

融合注意力机制的弱监督纸板表面缺陷检测

吴维崧¹, 涂福泉¹, 罗迎九¹, 杨家瑜¹, 韩天宇¹, 汪曙峰², 涂楚杰³

(1. 武汉科技大学 冶金装备及其控制教育部重点实验室, 武汉 430081; 2. 中冶南方武汉钢铁设计研究院有限公司, 武汉 430080; 3. 贺利氏创特(武汉)工程技术有限公司, 武汉 430070)

摘要: **目的** 针对目前表面缺陷检测方法因缺少实例级标签, 使神经网络在工业检测上的应用受到限制的问题。本文面向实际的纸板表面缺陷检测任务, 提出弱监督学习下融合卷积和注意力机制的神经网络算法。**方法** 该网络通过将通道注意力模块和梯度类激活映射模块相结合, 进一步提高类激活图的精细度, 实现纸板表面缺陷的精确定位; 同时通过倒残差结构和上采样层的组合操作, 进一步细化浅层特征提升网络的特征提取能力, 加快网络收敛速度。**结果** 通过在公开的纸板缺陷数据集上进行实验, 本文提出的算法在使用图像级标签训练的情况下, 分类正确率与定位正确率分别达到 99.0% 和 92.2%, 验证了该算法的有效性。**结论** 避免了实例级标签数量较少和过于主观的缺点, 为基于机器人的缺陷纸板剔除奠定了基础。

关键词: 弱监督学习; 对象定位; 深度学习; 纸板表面缺陷检测; 自注意力

中图分类号: TB487; TP391.41 文献标志码: A 文章编号: 1001-3563(2024)03-0201-07

DOI: 10.19554/j.cnki.1001-3563.2024.03.023

Weakly Supervised Cardboard Surface Defect Detection with Attention Mechanism

WU Weisong¹, TU Fuquan¹, LUO Yingjiu¹, YANG Jiayu¹, HAN Tianyu¹,
WANG Shufeng², TU Chujie³

(1. Key Laboratory of Metallurgical Equipment and Control, Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China; 2. WISDRI Engineering & Research Incorporation Limited, Wuhan 430080, China; 3. Heraeus TROT (Wuhan) Engineering and Technology Co., Ltd., Wuhan 430070, China)

ABSTRACT: The application of deep neural networks in industrial inspection is limited due to the lack of instance-level labels. To address this issue, the work aims to propose a neural network algorithm that combines convolution and attention mechanisms under weakly supervised learning for practical surface defect detection on cardboard. By integrating channel attention modules and gradient-based activation mapping modules, this network enhanced the precision of class activation maps and realized the precise localization of cardboard surface defects. Additionally, a combination of inverted residual structures and upsampling layers was utilized to refine shallow features and improve the network's feature extraction capabilities, thereby accelerating the convergence speed. Experiments were carried out on the publicly available cardboard defect dataset, achieving classification accuracy and localization accuracy of 99.0% and 92.2% respectively under the training with image-level labels and demonstrating the effectiveness of the proposed algorithm. The disadvantages of a small number of instance-level labels and excessive subjectivity are avoided, which lays a foundation for the removal of defective cardboard based on robots.

KEY WORDS: weakly supervised algorithm; object localization; deep learning; cardboard surface defect detection; self-attention

收稿日期: 2023-06-06

基金项目: 国家自然科学基金(51701145)

*通信作者

近年来,随着我国《推进快递业绿色包装工作实施方案》的出台,“限塑令”到“禁塑令”的转变,物流行业和包装行业的快速发展,纸板作为绿色环保包装材料的重要组成部分,其生产质量和智能化生产有着重要的影响,而在纸板制造的过程中,受到环境和生产过程因素的影响,可能导致纸板出现破损和裂纹等缺陷,如果不能及时地找出问题所在,将会造成一定的损失。因此,对生产纸板缺陷进行实时检测是保障纸板生产质量的关键技术之一。然而相较于其他类型的表面缺陷检测,不同类型的表面存在特征差异,而且算法的设计是基于不同类型表面缺陷的特点和需求进行设计和优化,因此需要对纸板表面缺陷检测方法进行相应改进。

采用传统的人工方式检测产品表面,不仅耗时费力且存在较大的主观性,严重影响检测的准确性。为了解决这个问题,有些工程技术人员将特征工程的方法(主要以数据预处理、手工设计特征和分类器设计组成)应用于缺陷检测之中,并取得了较好的效果^[1-5]。但该方法依赖设计者的专业知识,因此在很大程度上限制了这类方法的推广应用。

