

SAM-MR: 基于 SAM 的混合区域匹配专家适配布匹检测算法

罗其锋¹ 肖星¹ 温焯飞¹ 池明旻² 彭博³

1 南方电网广东中山供电局 广东 中山 528401

2 复旦大学计算机科学技术学院 上海 200438

3 上海海洋大学信息学院 上海 201306

(13824776366@139.com)

摘要 有监督异常检测因其精准的工业异常检测能力而广泛应用于布匹质量检测。现有的统一架构的异常检测方法,因其单一的特征适配能力,不能对多样化的,所以度较高的布匹瑕疵进行有效地区分,因此在布匹的多类别的异常检测中性能会显著下降。为此提出一种基于混合区域匹配专家适配方法(Mixture of Region Experts),通过 Mixture of Adapter Experts 模块来区别化不同类别的布匹瑕疵特征,使用 Align and Differencing 模块对齐模板图特征和瑕疵特征来进一步加强异常区域的划分,从而有效提高了模型分辨复杂多类型的布匹瑕疵的能力。同时,模型进一步集成成分检测任务,在完成瑕疵定位的基础上实现异常成分的语义识别。实验结果表明,SAM-MR 在布匹纤维材质和缺陷检测任务上取得了优于现有方法的性能,定性、定量分析及消融实验验证了所提出方法在多任务预测中的有效性。

关键词: 有监督学习;异常检测;混合专家

中图分类号 TP181

SAM-MR: SAM-based Mixed Region Matching Expert Adaptation Algorithm for Fabric Detection

LUO Qifeng¹, XIAO Xing¹, WEN Chaofei¹, CHI Mingmin² and PENG Bo³

1 Zhongshan Power Supply Bureau, Guangdong Power Grid Co., Ltd., Zhongshan, Guangdong 528401, China

2 School of Computer Science and Technology, Fudan University, Shanghai 200438, China

3 College of Information, Shanghai Ocean University, Shanghai 201306, China

Abstract Supervised anomaly detection has been widely applied to fabric quality inspection due to its high precision in industrial scenarios. However, existing unified-architecture methods often suffer from limited feature adaptation capabilities, making it difficult to distinguish diverse and highly similar fabric defects. This paper proposes a novel approach based on a Mixture of Region Experts (SAM-MR), which introduces a Mixture of Adapter Experts module to differentiate between various types of fabric defects. Additionally, an Align and Differencing module is employed to align features between template and defect images, further enhancing the localization of anomalous regions. The model is also extended to incorporate component detection, enabling semantic recognition of defect-related components on top of defect localization. Experimental results demonstrate that SAM-MR outperforms existing methods on fabric defect datasets, and qualitative, quantitative, and ablation studies validate the effectiveness of the proposed approach in multi-task prediction.

Keywords Supervised learning, Anomaly detection, Mixture of experts

1 引言

纺织品的质量影响着纺织行业的发展,布匹瑕疵直接影响产品的外观质量和使用性能,因此布匹瑕疵检测成为了产品质量控制的重要环节^[1],受到了广泛关注。传统的瑕疵检测方法往往依赖于人工经验或简单的图像处理技术,难以满足高效、准确的检测需求。近年来,深度学习技术的迅猛发展为图像处理领域带来了新的机遇,尤其是在异常检测任务中,深度学习模型展现出了优越的性能。然而,现有的深度学习方法在处理多类别瑕疵时,往往面临特征提取不充分、类别间

差异性不足等问题,导致检测效果不理想。

在实际工业环境的纺织生产线中,检查员或模型往往必须依赖有限的正常数据来确保过程的准确性。使用模型验证大长度面料的印刷和染色质量的方法存在着以下两种问题^[2-3]:1)产品设计的多样性和缺乏固定的生产模式;2)瑕疵通常较为微妙,涉及纹理和颜色的细微变化,呈现出复杂的背景。

为了解决上述问题,本文提出了一种基于 SAM^[4] (Segment Anything Model) 的混合区域匹配专家适配算法^[5-6]——SAM-MR (Mixture of Region Experts)。该算法的

基金项目:南方电网公司科技项目(032000KC23030007(GDKJXM20230265))

This work was supported by the Science and Technology Project of China Southern Power Grid Company(032000KC23030007(GDKJXM20230265)).

