

基于改进YOLO11的白坯布缺陷检测研究

史泽阳 张孝龙 崔传金 张元昭

Defect detection in greige fabric based on improved YOLO11

SHI Zeyang, ZHANG Xiaolong, CUI Chuanjin, ZHANG Yuanzhao

引用本文:

史泽阳, 张孝龙, 崔传金, 等. 基于改进YOLO11的白坯布缺陷检测研究[J]. 应用光学, 2026, 47(1): 202–212. DOI: 10.5768/JAO202647.0103004

SHI Zeyang, ZHANG Xiaolong, CUI Chuanjin, et al. Defect detection in greige fabric based on improved YOLO11[J]. Journal of Applied Optics, 2026, 47(1): 202–212. DOI: 10.5768/JAO202647.0103004

在线阅读 View online: <https://doi.org/10.5768/JAO202647.0103004>

您可能感兴趣的其他文章

Articles you may be interested in

基于改进YOLOv5s的曲面光学镜片缺陷检测

Defect detection of curved optical lenses based on improved YOLOv5s

应用光学. 2024, 45(4): 781–789 <https://doi.org/10.5768/JAO202445.0403003>

基于改进YOLO-V4的贴片二极管表面缺陷检测

Surface defect detection of patch diode based on improved YOLO-V4

应用光学. 2023, 44(3): 621–627 <https://doi.org/10.5768/JAO202344.0303007>

基于视觉SLAM和目标检测的语义地图构建

Semantic SLAM based on visual SLAM and object detection

应用光学. 2021, 42(1): 57–64 <https://doi.org/10.5768/JAO202142.0102002>

基于音视频信息融合的目标检测与跟踪算法

Object detection and tracking algorithm based on audio-visual information fusion

应用光学. 2021, 42(5): 867–876 <https://doi.org/10.5768/JAO202142.0502007>

基于知识蒸馏的夜间低照度图像增强及目标检测

Nighttime low-light image enhancement and object detection based on knowledge distillation

应用光学. 2023, 44(5): 1037–1044 <https://doi.org/10.5768/JAO202344.0502004>

基于IDOU-YOLO的红外图像无人机目标检测算法

Infrared image UAV target detection algorithm based on IDOU-YOLO

应用光学. 2024, 45(4): 723–731 <https://doi.org/10.5768/JAO202445.0402001>



关注微信公众号, 获得更多资讯信息

文章编号: 1002-2082 (2026) 01-0202-11

引用格式: 史泽阳, 张孝龙, 崔传金, 等. 基于改进 YOLO11 的白坯布缺陷检测研究 [J]. 应用光学, 2026, 47(1): 202-212.

SHI Zeyang, ZHANG Xiaolong, CUI Chuanjin, et al. Defect detection in greige fabric based on improved YOLO11[J]. Journal of Applied Optics, 2026, 47(1): 202-212.



在线阅读

基于改进 YOLO11 的白坯布缺陷检测研究

史泽阳, 张孝龙, 崔传金, 张元昭

(华北理工大学 电气工程学院, 河北 唐山 063210)

摘 要: 针对白坯布瑕疵缺陷识别难度大、精度低且漏检率高的问题, 对 YOLO11 模型进行了系统性的优化与改进, 以提升目标检测模型在复杂场景中的检测精度与速度。首先, 在主干网络中引入 C3K2-DCNv2 模块, 利用可变形卷积增强模型对目标形变与复杂背景的适应能力。其次, 在颈部网络部分, 融合自适应尺度融合 (adaptive scale fusion, ASF) 模块与 DySample 模块形成 ASF-DySample 模块, 实现了多尺度特征的动态采样与空间自适应融合, 保留细节信息, 进一步提升小目标检测精度, 有效提升了特征表达能力。最后, 检测头采用 MultiSEAMHead, 通过多分支语义解耦和上下文聚合策略, 进一步强化目标表征能力。实验结果表明, 改进后的 DAM-YOLO 模型的 mAP@0.5 较原始 YOLO11 模型提升了 2.3%, 验证了该模型在目标检测任务中的有效性与实用价值。

关键词: 目标检测; YOLO11; C3K2_DCNv2; ASF-Dysmple; MultiSEAMHead

中图分类号: TN206

文献标志码: A

DOI: 10.5768/JAO202647.0103004

Defect detection in greige fabric based on improved YOLO11

SHI Zeyang, ZHANG Xiaolong, CUI Chuanjin, ZHANG Yuanzhao

(School of Electrical Engineering, North China University of Science and Technology, Tangshan 063210, China)

Abstract: To address the challenges of low detection accuracy and high miss rate in defect recognition of greige fabric, this study systematically optimized and improved the YOLO11 model to enhance detection precision and speed in complex textile inspection scenarios. Firstly, a C3K2-DCNv2 module was introduced into the backbone network, leveraging deformable convolutions to improve the model's adaptability to target deformation and complex backgrounds. Secondly, in the neck network, the DySample and ASF modules were integrated to construct ASF-DySample module, achieved dynamic sampling and spatially adaptive fusion of multi-scale features, effectively preserving detail information, significantly improving small object detection accuracy and feature representation capability. Finally, the MultiSEAMHead module was adopted in the detection head, which further strengthened the target representation capability through multi-branch semantic decoupling and contextual aggregation strategies. Experimental results demonstrated that the improved DAM-YOLO model achieved a 2.3% increase in mAP@0.5 compared to the original YOLO11, validating the effectiveness and practical value of the proposed method in object detection tasks.

Key words: object detection; YOLO11; C3K2_DCNv2; ASF-Dysmple; MultiSEAMHead

收稿日期: 2025-07-25; 修回日期: 2025-08-20

基金项目: 河北省骨干人才项目 (B2024005010); 河北省省属高等学校基本科研业务费研究项目 (JQN2023026)

第一作者: 史泽阳, E-mail: szyeason@126.com

通信作者: 张孝龙, E-mail: zhangxiaolong@ncst.edu.cn

引言

随着智能制造和自动化技术的发展,白坯布作为纺织品生产中的重要原材料,其质量直接影响到后续产品的生产效率和市场竞争力。然而,目前大部分国内外企业仍依赖人工检测,这种方式存在检测效率低、漏检和误检率高等问题,同时人力成本也较高。少数企业已开始使用传统的机器视觉技术进行自动化缺陷检测,这种方法虽然提升了检测效率,但仍有一定的局限性。传统的机器视觉检测法^[1]对于瑕疵特征的提取太过于依赖人工设计的特征,对拍摄视角和环境光线较为敏感,面对白色胚布存在无法检测多种缺陷类别、错检率高、漏检率高等问题,无法达到布匹厂家实际生产中对于缺陷瑕疵检测的需求。因此,设计一种准确高效、实用性强的白坯布检测方法有重要意义。