随着深度学习(Deep Learning, DL)理论的发展和完善,基于深度学习的产品表面缺陷检测方法得到了成功应用。姜旭辉^[6]提出对 Yolo v5 算法的激活函数、损失函数和预测框筛选方式的改进,提高对纸板缺陷定位的准确度。叶宇星等^[7]提出使用 Ghost 卷积替换 CSP 模块中的卷积,提高 Yolo v5 算法骨干网络对腌制蔬菜真空包装的特征提取能力。然而,深度学习方法需要大量的实例级标签(标注缺陷图像的缺陷类别和位置)进行训练,但在实际生产中,实例级标签数据量少。为了解决这个问题,研究者们提出使用弱监督学习的方法,即使用图像级标签(仅标注缺陷图像的缺陷类别)代替实例级标签来进行缺陷定位。何彧等^[8]提出对卷积神经网络融合多层特征,提高弱监督方法对钢板缺陷定位的准确度。孙美君等^[9]提出对深度卷积神经网络融合注意力机制,提高弱监督方法对电机转向器缺陷定位的准确度。这些弱监督学习的方法广泛推进了缺陷检测的研究。然而,使用弱监督学习方法进行缺陷定位时仍然存在一些问题,包括卷积神经网络的卷积层不能建立全局图像的长距离连接,导致使用基于弱监督学习方法时,只能定位目标对象的粗略区域^[10]。

针对上述问题,本文提出融合注意力机制的弱监督纸板表面缺陷检测算法。该算法利用改进网络生成的类激活图进行定位,而且通过卷积模块与自注意力模块的结合,使网络不但能够学习全局表征,并且能够在参数更少的时候学习更多的表征。为了使生成的类激活图对目标区域的激活更加精密,算法将梯度类激活映射模块和通道注意力模块相结合。利用通道注意力模块对有效信息特征的增强和

对无效信息特征的抑制,从而对类激活图生成的目标区域进行聚焦,并且利用梯度类激活映射模块的类别梯度信息进一步对有效信息特征加强,达到对目标区域修正的目的,使纸板表面缺陷检测的定位性能得到提高。

1 成像系统设计

本研究对纸板表面缺陷检测的实验平台搭建如图 1 所示。图 1a 是纸板图片采集模块,待检测纸板经过光源的充分照射,使得背景信息与缺陷信息具备明显的区别,再通过工业相机,得到高分辨率的样本图片。图 1b 是纸板缺陷识别模块,将建设好的数据集传入用 Pytorch 设计出的纸板表面缺陷检测网络中,对缺陷图片进行训练和测试。

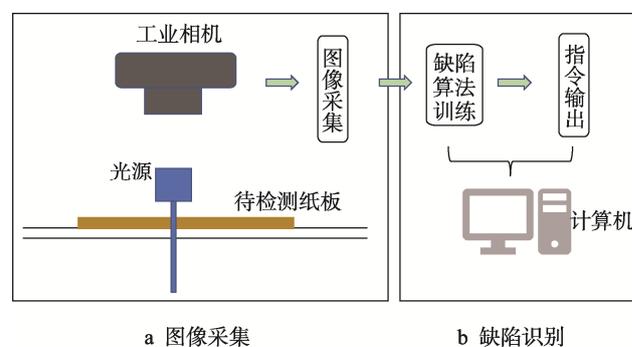


图 1 纸板表面缺陷识别过程
Fig.1 Cardboard surface defect identification process

2 网络搭建

2.1 整体架构

本文提出的弱监督学习缺陷检测网络如图 2 所示,整个网络由 2 个部分组成:主干特征提取网络和定位网络。其中主干特征提取网络包括了基于自注意力的模块(Swin Transformer Block, STB)^[11]和倒残差结构模块(Inverted Residual Block, IRB),定位网络包括了通道注意力模块(Squeeze and Excitation Block, SEB)和梯度类激活映射模块(Grad CAM Block, GCB)。

本文提出网络的特点体现在以下 3 点:基于自注意力的模块学习全局表征,再根据卷积的空间归纳偏差特性结合卷积模块,使网络在参数更少的时候学习更多的表征,并且收敛相对更快;引入上采样层后,提高特征图的分辨率,使得网络能够更好地捕捉图像中的细节信息;将通道注意力模块和梯度类激活映射模块相结合,对生成的类激活图激活的目标区域进行聚焦和修正,使网络的定位性能得到提高。

