

改进 YOLOv11n 的布匹织物缺陷检测算法^{*}

王海群 陈珊珊

(华北理工大学电气工程学院 唐山 063000)

摘 要: 布匹织物缺陷的自动检测是保障纺织品质量的关键环节,针对布匹织物缺陷检测中因特征提取困难导致的漏检以及缺陷间差异不明显造成的误检问题,提出了一种改进的 YOLOv11n 检测算法。在主干网络 C3k2 中引入 DRB,增强多尺度特征表达能力;构建 C2TDME 模块替代原 C2PSA 中 PSABlock 部分,提升模型小目标检测与全局上下文建模能力;在颈部网络中,设计 GC-MANet 模块替代 C3k2,强化缺陷形态与纹理特征的捕捉,缓解背景与缺陷特征混淆问题;通过 RepStem 结构改进下采样过程,扩大了感受野并增加特征提取路径。实验结果表明,改进后算法的精度达到 93.2%、召回率达 90%、mAP50 达 94.1%,相较于原 YOLOv11n 模型分别提升 1.30%、7.53%、4.21%,有效提高了织物缺陷检测的准确性,满足工业生产的实际需求。

关键词: 缺陷检测;YOLOv11n;重参数卷积;MANet;Mona layer

中图分类号: TP391.41;TN86.2 **文献标识码:** A **国家标准学科分类代码:** 510.80

Improved defect detection algorithm for YOLOv11n fabric

Wang Haiqun Chen Shanshan

(Faculty of Electrical Engineering, North China University of Science and Technology, Tangshan 063000, China)

Abstract: Automatic detection of fabric defects is a critical step for quality assurance in the textile industry. However, current methods often miss defects because of difficult feature extraction. They can also misclassify defects that have little contrast with the fabric background. To address these challenges, we proposed an improved algorithm based on the YOLO version 11 nano model. First, we introduced a parameterized block with dilated convolutions into the main network backbone to enhance its ability to represent features at multiple scales. We also developed a new module to improve small object detection and model the global context. In the network's neck, we designed a specialized attention module. This module strengthens the capture of defect shapes and textures, which reduces confusion between defects and the background. Finally, we refined the downsampling process using a re-parameterized stem structure. This change expanded the model's receptive field and added more paths for feature extraction. Experimental results show that our improved algorithm achieves 93.2% precision, 90% recall, and 94.1% mean average precision. These results are 1.3%, 7.53% and 4.21% higher than the original model, respectively. The proposed algorithm effectively improves the accuracy of fabric defect detection and meets the practical demands of industrial production.

Keywords: defect detection;YOLOv11n;re-parameterized convolution;MANet;Mona layer

0 引 言

布匹织物在生产过程中易受到人为失误、机器损坏等影响,出现纱线断裂、污渍等缺陷,影响布匹织物质量。织物质量检测作为保证产品性能和市场竞争力的重要环节,高效、精确的织物缺陷检测算法成为纺织企业提升自动化水平和生产质量的关键技术需求。

传统布匹织物缺陷检测依赖于人工目视检查,受限于

检测效率低、主观因素影响大且难以实现大规模工业自动化。机器学习发展提升了织物缺陷检测效率,任梦凡等^[1]提出通过粗糙度测量对织物图像分类,结合颜色距离进行缺陷粗定位,利用显著性检测和二值化处理优化检测结果;刘秀平等^[2]提出采用等价旋转不变的局部二值模式(local binary pattern,LBP)算法提取纹理特征,结合低秩分解模型进行缺陷检测,提高检测的准确性和鲁棒性;Kim 等^[3]提出结合甲虫触角搜索(beetle antennae search,BAS)和引力

搜索算法 (gravitational search algorithm, GSA) 优化 Gabor 滤波器参数, 提高缺陷检测的准确性和效率; Kang 等^[4]提出基于形态学操作的检测方法, 结合腐蚀和膨胀操作增强缺陷特征, 通过人工神经网络选择结构元素, 提高检测的灵活性和准确性。虽相较于人工目视方法检测效率有所提升, 但仍未满足生产实际中对于布匹复杂纹理、多样性缺陷的检测要求。

深度学习的发展使目标检测领域焕发新生机, 深度学习模型可自动从数据中学习高层次、鲁棒的特征表示, 无需依赖人工设计的特征提取方法^[5]。基于深度学习的目标检测算法主要分为以 Faster R-CNN 和 Cascade R-CNN^[6-7]等为主的两阶段算法和以 YOLO 系列为主的单阶段算法。曾秀云等^[8]提出基于改进 Faster R-CNN 的棉布包装缺陷检测方法, 引入特征金字塔网络和双线性插值改进, 解决了小目标缺陷检测难题, 并通过数据增强提高模型泛化能力; 莫惠文等^[9]提出基于 UNet++ 自编码器的织物缺陷检测方法, 引入通道注意力和空间注意力模块, 结合 MAE 损失和 SSIM 损失, 提高模型鲁棒性和检测精度; 朱磊等^[10]提出了一种基于改进 YOLOv5 的织物缺陷检测方法, 通过引入双级联注意力机制 (DDA)、鬼影混洗卷积 (GSC) 和 SIOU 损失函数, 在不增加网络参数数量的前提下提升了检测精度和效率; 陈森等^[11]提出基于改进 YOLOv5 的织物缺陷检测方法, 引入 CBAM 注意力机制和 Focal IOU Loss, 增强特征提取能力和模型收敛速度; Tang 等^[12]提出基于 YOLOv7-Tiny 的轻量化检测算法, 优化特征提取模块和损失函数, 提升检测性能。化春键等^[13]提出基于 YOLO-CFD 的棉布微小微弱缺陷检测方法, 引入双层路由注意力机制和空间深度卷积块, 增强对小目标的检测能力, 并设计新的边界框回归损失函数, 提升检测精度。白先浪等^[14]提出基于改进 YOLOv5s 的皮革表面缺陷检测方法, 通过在主干网络中引入坐标注意力机制增强对缺陷区域的感知能力, 重构颈部网络为加权双向特征金字塔以优化多尺度特征融合, 并增设小目标检测头以提升对细微缺陷的捕获性能。

目前基于深度学习的织物缺陷检测方法在一定程度上提升了检测精度, 但在复杂实际场景下仍存在明显局限: 小目标缺陷在多次下采样过程中易丢失关键语义信息, 导致检出率偏低; 织物表面纹理重复且对比度弱, 现有模型对缺陷与背景的判别能力不足, 易产生误检; 缺陷尺度跨度大, 现有多尺度特征融合机制对不同尺度目标的响应不均衡, 影响检测稳定性。上述问题导致目前主流模型难以满足工业在线检测需求, 针对上述问题, 本文基于 YOLOv11n 作出以下改进。

1) 引入 DRB 模块替换主干网络 C3k2 的关键卷积层, 在训练时使用并行结构增强模型的特征提取能力, 在推理时重参数化为单一、高效的大型卷积层, 增强模型的多尺度特征表达能力。

2) 使用 C2TDME 替换主干网络 C2PSA 的 PSABlock 部分, 提升模型对于小目标的检测能力以及全局上下文建模能力。

3) 设计 GC-MANet 重构颈部网络的 C3k2, 利用 GCBLOCK 的重参数化优势与 MANet 在单一模块特征提取的多样性和高效性, 提升模型的通道信息交互能力。

4) 剔除 YOLOv11n 的原下采样卷积, 设计 RepStem 改进模型下采样, 在训练时使用并行分支和多种卷积核增大模型的参数空间及特征提取路径, 在推理时所有冗余的并行分支被融合为一个单一的、等效的卷积层, 保证模型的检测效率及精度。