随着时代的发展和工业自动化的普及,深度学习模型^[2]已经被广泛应用到轻工业制造品的缺陷检测中,并且产生了良好的效果。目标检测算法主要包括双阶段算法^[3]和单阶段算法^[4]。双阶段目标检测算法先根据图像提取候选框,再基于候选区域做二次修正得到检测点结果,检测的准确度虽然较高,但检测效率比较低,此类算法的开山之作是 RCNN(regions with convolutional neural network)^[5]系列;而以 YOLO(you only look once)^[6]系列为代表的单阶段算法,直接从图像中提取特征,并同时回归出物体的位置和分类结果,大大提升了检测效率。

针对白坯布缺陷瑕疵检测速度和精度问题,众多研究者在多种深度学习模型上进行了多项改进。苏立鹏^[7]等在 Fsater-RCNN 网络中引入了可变性卷积,以增强捕捉传统卷积难以检测的复杂形状特征的能力,并采用了广义交并比(generalized intersection over union, GIoU)损失函数代替传统的交并比(intersection over union, IOU)损失函数,显著提高了边界框预测的准确性。但两阶段目标检测算法存在计算量大、检测速度慢的问题。徐春鸽^[8]通过融入卷积块注意力机制(convolutional block attention module, CBAM)和双向特征金字塔网络(bidirectional feature pyramid network, Bi FPN),以提升对布匹缺陷的聚焦效果,有效抑制了无效特征和噪声,提高了检测精度并取得了不错的效果。王千^[9]等提出了一种改进的 YOLOv8 模型,引

入组洗牌卷积(grouped shuffle convolution, GSConv)网络、双向特征金字塔和 TransViBlock 模块,有效提升了布匹残损检测的精度与实时性。郑雨婷^[10]引入均衡损失函数(equalization loss, EQL)和均衡损失函数 v2(equalization lossv2, EQLv2),优化了 YOLOv5 的分类损失和梯度权重,从而提高了召回率,并结合高效特征金字塔网络(efficient feature pyramid network, EFPN)和噪声过滤残差收缩模块(noise reduction-residual shrinkage building unit, NR-RSBU),提升对微小瑕疵的检测能力。毛龙妹^[11]等结合深度卷积神经网络和知识图谱,提出了一种新的针织物疵点检测方法。通过在深度学习模型中融合知识图谱,不仅提升了模型的语义理解能力,还增强了对不同类型针织物疵点的识别精度,但样本数量过少且改进后的模型参数量较大,在算力受限的设备上部署难度较大。

针对以上问题,为提高白坯布缺陷检测的精度,本文提出一种基于改进 YOLO11 的白坯布缺陷检测模型 DAM-YOLO。在主干网络中融入 C3K2-DCNv2 可变形卷积网络,通过调节感知输入特征的偏移量和不同的空间位置、采样点的输入特征幅值,来调整卷积操作中的采样点,从而使卷积核能够更加高效地感受图像中的形变特征。采用 ASF-DySample 上采样模块,降低背景和光照反射对瑕疵点检测的干扰。同时将检测头替换为 Multi-SEAMHead^[12]模块,提高了特征图的分辨率,增强对小目标的识别能力。

1 YOLO11 模型原理

1.1 YOLO11 模型简介

Ultralytics 团队于 2024 年 9 月发布了 YOLO11,相较此团队上一代作品 YOLOv8,其主要改变为:提出 C3K2 机制,替换掉原有 C2f,并在 SPPF 后加了一层类似于注意力机制的 C2PSA;在原本的解耦检测头内部增加了两个深度卷积模块(depthwise convolution, DWConv),DWConv 是一种高效的卷积方式,通过单独处理每个通道来减少计算量,结合 1×1 的点卷积(pointwise),形成深度可分离卷积(depthwise separable, DSConv),可以在保持网络性能的同时极大地减少模型的计算复杂程度和参数量;同时对 YOLO11 模型的深度和宽度参数进行了大幅度调整。YOLO11 网络结构主要包括输入端(In-put)、骨干网络(Backbone)、颈部(Neck)和检

测头(Head)4部分,本研究使用参数量和精准度较为平衡的YOLO11作为基线模型,其结构如图1所示。

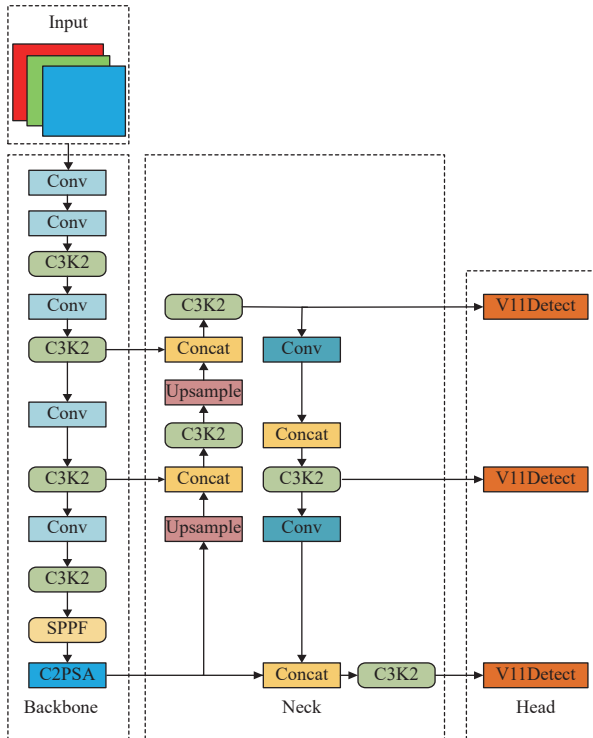


图1 YOLO11网络结构图

Fig. 1 Network structure of YOLO11

1.2 DAM-YOLO 设计思想

1.2.1 C3K2-DCNv2 模块

传统的卷积神经网络(convolutional neural network, CNN)使用固定的、规则的卷积核(如 3×3),在图像上按照固定的格点采样特征。这种固定的采样方式限制了模型对几何变形和目标形状变化的适应能力,尤其在处理复杂场景,如物体变形、姿态变化和视角变换时表现较差。可变形卷积(deformable convolution network, DCN)是对传统卷积的一种扩展,其核心思想是在固定采样点基础上,通过学习偏移量,使卷积核的采样位置能够根据输入特征进行自适应调整,从而动态适应形变。然而,变形卷积得到的特征可能会受到无关图像内容的干扰。第二代可变形卷积网络(deformable convolutional networks v2, DCNv2)通过提升几何建模能力和训练强度来改进DCN,以便更精确地提取目标对象的特征信息,DCN模块结构如图2所示。