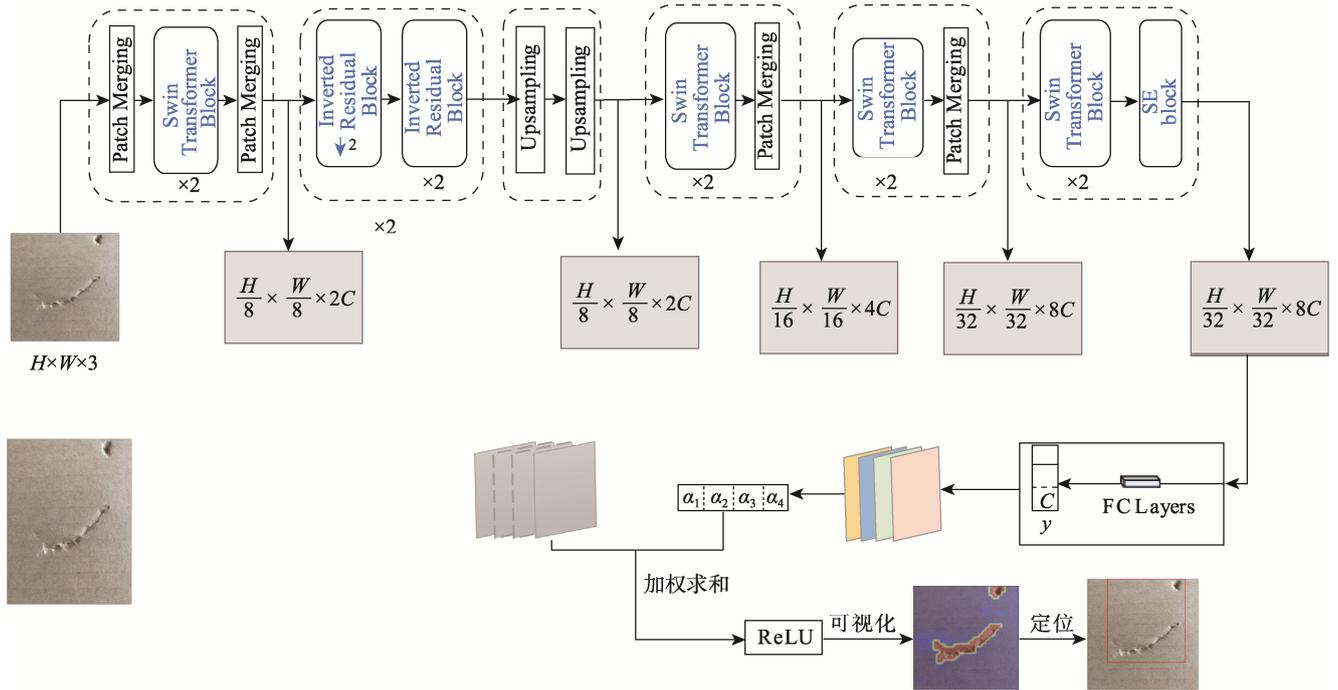


图 2 弱监督纸板表面缺陷检测网络
Fig.2 Weakly supervised cardboard surface defect detection network

2.2 主干网络

基于自注意力的网络通过具备学习全局表征的能力, 在图像领域的各种任务中都获得了优异的成绩。然而, 完全基于自注意力的弱监督网络需要更多的学习参数和学习成本来获得较好的性能, 这给缺陷检测带来了困难。因此, 本文的弱监督检测网络的主干网络由注意力模块和倒残差结构模块组成。

当缺陷样本输入到网络时, 通过 STB 层的自注意力机制获得整张图片的全局交互, 即整张图片上每个特征与图上其他的特征产生交互, 得到整张图片的全局语义信息。STB 层产生的特征图经 IRB 层的特征提取, 使网络通过更少的参数去学习表征, 降低模型的计算成本。上采样层通过对特征图进行空间尺寸的放大, 使得网络后面的 STB 层能够更快地学习到缺陷特征, 网络更容易收敛。Patch Merging 层是下采样层, 将整张缺陷图片分割成不重叠的图片块, 然后在通道方向进行拼接, 最后对每个像素的维度进行线性变换。

STB 层结构细节如图 3a 所示, 由 2 个模块串联组成, 其模块通过多头自注意力机制 (Multi-Head Self Attention, MSA) 计算每个像素之间的相关性。流程如下: 首先将输入特征图进行层归一化操作 (Layer Normalization, LN), 然后在窗口内进行多头自注意计算 (W-MSA), 最后在进行一次层归一化操作后传入多层感知器 (Multi-Layer Perceptron, MLP); 另一个模块是在移动后的窗口内进行多头自注意计算 (SW-MSA)。连续 Swin Transformer Block 计算如式

