



## 结合注意力的纺织品瑕疵检测方法研究

周在雍, 狄岚

引用本文:

周在雍, 狄岚. 结合注意力的纺织品瑕疵检测方法研究[J]. 智能系统学报, 2024, 19(4): 827-838.

ZHOU Zaiyong, DI Lan. Research on textile defect detection method combining attention mechanism[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(4): 827-838.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202304045>

## 您可能感兴趣的其他文章

### 双向特征融合与注意力机制结合的目标检测

Target detection based on bidirectional feature fusion and an attention mechanism

智能系统学报. 2021, 16(6): 1098-1105 <https://dx.doi.org/10.11992/tis.202012029>

### 注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN

智能系统学报. 2020, 15(1): 92-98 <https://dx.doi.org/10.11992/tis.201907023>

### 基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism

智能系统学报. 2020, 15(5): 956-963 <https://dx.doi.org/10.11992/tis.201903001>

### 基于注意力融合的图像描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740-749 <https://dx.doi.org/10.11992/tis.201910039>

### 基于光照预处理与特征提取的纺织品瑕疵检测方法

Fabric defect inspection based on illumination preprocessing and feature extraction

智能系统学报. 2019, 14(4): 716-724 <https://dx.doi.org/10.11992/tis.201805023>

### 基于双向消息链路卷积网络的显著性物体检测

Salient object detection based on bidirectional message link convolution neural network

智能系统学报. 2019, 14(6): 1152-1162 <https://dx.doi.org/10.11992/tis.201812003>

DOI: 10.11992/tis.202304045

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20240311.1010.002>

# 结合注意力的纺织品瑕疵检测方法研究

周在雍<sup>1</sup>, 狄岚<sup>2</sup>

(1. 江南大学人工智能与计算机学院, 江苏 无锡 214000; 2. 江南大学数字媒体学院, 江苏 无锡 214122)

**摘要:** 本文阐述了一种名为 SAAM-YOLOX 的基于改进 YOLOX 的纺织品瑕疵检测模型, 旨在解决纺织品瑕疵检测中针对犬牙花纹与格纹背景中出现的误检和漏检问题, 以及整体检测精度不高的问题。在特征提取阶段, 该模型引入了离散余弦变换所构建的多分支离散余弦注意力机制 (multi-branch discrete cosine attention, MDCA), 能够解决模型在犬牙花纹与格纹背景中出现的误检和漏检问题, 并且在检测精度上有一定的提高; 在特征融合阶段, 为了聚集和加强不同尺度的语义特征, SAAM-YOLOX 模型采用了尺度聚合技术和注意力机制来构建尺度聚合注意力模块 (scale aggregation attention module, SAAM)。在 SAAM 的上采样过程中, 使用双线性插值结合自注意力机制来增强特征信息的有效性, 从而进一步提高检测的精度。在完成尺度聚合后, 加入注意力模块来增强混合尺度的特征表示, 最终实现提高检测效果的目的。实验结果表明, 本文检测模型解决了犬牙花纹与格纹背景中出现的误检和漏检问题, 并且提高了瑕疵检测的精度。

**关键词:** 注意力机制; 尺度聚合; 双线性插值; 离散余弦变换; 多尺度特征; 特征融合; 纺织品瑕疵检测; 计算机视觉

中图分类号: TP391 文献标志码: A 文章编号: 1673-4785(2024)04-0827-12

中文引用格式: 周在雍, 狄岚. 结合注意力的纺织品瑕疵检测方法研究 [J]. 智能系统学报, 2024, 19(4): 827-838.

英文引用格式: ZHOU Zaiyong, DI Lan. Research on textile defect detection method combining attention mechanism [J]. CAAI transactions on intelligent systems, 2024, 19(4): 827-838.

## Research on textile defect detection method combining attention mechanism

ZHOU Zaiyong<sup>1</sup>, DI Lan<sup>2</sup>

(1. School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214000, China; 2. School of Digital Media, Jiangnan University, Wuxi 214122, China)

**Abstract:** This paper presents a textile defect detection model called SAAM-YOLOX, which is based on the improved YOLOX. The model aims to address the issues of false positives and false negatives in textile defect detection, particularly in detecting Houndstooth and Gingham backgrounds, as well as the problem of overall low detection accuracy. In the feature extraction stage, the model introduces a multi-branch discrete cosine attention mechanism (MDCA) based on the discrete cosine transform to address false positives and false negatives in the Houndstooth and Gingham backgrounds, and thereby improve the detection accuracy. In the feature fusion stage, the SAAM-YOLOX model adopts scale aggregation technology and attention mechanism to construct a scale aggregation attention module (SAAM) to aggregate and enhance semantic features of different scales. SAAM uses bilinear interpolation combined with self-attention mechanism to enhance validity of feature information in the upsampling process, further improving the detection accuracy. After completing the scale aggregation, an attention module is added to enhance the mixed-scale feature representation, ultimately achieving the goal of improving detection performance. Experimental results demonstrate that the proposed detection model can effectively solve the problem of false positives and false negatives in Houndstooth and Gingham backgrounds, and improve the accuracy of defect detection.

**Keywords:** attention mechanism; scale aggregation; bilinear interpolation; discrete cosine transform; multi-scale feature; feature fusion; textile defect detection; computer vision

收稿日期: 2023-04-22. 网络出版日期: 2024-03-11.

基金项目: 国家自然科学基金项目 (71971031).

通信作者: 狄岚. E-mail: [dilan@jiangnan.edu.cn](mailto:dilan@jiangnan.edu.cn).

在纺织品生产过程中, 瑕疵的形成会对纺织品的质量造成严重影响。目前, 传统的纺织品瑕

疵检测主要依靠人工完成,其检测质量取决于工人状态,且检测速度较慢、效率低。因此,工业生产需要自动高效的检测方法。

目前,纺织品瑕疵检测方法可以归纳为统计学方法、结构类方法、模型方法和基于神经网络的方法。统计学方法包括线性反投影算法(linear back projection, LBP)<sup>[1]</sup>和灰度共生矩阵(gray level CO-Occurrence matrix, GLCM)<sup>[2]</sup>等,这些方法基于相关函数和算子来提取图像特征,并确定瑕疵位置。结构类方法以纹理为基本单元提取纺织品特征,并根据结构相似性最小原则确定瑕疵图像块。模型方法则使用特定分布模型来进行纺织品瑕疵检测。此外,一些研究者提出了基于图像校正和模板分割的纺织品瑕疵检测方法<sup>[3]</sup>以及基于畸变矫正与视觉显著性特征的方法<sup>[4]</sup>,但这些方法大多适用于简单周期性图案,无法满足复杂的现实世界纺织品瑕疵检测的要求。因此,需要研究新的方法来解决这一问题,其中基于神经网络的方法已经成为研究热点。

随着人工智能技术的不断进步,卷积神经网络被广泛应用于纺织品瑕疵检测领域,在检测精度和速度方面相较于传统方法有着显著的提升。基于神经网络的目标检测算法一般分为一步式检测和两步式检测<sup>[5-6]</sup>。一般而言,一步式检测速度更快,两步式检测的精度更高。代表两步式检测算法的有 RCNN(region-based convolutional neural network)<sup>[7]</sup>、Fast-RCNN(fast region-based convolutional neural network)<sup>[8]</sup>和 Faster-RCNN(faster region-based convolutional neural network)<sup>[9]</sup>,这些算法具有强大的特征工程能力,在目标检测方面表现良好。代表一步式检测算法的有 SSD(single shot multibox detector)<sup>[10-11]</sup>、Cascade RCNN(cascade region-based convolutional neural network)<sup>[12]</sup>、YOLO(you only look once)<sup>[13]</sup>系列,采用端到端的检测算法生成预测信息。