在C3K2-DCNv2模块中,由输入通道通过一个 1×1 卷积调整至中间通道数 c ,随后依次堆叠 n 个带DCNv2的子模块,每个子模块内部包含 1×1 降维卷积、 3×3 可变形卷积DCNv2和 1×1 升

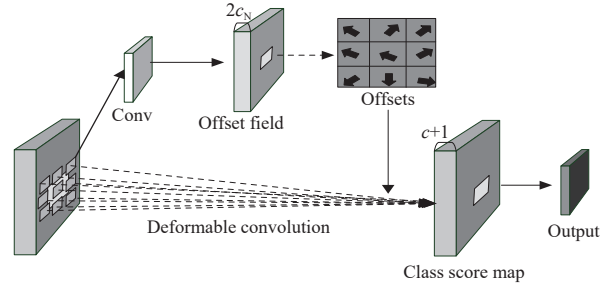


图2 DCN模块结构

Fig. 2 Structure of DCN

维卷积,模块间通过残差连接实现信息传递。最后,输出经过一个 1×1 卷积调整通道数至 c_2 ,完成特征变换。其原理图如图3所示。

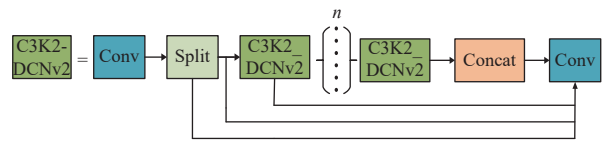


图3 C3K2-DCNv2工作原理图

Fig. 3 Schematic diagram of C3K2-DCNv2 working principle

DCNv2 的卷积公式为

$$y(a) = \sum_{k=1}^k W_k \cdot X(a + d_k + \Delta d_k) \cdot \Delta m_k \quad (1)$$

式中: X 为输入特征图上 a 像素点的特征; $y(a)$ 为输出特征图上像素点 a 的特征; W_k 为学习权重; d_k 为预先指定偏移量; Δd_k 为像素点 k 的可学习偏移值; Δm_k 为像素点 k 的调制权重,取值范围为 $[0,1]$ 。

1.2.2 ASF-Dysample 模块

在Neck网络中,融合了自适应尺度融合(adaptive scale fusion, ASF)模块^[13]与Dysample模块形成新的Neck网络。其中,ASF模块通过结合空间和尺度特征实现精确且快速的实例分割。该模块在YOLO分割框架的基础上进行扩展。首先,引入了尺度序列特征融合(scale sequence feature fusion, SSFF)模块,通过SSFF模块,网络能够有效地整合来自不同尺度的特征信息,提升对不同尺寸目标的识别和处理能力。为了进一步优化特征提取过程,模块还引入了三重特征编码器(triple feature encoder, TFE)模块。该模块通过融合来自大、中、小尺度的特征图,能够综合不同尺度的细节信息,进一步增强对目标的细节提取能力,尤其是在应对小尺寸目标检测和复杂背景时表现更为突出。通过这两种模块的结合,网络不仅能够更好地提取多尺度信息,还能在细节捕捉和目标识

别方面提升准确性,显著优化了模型在目标检测和分割任务中的表现。

SSFF 模块通过融合浅层细节与深层语义特征,解决了传统降采样中图像细节丢失的问题。有效保留了多尺度信息,提升了网络对细节的敏感度,在小目标或复杂背景下表现尤其出色。此外,SSFF 模块能保证尺度不变的特征稳定提取,增强了模型的鲁棒性与精度,在小目标检测和图像分割中取得了优异效果。设二维输入图像为 $f(w, h)$, 其宽度为 w 、高度为 h , c 为特征图的通道维度索引,输入特征图与二维高斯滤波器 $G_\sigma(w, h)$ 进行卷积运算后得到滤波结果 $F_\sigma(w, h)$ 。最后输入到 SSFF 的缩放图像可通过对原始输入图像进行尺度变换获得,其计算过程为

$$\begin{cases} F_\sigma(w, h) = G_\sigma(w, h) f(w, h) \\ G_\sigma(w, h) = \frac{1}{2\pi\sigma^2} e^{-\frac{w^2+h^2}{2\sigma^2}} \end{cases} \quad (2)$$

式中: $G_\sigma(w, h)$ 为一种常用的图像平滑工具,主要用于降低图像中的噪声和细节,以便提取更稳定的特征,其形状由标准偏差 σ 调节, σ 控制滤波器的扩展范围和模糊程度,卷积操作通过与图像像素值进行加权求和,减少图像的模糊化,降低噪声并使图像更加平滑,较小的 σ 会生成较窄的高斯函数,细节保留较多,而较大的 σ 会产生更广泛的高斯函数,使得图像更加平滑。

P3、P4 和 P5 分别表示特征金字塔网络中不同层级、不同下采样尺度的特征图。其中, P3 具有较高的空间分辨率和丰富的细节信息,更适合小尺度目标的特征表达。因此在 SSFF 模块中,首先,通过 1×1 卷积将 P4 和 P5 特征图的通道数统一为 256,以确保后续处理的一致性与高效性。其次,使用最近邻插值将 P4 和 P5 的尺寸调整至 P3 特征图

的级别,从而实现不同尺度特征图的统一尺度对齐。随后,通过 `unsqueeze` 操作将三维张量扩展为四维,增加深度维度,为后续特征融合提供结构支持。转换后的特征图在深度维度上拼接,形成融合多尺度信息的新特征图,作为三维卷积和后续处理的输入。最后,通过三维卷积操作对拼接后的特征图进行进一步处理,同时结合三维批标准化 (batch normalization, BN) 和 SiLU 激活函数 (sigmoid linear unit, SiLU) 提取尺度序列特征并增强模型的表达能力。通过这些操作, SSFF 模块不仅有效融合了不同尺度的特征,还提升了网络对细节和多尺度信息的敏感度。

对于瑕疵物检测中存在密集小目标的问题,传统特征融合方法往往仅对小尺寸特征图进行上采样或与前层特征直接合并,未能充分利用大尺寸特征层中的细节信息。为解决这一问题,提出了 TFE 模块。该模块通过有效分割和处理不同尺度的特征图,放大其细节特征,提升了细节质量与特征表达能力。TFE 模块不仅保留了各尺度特征的细节,还优化了特征融合,尤其在多尺度信息融合过程中,显著增强了对小目标和复杂细节的捕捉能力,提高了网络在不同尺度图像下的准确性。