1 YOLOv11n 算法介绍

YOLOv11n 的整体架构遵循经典的骨干网络 (backbone)、颈部网络 (neck) 和检测头 (head) 3 段式设计, 骨干网络沿用 YOLOv9 的通用高效层聚合网络 (generalized efficient layer aggregation network, GELAN), GELAN 为跨阶段局部网络 (cross stage partial network, CSPNet) 变体, 结合了 CSPNet 的跨阶段连接思想与高效层聚合网络 (efficient layer aggregation network, ELAN) 的高效层聚合能力, 在保持轻量化的同时, 具备强大的特征提取能力; 颈部网络使用路径聚合网络 (path aggregation network, PAN) 结构, 通过自顶向下和自底向上的双向路径, 将高层的强语义信息与低层级的丰富空间细节信息进行融合, 形成多个尺度的特征金字塔, 使得模型能够同时对不同大小的目标进行精准检测, 增强模型的多尺度检测能力; 检测头采用无锚框 (anchor-free) 的解耦检测头, anchor-free 机制直接预测目标的中心点、宽度和高度, 减少了与锚框相关的超参数设计, 不仅简化了训练流程, 还提高了模型对不规则形状目标的适应性, 解耦检测头将分类任务和回归任务在检测头的最后几层进行分离, 使用不同的分支分别进行处理, 分类任务和定位任务的目标存在一定的冲突, 将它们解耦可以有效避免二者之间的相互干扰, 从而提升模型的收敛速度和检测精度。

2 YOLOv11n 算法改进

本文改进 YOLOv11n 模型结构如图 1 所示。

2.1 多尺度重参数特征提取模块 C3k2-DRB

在布匹缺陷检测任务中, 由于织造工艺、光照条件及缺陷类型的多样性, 缺陷目标常呈现出显著的尺度变化与形态差异。YOLOv11n 主干的原始 C3k2 模块虽然具备一定的特征提取能力, 但其内部的卷积操作通常采用标准尺寸的卷积核, 这在处理跨度较大的条纹状缺陷或分布稀疏的微小斑点时, 存在特征感受野单一、多尺度信息捕获能力不足的固有局限性。其特征融合机制偏向于局部整合, 缺少对全局及稀疏上下文信息的有效建模, 进而可能影响对尺度变化显著缺陷的识别精度。

$$\mathbf{Y}_{\text{drb_out}} = \mathcal{F}_{\text{DRB-train}}(\mathbf{X}_{\text{drb_in}}) = \underbrace{\text{BN}_0(\text{Conv}_0(\mathbf{X}_{\text{drb_in}}))}_{\text{Main-Branch}} + \sum_{i=1}^3 \underbrace{\text{BN}_i(\text{Conv}_i(\mathbf{X}_{\text{drb_in}}))}_{\text{Enhancement-Branches}} \quad (1)$$

在推理阶段,DRB的多分支结构等效融合为一个单一的、密集的 7×7 卷积层,即结构重参数化过程。对于第 i 个分支,其Conv-BN结构等效为一个具有新权重 $\mathbf{W}'_i$ 和新偏置 b'_i 的单一卷积层,假设原始权重和偏置为 $\mathbf{W}_i, b_i$,BN层参数为 $\mu_i, \sigma_i^2, \gamma_i, \beta_i$,融合公式如式(2)、(3)所示。

$$\mathbf{W}'_i = \frac{\gamma_i}{\sqrt{\sigma_i^2 + \epsilon}} \mathbf{W}_i \quad (2)$$

$$b'_i = \frac{\gamma_i(b_i - \mu_i)}{\sqrt{\sigma_i^2 + \epsilon}} + \beta_i \quad (3)$$

其中, ϵ 为一个极小常数,用于维持数值稳定性。

将所有分支的卷积核变换并扩展至与主干分支相同的 7×7 尺寸,对于扩张率为 d_i 的空洞卷积核 $\mathbf{W}'_i$,需将其

变换为等效的系数标准卷积核 $\text{DilatedToStd}(\mathbf{W}'_i)$,将变换后的所有分支等效卷积核与偏置相加,融合卷积核 $\mathbf{W}_{\text{rep}}$ 与偏置 b_{rep} 的计算公式如式(4)、(5)所示。

$$\mathbf{W}_{\text{rep}} = \text{Pad}(\mathbf{W}'_0) + \sum_{i=1}^3 \text{Pad}(\text{DilatedToStd}(\mathbf{W}'_i)) \quad (4)$$

$$b_{\text{rep}} = b'_0 + \sum_{i=1}^3 b'_i \quad (5)$$

重参数化后的前向传播过程简化为式(6)所示。

$$\mathbf{Y}_{\text{drb_out}} = \mathcal{F}_{\text{DRB}}(\mathbf{X}_{\text{drb_in}}) = \text{Conv}_{\text{rep}}(\mathbf{X}_{\text{drb_in}}) \quad (6)$$

C3k2-DRB模块与一个使用 7×7 大核卷积的标准C3k2模块完全相同,保持了单路架构的高效性。改进后的模型在反向传播时,能够从多个感受野尺度上接受梯度,提升模型对缺陷的识别能力。C3k2-DRB模块如图3所示。

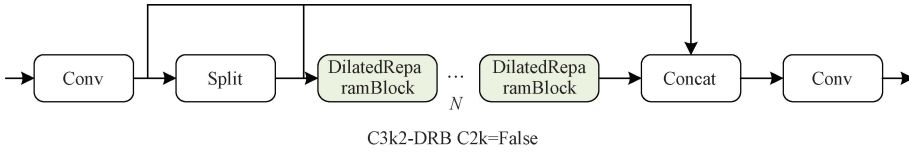


图 3 C3k2-DRB 结构

Fig. 3 C3k2-DRB structure

2.2 自适应边缘感知模块 C2TDME

跨阶段自注意力模块(cross-stage conv and positional self-attention, C2PSA)为YOLOv11n增强特征表达能力与跨层信息融合的关键组件,但原始主干网络的C2PSA对于带纱、断纱、脱纱等线性和高频特征主导的缺陷类型感知能力不足,易出现关键缺陷在深度传播中消失的问题。针对上述问题,本文设计了一种基于动态双曲正切函数(dynamic tanh, DyT)、令牌统计量自注意力机制(token statistics self-attention, TSSA)、高效判别式频域前馈网络(efficient discriminative frequency domain-based feedforward network, EDFFN)和多认知视觉适配器层(multi-cognitive visual adapter layer, Mona-layer)的改进Swin-Transformer Block来替代C2PSA的PSABlock部分^[16-20],提升模型对于小目标的检测能力以及全局上下文建模能力,减少模型对于布匹织物缺陷的漏检率、误检率。

Swin Transformer Block是Swin Transformer的核心计算单元,在保持可控计算量的前提下实现高效的局部-全局信息建模。Swin Transformer Block以分层视觉表示为设计基底,通过在局部窗口内进行自注意力计算(window-based multi-head self-attention, W-MSA),并在相邻的层之间通过对窗口划分的循环平移(shifted window, SW-MSA)实现跨窗口的信息交互,增大全局感受野的同时保证模型的计算效率。

在注意力机制层面,以TSSA取代Swin Transformer Block原有的多头注意力,传统注意力机制依赖于计算所

有令牌对间的相似度,计算复杂度为 $O(n^2)$,TSSA基于最大编码率降低(MCR2)原则,通过将输入令牌投影至多个子空间,计算其加权二阶统计量,生成一个所有令牌共享的、数据依赖的全局特征选择滤波器,该滤波器能够动态地识别并保留信息丰富的特征,同时抑制冗余或噪声,可将计算复杂度降低为 $O(n)$ 并为模型提供了一种更具全局视野且计算高效的上下文信息整合方式。

对于输入令牌序列 $\mathbf{X} \in \mathbf{R}^{N \times D}$, N 为令牌数, D 为维度,TSSA计算注意力输出如式(7)~(9)所示。

$$\Sigma_{\mathbf{x}} = \frac{1}{N} \mathbf{X}^T \mathbf{X} - \left(\frac{1}{N} \mathbf{X}^T \mathbf{1} \right) \left(\frac{1}{N} \mathbf{1}^T \mathbf{X} \right) \quad (7)$$

$$\mathbf{F}_{\text{global}} = \text{MCR2}(\Sigma_{\mathbf{x}}) \quad (8)$$

$$\mathbf{Z} = \mathbf{X} \cdot \mathbf{F}_{\text{global}} \quad (9)$$

其中, $\Sigma_{\mathbf{x}}$ 是输入令牌 $\mathbf{X}$ 的协方差矩阵; $\mathbf{1}$ 是全1向量;MCR2($\cdot$)为基于最大编码率降低原则导出的变换函数, $\mathbf{F}_{\text{global}} \in \mathbf{R}^{D \times D}$ 变换函数根据协方差矩阵生成的数据依赖全局特征选择滤波器; $\mathbf{Z}$ 为经过全局滤波器加权后的输出。