在 YOLO 系列中,YOLOX<sup>[14]</sup>以 YOLOv3 作为模型的原始框架,加载 Decoupled Head、Data Aug、Anchor Free 和 SimOTA 部件。尽管 YOLOX 在物体检测方面表现出色,但在纺织品瑕疵检测中其表现欠佳,由于纺织品瑕疵的特殊性质,即瑕疵尺寸和形状多种多样,传统的特征提取方法无法很好地捕捉到这些瑕疵的特征信息,最终导致检测精度较低。

针对上述问题,本文贡献如下:

1)为了获得更好的纺织品瑕疵预测图,在特征提取网络中使用多分支离散余弦注意力(multi-

branch discrete cosine attention, MDCA)机制加强对纺织品瑕疵特征的提取能力。

2)在特征融合网络模仿变焦策略,在多个尺度上同步考虑瑕疵与背景的差异化关系,采用尺度聚合技术和注意力机制来构建尺度聚合注意力模块(scale aggregation attention module, SAAM),从而充分感知瑕疵与混淆的背景,最终解决了犬牙花纹与格纹背景中出现的误检和漏检问题。

3)在尺度聚合的上采样阶段,使用双线性插值结合自注意力机制的上采样方法增强特征信息的有效性,进一步提高检测的精度。

实验结果表明,本文提出的 SAAM-YOLOX(scale aggregation attention module you only look once with overhaul X)在保证速度的同时具有较高的检测精度。

## 1 相关工作

### 1.1 注意力机制

注意力机制<sup>[15]</sup>是一种能够提升深度学习模型性能的强大工具,其中自注意力机制是最基本也是最常用的形式之一。自注意力机制最初被广泛应用于 Transformer 模型中,它可以有效地捕捉输入序列或特征图中的长程依赖关系,计算每个元素与其他元素之间的关联性,然后对不同元素进行加权,以得到具有不同重要性的特征表示。

除了自注意力机制之外,还有许多其他类型的注意力机制。通道注意力机制主要关注输入特征图中的通道信息,通过对通道维度进行加权来提高不同通道的重要性。例如,SENet(squeeze and excitation network)模型就是一种基于通道注意力机制的神经网络结构,它在图像分类等任务中取得了优异的性能。空间注意力机制主要关注输入特征图中的空间位置信息,通过对空间维度进行加权来提高不同空间位置的重要性,例如,CBAM(convolutional block attention module)模型就是一种基于空间注意力机制的神经网络结构。多头注意力机制是一种同时考虑多个注意力机制的方法,通过将输入特征图分成多个头,分别对不同的头进行自注意力或通道空间注意力等操作,然后将多个头的输出进行融合,得到最终的特征表示。

### 1.2 多尺度特征融合

多尺度特征融合<sup>[16]</sup>是一种常用的视觉任务中的技术,旨在将来自不同尺度的特征信息有效地结合起来,以提高模型的性能。常用的多尺度特征融合方法包括以下4种:

1)级联方法:一种将多个网络层级联起来的方法,每层负责处理不同尺度的特征<sup>[17]</sup>。

2)金字塔方法:一种通过对输入图像进行不同尺度的缩放来获得多尺度特征的方法<sup>[18]</sup>。

3)横向连接方法:一种将不同网络层的特征进行横向连接的方法,以获得不同尺度的特征。

4)注意力机制方法:一种通过对不同尺度的特征进行加权来获得多尺度特征的方法。

这些多尺度特征融合方法可以相互组合以提高模型性能。

### 1.3 纺织品瑕疵检测相关的目标检测算法

目前已经有一些基于深度学习的与纺织品瑕疵检测相关的目标检测算法。文献[19]提出了一种基于 Cascade Faster R-CNN 的织物瑕疵检测方法,采用级联式检测,能够进一步提高检测准确率和检测速度。该方法虽然在准确性和速度方面表现良好,但是仍存在一些局限性。例如当织物出现多种瑕疵时,该算法会出现漏检或误检的情况。文献[20]提出了一种基于 Cascade R-CNN 的瑕疵检测算法,并且在该算法的基础上引入了多项技巧以提高检测的准确率和效率。然而,该方法的主要局限性是采用了多项技巧,使其实现相对较为复杂,需要对算法进行多次调优和参数设置。文献[21]提出了一种改进的 YOLOv3 模型用于纺织品瑕疵检测,其中采用了一种基于特征聚合的注意力机制来增强模型对小目标的检测能力,并在损失函数中加入了一个负样本权重系数,以解决类别不平衡问题。然而,该方法仍存在一些局限性,例如对于一些复杂的瑕疵类型检测准确率不够高。文献[22]提出了一种基于改进的 YOLOv5 算法的布匹疵点检测方法。该方法使用了自适应池化模块和 CBAM 模块来增强网络的特征表达能力,使用了多任务学习方法的同时进行疵点分类和位置回归,通过改进的损失函数来优化模型性能。然而,该方法对特定类型的疵点检测不准确,同时在复杂场景下的表现会受到一定的限制。

为了保证在检测速度和检测精度之间达到平衡,本文采用了基于 YOLOX 网络的纺织品瑕疵检测算法。相比于双阶段目标检测算法,YOLOX 网络具有更快的检测速度,同时比单阶段目标检测算法具有更高的检测精度。

## 2 SAAM-YOLOX 纺织品瑕疵检测算法

本文提出的改进 YOLOX 的纺织品瑕疵检测算法 (scale aggregation attention module you only

look once with overhaul X, SAAM-YOLOX) 的整体网络结构如图 1 所示。该算法主要包括骨干网络模块、特征融合模块、检测头模块 3 个模块。其中,骨干网络模块用于提取输入图像的特征;特征融合模块将来自不同尺度的特征进行融合;检测头模块将融合后的特征用于瑕疵检测任务。

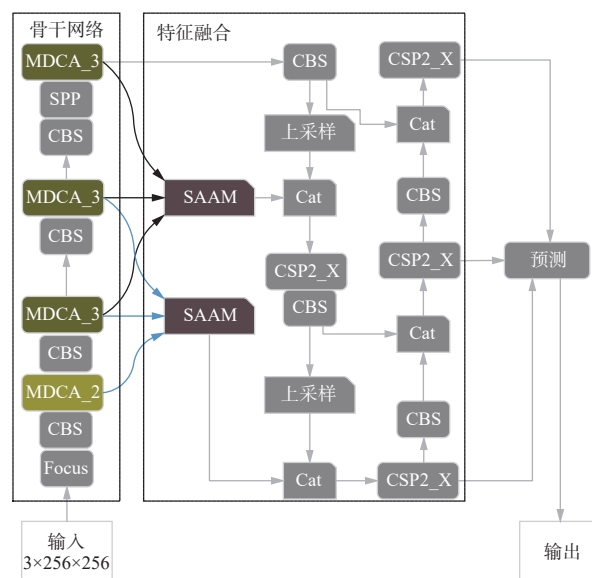


图 1 SAAM-YOLOX 网络结构

Fig. 1 SAAM-YOLOX network structure

本文在骨干网络模块中引入多分支离散余弦注意力机制模块 MDCA,对 YOLOX 的特征提取部分进行优化。在特征融合阶段采用尺度聚合技术和注意力机制来构建尺度聚合注意力模块 SAAM,从而充分感知瑕疵与混淆的背景,最终实现提高检测效果的目的。

### 2.1 引入多分支离散余弦注意力机制模块 MDCA

在本文提出的改进 YOLOX 的纺织品瑕疵检测算法中,为了解决模型对纺织品瑕疵中犬牙花纹与格纹背景所出现的误检和漏检问题,本文在 YOLOX 的特征提取网络中引入了一种名为多分支离散余弦注意力机制模块 MDCA。MDCA 模块基于离散余弦变换<sup>[23]</sup>,能够聚焦于不同尺度上的特征,增强对目标区域的感知,从而提高纺织品瑕疵检测的精度。在 MDCA 模块中,本文使用了多个尺度的特征图来提取不同尺度的特征,并通过一个注意力机制来加权融合这些特征,最终生成一个更具有区分性的特征图,为后续的瑕疵检测提供更加准确的信息。通过引入 MDCA 模块,本文能够在保证较高检测精度的同时,解决犬牙花纹与格纹背景中出现的误检和漏检问题。如图 2 为多分支离散余弦注意力机制 MDCA 模块结构。

