

DOI: 10.13382/j.jemi.B2508702

FDRT-DETR: 一种多类别布匹瑕疵检测算法*

金 民¹ 汪 军¹ 徐叶松¹ 刘文学²

(1. 安徽工程大学计算机与信息学院 芜湖 241000; 2. 安徽华烨特种材料有限公司 芜湖 241000)

摘 要:针对布匹瑕疵检测中多类别瑕疵形态差异大、微小瑕疵目标密集分布以及实时高精度检测需求等挑战,提出了 FDRT-DETR: 一种多类别布匹瑕疵检测算法。首先,在主干网络中使用反向残差移动块(inverted residual mobile block, iRMB)模块,强化多尺度特征提取能力并降低计算冗余。其次,引用了基于统计学特征的注意力机制(token statistics self attention, TSSA)来增强模型对瑕疵区域特征的捕捉精度,降低背景纹理的干扰。再者,在颈部网络中设计了并行空洞卷积注意力金字塔网络(parallel atrous convolution attention pyramid network, PACAPN),显著增强了对小目标特征的保留和识别能力。最后,针对微小瑕疵密集分布导致的低匹配质量问题,引入了可匹配性感知损失函数(matchability aware loss, MAL),提高检测性能。在阿里天池的布匹瑕疵数据集上的实验结果表明,提出的检测算法将布匹瑕疵检测的 mAP@0.5 提高了 3.7%,检测速度达到了 64.4 fps,参数量仅为 31.4×10^6 ,可以满足工业生产的实际需要。

关键词: 布匹瑕疵检测; RT-DETR; 注意力机制; 金字塔网络; 损失函数

中图分类号: TP391; TN911

文献标识码: A

国家标准学科分类代码: 520.60

FDRT-DETR: An multi-category fabric defect detection algorithm

Jin Min¹ Wang Jun¹ Xu Yesong¹ Liu Wenxue²

(1. College of Computing and Information Technology, Anhui Polytechnic University, Wuhu 241000, China;

2. Anhui Huaye Advanced Materials Co., Ltd., Wuhu 241000, China)

Abstract: Aiming at the challenges of large differences in the morphology of multi-category defects in fabric defect detection, the dense distribution of tiny defective targets, and the demand for real-time high-precision detection, FDRT-DETR: a multi-category cloth defect detection algorithm is proposed. Firstly, the inverted residual mobile block (iRMB) module is applied in the backbone network to strengthen multi-scale feature extraction and reduce computational redundancy. Secondly, the token statistics self attention (TSSA) mechanism is invoked to enhance the model's accuracy in capturing features in defective regions and reduce the interference of background texture. Further, the parallel atrous convolution attention pyramid network (PACAPN) is designed in the neck network to significantly improve the preservation and recognition of small target features. Lastly, for the problem of low matching quality due to the dense distribution of tiny imperfections, a matchability aware loss function (MAL) is introduced to improve the detection performance. Experimental results on Alibaba Tianchi's fabric defect dataset show that the improved model increases the mAP@0.5 for fabric defect detection by 3.7%, achieving a speed of 64.4 frames/s with only 31.4×10^6 parameters, thus satisfying the practical needs of industrial production.

Keywords: fabric defect detection; RT-DETR; attention mechanism; pyramid network; loss function

0 引言

纺织工业作为全球制造业体系的重要支柱,其产品广泛应用于衣着、家居及工业领域。然而,生产过程中由机械故障、纱线缺陷及人为操作引起的布匹瑕疵,已成为各国纺织行业亟待解决的共性难题^[1]。当前纺织品缺陷类型已超过80种,其不规则形态与尺寸差异显著增加了检测难度。传统人工检测存在效率低、主观性强及视觉疲劳导致的漏检率攀升等缺陷,面对现代工业对生产效率和产品质量的严苛需求,发展高精度、自适应的自动化检测技术已成为推动纺织业智能化升级的关键路径^[2-3]。

目前在布匹瑕疵检测领域,深度学习方法主要分为3类,基于候选区域的两阶段检测法、基于回归的一阶段检测法以及新兴的基于Transformer的检测技术。两阶段模型,如Faster R-CNN^[4]、Mask R-CNN^[5]和H2RBox-v2^[6],虽在精度上占优,但计算开销较大。而一阶段模型,如SSD^[7]、RTMDet^[8]和YOLO^[9-12]系列,凭借低资源需求和高实时性,成为研究主流方向。文献[13]提出一种改进的YOLOv5算法,通过嵌入坐标注意力模块、采用Mish激活函数以及引入平滑交并比(smoothed intersection over union, SloU)损失函数,有效解决了多类别瑕疵和样本不均衡的问题,显著提升了检测精准度。文献[14]提出一种名为Fab-ME的织物缺陷检测框架,基于YOLOv8s并引入C2F-VMamba模块和增强型多尺度通道注意模块,分别增强全局上下文捕捉和小目标检测能力,并有效平衡了准确性和实时性。文献[15]基于YOLOv10,创新性地构建了一个包含织物刚度信息的专用数据集,并运用数据增强技术提升模型泛化能力,有效解决了纺织行业中织物类型检测和撕裂路径识别问题。

上述基于YOLO系列的织物缺陷检测算法在检测精度和速度有了一定的进展,但仍存在因非极大值抑制(non-maximum suppression, MMS)而增加模型复杂度和小目标瑕疵误检漏检等问题。而以DETR^[16]、RT-DETR^[17]和DEIM^[18]代表的基于Transformer的目标检测算法具有很大的发展前景,其主要特点是实现了端到端的目标检测,不需要预先定义锚框和非极大值抑制的后处理策略,使得目标检测在保持检测精度的同时变得更加简单高效。文献[19]提出改进RT-DETR的织物缺陷检测算法,通过引入轻量级FasterNet-T0、可变形注意力机制和动态上采样算子,有效减少了模型参数数量和计算量,提升了检测速度和准确性。文献^[20]基于RT-DETR,通过部分卷积、高效多尺度注意力机制增强特征识别,同时优化多尺度特征融合模块和损失函数,进一步提升了对牛仔面料瑕疵的定位准确度与检测精度。

尽管现有研究在纺织检测领域取得了一定成果,但在多类别布匹瑕疵检测中,难以兼顾速度与精度。因此文中通过对RT-DETR的改进,提出一种高效的多类别布匹瑕疵检测算法。

本文引入了一种新的轻量化主干网络反向残差移动块(inverted residual mobile block, iRMB)^[21],在降低模型计算量和参数数量的同时,有效地扩大了感受野,加强了特征提取能力和特征融合效率;使用统计学特征的注意力机制(token statistics self attention, TSSA)模块^[22]代替由尺度特征交互(attention based intra-scale feature interaction, AIFI)模块,以增强模型对瑕疵区域特征的捕捉精度,降低背景纹理的干扰;创新性地设计了并行空洞卷积注意力金字塔网络(parallel atrous convolution attention pyramid network, PACAPN)模块对颈部网络进行优化,加强对小目标瑕疵特征的处理,提升了模型对微小尺寸瑕疵的检测能力;引入可匹配性感知损失函数(matchability aware loss, MAL),能够有效处理微小瑕疵密集分布导致的低质量匹配问题,提高了低质量匹配下的检测精度。