(1) 所示。

$$\begin{aligned} \hat{z}_i &= WMSA(\ln(z_{i-1})) + z_{i-1} \\ z_i &= MLP(\ln(\hat{z}_i)) + \hat{z}_i \\ \hat{z}_{i+1} &= SWMSA(\ln(z_i)) + z_i \\ z_{i+1} &= MLP(\ln(\hat{z}_{i+1})) + \hat{z}_{i+1} \end{aligned} \quad (1)$$

式中: z_i 和 z_{i+1} 分别为基于窗口的多头自注意和基于移动窗口的多头自注意的输出特征向量。

IRB 层结构细节如图 3b 所示, 其模块通过深度卷积 (Depthwise Convolution, DW) 提取特征, DW 层对输入特征图的每个输入通道应用单个卷积核操作, 并且每个卷积核的通道数为 1, 相较于普通卷积层, DW 层的卷积操作大大降低了计算量。IRB 层流程如下: 利用 1×1 的卷积对输入特征图进行升维, 然后传入 DW 层捕捉特征图中的缺陷特征, 最后通过 1×1 的卷积对特征图进行降维。值得注意的是, 升维的目的是降低信息通过 ReLU 激活函数丢失的信息量。

2.3 定位网络

类激活映射^[12]是属于注意力方法, 它常用于可视化卷积神经网络模型, 使得模型具备可解释性和透明性。然而, 使用完全基于类激活映射生成的类激活图进行判别区域的定位, 只能定位粗略的空间对象区域, 这给缺陷定位带来了困难。因此, 本文的弱监督检测网络的定位网络由通道注意力模块和梯度类激活映射模块组成。

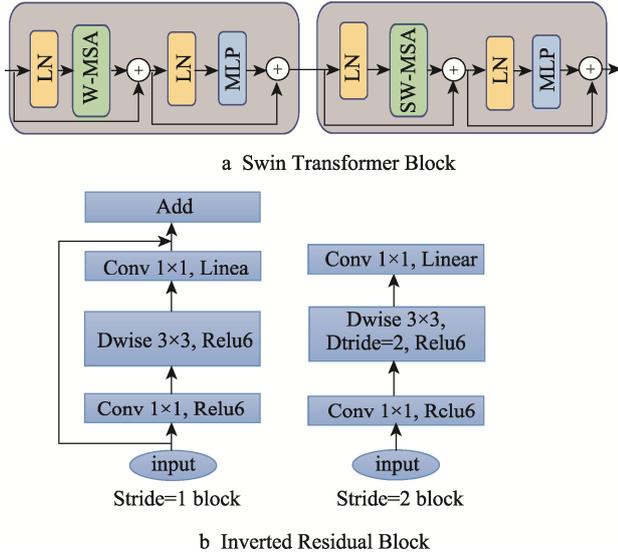


图3 主干网络的组成模块
Fig.3 Component modules of the backbone network

主干网络的输出特征图传到网络时，通过 SEB 层的注意力机制，会给正向传播得到输出特征图的各个通道赋予不同的权重信息，将权重分别与对应特征图各通道相乘后，对有效信息特征的关注度得到增强，对无效信息特征的关注度被抑制，使类激活图激活的目标区域聚焦。SEB 层产生的输出特征图传入 GCB 层后，GCB 层以反向传播的类别梯度信息给输出特征图的各个通道赋予不同的权重信息，将权重分别与对应特征图各通道相乘后，进一步对有效信息特征加强，达到对目标区域修正的目的。两模块的目的都是为蕴含高级语义信息的通道分配较大的权重信息，而给与检测类别无关特征的通道分配较小的权重信息。因此同时具备 SEB 层对目标区域聚焦和 GCB 层的目标区域修正，使网络的目标定位更加准确。

SEB 层结构细节如图 4a 所示，其模块通过增强输出特征图通道之间的交互，从而增强有用的特征并抑制无用的特征。SEB 层流程如下：首先利用全局平均池化层的操作得到输出特征图每个通道的重要性 w_c ；然后通过 Sigmoid 激活函数学习通道之间的非线性相互作用，调整每个通道的重要性；最终权重与对应通道的特征图相乘得到判别区域聚焦后的特征图。

$$w_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W A_c(i, j) \quad (2)$$

式中： A 为缺陷图片经过主干网络得到最终的输出特征图； H 和 W 分别为第 c 个特征图的高和宽； $A_c(i, j)$ 为第 c 个特征图上空间位置 (i, j) 的像素点。

