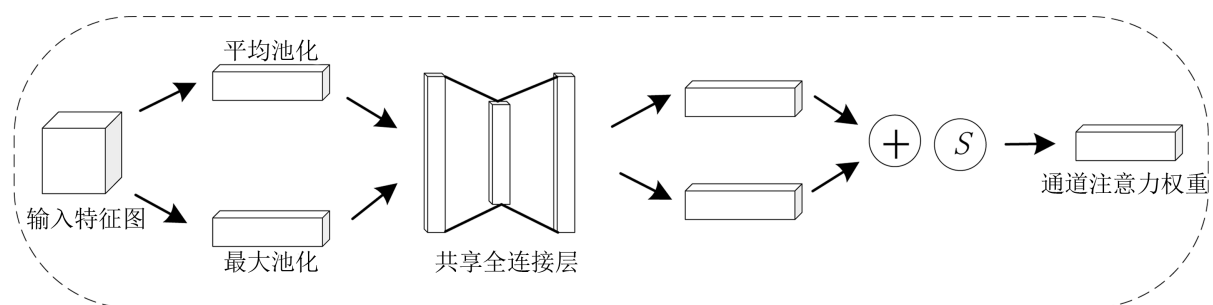


Figure 4. Channel attention chart

图 4. 通道注意力图

同样, 空间注意力机制通过对特征图在通道维度上进行压缩, 得到两个一维向量, 分别是平均池化和最大池化的结果。然后将这两个向量拼接起来, 再经过一个卷积层和一个 sigmoid 函数, 得到一个空间注意力图 M_S 。最后, 将空间注意力图与原始特征图相乘, 得到加权后的特征图。空间注意结构图如图 5 所示, 其数学表达式如式 2 所示。

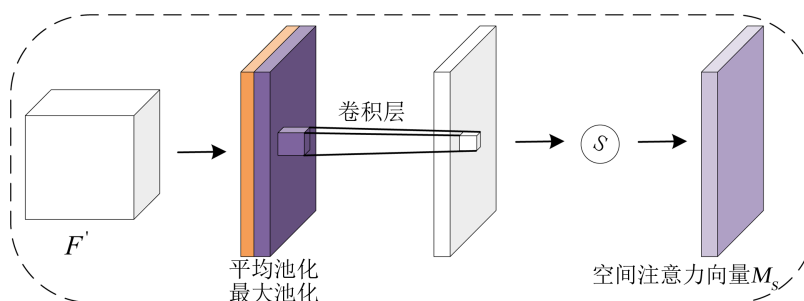


Figure 5. Spatial attention structure

图 5. 空间注意力结构

$$M_S(F) = \sigma(f^{7 \times 7}(Avgpool(F'); \max Pool(F'))) \quad (2)$$

其中 $f^{7 \times 7}$ 表示卷积核大小为 7×7 的卷积核。SAM 模块则通过在空间维度上自适应地对特征进行加权, 提高特征的局部感受野, 从而增强空间信息的表示能力。对于一张图像来说, 通道注意力可以自适应地调整每个通道的权重, 从而突出图像中重要的内容。空间注意力可以自适应地加权每个空间位置的特征, 从而突出感兴趣的区域, 并抑制背景信息的干扰, 进而提高检测模型的性能。

2.3. 自适应空间特征融合

在目标检测模型中结合高层的语义信息和低层的特征, 使得模型不仅能学习到低层次的纹理特征, 还能学习到高级的语义信息, 如单阶段多目标检测算法 SSD 首先在多级金字塔特征上检测目标, 它重用