1 RT-DETR 模型结构

RT-DETR是由百度提出的一种基于Transformer的实时端到端检测模型,由主干网络、颈部网络和头部网络组成。主干网络采用经典的残差网络(residual network, ResNet)架构,通过该网络从输入图像中提取其特征表示;颈部网络采用AIFI模块和基于卷积神经网络(convolutional neural network, CNN)的跨尺度特征融合模块(cnn based cross-scale feature fusion module, CCFM)组成的混合编码器,它通过尺度内交互和跨尺度融合两步操作,将多尺度特征转换为图像特征序列并捕获全局上下文信息,提升模型整体的检测性能;头部网络是由交并比(IoU)感知查询选择机制和检测头组成的变换解码器构成,它能够提供更高质量的初始查询,避免了非极大值抑制的过程,加快了模型实时检测速度。RT-DETR算法模型结构如图1所示。

2 改进的FDRT-DETR 模型结构

通过对RT-DETR网络知识的深入研究,并结合布匹瑕疵的特征,创新性地设计出一种专门用于多类别布匹瑕疵检测的FDRT-DETR网络模型,该改进后的模型结构如图2所示。首先,在主干网络的P2、P3、P4以及P5层,巧妙引入轻量级iRMB模块,使其能够高效地对瑕疵图像进行特征提取,为后续的检测工作奠定坚实基础;其次,采用TSSA模块,精准地捕捉特征间的上下文交互信

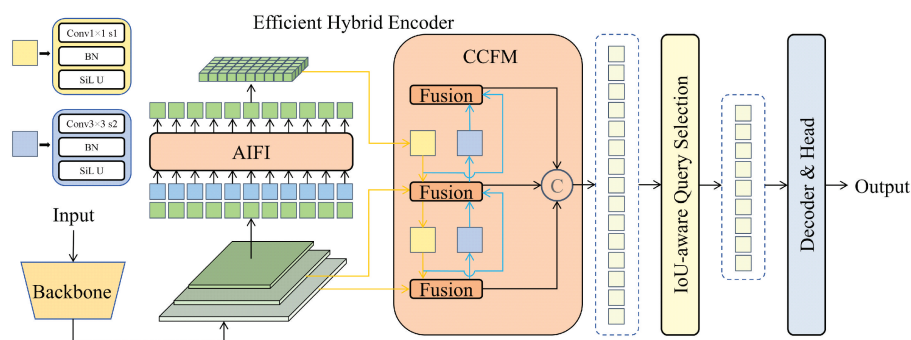


图 1 RT-DETR 模型结构
Fig. 1 RT-DETR model structure

息,从而实现对瑕疵关键区域信息的准确筛选与选择,有效提升检测的针对性;然后,在颈部网络部分,创新性地设计了 PACAPN 多尺度特征融合模块,能够将不同尺度的瑕疵特征进行有机融合,并确保关键特征得以有效保

留,进而显著增强了模型对于微小瑕疵的检测能力;最后,采用 MAI 感知损失函数,针对低质量匹配情况下的边界框回归精准度进行优化,并通过 Decoder 的多层次优化查询,最终精准地输出布匹瑕疵检测的结果。

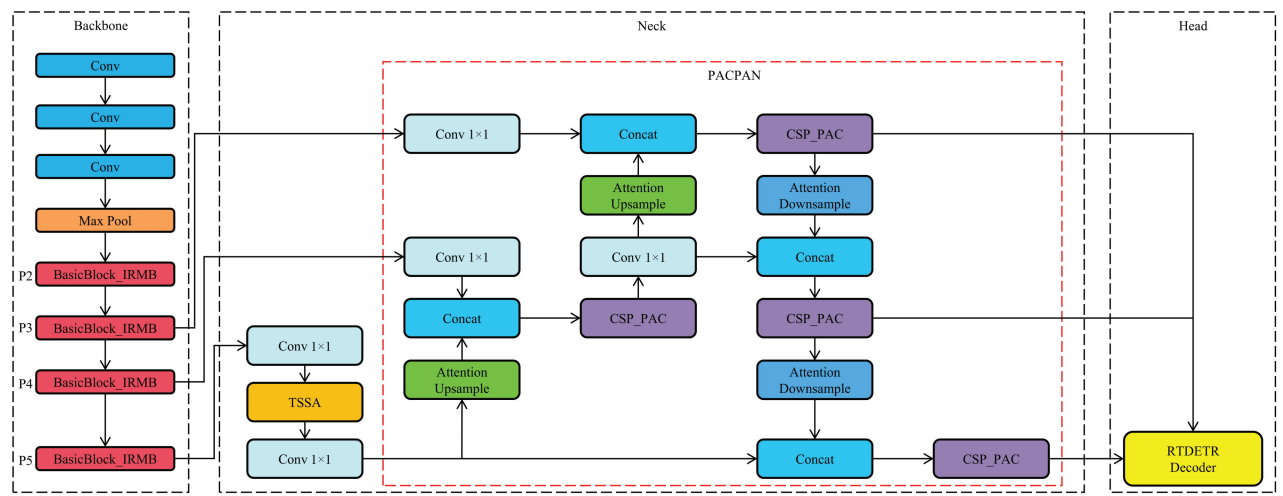


图 2 FDRT-DETR 模型结构
Fig. 2 FDRT-DETR model structure

2.1 iRMB 模块

布匹瑕疵检测模型的主干网络中引入了 iRMB 模块,其结合了卷积神经网络和 Transformer 的轻量化模型设计,旨在平衡参数效率、计算效率和模型性能,其模块结构如图 3 所示。

iRMB 通过其独特的设计提升了特征提取能力和特征融合效率。首先,输入特征图通过 1×1 卷积层进行通道扩展,增加通道数量,使模型能够捕获更丰富的特征信息。接着,扩展后的特征图经过 3×3 深度可分离卷积层处理,该卷积层在降低计算复杂度的同时,能够有效地提取局部特征。与此同时,注意力机制被应用于输入特征图,生成注意力权重矩阵,该矩阵能够突出瑕疵重要特征并抑制冗余特征。然后,注意力权重矩

阵与深度可分离卷积处理后的特征图进行逐元素相乘操作,实现特征的选择性增强。最后,增强后的特征图通过另一个 1×1 卷积层进行通道压缩,减少通道数量,从而降低计算负担。整个 iRMB 模块通过残差连接将输入直接与输出相加,形成了“倒置残差”结构,使得模型在保持高效特征提取的同时,也提高了特征融合的效率。

iRMB 以结合深度可分离卷积 (depthwise convolution, DW-Conv) 与改进的自注意力机制 (expanded window multi head self attention, EW-MHSA) 实现轻量化设计,通过残差结构增强特征传递,减少信息损失,采用逐渐增加的扩展率和通道数,适应不同计算需求。首先,输入特征图 $X \in R^{AC \times H \times W}$, 通过过扩展率 λ 控制的通道扩

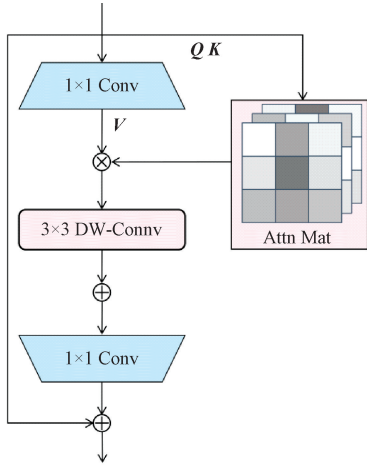


图3 iRMB 模块结构

Fig. 3 iRMB module structure

展操作,得到扩展后的特征图 $X_e \in \mathbf{R}^{AC \times H \times W}$;接着,依次通过改进的自注意力和深度可分离卷积,其中 EW-MHSA 公式为:

$$EW-MHSA(X) = Softmax\left(\frac{QK^T}{\sqrt{d}}\right) V \quad (1)$$

式中: Q 是查询矩阵; K 是键矩阵; V 是值矩阵; d 是特征维度,且 $Q = K = X, V = X_e$ 。

从而增强特征图 X_f ;然后通过收缩操作得 $X_s \in \mathbf{R}^{C \times H \times W}$;最后通过残差连接将输入特征图 X_f 与 X_s 结合,得到最终输出 $Y = X + X_s$ 。

iRMB 模块的设计优势在于其高效性、简洁性和灵活性,通过轻量化算子大幅降低计算复杂度,仅依赖标准卷积和自注意力,无需额外复杂模块,并可通过调整扩展率 λ 和通道数适应不同模型规模,从而在 RT-DETR 主干网络中实现轻量化与性能提升的平衡。