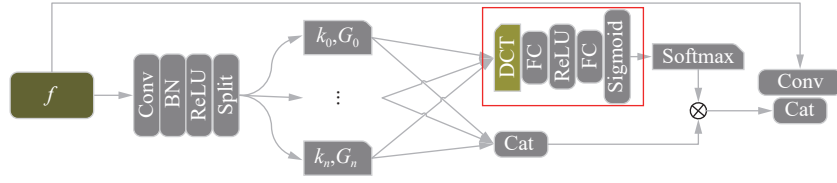


图 2 多分支离散余弦注意力机制 MDCA 模块

Fig. 2 Multi-branch discrete cosine attention(MDCA) module

首先, 本方法采用多分支的方式进行输入特征图的空间信息提取。这种多分支方式可以更全面地捕获位置信息, 同时可以在多个尺度上并行提取特征, 从而实现多尺度特征提取的目的。其次, 在 SEWeight 模块中引入离散余弦变换 (discrete cosine transform, DCT) 来增强频率域上的信息提取能力。DCT 的基本计算方式是: 将输入的  $N$  个一维信号分别乘以不同的频率因子, 再将结果相加得到一个  $N$  维的 DCT 系数向量, 其中每个元素代表一个频率成分的强度。通过使用 DCT, 可以提高特征图在频域上的表达能力, 进而增强模型的表达能力和性能, 计算公式为

$$B_{h,w}^{u,v} = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} x_{u,v}^{2d} \cos\left(\frac{\pi h}{H} \left(u + \frac{1}{2}\right)\right) \cos\left(\frac{\pi w}{W} \left(v + \frac{1}{2}\right)\right)$$

式中:  $H$  和  $W$  分别为图像的宽和高,  $u$  和  $v$  为频率参数,  $x$  为空域中原始数据,  $B$  为经过变换后的频率数据。对于原来的全局平均池化操作而言, 它仅仅是离散余弦变换在最低频率分量上得到的单频谱向量。具体计算如下:

$$B_{0,0}^{u,v} = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} x_{u,v}^{2d} \cos\left(\frac{0}{H} \left(u + \frac{1}{2}\right)\right) \cos\left(\frac{0}{W} \left(v + \frac{1}{2}\right)\right) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} x_{u,v}^{2d} = \text{GAP}(x^{2d})HW$$

式中 GAP 为全局平均池化。使用通道注意力权重与相应的特征图相乘, 得到经过注意力机制筛选后的特征图。为了避免在网络加深时梯度消失的问题, 采用残差结构将得到的特征图与最初的输入特征图相加。

图 3 给出了原始图像、YOLOX 模型预测的瑕疵图像和在骨干网络中加入 MDCA 模块后预测的瑕疵图像。从图 3 中可以看出, YOLOX 模型在纯色背景下的预测效果较好, 但对于方格纹案的背景, 瑕疵预测效果较差, 并且容易将整个纺织品预测为污渍类别的瑕疵。然而, 在骨干网络中加入 MDCA 模块后, 纺织品瑕疵的预测效果显著提高。本文提出的 MDCA 模块是一种注意力机制, 可以让网络更关注关键区域并减少噪声的干扰。因此, 在加入 MDCA 模块后, 网络可以更好地学习纺织品瑕疵的特征, 从而提高瑕疵的检测准确性。

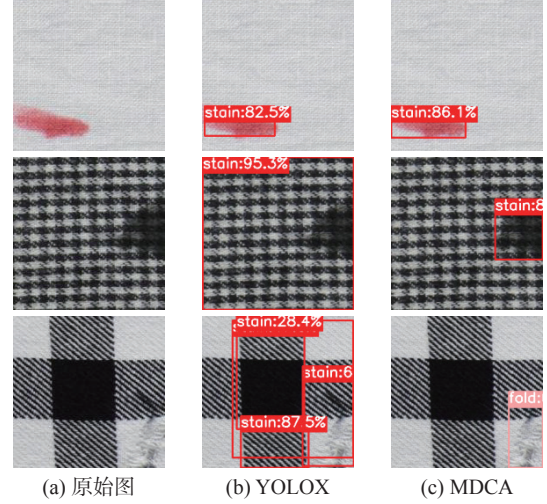


图 3 应用 MDCA 模块前后预测图对比

Fig. 3 Comparison of predicted images before and after applying the MDCA module

## 2.2 构造 SAAM 模块

本文提出了一种基于注意力机制的尺度聚合模块 SAAM, 用于加权和整合特定尺度的特征信息。如图 4 所示, SAAM 包含尺度变换层和尺度聚合层 2 个部分。尺度变换层用于调整不同尺度特征图的尺度以保持一致, 而尺度聚合层则将仅保留特定尺度特征信息的特征图输入到注意力生成器中, 以获得相应于每个尺度的权重。通过使用这些权重来加权特征图, 可以提取多尺度融合特征。

构造 SAAM 模块的具体步骤如下:

1) 提取骨干网络中不同尺度的特征图  $S$  (小尺度)、 $M$  (主尺度) 和  $L$  (大尺度), 通过卷积和上采样的方式, 将小尺度特征图  $S$  的尺度调整到与主尺度特征图  $M$  相同, 其中上采样采用双线性插值的方法; 通过卷积核下采样的方法, 将大尺度特征图  $L$  的尺度调整为主尺度, 其中下采样采用最大值池化的方法。

设小尺度特征图  $S$  的大小为  $h_S \times w_S \times c_S$ , 主尺度特征图  $M$  的大小为  $h_M \times w_M \times c_M$ , 大尺度特征图  $L$  的大小为  $h_L \times w_L \times c_L$ , 调整后的小尺度特征图为  $S'$ , 大尺度特征图为  $L'$ , 大小与主尺度特征图相同, 即  $S'$  和  $L'$  的大小为  $h_M \times w_M \times c_M$ 。卷积和上采样的方式可以表示为

$$S' = \text{Up}(C(S), s = [h_M, w_M])$$

式中:  $C(S)$  为小尺度特征图  $S$  进行卷积操作,  $\text{Up}$  为上采样操作,  $s$  参数为上采样后的大小为  $[h_M, w_M]$ 。上采样采用的双线性插值表示为

$$g_{i,j} = (1-dx)(1-dy)f_{x,y} + dx(1-dy)f_{x+1,y} + (1-dx)dyf_{x,y+1} + dxdyf_{x+1,y+1}$$

式中:  $f_{x,y}$  为原始图像中坐标为  $(x,y)$  的像素值,  $f_{x+1,y}$ 、 $f_{x,y+1}$ 、 $f_{x+1,y+1}$  分别为其相邻的 3 个像素的值。这样, 就可以通过双线性插值来将小尺度特征图