GCB 层结构细节如图 4b 所示，其模块通过生成的类激活图实现缺陷定位。GCB 层流程如下：利用某类别回流的梯度信息来生成输出特征图通道的重

要性 α_k^c ，在与对应通道的特征图相乘后得到判别区域修正后的特征图，再通过加权求和得到类激活图 $L_{\text{Grad-CAM}}^c$ 从而实现缺陷定位。

$$\alpha_k^c = \frac{1}{z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (4)$$

$$L_{\text{Grad-CAM}}^c = \text{ReLU}(\sum_k \alpha_k^c A^k) \quad (5)$$

式中： A_{ij}^k 为第 k 个特征图上空间位置为 (i, j) 的像素点； $\frac{\partial y^c}{\partial A_{ij}^k}$ 为主干网络关于检测类别 c 的预测评分 y^c 对第 k 个特征图上不同像素点的梯度，即得到特征图 A^k 每个像素对预测结果的重要程度； z 为特征图 A^k 像素点的个数。

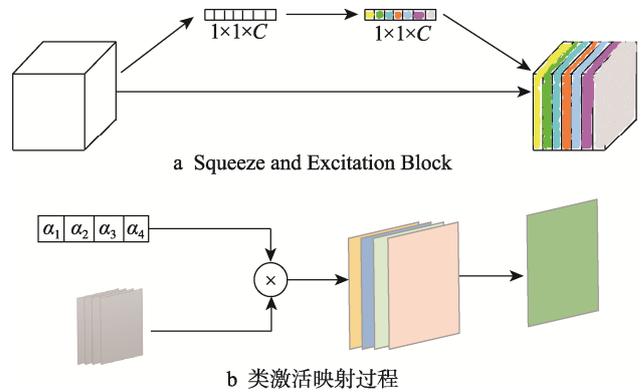


图4 定位网络的组成模块
Fig.4 Component modules of the localization network

3 实验结果分析

3.1 实验数据与评价标准

实验采用公开发表的纸板缺陷数据集，如图 5 所示，该数据集包括纸板破损、起泡、划痕和排骨痕等缺陷，但由于数据集将所有缺陷均化为一个类 (detect)，因此为了适合网络，收集了生产过程中没有缺陷的纸板图片作为另外的类 (no_detect)，形成了最终的数据集。数据集一共包含 1 539 张样本，其中训练集与测试集的比为 4 : 1，如表 1 所示。



图5 纸板缺陷示意图
Fig.5 Illustration of cardboard defects

表 1 训练集和测试集分布
Tab.1 Training set and test set distribution

纸板数据集类别	训练集数量	测试集数量
detect	845	211
no_detect	387	96

缺陷检测的精度主要取决于分类精度和定位精度。分类精度取决于分类结果的正确率, 定位精度的测试则使用阈值法来对缺陷进行定位。在实验中, 选择将类激活图里像素值大于 81 的最大连通区域作为纸板缺陷的识别区域, 并且将预测框与真实框 IOU^[13] 大于 0.5 的认定为识别成功。为了更准确地评估模型的性能, 综合考虑最终选择分类精度、定位精度、精确率 (Precision, P)、召回率 (Recall, R) 和交并比 (Intersection Over Union, I_{IOU}) 5 个指标。具体计算公式如下:

$$P = \frac{T_p}{T_p + F_p} \quad (6)$$

$$R = \frac{T_p}{T_p + F_N} \quad (7)$$

$$I_{IOU} = \frac{A \cap B}{A \cup B} \quad (8)$$

其中, 在本实验中包含无缺陷和有缺陷纸板 2 个分类: 用有缺陷纸板图像表示正样本, 用无缺陷纸板图像表示负样本。式中: T_p 表示正样本被预测为有缺陷纸板的样本数; F_p 表示负样本被预测为有缺陷纸板的样本数; F_N 表示正样本被预测为无缺陷纸板的样本数; A 表示预测的检测框; B 表示真实的检测框。

3.2 实验细节

实验环境为 Window11 操作系统, 选用 i7-12700KF CPU, 32 GB RAM 和 NVIDIA GeForce RTX 3080Ti GPU 显卡, 编程语言为 python3.9, 深度学习开发框架为 Pytorch。