通信作者:彭博(bpeng@shou.edu.cn)

核心思想是通过混合专家适配器,针对不同种类的布匹瑕疵^[7]和模板图进行特征提取,同时利用布匹的掩码信息标注瑕疵区域^[8-9]。通过度量并扩大瑕疵区域的差异,缩小非瑕疵区域的差异,SAM-MR 能够实现针对不同种类布匹瑕疵的专有学习,并确保图像非瑕疵部分的特征保持同一分布下。这一方法为后续的对齐和差分提供了优质的图像特征表示。此外,为进一步提升模型的多任务处理能力,本文将瑕疵检测扩展至成分测定任务,实现从异常定位到异常语义识别的统一框架。实验结果表明,所提出的 SAM-MR 在布匹瑕疵数据集上取得了领先性能,并在成分识别任务中展现出良好的泛化能力。

本文的贡献可以总结如下:

1) 提出了一种创新的混合区域匹配专家适配算法,能够有效处理多类别布匹瑕疵的检测问题,提升了检测的准确性和效率。

2) 通过实验验证了 SAM-MR 算法的有效性,展示了其在纺织生产线瑕疵检测中的潜力,为未来的智能化质量控制提供了新的思路和方法。

2 相关工作

在织物材料产品的生产过程中,因设备硬件和环境因素的影响,难免会出现异常。如果织物的不规则性导致消费者拒绝,该情况将被归类为织物缺陷。最近,深度学习技术广泛应用于在工业应用中,这主要得益于其从原始数据中解读特征和自动识别^[10]的能力。基于这种多层网络,深度学习引导提取深层特征,利用这些更深入的见解来表示数据中的嵌入语义信息,从而增强特征的鲁棒性。通过提高精度和效率,深度学习方法能够有效检测和分类织物中的缺陷。

在图像检测和识别领域,卷积神经网络(CNN)相比传统分类方法展现出更好的容错性和学习能力。这些网络特别擅长适应涵盖更广泛缺陷类别的织物。Han 等^[10]创新性地构建了一种特征浅层结构和双网络并行性的神经网络架构。基于此,Ghosh 等^[11]巧妙地利用 CNN 的特征提取能力,随后使用支持向量机进行精细分类。

基于深度学习的有监督变化检测(CD)方法通常可以分为两种不同类型:两阶段方法和单阶段方法。两阶段方法通常利用训练卷积神经网络^[12-13](CNN)或全卷积网络^[14-15](FCN)来分别分类异常图片和正常图片,随后比较它们的分类结果以做出变化决策。为了实现这一目标,必须提供布匹的正常标签和异常标签^[16-18]。

这些方法的局限性主要源于对空间上下文信息的关注不足,归因于缺乏对全局特征的提取。因此,获得的结果常常受到噪声和上下文因素的影响。因此,已经引入了几种算法来增强特征识别的能力。这些增强措施包括利用扩张卷积、增加卷积层的数量以及集成注意力机制以提高精度和精细度。

此外,许多研究未能考虑正常样本和异常样本的差异信息,主要依赖异常样本进行监督训练。Sindagi 等^[19]研究了在各种设置下用于异常检测的数据集领域转移问题。Qiu 等^[20]提出的双重加权主成分分析(DWPCA)是一种用于图像配准和表面缺陷检测的算法。Bhattacharya 等^[21]提出了交错深度伪影感知注意机制(iDAAM),用于分类异常图像中的多对象和多类别缺陷。Zeng 等^[22]将异常检测视为目标检测的

一个子集,设计了一个基于参考的缺陷检测网络(RDDN),利用模板参考和上下文参考来检测异常。

3 SAM-MR 方法

本文提出了一种基于 SAM 的混合区域匹配专家适配算法 SAM-MR(Mixture of Region Experts),其核心是使用混合专家适配器,针对不同种类的布匹瑕疵和模板图进行特征提取,同时使用布匹的掩码信息标注出瑕疵区域,通过度量并扩大瑕疵区域的差异,缩小非瑕疵区域的差异。该方法让混合专家适配器能针对不同种类的布匹瑕疵进行专有的学习,同时保证图像的非瑕疵部分的特征能处于同一分布下,这为后续的对齐和差分提供了较为优秀的图像特征表示。

