

基于改进YOLOX的编织袋表面缺陷检测算法

胡磊¹, 刘冬伟¹, 王可欣², 汪凯³, 王振伟^{4*}

(1 黄石邦柯科技股份有限公司, 湖北 黄石 435000; 2 中国地质大学(武汉) 资源学院, 武汉 430074;
3 湖北汽车工业学院 电气与信息工程学院, 湖北 十堰 442000; 4 大连理工大学 计算机科学与技术学院,
辽宁 大连 116024)

摘要 为解决工业生产线上编织袋缺陷检测仅依赖传统人工目视检测方法耗时费力且漏检率高等问题, 提出了一种用于编织袋生产线的轻量级实时缺陷检测算法GT-YOLOX (Ghost-TwoScale-YOLOX). 该算法利用GhostNet作为骨干网络, 嵌入了CA (Coordinate Attention) 注意力机制加强特征学习, 同时设计了一种双层特征融合策略, 以降低模型计算量并提升检测速度. 该算法已经在湖北黄石华新水泥袋智能剔除项目中完成有效验证. 结果表明: GT-YOLOX算法可对编织袋特定缺陷的平均检测精度达到了95.22%. 同时, 在单CPU设备上, 每张图像的检测速度缩短至67毫秒, 相较于标准YOLOX模型, 推理时间提升了34.31%, 模型权重减少了近50%. 说明GT-YOLOX算法能够有效满足生产线上编织袋缺陷检测的快速且高精度要求.

关键词 表面缺陷检测; YOLOX算法; CA注意力机制; 轻量级模型

中图分类号 TP391.4 文献标志码 A 文章编号 1672-4321(2025)05-0675-10

doi:10.20056/j.cnki.ZNMDZK.20250513

Surface defect detection algorithm for woven bags based on improved YOLOX

HU Lei¹, LIU Dongwei¹, WANG Kexin², WANG Kai³, WANG Zhenwei^{4*}

(1 Huangshi Bangke Technology Company Limited, Huangshi 435000, Hubei China; 2 School of Resources, China University of Geosciences (Wuhan), Wuhan 430074, China; 3 School of Electrical and Information Engineering, Hubei Institute of Automotive Industry, Shiyan 442000, Hubei China; 4 School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, Liaoning China)

Abstract In order to solve the problems of time-consuming and labor-intensive traditional manual visual inspection methods and high missed detection rate in woven bag defect detection on industrial production lines, a lightweight real-time defect detection algorithm GT-YOLOX (Ghost-TwoScale-YOLOX) for woven bag production lines is proposed. The algorithm uses GhostNet as the backbone network, embeds the CA (Coordinate Attention) attention mechanism to strengthen feature learning, and designs a two-layer feature fusion strategy to reduce the model calculation amount and improve the detection speed. The algorithm has been effectively verified in the Hubei Huangshi Huaxin cement bag intelligent rejection project. The results show that the GT-YOLOX algorithm can achieve an average detection accuracy of 95.22% for specific defects of woven bags. At the same time, on a single CPU device, the detection speed of each image is shortened to 67 milliseconds. Compared with the standard YOLOX model, the inference time is increased by 34.31%, and the model weight is reduced by nearly 50%. It shows that the GT-YOLOX algorithm can effectively meet the requirements of fast and high-precision defect detection of woven bags on production lines.

Keywords surface defect detection; YOLOX; CA attention mechanism; lightweight model

收稿日期 2024-07-17

* 通信作者 王振伟 (1997-), 男, 博士生, 研究方向: 深度学习, 计算机视觉、图像处理、视频行为识别等, E-mail: zhenweiwang@mail.dlut.edu.cn

基金项目 吉林省科技发展计划资助项目 (20230201111GX)

编织袋广泛用于农业、工业、运输和建筑等领域的物资外包装,与人类的生产生活密切相关.在编织袋的生产过程中,受生产工艺、机械故障、人为操作失误等多方面因素影响,会在表面产生孔洞、拉丝、污点等缺陷.这些缺陷降低了编织袋的质量,有缺陷的编织袋流入市场除了造成经济损失之外,还会对运输过程和人身安全造成威胁^[1].传统质量检测通常依赖人工方法,然而人工肉眼检测不仅耗时费力、漏检率高且易受操作员主观因素影响^[2].因此,本文提出了一种用于编织袋生产线的轻量级实时缺陷检测算法 GT-YOLOX (Ghost-TwoScale-YOLOX).首先,该算法利用 GhostNet 作为骨干网络,以增强特征提取和减小模型计算量.其次,引入了 CA (Coordinate Attention) 注意力机制加强特征细节位置信息提取与跨通道特征融合,提升了检测精度.最后,设计了一种双层特征融合策略,进一步降低模型计算量并提升检测速度.

1 引言

近些年,随着卷积神经网络 (Convolutional Neural Network, CNN) 的发展,基于深度学习的方法已逐步运用于目标检测等领域^[3-7].目前,广泛应用于目标检测领域的深度学习主要方法分为两大类:两阶段 (Two-stage) 目标检测算法和一阶段 (One-stage) 目标检测算法.两阶段目标检测算法也被称为基于区域的方法,其工作流程通常包括两个阶段:首先由算法生成一系列候选边界框作为样本,然后再通过卷积神经网络进行分类.两阶段目标检测算法的代表网络包括 R-CNN^[8]、Fast R-CNN^[9]、Faster R-CNN^[10]、R-FCN^[11]等;一阶段目标检测算法则直接将目标检测问题转化为回归预测问题,将整张图片作为网络的输入,直接输出预测框的类别和坐标信息.一阶段目标检测算法的代表网络包括 SSD^[12]、YOLO^[13]、YOLOv2^[14]、YOLOv3^[15]、YOLOv4^[16]、YOLOv5、YOLOX^[17]等.相比于两阶段网络,一阶段的 YOLO 系列网络具有模型轻量化、推理速度快、端到端等优势.因此,基于深度学习的一阶段目标检测算法已成为一种研究热点.