2.2 TSSA 模块

TSSA 模块是基于统计学线性复杂度 Transformer 架构(token statistics transformer, TOST)中的核心创新,采用基于统计自注意力机制(TSSA)替代 RT-DETR 中原有的基于注意力的特征交互机制,在保持竞争力的性能表现同时实现线性计算复杂度。TSSA 模块结构如图4所示。

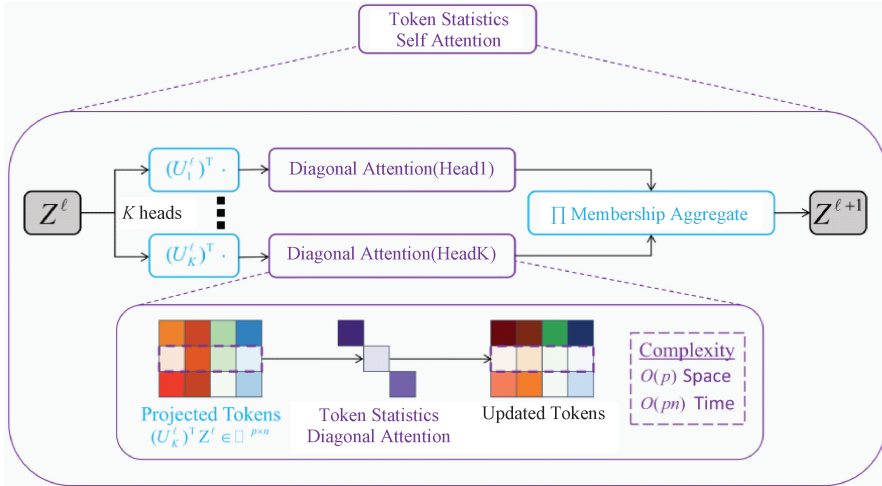


图4 TSSA 模块结构

Fig. 4 TSSA module structure

在传统的检测模型中,自注意力机制依赖于对输入 token 两两相似性的计算,这一过程虽然有效,但其资源开销显著,尤其当输入 token 数量极大时,传统的注意力机制(如 Transformer 中的全局注意力等)在计算复杂度和内存使用上的瓶颈问题愈发显著。而我们引入的基于统计学特征的注意力机制,它通过避免两两相似性的计算,仅依赖于 token 特征的统计量,显著降低了模型的计算复杂度,TSSA 通过对投影后的 token 进行行标量化变换,即降低分率实现线性时间注意力。TSSA 在 CRATE 的理论基础上,通过几何空间的结构化特征实现了 token 分组和映射。

TSSA 模块包含 3 个核心组件,令牌投影层、TSSA 注意力层以及特征重建层。给定输入特征图 $X \in \mathbf{R}^{d \times n}$ (包含 n 个 d 维令牌),TSSA 模块的处理流程如下。

首先经过令牌投影层,将输入特征映射到低维空间,其投影层公式如下:

$$Z = W_p X, W_p \in \mathbf{R}^{p \times d} \quad (2)$$

其中, $p \ll d$ 为投影后的令牌维度。

再经过 TSSA 令牌统计自注意力层,通过令牌统计量实现高效注意力计算,其 TSSA 算子定义为:

$$TSSA(Z | \{U_k\}_{k=1}^K) =$$

$$-\frac{\tau}{n} \sum_{k=1}^K U_k D(Z, \pi_k | U_k) U_k^T Z \text{Diag}(\pi_k) \quad (3)$$

$$\pi_k = \text{softmax}\left(\frac{1}{2\eta} \|U_k^T z_j\|_2^2\right) \quad (4)$$

$$D_{ii} = \nabla f\left(\frac{1}{n_k} (U_k^T Z \text{Diag}(\pi_k) Z^T U_k)_{ii}\right) \quad (5)$$

$$\nabla f(x) = (1 + x)^{-1} \quad (6)$$

式中: $U_k \in \mathbf{R}^{p \times p}$ 为每个注意力 k 的可学习正交投影矩阵; π_k 表示第 k 个头的分组隶属概率; D_{ii} 是 $D(Z, \pi_k | U_k)$ 对角矩阵的元素。

再经过特征重建层,将处理后的特征映射回原始维度,其公式为:

$$Y = W_r TSSA(Z) + b_r, W_r \in \mathbf{R}^{d \times p} \quad (7)$$

TSSA 的核心优点是通过特定的概率分布函数对输入序列进行建模,减少冗余信息并提取关键特征。

1) 统计特征提取,对序列中的每个 token 提取其统计特征。

2) 变分编码率缩减,利用变分编码率缩减框架对特征进行压缩,减少信息冗余。

3) 线性复杂度实现,通过一系列优化,其计算复杂度从 $O(n^2)$ 降低为 $O(n)$ 。

TSSA 模块在令牌数量 n 上实现线性时间复杂度 $O(pn)$ 和线性空间复杂度 $O(p)$,较标准自注意力机制显著提升效率,该设计使 RT-DETR 能够高效处理高分辨率特征图,同时保持模型捕捉全局依赖关系的能力。TSSA 算子基于统计特性的特征压缩机制,在提升计算效率的同时也增强了模型的可解释性。

2.3 PACAPN 模块

在布匹瑕疵检测中,微小瑕疵的占比突出,其检测难度高,为此,在颈部网络的 CCFM 结构中创新性地设计了一种 PACAPN 模块,通过结合跨阶段部分并行空洞卷积(cross stage partial-parallel atrous convolution, CSP_PAC)模块和并行上下采样分支模块,显著提升了对小目标特征的保留和识别能力。CSP_PAC 模块利用不同空洞率的卷积捕捉多尺度空间上下文信息,扩大感受野以提取不同尺度的特征,从而提高目标定位精度。并行上下采样分支模块则通过强化上采样中的特征表达和下采样中的细节保留,增强了对微小瑕疵的检测能力。PACAPN 的优势在于其多尺度特征捕获能力,能够同时关注大范围纹理特征和局部微小瑕疵;增强的小目标检测能力有效降低了小瑕疵被忽略的风险;高效的特征融合机制通过强化有意义特征和抑制冗余特征,提高了特征表达的有效性。此外,PACAPN 适应性强,能有效识别不同尺度和类型的瑕疵,同时保持了较高的计算效率,满足了工业生产中对检测速度的要求。

1) CSP_PAC 模块

RT-DETR 采用重参化卷积(reparameterization convolution, RepC3)结构,仅简单地将两分支特征元素相加,可能导致信息损失,削弱对小尺寸目标的特征提取与识别能力。鉴于跨阶段部分网络(cross stage partial, CSP)能通过分解与加载特征优化计算资源利用,实现检测速度提升,本文受此启发,结合其优势提出 CSP_PAC 模块,利用不同空洞率的卷积来扩展感受野,捕获多尺度信息并增强特征表达能力,其模块的结构如图 5 所示。