训练时, 缺陷检测网络采用 AdamW 梯度下降法^[14]进行训练, 其中学习率为 0.000 1, 权重衰减系数为 0.05, 最小批次尺寸为 14, 训练迭代次数为 200

轮。在模型搭建过程中, 由于 Swin Transformer Block 的输出是长度为长乘宽的序列, 因此在传入倒残差结构时需要对其进行尺寸调整, 调整后的尺寸为输入图片 ($H \times W$) 经过模型 n 次下采样后的尺寸 ($H/n \times W/n$)。在训练和验证过程中, 由于 GPU 内存有限, 因此对每张图片的分辨率都调整为 224 像素 \times 224 像素, 为了确保工业缺陷检测的准确性, 在进行生成类激活图进行缺陷定位时, 传入网络的每张图片的分辨率都为 1 024 像素 \times 1 024 像素。

3.3 网络结果分析

本文模型随着网络的迭代次数增加, 相较于作为骨干网络的 Swin minuscule (由深度为 (2, 2, 2, 2) 的 STB 模块组成), 引入 IRB 层和 SEB 层的网络能够更快地学习到良好的底层特征, 从而得到更精细的高层特征, 加快网络的收敛, 并且 Swin minuscule 需要接近 2 倍的训练次数才能达到优化后本文模型的学习性能。本文模型的识别率曲线、训练集损失曲线和验证集曲线如图 6 所示。

为评估本文模型的有效性, 在相同的训练条件下与当前常见的算法 ResNet34^[15]、Swin tiny^[11]进行对比。由于本文模型的原始骨干网络 (Backbone) 是深度为 (2, 2, 2, 2) 的 STB 模块组成, 因此选择加入深度为 (2, 2, 2, 2) 的 Swin minuscule 作为比较。从表 2 中可以看出, 本文模型具备良好的定位性能和分类性能, 在 100.0% 精确率和 98.5% 召回率的情况下, 在纸板表面缺陷数据集上的分类正确率达到了 99.0%, 定位正确率达到了 92.2%。相比之下, 完全基于卷积的 ResNet34 虽然在分类上有不错的性能, 但由于卷积对像素之间的长距离不能充分学习, 因此定位性能一般。而完全基于自注意力的 Swin tiny 算法, 虽然召回率为 100.0%, 但网络收敛速度慢。从表 3 中可以看出, 即使将预测框与真实框 IOU 大于 0.9 的认定为识别成功, 本文模型也有着 57.7% 的定位正确率, 相较于原模型的定位性能提高了 8.8%。图 7 给出了本模型在纸板表面缺陷数据集的各类别检测示例, 其中颜色越深代表对识别类别的贡献度越高。

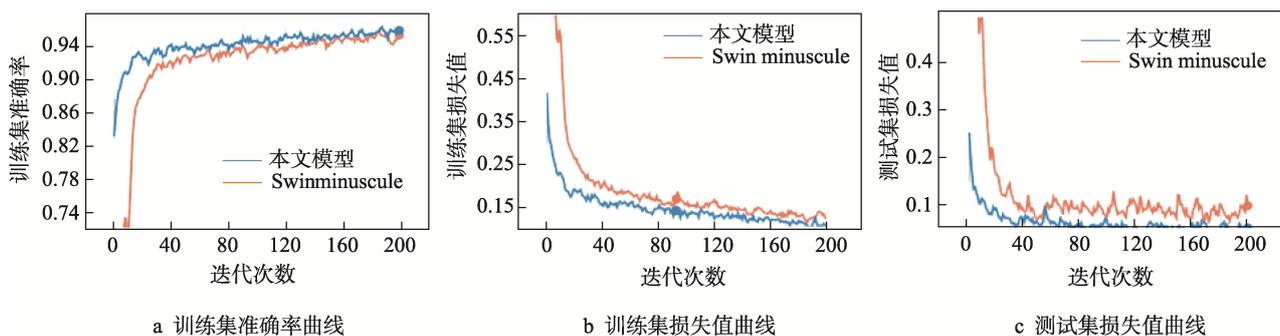


图 6 纸板缺陷图像分类准确率及损失曲线

Fig.6 Accuracy and loss curve of cardboard defect image classification

表2 各模型对比实验结果

Tab.2 Comparison of experimental results of each model

模型	分类 正确率/%	定位 正确率/%	精确率/%	召回率/%
ResNet34	94.5	75.6	91.3	99.5
Swin minuscule	97.7	84.7	97.0	99.0
Swin tiny	94.4	87.6	93.5	100.0
本文模型	99.0	92.2	100.0	98.5