鉴于深度学习方法具有泛化性强、精度高、运算速度快等优势,目前基于深度学习的缺陷检测方法已被广泛运用于工业领域中多个不同场景^[18-19].例如,HE 等人^[20]通过所提出的多级特征融合网络

(MFN) 融合多尺度特征,提取更多的位置细节信息,提升了钢材表面缺陷检测精度.ZHANG 等人^[21]采用 YOLOv3 检测桥梁缺陷,利用迁移学习与 Focal Loss 去进一步提升模型的检测性能.CHEN 等人^[22]利用密集连接卷积网络 DenseNet 与 YOLOv3 结合,增强网络的特征提取能力,用以检测 LED 表面缺陷.WANG 等人^[23]采用新的金字塔特征融合设计,提出基于改进 Mask R-CNN 的钢轨表面缺陷检测.LU 等人^[24]使用改进的 Shufflenetv2 主干进行特征提取,通过引入深度可分离卷积对 YOLOv5s 中 Bottleneck 模块进行优化,提出基于改进的 YOLOv5 的瓷砖表面缺陷智能检测方法.以上工作对基于深度学习的不同的工业场景上的缺陷检测进行了研究与探讨,但目前很少有研究者关注工业生产线上编织袋缺陷检测工作.为此,本文在上述研究工作的启发下,开展基于深度学习的工业生产线上编织袋缺陷检测方法研究.

与其他缺陷检测任务不同的是,工业生产线上编织袋缺陷检测需要考虑特定任务需求、数据样本以及模型部署环境的特点,这些对编织袋缺陷检测任务提出了不同的挑战.同时,编织袋容易受到环境温度、湿度和人工操作不当等因素的影响而发生变形,从而引入样本成像时的各种噪音,这进一步对检测算法的鲁棒性提出了要求.目前,针对编织袋缺陷检测方法主要是基于机器学习方法,基于机器学习的编织袋缺陷检测方法虽然理论简单、易于实现、模型可解释性较好,但分割精度差、训练时间较长且不适合大规模的数据样本.2021年,有研究者提出的基于遗传算法与二维最大熵的编织袋缺陷检测算法^[25],该方法虽然能够较为精准地检测出编织袋的质量缺陷,一定程度上提高了分割精度与速度,但是在实际应用中遗传算法需要大量样本才能求解,这导致了较高的时间成本.此外,遗传算法可能会陷入局部最优解,导致无法得到全局最优解.再如,ZHONG 等人^[29]提出的基于机器视觉的编织袋表面缺陷检测系统设计与实现方法也是基于遗传算法改进而来,该方法可有效提高编织袋缺陷检测的精度问题,但是在检测速度上较慢,处理一幅编织袋图像的时间约为 600 ms,这无法满足工业生产线上效率需求.2022年,GAO 等人^[4]提出的基于机器视觉编织袋缝合缺陷识别与检测方法,针对生产线编织袋密封检测技术的缺乏,设计研发了编织袋缝合缺陷识别与检测.该算法首先要通过二值化、去噪、边缘检测等算法进行图像校正,然后使用集成

模型中的RUSBoostedTrees算法进行缺陷分类.虽然在一定程度上提高检测效率,但是在移动端实际部署和缺陷检测时间上仍无法满足现代工业要求.基于机器学习的编织袋缺陷检测方法部署在移动设备上普遍存在参数量大、计算量大等问题,导致检测速度较慢.同时,仅追求网络结构轻量化的模型,虽检测速度快,但精度却不足,无法满足工业实时高精度检测的需求.

因此,本文针对工业生产线上编织袋缺陷检测的精度、效率和算法部署移动端上的不足,基于深度学习设计了一种编织袋缺陷检测的轻量级目标检测算法Ghost-TwoScale-YOLOX(GT-YOLOX),旨在精准识别编织袋缺陷,满足实际生产线上实时高精度检测需求.该模型以YOLOX为基础,集成了GhostNet作为骨干网络进行特征提取,同时引入了Coordinate Attention(CA)注意力机制进行特征提取与融合.本文主要创新点如下:(1)提出了一种高性能编织袋

缺陷检测的轻量级目标检测算法GT-YOLOX.利用轻量级网络GhostNet作为骨干网络进行特征提取,同时降低模型复杂度.(2)设计了嵌入CA注意力的双层特征融合模块.利用CA注意力强化重要通道的特征和细节位置信息,同时设计双层的特征融合,用以降低网络计算量并提升检测速度.(3)大量实验证明,本文方法在实际生产线上编织袋缺陷检测的精准度高、实时性强、鲁棒性好.同时,本方法的轻量级参数可更加便捷地部署到工业生产线上,以提高生产质量.

2 YOLOX 网络模型

YOLOX作为YOLO系列发展历程的重要版本之一,在实时检测方面拥有十分优越的性能.YOLOX网络结构如图1所示,其主要包括4个部分:输入端Input、Backbone主干网络、Neck网络和Prediction部分.