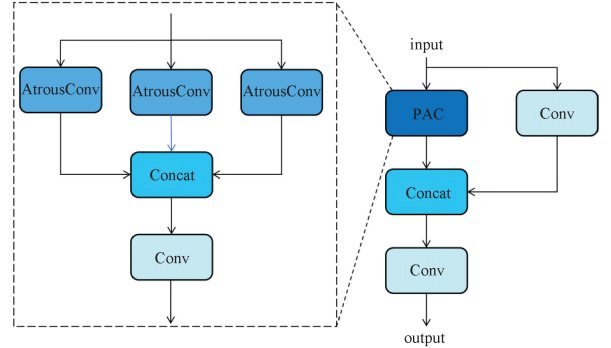


图 5 CSP_PAC 模块结构

Fig. 5 CSP_PAC module structure

CSP_PAC 模块构建出的平行空洞卷积结构,首先采用 3 个不同空洞率($R=0,1,2$)的 3×3 卷积捕捉不同尺度的空间上下文信息,经 1×1 卷积根据通道数量合并 3 个空洞卷积的输出。再利用一个 1×1 卷积调整输入特征通道数量,与经变化空洞率卷积的特征在通道维度结合,最终经单一 1×1 卷积生成更具表现力的特征图。通过使用具有不同膨胀率的并行空洞卷积扩大了感受野,有效地提取不同尺度的特征,这使得网络能够捕捉数据中局部和上下文信息,提高了模型目标定位精度。CSP_PAC 模块基于 CSP 瓶颈结构设计,在保证检测性能的同时,降低计算量与参数量,有效提升小目标检测性能,兼顾模型整体效率、准确性,聚焦特征表达与定位精度,为目标检测提供新方案。

2) 并行上下采样分支模块

RT-DETR 采用传统的最近领域插值上下采样方法来提取特征信息,然而在处理布匹瑕疵等小尺寸目标检测时,存在明显不足,容易在采样过程中丢失图像细节和纹理,影响检测精度和特征表达。因此,为了进一步加强对小目标的检测能力,提出一种新的并行上下采样分支模块,其模块的结构如图 6 所示。

在并行上采样分支中,针对上层输入特征进行全局平均池化操作,以提取整个特征图的全局上下文信息。然后,借助通道门控机制利用全局平均池化的结果,经由 1×1 卷积和 Hardsigmoid 激活函数生成,通道注意力图,

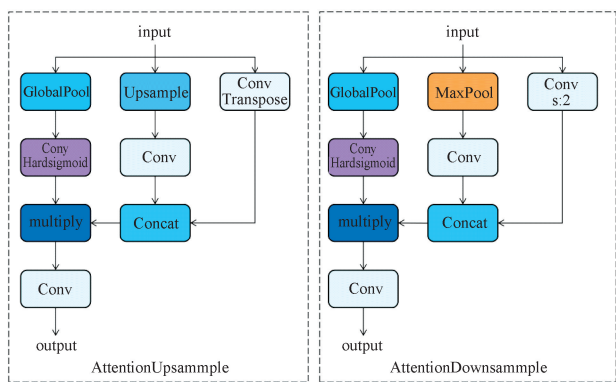


图6 并行上下采样分支模块结构

Fig. 6 Up and down sampling branch module structure

该图表征了每个通道的权重,这一过程有助于突出关键特征,同时抑制不相关的背景噪声,使上采样过程更具适应性。此外,上采样分支还结合了转置卷积以及通过平行路径进行特征图放大的最近邻插值,将转置卷积集成到网络中的优势在于它们的能在训练过程中通过反向传播调整参数,提高了上采样后的特征图质量,获得更丰富的细节。

并行下采样分支的主要结构与并行上采样分支相似,但它的区别在于其使用了两个不同的下采样分支。第1个分支使用最大池化层和 1×1 卷积,进一步降低空间分辨率,同时保留边缘的显著特征。第2个分支则在借助步长为2的 3×3 卷积,在空间维度上实现了下采样,以降低特征图的空间分辨率。最终,加权后的特征经过另一次卷积运算整合,生成高分辨率的特征图。这种方法提高了小目标检测的检测精度,并在一定程度上保留了上下文信息,从而提高了布匹瑕疵检测性能。并行上下采样分支为网络提供了多条特征提取路径,结合门控机制对采样后的特征进行特征选择,强化了有意义的特征,抑制了冗余和不相关特征,提升了特征表达的有效性。

2.4 可匹配性感知损失

在目标检测模型的训练过程中,损失函数作为关键优化指标,它通过计算预测值与真实值之间的误差来引导模型向收敛方向优化,对检测模型性能起到至关重要的作用。RT-DETR采用可变焦损失(varifocal loss, VFL)来降低分类置信度与框质量之间的不确定性,从而提升实时性能。然而,VFL主要针对传统检测器设计,传统检测器低质量匹配较少,且侧重于高IoU优化,导致低IoU匹配由于损失值较小且平坦而未能得到充分优化。并且,在布匹瑕疵检测任务中,像整经结、破洞、结头等微小瑕疵常密集出现,随机初始化的查询难以与微小瑕疵的空间分布对齐,致使匹配框的IoU普遍偏低,预测框高度重叠。重叠的预测框可能具有相似的特征,模型难以区

分不同目标,从而产生低质量匹配。为了解决这一问题,引入了MAL,与VFL相比,MAL在处理低质量水平的匹配时更加有效,其公式如式(8)所示。

$$MAL = \begin{cases} -q^\gamma \log(p) - (1 - q^\gamma) \log(1 - p), & y = 1 \\ -p^\gamma \log(1 - p), & y = 0 \end{cases} \quad (8)$$

式中: q 是模型预测的真实类别概率; p 是预测框与真实框的交并比; y 是真实标签; γ 是调节因子,控制低质量匹配的敏感度。

在前景样本($y=1$)中,第一项鼓励模型对高IoU的匹配(高质量匹配)赋予更高的预测概率, q^γ 将IoU加权,使得高IoU的匹配对损失的贡献更大;第二项对低IoU的匹配(低质量匹配)施加更强的惩罚,通过 $1 - q^\gamma$,低IoU的匹配会显著增加损失,迫使模型优化这些匹配。而在背景样本($y=0$)中,单一项 $-p^\gamma \log(1 - p)$ 与Focal Loss类似,通过 p^γ 降低简单负样本的权重,使模型更关注难分类的负样本。调节因子 γ 控制对低质量匹配的敏感度。较大的 γ 会放大低IoU匹配的损失,从而更强调优化低质量匹配。

MAL通过 q^γ 和 $1 - q^\gamma$ 显式区分不同质量的匹配,显著提升了低IoU匹配的优化效果。同时移除了用于平衡正负样本的超参数,仅通过 γ 调节,使公式更简洁且易于实现。MAL的这些改动有助于避免对高质量框的过度强调,根据匹配性调整惩罚力度,通过动态调整不同质量匹配的损失权重,有效解决了RT-DETR中因稀疏查询导致的低质量匹配问题,从而提升布匹瑕疵的检测性能。

3 实验与分析

3.1 实验环境和参数设置

本文所有实验在均Linux系统下实现,操作系统为Ubuntu,运行环境GPU显卡为NVIDIA GeForce RTX 3090,显存24 GB,基于Python3.8进行编程,所使用的深度学习框架为Pytorch。整个训练过程中,使用AdamW优化器进行优化网络,将初始学习率设置为0.001,批训练样本数设置为16,epoch设置为300。