调整到与主尺度特征图相同的尺度。

卷积和下采样的方式可以表示为

$$L' = \text{Do}(C(L), s = [h_M, w_M])$$

式中:  $C(L)$  为对大尺度特征图  $L$  进行卷积操作,  $\text{Do}$  为下采样操作,  $s$  参数为下采样后的大小为  $[h_M, w_M]$ 。下采样采用的最大值池化表示为

$$\max_{i=1}^k(x_i)ki$$

式中:  $k$  为池化窗口内的元素个数,  $x_i$  为池化窗口内第  $i$  个元素的值,  $\max$  为求最大值的操作。

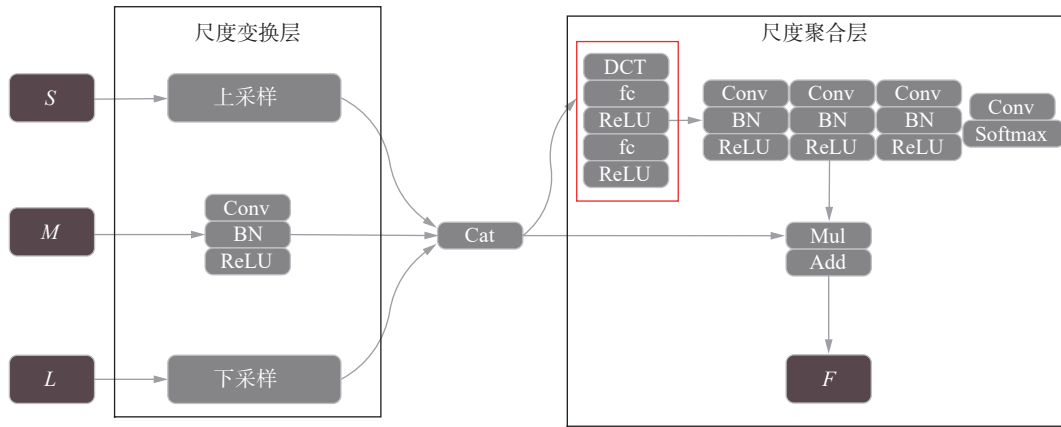


图4 尺度聚合注意力模块

Fig. 4 Scale aggregation attention module

2) 在将保留特定尺度特征信息的特征图通过上采样、下采样和卷积等操作后, 输入到注意力生成器中, 以获得对应于每个尺度的权重信息, 用于提取多尺度融合特征。首先将特征图输入到 DCT 变换中, 将结果经过 2 次全连接层 ( $f$ ) 和 ReLU 激活函数处理, 该过程可以表示为

$$y = R(f_2(R(f_1(D(x)))))$$

式中:  $f_1$  和  $f_2$  分别为 2 个全连接层,  $R$  为激活函数,  $D$  为离散余弦变换。然后通过一系列卷积层计算出三通特征图, 在 Softmax 激活层后获取分别对应每个尺度的注意力权重, 计算过程如下:

$$A = \text{Softmax}(\Psi([\text{Up}(S), M, \text{Do}(L)], \alpha))$$

$$F = A^S \cdot \text{Up}(S) + A^M \cdot M + A^L \cdot \text{Do}(L)$$

式中:  $\Psi(\beta, \alpha)$  为“Conv-BN-ReLU”层,  $[\beta]$  为连接操作,  $\alpha$  为参数,  $\text{Up}$  和  $\text{Do}$  分别为上文中上采样与下采样运算,  $A$  为对应尺度的注意力权重,  $F$  为最终输出结果。

### 2.3 引入自注意力机制构造上采样模块

上采样是一种常见的图像处理操作, 用于提高特征图的尺度。然而, 传统的上采样方法会导致图像细节的丢失和模糊, 因为它们只是简单地将像素复制或插值到更高分辨率的位置。为了解

决这个问题, 本文提出了一种结合自注意力机制的上采样方法, 即双线性插值结合自注意力机制。如图 5 所示, 这种方法通过使用自注意力生成器来突出特定尺度的特征表达, 确保了双线性插值后特征信息的有效性, 并提高了模型的性能和准确性。图 6 为引入上采样模块后的改进尺度聚合注意力模块。



图5 上采样模块

Fig. 5 Upsampling module

具体来说, 该方法首先使用双线性插值将低尺度的特征图扩展到与主尺度特征图相同的尺度大小。这样可以避免尺度不一致的问题, 并确保特定尺度的信息在尺度聚合过程中得到更好的保留和利用。然后, 将特征图输入到自注意力生成器中, 生成相应的权重系数, 以突出特定尺度的特征表达。最后, 将这些特征图与其他尺度的特征图相结合, 生成最终的特征表示。这种方法能够保留特定尺度的信息, 提高模型性能, 同时避免了丢失细节和模糊的问题, 确保了双线性插值后特征信息的有效性。

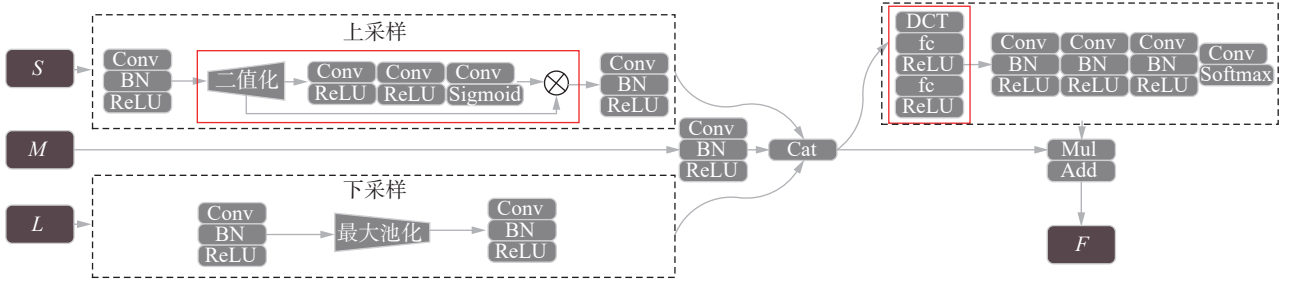


图 6 引入上采样模块的改进尺度聚合注意力模块

Fig. 6 Improved scale aggregation attention module with the introduction of the upsampling module

首先, 对该尺度的特定特征图进行双线性插值操作, 得到上采样后的特征图 $Y$ :

$$Y_{ij} = \sum_m \sum_n X_{mn} \cdot f(m-n) \cdot f(n-j)$$

式中:  $i$ 和 $j$ 分别为上采样后 $Y$ 中的行和列坐标,  $m$ 和 $n$ 分别为原始特征图 $X$ 中的行和列坐标,  $f(x)$ 为双线性插值函数。

接下来, 将 $Y$ 输入到一个自注意力结构中。该结构由 3 个卷积层组成, 每个卷积层后都跟着一个 ReLU 激活函数。其中, 第 1 个、第 2 个卷积层的输出通道数为 $C$ , 第 3 个卷积层的输出通道数为 1, 用于产生注意力权重。第 1 个卷积层的输出为 $Z_1$ :

$$Z_1 = \text{Conv}(Y)$$

式中: Conv 表示卷积操作;  $Z_1$ 的形状为 $[H, W, C]$ ,  $H$ 和 $W$ 分别为上采样后特征图 $Y$ 的高度和宽度。第 2 个卷积层的输出为 $Z_2$ :

$$Z_2 = \text{Conv}(\text{ReLU}(\text{Conv}(Z_1)))$$

式中: ReLU 为激活函数,  $Z_2$ 的形状和 $Z_1$ 相同。第 3 个卷积层的输出为注意力权重 $A$ :

$$A = \text{Sigmoid}(\text{Conv}(Z_2))$$

式中: Sigmoid 为激活函数,  $A$ 的形状为 $[H, W, 1]$ 。最终, 将注意力权重 $A$ 与原始特定尺度的特征图 $X$ 相乘, 得到加权后的特征图 $Z$ :

$$Z_{mn} = A_{mn} \cdot X_{mn}$$

### 3 实验与结果分析

ZJU-Leaper 纺织品瑕疵数据集由 Zhang 等<sup>[25]</sup>收集, 共 7435 张瑕疵图片, 图片大小统一为  $256 \times 256$ , 图片背景共有 7 种类型, 瑕疵类型分为污渍、褶皱 2 类。本文根据图像背景的复杂程度将 7 种图片分成 Group1、Group2 共 2 组。其中 Group1 为简单纹理图像, 为纯色、粗条纹、细条纹; Group2 为复杂纹理图像, 为点图案、犬牙花纹、格纹及打结花纹。样本纹理图像如图 7 所示。