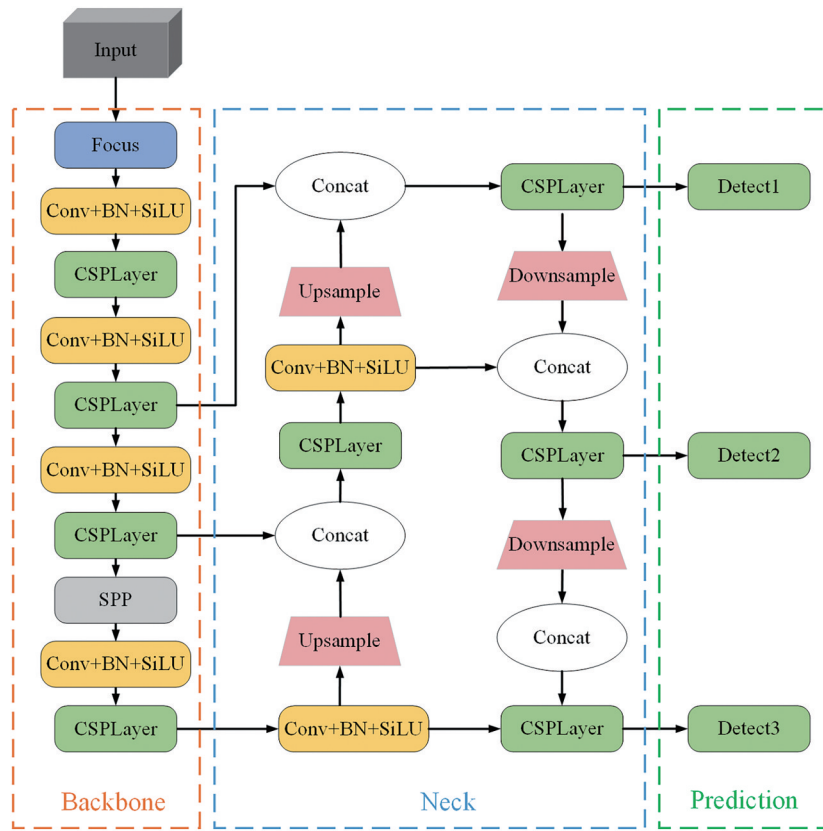


图1 YOLOX网络结构^[17]

Fig. 1 YOLOX Network Architecture^[17]

在输入端,YOLOX采用了Mosaic和Mixup作为数据增强方式.Mosaic通过随机缩放、裁剪等方式拼接提高对小目标的检测效果;Mixup通过随机线性插值混合训练样本和标签生成新的训练样本以扩充数据集,增强模型的鲁棒性.

在网络结构上,YOLOX选用CSPDarknet作为特征提取的主干网络,Backbone网络主要包括Focus、CSPLayer和SPP三个结构.其中,Focus结构的作用是对图像进行切片,选择图像中的其他像素,使高度和宽度信息集中在通道中,防止信息丢失;CSPLayer

主要利用了残差结构,将特征图分为两部分,一部分进入残差块进行特征提取,另一部分直接与特征提取后的部分进行 Concat 操作;SPP 部分延用了 YOLOv3-SPP 版的 SPP (Spatial Pyramid Pooling) 结构,SPP 结构借鉴了空间金字塔池化思想,使用不同尺度的池化层对输入特征图分块采样和池化操作,融合局部特征和全局特征。

在 Neck 部分,YOLOX 使用 FPN (Feature Pyramid Networks) + PAN (Path Aggregation Network) 的结构,具有较强的语义特征和定位特征.通过 FPN 自上而下的特征传递和 PAN 自下而上的上下文信息利用,使得网络能够更好地理解图像中的目标和背景关系。

在 Prediction 部分,YOLOX 在 YOLO 系列中首次采用解耦检测头提高准确率,不同的卷积层分别处理分类、定位、置信度预测.这种设计在仅增加少量网络参数的情况下提高了网络收敛速度。

3 GT-YOLOX 网络模型

为了满足工业生产线上实时检测速度和精度的需求,同时考虑到模型轻量化的需要,本文基于

YOLOX-s 进行了如下改进:首先,利用轻量级网络 GhostNet 替换原始 YOLOX-s 中的骨干网络,减少计算量和内存占用,提高模型运行速度.然后,对特征融合结构进行剪枝操作,设计双层特征融合模块以减少小目标尺度的特征融合层,进一步降低网络计算量并提高检测速度.最后,嵌入 CA 注意力机制以加强模型对缺陷位置细节信息的关注,进一步提升准确率。

本文所提出的 GT-YOLOX 基本结构如图 2 所示,GT-YOLOX 算法首先利用主干网络 GhostNet 进行特征提取,将输入大小 $640 \times 640 \times 3$ 的图像经过 12 层 G-bottleneck 模块进行特征提取,输出图像大小为 $20 \times 20 \times 160$ 的特征.然后,将大小分别为 $40 \times 40 \times 112$ 和 $20 \times 20 \times 160$ 的输出特征进行双层特征融合,进一步加强多尺度特征提取与融合,提高特征的表达能力.在双层特征融合中,输入的特征在上下两条路径中首先进行 CA 注意力的操作,加强关注特征的细节和位置信息,进一步提高算法的识别率.其次,上下两条路径为了进行特征融合,下方、上方路径中特征分别进行上、下采样操作.然后,上下路径特征分别送入 CBL 模块进行特征变化.最后,经过卷积操作、Sigmoid 激活函数和上下路径输出的 Contact 后,得到最终的输出。