3.2 数据集介绍

实验的布匹瑕疵检测数据集来源于天池平台,包含5 913张有瑕疵的素色布匹图像,瑕疵类别为纺织行业实际生产中常见的20类瑕疵,具体分为破洞、渍类疵、三丝、结头、花板跳、百脚、毛粒、粗经、松经、断经、吊经、粗维、纬缩、浆斑、整经结、跳类疵、断氨纶、档类疵、痕类疵、其他,每张瑕疵图像大小为 2446×1000 ,图像以8:1:1的比例划分为训练集、验证集和测试集。另外采用阿里天池花色布匹数据集验证改进模型的泛化能力,包含沾

污、错花、水印、花毛、缝头、缝头印、虫粘、破洞、褶子、织疵、漏印、蜡斑、色差、网折和其他这 15 种类型瑕疵,共计 4 371 张图片,图像尺寸为 4 096×1 800,同样以 8 : 1 : 1 的比例划分为训练集、验证集和测试集。

3.3 对比实验

1) 对比不同主干网络

为了验证本文在主干网络所引入的 iRMB 模块的有效性,分别与 Base、FasterNet^[23]、EfficientViT^[24]、RMT^[25]、

MobileNetv4^[26]、MambaOut^[27]不同主干网络进行比较,实验结果如表 1 所示。分析表 1 可知,iRMB 的平均精度均值(mAP) mAP@ 0.5 值为 52.2%,为最佳值,比 Base 提高 1.5%,同时其帧率(FPS)达到 54.0 fps,比其他主干网络改进略高,具备较高的实时性。与其他主干网络相比,iRMB 在保持较低的参数量、复杂度的同时,实现更高的检测精度和更快的检测速度,表明 iRMB 在进行布匹瑕疵检测时,展现出更高的适应性和准确性。

表 1 不同主干对比实验
Table 1 Comparative experiments of different backbones

Backbone	<i>P</i> /%	<i>R</i> /%	mAP@ 0.5/%	mAP@ 0.5;0.95/%	GFLOPs	Params/(×10 ⁶)	帧率/fps
Base	57.8	49.5	50.7	26.0	58.6	38.6	35.2
FasterNet	58.0	49.9	51.0	26.3	54.3	35.8	48.4
EfficientViT	58.3	50.0	51.5	26.7	53.1	35.3	50.2
RMT	57.2	49.2	50.5	25.5	56.2	36.2	48.7
MobileNetv4	58.1	50.1	51.4	26.6	53.8	35.0	51.4
MambaOut	58.5	50.2	51.8	26.9	52.8	34.9	52.9
iRMB	58.8	50.5	52.2	27.4	51.8	34.6	54.0

2) 对比不同颈部网络

为了验证本文在颈部网络中所引入的 PACAPN 模块改进的有效性,分别与 Base、ASF-

YOLO^[28]、SlimNeck^[29]、HS-FPN^[30]、UCTransNet^[31]、WFU^[32]不同颈部网络改进进行比较,实验结果如表 2 所示。

表 2 不同颈部对比实验
Table 2 Comparative experiments of different necks

Neck	<i>P</i> /%	<i>R</i> /%	mAP@ 0.5/%	mAP@ 0.5;0.95/%	GFLOPs	Params/(×10 ⁶)	帧率/fps
Base	57.8	49.5	50.7	26.0	58.6	38.6	35.2
ASF-YOLO	58.5	51.2	52.3	27.4	52.9	34.6	55.8
SlimNeck	58.7	50.6	51.8	27.0	55.2	35.8	48.5
HS-FPN	58.3	50.2	51.6	26.8	56.8	36.0	46.2
UCTransNet	58.5	50.8	51.9	27.1	55.3	35.6	50.5
WFU	58.8	51.1	52.2	27.3	56.5	36.1	51.4
PACAPN	59.0	51.5	52.9	27.6	54.9	35.4	52.8

分析表 2 可知,与 Base 相比,PACAPN 模块在参数量和计算复杂度略低的情况下,mAP@ 0.5 值提高 2.2%,同时帧率达到 52.8 fps。与其他颈部网络相比,PACAPN 在 *P*、*R*、mAP@ 0.5 及 mAP@ 0.5;0.95 等核心精度指标上均取得最优结果;进一步对比精度表现相近的 ASF-YOLO 可知,ASF-YOLO 的参数量为 34.6×10⁶、计算量为 52.9 GFLOPs,与 PACAPN 的 35.4×10⁶、54.9 GFLOPs 处于同一量级,二者参数量差值仅为 0.8×10⁶,计算量差值为 2.0 GFLOPs,模型复杂度差异极小。在此前提下,PACAPN 的 *P*、*R*、mAP@ 0.5 及 mAP@ 0.5;0.95 指标均优于 ASF-YOLO,充分证明 PACAPN 模块在保证轻量化特性的同时,具备更优的特征融合能力,凸显了本文改进方法的有效性。

3) 对比不同瑕疵检测结果

表 3 是基准 RT-DETR 模型与改进的 FDRT-DETR 模型对不同类别瑕疵的检测结果。与 RT-DETR 相比,FDRT-DETR 模型在多数瑕疵类别的 AP@ 0.5 值上均有提升。其中在破洞、三丝、结头等微小瑕疵(表 3 中加粗部分)上的性能提升尤为明显;同时,FDRT-DETR 模型对由细小斑点构成的渍类疵、百脚、档类疵等中大型瑕疵的 AP@ 0.5 值也实现了显著提升。这充分证明了 FDRT-DETR 模型通过引入 PACAPN 模块增强多尺度特征提取,并结合 MAL 优化训练的有效性,显著提升了对微小瑕疵及复合型瑕疵的检测能力。而或许因数据集类别不平衡,部分低频次瑕疵类别 AP@ 0.5 值未提升,但整体仍体现了改进模型在特征提取与检测能力上的增强。

表 3 不同瑕疵检测结果

瑕疵类型	RT-DETR	FDRT-DETR
破洞	54. 6	67. 4
渍类疵	29. 9	36. 1
三丝	66. 1	72. 0
结头	53. 1	61. 9
花板跳	57. 4	58. 8
百脚	81. 5	86. 5
毛粒	29. 5	34. 0
粗经	74. 1	78. 8
松经	54. 2	61. 7
断经	14. 1	20. 0
吊经	22. 7	25. 4
粗维	58. 9	51. 2
维缩	27. 0	18. 5
浆斑	77. 5	77. 6
整经结	61. 7	65. 4
跳类疵	59. 5	69. 5
断氨纶	45. 0	55. 3
档类疵	38. 6	44. 2
痕类疵	53. 4	53. 4
其他	55. 3	49. 9

4) 对比不同模型(SOTA)的实验结果

为了验证本文所提出的布匹瑕疵检测算法的优越性,将改进后的算法与 Faster R-CNN、Mask R-CNN、SSD、YOLOv5、YOLOv8s、YOLOv10s、YOLOv11、DETR、DEIM 及基准 RT-DETR 的算法在相同的数据集上进行对比,实验结果如表 4 所示,与部分主流算法检测效果对比如图 7 所示。