引入EDFFN克服标准前馈网络因位置无关性而导致的局部空间细节感知不足的缺陷,EDFFN的核心为其局部增强分支,该分支将序列化的令牌恢复至二维空间形态,利用轻量的深度可分离卷积(DW-Conv)高效聚合邻近信息。边缘增强模块(edge enhancement module, EEM)通过固定的Sobel高通滤波器生成边缘响应,引导门控机制,使模型自适应地增强特征图中边缘和轮廓等高频细节地

响应强度,提升模型的检测精度。

对于 TSSA 的输出 $Z \in \mathbf{R}^{N \times D}$, EDDFN 的处理过程如式(10)~(12)所示。

$$Z_{2D} = \text{Reshape}(Z) \quad (Z_{2D} \in \mathbf{R}^{H \times W \times D}) \quad (10)$$

$$E = \text{Sobel}(Z_{2D}) \quad (11)$$

$$Y_{\text{EDDFN}} = \text{MLP}(Z) + \sigma(E) \odot \text{DW-Conv}(Z_{2D}) \quad (12)$$

式中:Reshape($\cdot$)为将序列化的令牌恢复为二维空间形态;Sobel($\cdot$)为固定的 Sobel 算子,用于提取边缘响应 E ; $\sigma(\cdot)$ 为 Sigmoid 门控函数; $\odot$ 为逐元素乘法; DW-Conv($\cdot$)为聚合局部邻近信息的轻量级深度可分离卷积;MLP($\cdot$)为负责全局特征变换的标准多层感知机分支。

将 Mona Layer 引入到 Swin Transformer Block 中,在模块的关键节点嵌入轻量级适配器结构,每个适配器采用瓶颈式架构,利用 Mona Layer 的多认知卷积过滤器组通过不同尺寸的卷积核从多个尺度提炼视觉特征,使得改进后的模型仅需更新极少量的参数,极大地提升了模型地部署与迭代效率。

对于输入特征 Z_{in} , Mona Layer 计算方法如式(13)所示。

$$Y_{\text{Mona}} = Z_{\text{in}} + W_{\text{up}} \cdot \text{GELU}\left(\sum_{k \in K} \text{Conv}_k(W_{\text{down}} \cdot Z_{\text{in}})\right) \quad (13)$$

式中: W_{up} 、 W_{down} 分别为升维和降维的线性投影层,构成瓶颈结构; Conv_k 为多认知过滤器组中的一个卷积,其核尺寸 k 为预设集合;GELU 为激活函数。

使用 DyT 替换 Swin Transformer Block 传统归一化层。DyT 通过一个简单的、逐元素的 tanh 函数加可学习缩放参数的形式,模拟了归一化层抑制极端激活值的 S 型映射效果。这种方法完全不依赖于样本间的统计信息,不仅极大地减少了计算开销,还从根本上杜绝了由批次变化带来的不稳定性,提升了模型的稳定性。DyT 结构如图 4 所示。

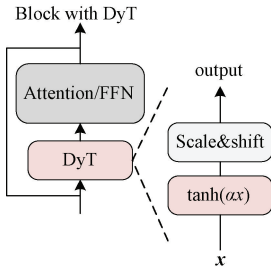


图 4 DyT 结构

Fig. 4 DyT structure

对于输入向量 x 中的每个元素 x_i , 计算公式如式(14)所示。

$$\text{DyT}(x_i) = \alpha_i \cdot \tanh(x_i) \quad (14)$$

其中, α_i 为与 x_i 相对应位置的可学习缩放参数。

改进后的 Swin Transformer Block 替代 C2PSA 的

PSABlock 部分,构成 C2TDME 模块,C2TDME 显著提升了模型对小目标缺陷的定位精度,增强了模型的上下文建模能力,提升了模型对布匹织物缺陷的检测精度与模型稳定性,有效降低了漏检率、误检率。C2TDME Block 结构如图 5 所示。

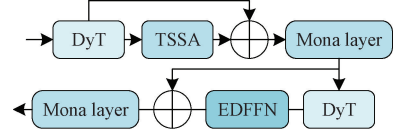


图 5 C2TDMEBlock 结构

Fig. 5 C2TDMEBlock structure

2.3 可重参数化混合聚合网络 GC-MANet

YOLOv11n 的颈部网络通过 C3k2 模块进行深层特征融合时,利用堆叠卷积操作来扩大感受野,但增强语义信息的同时,下采样和局部卷积操作易导致微小目标的精细空间信息和弱纹理特征丢失。为增强 YOLOv11n 的多尺度局部细节分辨能力和全局上下文感知能力,本文设计了可重参数化混合聚合网络 GC-MANet,并使用其对 YOLOv11n 的颈部 C3k2 进行重构,提升模型的检测精度和计算效率。

金箍棒模块(golden cudgel block, GCBLOCK)^[21] 作为一个可重参数化卷积,其核心思想是构建一个在训练阶段拥有复杂拓扑结构的模块以增强模型的特征表征能力;在训练完成后,通过数学等价变换,将这个复杂结构重参数化为一个用于推理的、简洁高效的单一结构。在训练阶段, GCBLOCK 使用结构解耦设计,将一个卷积操作分解为多个并行的分支,包含一个 3×3 卷积分支、一个 1×1 卷积分支以及一个恒等映射分支。这 3 个分支并行处理输入特征,并将输出逐元素相加, 3×3 卷积分支负责捕捉具有空间局部性的图案和纹理信; 1×1 卷积分支负责进行跨通道的信息融合与变换;恒等映射分支构建一条残差连接,保证了信息的无损流通,有效缓解了深度网络中的梯度消失问题。针对带纱、断纱等细微线性缺陷,使用该模块进行训练时, 3×3 卷积分支可精确捕捉缺陷的形态和纹理, 1×1 卷积分支可增强模型在通道维度上与背景织物的细微差异的敏感度,提升对于小目标的检测精度。训练阶段计算方法如式(15)所示。

$$Y_{\text{train}} = \text{BN}_1(\text{Conv}_{3 \times 3}(X)) + \text{BN}_2(\text{Conv}_{1 \times 1}(X)) + \text{BN}_{\text{identity}}(X) \quad (15)$$

在推理阶段,所有并行分支的参数可以被融合进一个单一的 3×3 卷积核中, 1×1 卷积核可以通过 Padding 变换为一个等效的 3×3 卷积核,恒等映射分支也可以变换为一个特殊的 3×3 单位卷积核。等效的 3×3 卷积核相加,形成一个功能等价但结构单一的最终 3×3 卷积层,当模型在部署推理时,完全摆脱了多分支带来的额外计算和内存开销,与普通单路卷积网络速度相同,使模型在保持

高检测精度。推理阶段计算方法如式(16)所示。

$$\mathbf{Y}_{\text{infer}} = \text{Conv}_{\text{final}}(\mathbf{X}) \quad (16)$$

GCBlock 的结构如图 6 所示。

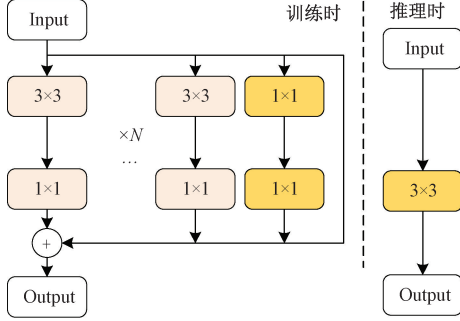


图 6 GCBlock 结构

Fig. 6 GCBlock structure

混合聚合网络(mixed aggregation network, MANet)^[22]为一种复合型卷积模块, MANet 可并行地运用多种不同类型地卷积操作, 并将它们的输出进行聚合, 实现在单一模块特征提取的多样性和高效性。

