

本文引用格式：王伊杰,蔡建扑,任志刚.基于高分辨率特征图的工业产品表面缺陷检测算法[J].自动化与信息工程,2025,46(2):1-8.

WANG Yijie, CAI Jianpu, REN Zhigang. Surface defect detection algorithm for industrial products based on high-resolution feature maps[J]. Automation & Information Engineering, 2025,46(2):1-8.

基于高分辨率特征图的工业产品表面缺陷检测算法*

王伊杰¹ 蔡建扑¹ 任志刚^{1,2}

(1.广东工业大学自动化学院, 广东 广州 510006

2.粤港澳离散制造智能化联合实验室, 广东 广州 510006)

摘要: 在工业制造领域,对加工后的产品进行表面缺陷检测能够有效提高产品出厂质量。针对工业产品表面缺陷通常具有对比度低、形状不规则、尺寸小且细长等特征,并伴随有明显噪声,导致检测任务充满挑战的问题,提出一种基于高分辨率特征图的工业产品表面缺陷检测算法。首先,提出一种高分辨率特征图实时网络(RHNet)模型,通过将每个阶段的输入输出特征图保持为原始图像分辨率的1/4,有效保留了更多的细节信息;然后,提出短期双分支模块(SDBM),实时处理高分辨率特征图;最后,设计一种快速并行聚合金字塔池化模块(FPAPPM),快速提取深层信息并进行多尺度上下文融合。实验结果表明,RHNet模型在表面缺陷建模能力和检测性能方面均表现较好,能够满足工业场景实时性与应用部署的要求。

关键词: 表面缺陷检测; 高分辨率特征图实时网络; 多尺度融合

中图分类号: TP391.41

文献标志码: A

文章编号: 1674-2605(2025)02-0001-08

DOI: 10.12475/aie.20250201

开放获取

Surface Defect Detection Algorithm for Industrial Products Based on High-resolution Feature Maps

WANG Yijie¹ CAI Jianpu¹ REN Zhigang^{1,2}

(1.School of Automation Guangdong University of Technology, Guangzhou 510006, China

2.Guangdong-HongKong-Macao Joint Laboratory for Smart Discrete Manufacturing Guangdong University of Technology, Guangzhou 510006, China)

Abstract: In the field of industrial manufacturing, surface defect detection of processed products can effectively improve the quality of finished products. Industrial product surface defects often exhibit characteristics such as low contrast, irregular shapes, small and slender sizes, and significant noise, making the detection task highly challenging. To address this issue, a surface defect detection algorithm for industrial products based on high-resolution feature maps is proposed. First, a real-time high-resolution network (RHNet) is introduced, which maintains the input and output feature maps of each stage at 1/4 of the original image resolution, effectively preserving more detailed information. Then, a short-term dual-branch module (SDBM) is proposed to process high-resolution feature maps in real time. Finally, a fast parallel aggregation pyramid pooling module (FPAPPM) is designed to rapidly extract deep level information and perform multi-scale context fusion. Experimental results demonstrate that the RHNet model performs well in both surface defect modeling capability and detection performance, meeting the real-time requirements and deployment needs of industrial scenarios.

Keywords: surface defect detection; real-time network based on high-resolution feature maps; multi-scale fusion

0 引言

工业产品表面缺陷检测能够有效提高产品质量并减少事故发生,在现代制造和生产领域具有重要的作用^[1]。然而,人工检测工业产品表面缺陷存在主观性强和结果一致性不足等问题;基于传统机器视觉的缺陷检测方法依赖缺陷特征提取算法,缺乏通用性和可扩展性。近年来,深度学习方法在表面缺陷检测精度和推理速度方面取得了显著进展,同时边缘端硬件性能的提升,使其部署在工业环境成为可能。

