

基于三重注意力机制的带钢表面显著性缺陷检测

张莉^{a,b}, 李锡瑞^a, 郭华平^{a,b*}, 孙艳歌^{a,b}

(信阳师范大学 a. 计算机与信息技术学院; b. 河南省教育大数据分析与应用重点实验室, 河南 信阳 464000)

摘要:带钢表面缺陷普遍存在低对比度、类内差异显著及类间相似性高的特点,对模型检测精度构成巨大挑战。为解决上述问题,提出一种基于三重注意力机制的显著性目标检测模型,旨在实现带钢表面缺陷检测。针对主干网络提取的粗缺陷特征,构建包含三个分支的特征细化模块,各分支从不同维度关注带钢表面缺陷的通道和空间特征。同时,为细化显著目标的边缘细节,构建一种局部注意力特征过滤模块,并将其嵌入到特征细化模块中,提高模型对缺陷区域的敏感性。实验结果表明,与现有方法相比,本文所提模型在显著性目标检测任务中展现显著优势,其平均绝对误差、加权F测度等评价指标均优于同类方法,验证了所提方法的有效性和鲁棒性。

关键词: 缺陷检测; 注意力机制; 特征提取; 特征融合

中图分类号: TP391.4

文献标志码: A

开放科学(资源服务)标识码(OSID):



Salient object detection for strip steel surface defects based on triple attention mechanism

ZHANG Li^{a,b}, LI Xirui^a, GUO Huaping^{a,b*}, SUN Yange^{a,b}

(a. College of Computer and Information Technology; b. Henan Provincial Key Laboratory of Education Big Data Analysis and Application, Xinyang Normal University, Xinyang 464000, China)

Abstract: Surface defects on strip steel have characteristics such as low contrast, large intra-class differences, and inter-class similarity, which bring significant challenges to the accuracy of model detection. To address these challenges, a salient object detection model based on a triple-attention mechanism was proposed to achieve more accurate detection of surface defects on strip steel. A feature refinement module, consisting of three branches, was constructed for the coarse defect feature maps extracted from the model backbone network. Each branch was designed to focus on the channel and spatial features of surface defects on the strip steel from different perspectives. Additionally, to refine the edge details of salient targets, a local attention feature filter module was built and embedded into the feature refine module as a plugin, which enhances the model's sensitivity in defect areas. The experimental results indicated that the proposed model had significant advantages over existing methods in salient object detection and outperforms similar methods in evaluation indicators such as mean absolute error and weighted F-measure, thereby verifying the effectiveness and robustness of the proposed method.

Key words: defect detection; attention mechanism; feature extraction; feature fusion

0 引言

在带钢生产流程中,产品质量管控是工业制

造过程中的关键环节。作为重要的金属原材料,带钢广泛应用于建筑、交通、机械等领域,其质量优劣直接影响工程结构的稳定性和使用寿命。为

收稿日期:2024-07-03;修回日期:2024-12-06;*.通信联系人,E-mail: hpguo@xynu.edu.cn

基金项目:国家自然科学基金项目(62403405);河南省科技计划项目(242102210092, 252102220046);河南省科学技术研究联合基金项目(20240012);河南省高等学校重点科研项目(25B520004);教育部智能群系统工程技术研究中心开放基金项目(ZZU-CISS-2024004)

作者简介:张莉(1987—),女,河南信阳人,讲师,博士,硕士生导师,主要从事人工智能、图像处理等方向研究。

引用格式:张莉,李锡瑞,郭华平,等.基于三重注意力机制的带钢表面显著性缺陷检测[J].信阳师范大学学报(自然科学版),2025,38(4):414-420.

ZHANG Li, LI Xirui, GUO Huaping, et al. Salient object detection for strip steel surface defects based on triple attention mechanism[J]. Journal of Xinyang Normal University (Natural Science Edition), 2025, 38(4):414-420.

