

# 基于GAN-Data数据增强与FD-DETR的织物疵点检测方法

陈孟元\*, 张坦坦, 唐哲

(安徽工程大学高端装备先进感知与智能控制教育部重点实验室, 安徽芜湖 241000)

**摘要:** 针对纺织工业实际生产中织物疵点样本获取困难且数量稀缺导致深度学习模型应用受限的问题, 本文提出了一种将上游数据增强与下游检测模型深度优化相结合的织物疵点检测方法。本文在上游数据增强阶段基于循环生成对抗网络(Cycle-consistent Generative Adversarial Network, Cycle-GAN)提出GAN-Data生成网络, 该网络利用掩码(Mask)引导机制实现了疵点特征与背景纹理的解耦处理, 在精准控制生成位置的同时解决了基准模型中疵点分布随机性过大的问题。为了应对织物疵点尺度差异显著的挑战, GAN-Data设计了增强疵点生成模块(Enhanced Defect Generation Module, EDGM), 通过四个并行的多尺度膨胀卷积分支使感受野能够灵活覆盖从1~17像素的范围, 实现了对点状、线状、块状及大面积疵点的自适应特征提取。同时, 本文针对背景纹理失真问题引入了基于VGG19网络的纹理保持损失函数, 确保了非疵点区域周期性纹理的完整性。在下游优化阶段, 本文在RT-DETR的基础上构建了FD-DETR检测网络, 在主干网络嵌入了基于Prewitt算子的四方向边缘增强模块以强化弱疵点轮廓捕获能力, 并设计了稀疏注意力机制(Sparse Attention-based Intra-scale Feature Interaction, SparseAIFI), 通过融合局部窗口、跨步采样及块级模式降低模型复杂度。此外, FD-DETR引入了长宽比感知IoU损失函数(Angle Ratio Aware-Intersection over Union, ARA-IoU), 通过中心距离归一化与自适应权重机制优化了不规则疵点的定位精度。实验部分结合了MVTec AD公开数据集、Industrial Textile Dataset (ITD)数据集及真实生产线自建数据集进行了多维度验证。本文首先在MVTec AD公开数据集上通过结构相似性指标(Structural Similarity Index Measure, SSIM)、峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)及弗雷歇初始距离(Fréchet Inception Distance, FID)等指标验证了GAN-Data生成的图像质量优于主流方法, 并通过工业织物数据集(Industrial Textile Dataset, ITD)验证了模型在多材质背景下的跨域泛化能力。随后, 通过在多个数据集上的对比实验证明, 以GAN-Data增强数据训练的FD-DETR模型在织物疵点识别上显著领先。最后, 通过协同优化实验的F1分数曲线分析证实, 相较于单一改进, GAN-Data与FD-DETR的结合在小样本场景下具有更快的收敛速度与更高的性能上限, 验证了上下游协同路线的可行性。

**关键词:** 织物疵点检测; 数据集增强; 生成对抗网络; 深度学习

**基金项目:** 安徽省重点研究与开发计划项目(No. 202304a05020073); 安徽省高校杰出青年科研项目(No. 2022AH020065); 安徽省高校哲学社会科学基金项目(No. 2023AH050881)

**中图分类号:** TP751.1; TP183 **文献标识码:** A **文章编号:** 0372-2112(2026)01-0417-16

**电子学报URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20250899

## Fabric Defect Detection Method Based on GAN-Data Augmentation and Improved RT-DETR

CHEN Mengyuan\*, ZHANG Tantan, TANG Zhe

(Key Laboratory of Advanced Perception and Intelligent Control of High-end Equipment, Ministry of Education, Anhui Polytechnic University, Wuhu, Anhui 241000, China)

**Abstract:** To address the limitations of deep learning models in the textile industry caused by the scarcity of fabric defect samples, this paper proposes a method combining upstream data augmentation with downstream detection model deep optimization. In actual production, the extreme scarcity of defect samples results in a "small-sample dilemma" that hinders model training. During the upstream stage, a GAN-Data generative network is designed based on the cycle-consistent generative adversarial network (Cycle-GAN) architecture. This network utilizes a mask-guided mechanism to decouple defect features from background textures, ensuring precise positioning and resolving distribution randomness. To handle significant scale variations, GAN-Data incorporates an enhanced defect generation module (EDGM), which employs parallel multi-scale dilated convolution branches to achieve adaptive feature extraction for various defect types. Furthermore, a tex-

ture-preservation loss function based on the VGG19 network and Gram Matrix constraints is introduced to maintain the integrity of periodic fabric textures in non-defect regions. In the downstream stage, the FD-DETR detection network is constructed. Its backbone embeds a four-directional edge enhancement module based on the Prewitt operator to strengthen the capture of weak defect contours. To improve efficiency, a sparse attention-based intra-scale feature interaction (SparseAIFI) mechanism is designed, which effectively reduces the computational complexity by fusing local window, striped sampling, and block-level sparse patterns. Additionally, an aspect ratio aware-IoU (ARA-IoU) loss function is introduced to optimize the localization accuracy for irregular defects through center distance normalization and an adaptive weighting mechanism. The method is validated using the MVTec AD dataset, the industrial textile dataset (ITD), and a self-built production line dataset. Initial evaluations using the structural similarity index measure (SSIM), peak signal-to-noise ratio (PSNR), and Fréchet inception distance (FID) demonstrate that GAN-Data achieves superior image quality and cross-domain generalization. Subsequent comparative experiments show that the FD-DETR model trained with GAN-augmented data significantly improves detection accuracy while meeting industrial requirements. Finally, collaborative optimization experiments confirm that integrating GAN-Data and FD-DETR achieves faster convergence and higher performance ceilings than single-stage improvements. In conclusion, this bidirectional synergistic route provides an efficient solution for fabric defect detection under small-sample conditions.

**Keywords:** fabric defect detection; dataset augmentation; generative adversarial network; deep learning

**Foundation Item(s):** Anhui Provincial Key Research and Development Program (No. 202304a05020073); Anhui Provincial University Distinguished Young Scholars Research Project (No. 2022AH020065); Philosophy and Social Science Research Project of Anhui Higher Education Institutions (No. 2023AH050881)

## 0 引言

织物疵点是纺织品生产期间,受原料、设备及工艺多重因素干扰产生的布面质量缺陷,此类问题长期制约纺织行业产品质量控制效率<sup>[1]</sup>。近年来,基于深度学习的织物疵点检测技术突破了传统机器学习依赖人工设计算法的局限,但其实际应用仍面临关键瓶颈:上游的高质量训练样本获取与下游的精准检测模型构建。

在上游数据层面,训练样本数量稀缺是制约下游检测模型性能的首要因素。工业生产中的高良品率使得疵点样本极其稀少,实际收集的数据往往仅有几张图像,呈现离散的“点”状分布,无法呈现其他领域具有统计意义的“面”状分布<sup>[2]</sup>。现有公开织物疵点数据集虽然在样本数量上较为充足,但织物材质、纹理结构、工艺参数等因素导致公开数据集与特定应用场景之间存在明显的域间分布偏移,下游检测模型在实际应用时检测性能会显著衰减。因此,数据增强成为研究人员提升下游检测模型性能的常用手段。传统数据增强方法如随机裁剪、翻转、对比度调整等,本质上仍依赖对现存图像的二次应用,难以生成新的疵点特征<sup>[3]</sup>。生成对抗网络(Generative Adversarial Network, GAN)为数据增强提供了新思路<sup>[4-5]</sup>,2017年Zhu等人<sup>[6]</sup>提出的循环一致性生成对抗网络通过无监督学习实现了两个不同域之间图像的相互转换,并在多个领域取得广泛应用<sup>[7-8]</sup>。随后Duan等人<sup>[9]</sup>在Style-GAN2的基础上提出DFM-GAN,通过缺陷感知特征操作实现不同领域疵点图像生成。这些方法在织物疵点生成应用中仍存在局限:固定卷积核难以适应

多尺度的疵点变化,易将图像中小尺寸疵点识别为纹理噪声进行修复而非生成;同时缺乏纹理约束,生成过程中易将正常织物纹理视为干扰信号,导致背景纹理失真,进而影响生成样本的真实性<sup>[10]</sup>。

在下游检测层面,数据增强虽能缓解样本稀缺问题,但如何有效利用这些样本进行精准检测仍是关键<sup>[11]</sup>。基于深度学习的检测模型为织物疵点识别提供了技术支撑。YOLO系列通过单阶段框架实现了检测速度的突破<sup>[12]</sup>;DETR将Transformer引入目标检测领域实现端到端检测<sup>[13]</sup>,Deformable DETR通过可变形注意力机制改善了小目标检测效果<sup>[14]</sup>,RT-DETR进一步实现了实时处理能力<sup>[15]</sup>。这些检测模型经过充足的样本训练后,在织物疵点检测领域展现出良好的应用前景,但在织物实际生产检测中仍存在三个问题:首先,部分织物疵点边缘特征微弱且易被复杂背景纹理干扰,传统主干网络难以有效区分疵点与正常纹理,导致疵点特征的提取与表达能力不足<sup>[16]</sup>;其次,工业生产线需要处理大量的织物图像,但现有模型面临计算开销大、内存消耗高的问题,难以满足实时检测需求<sup>[17]</sup>;最后,部分织物疵点呈现不规则形状且长宽比变化显著,传统IoU(Intersection over Union)损失函数对这类目标的边界框回归精度低、收敛速度慢,影响检测准确性<sup>[18]</sup>。单纯依靠模型改进仍难以突破性能瓶颈。