目前,学者们已提出多种针对表面缺陷检测的语义分割模型。文献[2]提出一种基于改进 U-Net 模型的手机玻璃屏幕缺陷语义分割算法,采用带有移动端倒残差瓶颈卷积模块的高效编码器-解码器结构,即高效 U-Net,在减少算法参数量的同时保持了较高的检测精度。文献[3]对快速语义分割卷积神经网络进行改进,并提出一种快速表面缺陷分割网络,通过引入双分支架构和全局上下采样模块,实现对多尺度远程上下文信息的有效捕获,解决了缺陷的边界细节不完整、局部相似及与背景对比度低等问题。文献[4]提出一种用于表面缺陷检测的多层特征融合网络,采用单解码器和多编码器结构,利用预训练的 ResNet34 模型捕捉语义信息,并引入全局通道注意力模块进行多层次的特征融合,以更好地恢复缺陷细节。文献[5]提出一种基于深度特征融合的分割网络,利用 ResNet50 提取低、中、高级特征,并通过多级特征融合模块将不同层次的特征图进行融合,以提取更多的上下文信息,从而提升缺陷检测精度。文献[6]通过引入多尺度注意力特征融合模块,有效融合了不同层次的特征信息,以更好地获取图像的缺陷特征;采用归一化均方频率类别权重策略,抑制小缺陷类别的过度权重和假阳性问题,提升了金属表面缺陷检测的准确性。然而,上述模型存在一些共性问题:1) 模型强调利用上下文信息,而忽略了空间信息,导致检测结果的边界轮廓和细节不够完整;2) 模型在检测精度和推理速度方面难以平衡,限制了模型的通用性和应用范围。在工业制造过程中,推理速度仅需与产品制造的速度匹配

即可,但检测精度需要尽可能地高。

现有的语义分割模型大多通过下采样来减小特征图的尺寸,以有效降低计算复杂度。但这种做法会导致空间信息丢失,尤其是缺陷边界附近的区域,致使模型检测的缺陷轮廓不完整,甚至可能漏检一些细小缺陷。然而,高分辨率特征图能够通过精确捕获缺陷轮廓来保留关键空间信息,从而有效避免模型在表面缺陷检测过程中遗漏重要细节,这种特性使其在细小缺陷识别方面展现出显著优势。

本文提出一种基于高分辨率特征图的工业产品表面缺陷检测算法。首先,提出一种高分辨率特征图实时网络(real-time network based on high-resolution feature maps, RHNet)模型,通过保持高分辨率特征图,提高了模型对细小缺陷和复杂边界轮廓的检测精度;然后,提出短期双分支模块(short-term dual-branch module, SDBM),通过上下文分支和空间分支的协同作业,在保证高分辨率特征图的同时有效提升了推理速度;最后,设计一种快速并行聚合金字塔池化模块(fast parallel aggregation pyramid pooling module, FPAPPM),加速了多尺度特征的提取与融合,确保了检测精度与推理速度的平衡。

1 研究方法

1.1 RHNet 模型

RHNet 模型由 CB(R)模块、Basic Block、SDBM、主分割头和辅助监督机制(辅助分割头)组成,结构如图 1 所示。

CB(R)模块由卷积层 Conv、批归一化层 BN 和 ReLU 激活函数组成,用于对输入特征进行基本变换和初步特征提取。其中,(R)表示在某些情况下,未使用 ReLU 激活函数。

Basic Block 是 ResNet^[7]的基础模块,由 2 个 CB(R)模块和 1 个残差连接组成,能够在较低计算量的前提下,提升模型的特征表达能力。

主分割头和辅助分割头采用相同的网络结构,均由 BRC(批归一化层 BN、ReLU 激活函数和卷积层 Conv 的序列)、RC(ReLU 激活函数和卷积层 Conv