SAM-MR 算法的整体框架如图 1 所示,其中包括以下 3 个部分。首先是特征提取部分,使用微调的 SAM 模型在文本提示下提取出异常图和模板图的样本特征。其次是混合专家适配器 MOAE(Mixture of Adapter Experts)模块,MOAE 模块在训练过程中学习不同瑕疵类别的特征分布。在瑕疵区域和非瑕疵区域分别采用余弦相似度^[23]作为相似性度量:对于瑕疵区域,通过最小化模板与缺陷特征之间的余弦相似度以增强差异;对于非瑕疵区域,则通过最大化其相似度以实现特征对齐。该相似度被用于构建训练损失函数。最后是 AD(Align and Differencing)模块,该模块将模板图和瑕疵图特征进行对齐,并得出明显的瑕疵区域提示,保证模型更专注于瑕疵区域的处理,从而有效地提升模型性能。

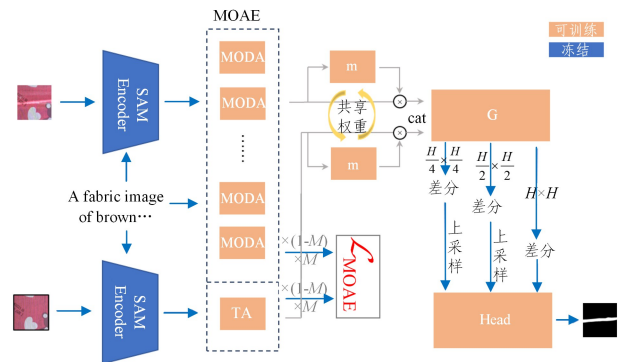


图 1 SAM-MR 算法的整体框架图

Fig. 1 Overall framework of SAM-MR method

3.1 MOAE 模块

为了确保 SAM 提取的样本特征能在不同的瑕疵类别上具有更明显的分布差异,MOAE 模块同时适配不同类别的异常样本特征。如图 2 所示,MOAE 模块在训练阶段可以分为 4 个部分:MODA (Mixture of Defect Adapters)、TA (Template Adapter)、非瑕疵区域对齐(Align of Non-defective Region)、瑕疵区域差异化(Differentiation of Defective Region)。其中 MODA, TA 部分受到混合专家模型的启发。为了更好地利用各类瑕疵之间的差异信息,Task Aware 使用文本信息^[24]来确定瑕疵的类别 k ,从而让 MODA 针对不同类别的瑕疵分别选择一个匹配的适配器。该适配器会在训练中学会对应瑕疵类别 k 的特征 $F_k \in \mathbb{R}^{C_k \times H_k \times W_k}$:

$$F_k = MODA_k(SAM(T_k, I_k), T_k) \quad (1)$$

其中, $MODA_k$ 是瑕疵类别 k 对应的 MODA; T_k 和 I_k 分别是文本标注和带有瑕疵的印花图; C_k 是每个瑕疵类别的适配特征通道数; H_k 和 W_k 分别是每个瑕疵类别的适配特征高度和

宽度。而 TA 通过训练得到模板图的特征 $F' \in \mathbb{R}^{C_t \times H_t \times W_t}$, 其中 C_t 是模板图的适配特征通道数, H_t 和 W_t 分别是模板图的适配特征高度和宽度。 F_k 和 F' 将用于 MOAE 的训练和后续的特征对齐与差分。

在 MOAE 的训练阶段, 为了确保模型能更精确的学习瑕疵区域和非瑕疵区域, 通过真实标签得出瑕疵区域掩码 M_d 和非瑕疵区域掩码 $M_{\text{non-d}}$, 从而更精确地分割出两个区域。首先根据瑕疵区域掩码 M_d 分别在瑕疵特征图 F_k 和模板特征图 F' 中裁剪出对应的瑕疵区域 D_k 和 D_t , 然后通过余弦相似度计算出该部分对应的损失 $\mathcal{L}_{\text{difference}}$:

$$\mathcal{L}_{\text{difference}} = \frac{D_k \cdot D_t}{\|D_k\| \|D_t\|} \quad (2)$$