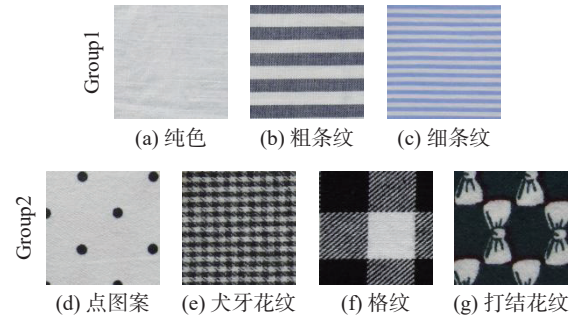


图 7 ZJU-Leaper 纺织品瑕疵数据集

Fig. 7 ZJU-Leaper textile defect dataset

#### 3.1 实验环境及评价指标

本文实验 PyTorch 框架下完成, 软件环境为 PyTorch1.9.1、cuda11.4、Python3.8.12。

本文选择  $\text{mAP}@0.5$  和  $\text{mAP}@0.5:0.95$  作为评价指标,  $\text{mAP}@0.5$  为损失函数 (intersection over union, IoU) 阈值在 0.5 时, 所有类别的平均精度  $A$  (average precision, AP),  $\text{mAP}@0.5:0.95$  为不同 IoU 阈值上平均精度的平均值 (mean average precision, mAP), 下文中 mAP 均表示  $\text{mAP}@0.5:0.95$ 。因此需要计算模型的精确率  $P$  (precision) 和召回率  $R$  (recall)。

$$P = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}$$

$$R = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}$$

$$A = \int_0^1 P(R) dR$$

$$m_{\text{AP}} = \frac{\sum_{i=1}^C A_i}{C}$$

式中:  $N_{\text{TP}}$  (true positive, TP) 为正确预测的正样本个数,  $N_{\text{FP}}$  (false positive, FP) 为错误预测的正样本个数,  $N_{\text{FN}}$  (false negative, FN) 为错误预测的负样本个数。

#### 3.2 消融实验

为了验证本文方法的有效性和泛化性, 本文在 ZJU-Leaper 纺织品瑕疵数据集上进行多个背景

消融实验, 以此来评估多分支离散余弦注意力机制模块 MDCA、尺度聚合模块 SAAM 以及引入上采样模块后改进的 SAAM\_up 模块对纺织品瑕疵检测的效果。Group1、Group2 实验结果见表 1 和表 2。表 1 中 G1\_1 为 Group1 中纯色, G1\_2 为粗

条纹, G1\_3 为细条纹, G1 为 Group1 整体参与训练及测试; 表 2 中 G2\_1 为 Group2 中点图案, G2\_2 为犬牙花纹, G2\_3 为格纹, G2\_4 为打结花纹, G2 为 Group2 整体参与训练及测试。GFLOPs 为每秒浮点运算量, Size 表示模型大小。

表 1 Group1 数据集消融实验结果  
Table 1 Results of Group1 dataset ablation experiment

分组	MDCA	SAAM	SAAM_up	mAP@0.5/%			mAP/%			运算量/GFLOPs	Size/MB
				污渍	褶皱	总	污渍	褶皱	总		
G1_1	—	—	—	80.57	52.04	66.31	45.37	23.16	34.50	4.26	68.00
	√	—	—	79.59	69.10	74.34	47.23	34.61	40.92	4.23	64.97
	√	√	—	78.48	<b>69.46</b>	73.97	47.46	35.28	41.37	8.10	129.81
	√	√	√	<b>81.59</b>	67.70	<b>74.64</b>	<b>48.55</b>	<b>37.48</b>	<b>43.01</b>	8.80	136.33
G1_2	—	—	—	75.13	56.48	65.81	40.61	27.38	34.00	4.26	68.00
	√	—	—	76.32	66.33	71.33	46.34	37.78	42.06	4.23	64.97
	√	√	—	<b>78.02</b>	65.56	71.79	46.11	37.08	41.59	8.10	129.81
	√	√	√	76.28	<b>69.14</b>	<b>72.71</b>	<b>46.58</b>	<b>38.17</b>	<b>42.37</b>	8.80	136.33
G1_3	—	—	—	81.41	63.37	72.39	44.99	31.1	38.05	4.26	68.00
	√	—	—	87.83	68.27	78.05	57.83	35.92	46.87	4.23	64.97
	√	√	—	88.54	67.74	78.14	59.72	35.71	47.72	8.10	129.81
	√	√	√	<b>89.15</b>	<b>70.82</b>	<b>79.98</b>	<b>60.88</b>	<b>37.07</b>	<b>48.98</b>	8.80	136.33
G1	—	—	—	74.34	75.00	64.43	40.88	44.98	33.92	4.26	68.00
	√	—	—	81.29	<b>67.66</b>	74.48	49.51	36.26	42.89	4.23	64.97
	√	√	—	81.67	67.57	74.62	50.11	36.89	43.50	8.10	129.81
	√	√	√	<b>82.88</b>	67.23	<b>75.05</b>	<b>50.16</b>	<b>36.98</b>	<b>43.57</b>	8.80	136.33

表 2 Group2 数据集消融实验结果  
Table 2 Results of Group2 dataset ablation experiment

分组	MDCA	SAAM	SAAM_up	mAP@0.5/%			mAP/%			运算量/GFLOPs	Size/MB
				污渍	褶皱	总	污渍	褶皱	总		
G2_1	—	—	—	76.79	51.98	64.38	47.57	26.60	37.09	4.26	68.00
	√	—	—	77.53	64.63	71.08	48.73	35.58	42.16	4.23	64.97
	√	√	—	78.45	65.63	72.04	48.92	36.65	42.78	8.10	129.81
	√	√	√	<b>79.76</b>	<b>66.57</b>	<b>73.16</b>	<b>50.05</b>	<b>39.22</b>	<b>44.62</b>	8.80	136.33
G2_2	—	—	—	82.92	55.04	68.98	51.79	28.59	40.19	4.26	68.00
	√	—	—	86.53	64.66	75.59	55.69	34.61	45.15	4.23	64.97
	√	√	—	87.30	69.23	78.26	57.29	36.67	46.98	8.10	129.81
	√	√	√	<b>88.28</b>	<b>69.73</b>	<b>79.01</b>	<b>57.55</b>	<b>37.15</b>	<b>47.35</b>	8.80	136.33
G2_3	—	—	—	77.24	60.10	68.67	47.44	29.46	38.45	4.26	68.00
	√	—	—	78.44	64.16	71.30	49.29	33.36	41.32	4.23	64.97
	√	√	—	79.44	64.35	71.89	50.14	35.92	43.03	8.10	129.81
	√	√	√	<b>80.76</b>	<b>69.23</b>	<b>74.99</b>	<b>52.11</b>	<b>37.39</b>	<b>44.75</b>	8.80	136.33

续表 2

分组	MDCA	SAAM	SAAM_up	mAP@0.5/%			mAP/%			运算量/GFLOPs	Size/MB
				污渍	褶皱	总	污渍	褶皱	总		
G2_4	—	—	—	56.31	<b>73.96</b>	65.14	27.41	39.29	33.35	4.26	68.00
	√	—	—	56.34	72.42	64.38	31.69	41.57	36.72	4.23	64.97
	√	√	—	60.78	73.40	67.09	31.32	42.61	36.96	8.10	129.81
	√	√	√	<b>61.08</b>	73.81	<b>67.44</b>	<b>33.31</b>	<b>42.73</b>	<b>38.02</b>	8.80	136.33
G2	—	—	—	68.57	56.85	62.71	39.56	46.78	33.88	4.26	68.00
	√	—	—	75.05	68.99	72.02	47.38	36.18	41.78	4.23	64.97
	√	√	—	74.95	69.67	72.31	47.51	39.00	43.25	8.10	129.81
	√	√	√	<b>75.45</b>	<b>69.77</b>	<b>72.61</b>	<b>47.57</b>	<b>39.52</b>	<b>43.55</b>	8.80	136.33