由表 4 可知,Faster R-CNN、Mask R-CNN、SSD 算法与 YOLO 系列算法相比,存在检测精度低、检测速度慢等问题。DETR、DEIM 算法在检测精度上虽略高于 YOLO 系列,但其模型相对复杂,检测速度偏慢。而 FDRT-DETR 算法与 Faster-RCNN、Mask-RCNN、SSD、YOLOv5、YOLOv8s、DETR 以及基线模型 RT-DETR 相比,在多个关键指标上均表现出显著优势,其 P 值和 R 值分别为 62.6%和 52.5%, $mAP@0.5$ 值和 $mAP@0.5:0.95$ 值分别达到了 54.4%和 28.2%,参数量仅为 31.4×10^6 ,计算复杂度仅为 50.6 GFLOPs,检测速度达到 64.4 fps,均优于其他对比算法。在对比分析下,YOLOv10s、YOLOv11 算法虽然在参数量与计算复杂度方面较低,其检测速度分别为 72.2 和 78.6 fps,与 FDRT-DETR 大致处于同一

表 4 不同主流方法检测效果对比实验

模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	GFLOPs	Params/ $(\times10^6)$	帧率/fps
Faster R-CNN	42. 4	34. 6	35. 9	17. 8	220. 2	119. 1	19. 8
Mask R-CNN	47. 1	38. 2	40. 8	19. 5	168. 5	102. 1	20. 9
SSD	38. 6	31. 2	32. 3	17. 0	62. 7	26. 3	54. 5
Yolov5	55. 8	47. 3	48. 7	24. 4	53. 2	33. 1	61. 0
Yolov8s	56. 3	47. 7	49. 8	25. 8	51. 7	32. 8	63. 8
Yolov10s	57. 1	48. 2	50. 2	25. 5	48. 2	25. 2	72. 2
Yolov11	57. 4	48. 4	50. 5	25. 9	45. 1	26. 3	78. 6
DETR	52. 7	44. 9	46. 5	22. 7	88. 2	41. 3	24. 4
DEIM	58. 9	51. 9	51. 4	26. 4	72. 3	40. 8	38. 6
RT-DETR	57. 8	49. 5	50. 7	26. 0	58. 6	38. 6	35. 2
FDRT-DETR	62. 6	52. 3	54. 4	28. 2	50. 6	31. 4	64. 4

层级,然而该算法在检测精度这一关键性能指标上的表现却相对较差,未能达到较高水平。尽管 DEIM 在检测精度上表现出色,但该算法的计算量和参数量较高,导致检测速度较慢。

由图 7 可见,改进的 FDRT-DETR 方法在图像细节和纹理信息完整性上对比其他方法的优势,图 7(d)中,FDRT-DETR 能精准定位并清晰勾勒其轮廓,突出纹理特征,而在 Faster R-CNN 的检测的瑕疵图像细节丢失,类型判断出错,轮廓不完整;图 7(e)中,FDRT-DETR 能够完整保留了结头边缘的纤维纹理细节,而其他方法在结头边缘的检测上存在遗漏。综合分析,FDRT-DETR 算法在计算成本、检测精度以及效率等方面均展现出卓越性能,

相较于其他算法优势显著。表明该算法在布匹瑕疵缺陷检测中精度高,且在检测速度与精度上平衡良好,更适配布匹瑕疵检测任务,为该领域提供了更理想的解决方案。

3.4 消融实验

为了探究本文提出的 FDRT-DETR 模型的各个模块对布匹检测性能的影响,对 4 个改进模块进行了消融实验的研究,依次定义基准模型 Base、改进模型 A(主干)、改进模型 B(注意力机制)、改进模型 C(颈部)、改进模型 D(损失函数)、改进模型 A+B(主干、注意力机制)、改进模型 A+B+C(主干、注意力机制、颈部)、改进模型 A+B+C+D(主干、注意力机制、颈部、损失函数),其研究的实验结果如表 5 所示。

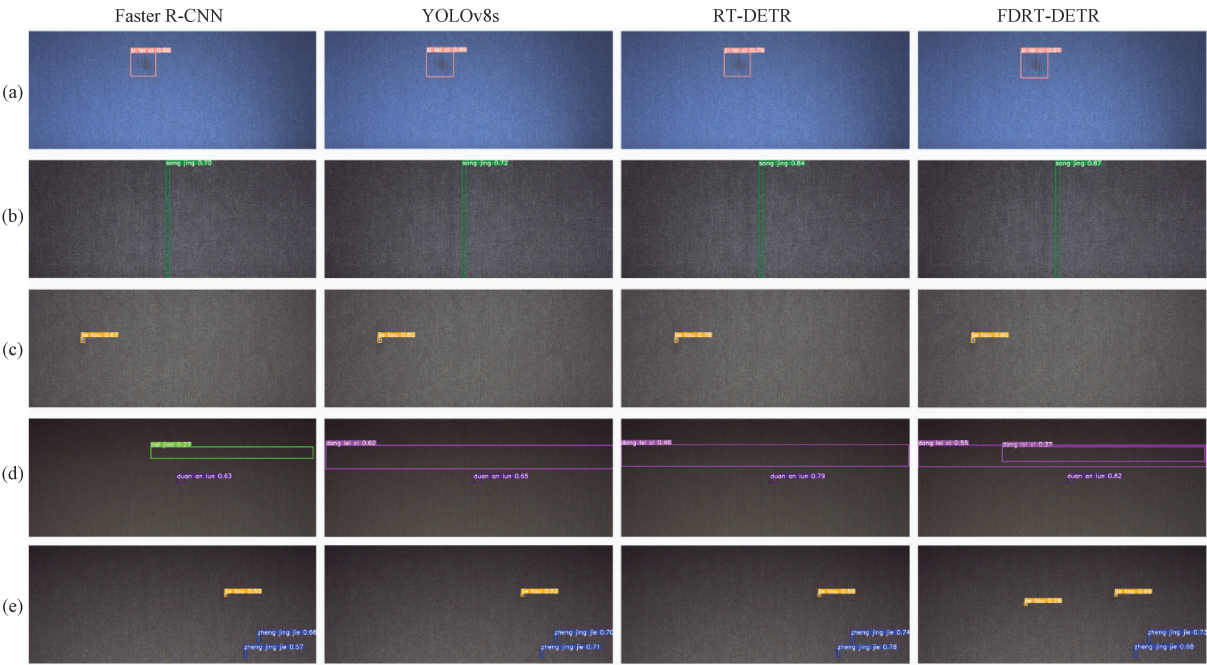


图 7 主流算法检测效果对比

Fig. 7 Comparison of detection effects of mainstream algorithms

表 5 布匹瑕疵消融实验

Table 5 Fabric defect ablation experiment

编号	模型	iRMB	TSSA	PACAPN	MAL	P/%	R/%	mAP@ 0. 5/%	mAP@ 0. 5; 0. 95/%	GFLOPs	Params/(× 10 ⁶)	帧率/fps
1	Base					57. 8	49. 5	50. 7	26. 0	58. 6	38. 6	35. 2
2	A	✓				58. 8	50. 5	52. 2	27. 4	51. 8	33. 6	56. 8
3	B		✓			58. 0	49. 9	51. 9	27. 2	58. 3	38. 4	52. 6
4	C			✓		59. 0	51. 5	52. 9	27. 6	54. 9	35. 4	52. 8
6	D				✓	58. 5	49. 6	51. 5	27. 0	58. 6	38. 6	55. 0
7	A+B	✓	✓			59. 8	50. 4	52. 7	27. 5	51. 6	33. 6	57. 4
8	A+B+C	✓	✓	✓		61. 9	51. 9	54. 0	28. 0	50. 6	31. 4	62. 2
9	A+B+C+D	✓	✓	✓	✓	62. 6	52. 3	54. 4	28. 2	50. 6	31. 4	64. 4