在 TFE 模块中原本使用的近邻上采样法锯齿效应明显,细节丢失严重,仅适用于整数倍放大。因此使用 `Dysample` 动态上采样替换原有上采样,使低分辨率图像的特征丰富性显著提高;中等分辨率在其中取得平衡,模型的综合分析能力得到了增强。针对大、中、小 3 个不同尺度的特征图,首先分别进行卷积操作以提取关键信息,随后在通道维度上将其拼接,从而实现了多尺度特征的有效整合与融合。改进后的 TFE 模块的结构如图 4 所示。

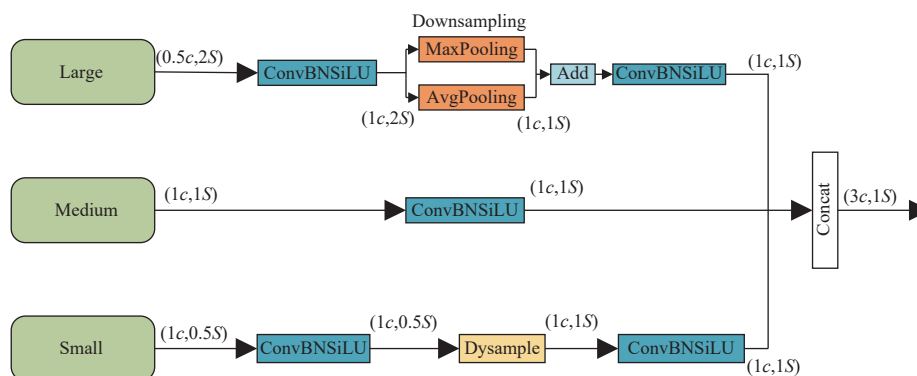


图 4 引入动态上采样的 TFE 模块

Fig. 4 TFE module with dynamic upsampling

图 4 中 c 为通道数, S 为特征图大小, 每个 TFE 模块将 3 种不同尺寸的特征图作为输入。在处理大尺度特征图 (large) 时, 首先, 利用卷积操作将输入特征图的通道数调整为 1 通道, 以便为后续处理统一维度。随后, 分别采用最大池化与平均池化策略进行下采样, 进一步压缩空间信息, 同时保留高分辨率特征的关键内容。在对小尺寸特征图 (small) 的处理过程中, 同样通过卷积模块对其通道数进行统一调整, 以保证整体特征融合的一致性与有效性。最终, 通过拼接大、中、小尺寸的特征图 (F_l , F_m 和 F_s), 得到 TFE 模块的输出特征图 $F_{TFE} = \text{Concat}(F_l, F_m, F_s)$, 输出特征图 F_{TFE} 的分辨率与中尺度特征图 F_m 相同, 但通道数是 F_m 的三倍。通过这种方式, TFE 模块成功实现了不同尺度特征的融合, 增强了多尺度信息的表达能力。

Dysample 模块^[14]是一款超轻量级且高度有效的动态上采样模块, 提出了一种绕过传统动态卷积的新思路, 利用 PyTorch 的标准内置功能来执行点采样, 能够根据特征图动态调整采样方式, 是一种既有轻量特性又保持高性能的解决方案。DySample 模块的结构如图 5 所示, 动态采样点生成器的原理如图 6 所示。

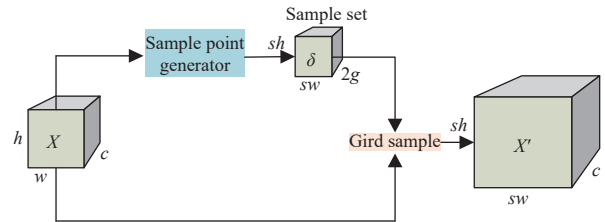


图 5 Dysample 模块结构图

Fig. 5 Dysample module structure diagram

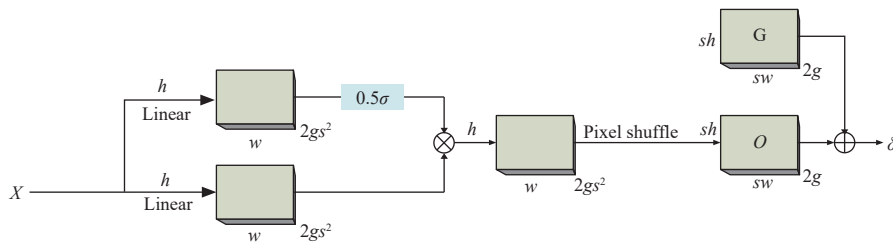


图 6 动态采样点生成器

Fig. 6 Dynamic sampling point generator

在图 5 中, X 为输入特征, X' 为输出特征, g 为每个输出像素对应的可学习的采样点。工作时首先通过动态采样点生成器 (sample point generator) 预测每个采样点的位置, 使采样点能够适应特征图的空间变化。接着, 采样函数 δ 利用这些生成的采样点对特征图通过网络样本函数 (grid sampling) 进行重新采样, 确保网络能够更好地捕捉和利用图像中的细节信息。最终, 通过这一系列操作, 上采样后的特征图能够提供更高的分辨率和更多的细节, 为后续的处理步骤提供更为丰富的信息。这种基于动态采样的上采样方法在处理多尺度图像时具有很大的优势, 能够提升模型对小目标或细节的识别能力。

图 6 为动态采样点生成器原理图, 其核心组件包括采样网络 (G) 和生成的偏移量 (O), s 为上采样比例因子, 采样点生成器通过 Sigmoid 函数 σ 对输入特征进行范围因子的生成。同一个采样点生成器生成的采样点具有相同的初始位置, 为了增加图像尺寸的同时保持其细节, DySample 模块利

用线性投影层为特征图中的每一个位置生成对应的初始偏移向量。该偏移量描述了从原始采样点到目标位置之间的空间位移关系, 反映了采样格点的空间变换信息。然后, 采用像素混洗技术^[15] (pixel shuffle) 将这些初始偏移量重组为更高分辨率的偏移量图。通过这种方式, 偏移量图能够更精确地对应输入特征图中的每个位置, 确保精确的空间对齐。为解决采样点特征细节信息重构问题, DySample 模块对输入特征图进行重新采样, 该过程通过精确的空间变换结合高分辨率偏移量生成新的特征图, 从而在采样过程中有效保留特征细节并提升图像分辨率, 进一步增强对小尺度目标区域特征恢复的准确性。