如表 1、表 2 所示, 本文于 YOLOX 基础上改进, 在骨干网络中加入 MDCA, 7 种背景的纺织品瑕疵检测精度皆有提升, 其中 Group1 的 mAP@0.5% 提升 10.05%, Group2 提升了 9.31%, 提升效果较为明显, 从数据可以看出 MDCA 对于瑕疵特征的提取能力有较大提升; 在特征融合网络中加入 SAAM 模块后, Group1 的 mAP@0.5% 提升 0.14%, Group2 提升了 0.29%; 而在改进原 SAAM 模块为 SAAM\_up 后, 增强了特征信息的有效性, Group1 的 mAP 提升 0.07%, Group2 提升了 0.3%; 可以看出本文的 SAAM 及改进后的 SAAM\_up 模块对检测 Group2 这类背景更为复杂的瑕疵提升能力更大。

图 8 为消融实验特征图对比,  $\times 0$ 、 $\times 1$ 、 $\times 2$  为骨干网络输出的不同尺度特征图。从对比图可以看出 MDCA 模块大大加强了对瑕疵特征的提取能力, 特征图  $\times 1$  和  $\times 2$  中瑕疵特征更为突出; 而在特征融合网络中加入 SAAM 模块后, 特征图  $\times 1$  的瑕疵区域与背景区域的差异对比效果有一个明显的增强; 当引入改进后的 SAAM\_up 后, 特征图中瑕疵区域的边界与原图中瑕疵的边界更加吻合, 大大提升了检测精度。

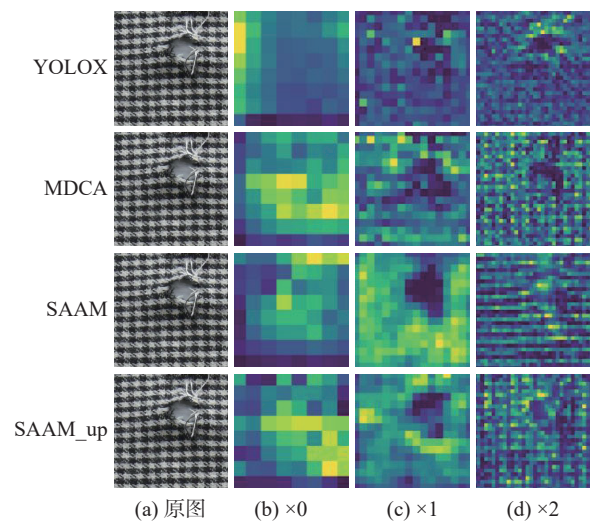


图 8 消融实验特征图对比

Fig. 8 Comparison of feature maps in ablation experiments

### 3.3 对比实验

为了进一步衡量本文模型对于纺织品瑕疵检测的性能, 基于 ZJU-Leaper 纺织品瑕疵数据集, 将本文模型与 faster\_RCNN<sup>[9]</sup>、cascade\_RCNN<sup>[12]</sup>、YOLOv5<sup>[26]</sup>、YOLOv7<sup>[27]</sup>、PPYOLOE<sup>[28]</sup>、YOLOX<sup>[14]</sup>、YOLOv7<sup>[29]</sup> 和 YOLOv6<sup>[30]</sup> 等现有目标检测模型进行对比。实验图片输入大小统一为  $256 \times 256$ , 实验结果见表 3 及表 4。

表 3 Group1 数据集实验对比

Table 3 Experimental comparison of Group1 dataset

模型	mAP@0.5/%			mAP/%			运算量/GFLOPs	Size/MB	传输速率/(f/s)
	污渍	褶皱	总	污渍	褶皱	总			
faster_RCNN	80.90	62.80	71.80	44.40	29.60	37.00	131.40	158.40	65.20
cascade_RCNN	81.20	62.10	71.70	45.40	30.30	37.80	158.40	265.00	55.60
YOLOv5_s	78.24	59.27	68.66	48.15	30.50	39.36	15.80	<b>13.70</b>	101.90
YOLOX	81.69	61.29	71.49	47.13	32.95	40.04	<b>4.26</b>	68.00	214.20

续表 3

模型	mAP@0.5/%			mAP/%			运算量/GFLOPs	Size/MB	传输速率/(f/s)
	污渍	褶皱	总	污渍	褶皱	总			
YOLOR_p6	82.77	61.31	72.04	49.01	34.29	41.65	81.56	142.68	122.50
YOLOR_csp	81.68	64.16	72.92	50.75	35.83	43.29	120.63	202.36	141.80
PPYOLOE	77.20	64.80	71.00	44.10	33.50	38.80	17.36	58.53	79.22
YOLOv7	<b>84.03</b>	62.89	73.46	48.60	32.30	40.45	103.30	71.30	106.70
YOLOv6_s	79.60	63.10	71.40	<b>52.00</b>	35.80	<b>43.90</b>	44.11	36.25	<b>359.07</b>
SAAM_YOLOX	82.88	<b>67.23</b>	<b>75.05</b>	50.16	<b>36.98</b>	43.57	8.80	136.33	257.65

表 4 Group2 数据集实验对比  
Table 4 Experimental comparison of Group2 dataset

模型	mAP@0.5/%			mAP/%			运算量/GFLOPs	Size/MB	传输速率/(f/s)
	污渍	褶皱	总	污渍	褶皱	总			
faster_RCNN	77.20	67.00	72.10	42.50	32.90	37.70	131.40	158.40	65.20
cascade_RCNN	77.30	66.00	71.60	44.60	33.70	39.10	158.40	265.00	55.60
YOLOv5_s	72.64	61.15	66.83	45.19	32.46	38.92	15.80	<b>13.70</b>	101.90
YOLOX	75.35	63.83	69.59	45.61	34.39	40.00	<b>4.26</b>	68.00	214.20
YOLOR_p6	76.34	63.90	70.12	47.43	35.79	41.61	81.56	142.68	122.50
YOLOR_csp	75.83	66.13	70.98	48.02	37.58	42.80	120.63	202.36	141.80
PPYOLOE	75.50	62.70	69.10	43.80	32.60	38.20	17.36	58.53	79.22
YOLOv7	<b>77.51</b>	65.49	71.50	47.03	33.79	40.41	103.30	71.30	106.70
YOLOv6_s	73.90	65.10	69.50	<b>48.80</b>	38.10	43.40	44.11	36.25	<b>359.07</b>
SAAM_YOLOX	75.45	<b>69.77</b>	<b>72.61</b>	47.57	<b>39.52</b>	<b>43.55</b>	8.80	136.33	257.65

如表 3, 在 Group1 中本文模型在保证速度的同时在检测精度 mAP@0.5 中对比其他模型达到最高值 75.05%, 特别是褶皱类这种与背景难以区分、检测较难的瑕疵检精度更是达到 67.23%; 而在表 4 中, 本文模型不管是 mAP@0.5 还是 mAP 都达到最高值, 分别为 72.61%、43.55%。从数据可分析出本文模型在背景更加复杂的情况下检测效果比其他模型更加出色。

本文模型在 2 个组别中都具有较强的检测效果, 且模型尺度小, 有较强的适用性; 对于多种纹理背景仍具备较高的检测精度, 具备较强的泛化性。

图 9 给出了本文模型与其他模型的检测结果对比, 其中 YOLOR 模型复杂度较低, 对背景干扰的抵抗能力不足, 无法很好地捕捉到瑕疵的特征信息, 导致漏检问题的出现; YOLOv5\_s 感受野不足, 模型无法充分理解和捕捉到图像中较远的对象或较大的瑕疵, 从而导致部分瑕疵未被完整地

检测到, 存在检测不完整的问题; YOLOv7 网络架构没有充分考虑到纺织品瑕疵检测任务的复杂性, 提取到的瑕疵特征表达能力不足, 导致漏检、检测不完全的问题; 原 YOLOX 对于纺织品瑕疵的尺度适配性不佳, 在犬牙花纹与格纹背景出现误检和漏检情况; YOLOv6\_s 检测效果较好, 但模型对于小目标瑕疵和瑕疵边缘的信息提取不完全, 在小目标瑕疵以及瑕疵边缘检测效果弱于本文模型。可以看出本文模型在检测效果上超越其他模型。针对犬牙花纹与格纹背景的瑕疵图, 图 10 给出了本文模型与其他模型的检测对比, 表 5 为相对应的精度数据。