了满足市场需求,工业生产必须在提高带钢生产效率的同时,确保产品的质量。

受制造工艺、生产环境等多因素叠加影响,带钢表面缺陷呈现出显著的多样性特征,给自动缺陷检测带来三大核心挑战:低对比度、类内差异和类间相似性。在此背景下,确保带钢表面缺陷检测模型在复杂多变的生产环境中,对各类缺陷实现精准且高效识别至关重要。

构建高精度缺陷检测模型的核心在于有效提取缺陷特征,而注意力机制是获取兼具丰富语义信息与空间信息特征图的主流方法之一。现有的基于注意力机制的相关方法可划分为两类:一类是基于通道注意力机制,另一类基于空间注意力机制。这两类注意力机制能明确构建通道或空间维度上的依赖关系,通过学习注意力权重,引导模型主动聚焦于待检测的目标区域。例如,TANG等^[1]提出一种轻量级印刷电路板(PCB)缺陷检测模型,该模型引入双域注意力机制,可有效提取微小尺寸PCB缺陷特征,显著提升缺陷识别的准确性;LEI等^[2]设计了一种新型多通道时空图卷积网络用于交通预测,该网络借助注意力机制为不同通道分配差异化权重,使模型能够精准捕捉并利用交通流的核心特征。VIEIRA等^[3]提出基于注意力的新型差分网络(AttentDifferNet)用于工业检测任务,在三个异常检测数据集的全对象测试中,检测性能均优于标准差分网络。然而,当前这些方法通过计算通道注意力或者空间注意力权重获取丰富的特征表示,但在提取缺陷特征的过程中,对于通道维度和空间维度之间的交互关联挖

掘仍存在明显的局限性。

直观角度出发,在提取带钢表面缺陷图像特征时,若能通过注意力权重计算获取丰富的特征表示,同时着重强调通道与空间维度间的跨维度交互信息捕获,将为解决自动缺陷检测任务中面临的低对比度、类间相似性等问题提供有效路径。为此,提出了一种基于三重注意力的带钢表面缺陷显著性目标检测模型。首先,以ResNet作为主干网络,对带钢图像进行多尺度特征提取;其次,针对每个尺度下的特征图,构建三重注意力模块,以便从三个不同角度(即通道-高度、通道-宽度和宽度-高度)捕获特征图在通道与空间维度上的交互信息;同时,构建局部注意力特征过滤模块,并将其嵌入到特征细化模块中,提高模型在缺陷区域的敏感性;最后,融合细化的多尺度特征图,生成具有更强表达能力的融合特征,提升模型对带钢表面缺陷的检测精度。

1 显著性目标检测模型架构

提出一种基于三重注意力的带钢表面缺陷显著性目标检测模型,总体架构如图1(a)所示。该模型由特征提取模块、特征细化模块和特征聚合模块三部分构成。首先,在特征提取模块中,采用ResNet作为主干网络,对带钢缺陷图像进行多尺度特征提取;然后,通过特征细化模块获取缺陷特征的全局信息,并对粗尺度特征进行细化;最后,特征聚合模块对经过细化的多尺度特征进行融合,以获得丰富的语义信息,从而解决自动缺陷检测任务中低对比度、类间相似性等问题。