同理, 根据非瑕疵区域掩码 $M_{\text{non-d}}$ 分别在瑕疵特征图 F_k 和模板特征图 F' 中裁剪出对应的非瑕疵区域 N_k 和 N_t , 计算得出 $\mathcal{L}_{\text{same}}$, 其定义为:

$$\mathcal{L}_{\text{same}} = \frac{N_k \cdot N_t}{\|N_k\| \|N_t\|} \quad (3)$$

最后, 用于训练 MOAE 并实现非瑕疵区域对齐和瑕疵区域差异化的损失函数 $\mathcal{L}_{\text{MOAE}}$ 形式化表示为:

$$\mathcal{L}_{\text{MOAE}} = \frac{1}{\mathcal{L}_{\text{difference}} + \epsilon} + \alpha \mathcal{L}_{\text{same}} \quad (4)$$

其中 ϵ 为确保 $\mathcal{L}_{\text{difference}} + \epsilon$ 不小于 0 的超参数, α 是表示 MOAE 对于非瑕疵区域的关注程度的超参数。由于非瑕疵区域面积大于瑕疵区域, MOAE 更容易关注非瑕疵区域的统一性, 忽视了瑕疵区域的差异性, 这对瑕疵的检出是不利的。因此需要一个 α 来平衡 MOAE 对非瑕疵区域和瑕疵区域的关注程度。

如图 2 所示, 在推理阶段, MOAE 仅有两部分, 包括 MODA 和 TA。训练好的 MODA 可以在文本的提示下有效地提取出对应瑕疵种类的特征 F_k , 并在瑕疵区域很好地保持了和 TA 提取的特征 F' 的差异, 同时也在非瑕疵区域尽量保证了特征分布的统一性。

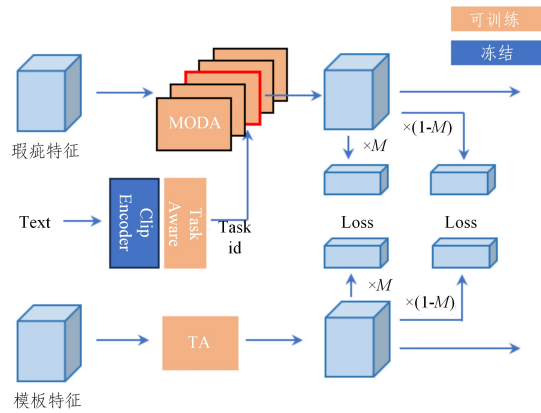


图 2 MOAE 模块架构图

Fig. 2 Architecture diagram of MOAE module

3.2 AD 模块

如图 1 所示, 为了让模型能更好的注意到瑕疵特征和非瑕疵特征的各个区域的匹配程度, AD (Align and Differencing) 模块使用共享的子网络 m 在瑕疵特征图和非瑕疵特征图上生成权重图 W_k 和 W_t 。

$$W_k = m(F_k) \quad W_t = m(F_t) \quad (5)$$

权重图 W 的大小与特征图 F 大小保持一致, 将 W 与 F

相乘, 突出显示特征中的同源区域。随后进一步对 MOAE 提取的瑕疵特征和非瑕疵特征进行加权, 再将两个特征连接后送入交互网络 G 。

$$I_i = G_i(\text{concat}(F_k \cdot W_k, F_t \cdot W_t)) \quad (6)$$

为了兼顾对大瑕疵和小瑕疵的尺度信息, 通过交互网络 G 得出 3 个尺度的交互信息 I_1, I_2 和 I_3 , 其大小分别是 $H \times H, \frac{H}{2} \times \frac{H}{2}, \frac{H}{4} \times \frac{H}{4}$, 同时在不同尺度上对瑕疵特征和非瑕疵特征进行差分得出 3 个尺度的差分特征图。为了统一 3 个尺度的信息, 将 $\frac{H}{2} \times \frac{H}{2}$ 和 $\frac{H}{4} \times \frac{H}{4}$ 尺度的差分特征图上采样到与 $H \times H$ 一致大小, 最后送入头部网络进行信息处理输出。