基于上述问题,本文通过上下游模型相结合的方式来提升织物疵点检测技术性能。在上游数据增强方面,设计GAN-Data生成网络,通过Mask引导机制、多尺度膨胀卷积模块和纹理保持损失函数,实现高质

量疵点样本生成,有效解决背景失真和多尺度疵点生成难题。在下游模型优化方面,构建 FD-DETR 检测模型,针对织物疵点边缘特征微弱、不规则形状边界框回归精度低、检测模型图像计算开销大等问题,分别提出多尺度边缘增强主干网络、稀疏注意力机制和长宽比感知 IoU 损失函数。通过上下游双向协同,实现织物疵点检测性能的显著提升。

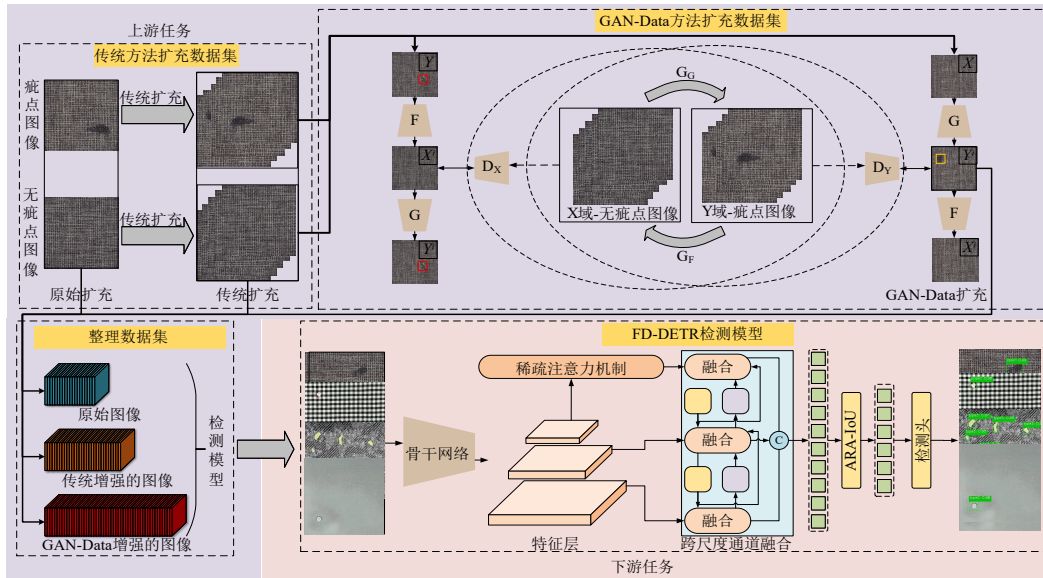
### 1 系统总体框架设计

GAN-Data 数据增强与 FD-DETR 的织物疵点检测方法框架如图 1 所示。传统数据增强方法虽能扩充样本数量,但生成样本的质量和多样性受限,而先进的检测模型如果仅依赖有限的真实数据训练,容易陷入过拟合,导致泛化性能下降。因此,本文采用上下游协同的思路:上游通过生成对抗网络进行数据增

强,下游通过优化检测模型来提升织物疵点检测性能。传统增强图像和 GAN-Data 生成的图像共同构成“面”状分布的数据集。

下游检测模型的核心任务是利用上游整理的数据集对织物疵点特征进行学习。因此,本文针对织物疵点检测中的前述三个问题,在主干网络、注意力机制和损失函数三个方面进行了优化设计,使得模型在充足样本训练的基础上更进一步提升检测性能。

上游任务中生成的样本为下游模型提供了充足的训练数据,下游检测模型的优化验证了上游增强的有效性,相较于单独开展数据增强或模型改进的方案,上游与下游的协同设计能够更有效地突破小样本条件下的检测性能瓶颈。下面将详细阐述各部分的核心模块与优化策略。



注:G<sub>C</sub>和G<sub>F</sub>为风格转换的生成器;X和Y为原始数据集;X<sup>i</sup>和Y<sup>i</sup>为翻译后图像;D<sub>X</sub>和D<sub>Y</sub>为判别器;G为正向生成器,F为反向生成器。

图 1 系统总体框图

Figure 1 Overall system block diagram

### 2 GAN-Data 数据增强

GAN-Data 是在 Cycle-GAN 网络的基础上改进特征提取、纹理损失与分区处理的新模型。模型的整体结构如图 2 所示,包含生成器 G、F 和判别器 D<sub>X</sub>、D<sub>Y</sub>,通过对抗学习实现无织物疵点图像与织物疵点图像之间的双向转换,以此生成高质量的织物疵点图像。原 Cycle-GAN 网络在生成织物疵点图像时,存在疵点生成位置随机、疵点特征提取不足、织物纹理不真实等问题。首先,在生成器中融入增强疵点生成模块;其次,在原有的损失函数基础上引入基于 VGG19 的纹理保持损失函数,专门保护非疵点区域的织物纹理

特征,避免纹理失真;最后,通过 Mask 掩码进行疵点生成与分区处理。

#### 2.1 生成器优化

GAN-Data 生成器框架如图 2 右侧所示。第一部分为 Mask 引导机制。传统 Cycle-GAN 生成器仅接收单一 RGB 图像输入,而 GNA-Data 采用扩展输入形式。图 2 中生成器输入框中的图像采用 RGB 与 Mask 叠加方式可视化展示,实际网络输入为两者的通道拼接,如式(1)所示:

$$X_{input} = \begin{cases} I_{fabric} \in \mathbb{R}^{H \times W \times 3} \\ M_{defect} \in \mathbb{R}^{H \times W \times 1} \end{cases} \quad (1)$$



码。 $B(x, y)$  通过高斯函数进行平滑处理,生成软边界引导信息:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\mu_x)^2 + (y-\mu_y)^2}{2\sigma^2}} \quad (8)$$

$$\mathbf{F}_{\text{boundary}}(x, y) = B(x, y) \otimes G(x, y) \quad (9)$$

其中,  $\otimes$  表示卷积操作;  $G(x, y)$  表示位置  $(x, y)$  处的高斯函数值;  $\mu_x$  和  $\mu_y$  分别表示高斯函数的中心坐标, 对应于边界像素的位置索引;  $\sigma$  为高斯函数的标准差, 控制高斯函数的衰减速度和影响范围, 系数  $\frac{1}{2\pi\sigma^2}$  为归一化因子, 确保高斯函数在整个二维空间的积分为 1, 保证生成的软边界引导信息的数值稳定性。经过高斯平滑后的边界特征  $\mathbf{F}_{\text{boundary}}$  不仅保留了疵点的边界位置信息, 还通过平滑过渡减少了硬边界带来的伪影, 为后续注意力机制提供边界引导。织物表面具有复杂的纹理结构, 包括经纬线交织模式、色彩变化、光照反射等, 这些背景特征容易与疵点特征产生混淆, 影响生成质量。引入注意力机制通过学习空间权重分布, 网络能够区分织物背景纹理和疵点区域, 重点关注需要生成疵点的区域。注意力权重的计算结合了多尺度特征和边界信息:

$$A = \sigma_1 \left( \text{Conv3} \times 3 \left( [\text{Conv1} \times 1(\mathbf{F}_{\text{fused}}), \mathbf{F}_{\text{boundary}}] \right) \right) \quad (10)$$

其中,  $\sigma_1$  为 Sigmoid 激活函数, 确保权重值在  $[0, 1]$  范围内。权重分布  $A(x, y)$  接近 1 的区域表示网络应重点关注的疵点生成区域, 接近 0 的区域表示应保持原有织物纹理区域。最终的特征增强通过空间自适应加权和残差连接实现, 整个模块在编码器产生的通道特征基础上进行处理, 为后续的疵点生成变换提供包含空间引导信息的增强特征, 从而实现从随机疵点生成向 Mask 掩码控制高质量疵点生成的转变。

## 2.2 损失函数改进

传统的 Cycle-GAN 损失函数仅依靠对抗损失  $\mathcal{L}_{\text{GAN}}$ 、循环一致性损失  $\mathcal{L}_{\text{cyc}}$  与身份损失  $\mathcal{L}_{\text{id}}$ , 其中对抗损失函数  $\mathcal{L}_{\text{GAN}}$  用于评估生成样本的真实性,  $\mathcal{L}_{\text{cyc}}$  确保转换的可逆性,  $\mathcal{L}_{\text{id}}$  确保在循环一致性中生成的图像能够保持源图像的某些特征, 以上损失函数基本满足了 Cycle-GAN 生成图像方面的要求, 但生成的图像存在大量噪声, 因此引入 Mask 与 VGG19 网络的纹理保持损失  $\mathcal{L}_{\text{Texture}}$  来保持织物纹理效果的一致性, 如式(11)所示:

$$\mathcal{L}_{\text{Texture}} = \frac{1}{N} \sum_{i=1}^N \left\| G_i(G(x, M_{\text{defect}}) \odot (1 - M_{\text{defect}})) - G_i(x \odot (1 - M_{\text{defect}})) \right\|_{\text{F}}^2 \quad (11)$$

其中,  $N=4$  为选择的 VGG19 网络特征层数量;  $G(x, M_{\text{defect}})$  表示生成器输出的完整图像;

$G(x, M_{\text{defect}}) \odot (1 - M_{\text{defect}})$  表示提取生成图像的背景区域;  $x \odot (1 - M_{\text{defect}})$  表示提取原始图像的背景区域;  $G_i$  为通过 VGG19 网络第  $i$  层特征图计算的 Gram 矩阵函数;  $\|\cdot\|_{\text{F}}$  表示 Frobenius 范数。该损失函数通过约束生成图像背景区域与原始图像背景区域在 VGG19 特征空间的纹理一致性, 确保背景纹理不失真。Gram 矩阵用于捕获特征图中不同通道之间的相关性, 能够有效表征图像的纹理特征。对于 VGG19 网络第  $i$  层提取的特征图, 其 Gram 矩阵计算公式如式(12)所示:

$$G_i(\phi_i(I))_{c_1, c_2} = \frac{1}{H_i \times W_i} \sum_{h=1}^{H_i} \sum_{w=1}^{W_i} \phi_i(I)(c_1, h, w) \cdot \phi_i(I)(c_2, h, w) \quad (12)$$