假设输入特征 $\mathbf{X} \in \mathbf{R}^{C_{\text{in}} \times H \times W}$, 经过一个 1×1 卷积层, 将其通道维度扩展至 $2c$, 扩展后的特征图被送入多个并行处理的分支, 分别为 1×1 卷积分支、深度可分离卷积分支(depthwise separable convolution, DSConv)和由 n 个 ConvNeck 模块串联组成的深度变换分支。 1×1 卷积分支, 负责进行通道维度的信息交互和极小感受野的特征提取; 深度可分离卷积分支通过深度卷积(depthwise conv, DWConv)和逐点卷积(pointwise conv, PWConv)的组合, 以较低的参数量和计算成本高效地提取空间特征; 每个 ConvNeck 模块由两个 3×3 卷积构成, 通过堆叠加深局部区域的非线性特征变换能力。

所有并行分支的输出特征图在通道维度上被拼接起来, 形成一个包含了多种感受野和抽象层次的特征图, 再由一个 1×1 卷积对该聚合特征进行降维和信息融合, 得到最终输出 $\mathbf{X}_{\text{out}} \in \mathbf{R}^{C_{\text{out}} \times H \times W}$ 。MANet 通过这种混合聚合的设计, 使网络能够自适应地从不同复杂度的路径中学习和组合最有效的特征表示。MANet 虽考虑了特征多样性, 但在深度变换分支上仍采取传统卷积堆叠, 推理效率较慢, 本文引入可重参数化卷积 GCBlock 对 MANet 进行改进, 设计了一种可重参数化混合聚合网络 GC-MANet。

GC-MANet 保留了 MANet 并行的多分支架构, 包括高效的 1×1 卷积分支和 DSConv 分支, 将由 n 个 ConvNeck 模块串联构成的第 3 个分支, 替换为由 n 个 RepConv 模块串联构成的新分支。保持通道信息交互和轻量化特征提取的优势下, 提升模型的计算效率和检测精度。

对于输入特征 $\mathbf{F}_{\text{in}} \in \mathbf{R}^{C \times H \times W}$, GC-MANet 结构公式如式(17)~(22)所示。

$$\mathbf{F}_{\text{exp}} = \text{Conv}_{1 \times 1}^{\text{exp}}(\mathbf{F}_{\text{in}}) \quad (\mathbf{F}_{\text{exp}} \in \mathbf{R}^{2C \times H \times W}) \quad (17)$$

$$\mathbf{O}_1 = \text{Conv}_{1 \times 1}^{\text{br1}}(\mathbf{F}_{\text{exp}}) \quad (18)$$

$$\mathbf{O}_2 = \text{DSConv}^{\text{br2}}(\mathbf{F}_{\text{exp}}) \quad (19)$$

$$\mathbf{O}_3 = (\text{GCBlock}_n \circ \dots \circ \text{GCBlock}_1)(\mathbf{F}_{\text{exp}}) \quad (20)$$

$$\mathbf{F}_{\text{cat}} = \text{Concat}(\mathbf{O}_1, \mathbf{O}_2, \mathbf{O}_3) \quad (21)$$

$$\mathbf{F}_{\text{out}} = \text{Conv}_{1 \times 1}^{\text{fuse}}(\mathbf{F}_{\text{cat}}) \quad (\mathbf{F}_{\text{out}} \in \mathbf{R}^{C \times H \times W}) \quad (22)$$

其中, $\text{Conv}_{1 \times 1}^{\text{exp}}$ 和 $\text{Conv}_{1 \times 1}^{\text{fuse}}$ 分别为用于通道扩展和最终融合的 1×1 卷积; $\mathbf{O}_1, \mathbf{O}_2, \mathbf{O}_3$ 分别为 3 个并行分支的输出; $\circ$ 表示函数复合, 级 GCBlock 的串联; Concat 为沿通道维度的拼接操作。

在训练阶段, GC-MANet 的第 3 个分支与前两个分支协同, 极大地丰富了整个模块的特征学习空间, 使网络能更精细地捕捉各类织物瑕疵的特征。在推理阶段, 第 3 个分支被重参数化为一条简洁的 3×3 卷积链, 使得整个 GC-MANet 模块在保持强大性能的同时, 提高了模型的计算效率。GC-MANet 结构如图 7 所示。

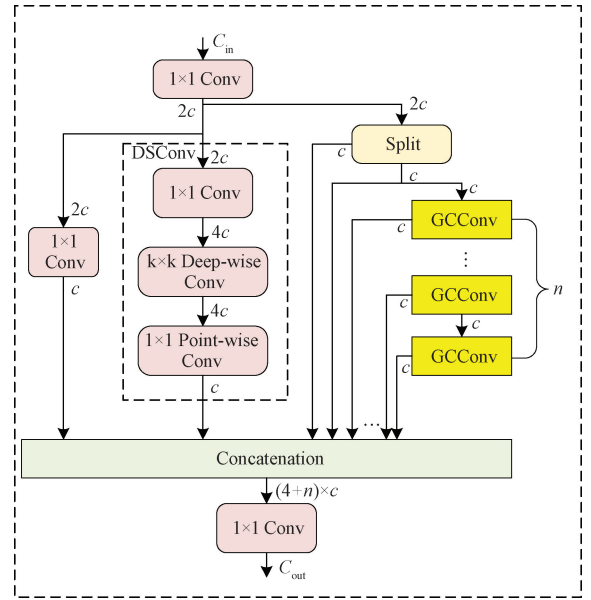


图 7 GC-MANet 结构

Fig. 7 GC-MANet structure

2.4 RepStem 下采样

YOLOv11n 原始下采样使用单一卷积, 难以捕捉输入图像的丰富特征, 且感受野固定, 易丢失大量细节信息, 布匹织物中含有大量例如破洞、脱纱等小目标缺陷, 使用原始 YOLOv11n 进行检测, 导致漏检、误检。使用结构重参数模块 RepStem^[23]改进 YOLOv11n 模型的下采样部分, 在训练时通过并行分支和多种卷积核, 使模型拥有更大的参数空间和更丰富的特征提取路径, 且不同大小的卷积核在下采样的同时, 可以捕获多尺度信息, 提升模型的检测精度; 在推理时所有冗余的并行分支被融合为一个单一的、等效的卷积层, 保证精度的同时, 提升模型的推理速度。

训练阶段 RepStem 的采样阶段以 stride=2 实现空间

下采样,模型主干部分采用经典的多分支设置,主干分支的 3×3 卷积提取局部特征, 1×1 卷积提供通道混合与参数化冗余,经由求和、激活将并行的结果进行融合;两路并行深度卷积的 3×3 和 1×1 深度可分离卷积进一步提高计算效率,同样经由求和、激活进行融合;最后经过 1×1 点卷积用于混合通道信息并作为这部分的最终输出。

对于输入特征图 $\mathbf{X}$,多分支的计算方式如式(23)~(26)所示。

$$\mathbf{Y}_1 = \text{BN}_1(\text{Conv}_{3\times 3}^{s=2}(\mathbf{X}) + \text{Conv}_{1\times 1}^{s=2}(\mathbf{X})) \quad (23)$$

$$\mathbf{Y}_2 = \text{BN}_2(\text{DWConv}_{3\times 3}^{s=2}(\mathbf{X}) + \text{DWConv}_{1\times 1}^{s=2}(\mathbf{X})) \quad (24)$$

$$\mathbf{Y}_2 = \text{BN}_2(\text{DWConv}_{3\times 3}^{s=2}(\mathbf{X}) + \text{DWConv}_{1\times 1}^{s=2}(\mathbf{X})) \quad (25)$$

$$\mathbf{Y}_{\text{train}} = \text{Conv}_{1\times 1}^{\text{final}}(\text{Activation}(\mathbf{Y}_1 + \mathbf{Y}_2)) \quad (26)$$

式中: $s=2$ 为卷积层步长为2; BN_1 、 BN_2 为各自块内的批归一化层; Activation 为激活函数; $\text{Conv}_{1\times 1}^{\text{final}}$ 为最后的 1×1 点卷积。