3.3 训练与损失函数

SAM-MR 模型的训练损失包括两部分, 一个是 MOAE 的对齐和差异化损失 $\mathcal{L}_{\text{MOAE}}$, 另一个是模型的预测结果与真实标签的损失 \mathcal{L}_{MSE} 。

$$\mathcal{L}_{\text{MSE}} = \frac{\|P_k - GT_k\|_2}{H \times W} \quad (7)$$

其中, P_k 是模型对瑕疵类别 k 的预测结果, GT_k 是瑕疵类别 k 的真实标签, H 和 W 分别是真实标签图的高和宽。

\mathcal{L}_{MSE} 能确保模型正确学习到合适的瑕疵数分布的信息, 并同时使 MOAE 模块中对应的瑕疵种类适配器 MODA 可以针对预测结果进行参数迭代, 促使 MOAE 的参数靠近全体参数的最优解。

将上述两个损失进行加权求和, 构成模型训练的总损失:

$$\mathcal{L} = \lambda_{\text{MOAE}} \mathcal{L}_{\text{MOAE}} + \lambda_{\text{MSE}} \mathcal{L}_{\text{MSE}} \quad (8)$$

其中, λ_{MOAE} 和 λ_{MSE} 分别为 $\mathcal{L}_{\text{MOAE}}$ 损失和 \mathcal{L}_{MSE} 损失的权重超参数。在全类别的瑕疵样本上进行训练, MOAE 模块和 AD 模块可以协调学习到适配于全种类瑕疵预测的参数, 以此保证所有 MODA 学习到的样本特征可以同时适配到 AD 模块。

4 实验

4.1 数据集与指标

为了验证 SAM-MR 的有效性, 将本文方法与 PromptAD^[25], WinCLIP^[26] 在 FabricEyes 上进行多类别的实验对比。

FabricEyes 是从严格的印刷瑕疵检测生产线中精心收集而成, 按照各类型瑕疵的特点, 将瑕疵分为 12 个不同种类。该数据集总计拥有 40000 个样本, 同时对纺织品的材质进行了标记, 分为棉和非纯棉, 并标定了棉含量。

如图 3 所示, 在训练集和测试集中, 每个样本都有对应的模板印花图、带有瑕疵的印花图、文本标注和瑕疵标注, 如变形、模糊印刷、色差、色彩扩散、色彩缺失、孔洞、缝头、伪色、去浆、污渍、水渍和白条等类别。

在评估指标方面, 与当前主流的其他异常检测研究一致, 使用 AUROC^[27] (Area Under the Receiver Operating Curve) 指标来衡量模型检测异常样本的性能。同时使用 F1-max^[28] (F1 Score Maximization) 来评估模型的最佳类别检测性能。

在训练方面, $\mathcal{L}_{\text{MOAE}}$ 中的 ϵ 设置为 1.5, α 设置为 1。使用 Adam 优化器进行训练, 训练的轮数为 100 轮, 初始学习率设置为 1×10^{-3} , 并采用指数衰减策略, 衰减率设为 0.8, 批量大小设为 8。



A fabric image of brown bear.pink background with Seam Pucker anomaly. The textile prints include white hearts and white polka dots. The heart pattern presents a dense arrangement.

图3 FabricEyes 数据集形式

Fig. 3 FabricEyes dataset format

4.2 实验结果与分析

在不同样本量下异常检测方法的比较如表 1 所列, 本文分别使用少量的样本和使用全量的样本对所提出的方法进行了训练, 然后在 30 个小类别上进行了实验, 并展示了平均性能和最佳的类别性能, 并以加粗的形式表示最好的结果。

表 1 不同样本量下异常检测方法的性能比较

Table 1 Performance comparison of anomaly detection methods across different shots

样本量	方法	AUROC	F1-max	ρ AUROC	ρ F1
少量 样本	PromptAD	59.8	85.8	78.9	77.5
	WinCLIP	59.6	84.2	71.9	72.1
	Ours	68.7	86.6	79.4	78.2
全量 样本	PromptAD	60.6	86.4	79.8	82.1
	WinCLIP	60.9	86.7	74.7	79.5
	Ours	73.2	87.3	80.5	82.4

SAM-MR 在模型的总体性能以及最佳的单类别性能下均高于之前的方法。在少量样本, 训练不充分的情况下, SAM-MR 在对比中获得了最佳的表现效果。具体来说, 相较于现有的多类别异常检测方法 PromptAD 和 WinCLIP 在 FabricEyes 上将 AUROC 从 59.8, 59.6 提升到 68.7, 这充分说明了本文算法架构对于特定任务的异常有着更良好的适配