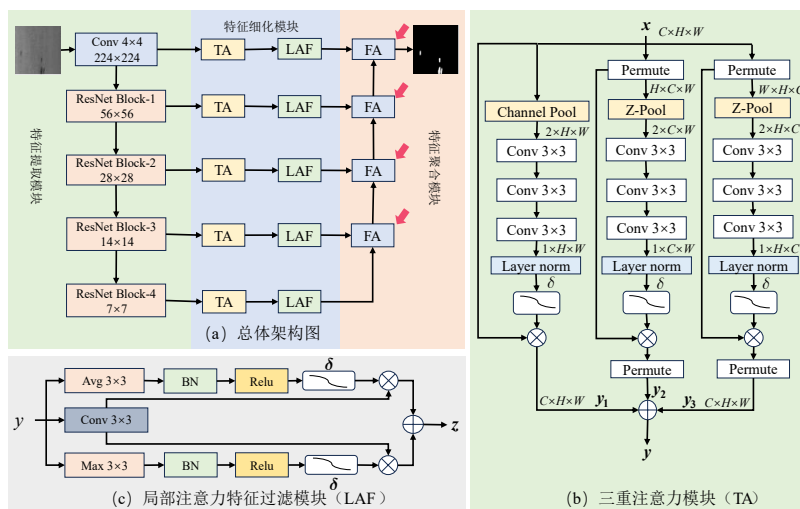


图1 总体结构图

Fig. 1 Overall structure diagram

1.1 特征提取模块

选取 ResNet-50 作为骨干特征提取器,用于提取带钢表面缺陷图像的粗尺度特征。该模型的总体架构如图 1(a)所示,其中特征提取模块由 1 个卷积层和 4 个残差学习块(分别记为 ResNet Block-1、ResNet Block-2、ResNet Block-3 和 ResNet Block-4)组成,其输出特征图统一表示为 $R_i, i \in \{0, 1, 2, 3, 4\}$,其中 R_0 表示卷积输出。与原始的 ResNet-50 不同,采用 Swin Transformer 的相似思路,对 4 个残差学习块的层数进行调整,最终确定各块的层数配置为(3, 3, 27, 3)。此外,为进一步优化特征提取效果,还对原始网络的关键组件做了两处改进:一是将原始尺寸为 7×7 、步长为 2 的卷积层,替换为尺寸 4×4 、步长为 2 的卷积层;二是采用 Layer Norm 替代批处理归一化操作,对卷积层的输出特征进行规范化处理。上述调整能够更有效地捕获缺陷图像的细节信息。

1.2 特征细化模块

在特征提取阶段,采用 ResNet 提取缺陷图像的多尺度特征。然而,受限于卷积操作的固有感受野,ResNet 在全局语义信息的捕捉与提取上存在明显不足。为解决该问题,设计针对特征的特征细化模块,具体包括三重注意力模块与局部注意力特征过滤模块(如图 1(a)所示)。其中,三重注意力模块用于捕获特征的全局语义信息,局部注意力特征过滤模块用于优化缺陷区域的细节信息。

1.2.1 三重注意力模块

三重注意力模块借助相关性计算分配注意力权重,能够有效地捕获与缺陷相关的关键特征,同时有效抑制无关背景特征的干扰。此外,该模块可实现特征在通道-高度、通道-宽度和高度-宽度三个不同维度视角下的有机交互。

图 1(b)呈现了三重注意力模块的结构,其具体实现如下:该模块由 3 个并行分支构成,旨在捕获输入张量 x 在不同维度上的交互信息。对于尺寸为 $C \times H \times W$ 的输入张量 x ,将其传入 3 个分支,每个分支分别负责在特定维度间建立关联。

在第一个分支中,构建高度维度 H 和宽度维度 W 的交互关系。具体而言,先对通道维度 C 进行池化操作,将通道数降至 2,得到张量 x_1 (尺寸为 $2 \times H \times W$),如式(1)所示。随后,对张量 x_1 依次进行 3 次卷积核大小为 3×3 的卷积运算,再经过 Layer Norm 层归一化处理,最终通过 sigmoid 激活

函数生成注意力权重,并与原始输入张量 x 加权相乘。具体计算如式(2)所示。

$$x_1 = \text{channelpool}(x), \quad (1)$$

$\hat{x}_1 = x_1 \otimes \delta(\text{LN}(\text{Conv}_{3 \times 3}(\text{Conv}_{3 \times 3}(\text{Conv}_{3 \times 3}(x_1))))),$ (2)
式中:channelpool 表示池化操作将通道维度降至 2, LN 表示层归一化操作, $\text{Conv}_{k \times k}$ 表示卷积核大小为 $k \times k$ 的卷积操作。

第二个分支聚焦于通道维度 C 和宽度维度 W 间的关联。首先,将输入张量 x 绕 W 轴逆时针旋转 90° ,得到张量 x_2 (尺寸为 $H \times C \times W$)。随后,通过 Z -pool 池化操作,将高度维度 H 缩减至 2,具体过程如式(3)所示。接着,对降维后的张量依次执行 3 次卷积核大小为 3×3 的卷积运算、Layer Norm 层归一化,再经过 Sigmoid 激活函数生成注意力权重。最后,将该权重与 x_2 进行加权相乘(计算过程如式(4)所示),并沿 W 轴顺时针旋转 90° 得到 y_2 ,以保持与原始输入 x 相同的尺寸。

$$x_2 = \text{Per}(Z\text{-pool}(x)), \quad (3)$$

$$\hat{x}_2 = x_2 \otimes \delta(\text{LN}(\text{Conv}_{3 \times 3}(\text{Conv}_{3 \times 3}(\text{Conv}_{3 \times 3}(x_2))))), \quad (4)$$