推理阶段将训练阶段的多分支子结构合并为等效的单一卷积 $\text{Conv}_{\text{equivalent}}$,显著降低计算开销,同时保留训练阶段的检测精度。 $\text{Conv}_{\text{equivalent}}$ 计算方法如式(27)所示。

$$\mathbf{Y}_{\text{infer}} = \text{Conv}_{\text{equivalent}}(\mathbf{X}) \quad (27)$$

RepStem结构如图8所示。

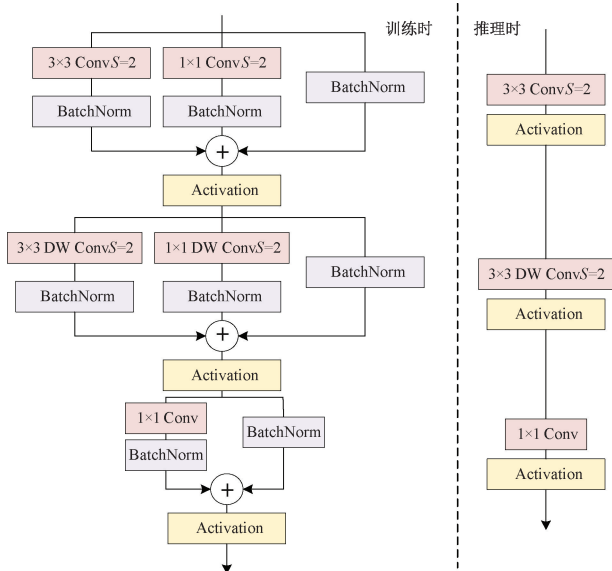


图8 RepStem结构

Fig. 8 RepStem structure

3 实验结果与分析

3.1 实验环境及参数配置

本实验环境均为 Ubuntu20.04,深度学习框架为 PyTorch 1.11.0,python 版本为 3.8,CUDA 版本为 11.3,训练 GPU 为 NVIDIA Geforce RTX 3090,实验的训练关键参数如表1所示。

经实验表明,当输入图像为 640×640 时,优化器选用

表1 参数配置

Table 1 Parameter configuration

函数	设置值
优化器	SDG
输入图像大小	640×640
学习率	0.01
Epochs	200
批处理大小	48

SGD,学习率为 0.01,训练轮次为 200,批处理大小为 48 时,模型训练最稳定。

3.2 数据集

本文所使用的数据集来源于真实织物生产车间的在线检测场景,使用工业摄像头 Hikvision 在实际生产工况下进行采集,数据覆盖设备振动、环境干扰等典型工业环境因素。为消除环境散光对图像采集的影响,采用同轴环形 LED 光源继续正面均匀照明,拍摄距离固定为 20 cm,以保证图像尺度的一致性。采集的原始图像分辨率为 $3\,072\times 2\,048$ pixels,图像中包含多种形态、尺度及对比度差异明显的真实织物缺陷,具有显著的小目标与弱纹理特征。为适应模型输入要求并保留缺陷细节,所有图像预处理阶段被统一裁剪并缩放至 640×640 。缺陷类别包括带纱(DaiSha)、断纱(DuanSha)、棉球(MianQiu)、破洞(PoDong)、脱纱(TuoSha)、污渍(WuZi)6类缺陷,如图9所示。使用 LabelImg 标注软件将织物缺陷数据标注为 YOLO 格式,共有图像 2 115 张,其中,带纱缺陷有 328 张图像,断纱缺陷有 244 张图像,棉球缺陷有 138 张图像,破洞缺陷有 89 张图像,脱纱缺陷有 522 张图像,污渍缺陷有 329 张图像,各类样本数量分布不均,整体特性与实际工业生产场景高度一致,按照 8:2 的比例划分为 1 650 张训练集和 465 张验证集。部分缺陷效果如图 9 所示。

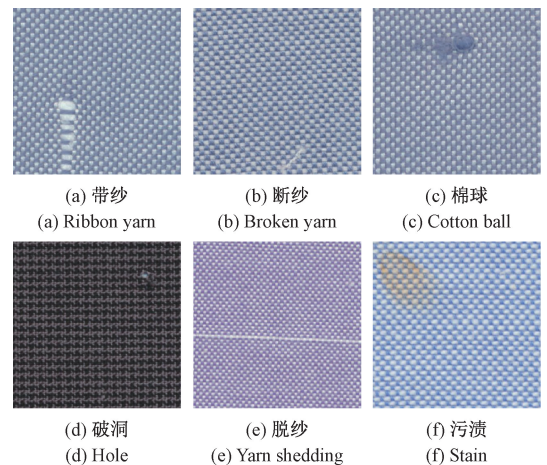


图9 缺陷效果

Fig. 9 Defect effect

3.3 评价指标

本文使用准确率 (precision, P)、召回率 (recall, R)、平均精度 (average precision, AP)、平均精度均值 (mean average precision, mAP)、浮点运算次数 (GFLOPs)、参数量 (Params) 来评估模型性能。各评价指标的计算公式如式 (28)~(31) 所示。

$$P = \frac{TP}{TP + FP} \tag{28}$$

$$R = \frac{TP}{TP + FN} \tag{29}$$

$$AP = \int_0^1 P(r) dr \tag{30}$$

$$mAP = \frac{\sum_{i=0}^n AP_i}{n} \tag{31}$$

其中, TP 为模型正确预测的正样本, FP 为模型错误预测的正样本, FN 为模型未能检测出的真实目标。AP 为单个类别在不同召回率水平下的平均精度, 是 P-R 曲线下的面积, mAP 为所有类别 AP 值的平均值, n 为所有类别数, AP_i 为第 i 类的平均精度值。

3.4 主干网络 C3k2 模块对比实验

为验证使用 DRB 改进 C3k2 的优势, 将 DynamicConv、ODConv、gConv、PConv^[24-27] 几个卷积模块分别引入到模型的相同位置与本文所使用模块进行实验对比, 实验结果如表 2 所示。

表 2 主干网络 C3k2 模块对比实验
Table 2 Comparative experiment on the C3k2 module in the backbone network

模型	mAP50	GFLOPs	Params/($\times 10^6$)
基线	0.903	6.3	2.58
+DynamicConv	0.912	6.3	2.91
+ODConv	0.843	6.0	2.62
+gConv	0.863	6.1	2.43
+PConv	0.875	6.0	2.41
+DRB	0.917	6.3	2.50

根据表 2 数据可知, 在 C3k2 中引入 ODConv、gConv、PConv 后检测的平均精度 mAP50 分别下降了 6.62%、4.43%、3.1%, 虽然计算量相比原网络有所降低, 但不符合工业生产对于检测精度的要求, 易出现漏检、误检等问题; 引入 DynamicConv 后检测 mAP50 上升了 1%, 但参数量上升了 0.33×10^6 , 增大了模型对于检测设备的算力要求; 引入本文所使用的 DRB 后, 模型的 mAP50 上升了 1.55%, 且参数量下降了 0.08×10^6 , 证明了本文改进的有效性。

3.5 主干网络 C2PSA 模块对比实验

为验证 C2TDME 模块替代 YOLOv11n 主干 C2PSA 的有效性, 将 CGLU^[28]、FMFFN、DyT、EDFFN、Mona 引

入到模型的相同位置与本文使用模块进行对比实验, 实验结果如表 3 所示。

表 3 主干网络 C2PSA 模块对比实验
Table 3 Comparative experiment on the C2PSA module in the backbone network

模型	mAP50	GFLOPs	Params/($\times 10^6$)
C2PSA	0.903	6.3	2.58
+CGLU	0.915	6.3	2.58
+FMFFN	0.921	6.3	2.59
+DyT	0.917	6.3	2.58
+EDFFN	0.919	6.3	2.63
+Mona	0.905	6.4	2.58
C2TDME	0.924	6.4	2.68