表 2 主流方法在多类异常检测和成分检测性能上的比较

Table 2 Comparison of main stream methods on multi-class anomaly detection performance

模式名	PromptAD			WinCLIP			Ours		
	AUROC	ρ AUROC	ACC	AUROC	ρ AUROC	ACC	AUROC	ρ AUROC	ACC
小白色矩形, 灰色背景	98.44	90.80	95.1	84.38	75.69	80.2	97.66	95.49	87
白色条纹, 棕色背景	95.83	82.79	90.5	80.42	86.06	79.5	88.33	88.30	79
小白点, 深红色背景	94.44	76.25	82.3	87.78	67.17	59.3	100.00	90.36	80.2
蓝色条纹, 绿色背景	93.58	74.93	84.2	64.73	63.42	93.6	99.66	93.38	93.1
白色叶子方块, 黑色背景	90.48	70.13	81.1	70.24	26.83	63.2	100.00	94.18	93.4
黑色圆圈, 白色背景	87.00	51.14	93.6	72.67	66.38	84.5	87.67	54.08	84.2
红绿花朵, 白色背景	84.87	59.71	67.4	82.24	44.59	71.7	100.00	95.21	71.8
红色图案, 黑色背景	76.63	61.46	81.8	84.24	57.55	66.1	82.61	94.32	68.8
黑色背景	75.78	96.80	77.5	46.09	74.46	85.8	83.59	97.35	81.9
黄色蜜蜂, 蓝色背景	75.36	72.73	68.1	76.33	86.01	73.4	89.37	93.84	73.6
均值	87.24	73.67	79.2	74.91	64.82	76.2	92.89	89.65	0.082

注: 由于空间限制, 仅展示了前 10 种小类别的结果。

PromptAD 和 WinCLIP 的 AUROC 在差异较大的小类别之间, 预测性能相差大, 对于部分类型差异大的瑕疵不能实现相近的预测能力。在 PromptAD 中, 其 AUROC 在小白色矩形, 灰色背景的小类别下高达 98.44, 然而在黄色蜜蜂, 蓝色背景小类别下只有 75.36。WinCLIP 也表现出了同样的问题, 其 AUROC 最高值为 87.78, 而最低值为 46.09。本文方法则保持了相对较小的性能差异, 对于各小类的瑕疵检测性能更佳。此外, 表 2 展示了 SAM-MR 模型与其他主流方法 (PromptAD 和 WinCLIP) 在多类别异常检测以及成分检测性

效果, 更容易达到更准确、更精细的分割结果, 而另外两个模型由于架构问题, 很难适用于下游任务特定的多类别瑕疵检测, 特别是在检测细微和小瑕疵上。

图 4 展示了在全量样本的训练情况下的布匹瑕疵效果可视化的结果。同时, 表 2 列出的数据也表明 SAM-MR 的效果提升远大于其他两种模型, PromptAD 和 WinCLIP 的 AUROC 提升分别是 0.8 和 1.3, 而 SAM-MR 的 AUROC 提升高达 4.5。SAM-GMR 在各类别上的性能差异最小, 表明该方法在多类别布匹瑕疵检测中具有更均衡的泛化能力。

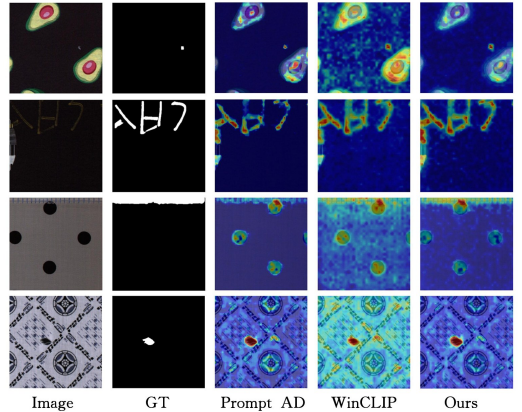


图 4 各模型在 FabricEyes 数据集的结果

Fig. 4 Results of various models on the FabricEyes dataset

表 2 列出了 3 种方法在本文收集的 Fabric-Composition 数据集上的表现最好的前 10 个小类别的缺陷检测和成分测定性能。前 10 种产品的表现范围适中, 平均 AUROC 分数处于 74.91 (WinCLIP) 和 92.89 (SAM-MR) 之间。模型在具有多样纹理、结构和颜色变化的复杂环境中, 其有效性有所下降。复杂多样的瑕疵类别突显了当前各类模型面临的问题, 表明了不同类别异常差别的重要性。