式中:Per 表示置换操作, Z -pool 表示将高度维度 H 降至 2 的池化操作, δ 表示 Sigmoid 激活函数操作, LN 表示归一化操作。

第三个分支以构建通道维度 C 和高度维度 H 的关联为目标,其逻辑与第二个分支类似。在完成三个分支的独立计算后,可得到对应不同维度关联的三类特征信息,将这三类信息进行元素相加,即可生成三重注意力模块的最终输出,具体计算过程如式(5)所示。该设计助力模型更高效地捕捉特征的全局语义信息,实现通道和空间之间的跨维度交互,提高对目标区域的关注能力。

$$y = \frac{1}{3}(y_1 \oplus y_2 \oplus y_3). \quad (5)$$

1.2.2 局部注意力特征过滤模块

在三重注意力模块中,建立通道-高度、通道-宽度和高度-宽度 3 个不同维度间的关联,同时,引入新的局部注意力特征过滤模块,以进一步细化缺陷的边缘细节信息。

图 1(c)为局部注意力特征过滤模块。首先,对特征 y 进行局部平均池化和局部最大池化,以捕获缺陷特征的局部信息。其次,通过归一化、ReLU 激活和 Sigmoid 操作,生成两个注意力权重 (W_{avg} 和 W_{max})。具体生成过程如式(6)和式(7)所示。接着,对特征 y 进行 3×3 卷积操作,并分别与

W_{avg} 和 W_{max} 相乘,得到的结果相加得到最终的输出 z ,具体计算如式(8)所示。

$$W_{\text{avg}} = \delta(\text{Relu}(\text{BN}(\text{Avg}_{3 \times 3}(y))))), \quad (6)$$

$$W_{\text{max}} = \delta(\text{Relu}(\text{BN}(\text{Max}_{3 \times 3}(y))))), \quad (7)$$

$$z = (W_{\text{avg}} \otimes \text{Conv}_{3 \times 3}(y)) \oplus (W_{\text{max}} \otimes \text{Conv}_{3 \times 3}(y)), \quad (8)$$

式中:BN表示批归一化操作。

1.3 特征聚合模块

尽管通过引入特征提取模块和特征细化模块,模型对目标区域和背景区域的区分能力已得到显著提升,但受同类缺陷尺度差异较大的影响,模型仍面临域内差异性带来的挑战。为此,设计了特征聚合模块,通过融合不同尺度的特征,充分利用低层特征所蕴含的丰富空间信息与高层特征具备的强大语义抽象信息,实现两类信息的高效交互,以解决缺陷尺度不一的问题。图2为特征聚合模块,具体实现如下:

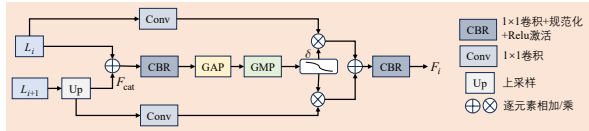


图2 特征聚合模块

Fig. 2 Feature aggregation module

首先,输入低级特征 L_i 和高级特征 L_{i+1} 。为使高层特征分辨率与低层特征保持一致,需要对高层特征执行2倍上采样操作,随后将其与低层特征进行融合,得到特征 F_{cat} 。具体计算过程如式(9)所示。

$$F_{\text{cat}} = L_i \oplus \text{Up}(L_{i+1}), \quad (9)$$

式中:Up表示上采样操作。

然后,对 F_{cat} 进行 1×1 卷积操作、归一化、ReLU激活、全局平均池化和全局最大池化,再通过激活函数 Sigmoid 生成注意力权重 W_i ,如式(10)所示。

$$W_i = \delta(\text{GMP}(\text{GAP}(\text{CBR}(F_{\text{cat}}))))), \quad (10)$$

式中:GAP和GMP表示平均、最大全局池化,CBR表示 1×1 卷积、规范化和ReLU激活操作。

接着,对低层特征和高层特征进行卷积操作,将生成的注意力权重 W_i 分别与这两个特征加权相乘后求和,之后再行 1×1 卷积操作、规范化和ReLU激活函数,得到最终的融合特征 F_i ,如式(11)所示。

$$F_i = \text{CBR}((W_i \otimes \text{Conv}(L_i)) \oplus$$

$$(W_i \otimes \text{Conv}(\text{Up}(L_{i+1}))))). \quad (11)$$

1.4 损失函数

为生成高质量的显著性图,使用融合损失与深度监督相结合的方法来训练模型,包括二值交叉熵(Binary Cross Entropy, BCE^[4])损失和边界交并比(Intersection Over Union, IOU^[5])损失。总的损失函数如式(12)所示。

$$L_{\text{all}} = \sum_{i=1}^4 L_{\text{bce}}(S_i, G) + L_{\text{iou}}(S_i, G), \quad (12)$$