根据表 3 数据可知, 引入不同模块改进后, 模型的检测精度均有所上升, 引入 CGLU、DyT 后, 模型的计算量和参数量没有发生变化, 较 YOLOv11n 基础模型 mAP50 分别上升了 1.33%、1.55%; 引入 FMFFN 后, 模型的参数量上升了 0.01×10^6 , mAP50 上升了 1.99%, 相较于其他模块, 精度提升最为明显; 引入 EDFFN 后, mAP50 上升 1.77%, 参数量上升 0.05×10^6 ; 引入 Mona 后, mAP50 上升了 0.22%, 计算量上升了 0.1 GFLOPs; 几个模块引入至 C2PSA 后, 参数量和计算量均无较明显变化, 没有增大模型对于算力的要求。使用本文所改进模块, 模型的 mAP50 上升了 2.33%, 计算量和参数量仅分别上升 0.1 GFLOPs、 0.1×10^6 , 符合工业生产对于检测精度和计算开销的要求, 证明了本文改进的有效性。

3.6 颈部网络 C3k2 模块对比实验

为验证使用 GC-MANet 替代颈部 C3k2 模块的有效性, 本文将 CFBlock、ELGCA、EMBC、GCCConv、SFSCConv^[29-32] 引入到模型的相同位置与之进行实验对比, 实验结果如表 4 所示。

表 4 颈部网络 C3k2 模块对比实验
Table 4 Comparative experiment on the C3k2 module in the neck network

模型	mAP50	GFLOPs	Params/($\times 10^6$)
基线	0.903	6.3	2.58
+CFBlock	0.926	6.7	2.93
+ELGCA	0.919	6.3	2.52
+EMBC	0.901	6.2	2.78
+GCCConv	0.892	6.7	2.67
+SFSCConv	0.924	6.1	2.48
GC-MANet	0.930	6.4	2.77

根据表 4 数据可知, 引入 CFBlock 后, 模型的 mAP50 上升了 2.55%, 计算量上升了 0.4 GFLOPs, 参数量上升了 0.35×10^6 , 不符合模型对于低算力要求; 引入 ELGCA、

SFSCConv后,参数量分别降低 0.06×10^6 、 0.1×10^6 ,mAP50分别上升1.77%、2.33%,虽模型参数量有所降低,但mAP50提升并不明显;引入EMBC后,mAP50下降且参数量上升了 0.2×10^6 ,改进效果较差;使用本文GC-MANet改进后,参数量上升了 0.19×10^6 ,计算量仅上升了0.1 GFLOPs,mAP50上升了3%,在没有明显提升计算量的同时,保证了检测精度的最大化提升,证明了本文改进的有效性和先进性。

3.7 下采样模块对比实验

为验证改进下采样对于模型性能的影响,将ContextGuidedDown、v7DS、SRFD^[33-35]分别引入至模型的不同位置与本文所使用的RepStem进行实验对比,实验结果如表5所示。

根据表5数据可知,引入CGDown下采样后,模型的mAP50提升最高为3.32%,但计算量上升至9.0 GFLOPs,参数量上升至 3.53×10^6 ,不符合模型的实时性检测需求;引入v7DS下采样,模型的计算量和参数量虽有所下降,但检测精度mAP50相较于基线变低,不符合模型的精度要求;引入SRFD下采样后,mAP50上升了2.77%,参数量下降

表5 下采样模块对比实验

Table 5 Comparative experiment on downsampling modules

模型	mAP50	GFLOPs	Params/($\times 10^6$)
基线	0.903	6.3	2.58
CGDown	0.933	9.0	3.53
v7DS	0.900	5.7	2.23
v7DS	0.928	7.6	2.56
RepStem	0.927	6.1	2.55

了 0.02×10^6 ,但计算量提升较多,仍不符合模型的实时性检测要求;使用本文RepStem改进模型下采样后,mAP50提升2.66%,计算量下降了0.2 GFLOPs,参数量下降了 0.03×10^6 ,综合性能相较于其他模块最优,符合工业生产实际对于模型检测精度和实时性要求。

3.8 消融实验

为验证本文改进方案对于YOLOv11n模型的改进效果,在布匹织物缺陷数据集上进行消融实验。其中,“√”表示使用该模块,“×”表示未使用该模块。实验结果如表6所示。

表6 消融实验

Table 6 Ablation study

模型	C3k2-DRB	C2TDME	GC-MANet	RepStem	P	R	mAP50	Params/($\times 10^6$)	GFLOPs	FPS/fps
1	×	×	×	×	0.920	0.837	0.903	2.58	6.3	221.65
2	√	×	×	×	0.892	0.872	0.917	2.50	6.3	225.61
3	×	√	×	×	0.901	0.878	0.924	2.68	6.4	220.51
4	×	×	√	×	0.916	0.867	0.930	2.77	6.4	217.71
5	×	×	×	√	0.908	0.882	0.927	2.55	6.1	209.05
6	√	√	×	×	0.910	0.839	0.920	2.58	6.5	225.32
7	√	√	√	×	0.925	0.882	0.939	2.50	6.5	229.72
8	√	√	√	√	0.932	0.900	0.941	2.65	6.5	208.51

根据表6数据可知,引入各模块后模型的检测效果均有提升,实验1为基础YOLOv11n模型,实验2为使用DRB改进主干C3k2,相较于基础模型,R上升4.18%,mAP50提升1.55%,参数量降低了 0.08×10^6 ,FPS提升了3.96 fps,证明该模块提升了主干网络捕捉不同尺度上下文信息的能力;实验3为仅使用C2TDME替换主干C2PSA,改进后模型的R相较于基础模型提升了4.9%,mAP50提升了2.33%,参数量和计算量略有上升,但仍符合工业生产要求,证明该模块提升了模型对于织物小目标缺陷的定位精度;实验4为仅使用GC-MANet改进颈部C3k2,较原模型R提升了5.26%,mAP50提升了3%,证明该模块提升了模型的多尺度特征融合能力和全局上下文感知能力;实验5为使用RepStem改进网络下采样,较原模型R提升了5.38%,mAP50提升了2.66%,证明该模块扩大了模型的参数空间,提高了检测精度;实验6为

融合主干网络的两个改进,相较于实验2,mAP50提升了0.32%,相较于基础模型,mAP50提升了1.88%,FPS提升了3.67 fps,计算量仅上升了0.2 GFLOPs,表现出良好的检测性能;实验7为同时引入3个不同的模块,相较于实验6,R提升了5.13%,mAP50提升了2.07%,参数量下降了 0.08×10^6 ,FPS提升了4.4 fps,相比于基础模型mAP50提升了3.99%,参数量下降了 0.08×10^6 ,FPS上升了8.07 fps;实验8为同时引入4个改进模块,相较于基础模型,R提升了7.53%,mAP50提升了4.2%,参数量和计算量略有上升,但仍满足生产实际对于模型部署的要求,证明了改进方法对于提升模型检测性能的有效性。

为更清晰地展示引入各模块后对布匹织物各缺陷类别精度的影响,对比了引入各模块后不同类别缺陷的精度变化,实验结果如表7所示,其中A为使用DRB改进主干C3k2;B为使用C2TDME替换主干C2PSA;C为使用GC-

表 7 各缺陷类别精度对比

Table 7 Comparison of accuracy per defect category

模型	AP-DaiSha	AP-DuanSha	AP-MianQiu	AP-PoDong	AP-TuoSha	AP-WuZi	mAP50
YOLOv11n	0.894	0.965	0.928	0.849	0.943	0.837	0.903
+A	0.966	0.975	0.897	0.831	0.963	0.871	0.917
+B	0.948	0.970	0.912	0.911	0.954	0.850	0.924
+C	0.942	0.976	0.923	0.947	0.952	0.839	0.930
+D	0.963	0.978	0.917	0.878	0.95	0.873	0.927
+A+B	0.958	0.971	0.921	0.869	0.962	0.838	0.920
+A+B+C	0.977	0.962	0.919	0.907	0.976	0.893	0.939
+A+B+C+D	0.962	0.983	0.937	0.908	0.973	0.885	0.941