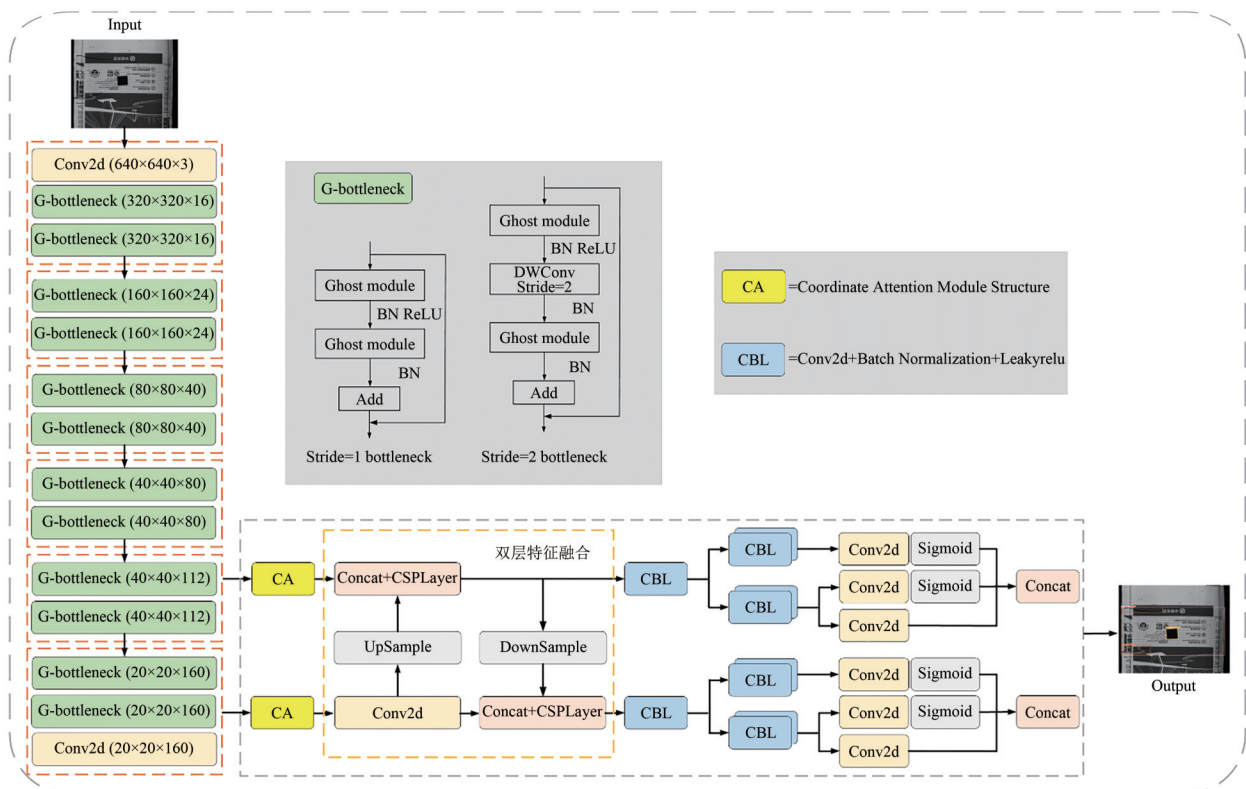


图 2 GT-YOLOX 结构

Fig. 2 The structure of GT-YOLOX

3.1 基于 GhostNet 的轻量级骨干网络

当前,传统的特征提取网络由于结构较深、计算复杂等问题,在处理特征较为明显的图像数据时会存在大量的特征冗余,这会造成时间和内存空间的浪费.因此,本文采用更轻量级的特征提取网络作为主干效率会更高.GhostNet 是一种轻量级的卷积神经网络,具有运行快、准确率高等优点.GhostNet 引入了 Ghost 模块, Ghost 模块会利用普通卷积操作生成一组特征图,再对生成的特征图进行线性变换生成冗余特征图,最后将普通卷积生成的特征图与线性变换生成的特征图进行连接操作,确保最终输出的内部特征图与普通卷积输出的特征图保持一致.GhostNet 能够通过这种方式使用更少的参数生成更多有效的特征.以 n 个特征映射的输出为例,假设输入特征映射的大小为 $h \times w \times c$,输出特征层的大小为 $h' \times w' \times n$,卷积核的大小为 k .首先,通过普通卷积得到 $m = \frac{n}{s} (m < n)$ 个特征层;第二步,通过线性运算获得 $m \times (s - 1) = \frac{(s - 1)n}{s}$ 个特征层;第三步,将前两部分的通道数相加,得到 $m + m(s - 1) = m \times s$ 个输出特征图.由于 $m = \frac{n}{s}$,所以输出的特征映射仍然是 n , s 用于调整传统卷积特征映射的个数.普通卷积所需的参数量为 cost_1 , Ghost 模块所需的参数量为 cost_2 ,压缩比为 r_c ,计算公式如下:

$$\text{cost}_1 = h' \times w' \times n \times k \times k \times c, \quad (1)$$

$$\text{cost}_2 = h' \times w' \times \frac{n}{s} \times k \times k \times c + (s - 1) \times h' \times w' \times \frac{n}{s} \times k \times k, \quad (2)$$

$$r_c = \frac{\text{cost}_1}{\text{cost}_2} = \frac{n \times c}{\frac{n}{s} \times c + (s - 1) \frac{n}{s}} = \frac{s \times c}{c + s - 1} \approx s, \quad (3)$$

根据上述公式和大部分实验结果显示,使用 Ghost 模块理论上可以节省 5 倍的训练时间和参数.因此,使用 Ghost 模块构建的 GhostNet 更适合轻量级的缺陷检测任务.GhostNet 骨干网框架及参数如下表 1 所示,其中, G-bottleneck 是 GhostNet 中的瓶颈结构,由

表 1 GhostNet 参数配置
Tab. 1 GhostNet configuration

Input Size	Structure	Step
$640 \times 640 \times 3$	Conv2d(3 × 3)	2
$320 \times 320 \times 16$	G-bottleneck × 2	1, 2
$160 \times 160 \times 24$	G-bottleneck × 2	1, 2
$80 \times 80 \times 40$	G-bottleneck × 2	1, 2
$40 \times 40 \times 80$	G-bottleneck × 2	1, 1
$40 \times 40 \times 112$	G-bottleneck × 2	1, 2
$20 \times 20 \times 160$	G-bottleneck × 2	1, 1
$20 \times 20 \times 160$	Conv2d(3 × 3)	1

Ghost 模块、批标准化层叠加而成其用计算量很少的线性运算代替部分卷积运算,将得到的冗余特征图与本征特征图进行拼接,提高网络的特征表达能力.