因此, 本文将 ASF 模块与 Dysample 模块融合至 Neck 网络中, 在保证较高实用性的同时, 有效提升了对白坯布瑕疵的准确检测能力。

1.2.3 MultiSEAMHead 检测头

MultiSEAMHead 检测头由回归分支和分类分支两部分组成。其中回归分支引入 MultiSEAM 模

块强化空间特征建模能力, 从而提升目标定位精度。检测头通过多尺度卷积和自适应池化操作有效增强了特征表示能力, 并结合权重分配机制自适应调整不同特征的重要性, 引导模型更关注关键信息。通过分布焦点损失^[16] (distribution focal loss, DFL) 模块将离散分布转换为连续边界框坐标。DFL 的核心思想是将边界框的回归问题转化为一个分布学习问题。具体来说, DFL 将边界框的坐标 (如中心点、宽度、高度) 视为一个概率分布, 并通过最小化预测分布与真实分布之间的差异来优化模型。DFL 公式如式(3)所示:

$$\text{DFL}(p, q) = - \sum_i q_i \ln(p_i) \quad (3)$$

式中: p_i 是模型预测的分布; q_i 是真实分布。分类分支首先对输入特征应用一个深度可分离卷积 (depthwise separable convolution, DWConv) 以实现高效的特征提取, 接着通过一个卷积提升特征维度, 再利用 MultiSEAM 模块增强分类语义信息, 最终通过一个卷积映射为类别数维输出。MultiSEAM-Head 结构如图 7 所示。

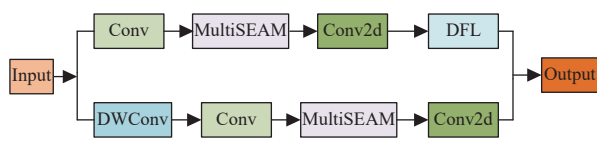


图 7 MultiSEAMHead 整体架构

Fig. 7 Overall architecture of MultiSEAMHead

MultiSEAM 是一种多头注意力机制, 其结构如图 8 所示, 包括 3 个不同尺寸 (patch-3、patch-5、patch-7) 的通道与空间混合 (channel and spatial mixing module, CSMM) 模块。首先通过平均池化 (average pooling) 操作对这些模块的输出进行处理, 然后进行通道扩展 (channel expansion) 操作, 最后相乘以提供增强的特征表示。图 8 右侧为 CSMM 模块的结构布局。该模块通过采用不同尺寸的 patch 来有效提取多尺度特征, 从而确保网络能够处理来自多个尺度的信息。此外, 模块还利用深度可分离卷积来学习空间维度与通道之间的关系。深度可分离卷积能够显著降低计算复杂度, 并帮助网络捕捉到更加细致的空间特征和通道依赖关系。通过这种方式, CSMM 模块不仅提升了特征提取的多样性, 还增强了模型对不同尺度特征的处理能力, 从而提高了整体的特征表示质量。CSMM 模块中, Patch embedding 将输入的图像块 (patch) 进行嵌入的操作, 高斯误差线性单元^[17]

(gaussian error linear unit, GELU) 是一种神经网络激活函数, 旨在更好地模拟人类神经元的非线性特性, 提升模型的表达能力, BatchNorm 进行批量归一化处理, 用于加速训练过程, 提高性能并减少过拟合的发生, Depthwise convolution 为深度可分离卷积^[18], 对每个输入通道分别进行卷积操作, Pointwise convolution 为逐点卷积^[19], 使用尺寸为 1×1 的卷积核来融合深度可分离卷积的特征, 对输入特征图进行处理。该模块设计旨在通过对空间维度和通道的细致处理, 增强网络对小目标的注意力和捕捉能力。通过综合利用多尺度特征和深度可分离卷积, CSMM 在保持计算效率的同时, 提高了特征提取的精确度。这对于小目标检测尤其重要。

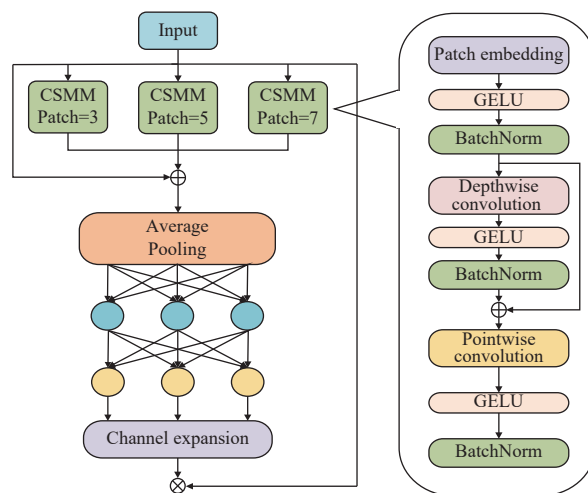


图 8 MultiSEAM 整体架构

Fig. 8 Overall architecture of MultiSEAM

通过将 MultiSEAMhead 融入到 YOLO11 的检测头中, 提高了模型对复杂场景中多种白坯布瑕疵特征的识别能力。

1.3 DAM-YOLO 模型

基于上述分析设计, 本文以 YOLO11 作为基础模型提出了改进模型 DAM-YOLO。首先, 在 YOLO11 的 Backbone 网络中, 将 C3K2 替换为了 C3K2_DCNv2, DCNv2 在可变形卷积中引入了“调制机制” (modulation), 通过额外的权重学习每个采样点的偏移量和缩放因子。通过学习偏移量, 卷积核的感受野不再局限于固定矩形, 而是能根据输入特征的语义信息动态调整, 从而更准确地提取目标对象的特征信息。其次, 在 Neck 网络中添加了 ASF 模块, 将 Zoom_cat 中 TFE 模块的近邻上采样法改为动态上采样法, 并在 Neck 网络中添加了上采样模

块 Dysmple, 有效提升了对白坯布缺陷的识别能力。最后, 将检测头替换为 Detect_MultiSEAM, 其利用多尺度特征^[20]和深度可分离卷积来学习空间