MANet 改进颈部 C3k2;D 为使用 RepStem 改进网络下采样。

根据表 7 数据可知,使用 DRB 改进主干 C3k2 后,相较于基础模型,带纱、断纱、脱纱和污渍缺陷的检测精度均有提升,证明该模块对于线性缺陷的检测效果提升明显;使用 C2TDME 替换主干 C2PSA 后,较于基础模型,绝大部分类别的检测精度得到了提升,证明使用该模块改进 C2PSA 的 PSABlock 后,大大提升了模型对于各类别缺陷的定位精度;使用 GC-MANet 改进颈部 C3k2 后,相较于基础模型,仅棉球缺陷的检测精度略有下降,且破洞缺陷的检测精度相较于其他几个模型最高,AP 达到了 0.947;使用 RepStem 改进网络下采样后,较于基础模型,

大部分缺陷类别的检测精度均有明显提升;融合主干两模块后,相较于基础模型,各类缺陷的检测精度大都有所提升;融合 3 模块后,带纱缺陷的检测精度较于其他模型最高,AP 达到了 0.977,相较于两模块融合,带纱、破洞、脱纱和污渍缺陷的检测精度均有提升,证明了模型融合后的有效性;同时引入 4 个改进模块后,相较于基础模型,织物缺陷所有类别的检测精度均有明显提升,大部分类别的检测精度达到了峰值,证明了 4 模块融合后的优异检测性能。

为更直观地展示本文改进算法地有效性与先进性,选取布匹织物缺陷数据集中图像分别在改进前后的 YOLOv11n 模型上进行检测,检测结果如图 10 所示。

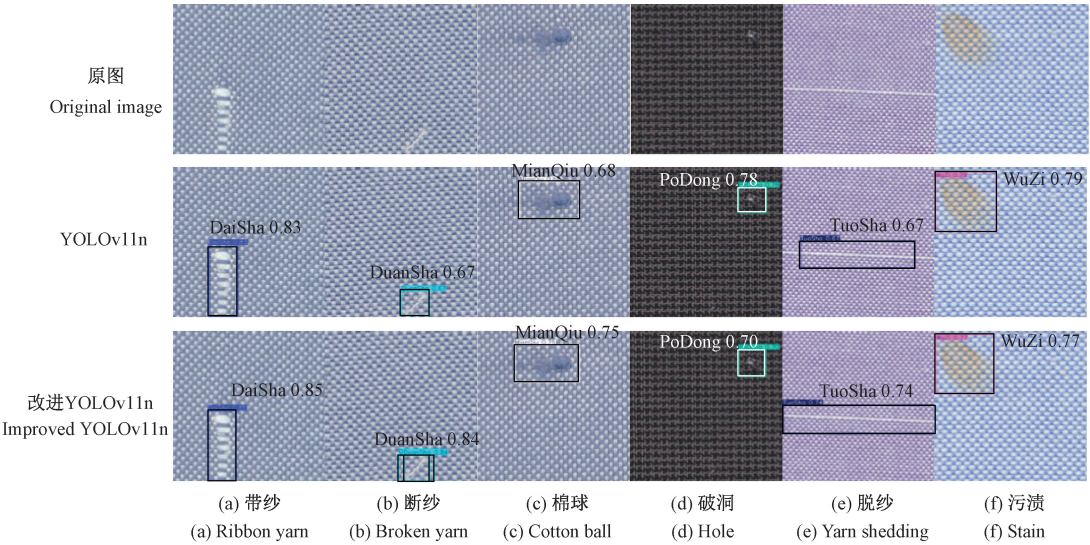


图 10 改进前后算法检测对比

Fig. 10 Before-and-after comparison of algorithm detection

对比算法改进前后的可视化结果,改进后的 YOLOv11n 在带纱、断纱、棉球、脱纱等缺陷上的检测精度明显优于改进前算法,如图 10(a)~(e)所示。相较于改进前算法,改进后算法提升了重叠缺陷的检测精度;对于布匹织物中的带纱、断纱和脱纱等受背景信息干扰的缺陷类别的检测效果增强;提升了破洞这一小目标缺陷的检测精

度,如图 10(d)所示;改进后模型提升了不同尺度大小的棉球、污渍等块状缺陷的检测精度,如图 10(c)、(f)所示,证明改进后算法对于提升布匹织物缺陷检测性能的有效性。

3.9 对比实验

为进一步验证本文改进算法对于布匹织物缺陷检测的优越性,将本文算法与主流目标检测算法在相同条件下

进行对比实验,实验结果如表 8 所示。根据表 8 数据可知,本文所改进算法对于 YOLOv11n 模型的检测精度提升明显,相比于其他几种主流目标检测算法,本文改进算法的 mAP50 分别提升了 11.09%、6.69%、1.29%、2.73%、40.45%、1.84%、7.05%、3.75%、4.21%、12.02%、3.29%,且计算量和参数量仅略高于 YOLOv11n、YOLOv12 和 YOLOv13,表现出本文改进算法的优越性能。

表 8 对比实验
Table 8 Comparative experiment

算法	GFLOPs	Params/($\times 10^6$)	mAP50
Faster_RCNN	206.69	41.15	0.847
Cascade_RCNN	234.48	68.64	0.882
YOLOv3	14.50	116.2	0.929
YOLOv5n	7.00	15.5	0.916
YOLOv6	11.60	4.16	0.670
YOLOv8n	6.90	3.05	0.924
YOLOv9-tiny	6.70	1.76	0.879
YOLOv10n	8.20	2.75	0.907
YOLOv11n	6.30	2.58	0.903
YOLOv12	5.80	2.51	0.840
YOLOv13	6.10	2.45	0.911
本文	6.50	2.65	0.941

综上所述,通过将本文改进算法与主流目标检测算法在布匹织物缺陷数据集下进行实验对比,证明本文算法在织物缺陷检测方面的有效性和先进性,改进后算法相较于原 YOLOv11n 算法虽有增加,但仍满足低算力设备的部署要求,实现了检测精度和计算开销的平衡,综合性能优于其他算法,满足纺织工业生产对于织物缺陷检测的实时性和高精度需求。

4 结 论

为解决布匹织物缺陷检测中存在的特征提取困难造成的漏检、个别类别缺陷间差异不明显所造成的误检等问题,本研究提出了基于 YOLOv11n 的改进织物缺陷检测算法。在网络主干 C3k2 中引入膨胀卷积参数化模块 DRB,提高模型的多尺度特征表达能力,提升了模型特征提取能力,降低漏检率;构建 C2TDME 模块替代 YOLOv11n 主干的 C2PSA,提升模型对于小目标的检测能力和全局上下文建模能力,解决个别类别缺陷差异较小引起的误检问题;构建 GC-MANet 模块替换 YOLOv11n 颈部网络的 C3k2,增强模型捕捉缺陷形态和纹理特征能力,解决模型因背景与缺陷特征混淆所引起的检测精度低问题;采用 RepStem 改进模型的下采样,扩大模型感受野,增加特征提取路径,提高模型检测精度。实验结果表明,改进后算法在布匹织

物缺陷数据集上相较于 YOLOv11n 的平均检测精度提升了 4.21%,证明了模型性能的优越性,满足纺织工业生产对于织物缺陷检测要求。

参考文献

[1] 任梦凡,朱磊,马晓敏,等. 基于粗糙度测量和颜色距离的织物缺陷检测方法[J]. 激光与光电子学进展, 2021, 58(4): 155-162.
Ren Mengfan, Zhu Lei, Ma Xiaomin, et al. Fabric defect detection method based on roughness measurement and color distance [J]. Laser & Optoelectronics Progress, 2021, 58(4): 155-162.

[2] 刘秀平,冯奇,袁皓,等. LBP 与低秩分解的网纹织物纹理缺陷检测方法[J]. 电子测量与仪器学报, 2021, 35(1): 135-141.
Liu Xiuping, Feng Qi, Yuan Hao, et al. Defect detection method of mesh fabric texture based on LBP and low-rank decomposition[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(1): 135-141.

[3] Kim J C, Jong H, Ryu J M, et al. Automatic fabric defect detection using optimal Gabor filter based on hybrid beetle antennae search-gravitational search algorithm[J]. Journal of Optics, 2023, 52(4): 1667-1675.