3.2 双层特征融合

双层特征是由对原始图像进行卷积下采样等操作得到高分辨率和低分辨率的特征层.高分辨率特征层包含更多的位置和细节信息,但其语义更低,噪声更大.低分辨率特征层的特征信息更多,语义信息更强,但对细节的感知能力较差.为了使执行预测输出的特征层信息量更大,融合不同尺度的特征层有利于提高检测性能.

在标准的 YOLOX 颈部网络(Neck)中,包含三层不同大小的特征: 80×80 、 40×40 以及 20×20 ,其中, 80×80 层主要用于较小目标的检测.根据对本实验数据集的观察,编织袋上需要检测的缺陷特征明显,缺陷目标边界框面积与原图面积之比较大,不属于小目标缺陷.因此,本实验中 80×80 分辨率的有效特征层不参与推理计算.此外,大量实验表明^[25-27],在大型目标检测任务上,三层尺度的输出有大量的计算冗余等不足,而采用两层尺度结构将减少这种冗余现象,从而加快推理速度并保证精度与之相当.

本文改进标准 YOLOX 颈部网络(Neck)中结构,采用嵌入 Coordinate attention(CA)注意力的双层特征融合模块进行特征的进一步提取与融合操作,其结构如图 3 所示.

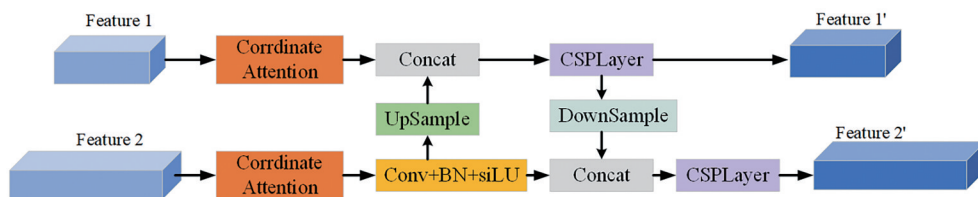


图 3 融合 CA 注意力机制的双层特征融合结构图

Fig. 3 Structure of two-layer feature fusion with fused Coordinate Attention mechanism

在图 3 中, feat1 代表下采样为 16 倍的特征层, 大小为 40×40 , feat2 代表下采样为 32 倍的特征层, 大小为 20×20 . Conv2D 模块为普通卷积模块, 用于降低通道数; UpSample 为双线性差值的上采样操作, 用于扩大特征层宽高; CSPLayer 结构代表跨阶段局部网络(Cross Stage Partial Network), 是将原输入分成两个分支, 分别进行卷积操作使得通道数减半, 一个分支进行多次残差结构操作. 然后将两个分支拼接, 使模型学习到更多的特征; DownSample 为下采样操作, 用于压缩特征层宽高. 相比于三层尺度输出, 可以减少 6400 个网格解码计算操作, 计算量上可以

减少 70%.

3.3 CA 注意力机制层特征融合

为了获得对图像宽度和高度的关注, 并对精确的位置信息进行编码, 本文采用了 Coordinate attention (CA) 注意力机制. 通过将位置信息嵌入到通道注意力中, CA 注意力机制使轻量级网络能够在更大的区域上进行注意力, 避免产生大量计算开销. 同时, CA 注意力机制还能捕获方向感知和位置感知的信息, 使模型更加关注目标所在区域, 减少背景干扰, 提高模型精度. CA 模块结构图如图 4 所示.

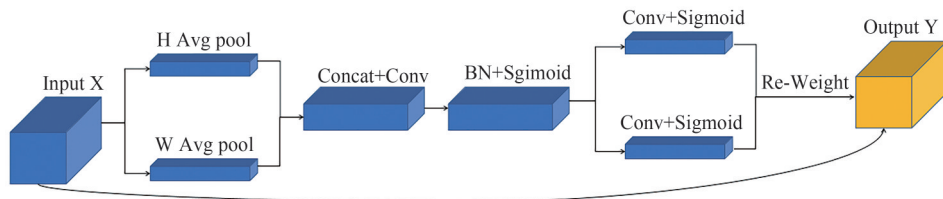


图 4 CA 模块结构图

Fig. 4 Coordinate Attention module structure

如图 4 所示, CA 注意力机制首先将输入的特征图分成宽度和高度两个方向进行全局平均池化. 设输入特征层为 $X, X = [x_1, x_2, \dots, x_c] \in R^{H \times W \times C}$, 输出特征层为 $Y, Y = [y_1, y_2, \dots, y_c] \in R^{H \times W \times C}$, H 为特征层高度, W 为特征层宽度, C 为特征层数. 首先, 分别获得宽度和高度两个方向上的特征图 z^h 和 z^w :

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i), \quad (4)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w), \quad (5)$$