其中,  $\phi_i(I)$  表示输入图像  $I$  经过 VGG19 网络在第  $i$  层提取的特征图, 其维度为  $C_i \times H_i \times W_i$ ;  $C_i$  为该层的通道数;  $H_i$  和  $W_i$  分别为特征图的高度和宽度;  $c_1, c_2 \in \{1, 2, \dots, C_i\}$  表示特征图的通道索引; Gram 矩阵  $G_i(\phi_i(I))$  的维度为  $C_i \times C_i$ , 其元素  $G_i(\phi_i(I))_{c_1, c_2}$  表示通道  $c_1$  和通道  $c_2$  之间的相关性。通过计算不同 VGG19 层的 Gram 矩阵差异, 可以约束生成图像在多个尺度上保持与原图一致的背景纹理。最终的总损失函数  $\mathcal{L}_{\text{total}}$  可以结合上述的对抗损失、循环一致性损失、身份损失与纹理保持损失, 如式(13)所示:

$$\mathcal{L}_{\text{total}}(G, D, F, X, Y) = \mathcal{L}_{\text{GAN}} + \lambda_{\text{cyc}} \mathcal{L}_{\text{cyc}} + \lambda_{\text{id}} \mathcal{L}_{\text{id}} + \lambda_{\text{Texture}} \mathcal{L}_{\text{Texture}} \quad (13)$$

各损失项的权重参数  $\lambda_{\text{cyc}}=12$ 、 $\lambda_{\text{id}}=1.0$ 、 $\lambda_{\text{Texture}}=0.05$ 。

## 3 FD-DETR 网络

织物疵点具有尺寸微小、边缘特征微弱、易受纹理背景干扰等特点。针对上述问题, 本文提出 FD-DETR 检测模型。在主干网络层面, 设计基于 Prewitt 算子的多尺度边缘增强 ResNet 网络, 通过四个方向的边缘检测算子, 增强模型对疵点边缘特征的提取能力。在编码器层面, 引入稀疏注意力机制, 采用局部窗口、跨步和分块的混合稀疏模式, 在降低计算复杂度的同时减少复杂背景纹理的干扰, 提升模型对疵点区域的聚焦能力。在损失函数层面, 针对织物疵点边界框回归精度不足的问题, 设计 ARA-IoU 损失函数替代传统 IoU 损失, 通过中心距离归一化、长宽比感知和自适应权重机制优化定位精度。模型整体架构如图 4 所示。

### 3.1 多尺度边缘增强 ResNet 主干网络

传统 ResNet 网络在提取织物疵点特征时, 疵点边缘信息微弱且容易被复杂背景纹理干扰, 导致特征表达能力不足。为了增强模型对疵点边缘特征的提取能力, 本文设计了多尺度边缘增强 ResNet 主干网络。该网络在 ResNet 的关键特征提取层级嵌入 Prewitt 四向边缘增强模块, 通过水平、垂直、对角和反对角四

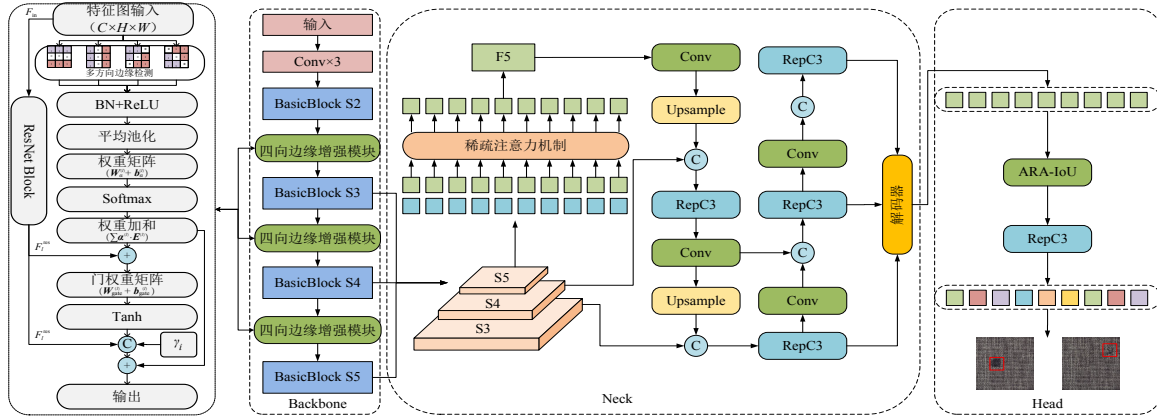


图4 FD-DETR网络结构图

Figure 4 FD-DETR network architecture diagram

个方向的Prewitt边缘算子,实现多方向边缘特征并行提取。

Prewitt四向边缘增强模块首先对输入特征图进行多方向边缘特征提取。设输入特征图为 $F_{in}$ 四个方向的边缘特征通过深度可分离卷积计算:

$$E_i^{(l)} = \sigma_2(\text{BN}(F_{in}^{(l)} \otimes P_i + b_i)), \quad i \in \{h, v, d, ad\} \quad (14)$$

其中, $P_i$ 为第 $i$ 个方向的深度Prewitt算子; $\otimes$ 表示卷积操作; $b_i$ 为偏置项;BN为批归一化操作; $\sigma_2$ 为ReLU激活函数;上标( $l$ )表示第 $l$ 个网络层级。为了自适应地平衡不同方向算子的贡献并考虑通道间的相互作用,引入多头注意力权重分配机制:

$$\alpha^{(l)} = \text{Softmax}\left(W_{\alpha}^{(l)} \cdot \text{Pool}\left(\text{Concat}\left[E_h^{(l)}, E_v^{(l)}, E_d^{(l)}, E_{ad}^{(l)}\right]\right) + b_{\alpha}^{(l)}\right) \quad (15)$$

其中, $W_{\alpha}^{(l)}$ 为第 $l$ 层的权重学习矩阵; $b_{\alpha}^{(l)}$ 为偏置向量,Pool表示全局平均池化操作; $\alpha^{(l)} = [\alpha_h^{(l)}, \alpha_v^{(l)}, \alpha_d^{(l)}, \alpha_{ad}^{(l)}]^T$ 为归一化的方向权重向量。多尺度特征增强通过门控残差连接机制将边缘增强特征与原始ResNet特征融合:

$$F_l^{\text{out}} = F_l^{\text{res}} + \gamma_l \cdot \tanh\left(W_{\text{gate}}^{(l)} \cdot \left[\sum_i \alpha_i^{(l)} \cdot E_i^{(l)} \oplus F_l^{\text{res}}\right] + b_{\text{gate}}^{(l)}\right) \odot F_{\text{edge}}^{(l)} \quad (16)$$

其中, $F_l^{\text{res}}$ 为第 $l$ 层ResNet块的输出特征; $\gamma_l \in [0, 1]$ 为可学习缩放因子; $W_{\text{gate}}^{(l)}$ 为门控权重矩阵; $b_{\text{gate}}^{(l)}$ 为门控偏置; $\oplus$ 表示特征拼接操作; $\odot$ 是元素乘法; $F_{\text{edge}}^{(l)} = \sum_i \alpha_i^{(l)} \cdot E_i^{(l)}$ 为加权融合的边缘特征。该门控机制通过tanh激活函数提供非线性门控信号,能够根据原始特征和边缘特征的相关性自适应地控制边缘信息的融入程度,避免边缘增强对原有语义特征的过度干扰。

### 3.2 稀疏注意力机制

传统的AIFI模块虽然能够有效建模长距离依赖

关系,但其计算复杂度较大,在处理高分辨率特征图时面临计算量大、内存消耗高的问题。因此,本文设计稀疏注意力机制SparseAIFI,具体结构如图5所示,通过混合稀疏注意力模式降低计算复杂度。

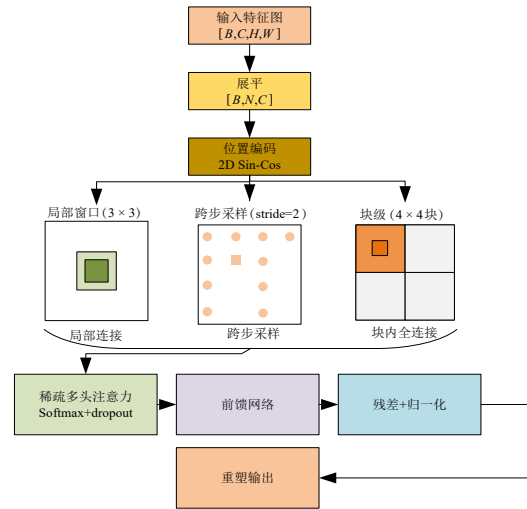


图5 稀疏注意力机制结构图

Figure 5 Structure diagram of the sparse attention mechanism

稀疏注意力机制首先生成混合稀疏掩码,将三种注意力模式通过逻辑或操作融合。局部窗口注意力保证了空间邻近位置的密切交互,跨步注意力维持了全局信息流动,块级注意力增强了中等尺度的特征提取。设特征图维度为 $H \times W$ ,序列长度 $n = H \times W$ ,混合稀疏掩码 $M \in \{0, 1\}$ 为

$$M(i, j) = M_{\text{local}}(i, j) \vee M_{\text{stride}}(i, j) \vee M_{\text{block}}(i, j) \quad (17)$$