式中: S_i 是由预测的 F_i 获得的显著性图,该图对应所提方法的第 i 层。

二值交叉熵损失常用于二分类任务,记作 L_{bce} ,如式(13)所示。

$$L_{\text{bce}} = - \sum_{(x,y)} [G(x,y) \ln(S(x,y)) + (1-G(x,y)) (\ln(1-S(x,y)))] \quad (13)$$

式中: $G(x,y)$ 表示图像中像素点 (x,y) 的真实标签,其中0表示背景,1表示缺陷对象; $S(x,y)$ 表示在相应像素点的预测概率。

边界交并比损失用来惩罚分类的不准确,记作 L_{iou} ,如式(14)所示。

$$L_{\text{iou}} = 1 - \frac{\sum_{(x,y)} S(x,y)G(x,y)}{\sum_{(x,y)} S(x,y) + G(x,y) - S(x,y)G(x,y)}. \quad (14)$$

2 实验

2.1 实验设置

为证明所提出模型在带钢表面缺陷检测中的有效性,在公共数据集SD-saliency-900^[6]上进行相关实验。该数据集包含900张经过裁剪处理的图像,涵盖夹杂物、斑块和划痕3类缺陷类型,如图3所示。此外,SD-saliency-900数据集为每张图像提供了像素级标签注释,且所有图像均统一采用 200×200 像素的分辨率规格。

1) 实验细节

所提方法的实验环境配置如下:基于PyTorch深度学习框架搭建,硬件采用Intel(R)Core(TM)i7-8750H CPU(主频2.20 GHz/2.21 GHz)、TITAN Xp GPU(显存12 G),操作系统为Ubuntu。训练集构建细节为:总计包含810张图像,其中540张源自SD-saliency-900数据集(每种缺陷类型各180张),剩余270张为受到椒盐噪声干扰的图像(噪声密度 $\rho = 20\%$,每种缺陷类型各90张)。训练阶段的数据预处理流程为:将每张图

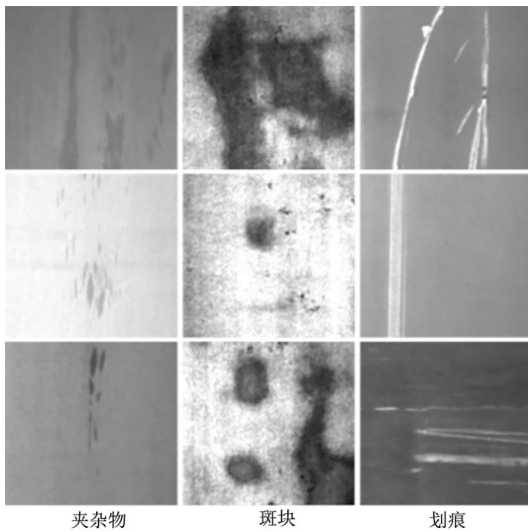


图 3 在 SD-saliency-900 上的 3 种缺陷

Fig. 3 Three defects on SD-saliency-900

像 I 的尺寸调整至 256×256 , 并执行 Zscore 归一化操作(计算公式为 $(I - \mu) / \delta$, 其中 $\mu = 0.4669$, $\delta = 0.2437$). 测试阶段则先将图像尺寸调整为 256×256 , 并将输出的预测显著性图调整为 200×200 进行评估。模型参数初始化与训练设置如下: 特征提取模块参数采用在 ImageNet^[7] 上预训练的 ResNet-50 参数, 其余模块的参数则使用 PyTorch 默认的配置; 优化器选择 Adam^[8], 初始学习率设为 $lr = 0.001$, 动量参数 $\alpha = 0.9$, 训练批大小 $batch_size = 8$, 其余参数设置为默认值。

2) 评价指标

为评估本文提出方法的性能, 并与其他主流检测方法展开对比, 实验采用 3 个学术界广泛认可的评价指标, 包括 F-measure^[9]、平均绝对误差 (Mean Absolute Error, MAE)^[10] 和结构度量 (Structure Measure, SM)^[11]。其中, F-measure 作为综合考虑显著性检测结果精确性 (Precision) 与召回率 (Recall) 的评价指标, 其计算方式如式 (15) 所示。

$$F_{\beta} = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}}, \quad (15)$$