网络前馈中计算出的多尺度特征图来预测不同尺寸的对象。YOLOv5s 颈部结构只是简单地将不同层级的特征层调进行相加, 这样会忽略了不同尺度特征之间的差异, 导致融合后的特征图质量下降和噪声增加, 影响检测效果[16]。本文引入了自适应空间特征融合模块[17] (ASFF), 可以根据不同位置 and 不同尺度的目标对象自适应地调整不同层级特征之间的权重分配, 从而避免空域冲突并提高检测精度。具体来说, 对于某个层级的特征图, 首先将其他层级的特征图调整为相同的分辨率, 并简单地将它们相加得到一个初始融合特征图。在每个空间位置, 利用一个可微分的网络模块来学习不同尺度特征图的空间重要性权重, 并将它们加权相加得到融合后的特征图。例如, 如果某个位置上携带了矛盾或冲突的信息, 则这些信息会被赋予较低的权重; 如果某个位置上携带了更多有助于区分目标类别或位置的信息, 则这些信息会被赋予较高的权重。这样就可以有效地抑制不一致性问题, 并增强目标检测性能。通过这种方式, ASFF 可以自适应地融合不同尺度的特征图, 使得模型更加灵活和鲁棒, 其结构如图 6 所示。

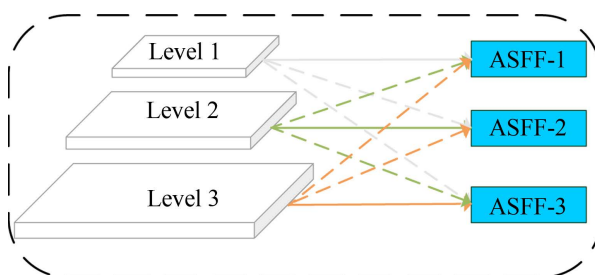


Figure 6. ASFF module
图 6. ASFF 模块

ASFF 模块以来自 FPN 三个层级的特征 x_1 、 x_2 、 x_3 作为输入, 分别与自适应权重系数 α_{ij}^l 、 β_{ij}^l 、 γ_{ij}^l 相乘得到新的融合特征。其数学表达式如公式 3 所示。

$$\begin{cases} out = x_1 \times \alpha_{ij}^l + x_2 \times \beta_{ij}^l + x_3 \times \gamma_{ij}^l \\ \alpha_{ij}^l + \beta_{ij}^l + \gamma_{ij}^l = 1 \\ \alpha_{ij}^l = \frac{e^{\lambda_{\alpha_{ij}}^l}}{e^{\lambda_{\alpha_{ij}}^l} + e^{\lambda_{\beta_{ij}}^l} + e^{\lambda_{\gamma_{ij}}^l}} \end{cases} \quad (3)$$

其中 $e^{\lambda_{\alpha_{ij}}^l}$ 、 $e^{\lambda_{\beta_{ij}}^l}$ 、 $e^{\lambda_{\gamma_{ij}}^l}$ 分别为三个特征图的控制参数, α_{ij}^l 、 β_{ij}^l 、 γ_{ij}^l 是经过归一化函数得到的, 因此它们的和是 1。因为采用相加的方式, 所以需要相加时的特征大小相同, 且通道数也要相同, 因此本文将不同层的特征上采样或下采样至 7×7 大小, 并调整通道数为 256。

2.4. Decouple 模块

目标检测是一个需要同时定位和分类目标的任务, 在定位和分类子任务中, 两者关注和感兴趣的特征各不相同, 分类更侧重于类别相关的特征, 而定位更侧重于与边界框的位置坐标匹配的特征。因此, 如果使用同一个特征图进行分类和定位, 效果往往不理想。对于分类任务而言, 类似于全连接层的分类头更具有优势。而对于定位子任务而言, 由于需要知道目标的具体空间位置, 并在边界框上计算 IoU 值, 此时, 卷积操作更占优势。

本文引入 Decouple head [18]来提升 YOLOv5 在卷纸缺陷中的检测效果。Decouple 模块如图 7 所示, 其以 ASFF 模块的输出作为输入, 经过两个 3×3 卷积得到三个分支后, 分别使用 1×1 卷积对每个分支进行降维, 最上面的分支用于预测对应像素的类别概率; 中间分支用于计算预测框的回归参数; 最下面的