其次, 拼接两个特征图并送入卷积核为 1×1 的卷积模块中进行降维. 降维后, 将经过批处理归一化处理的特征图 F_1 送入 Sigmoid 激活函数进行激活, 得到特征图 f :

$$f = \sigma(F_1([z^h, z^w])), \quad (6)$$

将 f 按原来的高和宽分别进行 1×1 卷积得到的特征图送入 Sigmoid 激活函数进行激活:

$$g^h = \sigma(F_h(f^h)), \quad (7)$$

$$g^w = \sigma(F_w(f^w)), \quad (8)$$

经过上述计算, 将得到输入特征图高度方向上的注意注意力权重 g^h 和宽度方向上的注意权重 g^w . 最后, 对原始特征图进行乘法加权计算, 并将最终的宽度和高度方向与特征图的注意权重相结合:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j). \quad (9)$$

3.4 损失函数

YOLOX 在 Head 部分包括分类分支、边界框回归分支和置信度分支. 其中, 分类分支和置信分支采用 BCELoss 损失函数, 如式 (10) 所示.

$$L = \frac{1}{N} \sum_{i=1}^N L_i = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^M y_{ic} \log(p_{ic}), \quad (10)$$

其中, N 表示样本数量, M 为样本的类别数, y_{ic} 表示符号函数 (即如果样本 i 的实类别等于 c 类, 则值为 1, 否则值为 0), p_{ic} 为观测样本 i 属于 c 类的预测概率.

边界框回归使用 GIOU loss 损失函数, 如式 (11)-(13) 所示:

$$\text{Loss}_{\text{coord}} = \sum_{i=0}^S \sum_{j=0}^B I_{ij}^{\text{obj}} (1 - \text{GIOU}_{ij}), \quad (11)$$

$$\text{GIOU} = \frac{J}{U} - \frac{A - U}{A}, \quad (12)$$

$$U = \hat{w}_i \hat{h}_i + w_i h_i - J, \quad (13)$$

其中, $\text{Loss}_{\text{coord}}$ 为目标位置损失函数, I_{ij}^{obj} 表示由单元格 i 生成的先验框 j 包含目标, J 为预测框和真实框的相交面积, U 为预测框和真实框的并集面积, A 为包含预测框和真实框的最小矩形面积, \hat{w}_i 和 \hat{h}_i 表示预测框的宽度和高度, S 表示真实框和预测框的最小限定矩形面积, B 表示真实框和预测框的组合面积, \hat{w}_i 和 \hat{h}_i 表示 GIOU 损失函数中真实框的宽和高.

4 实验结果及分析

4.1 数据集制作与实验设置

由于编织袋缺陷检测任务缺少公开数据集,为保证数据集的有效性,本文利用湖北黄石华新水泥袋智能剔除项目中生产线上的水泥编织袋构建了数据集.本文使用工业相机,共采集3000张图像样本进行实验分析.

其中,经过工业相机采集的原图如图5(a)所示,

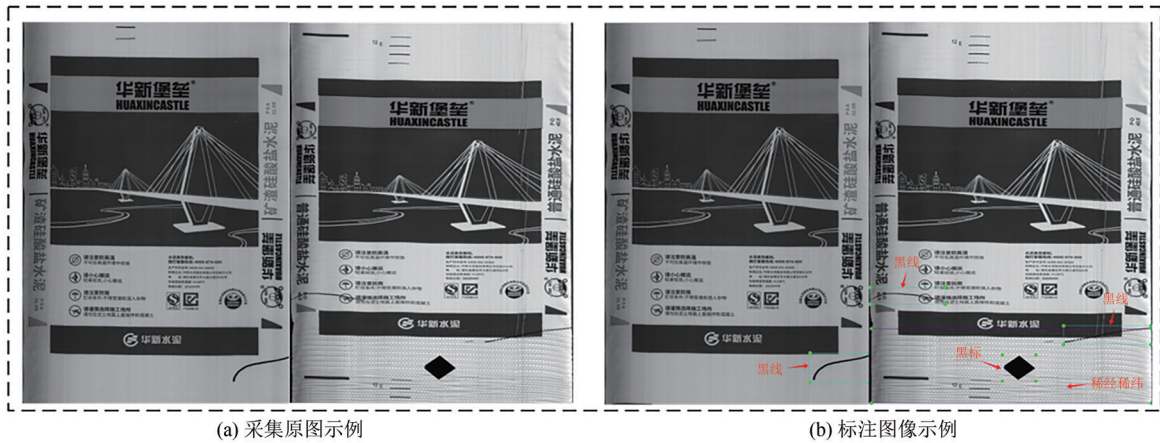


图5 采集的原图与标注图示例

Fig. 5 Examples of collected original and annotated images

在本研究中,模型使用的CPU型号为Intel Core i9-9900KF,GPU型号为NVIDIA RTX3090.深度学习框架基于PyTorch 1.10搭建,Python版本为3.8.为保证模型对比结果科学性,在相同硬件平台下使用相同训练参数进行对比实验,使用SGD优化器,batchsize设为32,epoch设为300,输入图像尺寸为640×640.