表3 不同模型在 IOU 指标下的定位性能

Tab.3 Localization performance of the model under different IOU metrics

模型	IOU(0.5) 值/%	IOU(0.6) 值/%	IOU(0.8) 值/%	IOU(0.9) 值/%
Backbone	84.7	77.5	62.5	48.9
本文模型	92.2	89.6	75.2	57.7

3.4 消融实验结果分析

为验证 IRB 层和 SEB 层对网络整体性能上的贡献,

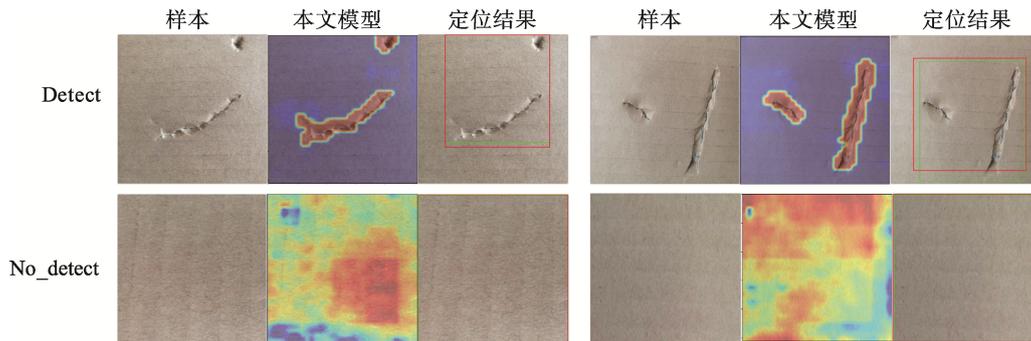


图7 纸板缺陷目标检测结果示例

Fig.7 Example of cardboard defect target detection results

注: 红框是预测框, 绿框是真实框。

本节对网络进行拆分实验, 分别验证 IRB 层和 SEB 层对网络的影响, 测试结果如表 4 所示。引入 IRB 层和 SEB 层后, 网络在分类和定位上的性能均有所提高。其中, 通过 IRB 层提高骨干网络的特征提取能力, 使网络定位性能提高了 1.6%, 通过 SEB 层对判别区域进行聚焦, 使网络定位性能提高了 3.2%, 结合 2 个模块后使网络定位性能提高了 7.5%。通过可视化纸板排骨痕缺陷进行分析, 结果如图 8 所示, IRB 层使网络能学习到更复杂和抽象的特征表示, 网络的注意力更加聚集在缺陷处, SEB 层则进一步提高了提取缺陷特征的能力。证明了使用 IRB 层和 SEB 层可

以提高分类准确率, 并且划分出有效的判别区域进行定位, 从而提高模型的性能。

表4 实验结果

Tab.4 Experimental result

模型	分类正确率/%	定位正确率/%
Backbone	97.7	84.7
Backbone +IRB	98.7	86.3
Backbone +SEB	99.0	87.9
Backbone +IRB+SEB	99.0	92.2

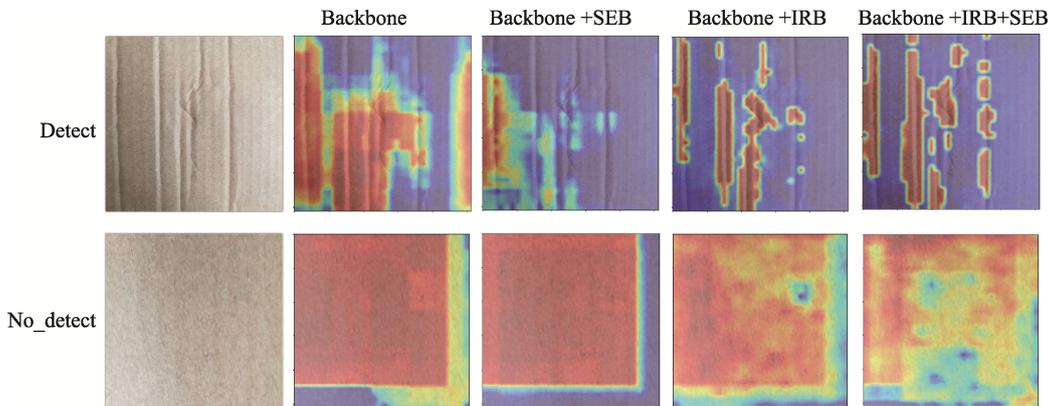


图8 分解实验的类激活图可视化结果

Fig.8 Visualization results of class activation diagram of decomposition experiment