其中, $i, j \in \{1, 2, \dots, n\}$ 表示特征序列中的位置索引; $\vee$ 表示逻辑或运算。将一维序列索引 $i$ 映射回二维特征图坐标 $(r_i, c_i)$ ,其中 $r_i = \lfloor (i-1)/w \rfloor + 1$ 表示行坐标, $c_i = (i-1) \bmod w + 1$ 表示列坐标,mod表示取模运算, $w$

为局部窗口大小。

各稀疏模式定义如下。

$M_{\text{local}}(i,j)$ : 当  $\lfloor r_i - r_j \rfloor \leq w/2$  且  $\lfloor c_i - c_j \rfloor \leq w/2$  时,  $M_{\text{local}}(i,j)=1$ , 其他情况为 0, 用于捕获疵点的局部边缘和纹理细节。

$M_{\text{stride}}(i,j)$ : 当  $\lfloor r_i/s \rfloor = \lfloor r_j/s \rfloor$  且  $\lfloor c_i/s \rfloor = \lfloor c_j/s \rfloor$  时,  $M_{\text{stride}}(i,j)=1$ , 其他情况为 0, 用于建立长距离依赖关系, 区分疵点与周期性织物纹理, 其中  $s$  为跨步间隔。

$M_{\text{block}}(i,j)$ : 当  $\lfloor r_i/b_s \rfloor = \lfloor r_j/b_s \rfloor$  且  $\lfloor c_i/b_s \rfloor = \lfloor c_j/b_s \rfloor$  时,  $M_{\text{block}}(i,j)=1$ , 其他情况为 0, 处理大中尺寸疵点如线状断纱, 其中  $b_s$  为块大小。为适配多头注意力计算, 将掩码  $M \in \mathbb{R}^{n \times n}$  扩展为  $M_{\text{expanded}} \in \mathbb{R}^{B \times H \times N}$ , 其中  $B$  为批大小,  $H$  为注意力头数,  $N=H \times W$  为序列长度, 扩展规则为对每个批次和每个注意力头复制掩码  $M$ , 确保共享稀疏结构:

$$M_{\text{expanded}}(b, h, n) = M, \quad \forall b \in \{1, 2, \dots, B\}, h \in \{1, 2, \dots, H\}, n \in \{1, 2, \dots, N\} \quad (18)$$

稀疏多头注意力计算通过选择性地仅计算掩码为 1 的位置对来降低计算复杂度。对于查询  $Q, K, V \in \mathbb{R}^{B \times H \times N \times d_k}$ , 其中  $d_k$  为键的特征维度, 稀疏注意力输出计算为

$$O = \text{Softmax} \left( \frac{(QK^T) \odot M_{\text{expanded}} + (1 - M_{\text{expanded}}) \cdot (-\infty)}{\sqrt{d_k}} \right) V \quad (19)$$

其中,  $(QK^T) \odot M_{\text{expanded}}$  保留掩码为 1 的注意力权重,  $(1 - M_{\text{expanded}}) \cdot (-\infty)$  将不需要计算的位置设为负无穷, 经过 Softmax 后权重变为 0。根据式(17)进行理论分析, 三种不同模式的稀疏掩码非零元素数量分别为  $|M_{\text{local}}| = n \cdot w^2$ ,  $|M_{\text{stride}}| = ns^2$ ,  $|M_{\text{block}}| = n \cdot b_s^2$ , 三种模式通过逻辑或运算融合, 忽略部分位置重叠, 非零元素的总数为

$$|M| \approx n \cdot w^2 + ns^2 + n \cdot b_s^2 = n(w^2 + s^2 + b_s^2) \quad (20)$$

掩码矩阵中非零元素的数量  $|M|$  直接对应需要实际计算注意力的位置对数量, 当  $w, s, b_s$  为常数时, 稀疏掩码的非零元素数量为  $|M| = O(n)$ 。标准自注意力需要计算所有  $n^2$  个位置对, 因此稀疏注意力的计算复杂度从标准注意力的  $O(n^2 \cdot d_k)$  降低至  $O(n \cdot d_k)$ , 实现了线性复杂度。混合稀疏注意力模式的设计确保了在大幅降低计算成本的同时保持有效的长短距离依赖建模能力。

### 3.3 ARA-IoU 损失函数

传统的 IoU 损失函数在处理织物疵点检测时存在收敛慢、对长宽比敏感度低等问题。本文提出长宽比感知 IoU 损失 (ARA-IoU), 通过引入中心距离归一化、长宽比感知和自适应权重机制来改善边界框回归

效果。ARA-IoU 损失函数的形式为

$$\mathcal{L}_{\text{ARA}} = (1 - \text{IoU}) + \frac{\rho^2}{c^2} + \beta \times R_{\text{aspect}} \quad (21)$$

其中,  $(1 - \text{IoU})$  为重叠度损失, 使预测框的位置和大小向真实框靠近;  $\rho$  为框中心距离;  $c$  为外接矩形对角线长度;  $\frac{\rho^2}{c^2}$  为中心距离损失, 进一步精细化位置优化, 加快收敛速度;  $R_{\text{aspect}}$  为长宽比不匹配度;  $\beta$  为长宽比感知项,  $\beta \times R_{\text{aspect}}$  在框的重叠度较高时, 自动增大权重, 优化框的长宽比, 确保形状的一致性。 $\frac{\rho^2}{c^2}$  中心距离项通过归一化处理提供直接的位置优化梯度具体公式为

$$\frac{\rho^2}{c^2} = \frac{(\hat{x}_c - x_c)^2 + (\hat{y}_c - y_c)^2}{(x_{\max} - x_{\min})^2 + (y_{\max} - y_{\min})^2} \quad (22)$$

其中,  $(\hat{x}_c, \hat{y}_c)$  和  $(x_c, y_c)$  为预测框和真实框的中心点。

长宽比不匹配度  $R_{\text{aspect}}$  量化预测框与真实框在形状上的差异, 具体公式为

$$R_{\text{aspect}} = \frac{4}{\pi^2} \sin^2 \left( \arctan \frac{w_{\text{gt}}}{h_{\text{gt}}} - \arctan \frac{w_{\text{pred}}}{h_{\text{pred}}} \right) \quad (23)$$

其中,  $w_{\text{gt}}$  和  $h_{\text{gt}}$  分别表示框的宽度和高度;  $w_{\text{pred}}$  和  $h_{\text{pred}}$  分别表示预测边界框的宽度和高度。该项通过角度差异的正弦平方函数平滑地惩罚长宽比不匹配, 避免了直接比值计算可能出现的数值不稳定问题。自适应权重  $\beta$  根据训练阶段动态调整长宽比项的重要性:

$$\beta = \frac{R_{\text{aspect}}}{1 - \text{IoU} + R_{\text{aspect}} + \epsilon} \quad (24)$$

其中,  $\epsilon$  为防止分母为零的平滑参数; 当 IoU 较低时,  $\beta$  较小, 优化重点在于提升重叠度; 当 IoU 较高时,  $\beta$  增大, 重点转向形状一致性优化。

## 4 实验与分析

织物疵点检测领域现有公开数据集普遍存在域间迁移问题, 其核心原因是数据集的采集环境、织物材质、纹理结构等与实际生产场景存在显著差异。为全面评估所提方法的有效性和实用性, 本文采用公开数据集与自建数据集相结合的方式验证 GAN-Data 的数据增强能力和 FD-DETR 的检测性能。实验在表 1 所列的高性能云计算平台上进行, 首先选用 MVTec AD 数据集<sup>[19]</sup>验证 GAN-Data 在同类织物图像上生成不同疵点的能力, 该数据集包含大量正常织物图像和少量疵点样本, 符合实际工业生产中疵点样本稀缺的真实情况。其次, 采用 Industrial Textile Dataset<sup>[20]</sup>进行泛化实验, 该数据集包含 10 种不同类型的疵点图像, 验证 GAN-Data 在不同织物纹理间的迁移能力。最后, 构建真实生产环境数据集, 采集自实际纺织生产线,

用于验证所提方法在实际生产场景下的综合检测性能及上下游协同优化效果。本节通过数据集增强和网络架构改进两个层面的实验,验证所提方法在实际生产场景下的检测性能。

表1 软硬件配置参数表

Table 1 Software and hardware configuration

软硬件	参数配置
CPU	18 vCPU AMD EPYC 9754
GPU	RTX 3090
内存	64 GB
操作系统	ubuntu18.04
CUDA版本	Cuda11.1
深度学习框架	PyTorch1.9.0
Python版本	Python3.8

#### 4.1 评价指标

本研究建立了不同评价指标体系。数据增强方面,采用SSIM (Structural Similarity Index Measure)<sup>[21]</sup>和PSNR (Peak Signal-to-Noise Ratio)<sup>[22]</sup>评估生成图像的结构相似性和峰值信噪比,数值越高表明质量越好,采用LPIPS (Learned Perceptual Image Patch Similarity)<sup>[23]</sup>和FID (Fréchet Inception Distance)<sup>[24]</sup>评估生成图像的感知质量和真实性,数值越低表明效果越好。检测模型方面,采用精确率(Precision)和召回率(Recall)评估基础性能,mAP50%和mAP50%-95%作为综合指标,分别反映模型在单一阈值和多阈值条件下的平均检测精度。上下游协同作业时采用F1分数验证所提方法的完整性。

#### 4.2 上游数据增强实验

实验中统一将输入图像尺寸设定为 $256 \times 256$ 像素,GAN-Data模型训练轮次设置为200,初始学习率设为0.000 2,100轮次后采用线性衰减。