4.2 评价指标

在目标检测中,常用的评价指标包括:精确率(Precision)、召回率(Recall)、平均精度(AP)和平均精度均值(mAP).精确率(Precision)表示预测为正的所有样本中真正为正样本所占的比例,用于判断模型误检率,计算公式为:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (14)$$

召回率(Recall)表示真正为正的样本中被预测为正的样本所占比例,用于判断模型漏检率,计算公式为:

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (15)$$

平均精度(AP)表示精确率-召回率曲线(Precision Recall curve)下的面积,Total images代表数据集中总图片数目,计算公式为:

对采集到的图像样本使用LabelImg工具进行图像标注,标注格式为Pascal voc,标注图片示例如图5(b)所示.袋体中黑线(Blackline)和黑标(Blackblock)的类别表示袋面有破洞等缺陷,将其标记为废品,属于重点剔除对象;稀经稀纬(Sparse)类别表示织布过程中断经丝、纬丝,导致经纬丝缺失、布面稀疏,形成残次品.其中黑线类别标注2360个,黑标类别标注987个,稀经稀纬类别标注267个.针对建立的数据集图像样本按照7:2:1的比例划分为训练集、验证集和测试集.

$$\text{AP} = \frac{\sum \text{Precision}}{N(\text{Total images})}, \quad (16)$$

均精度均值(mAP)表示所有待检测类别(Class)的AP的平均值,Total classes代表数据集中待检测目标总类别数,计算公式为:

$$\text{mAP} = \frac{\sum \text{AP}}{N(\text{Total images})}, \quad (17)$$

其中,TP为真正类数量、FN为假负类数量、FP为假正类数量、N为样本总数量.

4.3 实验结果及分析

为分析本文改进部分对模型性能的影响,本文首先设计了三组消融实验:(1)双层尺度特征融合;(2)使用轻量级骨干网络GhostNet替换原始CSPDarknet骨干网络;(3)特征融合中结合Coordinate Attention机制.其次,为了符合模型在边缘设备上部署要求,本次实验选取YOLOX-s轻量级规格进行改进和实验,每组实验设置相同训练参数进行测试并与标准YOLOX-s的实验结果对比.最后,为了体现本文所提出的算法的有效性和竞争力,将其与主流目标检测模型实验结果对比.

消融实验1:根据水泥编织袋缺陷特征判断,需

要识别的目标存在特征明显、目标框与原图面积占比较小等特点.此情况下,三层特征融合结构存在大量特征冗余和计算冗余.为此,第一组实验设计使用两层特征融合结构替换原始的三层结构进行训练.

如表 2 所示,使用双层尺度的特征融合结构后,在黑线(blackline)、黑标(blackblock)和布面稀经稀纬(spare)这三个类别的精度对比原始 YOLOX 模型都有了提升,其中黑线和布面稀经缺纬分别提高了 0.69% 和 2.04%.因此说明仅使用两层的特征融合策略也可以达到优秀的性能同时可以减少冗余特征的计算量,且能提高模型拟合效率;其次,推理时间相比 YOLOX-s 在 CPU 设备上快了 25 ms,说明在此次检测任务中,设计双层特征融合策略是可以提升此类特征明显的缺陷检测精度和效率.

表 2 两层尺度特征融合对比实验结果

Tab. 2 Comparison experimental results of two-layer scale feature fusion

Model	AP/%			Fps	Weights /M	Inference-time /ms
	Blackline	Blackblock	Sparse			
YOLOX-s	92.75	98.91	90.40	9.81	36.0	102
Two-Scale	93.44	98.97	92.44	12.94	36.0	77

消融实验 2:为满足实际工业生产线上低成本实时检测,使用对 CPU 计算资源更加友好的轻量级骨干网络 GhostNet 替换原始的 CSPDarknet53 进行图像的特征提取.

表 4 添加 CA 注意力模块对比实验

Tab. 4 Comparison experiment of adding CA attention module

Model	AP/%			Fps	Weights/M	Inference-time/ms
	Blackline	Blackblock	Sparse			
YOLOX-s	92.75	98.91	90.40	9.81	36.0	102
YOLOX-s+CA	93.44	99.14	93.08	15.45	19.0	67

对比实验:为验证改进后模型的有效性和竞争力,分别对 YOLOv5-s、YOLOX-s、CenterNet 等经典算法和基于 YOLO 改进的算法在建立的水泥编织袋数据集上进行对比实验.为了对比的公平性和实验结果的权威性,所有实验均未使用任何预训练权重,且设置相同的 batch-size 和迭代次数,实验结果如表 5 所示.

如表 5 所示,本文提出的 GT-YOLOX 模型相比其他主流目标检测模型精度和速度都具有一定的竞争优势.其中,在水泥编织袋缺陷数据集上黑线平均检测精度为 93.44%,略低于最高精度的 PP-YOLO 算法;在黑标平均检测精度为 99.14%,布面稀经稀

纬的平均检测精度为 93.08%,这两个缺陷类别上精度是最优的.GT-YOLOX 算法在 CPU 端每秒可以检测 15.45 帧图像,平均检测速度可以达到 67 ms,权重参数量最优,大小仅有 19M.该模型不仅能满足边缘设备的部署,还能实现边缘设备的实时高精度检测.

表 3 替换 GhostNet 骨干网络主干对比实验结果

Tab. 3 Replacement of GhostNet backbone network backbone comparison experimental results

Model	AP/%			Fps	Weights /M	Inference-time /ms
	Blackline	Blackblock	Sparse			
YOLOX-s	92.75	98.91	90.40	9.81	36.0	102
GhostNet	91.53	99.00	92.98	15.45	18.0	65

消融实验 3:水泥编织袋生产过程中,由于编织袋印刷图案不唯一,检测背景会发生变化,但缺陷特征高度类似.为保证生产线上检测准确性,使用 CA 注意力机制对特征进行加强,减少背景干扰,降低误检率和漏检率.

添加 CA 注意力机制后与 YOLOX-s 算法对比结果如表 4 所示.可以得出引入 CA 注意力机制后三个类别精度整体总共上提升了 3.6%,说明引入 CA 注意力机制可以让模型对缺陷特征区域更加关注,进一步提升模型检测精度.虽然在模型的推理时间上有略微的增加,但 CA 注意力机制的引入,更能提升模型对复杂背景下的特征检测性能.