能上的比较结果。从表中的 ACC (准确率) 指标可以看出, 本文方法在成分判别上表现出色。例如, 在均值方面, SAM-MR 的 AUROC 达到了 92.89%, 显著高于 PromptAD 的 79.2% 和 WinCLIP 的 76.2%。这表明 SAM-MR 模型在识别布料成分方面具有很高的准确性, 例如能够精准判断材质是否为纯棉。这得益于模型集成的成分检测任务, 使其能够在完成瑕疵定位的同时, 实现对异常成分 (如纺织品材质) 的语义识别。

Fabric Composition 数据集中的图像示例如图 5 所示。

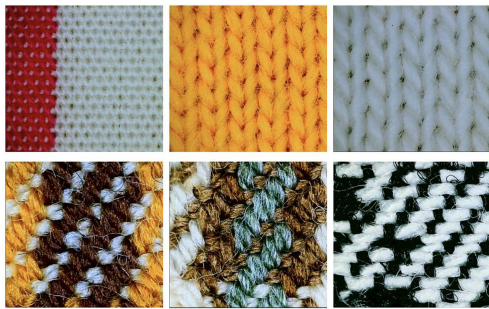


图5 Fabric Composition 数据集中的图像示例

Fig. 5 Image examples in the Fabric Composition dataset

在 FabricEyes 上对 SAM-MR 算法进行了消融实验,以验证每个模块的有效性,结果如表 3 所列。

表 3 SAM-MR 在 FabricEyes 上各个模块的消融实验

Table 3 Ablation experiments of each module of SAM-MR on the FabricEyes dataset

MODA&TA	模块		指标	
	对齐 & 差异化	AD	AUROC	pF1
✓			71.1	71.9
✓		✓	74.4	76.1
✓	✓		77.6	79.3
✓	✓	✓	80.5	82.4

从实验结果可以看出,MOAE 中的非瑕疵区域对齐和瑕疵区域差异化方法,让模型的 MODA 模块和 TA 模块学习到了瑕疵和非瑕疵的特征差别,有效提高了模型对瑕疵的区分能力,同时在 AD 模块的进一步特征对齐和差分下能更有效,更精确地得出瑕疵区域。而在缺少对齐差异化方法后,没有拉开正常特征和异常特征的差异情况下,AD 模块对于异常和正常区分度不大的特征,较难得出瑕疵的显著区域。

结束语 本文提出的基于 SAM 的混合区域匹配专家适配算法 SAM-MR 在布匹瑕疵检测任务中展现了显著的优势。通过引入混合专家适配器和对瑕疵区域与非瑕疵区域的精细化处理,SAM-MR 能够有效提取和对齐不同类别的特征,从而提升模型在多类别异常检测中的性能。

未来的研究可以进一步探索 SAM-MR 在其他领域的应用潜力,并考虑引入更多的上下文信息和多模态数据,以进一步提升模型的鲁棒性和适应性。此外,针对复杂环境下的瑕疵检测,如何优化模型以应对多样化的纹理和颜色变化,仍然是一个值得深入探讨的方向。

参考文献

[1] KIRILLOV A, MINTU E, RAVI N, et al. Segment anything [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023:3992-4003.

[2] CAO Y K, XU X H, ZHANG J N, et al. A Survey on Visual Anomaly Detection: Challenge, Approach, and Prospect [J]. arXiv:2401.16402, 2024.

[3] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018:801-818.

[4] SAEED M, EBRAHIMPOUR R. Mixture of experts: a literature

survey [J]. Artificial Intelligence Review, 2014, 42: 275-293.

[5] ZHOU Y Q, LEI T, LIU H X, et al. Mixture-of-experts with expert-choice routing [C]// Advances in Neural Information Processing Systems. 2022:7103-7114.

[6] RÜCKLÉ A, GEIGLE G, GLOCKNER M, et al. AdapterDrop: On the efficiency of adapters in transformers [J]. arXiv:2010.11918, 2020.

[7] CHEN Z, DUAN Y, WANG W, et al. Vision transformer adapter for dense predictions [J]. arXiv:2205.08534, 2022.