##### 4.2.1 GAN-Data图像生成质量实验

本实验定性结果如图6所示,定量结果如表2所示。实验涵盖五种典型织物疵点类型:color污渍疵点、cut剪切疵点、hole孔洞疵点、metal金属异物和thread断纱疵点。每种类型的(a)行展示带有位置标注的疵点生成结果,红色框标注疵点位置,橙色框标注背景纹理区域;(b)行为橙色区域的纹理放大图,用于评估背景纹理生成质量。

图6的图像生成结果能够观察到对比模型均存在不同的共性问题。在Cycle-GAN模型生成的图像中存在大量的噪声点;DFM-GAN模型生成的图像中疵点与背景纹理过渡十分生硬;Dual-GAN模型生成的图像中背景纹理模糊且均未能有效迁移疵点特征;DDPM生成的背景当中其纹理布局与真实图像存在较大差异;Style-GAN3生成的图像当中虽然纹理密度

表2 GAN-Data图像生成图像评价指标

Table 2 GAN-Data image generation evaluation metrics

评价指标	SSIM ↑	PSNR ↑	LPIPS ↓	FID ↓
Cycle-GAN	0.045	14.241	0.158	55.450
DFM-GAN	0.042	13.791	0.172	58.130
Dual-GAN	0.041	14.256	0.164	55.179
DDPM	0.037	13.511	0.196	127.012
Style-GAN3	0.035	13.963	0.153 (best)	59.668
GAN-Data	0.046 (best)	14.261 (best)	0.157	48.893 (best)

布局合理,但存在纹理扭曲现象。此外,各对比模型还存在不同的特性问题,除了DFM-GAN外,其他四种对比模型均未能完全生成五种疵点。DFM-GAN在color(a)、cut(a)当中生成的疵点仍具有织物纹理。相比之下GAN-Data实现了形态准确、位置精确的生成,疵点边界呈现自然过渡,且生成图像的背景纹理无高密度随机噪声、纹理扭曲或模糊现象。从实验结果来看,对比模型在织物疵点生成任务中表现出三种不足。首先,多尺度疵点生成能力不足是普遍存在的问题。Cycle-GAN和Dual-GAN生成的疵点真实性不足,且部分疵点无法生成,这可能是由于其卷积结构的感受野有限,无法适配从点状污渍到大面积孔洞的多尺度跨度。其次,疵点与背景纹理的耦合干扰导致生成质量下降。DFM-GAN在color(a)和cut(a)的疵点内部仍保留原始织物纹理,呈现“半透明”效果,这说明了现有方法在疵点生成时未能有效消除背景纹理的影响。最后,背景纹理真实性不足制约了生成样本的实用价值。实验观察到DDPM虽能生成较为细腻的纹理细节,但在背景区域出现纹理过渡平滑偏离真实分布的问题,反映出扩散模型的生成机制在疵点与背景的差异化处理上存在局限。此外,Cycle-GAN、Dual-GAN和Style-GAN3生成图像的背景纹理存在噪声点、模糊和扭曲。对比模型生成的背景纹理分别表现出以上的失真现象,这反映出有损失函数体系缺乏对织物背景当中周期性纹理的有效约束机制。

定量评估结果从统计层面验证了上述观察,表2显示,本文方法在SSIM、PSNR和FID三项指标上均取得最优性能,分别达到0.046、14.261和48.893,这反映GAN-Data在结构相似性、峰值信噪比和真实性三个维度的综合优势。深入分析各模型的指标可以发现,对比模型普遍存在指标间的失衡现象。Style-GAN3虽在LPIPS指标上表现最佳,但FID得分较高,结合定性结果来看,其在提升感知质量的同时存在细小尺度疵点生成缺失的问题。Cycle-GAN和Dual-GAN在SSIM和PSNR上相对较优,但LPIPS和FID得分较差,表明生成结果虽在像素级结构上接近原图,但在感知质量和真实性上存在明显不足,本文方法虽在LPIPS

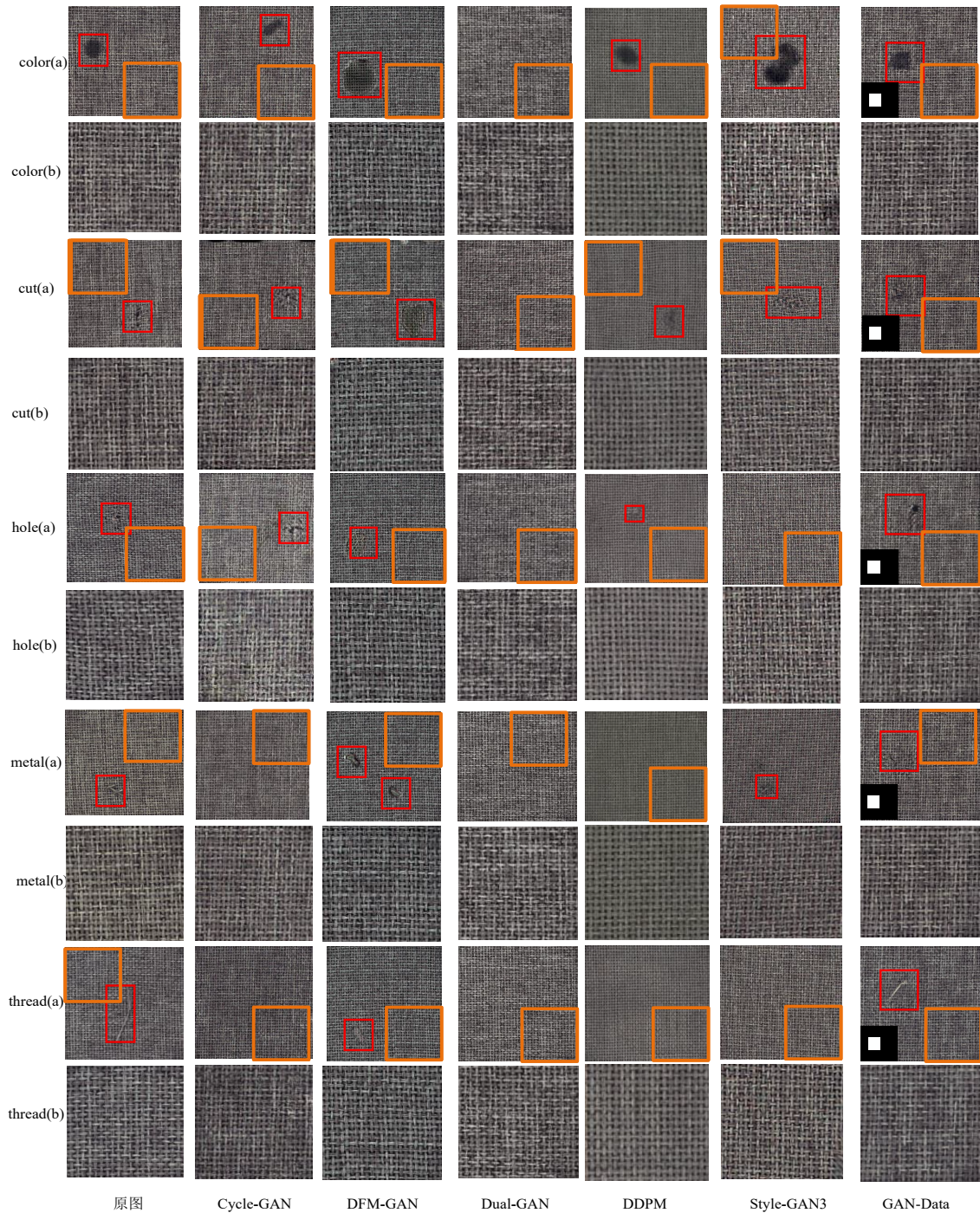


图6 GAN-Data 图像生成质量实验结果图

Figure 6 GAN-Data image generation quality experiment result

得分上未取得最佳,但疵点的准确生成和背景纹理的高保真度,证明所提方法较对比模型具有优势。

#### 4.2.2 GAN-Data 方法泛化实验

前述 MVTEC AD 数据集实验验证了 GAN-Data 在单类织物生成不同疵点类型的能力,但为充分评估模型的实际应用价值,还需验证其不同织物背景和多样化疵点分布下的泛化能力。为此,本文选用 Indus-

trial Textile Dataset 进行跨域泛化实验。实验涵盖 10 种不同织物背景和疵点类型,每种使用不同形状和尺寸的 Mask 进行测试。实验结果如图 7 所示,从左至右分别为原始图像(A 列)及不同测试条件下的生成结果(B、C、D 列)。