5 结语

针对编织袋生产中人力成本不断上升、人工检测效率低、人工劳动强度高、人工漏检率高等问题,本文提出了一种基于 YOLOX 模型的改进方法,称为

表5 主流目标检测模型对比结果
Tab. 5 Comparison results of mainstream target detection models

Model	AP/%			Fps	Weights/M	Inference-time/ms
	Blackline	Blackblock	Sparse			
YOLOv5-s	86.27	98.46	76.23	11.49	28.5	87
YOLOX-s	92.75	98.91	90.40	9.81	36.0	102
Centernet	91.26	99.00	88.78	3.61	131.0	276
YOLOv5-EfficientNet	93.29	98.96	89.86	13.69	45.0	73
YOLOv5-MobileNet	85.93	97.31	87.59	19.23	23.5	52
PP-YOLO	93.67	99.00	92.34	9.17	112.0	109
GT-YOLOX	93.44	99.14	93.08	15.45	19.0	67

GT-YOLOX.该模型将GhostNet与经剪枝优化的双层特征融合模块结合作为特征提取网络,并引入CA注意力机制加强关注特征细节位置信息.该算法在湖北黄石华新水泥袋智能剔除项目实际生产线上进行了有效验证,实验结果表明:GT-YOLOX在目前主流的目标检测模型中效率最高,能够有效满足生产线上编织袋缺陷识别的高精度要求.同时,GT-YOLOX在CPU设备上单个目标识别速度可以降低至67ms,权重参数优化到19M左右,适应于部署到低成本的边缘设备上,极大降低了工业检测设备的成本.

参 考 文 献

- [1] 张昌芳. 浅谈影响编织袋外观质量的因素及解决办法[J]. 化工管理, 2016(24): 113.
- [2] 孙国栋, 林松, 艾成汉, 等. 基于灰度直方图反向投影的织物疵点图像分割[J]. 制造业自动化, 2016, 38(4): 20-21, 35.
- [3] 郝正阳. 基于深度学习的去除运动模糊与布匹瑕疵检测算法研究与实现[D]. 开封: 河南大学, 2022.
- [4] 高举, 李永祥, 徐雪萌. 基于机器视觉编织袋缝合缺陷识别与检测[J]. 包装与食品机械, 2022, 40(3): 51-56.
- [5] 迟欢. 基于机器视觉的编织袋缺陷在线检测方法[J]. 激光与光电子学进展, 2020, 57(20): 320-327.
- [6] 葛万凯, 赵世海, 范雨佳. 基于对比度受限直方图均衡化和非锐化掩模的织物表面瑕疵图像增强算法[J]. 毛纺科技, 2021, 49(12): 68-74.
- [7] 李顺平. 基于深度学习的轻量级表情识别算法[D]. 乌鲁木齐: 新疆师范大学, 2022.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580-587.
- [9] GIRSHICK R. Fast R-CNN [C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015: 1440-1448.
- [10] REN S, HE K, 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.
- [11] DAI J, LI Y, HE K, et al. R-FCN: Object detection via region-based fully convolutional networks [C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. Red Hook: ACM, 2016: 379-387.
- [12] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C]//Computer Vision-ECCV 14th European Conference, Berlin: Springer, 2016: 21-37.
- [13] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 779-788.
- [14] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 6517-6525.
- [15] REDMON J, FARHADI A. YOLOv3: An incremental improvement [EB/OL]. 2018: 1804.02767. <https://arxiv.org/abs/1804.02767v1>
- [16] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [EB/OL]. 2020: 2004.10934. <https://arxiv.org/abs/2004.10934v1>
- [17] GE Z, LIU S, WANG F, et al. YOLOX: Exceeding YOLO series in 2021 [EB/OL]. 2021: 2107.08430. <https://arxiv.org/abs/2107.08430v2>
- [18] 舒军, 王祥, 李灵, 等. 基于改进GANomaly网络的旋开盖缺陷检测方法[J]. 中南民族大学学报(自然科学版), 2023, 42(6): 788-798.
- [19] 王岩, 胡睿甫, 吕传景, 等. 基于掩膜迭代ROI改进LBP算法的齿轮缺陷检测方法研究[J/OL]. 光电子·激光, 2024 (2024-03-21). <https://kns.cnki.net/kcms/>

- detail/12.1182.04.20240320.1541.024.html.
- [20] HE Y, SONG K, MENG Q, et al. An end-to-end steel surface defect detection approach *via* fusing multiple hierarchical features [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69 (4) : 1493-1504.
- [21] ZHANG C, CHANG C C, JAMSHIDI M. Concrete bridge surface damage detection using a single-stage detector[J]. Computer-Aided Civil and Infrastructure Engineering, 2020, 35(4) : 389-409.
- [22] CHEN S H, TSAI C C. SMD LED chips defect detection using a YOLOv3-dense model[J]. Advanced Engineering Informatics, 2021, 47: 101255.
- [23] WANG H, LI M, WAN Z. Rail surface defect detection based on improved Mask R-CNN [J]. Computers and Electrical Engineering, 2022, 102: 108269.
- [24] LU Q, LIN J, LUO L, et al. A supervised approach for automated surface defect detection in ceramic tile quality control [J]. Advanced Engineering Informatics, 2022, 53: 101692.
- [25] 黄露, 夏军勇, 吴庆华, 等. 基于遗传算法与二维最大熵的编织袋缺陷检测[J]. 轻工机械, 2021, 39(5): 69-73, 78.

(责编&校对 雷建云)