分支用来预测边界框内是否存在目标。

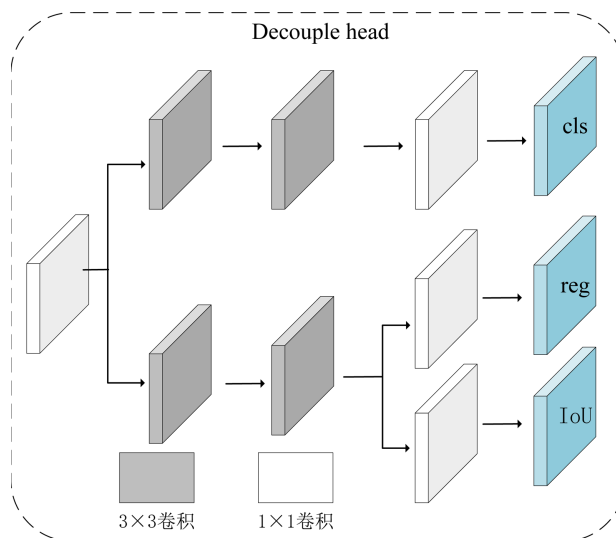


Figure 7. Decouple module
图 7. Decouple 模块

2.5. 边界框损失函数

目标检测回归和定位的基础是边界框回归, YOLOv5 原始模型所使用的边界框损失函数为 GIoU, 但在目标框被预测框包含时, GIoU 会完全退化成 IoU, 无法有效评价边界框的回归效果, 如图 8 所示, a 、 b 、 c 三种情况的 GIoU 相同, 但是回归效果不一致, 这会导致网络收敛变慢。其中 $b > a > c$ 。

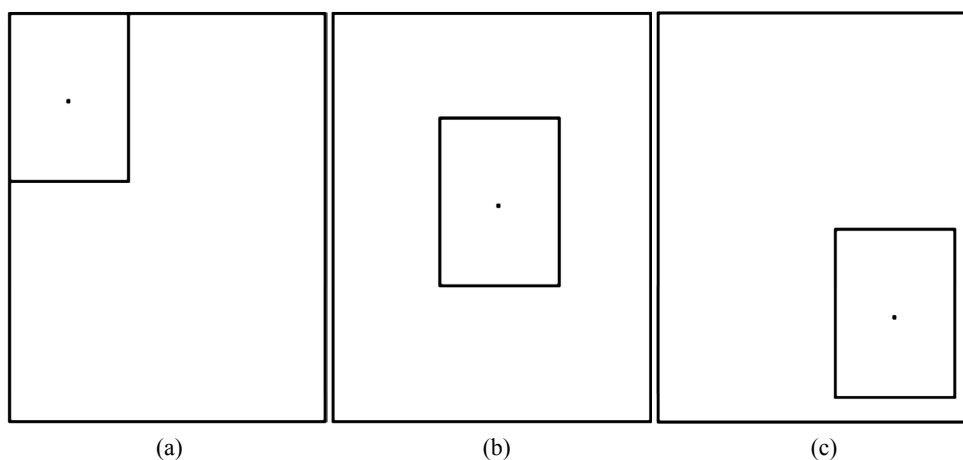


Figure 8. The return of GIoU
图 8. GIoU 回归情况

因此本文引入了 α IoU [19] 作为边界框损失函数, 该损失函数在现有的 IoU 损失函数中引入加权因子, 用于控制 IoU 损失对总损失函数的贡献, 计算公示如式 4 所示。

$$L_{\alpha\text{-IoU}} = \frac{1 - \text{IoU}^\alpha}{\alpha} \quad (4)$$

其中 α 是一个大于 0 的功率参数, 可以通过调整功率参数使模型探测器获得更高的回归精度, 提升检测

模型的鲁棒性。因为上式中的 α 是一个正常数, 所以上式可以简化为式 5:

$$L_{\alpha-IoU} = \begin{cases} -\log(IoU) & \alpha \rightarrow 0 \\ 1 - IoU^\alpha & \text{其他} \end{cases} \quad (5)$$

本文将其与 CIoU 函数相结合作为卷纸包装检测模型的边界框损失函数, 其公示如式 6 所示。相当于在 CIoU 损失函数中引入加权因子惩罚项, 使检测模型能够更加关注于高 IoU 目标, 进而提高边界框回归的精度, 同时能够加快模型的收敛速度。

$$L_{\alpha-CIoU} = 1 - IoU^\alpha + \frac{\rho^2(b, b^{gt})}{c^2} + (\beta v)^\alpha \quad (6)$$