从定性结果来看,GAN-Data 在跨域场景下展现出三方面显著优势。织物背景适应性方面,模型能够

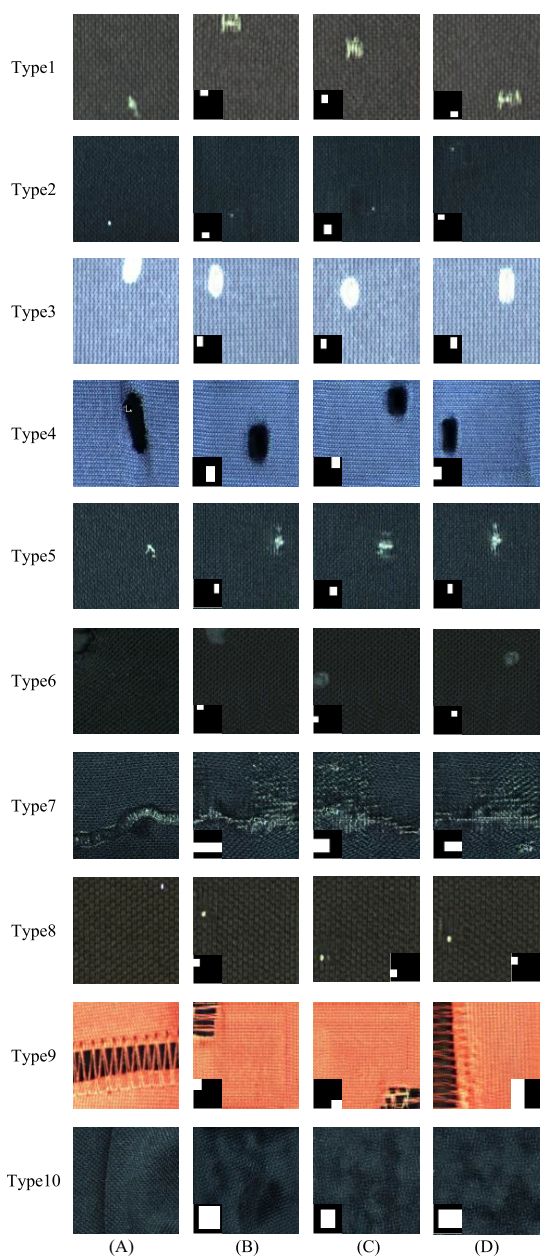


图7 GAN-Data泛化实验结果图

Figure 7 GAN-Data generalization experiment results

适配不同材质、颜色和纹理密度的织物。Type3、Type4蓝色平纹织物测试显示,生成的白色与黑色疵点边界清晰,与背景纹理实现自然融合。Type9橙色织物的测试表明,模型对于鲜艳的织物疵点仍具备一定的适应能力。Type2、Type6、Type8等深色织物测试进一步验证了模型对低亮度、高纹理密度背景的处理能力,即使在深色或低对比度环境下仍能生成可见度良好的疵点。疵点形态多样性方面,模型成功生成了从小尺度到大尺度、从规则到不规则的多种疵点类型。Type2、Type5、Type8展示了模型对点状小疵点的精准生成能力,Type7生成的边缘渐变形态保持了真

实特征,Type10大域疵点验证了模型对大面积缺陷的完整覆盖能力。位置灵活性方面,Type1的多组对比结果显示,相同织物背景下中心、边缘或任意位置的疵点生成均保持一致的质量水平,表明模型能够根据Mask掩码实现高质量织物疵点样本生成。

定量评估结果进一步验证了GAN-Data的跨域泛化能力。在ITD数据集上,模型的平均SSIM、PSNR、LPIPS、FID分别为0.251、19.735、0.351、117.268。相比MVTec AD训练集,SSIM、LPIPS和FID指标均有所下降,这反映了跨域场景中织物纹理类型差异带来的结构重建难度增加,但生成图像在真实性方面仍保持可接受范围。这一结果表明,GAN-Data在有限训练数据的前提下,仍可实现良好的跨域泛化效果。

### 4.3 下游织物疵点识别实验

为评估GAN-Data生成的织物样本图像对下游检测网络的实际影响,采用YOLO、Efficient-DETR和RT-DETR网络设计对比实验<sup>[25]</sup>,使用数据集如表3所示,其中Data1~Data3为自建数据集,Data4为MVTec AD数据集中的地毯类型,考虑到在GAN-Data数据增强实验中仅有DFM-GAN和GAN-Data两种方法生成了MVTec AD地毯数据集全部疵点,因此实验采用传统数据增强、DFM-GAN和GAN-Data三种方式生成的数据集。训练集使用DFM-GAN/GAN-Data生成的合成疵点图像,用来评估不同生成模型的疵点图像对下游的影响,验证集采用真实疵点图像,在统一的训练配置下进行评估。输入图像尺寸为 $256 \times 256$ 像素,每个检测模型的训练轮次设置为100。

表3 数据集增强前后数量对比表

Table 3 Table comparing dataset size before and after augmentation

类型	数据集名称	疵点类别	训练集	验证集	测试集	合计
传统增强	Data1-ARG	4	588	168	84	840
	Data2-ARG	3	275	78	40	393
	Data3-ARG	4	170	48	26	244
	Data4-ARG	5	400	113	61	574
GAN-Data/DFM-GAN增强	Data1-GAN/DFM	4	882	252	126	1 260
	Data2-GAN/DFM	3	825	234	120	1 179
	Data3-GAN/DFM	4	340	96	80	516
	Data4-GAN/DFM	5	1 000	283	153	1 436

实验结果如表4所示,GAN-Data相比其他方法展现出三方面显著优势。首先是生成样本的完整性,DFM-GAN在生成Data3数据集时,某一类疵点未被成功生成,导致对应训练集中缺少该疵点类型的合成样本,进而造成精确率和mAP指标较大幅度的下降。其次是生成疵点的真实性,DFM-GAN生成的样本精确率和召回率相差过大,在多个网络上存在严重的漏检现象,反映出生成的疵点特征表达不足。

相比之下, GAN-Data 实现了全类型疵点的完整生成, 精确率和召回率保持平衡, 为训练集提供了充足学习的疵点数据。最后是稳定性表现, DFM-GAN 在不同数据集间波动较大, 其中 Data2-DFM 相对最佳, 但其精确率和召回率仍明显低于传统增强的 Data-ARG, 而 GAN-Data 在所有数据集配置上都相比传统增强实现了显著提升, 精确率、召回率和 mAP 指标均保持稳定的高性能水平, 特别是在 Data2 和 Data4 数据集配置上, GAN-Data 相比 Data-ARG 在不同的检测模型当中的精确率和召回率均有较大的提升。FD-DETR 检测模型对检测性能的提升也具有贡献, 在 FD-DETR 网络上, GAN-Data 相比 Data-ARG 和 Data-DFM 的性能优势最为突出, 如 Data4-ARG 的 mAP50% 达到 0.970, mAP50%-95% 为 0.671 的较高水平, 但使用 GAN-Data 数据的 Data4-GAN 的 mAP50% 和 mAP50%-95% 仍能更进一步, 尤其 mAP50%-95%

提升显著。在相同数据集下, 使用传统增强数据集的 FD-DETR 相比对比模型大部分情况下均有提升, 但在部分数据集上的性能仍与对比模型存在一定差距, 如 YOLOV11 的 Data1-ARG 数据集上 FD-DETR 存在一定的差距, 可能是 Data1 数据集在扩充数量上最少, 导致 FD-DETR 模型性能未得到完全释放。使用 GAN-Data 数据集的情况下, FD-DETR 普遍表现最佳, 表明 FD-DETR 相较于对比模型具有一定的优势, 更表明了若有上游数据增强为其提供充足的训练样本, 则检测模型的性能可以更进一步。在模型效率上, FD-DETR 也具有一定的优势, 其各模型的效率如表 5 所示, 在六种检测网络中, FD-DETR 的帧率仅次于 YOLO 系列, 但相比其他检测模型有明显提升。同时, FD-DETR 的参数量为所有网络中最少, 计算复杂度也最低, 73.7 FPS(帧/s)能满足工业生产对实时检测的需求。

表 4 下游网络不同数据实验结果表

Table 4 Experimental results of different data on downstream networks

模型 Model	数据集 Datasets	准确率 Precision	召回率 Recall	平均精度 mAP50%	平均精度 mAP50%-95%
YOLOV5-l	Data1-ARG	0.849	0.837	0.884	0.542
	Data1-DFM	0.750	0.218	0.239	0.108
	Data1-GAN	0.919	0.903	0.936	0.576
	Data2-ARG	0.894	0.893	0.954	0.668
	Data2-DFM	0.682	0.579	0.596	0.347
	Data2-GAN	0.951	0.939	0.956	0.541
	Data3-ARG	0.836	0.893	0.940	0.766
	Data3-DFM	0.480	0.203	0.195	0.078
	Data3-GAN	0.951	0.939	0.965	0.716
	Data4-ARG	0.870	0.819	0.834	0.550
	Data4-DFM	0.545	0.313	0.356	0.193
	Data4-GAN	0.916	0.904	0.935	0.627
YOLOV8-l	Data1-ARG	0.891	0.817	0.899	0.586
	Data1-DFM	0.722	0.204	0.256	0.105
	Data1-GAN	0.945	0.901	0.922	0.625
	Data2-ARG	0.883	0.888	0.917	0.528
	Data2-DFM	0.708	0.617	0.657	0.357
	Data2-GAN	0.955	0.977	0.977	0.594
	Data3-ARG	0.912	0.802	0.879	0.653
	Data3-DFM	0.440	0.230	0.165	0.066
	Data3-GAN	0.921	0.829	0.912	0.647
	Data4-ARG	0.886	0.812	0.839	0.529
	Data4-DFM	0.512	0.265	0.339	0.165