从图 10 和表 5 可以看到, YOLOv5\_s、YOLOX 和 YOLOv7 受到背景影响较大, 存在大量的误检、检测不完整等问题, 表现在数据中为 mAP 值总体偏低; 而 YOLOv6\_s 与本文模型 SAAM\_YOLOX 检测效果较好、精度较高, 其中 YOLOv6 在小目标及瑕疵边缘部分未能达到本文模型的检测效果。

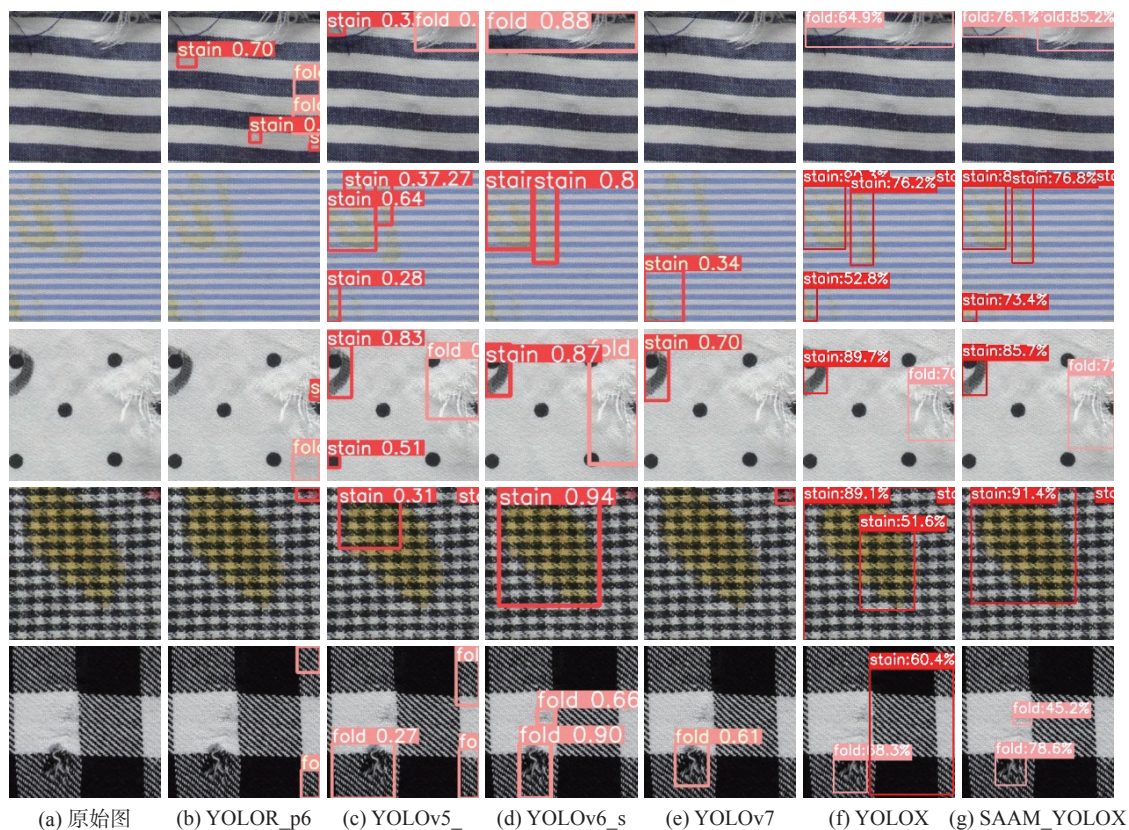


图 9 检测结果对比

Fig. 9 Comparison of test results

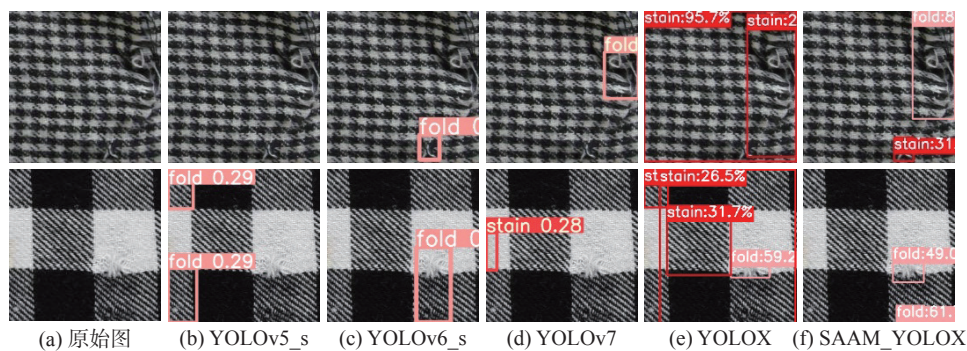


图 10 犬牙花纹与格纹背景的检测结果对比

Fig. 10 Comparison of detection results between houndstooth and gingham backgrounds

表 5 针对犬牙花纹与格纹背景的实验对比

Table 5 Experimental comparison between houndstooth and gingham backgrounds

%

模型	mAP@0.5			mAP		
	污渍	褶皱	总	污渍	褶皱	总
YOLOv5_s	65.20	39.80	52.50	38.20	15.70	26.90
YOLOv6_s	82.50	64.10	73.30	<b>55.80</b>	36.70	<b>46.30</b>
YOLOv7	50.90	52.0	51.50	29.80	23.50	26.70
YOLOX	70.16	36.13	53.14	26.43	13.52	20.47
SAAM_YOLOX	<b>84.62</b>	<b>68.88</b>	<b>76.75</b>	54.22	<b>37.58</b>	45.90

## 4 结束语

本文针对纺织品瑕疵中针对犬牙花纹与格纹背景出现的误检和漏检、整体精度较差的问题,提出 SAAM-YOLOX 纺织品瑕疵检测算法。首先,在骨干网络中引入 MDCA 模块来加强网络对纺织品瑕疵特征的提取能力,改善原 YOLOX 在特定背景中误检、漏检情况;其次,在特征融合网络加入 SAAM 模块,充分感知瑕疵区域与背景区域的不同特征信息;最后,在 SAAM 模块的上采样阶段引入双线性插值与自注意力机制,在尺度变换时采用改进后的上采样模块以保证特征信息的有效性,最终实现提高检测效果。经实验对比,本文模型在纺织品瑕疵检测中优于其他主流目标检测模型。