式中 IoU 表示预测框和真实框之间的重叠度, $\rho^2(b, b^{gt})$ 表示中心点之间的欧式距离, c 表示预测框和真实框最包围框的对角线长度, β 是一个正权衡参数, 用来控制长宽比惩罚项 v 的权重, v 是一个长宽比惩罚项, 用来度量预测框和真实框之间长宽比差异, α 是一个大于 0 的功率参数。

3. 实验结果与分析

3.1. 数据集介绍

本文使用的数据集是在卷纸包装现场拍摄的, 主要用于检测卷纸包装缺陷。数据集中的部分图像是由工业相机 MV-GE231GC-T-CL 和定焦镜头 MV-LD4-6M-G 拍摄的, 部分是由手机拍摄的。采集后的图像由人工进行标注, 生成对应的 xml 文件。数据集共包含 10416 张标注图像, 其中训练集 8332 张, 测试集 2084 张。每张图像中可能有多个检测目标(多个标签)。检测包装共分为 7 个类别, 分别为: 正常包装(good)、卷纸侧倒(dao)、顶部不良(ding)、顶部露白(lu)、侧面拖纸(tuo)、未包装(bai)、侧面不良(ce)。标注示例如图 9 所示。

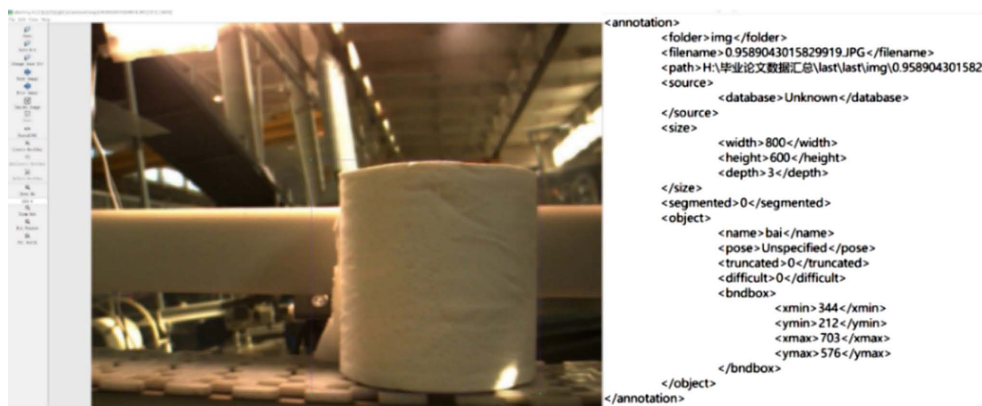


Figure 9. Example of a label

图 9. 标注示例

3.2. 数据增强

为了提高模型的鲁棒性, 在训练之前, 对部分训练图像数据进行了随机数据增强, 同时根据数据增强的方法调整了图片对应的标签。采用的数据增强方法包括: 随机翻转、随机图像缩放和色彩抖动。

3.3. 实验环境与训练策略

本文使用的软件环境: Ubuntu 18.04、CUDA 11.5; 硬件环境: GPU 是 NVIDIA GeForce RTX 3090 24

G, CPU 是至强金牌 Xeon-E5 5118 2.3 GHz。

不同的训练策略可能带来不一样的训练效果。YOLOV5 也根据不同的任务提出了不同的训练策略。本文参考了 YOLOv5 的训练预热策略, `batch_size` 设置为 16, 初始化学学习率为 0.01, 学习率在第 150 个 epoch 学习率线性增加至 0.1, 在第 150 至 300 个 epoch 内学习率以余弦方式递减至 0.01。目标检测可以理解为回归与分类, 本文沿用的 YOLOv5 自动调整锚框策略, 唯一不同的是最小锚框在原有的基础上放大了 2 倍, 因为卷纸包装一般较大, 过小的锚框在本文并不适用, 其他训练策略在 YOLOv5 的基础上没有做改动。