续表

模型 Model	数据集 Datasets	准确率 Precision	召回率 Recall	平均精度 mAP50%	平均精度 mAP50%-95%
	Data4-GAN	0.938	0.886	0.939	0.604
YOLOV11-l	Data1-ARG	0.903	0.806	0.899	0.559
	Data1-DFM	0.739	0.215	0.225	0.102
	Data1-GAN	0.903	0.939	0.928	0.599
	Data2-ARG	0.895	0.986	0.964	0.675
	Data2-DFM	0.787	0.663	0.697	0.389
	Data2-GAN	0.980	0.984	0.987	0.624
	Data3-ARG	0.914	0.947	0.937	0.687
	Data3-DFM	0.700	0.139	0.131	0.061
	Data3-GAN	0.934	0.857	0.929	0.637
	Data4-ARG	0.870	0.771	0.824	0.506
	Data4-DFM	0.631	0.162	0.237	0.130
	Data4-GAN	0.956	0.867	0.931	0.570
	Efficient-DET	Data1-ARG	0.811	0.833	0.872
Data1-DFM		0.549	0.245	0.256	0.102
Data1-GAN		0.943	0.922	0.924	0.546
Data2-ARG		0.957	0.945	0.858	0.449
Data2-DFM		0.648	0.620	0.615	0.355
Data2-GAN		0.971	0.964	0.973	0.578
Data3-ARG		0.862	0.912	0.892	0.548
Data3-DFM		0.464	0.242	0.178	0.069
Data3-GAN		0.943	0.906	0.938	0.578
Data4-ARG		0.835	0.838	0.823	0.514
Data4-DFM		0.603	0.185	0.225	0.133
Data4-GAN		0.956	0.867	0.931	0.570
RT-DETR-R50	Data1-ARG	0.861	0.786	0.794	0.417
	Data1-DFM	0.692	0.392	0.349	0.166
	Data1-GAN	0.876	0.869	0.871	0.477
	Data2-ARG	0.887	0.696	0.752	0.414
	Data2-DFM	0.684	0.716	0.665	0.366
	Data2-GAN	0.918	0.967	0.950	0.533
	Data3-ARG	0.711	0.625	0.592	0.340
	Data3-DFM	0.248	0.238	0.163	0.071
	Data3-GAN	0.800	0.619	0.662	0.371
	Data4-ARG	0.827	0.873	0.864	0.620
	Data4-DFM	0.533	0.364	0.301	0.170
	Data4-GAN	0.965	0.966	0.959	0.640

续表

模型 Model	数据集 Datasets	准确率 Precision	召回率 Recall	平均精度 mAP50%	平均精度 mAP50%-95%
FD-DETR	Data1-ARG	0.881	0.871	0.857	0.447
	Data1-DFM	0.693	0.121	0.114	0.032
	Data1-GAN	0.962	0.912	0.929	0.603
	Data2-ARG	0.899	0.897	0.945	0.624
	Data2-DFM	0.648	0.630	0.653	0.358
	Data2-GAN	0.986	0.971	0.954	0.627
	Data3-ARG	0.919	0.891	0.949	0.605
	Data3-DFM	0.659	0.139	0.128	0.040
	Data3-GAN	0.974	0.943	0.965	0.655
	Data4-ARG	0.979	0.943	0.970	0.671
	Data4-DFM	0.627	0.095	0.094	0.032
	Data4-GAN	0.985	0.986	0.989	0.740

表 5 各模型计算效率

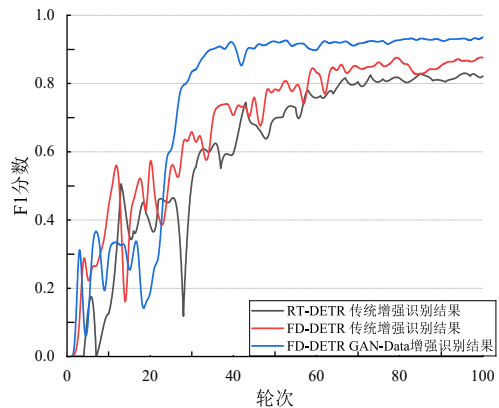
Table 5 Computational efficiency of each model

模型	帧率/FPS	参数量/M	Flops/G
YOLOV5-L	54.3	46.2	109.7
YOLOV8-L	71.5	43.7	164.8
YOLOV11-L	85.7	25.4	86.9
Efficient-DET	22.5	35.2	159.0
RT-DETR-R50	42.8	42.3	129.6
FD-DETR	73.7	32.9	81.7

#### 4.4 数据增强与检测模型的协同优化实验

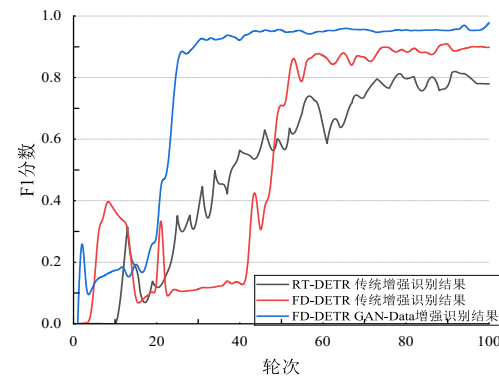
前述实验采用多个数据集的评价指标验证了GAN-Data和FD-DETR方法的有效性,但为了探究数据增强与检测模型在训练过程中的细节以及验证阶段的具体检测效果。本节通过F1分数曲线和检测框可视化对比,分析GAN-Data数据增强与FD-DETR模型在训练和检测中的协同效应。

图8展示了三种模型配置在4个数据集上的F1-Score随训练轮次的变化曲线。图8(a)、图8(b)、图8(c)、图8(d)分别对应Data1-Data4数据集的训练过程,从4个子图可以观察到一致的性能递进规律,在所有数据集上,FD-DETR的GAN-Data配置均表现最优,通常在第20~40轮内快速收敛至0.9以上的F1分数,且后续训练过程稳定、波动小。去除GAN-Data数据增强使用传统增强后FD-DETR的F1分数有所下降,收敛速度变慢。若进一步替换为基础检测模型,F1分数进一步下降,且训练曲线波动幅度大。这表明,GAN-Data数据增强与FD-DETR模型架构的协同方案,相比仅依赖数据增强或仅改进检测模型的方案,在织物小样本训练场景下可实现更快的收敛速度与更高的检测性能上限。



(a) Data1数据集不同配置下的训练过程

(a) Training process of Data1 dataset under different configurations

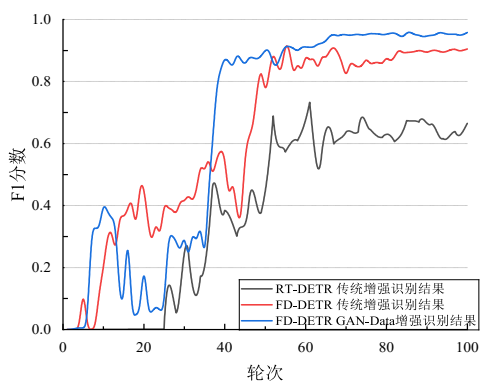


(b) Data2数据集不同配置下的训练过程

(b) Training process of Data2 dataset under different configurations

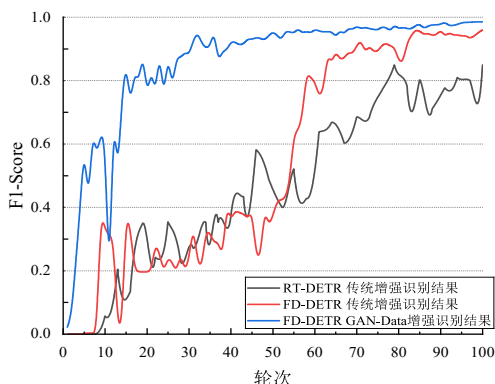
图8 改进前后与不同图像的识别结果

Figure 8 Recognition results with different images before and after improvement



(c) Data3数据集不同配置下的训练过程

(c) Training process of Data3 dataset under different configurations



(d) Data4数据集不同配置下的训练过程

(d) Training process of Data4 dataset under different configurations

续图8

图9进一步展示了各场景下的检测效果对比,提供定性可视化分析。图中选取了三种配置均成功检测的典型样本进行展示。每行对应一种疵点类型,从左至右依次为原始图像图9①、RT-DETR 普通增强识别结果图9②、FD-DETR 普通增强识别结果图9③、FD-DETR 与 GAN-Data 增强识别结果图9④。使用RT-DETR 普通增强识别结果在多数场景中能够识别出疵点目标,展现了基础检测能力。从图9②)列可见,该方案在各数据集场景下均能完成疵点识别。但图9②)列当中边界定位准确性不足,在Data3(b)和Data4(c)、Data4(d)中检测框略显宽松或者紧凑,在Data2(c)类型未能检测到全部疵点,这些问题反映出传统数据增强方法在样本多样性方面的局限性。FD-DETR模型通过织物的针对性优化,在检测精度上实现显著提升,其中Data3(b)和Data4(c)可以明显观察到检测框更加合理,Data2(c)中更多的疵点被检测到,检测置信度得到提升。当FD-DETR与GAN-Data增强结合使用时,检测系统达到最优性能。

Data4(d)的检测框相较于(B)(C)两列得到明显优化,Data2(c)中的疵点被完全检出。相比(C)列,GAN-Data增强进一步提升了疵点检测场景中的准确率,检测置信度和边界框定位精度。这进一步验证了图8当中的结论,为工业生产中布匹疵点检测提供了一种有效的技术解决方案。