由表 5 可知,本文改进的 4 个模块在不同指标上均有提升。与 Base 相比,在引入 iRMB 模块改进主干网络之后,mAP@ 0. 5 值和 mAP@ 0. 5; 0. 95 分别提高了 1. 5% 和 1. 4%,参数量降低至 33. 6×10⁶,有效增强了主干网络的特征提取能力。继续引入 TSSA 模块时,在模型的 mAP@ 0. 5 值和 mAP@ 0. 5; 0. 95 分别达到了 52. 7% 和 27. 5%,表明了其有效增强模型对瑕疵区域的关注度,降低背景噪声干扰。紧接着在引入 PACAPN 模块时,模型的 mAP@ 0. 5 值和 mAP@ 0. 5; 0. 95 分别达到了 54. 0% 和 28. 0%,使得模型在多尺度融合时能够更好地识别并保留微小瑕疵的信息。最后在引入 MAL 时,模型的各项指标达到了最佳,其 P 值和 R 值分别为 62. 6% 和 52. 3%,mAP@ 0. 5 值和 mAP@ 0. 5; 0. 95 分别达到了 54. 4% 和 28. 2%,参数量仅为 31. 4×10⁶,帧率达到了 64. 4 fps。实验表明,本研究所提出的 FDRT-DETR 算法

在进行布匹瑕疵检测时,在检测精度、模型复杂度及检测速度等方面达成了较好的平衡。

3. 5 可视化实验

为更加直观的展现本文所提出的 FDRT-DETR 模型的有效性,对模型改进前后所提取的到特征图进行了热力图可视化,结果如图 8 所示。

由图 8 可见模型改进前后在瑕疵特征提取方面的差异,改进后的 FDRT-DETR 模型能够更好地聚焦于图像中瑕疵的关键区域,热力图在瑕疵位置呈现出更集中、更强烈的响应,而背景纹理区域的响应则显著降低。这充分说明了提出的改进方法能够有效地增强模型对瑕疵特征的捕捉精度,使模型更加专注于瑕疵区域的特征提取,减少了背景干扰,从而提高了模型的特征表达能力和检测性能。例如,在含有“痕类疵”的图像中,改进后的模型热力图能够清晰地勾勒出痕类疵的轮廓和主要分布区

域,而原始模型的热力图则相对较为分散,对渍类疵的聚焦不够精准,这体现了改进方法在特征提取针对性和有效性的提升。此外,FDRT-DETR 在对像结头、整经结等微小瑕疵的识别能力更突出,能够更加准确地定位到小目标缺陷,进一步证明了改进方法对于增强 RT-DETR 性能的有效性。

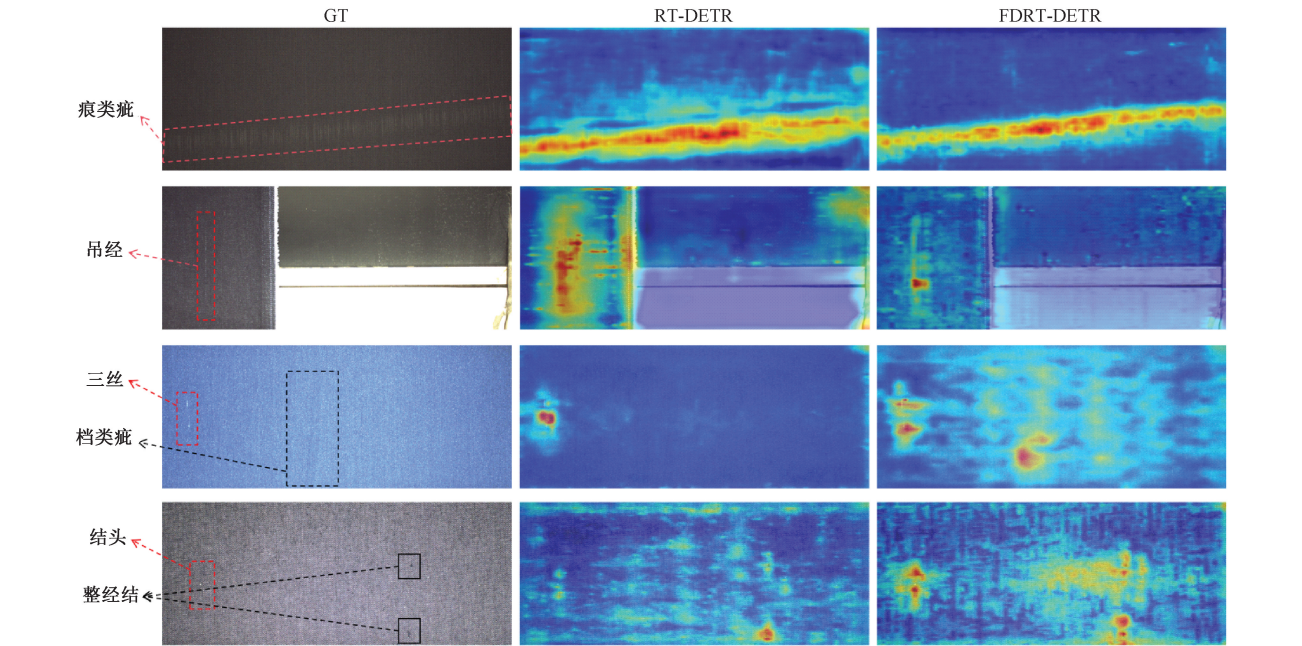


图 8 模型热力图
Fig. 8 Model heatmap

3.6 泛化实验

为进一步验证 FDRT-DETR 算法的泛化能力与鲁棒性,本文在阿里天池花色布匹数据集上进行泛化实验,并对比了不同的检测算法,其结果如表 6 所示。可以看出,在复杂花色干扰下,FDRT-DETR 依然保持最优性能,其 P 值、 R 值、 $mAP@0.5$ 和 $mAP@0.5:0.95$ 显著高于其他

对比方法,且推理速度(62.5 fps)保持均衡。相比之下,YOLO 系列精度下降明显,DEIM 虽精度较高但速度偏低。实验结果表明,改进算法不仅能够有效克服复杂花色纹理的干扰,也展现出对未见瑕疵类型良好的适应性,与检测稳定性,进一步验证了其在复杂工业场景中的泛化能力和鲁棒性。

表 6 泛化实验
Table 6 Generalization experiment

模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	GFLOPs	Params/($\times 10^6$)	帧率/fps
RT-DETR	53.6	45.2	46.8	22.9	58.6	38.6	34.8
Faster R-CNN	38.9	32.1	33.5	16.2	220.2	102.1	18.5
Yolov8s	52.1	44.3	45.6	22.4	51.7	32.8	60.2
Yolov10s	52.8	44.7	45.8	22.1	48.2	25.2	68.5
Yolov11	53.2	45.0	46.2	22.7	45.1	26.3	75.3
DEIM	55.4	47.6	49.5	24.8	72.3	40.8	36.2
FDRT-DETR	58.2	49.1	51.8	26.3	50.6	31.4	62.5

4 结 论

本文提出了一种用于布匹瑕疵检测的深度学习方法,该方法基于当前在目标检测任务上取得优异成绩的 RT-DETR 网络作为基准网络,通过引入 iRMB 模块、

TSSA 模块、PACAPN 模块及 MAL 损失函数,对其模型进行改进,使其更加适用于布匹瑕疵检测任务。实验结果表明,该算法在布匹瑕疵检测任务中,在保持较低参数量和计算量的同时,显著提升了检测精度和速度,能够有效实现多类别瑕疵的精准检测,并对小目标缺陷检测性能有明显改进。所提模型在布匹瑕疵检测中展现出快速识