## 参考文献:

- [1] 胡娜,马慧,湛涛.融合 LBP 纹理特征与 B2DPCA 技术的手指静脉识别方法[J].智能系统学报,2019,14(3): 533-540.  
HU Na, MA Hui, ZHAN Tao. Finger vein recognition method combining LBP texture feature and B2DPCA technology[J]. CAAI transactions on intelligent systems, 2019, 14(3): 533-540.
- [2] 狄岚,赵树志,何锐波.基于光照预处理与特征提取的纺织品瑕疵检测方法[J].智能系统学报,2019,14(4): 716-724.  
DI Lan, ZHAO Shuzhi, HE Ruibo. Fabric defect inspection based on illumination preprocessing and feature extraction[J]. CAAI transactions on intelligent systems, 2019, 14(4): 716-724.
- [3] 马明寅,狄岚,梁久祯.基于图像校正和模板分割的纺织品瑕疵检测[J].南京大学学报(自然科学版),2021,57(1): 29-41.  
MA Mingyin, DI Lan, LIANG Jiuzhen. Fabric defect detection based on image correction and template segmentation[J]. Journal of Nanjing University (natural science edition), 2021, 57(1): 29-41.
- [4] 龙涵彬,狄岚,梁久祯.基于畸变校正与视觉显著特征的纺织品瑕疵检测[J].模式识别与人工智能,2020,33(12): 1122-1134.  
LONG Hanbin, DI Lan, LIANG Jiuzhen. Fabric defect detection based on distortion correction and visual salient features[J]. Pattern recognition and artificial intelligence, 2020, 33(12): 1122-1134.
- [5] WU Xiongwei, SAHOO D, HOI S C H. Recent advances in deep learning for object detection[J]. *Neurocomputing*, 2020, 396: 39-64.
- [6] 赵永强,饶元,董世鹏,等.深度学习目标检测方法综述[J].中国图象图形学报,2020,25(4): 629-654.  
ZHAO Yongqiang, RAO Yuan, DONG Shipeng, et al. Survey on deep learning object detection[J]. Journal of image and graphics, 2020, 25(4): 629-654.
- [7] ROSS Girshick, JEFF Donahue, TREVOR Darrell, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE Computer Society, 2014: 580-587.
- [8] ROSS G. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago: IEEE Computer Society, 2015: 1440-1448.
- [9] 赵文清,程幸福,赵振兵,等.注意力机制和 Faster RCNN 相结合的绝缘子识别[J].智能系统学报,2020,15(1): 92-98.  
ZHAO Wenqing, CHENG Xingfu, ZHAO Zhenbing, et al. Insulator recognition based on attention mechanism and Faster RCNN[J]. CAAI transactions on intelligent systems, 2020, 15(1): 92-98.
- [10] HE Xinying, WU Liming, SONG Feiyang, et al. Research on fabric defect detection based on deep fusion DenseNet-SSD network[C]//Proceedings of the International Conference on Wireless Communication and Sensor Networks. Warsaw Poland: Association for Computing Machinery, 2020: 60-64.
- [11] CAO Guimei, XIE Xuemei, YANG Wenzhe, et al. Feature-fused SSD: fast detection for small objects[C]//Ninth International Conference on Graphic and Image Processing. Qingdao: International Society for Optics and Photonics, 2018: 106151E.
- [12] 王晓林,苏松志,刘晓颖,等.一种基于级联神经网络的飞机检测方法[J].智能系统学报,2020,15(4): 697-704.  
WANG Xiaolin, SU Songzhi, LIU Xiaoying, et al. Cascade convolutional neural networks for airplane detection[J]. CAAI transactions on intelligent systems, 2020, 15(4): 697-704.
- [13] 陈丽,马楠,逢桂林,等.多视角数据融合的特征平衡 YOLOv3 行人检测研究[J].智能系统学报,2021,16(1): 57-65.  
CHEN Li, MA Nan, PANG Guilin, et al. Research on multi-view data fusion and balanced YOLOv3 for pedestrian detection[J]. CAAI transactions on intelligent systems, 2021, 16(1): 57-65.
- [14] GE Zheng, LIU Songtao, WANG Feng, et al. YoloX: Exceeding yolo series in 2021[EB/OL]. (2021-07-18) [2023-04-22]. <https://arxiv.org/abs/2107.08430>.
- [15] 毛明毅,吴晨,钟义信,等.加入自注意力机制的 BERT 命名实体识别模型[J].智能系统学报,2020,

- 15(4): 772–779.
- MAO Mingyi, WU Chen, ZHONG Yixin, et al. BERT named entity recognition model with self-attention mechanism[J]. CAAI transactions on intelligent systems, 2020, 15(4): 772–779.
- [16] 王凤随, 陈金刚, 王启胜, 等. 自适应上下文特征的多尺度目标检测算法 [J]. 智能系统学报, 2022, 17(2): 276–285.
- WANG Fengsui, CHEN Jingang, WANG Qisheng, et al. Multi-scale target detection algorithm based on adaptive context features[J]. CAAI transactions on intelligent systems, 2022, 17(2): 276–285.
- [17] 王召新, 续欣莹, 刘华平, 等. 基于级联宽度学习的多模态材质识别 [J]. 智能系统学报, 2020, 15(4): 787–794.
- WANG Zhaoxin, XU Xinying, LIU Huaping, et al. Cascade broad learning for multi-modal material recognition[J]. CAAI transactions on intelligent systems, 2020, 15(4): 787–794.
- [18] 杜艳玲, 王丽丽, 黄冬梅, 等. 融合密集特征金字塔的改进 R2CNN 海洋涡旋自动检测 [J]. 智能系统学报, 2023, 18(2): 341–351.
- DU Yanling, WANG Lili, HUANG Dongmei, et al. Improved R2CNN ocean eddy automatic detection with a dense feature pyramid[J]. CAAI transactions on intelligent systems, 2023, 18(2): 341–351.
- [19] ZHAO Zhiyong, GUI Kang, WANG Peimao. Fabric defect detection based on cascade faster R-CNN[C]//Proceedings of the 4th International Conference on Computer Science and Application Engineering. Sanya: Association for Computing Machinery, 2020: 1–6.
- [20] LI Feng. Bag of tricks for fabric defect detection based on Cascade R-CNN[J]. Textile research journal, 2021, 91(5-6): 599–612.
- [21] JING Junfeng, ZHUO Dong, ZHANG Huanhuan, et al. Fabric defect detection using the improved YOLOv3 model[J]. Journal of engineered fibers and fabrics, 2020, 15: 1558–9250.
- [22] ZHENG Liaomo, WANG Xiaojie, WANG Qi, et al. A fabric defect detection method based on improved yolov5[C]//2021 7th International Conference on Computer and Communications. Chengdu: Institute of Electrical and Electronics Engineers, 2021: 620–624.
- [23] 伊力哈木·亚尔买买提. 一种新融合算法的维吾尔族人脸识别 [J]. 智能系统学报, 2018, 13(3): 431–436.
- YILIHAMU·Yaermaimaiti. A new fusion algorithm for uyghur face recognition[J]. CAAI transactions on intelligent systems, 2018, 13(3): 431–436.
- [24] PANG Youwei, ZHAO Xiaoqi, XIANG Tianzhu, et al. Zoom in and out: A mixed-scale triplet network for camouflaged object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: Institute of Electrical and Electronics Engineers, 2022: 2160–2170.
- [25] ZHANG Chenkai, FENG Shaozhe, WANG Xulongqi, et al. ZJU-leaper: a benchmark dataset for fabric defect detection and a comparative study[J]. IEEE transactions on artificial intelligence, 2020, 1(3): 219–232.
- [26] ZHAO Ziyu, YANG Xiaoxia, ZHOU Yucheng, et al. Real-time detection of particleboard surface defects based on improved YOLOV5 target detection[J]. Scientific reports, 2021, 11(1): 1–15.
- [27] WANG Chienyao, YEH I-hau, MARK Liaohongyuan. You only learn one representation: unified network for multiple tasks[EB/OL]. (2021–05–10)[2023–04–22]. <https://arxiv.org/abs/2105.04206>.
- [28] XU Shangliang, WANG Xinxin, LYU Wenyu, et al. PP-YOLOE: an evolved version of YOLO[EB/OL]. (2022–03–30)[2023–04–22]. <https://arxiv.org/abs/2203.16250>.
- [29] WANG Chienyao, BOCHKOVSKIY A, LIAO H Y M, et al. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Conference on Computer Vision and Pattern Recognition 2023. Vancouver: IEEE Computer Society, 2023: 7464–7475.
- [30] LI Chuyi, LI Lulu, JIANG Hongliang, et al. YOLOv6: A single-stage object detection framework for industrial applications[EB/OL]. (2022–09–07)[2023–04–22]. <https://arxiv.org/abs/2209.02976>.

#### 作者简介:



周在雍, 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: 6213113157@stu.jiangnan.edu.cn。



狄岚, 副教授, 主要研究方向为数字图像处理、计算机视觉、神经网络算法及其应用研究。申请国家发明专利 10 余项, 发表学术论文 50 余篇。  
E-mail: dilan@jiangnan.edu.cn。