3.4. 实验对比与分析

本文采用准确率(P), 召回率(R), IoU 阈值大于 0.5 到 0.95 步长为 0.05 的 mAP 平均值($mAP@50:5:95$), 每秒模型可处理的帧数(FPS), 模型大小对各个模型的综合性能进行分析。其中 P 表示预测准确的结果占所有预测结果的比率, R 表示预测准确的结果占正样本的比例, P 、 R 、 $mAP@50:5:95$ 越高代表模型的检测性能越好, FPS 越高说明模型的检测速度越快, 模型大小越小则说明模型的参数越少便于部署到实际生产中。 P 、 R 和 $F1$ 的计算式如式(7)-(9)所示:

$$P = \frac{TP}{TP+FP} \times 100\% \quad (7)$$

$$R = \frac{TP}{TP+FN} \times 100\% \quad (8)$$

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (9)$$

式中 TP (True Positives)代表边界框被正确预测的卷纸数量, FP (False Positives)代表不是卷纸但被预测为卷纸的数量, FN (False Negatives)代表是卷纸但没有检测出来的数量。

在计算 mAP (mean Average Precision)时, 首先需要对每个类别单独绘制 Precision-Recall 曲线。接着, 通过计算 Precision-Recall 曲线下的积分(即 AP), 得到该类别的平均精度。最后, 对所有类别的 AP 求算术平均, 就是 mAP。

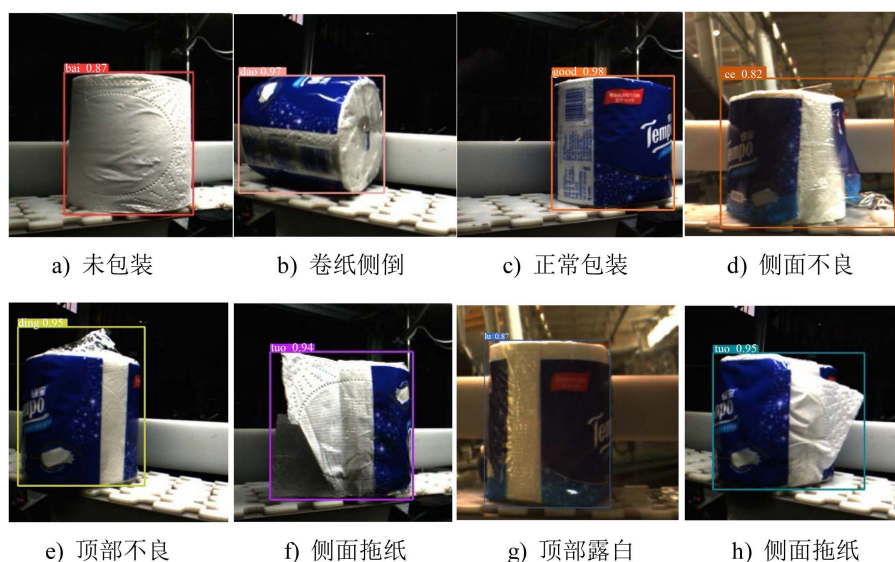


Figure 10. Visualisation of results
图 10. 可视化结果

实验过程中,我们将部分测试结果进行了可视化,结果如图 10 所示。我们将卷纸包装分为以下七类,具体有未包装(如图 10(a)所示)、包装后卷纸倒下(如图 10(b)所示)、正常包装(如图 10(c)所示)、侧面不良问题(如图 10(d)所示),顶部包装问题(如图 10(e)所示),侧面拖纸问题如图 10(f)和图 10(h)所示,顶部露白问题(如图 10(g)所示)。从图中可以看出,本文提出的方法基本能识别出常见的卷纸包装缺陷类别。

为了验证模型的有效性,接下来将本文设计的模型均与单阶段目标检测网络进行对比,分别是 YOLOv3、SSD 及李志诚等人提出的卷纸包装检测算法,从多个维度上进行对比并对结果进行分析。在卷纸包装数据集上进行实验,对比实验结果如表 1 所示。