别的优势,优于一些主流模型,为智能布匹瑕疵检测的进一步发展提供了有力支持。未来的研究将持续优化布匹瑕疵检测模型,致力于实现模型架构的进一步轻量化,以加速算法迭代并增强对各类瑕疵的精准识别能力。这不仅有助于在实际工作环境中高效部署相关模型,而且能够更好地满足企业对自动化瑕疵检测日益增长的需求。

参考文献

- [1] 郑小虎,刘正好,陈峰,等. 纺织工业智能发展现状于展望[J]. 纺织学报,2023,44(8):205-216.
Zheng Xiaohu, Liu Zhenghao, Chen Feng, et al. Current status and prospect of intelligent development in textile industry[J]. Journal of Textile Research, 2023, 44(8): 205-216.
- [2] 赵佰亭,张晨,贾晓芬. ECC-YOLO:一种改进的钢材表面缺陷检测方法[J]. 电子测量与仪器学报,2024,38(4):108-116.
Zhao Baiting, Zhang Chen, Jia Xiaofen. ECC-YOLO: An improved method for detecting surface defects in steel[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4): 108-116.
- [3] 邓王宁,智敏. 深度学习下的单阶段通用目标检测算法研究综述[J]. 计算机科学与探索,2025,19(5):1115-1140.
Deng Wangning, Zhi Min. Review of one-stage universal object detection algorithms in deep learning[J]. Journal of Frontiers of Computer Science and Technology, 2025, 19(5): 1115-1140.
- [4] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. ArXiv preprint arXiv:1506.01497, 2015.
- [5] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN[C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [6] Yu Y, Yang X, Li Q, et al. H2RBox-v2: Incorporating symmetry for boosting horizontal box supervised oriented object detection[J]. Advances in Neural Information Processing Systems, 2023, 36: 59137-59150.
- [7] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[C]//Computer Vision-ECCV 2016. Berlin: Springer International Publishing, 2016: 21-37.
- [8] Lyu C Q, Zhang W W, Huang H A, et al. RTMDet: An empirical study of designing real-time object detectors[J]. ArXiv preprint arXiv:2212.07784, 2022.
- [9] Wu W T, Liu H, Li L L, et al. Application of local fully convolutional neural network combined with YOLO v5 algorithm in small target detection of remote sensing image[J]. PloS One, 2021, 16(10): e0259283.
- [10] Reis D, Kupec J, Hong J, et al. Real-time flying object detection with YOLOv8[J]. ArXiv preprint arXiv:2305.09972, 2023.
- [11] Wang A, Chen H, Liu L H, et al. YOLOv10: Real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.
- [12] Khanam R, Hussain M. YOLOv11: An overview of the key architectural enhancements[J]. ArXiv preprint arXiv:2410.17725, 2024.
- [13] Li F, Xiao K, Hu Z P, et al. Fabric defect detection algorithm based on improved YOLOv5[J]. The Visual Computer, 2024, 40(4): 2309-2324.
- [14] Wang S, Kong H Y, Li B T, et al. Fab-ME: A vision state-space and attention-enhanced framework for fabric defect detection[J]. ArXiv preprint arXiv:2412.03200, 2024.
- [15] Mao M, Lee A, Hong M. Efficient fabric classification and object detection using YOLOv10[J]. Electronics, 2024, 13(19): 3840.
- [16] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]. European Conference on Computer Vision. Cham: Springer International Publishing, 2020: 213-229.
- [17] Zhao Y A, Lyu W Y, Xu S L, et al. Detrs beat YOLOs on real-time object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [18] Huang S H, Lu Z C, Cun X D, et al. DEIM: DETR with improved matching for fast convergence[J]. ArXiv preprint arXiv:2412.04234, 2024.
- [19] 宋荣,邱夷平,夏克尔·赛塔尔. 基于改进 RT-DETR 的织物缺陷检测算法[J]. 棉纺织技术, 2025, 53(7): 25-31.
Song Rong, Qiu Yiping, Shakir Saitar. Fabric defect detection algorithm based on improved RT-DETR[J]. Cotton Textile Technology, 2025, 53(7): 25-31.
- [20] 梁耕良,韩曙光. 基于改进 RT-DETR 的牛仔面料疵点检测算法[J]. 浙江大学学报(工学版), 2025, 59(6): 1169-1178, 1190.
Liang Gengliang, Han Shuguang. Denim fabric defect detection algorithm based on improved RT-DETR[J]. Journal of Zhejiang University (Engineering Science), 2025, 59(6): 1169-1178, 1190.
- [21] Zhang J N, Li X T, Li J, et al. Rethinking mobile block for efficient attention-based models[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE Computer Society, 2023: 1389-1400.
- [22] Wu Z Y, Ding T J, Lu Y F, et al. Token statistics

- transformer: Linear-time attention via variational rate reduction[J]. ArXiv preprint arXiv:2412.17810, 2024.
- [23] Chen J R, Kao S H, He H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [24] Liu X Y, Peng H W, Zheng N X, et al. Efficientvit: Memory efficient vision transformer with cascaded group attention[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 14420-14430.
- [25] Fan Q H, Huang H B, Chen M R, et al. RMT: Retentive networks meet vision transformers[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 5641-5651.
- [26] Qin D F, Leichner C, Delakis M, et al. MobileNetV4: universal models for the mobile ecosystem[C]//European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 78-96.
- [27] Yu W H, Wang X C. Mambaout: Do we really need mamba for vision? [J]. ArXiv preprint arXiv: 2405.07992, 2024.
- [28] 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.
- [29] Li H L, Li J, Wei H B, et al. Slim-neck by GSConv: A lightweight-design for real-time detector architectures[J]. Journal of Real-Time Image Processing, 2024, 21(3): 62.
- [30] Chen Y F, Zhang C Y, Chen B, et al. Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases[J].

Computers in Biology and Medicine, 2024, 170: 107917.

- [31] Wang H N, Cao P, Wang J Q, et al. Uctransnet: Rethinking the skip connections in u-net from a channel-wise perspective with transformer[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(3): 2441-2449.
- [32] Li W J, Guo H, Liu X N, et al. Efficient face super-resolution via wavelet-based feature enhancement network[C]//Proceedings of the 32nd ACM International Conference on Multimedia, 2024: 4515-4523.

作者简介



金民, 2021 年于安徽工程大学获得学士学位, 现为安徽工程大学硕士研究生, 主要研究方向为计算机视觉、数字图像处理。
E-mail: 2230911109@stu.ahpu.edu.cn

Jin Min received his B. Sc. degree from Anhui Polytechnic University in 2021. He is currently a M. Sc. candidate at Anhui Polytechnic University. His main research interests include computer vision and digital image processing.



汪军(通信作者), 1999 年于安徽机电学院学士学位, 2010 年于安徽工程大学获得硕士学位, 现为安徽工程大学硕士生导师、教授、院长, 主要研究方向为机器视觉与智能控制。

E-mail: wangjun@ahpu.edu.cn

Wang Jun (Corresponding author) received his B. Sc. degree from Anhui Institute of Mechanical and Electrical Engineering in 1999 and his M. Sc. degree from Anhui Polytechnic University in 2010. He is currently a professor, M.Sc. supervisor, and the Dean of the School of Computer and Information at Anhui Polytechnic University. His main research interests include machine vision and intelligent control.