[8] SUNG Y L, CHO J, BANSAL M. V-Adapter: Parameter-efficient transfer learning for vision-and-language tasks [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022:5227-5237.

[9] CHAN C J, PANG G K H. Fabric defect detection by Fourier analysis [J]. IEEE Transactions on Industry Applications, 2000, 36(5): 1267-1276.

[10] HAN Y J, YU H J. Fabric defect detection system using stacked convolutional denoising auto-encoder trained with synthetic defect data [J]. Applied Sciences, 2020, 10(7): 2511.

[11] GHOSH A, GUHA T, BHAR R B, et al. Pattern classification of fabric defects using support vector machines [J]. International Journal of Clothing Science and Technology, 2011, 23(2/3): 142-151.

[12] KATTENBORN T, LEITLOFF J, SCHIEFER F, et al. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 173: 24-49.

[13] HE K, GKIOXARI G, DOLL? R P, et al. Mask R-CNN [C]// Proceedings of the IEEE International Conference on Computer Vision. 2017:2961-2969.

[14] DAI J, LI Y, HE K, et al. R-FCN: Object detection via region-based fully convolutional networks [C]// Advances in Neural Information Processing Systems. 2016.

[15] VILLA M, DARDENNE G, NASAN M, et al. FCN-based approach for the automatic segmentation of bone surfaces in ultrasound images [J]. International Journal of Computer Assisted Radiology and Surgery, 2018, 13: 1707-1716.

[16] JI S P, SHEN Y Y, LU M, et al. Building Instance Change Detection from Large-Scale Aerial Images using Convolutional Neural Networks and Simulated Samples [J]. Remote Sensing, 2019, 11(11): 1343.

[17] LIU R Y, KUFFER M, PERSELLO C. The Temporal Dynamics of Slums Employing a CNN-Based Change Detection Approach [J]. Remote Sensing, 2019, 11: 2844.

[18] NEMOTO K, HAMAGUCHI R, SATO M, et al. Building change detection via a combination of CNNs using only RGB aerial imageries [C]// Proceedings of the SPIE. 2017.

[19] SINDAGI V A, SRIVASTAVA S. Domain adaptation for automatic OLED panel defect detection using adaptive support vector data description [J]. International Journal of Computer Vision, 2017, 122(2): 193-211.

[20] QIU K, TIAN L, WANG P. An effective framework of automated visual surface defect detection for metal parts [J]. IEEE Sensors Journal, 2021, 21(18): 20412-20420.

[21] BHATTACHARYA G, MANDAL B, PUHAN N B. Interleaved

- deep artifacts aware attention mechanism for concrete structural defect classification [J]. *IEEE Transactions on Image Processing*, 2021, 30: 6957-6969.
- [22] ZENG Z, LIU B, FU J, et al. Reference-based defect detection network [J]. *IEEE Transactions on Image Processing*, 2021, 30: 6637-6647.
- [23] RAHUTOMO F, KITASUKA T, ARITSUGI M. Semantic cosine similarity [C]// *The 7th International Student Conference on Advanced Science and Technology ICAST*. 2012.
- [24] SINGH A, HU R, GOSWAMI V, et al. FLAVA: A foundational language and vision alignment model [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022:15638-15650.
- [25] LI X, ZHANG Z, TAN X, et al. PromptAD: Learning prompts with only normal samples for few-shot anomaly detection [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024:16838-16848.
- [26] JEONG J, ZOU Y, KIM T, et al. WinCLIP: Zero-/few-shot anomaly classification and segmentation [C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023:19606-19616.
- [27] HANLEY J A, MCNEIL B J. The meaning and use of the area under a receiver operating characteristic(ROC) curve [J]. *Radiology*, 1982, 143(1): 29-36.
- [28] FUJINO A, ISOZAKI H, SUZUKI J. Multi-label text categorization with model combination based on F1-score maximization [C]// *Proceedings of the Third International Joint Conference on Natural Language Processing*. 2008.



LUO Qifeng, born in 1982, master, senior engineer. His main research interests include the application of new technology in power system, artificial intelligence, substation equipment operation, maintenance management, etc.



PENG Bo, born in 1994, doctoral student, lecturer. His main research interests include are machine learning and deep learning.