其中 YOLOv3 和 iYOLOv3 的图片输入分辨率为 608×608 。目标检测任务的核心任务可以概括为识别物体的类别、物体的位置。由表 1 可知,与其他单阶段检测模型相比,在准确率 P 、平均准确率 mAP 和 $F1$ 值上均取得了最优的结果。与原始模型相比,改进后的 YOLOv5 的 P 和 mAP 分别提升了 2.1%和 3.8%,帧率 FPS 值虽然低于原始模型,但是仍然能够达到较高的帧率,这可能是因为在原始模型的基础上集成了多个模块,增加了一定的计算量,使得帧率有所下降。相比于李志诚等人提出的 iYOLOv3 卷纸包装缺陷算法,本文所提出的 FYOLOv5 在准确率和平均准确率上分别提升了 4.1%和 5.8%,数据集数量也有所提升,数据集中所包含的缺陷种类也增加了两种,能够检测的卷纸包装缺陷更加全面。

Table 1. Performance test results of roll paper packaging defect detection model

表 1. 卷纸包装缺陷检测模型性能测试结果

Method	$P/\%$	$R/\%$	$mAP@50:5:95/\%$	$F1/\%$	FPS
YOLOv3	85.6	87.6	81.7	86.6	62
YOLOv5	93.2	92.9	88.6	92.6	81
SSD300	80.2	85.2	83.4	82.9	36
iYOLOv3	91.2	90.4	86.6	90.8	96
FYOLOv5	95.3	93.2	92.4	94.2	71

Table 2. Effect of different modules on model checking performance

表 2. 不同模块对模型检测性能的影响

Num	De	αIoU	CBAM	ASFF	$mAP@50:5:95/\%$
1	×	×	×	×	88.6
2	√	×	×	×	88.9
3	√	√	×	×	89.1
4	√	√	√	×	91.6
5	√	√	√	√	92.4

为了验证各个模块的有效性,我们系统分析了不同模块组合对各个评价指标的影响。其中表中“√”表示在网络模型中添加了该模块,“×”表示在网络模型中未使用该模块,ASFF 表示自适应空间特征融合模块,CBAM 表示卷积注意力块,De 表示 Decouple 模块, αIoU 表示 αIoU 损失函数。

由表 2 可知,通过对比第 1、2 组实验可知,Decouple 模块降低了模型处理问题的难度,给 mAP 值带来了 0.3%的提升,但也提升了模型的计算量,使模型检测速度下降了 7 frames/s。通过对比第 2、3 组

实验可知, α IoU 损失函数的引入增强了模型的鲁棒性, 在预测框定位时更加精准, 给检测模型带来了 mAP 值 0.2% 的提升, 略微提升了模型的检测速度。通过对比 3、4 组实验可知, 添加了卷积注意力模块后, 模型 mAP 值提升了 1.5%, 这说明卷积注意力模块能够帮助模型更好地关注到一些重要特征, 提升模型的特征提取能力。通过对比 4、5 组实验可知, 自适应空间特征融合的引入, 使得模型能从多个尺度学习到不同缺陷的特征, 给 mAP 值 0.8% 的提升。综合以上模块的改进, 相比于原始检测模型的 mAP 值提升了 3.8%, 检测速度达到了 71 frames/s, 满足实际检测需求。因此, 改进后的 FYOLOv5 模型能够较好的适用于卷纸包装缺陷检测。

4. 总结

针对卷纸包装缺陷检测任务, 本文在现有目标检测算法的基础上, 提出了一种结合注意力机制和多尺度特征融合的卷纸包装缺陷检测算法。使用卷积注意力模块提高模型对空间和通道特征的建模能力, 增强模型对合格品与不合格品之间差异的关注度; 利用自适应空间特征融合模块让模型具备更强的多尺度特征融合能力, 从不同层次学习到缺陷的特征; 引入 Decouple 模块对目标检测任务进行分解, 使得模型更加关注任务本身; 最后引入 α IoU 对原边界框损失函数进行优化, 提升模型的鲁棒性。实验证明, 改进后的检测模型准确率 95.3%, 检测速度 71 frames/s, 能够满足卷纸包装缺陷检测的任务需求。同时, 我们的实验也为深度学习在工业应用中提供了一定的借鉴作用。