[4] Kang J H, Jwong K M, Kim H J, et al. Fabric defects detection for multicolor yarn shoe upper using morphological operations[J]. International Journal of Precision Engineering and Manufacturing, 2025, 26(4): 1449-1456.

[5] 杨洋,吴一全. 视觉 Transformer 在表面缺陷检测中的应用:研究进展与挑战[J]. 仪器仪表学报, 2025, 46(12): 1-22.
Yang Yang, Wu Yiquan. Application of vision Transformer in surface defect detection: Research progress and challenges [J]. Chinese Journal of Scientific Instrument, 2025, 46(12): 1-22.

[6] Maity M, Banerjee S, Sinha Chaudhuri S. Faster R-CNN and YOLO based vehicle detection: A survey[C]//2021 5th International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2021: 1442-1447.

[7] Cai Z W, Vasconcelos N. Cascade R-CNN: High quality object detection and instance segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(5): 1483-1498.

[8] 曾秀云,陆华才,吕禾丰. 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.

- Zeng Xiuyun, Lu Huacai, Lyu Hefeng. Research on cotton packaging defect detection method based on improved Faster R-CNN[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(4): 179-186.
- [9] 莫惠文, 杨乐琪, 曹锦纲. 基于 UNet++ 自编码器的织物缺陷检测[J]. 棉纺织技术, 2023, 51(11): 22-29.
- Mo Huiwen, Yang Leqi, Cao Jingang. Fabric defect detection based on UNet++ autoencoder[J]. Cotton Textile Technology, 2023, 51(11): 22-29.
- [10] 朱磊, 王倩倩, 姚丽娜, 等. 改进 YOLOv5 的织物缺陷检测方法[J]. 计算机工程与应用, 2024, 60(20): 302-311.
- Zhu Lei, Wang Qianqian, Yao Lina, et al. Improved YOLOv5 fabric defect detection method[J]. Computer Engineering and Applications, 2024, 60(20): 302-311.
- [11] 陈森, 张胜利, 季坚堯. 基于改进 YOLOv5 的织物缺陷检测[J]. 毛纺科技, 2024, 52(1): 99-106.
- Chen Miao, Zhang Shengli, Ji Jianwan. Fabric defect detection based on improved YOLOv5 [J]. Wool Textile Journal, 2024, 52(1): 99-106.
- [12] Li T, Mei S Q, Shi Y S, et al. Research on fabric defect detection algorithm based on lightweight YOLOv7-Tiny[J]. Journal of Natural Fibers, 2024, 21(1): 1-15.
- [13] 化春键, 李秀琴, 蒋毅, 等. 基于 YOLO-CFD 的棉布微小微弱缺陷检测研究[J]. 电子测量与仪器学报, 2025, 39(4): 152-162.
- Hua Chunjian, Li Xiuqin, Jiang Yi, et al. Research on micro and weak defect detection of cotton fabric based on YOLO-CFD[J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(4): 152-162.
- [14] 白先浪, 张群利, 辛志强. 基于深度学习的皮革缺陷检测方法[J]. 电子测量技术, 2026, 49(3): 175-184.
- Bai Xianlang, Zhang Qunli, Xin Zhiqiang. Leather defect detection method based on deep learning[J]. Electronic Measurement Technology, 2026, 49(3): 175-184.
- [15] Ding X B, Zhang Y Y, Ge Y X, et al. UniRepLKNet: A universal perception large-kernel ConvNet for audio, video, point cloud, time-series and image recognition[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2024: 5513-5524.
- [16] Liu Z, Lin Y T, Cao Y, et al. Swin Transformer: Hierarchical vision Transformer using shifted windows [PP/OL]. V2. (2021-08-17) [2025-11-24]. <https://doi.org/10.48550/arXiv.2103.14030>.
- [17] Wu Z Y, Ding T J, Lu Y F, et al. Token statistics transformer: Linear-time attention via variational rate reduction[PP/OL]. V1. (2024-12-23) [2025-11-24]. <https://doi.org/10.48550/arXiv.2412.17810>.
- [18] Maaz M, Shaker A, Cholakkal H, et al. EdgeNeXt: Efficiently amalgamated CNN-transformer architecture for mobile vision applications[PP/OL]. V3. (2022-10-22) [2025-11-24]. <https://doi.org/10.48550/arXiv.2206.10589>.
- [19] Yin D S, Hu L Y, Li B, et al. 5% > 100%: Breaking performance shackles of full fine-tuning on visual recognition tasks[PP/OL]. V2. (2024-08-27) [2025-11-24]. <https://doi.org/10.48550/arXiv.2408.08345>.
- [20] Zhu J C, Chen X L, He K M, et al. Transformers without normalization [C]//2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2025: 14901-14911.
- [21] Yang G Y, Wang Y, Shi D M, et al. Golden cudgel network for real-time semantic segmentation [PP/OL]. V1. (2025-03-05) [2025-11-24]. <https://doi.org/10.48550/arXiv.2503.03325>.
- [22] Feng Y F, Huang J G, Du S Y, et al. Hyper-YOLO: When visual object detection meets hypergraph computation[PP/OL]. V2. (2024-10-16) [2025-11-24]. <https://doi.org/10.48550/arXiv.2408.04804>.
- [23] Vasu P K A, Gabriel J, Zhu J, et al. FastViT: A fast hybrid vision transformer using structural reparameterization[PP/OL]. V2. (2023-08-17) [2025-11-24]. <https://doi.org/10.48550/arXiv.2303.14189>.
- [24] Han K, Wang Y H, Guo J Y, et al. ParameterNet: Parameters are all you need[PP/OL]. V2. (2024-01-14) [2025-11-24]. <https://doi.org/10.48550/arXiv.2306.14525>.
- [25] Li C, Zhou A J, Yao A B. Omni-dimensional dynamic convolution[PP/OL]. V1. (2022-09-16) [2025-11-24]. <https://doi.org/10.48550/arXiv.2209.07947>.
- [26] Song Y D, Zhou Y, Qian H, et al. Rethinking performance gains in image dehazing networks[PP/OL]. V1. (2022-09-23) [2025-11-24]. <https://doi.org/10.48550/arXiv.2209.11448>.
- [27] Chen J R, Kao S H, He H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[PP/OL]. V3. (2023-05-21) [2025-11-24]. <https://doi.org/10.48550/arXiv.2303.03667>.
- [28] Shi D Q. TransNeXt: Robust foveal visual perception for vision transformers [C]//2024 IEEE/CVF Conference on Computer Vision and Pattern

- Recognition(CVPR). IEEE, 2024: 17773-17783.
- [29] Xu Z Z, Wu D Y, Yu C Q, et al. SCTNet: Single-branch CNN with transformer semantic information for real-time segmentation[J]. AAAI Conference on Artificial Intelligence, 2024, 38(6): 6378-6386.
- [30] Noman M, Fiaz M, Cholakkal H, et al. ELGC-Net: Efficient local-global context aggregation for remote sensing change detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-11.
- [31] Tan M X, Le Q. EfficientNet: Rethinking model scaling for convolutional neural networks[PP/OL]. V5. (2020-09-11)[2025-11-24]. <https://doi.org/10.48550/arXiv.1905.11946>.
- [32] Li K, Wang D, Hu Z Y, et al. Unleashing channel potential: space-frequency selection convolution for SAR object detection [C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2024: 17323-17332.
- [33] Wu T Y, Tang S, Zhang R, et al. CGNet: A light-weight context guided network for semantic segmentation [J]. IEEE Transactions on Image Processing, 2021, 30: 1169-1179.
- [34] Wang C Y, Bochkovskiy A, Liao Hong-yuan Mark. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2023: 7464-7475.
- [35] Lu W, Chen S B, Tang J, et al. A robust feature downsampling module for remote-sensing visual tasks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-12.

作者简介

王海群, 硕士, 主要研究方向为智能控制与应用、深度学习。

E-mail: wanghq0604@163.com

陈珊珊(通信作者), 硕士研究生, 主要研究方向为图像处理、目标检测。

E-mail: 1400283011@qq.com