维度和通道之间的相关性。DAM-YOLO 模型能够更精准地检测白坯布瑕疵, 检测精度和适用性都有显著提升, DAM-YOLO 结构如图 9 所示。

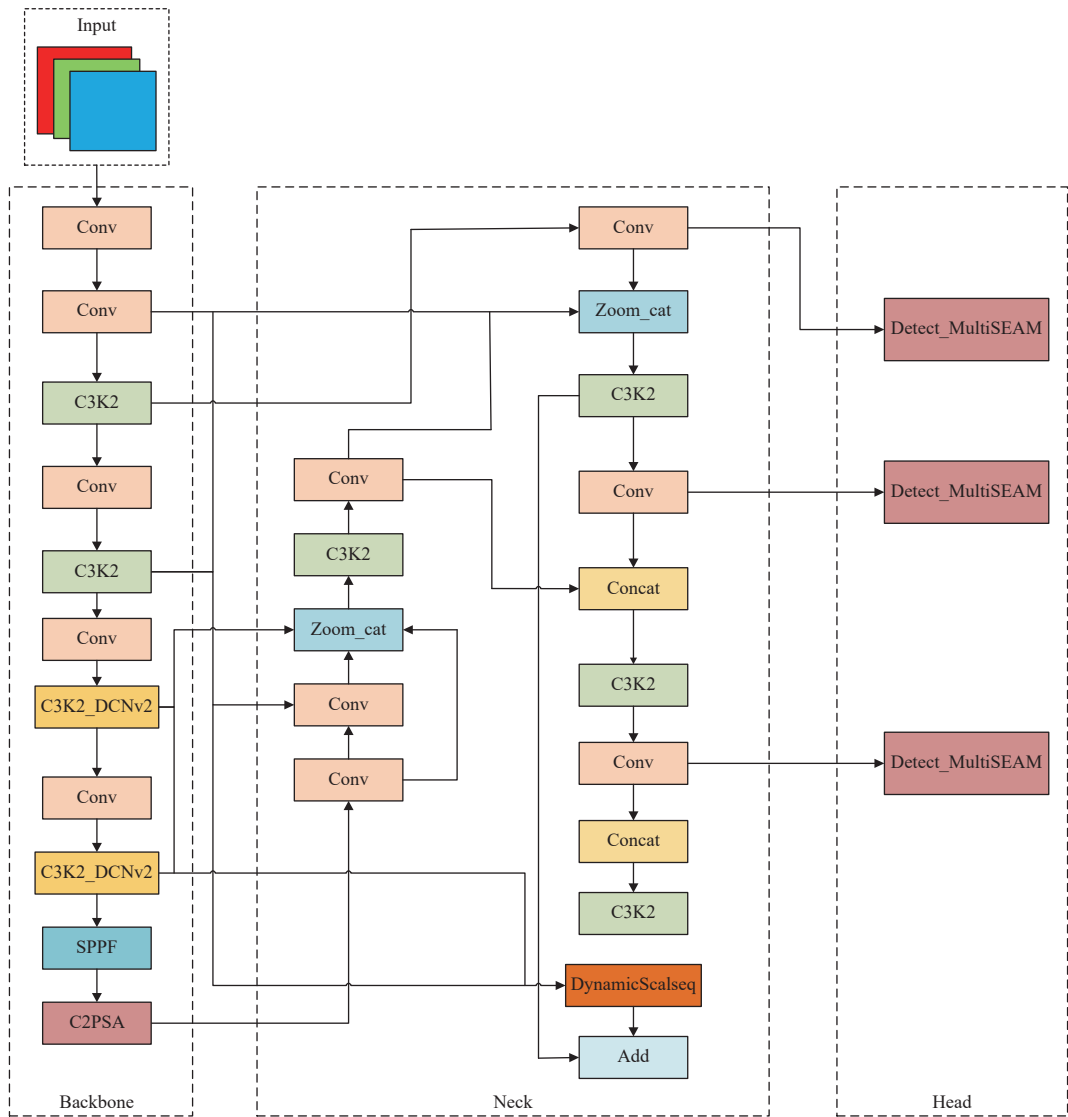


图 9 DAM-YOLO 模型结构图

Fig. 9 Structural diagram of DAM-YOLO model

2 实验结果与分析

2.1 白坯布缺陷数据来源及收集

因白坯布公共数据集较少, 所以本实验采用自建数据集, 制作数据集的白坯布材质为桑蚕丝, 另外为了丰富数据集的多样性, 本实验采用了一部分 TILDA 数据集, 其为从织物生产过程中获得的真实样本。自建数据集共有 2 107 张图片, 均使用 jpg 格式存储, 包括从 TILDA 中选取的 800 张图片和 1 307 张自制图片。将整个数据集按 8 : 1 : 1 的比例划分为训练集、测试集和验证集。数据集包括 4 种瑕疵, 通过对数据集样本的分析处理, 将数

据集分为破洞(hole)、划痕(scratch)、异物(objects)、污渍(stain)4 类, 如图 10 所示。

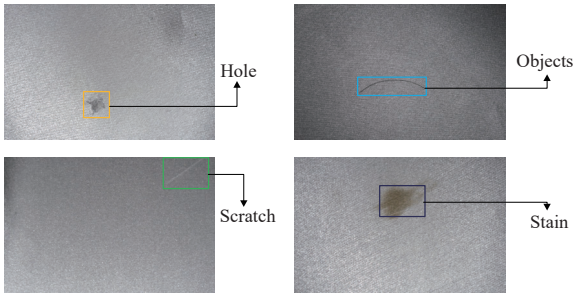


图 10 白坯布数据集样例

Fig. 10 Sample of greige fabric dataset

2.2 实验环境与参数设置

实验采用 Pytorch1.10.0 深度学习框架进行 YOLO11 模型的构建、训练和测试。实验采用 NVIDIA GeForce RTX 4090 显卡, batch 大小设置为 16, 训练轮数设置为 300 轮, 基于 CUDA11.3 (compute unified device architecture) 进行图形处理器 (graphics processing unit, GPU) 加速, 以提升计算机图形处理能力。

2.3 评价指标

本文实验使用精确率 P 、召回率 R 和平均精度均值 (mean average precision, mAP) P_{mAP} , 作为模型对白坯布缺陷检测的主要评价指标。mAP 为本文评估精确率的核心指标。 P 、 R 、 P_{mAP} 的计算公式分别如式 (4)~式 (6) 所示:

$$P = \frac{D_T}{D_T + D_F} \quad (4)$$

$$R = \frac{D_T}{D_T + D_N} \quad (5)$$

$$P_{\text{mAP}} = \frac{1}{n} \sum_{i=1}^n P_{\text{AP},i} = \frac{1}{n} \sum_{i=1}^n \int P dR \quad (6)$$

式中: D_T 为被正确识别到的缺陷数量; D_F 为错误检测到的缺陷数量; D_N 为未能检测到缺陷的数量; P_{AP} 为每个类别的平均精度; n 为目标检测任务中涉及的类别总数。此外, 选用参数量 (Params/ 10^6) 作为模型效率的评价指标之一, 参数量越小, 模型的计算量越小。

2.4 消融实验

为了验证本文改进模型对白坯布缺陷检测的影响, 以 YOLO11 原模型为基础进行了消融实验, 分析了各项单独以及组合融合的效果。消融实验结果如表 1 所示, “√”表示添加该改进。

表 1 消融实验结果

Table 1 Results of ablation experiment

C3k2_DCNv2	ASF-Dysample	MultiSEAM	$P/\%$	$R/\%$	mAP@0.5/%	FPS
			82.5	74.3	80.2	186.5
√			83.0	74.7	80.6	183.2
	√		84.2	75.0	81.2	165.6
		√	82.9	74.9	80.8	175.8
√	√		84.5	76.0	81.6	173.5
√		√	83.3	75.1	80.9	178.3
	√	√	85.0	76.5	82.1	170.2
√	√	√	86.8	76.9	82.5	172.3