式中: β 是调节 Precision 与 Recall 相对重要的参数。选择不同的 β 值, 可灵活调整二者在 F-measure 综合评估中的权重占比。

平均绝对误差是一种常用的评价指标, 用于衡量预测结果与真实值的绝对误差程度, 记作 MAE。计算过程如式 (16) 所示。

$$\text{MAE} = \frac{1}{n} \times \sum_i |S_i - G|, \quad (16)$$

式中: n 为样本数量, S_i 为预测值, G 为真实值。

结构评价指标用于衡量图像分割结果的结构相似性, 即分割结果与真实分割之间的结构一致性程度, 记作 SM。计算过程如式 (17) 所示。

$$\text{SM} = \alpha \times F_{\text{obj}} + \beta \times F_{\text{reg}}, \quad (17)$$

式中: F_{obj} 表示分割结果中对象的结构相似性; F_{reg} 表示分割结果中区域的结构相似性; α 、 β 表示权重系数, 用于平衡不同结构对总体结构相似性的贡献。

2.2 消融实验

为验证所提方法的有效性, 在 SD-saliency-900 数据集上开展全面实验。实验采用 MAE、SM 及 F_{β} 值作为评价指标, 对相关架构的定量评估结果如表 1 所示。

表 1 消融实验

Tab. 1 Ablation experiments

方法	MAE ↓	F_{β} ↑	SM ↑
ResNet Block	0.014 5	0.925 2	0.911 8
ResNet Block+TA	0.013 8	0.927 2	0.929 8
ResNet Block+LAF	0.013 5	0.927 0	0.931 2
ResNet Block+TA+LAF	0.012 8	0.930 2	0.945 2
ResNet Block+TA+LAF+FA	0.009 4	0.937 8	0.950 2

注 ↓ 表示值越小, 性能越好; ↑ 表示值越大, 性能越好。

为了验证模型中各核心组件的必要性, 设计并执行了一系列消融实验, 实验结果如表 1 所示。实验仅以 ResNet 作为基准模型, 在此基础上衍生出 4 种不同组件配置的变体模型。从表 1 可清晰观察到: 当模型集成所有设计组件 (即 ResNet Block、TA、LAF 和 FA) 时, 性能达到最佳, 这直接证明了所提模型中关键组件 TA、LAF 及 FA 的有效性。值得关注的是, 单独引入 TA 或 LAF 时, 模型性能的提升幅度并不显著, 甚至低于 TA 与 LAF 协同集成时的性能表现, 这也进一步表明, TA 与 LAF 属于互补模块, 需要同时考虑以实现最佳效果。

2.3 与其他方法比较

为进一步验证本文所提方法的有效性与鲁棒性, 在 SD-saliency-900 数据集上开展了系统且全面的实验。将所提方法与 11 种当前主流的显著性检测方法进行对比, 包括 RCRR^[12]、2LSG^[13]、

BC^[14]、PFANet^[15]、NLDF^[16]、R3Net^[17]、EDRNet^[22]。定量评估结果如表2所示,可视化对比结果如图4呈现。

表2 与11种方法在MAE、 F_β 、SM上比较的结果
Tab. 2 Results compared with 11 methods on MAE, F_β , SM