基金项目

广东省基础与应用基础研究基金(2020A1515011468); 广东省普通高校特色创新类项目(2019KTSCX189)。

参考文献

- [1] 钱桂敬. 新形势下我国造纸工业高质量发展面临的挑战[J]. 造纸信息, 2022(2): 18-21.
- [2] Zhang, X.-W., Ding, Y.-Q., Lv, Y.-Y., Shi, A.-Y. and Liang, R.-Y. (2011) A Vision Inspection System for the Surface Defects of Strongly Reflected Metal Based on Multi-Class. *Expert Systems with Applications*, **38**, 5930-5939. <https://doi.org/10.1016/j.eswa.2010.11.030>
- [3] Kuang, H., Ding, Y., Li, R. and Liu, X. (2018) Defect Detection of Bamboo Strips Based on LBP and GLCM Features by Using SVM Classifier. 2018 *Chinese Control and Decision Conference (CCDC)*, Shenyang, 9-11 June 2018, 3341-3345. <https://doi.org/10.1109/CCDC.2018.8407701>
- [4] 刘贺, 马小燕, 张静. 基于数据挖掘技术的牙刷包装缺陷检测方法研究[J]. 佳木斯大学学报(自然科学版), 2020, 38(3): 48-51.
- [5] 于惠钧, 吴婉, 成运. 改进 Otsu 算法在铝塑泡罩药品包装缺陷检测中的应用[J]. 包装工程, 2014, 35(15): 15-18.
- [6] Wang, C.-Y., Liao, H.-Y.M., Wu, Y.-H., et al. (2020) CSPNet: A New Backbone That Can Enhance Learning Capability of CNN. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, 14-19 June 2020, 1571-1580. <https://doi.org/10.1109/CVPRW50498.2020.00203>
- [7] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. ArXiv Preprint ArXiv: 1804.02767.
- [8] 刘志赢, 谢春思, 李进军, 桑雨. Cascade R-CNN 和 YOLOv3 在导弹目标识别中的应用[J]. 图像与信号处理, 2020, 9(2): 102-110.
- [9] 曾秀云, 陆华才, 吕丰. 基于改进 Faster R-CNN 的棉布包装缺陷检测的方法研究[J]. 电子测量与仪器学报, 2022, 36(4): 179-186.
- [10] 李建明, 杨挺, 王惠栋. 基于深度学习的工业自动化包装缺陷检测方法[J]. 包装工程, 2020, 41(7): 175-184.
- [11] 李志诚, 曾志强. 基于改进 YOLOv3 的卷纸包装缺陷实时检测算法[J]. 中国造纸学报, 2022, 37(2): 87-93.
- [12] 曹义亲, 伍铭林, 徐露. 基于改进 YOLOv5 算法的钢材表面缺陷检测[J]. 图学学报, 2023, 44(2): 335-345.
- [13] 许思昂, 李艺杰, 梁桥康, 杨彬. 基于改进 YOLOv5 算法的 PCB 裸板缺陷检测[J]. 包装工程, 2022, 43(15): 33-41.

- [14] Woo, S., Park, J., Lee, J.-Y. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018. ECCV 2018. Lecture Notes in Computer Science*, Vol. 11211, Springer, Cham, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [15] Alzubaidi, L., Zhang, J., Humaidi, A.J., *et al.* (2021) Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *Journal of Big Data*, **8**, Article No. 53. <https://doi.org/10.1186/s40537-021-00444-8>
- [16] Zhu, X., Lyu, S., Wang, X. and Zhao, Q. (2021) TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-Captured Scenarios. *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, 11-17 October 2021, 2778-2788. <https://doi.org/10.1109/ICCVW54120.2021.00312>
- [17] Liu, S., Huang, D. and Wang, Y. (2019) Learning Spatial Fusion for Single-Shot Object Detection. ArXiv Preprint ArXiv: 1911.09516.
- [18] Ge, Z., Liu, S., Wang, F., Li, Z. and Sun, J. (2021) Yolox: Exceeding Yolo Series in 2021. ArXiv Preprint ArXiv: 2107.08430.
- [19] He, J., Erfani, S., Ma, X., *et al.* (2021) Alpha-IoU: A Family of Power Intersection over Union Losses for Bounding Box Regression. <https://arxiv.org/abs/2110.13675>