由表 1 可知, 原始模型的 mAP@0.5 为 80.2%, 加入 C3K2_DCNv2 模块后检测精度和召回率都有所提升, 验证了可变形卷积在提高缺陷检测性能方面的有效性, 但 mAP@0.5 仅提升至 80.6%, 提升效果不是很明显。单独引入 ASF-Dysample 后模型检测精度得到了不错的提升, mAP@0.5 提升至 81.2%, 比单独引入 C3K2_DCNv2 更高, 证明 Dysample 动态上采样替换原有的上采样取得了不错的效果, 该模块有效增强了低分辨率特征的表达能力, 同时在中分辨率特征上实现了较好的平衡, 从而整体提升了模型的综合分析能力。进一步验证了 Dysample 模块在更好地保留和恢复瑕疵细节特征方面所发挥的重要作用。采用 MultiSEAM 替换原有的检测头后通过使用多尺度卷积和自适应池化有效提升了特征表示能力。进一步地, 当联合使

用 C3k2_DCNv2 与 ASF-Dysample 时, 召回率增加了 1.7%, mAP@0.5 增加了 1.4%, 表明 ASF-Dysample 能更有效地提升模型性能, 其中 SSFF 模块提取尺度序列特征并增强了模型的表达能力, TFE 模块融合 DySample 后, 通过精细的偏移控制, 对于不同尺度的特征图, 有效保留了小目标的细节信息, 显著提高了小目标瑕疵的检测效果。而 ASF-Dysample 与 MultiSEAM 联合使用时, 召回率提升了 2.2%, mAP@0.5 提升了 1.9%, 结合后的模块通过空间和通道维度的交互, 能够对多尺度语言信息进行有效融合, 既突出了目标的细节特征, 又提高了目标检测的精度。最终, 将各个模块整合后, 模型的精确率提升了 4.3%, 召回率提升了 2.7%, 平均精度提升了 2.3%, 取得了显著的性能提升。该结果充分验证了所提出的 3 个模块在结构设计上的

有效性与协同优势,最终全模块融合方案达到了最优的检测精度和整体性能,满足对白坯布缺陷检测的准确性要求。

2.5 不同模型实验结果对比分析

为证明所提出的改进模型 DAM-YOLO 的优越

性,将该模型与 SSD(single shot multibox detector)、Faster-RCNN、RT-DETR-1、YOLOv5s、YOLOv8n、YOLOv10n、YOLO11 和 YOLOv12n 进行了对比。在对比实验中均采用相同的实验配置、训练集、测试集和验证集,实验结果如表 2 所示。

表 2 不同模型实验结果对比

Table 2 Comparison of experimental results among different models

Model	P/%	R/%	mAP@0.5/%	Params/ 10^6	FPS
SSD	72.5	62.3	72.3	27.2	46.8
Faster-RCNN	72.9	63.5	73.1	136.3	27.2
RT-DETR-1	70.3	63.2	74.8	28.6	88.6
YOLOv5s	84.4	69.6	78.2	9.1	148.3
YOLOv8n	80.9	72.2	78.4	3.1	174.3
YOLOv10n	81.1	66.4	75.1	7.4	175.6
YOLO11	82.7	74.3	80.2	2.6	186.5
YOLOv12n	82.5	73.6	79.8	2.58	180.5
DAM-YOLO	86.8	76.9	82.5	5.3	172.3

由表 2 可知,SSD 模型的精确率、召回率和 mAP@0.5 相对于 YOLO 系列都较为一般,并且参数量相对较大,不适合计算资源较低的场景。相比 SSD, Faster-RCNN 的各项性能有所提升,但参数量达到了 136.3M,计算资源需求较高,模型较为复杂,不适合实时应用。YOLO 系列模型在精度和速度之间达到了较好的平衡, YOLOv5s 的精确率和召回率较高,分别为 84.4% 和 69.6%, mAP@0.5 为 78.2%, 尽管性能较好,但相比于 DAM-YOLO 等更先进的模型,精度稍有不足。YOLOv8n 的 mAP@0.5 相比 YOLOv5s 略有提升,但参数量远小于 YOLOv5s。YOLOv10n 各项指标相较于 DAM-YOLO 差距较大。YOLOv12n 各方面表现良好,但精确率和召回率以及 mAP@0.5 都略低于 YOLO11。YOLO11 是 YOLO 系列中表现最好的一个,在精度和召回率上都有明显的提升,并且参数量也较小。DAM-YOLO 模型表现出色,精确率为 86.8%,召回率为 76.9%, mAP@0.5 为 82.5%,参数量虽然略高于 YOLO11,但对检测速度并未产生较大影响,并且改进后的模型相较于原始模型的精确率、召回率、平均精度均有提升。对比其他模型,综合主要性能指标进行分析,改进模型具有最好的综合性能。

本文所提出的改进模型在精确率、平均精度和计算量上均达到了最佳效果。改进模型虽比原

始模型参数量稍高,但仍适合部署在对精度要求高且计算资源有限的环境中。

2.6 改进模型在自建数据集上的实验效果

为验证改进模型的可行性,在自建数据集上进行可视化比较,图 11 所示为 Faster R-CNN、YOLO11、YOLOv12n 与改进模型在 scrach(划痕)、objects(异物)、stain(污渍)和 hole(破洞)4 类缺陷上的检测效果。图 11(a)为“划痕”缺陷的检测效果,“划痕”检测易受到胚布材质和环境灯光的影响,可以看到,“划痕”检测的置信度由原始的 0.78 提升至 0.85,检测效果有一定提升。图 11(b)中“异物”缺陷检测效果也有明显提升,改进后的模型更加精确地定位了目标。图 11(c)为“污渍”缺陷的检测效果,“污渍”易受亮面布反光的影响,改进模型检测效果亦有小幅提升。图 11(d)为“洞”缺陷检测效果,虽然“洞”这种小目标缺陷的特征和细节难以捕捉,但改进后的模型检测效果提升较为显著。对 4 类缺陷检测结果进行对比,可以看到, DAM-YOLO 具有良好的检测效果,检测精度高,并且能够通过合理调整边界框的距离来更好地关注不同尺寸目标,尤其是小目标缺陷的关键特征,有效减少了漏检情况,相比其他模型,对 4 类缺陷的检测置信度均有提升,表现最优秀。充分验证了改进后模型的可行性。