检测	$\rho=0$			$\rho=10\%$			$\rho=20\%$		
	MAE↓	F_β ↑	SM↑	MAE↓	F_β ↑	SM↑	MAE↓	F_β ↑	SM↑
RCRR ^[12]	0.243 9	0.259 1	0.534 2	0.255 2	0.255 7	0.530 2	0.284 2	0.243 1	0.514 7
2LSG ^[13]	0.247 4	0.312 0	0.551 8	0.258 7	0.300 7	0.536 8	0.261 9	0.294 6	0.534 1
BC ^[14]	0.155 4	0.385 5	0.594 2	0.151 9	0.373 3	0.588 1	0.175 3	0.332 7	0.562 3
PFANet ^[15]	0.084 1	0.535 8	0.741 1	0.101 1	0.462 3	0.696 1	0.107 9	0.428 1	0.678 2
NLDF ^[16]	0.047 4	0.713 4	0.802 8	0.117 5	0.419 6	0.633 9	0.125 3	0.363 7	0.589 2
R3Net ^[17]	0.025 4	0.816 0	0.839 7	0.038 2	0.696 1	0.766 0	0.043 0	0.617 0	0.725 8
PiCANet ^[18]	0.025 9	0.828 9	0.896 3	0.035 1	0.752 1	0.849 0	0.040 4	0.707 3	0.822 8
DACNet ^[19]	0.011 8	0.925 0	0.942 8	0.012 5	0.922 3	0.937 8	0.014 4	0.911 1	0.924 6
BASNet ^[20]	0.015 2	0.909 2	0.927 6	0.016 0	0.903 3	0.923 5	0.016 0	0.901 4	0.921 9
CPD ^[21]	0.021 1	0.819 2	0.903 3	0.035 3	0.723 5	0.830 8	0.032 4	0.732 5	0.838 4
EDRNet ^[22]	0.013 0	0.922 5	0.937 4	0.013 9	0.912 5	0.929 9	0.014 6	0.905 6	0.924 4
Ours	0.009 4	0.937 8	0.950 2	0.009 4	0.937 4	0.948 7	0.015 2	0.914 2	0.928 4

注 ↓表示值越小,性能越好;↑表示值越大,性能越好。

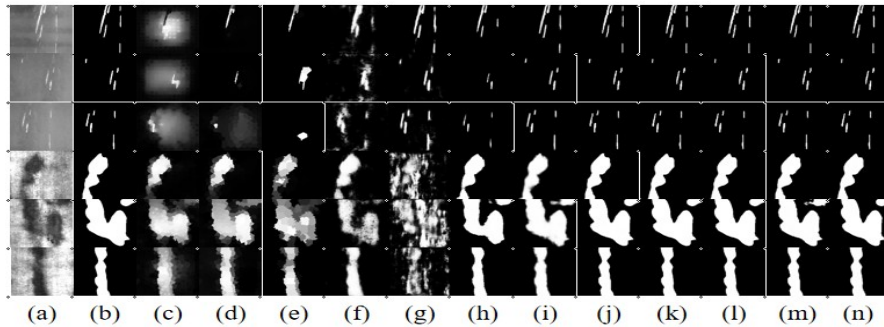


图4 显著性图的视觉比较结果

Fig. 4 Visual comparison results of salient maps

(a)输入图像,(b)真实标签,(c)RCRR,(d)2LSG,(e)BC,(f)PFANet,(g)NLDF,
(h)R3Net,(i)PiCANet,(j)DACNet,(k)BASNet,(l)CPD,(m)EDRNet,(n)Ours

图4展示了所提方法与其他先进方法的视觉对比结果。显然,本文的模型在图4的最后一列呈现出最优的缺陷检测结果,能够完整、准确地分割出缺陷区域。具体来看,前3行图像中缺陷区域与背景对比度较低,部分模型存在错误突出背景区域或缺陷分割不完整的问题;相比之下,本文方法在不引入任何背景干扰的情况下,能够准确完整地检测缺陷区域。此外,后3行图像属于背景复杂且具有挑战性的场景,多数现有模型生成的显著性图质量欠佳,存在边缘细节模糊或目标区域误突出等问题。而本文提出的模型仍能够有效地检测出缺陷区域。综上所述,所提方法在不同场景

下均能实现缺陷区域的完整分割,进一步验证了其在缺陷检测任务中的有效性。

表2呈现了不同模型的定量评估结果。与其他模型相比,所提出的模型在MAE、SM和 F_β 性能方面获得最优结果。具体而言,与DACNet相比,本文方法的MAE降低了0.24个百分点,SM与 F_β 分别提升了0.74和1.28个百分点。相较于带钢显著性目标检测模型EDRNet,本文模型的SM和 F_β 分别提高了1.28和1.53个百分点,MAE同步降低了0.36个百分点。这表明,所提方法相较现有先进模型实现了显著性改进。此外,即便在训练集引入10%与20%椒盐噪声的情况下,仍优于所有

对比模型,充分证明其在钢材表面缺陷检测任务中具备更优的性能与鲁棒性。

3 结论

提出了一种基于三重注意力机制的缺陷检测模型,用于钢材表面缺陷检测任务。首先,采用 ResNet-50 提取缺陷图像的多尺度特征;接着,通过设计的新三重注意力模块和局部注意力特征过