4 结语

针对工业行业的实例级标签数量较少和过于主观的问题, 本文面向实际的纸板表面缺陷检测任务, 提出了弱监督学习下融合卷积和注意力机制的神经网络算法。该网络仅使用图像级标签实现对缺陷的分类和定位, 并取得优异的检测性能, 避免了专业性强的实例级标签。在纸板缺陷数据集上的实验结果证明, 本文提出的模型在分类正确率与定位正确率上分别达到 99.0% 和 92.2%, 胜过其他模型。本研究为基于自动化缺陷纸板剔除系统的构建奠定基础。

参考文献:

- [1] BOWEN Z, HUACAI L, XIUYUN Z. Cotton Packaging Defect Detection Based on Improved SVM Algorithm[C]// 2022 5th World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM), IEEE, 2022: 1088-1092.
- [2] SHA Y, HE Z, DU J, et al. Intelligent Detection Technology of Flip Chip Based on H-SVM Algorithm[J]. Engineering Failure Analysis, 2022, 134: 106032.
- [3] GUAN S Y, WANG X K, HUA L, et al. Quantitative Ultrasonic Testing for Near-Surface Defects of Large Ring Forgings Using Feature Extraction and GA-SVM[J]. Applied Acoustics, 2021, 173(1/2/3/4): 107714.
- [4] ZHAO M, QIU W Y, WEN T X, et al. Feature Extraction Based on Gabor Filter and Support Vector Machine Classifier in Defect Analysis of Thermoelectric Cooler Component[J]. Computers & Electrical Engineering, 2021, 92(12): 107188.
- [5] GAO C R. Design of Tourism Package with Paper and the Detection and Recognition of Surface Defects-Taking the Paper Package of Red Wine as an Example[J]. Journal of Intelligent Systems, 2021, 30(1): 720-727.
- [6] 姜旭辉. 基于机器视觉的纸板缺陷检测控制系统研究及应用[D]. 青岛: 青岛科技大学, 2021.
JIANG X H. Research and Application of Paperboard Defect Detection Control System Based on Machine Vision[D]. Qingdao: Qingdao University of Science & Technology, 2021.
- [7] 叶宇星, 孙志锋, 马风力, 等. 基于改进 YOLOv5s 的腌制蔬菜真空包装缺陷检测[J]. 包装工程, 2023, 44(9): 45-53.
- [8] YE Y X, SUN Z F, MA F L, et al. Vacuum Packaging Defect Detection of Pickled Vegetables Based on Improved YOLOv5s[J]. Packaging Engineering, 2023, 44(9): 45-53.
- [8] 何彧, 宋克臣, 张德富, 等. 融合多层次特征的弱监督钢板表面缺陷检测算法[J]. 东北大学学报(自然科学版), 2021, 42(5): 687-692.
- [8] HE Y, SONG K C, ZHANG D F, et al. Weakly-Supervised Steel Plate Surface Defect Detection Algorithm by Integrating Multiple Level Features[J]. Journal of Northeastern University (Natural Science), 2021, 42(5): 687-692.
- [9] 孙美君, 吕超章, 韩亚洪, 等. 弱监督学习下的融合注意力机制的表面缺陷检测[J]. 计算机辅助设计与图形学学报, 2021, 33(6): 920-928.
- [9] SUN M J, LYU C Z, HAN Y H, et al. Weakly Supervised Surface Defect Detection Based on Attention Mechanism[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(6): 920-928.
- [10] JIANG P T, ZHANG C B, HOU Q B, et al. LayerCAM: Exploring Hierarchical Class Activation Maps for Localization[J]. IEEE Transactions on Image Processing, 2021, 30: 5875-5888.
- [11] LIU Z, LIN Y, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10012-10022.
- [12] ZHOU B, KHOSLA A, LAPEDRIZA A, et al. Learning Deep Features for Discriminative Localization[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2921-2929.
- [13] YU J, JIANG Y, WANG Z, et al. Unitbox: An Advanced Object Detection Network[C]// Proceedings of the 24th ACM International Conference on Multimedia, 2016: 516-520.
- [14] LOSCHILOV I, HUTTER F. Decoupled Weight Decay Regularization[J]. arXiv preprint arXiv:1711.05101, 2017: 1-11.
- [15] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.