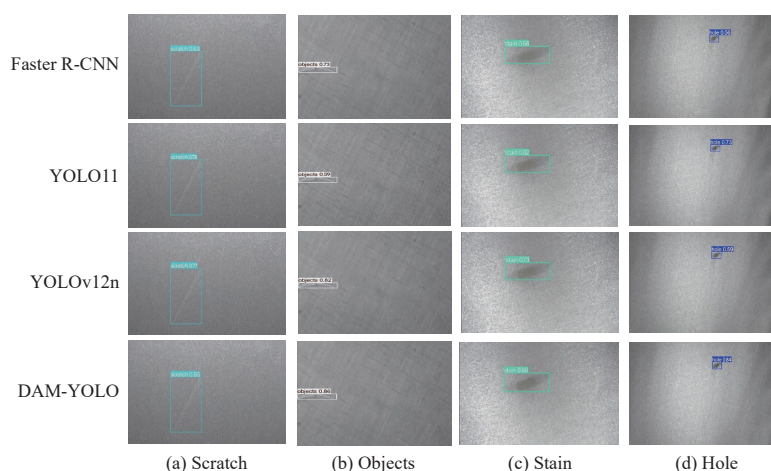


图 11 检测结果可视化

Fig. 11 Visualization of detection results

3 结论

针对白坯布瑕疵检测中小目标检测精度低、漏检率高以及白坯布数据资源较少等问题, 本文在 YOLO11 的基础上提出了 DAM-YOLO, 在优化模型检测性能的同时, 兼顾平台资源的消耗。首先, 在 Backbone 网络添加了 C3K2_DCNv2 模块, 设计了多尺度特征提取网络, 显著提升了几何建模能力, 并增强了对形态差异较大目标的特征提取能力。其次, 在 Neck 部分将 ASF 模块与 DySample 模块进行融合, 从而有效保留特征图的细节信息, 并提升了图像分辨率, 优化了特征融合效果。最后, 通过更换检测头和使用 MultiSEAM 模块, 提升了对小目标的定位精度。实验结果显示, DAM-YOLO 模型在保证检测效率的前提下, mAP@0.5 相较原始 YOLO11 提升了 2.3%, 验证了改进模型的有效性和实用性。未来的研究将重点优化模型结构, 提升检测精度, 并探索其在其他工业应用中的潜力。

参考文献:

- [1] 张凯俊, 张本国, 马棒棒, 等. 改进的混沌人工蜂群算法-支持向量机漏钢预报模型[J]. 材料与冶金学报, 2025, 24(3): 237-242.
- ZHANG Kaijun, ZHANG Benguo, MA Bangbang, et al. Breakout prediction model with improved chaotic bee colony algorithm optimised-support vector machine[J]. Journal of Materials and Metallurgy, 2025, 24(3): 237-242.
- [2] 何俊, 张彩庆, 李小珍, 等. 面向深度学习的多模态融合技术研究综述[J]. 计算机工程, 2020, 46(5): 1-11.

HE Jun, ZHANG Caiqing, LI Xiaozhen, et al. Review of multi-modal fusion technology for deep learning[J]. Computer Engineering, 2020, 46(5): 1-11.

- [3] DU L X, ZHANG R Y, WANG X T. Overview of two-stage object detection algorithms[J]. Journal of Physics: Conference Series, 2020, 1544(1): 012033.
- [4] TIAN Z, SHEN C H, CHEN H, et al. FCOS: fully convolutional one-stage object detection[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea: IEEE, 2019: 9626-9635.
- [5] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 1440-1448.
- [6] 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, NV, USA: IEEE, 2016: 779-788.
- [7] 苏立鹏, 娄益凡, 杨吴奔, 等. 基于改进 Faster-RCNN 的起重机钢丝绳表面缺陷识别方法[J]. 机电工程, 2025, 42(7): 1341-1349.
- SU Lipeng, LOU Yifan, YANG Wuben, et al. Method of surface defect recognition of crane wire rope based on improved Faster-RCNN[J]. Journal of Mechanical & Electrical Engineering, 2025, 42(7): 1341-1349.
- [8] 徐春鸽. 改进的 YOLOv5s 布匹疵点检测算法研究[J]. 现代工程科技, 2025, 4(2): 1-4.
- XU Chungu. Research on improved YOLOv5s fabric defect detection algorithm[J]. Modern Engineering Technology, 2025, 4(2): 1-4.

- [9] 王千, 王国栋, 李军, 等. 基于改进的 YOLOv8 布匹残损检测算法 [J]. 青岛大学学报 (工程技术版), 2024, 39(2): 47-52.
WANG Qian, WANG Guodong, LI Jun, et al. Fabric damage detection algorithm based on improved YOLOv8[J]. *Journal of Qingdao University (Engineering & Technology Edition)*, 2024, 39(2): 47-52.
- [10] 郑雨婷. 基于 YOLOv5 的织物瑕疵检测算法研究[D]. 杭州: 浙江理工大学, 2023.
ZHENG Yuting. Research on algorithms of fabric defect detection based on YOLOv5[D]. Hangzhou: Zhejiang Sci-Tech University, 2023.
- [11] 毛龙妹, 邱莉, 陈晓东, 等. 基于深度学习和知识图谱的针织物疵点检测方法研究[EB/OL]. (2025-05-13)[2025-05-28]. <https://link.cnki.net/urlid/61.1132.TS.20250512.1347.002>.
MAO Longmei, QIU Li, CHEN Xiaodong, et al. Research on knitted fabric defect detection method based on deep learning and knowledge graph[EB/OL]. (2025-05-13)[2025-05-28]. <https://link.cnki.net/urlid/61.1132.TS.20250512.1347.002>.
- [12] YU Z P, HUANG H B, CHEN W J, et al. YOLO-FaceV2: a scale and occlusion aware face detector[J]. *Pattern Recognition*, 2024, 155: 110714.
- [13] KANG M, TING C M, TING F F, et al. ASF-YOLO: a novel YOLO model with attentional scale sequence fusion for cell instance segmentation[J]. *Image and Vision Computing*, 2024, 147: 105057.
- [14] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE, 2023: 6004-6014.
- [15] HUANG C K, NIEN H H. Multi chaotic systems based pixel shuffle for image encryption[J]. *Optics Communications*, 2009, 282(11): 2123-2127.
- [16] LI X, WANG W H, WU L J, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection[J]. *arXiv*, 2020: 2006.04388
- [17] HENDRYCKS D, GIMPEL K. Gaussian error linear units (GELUs)[J]. *arXiv*, 2023: 1606.08415.
- [18] CAO J M, LI Y Y, SUN M C, et al. DO-conv: depthwise over-parameterized convolutional layer[J]. *IEEE Transactions on Image Processing*, 2022, 31: 3726-3736.
- [19] HUA B S, TRAN M K, YEUNG S K. Pointwise convolutional neural networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 984-993.
- [20] GUO C X, FAN B, ZHANG Q, et al. AugFPN: improving multi-scale feature learning for object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2020: 12592-12601.