滤模块,有效挖掘多尺度特征全局和局部信息之间的关联,从而实现通道-高度、通道-宽度和高度-宽度的跨维交互,增强了缺陷区域和背景区域的区分能力。同时,使用特征聚合模块逐步融合细化的多尺度特征,实现了高层和低层特征之间的互补,有效解决了域内差异性的问题。在 SD-saliency-900 数据集上进行大量的实验,结果表明所提模型在解决缺陷检测问题上的有效性。

参考文献:

- [1] TANG Jun, WANG Zidong, ZHANG Hongyi, et al. A lightweight surface defect detection framework combined with dual-domain attention mechanism[J]. *Expert Systems with Applications*, 2024, 238(Part A): 121726.
- [2] LEI Tianyang, YANG Kewei, LI Jichao, et al. Multichannel spatial-temporal graph convolution network based on spectrum decomposition for traffic prediction[J]. *Expert Systems with Applications*, 2024, 238(Part F): 122281.
- [3] VIEIRA E SILVA A L, SIMOES F, KOWERKO D, et al. Attention modules improve image-level anomaly detection for industrial inspection: A DifferNet case study[C]//2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, 2024: 8231-8240.
- [4] DE BOER P T, KROESE D P, MANNOR S, et al. A tutorial on the cross-entropy method [J]. *Annals of Operations Research*, 2005, 134(1): 19-67.
- [5] RAHMAN M A, WANG Yang. Optimizing intersection-over-union in deep neural networks for image segmentation [C]//*Advances in Visual Computing*, Cham, 2016: 234-244.
- [6] SONG Guorong, SONG Kechen, YAN Yunhui. Saliency detection for strip steel surface defects using multiple constraints and improved texture features[J]. *Optics and Lasers in Engineering*, 2020, 128: 106000.
- [7] DENG Jia, DONG Wei, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, 2009: 248-255.
- [8] KINGMA D, BA J. Adam: A method for stochastic optimization[J]. *Computer Science*, 2014, 6:980-991.
- [9] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency tuned salient region detection [C]//2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, 2009: 1597-1604.
- [10] PERAZZI F, KRÄHENBÜHL P, PRITICH Y, et al. Saliency filters: Contrast based filtering for salient region detection[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, 2012: 733-740.
- [11] CHENG Mingming, FAN Dengping. Structure-measure: A new way to evaluate foreground maps[J]. *International Journal of Computer Vision*, 2021, 129(9): 2622-2638.
- [12] YUAN Yuchen, LI Changyang, KIM J, et al. Reversion correction and regularized random walk ranking for saliency detection[J]. *IEEE Transactions on Image Processing*, 2018, 27(3): 1311-1322.
- [13] ZHOU Li, YANG Zhaohui, ZHOU Zongtan, et al. Salient region detection using diffusion process on a two layer sparse graph[J]. *IEEE Transactions on Image Processing*, 2017, 26(12): 5882-5894.
- [14] ZHU Wangjiang, LIANG Shuang, WEI Yichen, et al. Saliency optimization from robust background detection[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 2014: 2814-2821.
- [15] ZHAO Ting, WU Xiangqian. Pyramid feature attention network for saliency detection [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, 2019: 3080-3089.
- [16] LUO Zhiming, MISHRA A, ACHKAR A, et al. Non-local deep features for salient object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, 2017: 6593-6601.
- [17] DENG Zijun, HU Xiaowei, ZHU Lei, et al. R3net: Recurrent residual refinement network for saliency detection[C]//*Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, Stockholm, 2018: 684-690.
- [18] LIU Nian, HAN Junwei, YANG M H. PiCANet: Learning pixel wise contextual attention for saliency detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018: 3089-3098.
- [19] ZHOU Xiaofei, FANG Hao, LIU Zhi, et al. Dense attention-guided cascaded network for salient object detection of strip steel surface defects[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-14.
- [20] QIN Xuebin, ZHANG Zichen, HUANG Chenyang, et al. BASNet: Boundary-aware salient object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, 2019: 7471-7481.
- [21] WU Zhe, SU Li, HUANG Qingming. Cascaded partial decoder for fast and accurate salient object detection[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, 2019: 3902-3911.
- [22] SONG Guorong, SONG Kechen, YAN Yunhui. EDRNet: Encoder-decoder residual network for salient object detection of strip steel surface defects[J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(12): 9709-9719.

责任编辑:郭红建 陈松楠