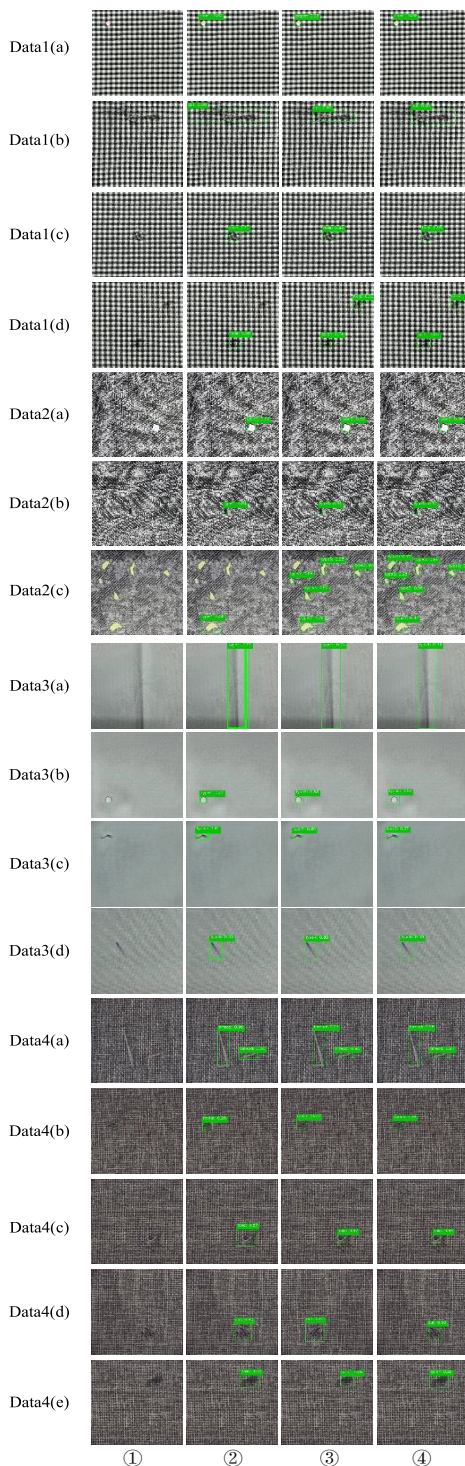


图9 不同方法下图像的识别结果

Figure 9 Recognition results of images using different methods

## 5 结论与展望

本文提出一种针对工业生产中织物疵点小样本检测方案,包括上游数据增强和下游模型优化两个环节。

在数据增强方面,本文提出的GAN-Data模型通过精准的疵点生成机制与纹理保真设计,为小样本学习条件下的检测模型提供了高质量训练样本支撑。实验验证,GAN-Data在MVTec AD地毯子类数据集上可精准生成亚麻布织物的多类型疵点样本,且背景纹理保真度高;ITD的跨域泛化实验进一步验证了GAN-Data在多织物类型、多疵点形态下的高质量样本生成能力。下游织物疵点检测实验中,不同对比检测模型在4个数据集上的结果证实,GAN-Data不仅具备高质量织物疵点样本生成能力,其生成的样本还能作为下游检测模型带来显著的性能提升效果,充分验证了上游数据增强方法的有效性。

在检测模型优化环节,本文提出的FD-DETR模型通过多尺度边缘增强主干网络、稀疏注意力机制及长宽比感知IoU损失函数,实现了检测准确率、mAP等核心指标的显著提升,且模型检测效率满足工业现场实时检测的部署要求,验证了下游检测模型优化方案的有效性。

最后,通过数据增强与检测模型的协同优化实验,结合逐步叠加改进模块的对照实验及F1分数训练曲线分析,证实相比单一的数据增强或模型改进,GAN-Data与FD-DETR的协同方案可实现检测性能的叠加提升,验证了“数据增强+模型优化”双向协同技术路线的可行性,为工业织物疵点小样本检测领域提供了可落地的技术方案与新的研究思路。

本文的工作仍存在一定的局限性。在具体的泛化实验中可以发现,在全图覆盖的超大范围疵点上生成的图像效果不佳,导致泛化实验中平均FID指标较高,这表明上游数据增强部分仍有改进空间。GAN-Data和FD-DETR方法在本文中主要在棉麻等常见织物纹理上进行训练和评估,对于特殊纹理的高端织物,如蕾丝、绸缎的推广能力有待验证,后续将针对这两个方面展开具体的研究。

### 参考文献

- [1] 王斌,李敏,雷承霖,等.基于深度学习的织物疵点检测研究进展[J].纺织学报,2023,44(1):219-227.  
Wang Bin, Li Min, Lei Chenglin, et al. Research progress in fabric defect detection based on deep learning[J]. Journal of Textile Research, 2023, 44(1): 219-227. (in Chinese)
- [2] 葛轶洲,刘恒,王言,等.小样本困境下的深度学习图像识别综述[J].软件学报,2022,33(1):193-210.  
Ge Yizhou, Liu Heng, Wang Yan, et al. Survey on deep learning image recognition in dilemma of small samples[J]. Journal of Software, 2022, 33(1): 193-210. (in Chinese)
- [3] Cheng Ling, Zhao Lianying, Chen Li, et al. Digital image processing of cotton yarn seriplane[C]//2010 International Conference on Computer and Information Application. Piscataway: IEEE, 2010: 274-277.
- [4] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [5] 王伟,张静宜,温玉辉,等.基于神经网络的图像风格迁移算法综述[J].电子学报,2025,53(5):1692-1712.  
Wang Wei, Zhang Jingyi, Wen Yuhui, et al. Neural network based image style transfer: A survey[J]. Acta Electronica Sinica, 2025, 53(5): 1692-1712. (in Chinese)
- [6] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2242-2251.
- [7] 李云红,张蕾涛,李丽敏,等.基于CycleGAN-IA方法和M-ConvNext网络的苹果叶片病害图像识别[J].农业机械学报,2024,55(4):204-212.  
Li Yunhong, Zhang Leitao, Li Limin, et al. Image recognition of apple leaf disease based on CycleGAN-IA method and M-ConvNext network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 204-212. (in Chinese)
- [8] 牛玉贞,张凌昕,兰杰,等.基于分频式生成对抗网络的非成对水下图像增强[J].电子学报,2025,53(2):527-544.  
Niu Yuzhen, Zhang Lingxin, Lan Jie, Lan J, et al. FD-GAN: Frequency-decomposed generative adversarial network for unpaired underwater image enhancement[J]. Acta Electronica Sinica, 2025, 53(2): 527-544. (in Chinese)
- [9] Duan Yuxuan, Hong Yan, Niu Li, et al. Few-shot defect image generation via defect-aware feature manipulation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(1): 571-578.
- [10] Zhang Cheng, Wang Yuanhao, Vicente F, et al. Fabric-Diffusion: High-fidelity texture transfer for 3D garments generation from in-the-wild images[C]//SIGGRAPH Asia 2024 Conference Papers. New York: ACM, 2024: 1-12.
- [11] 秦嘉奇,江泽涛,雷晓春.基于ICFIE-YOLO的低照度图像目标检测方法[J].电子学报,2025,53(2):514-526.  
Qin Jiaqi, Jiang Zetao, Lei Xiaochun. Low illumination image object detection method based on ICFIE-YOLO[J]. Acta Electronica Sinica, 2025, 53(2): 514-526. (in Chinese)
- [12] 邵延华,张铎,楚红雨,等.基于深度学习的YOLO目标检测综述[J].电子与信息学报,2022,44(10):3697-3708.  
Shao Yanhua, Zhang Duo, Chu Hongyu, Chu H Y, et al.

- A review of YOLO object detection based on deep learning[J]. Journal of Electronics & Information Technology, 2022, 44(10): 3697-3708. (in Chinese)
- [13] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]//Computer Vision - EC-CV 2020. Cham: Springer, 2020: 213-229.
- [14] CHEN Songle, SUN Hongbo, WU Yuxin, et al. A helmet detection algorithm based on transformers with deformable attention module[J]. Chinese Journal of Electronics, 2025, 34(1): 229-241.
- [15] Zhao Yian, Lv Wenyu, Xu Shangliang, et al. DETRs beat YOLOs on real-time object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2024: 16965-16974.
- [16] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 936-944.
- [17] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[M]//Computer Vision - ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [18] Motamedi M, Sakharnykh N, Kaldewey T. A data-centric approach for training deep neural networks with less data[PP/OL]. V2.arXiv (2021-10-29)[2026-01-06]. <https://doi.org/10.48550/arXiv.2110.03613>.
- [19] Bergmann P, Fauser M, Sattlegger D, et al. MVTec AD: A comprehensive real-world dataset for unsupervised anomaly detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019. DOI:10.1109/cvpr.2019.00982.
- [20] Simon T. Industrial textile dataset[DS/OL]. IEEE Dataport (2025-7-19) [2026-01-06]. <https://iee-dataport.org/documents/industrialtextiledataset>.
- [21] Wang Tianyu, Yang Xin, Xu Ke, et al. Spatial attentive single-image deraining with a high quality real rain dataset[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 12262-12271.
- [22] Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment[J]. Electronics Letters, 2008, 44(13): 800-801.
- [23] Zhang R, Isola P, Efros A A, et al. The unreasonable effectiveness of deep features as a perceptual metric[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 586-595.
- [24] Heusel M, Ramsauer H, Unterthiner T, et al. GANs trained by a two time-scale update rule converge to a local Nash equilibrium[J]. Neural Information Processing Systems, 2017, 30: 6629-6640.
- [25] Jiang Peiyuan, Dajiang Ergu, Liu Fangyao, et al. A review of yolo algorithm developments[J]. Procedia Computer Science, 2022, 199: 1066-1073.

### 作者简介



**陈孟元** 男, 1984年1月出生于安徽省芜湖市。现为安徽工程大学电气工程学院教授、博士生导师, 安徽省学术和技术带头人后备人选, 安徽省高校杰出青年基金获得者, 第一完成人获得安徽省科学技术进步奖一等奖。主要研究方向为移动机器人视觉SLAM与智能制造技术。

E-mail: mychen@ahpu.edu.cn



**张坦坦** 男, 2000年5月出生于安徽省宿州市。现为安徽工程大学电子信息专业硕士研究生。主要研究方向为风格迁移与缺陷检测。

E-mail: 2230342256@stu.ahpu.edu.cn



**唐哲** 男, 2000年6月出生于安徽省宿州市。现为安徽工程大学电子信息专业硕士研究生。主要研究方向为图像处理。

E-mail: 2230342287@stu.ahpu.edu.cn