的序列) 模块组成, 最后通过 Argmax 函数按通道维度选取最大值对应的类别索引生成分割结果。

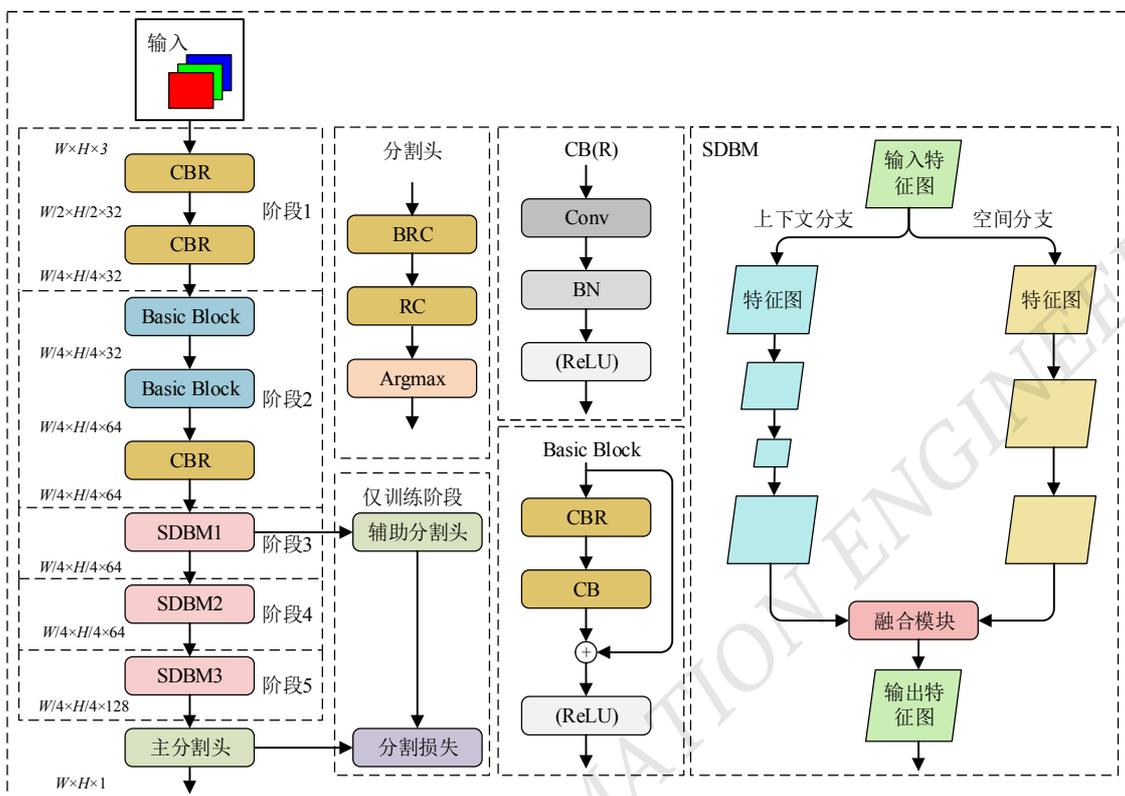


图 1 RHNNet 模型结构图

RHNNet 模型分为五个阶段:

第一阶段由 2 个步距为 2 的 CBR 模块组成, 负责将输入图像快速下采样至原始分辨率的 1/4, 以有效减少计算量并获得初步的特征图;

第二阶段通过 2 个 Basic Block 和 1 个 CBR 模块调整特征图的通道数至 64, 以进一步提升特征的表达能, 增加特征的维度和多样性;

第三至第五阶段由 3 个结构不同的 SDBM (SDBM1、SDBM2 和 SDBM3) 组成, 用于进一步挖掘深层特征, 并在保留空间信息的同时, 通过扩大感受野和提取上下文信息来增强模型的表达能。

最后, 添加了 1 个主分割头, 用于对像素进行类别预测, 从而完成细粒度的缺陷检测任务。

为了进一步提高模型的学习能力, RHNNet 在训练阶段引入辅助监督机制, 即在第三阶段的 SDBM1 后添加一个辅助分割头 (仅在训练阶段启用), 对中间层特征进行监督学习, 以加速网络收敛并提升检测

性能。

1.2 SDBM

RHNNet 模型的第二至第五阶段的输入、输出特征图的分辨率均为原始图像的 1/4, 若仅是简单堆叠卷积层, 会导致计算量显著增加。在语义分割领域, 多分支网络模型虽然在检测精度和推理速度方面表现出色, 但它们在利用空间信息时, 仅依赖一个堆叠了若干卷积层的分支, 难以充分保留空间细节信息, 从而影响目标的精细建模。同时, 高分辨率特征图能够有效保留更多的空间细节信息, 增强了对细小、狭长缺陷以及具有丰富细节缺陷边缘轮廓的建模能, 但也会显著增加计算复杂度。为此, 本文提出了轻量化 SDBM。

SDBM 主要由空间分支、上下文分支和融合模块组成, 结构如图 1 所示。

空间分支通过保持输入特征图的分辨率不变, 尽可能地保留细粒度的空间细节信息, 这对于细小、狭

长的缺陷检测尤为有效，并且能够更好地建模复杂的缺陷轮廓。

上下文分支通过逐步下采样来扩大感受野，提取深层语义信息，增强模型的抽象表达能力，并在输入融合模块前，通过插值调整将输出特征图恢复至与输入特征图相同的大小，以确保空间信息与上下文信息充分融合。

融合模块将空间分支和上下文分支的信息高效融合，实现语义信息和空间信息的互补。

SDBM 的结构（空间分支、上下文分支和融合模块）可根据任务需求灵活调整，从而在检测精度与推理速度之间实现平衡。本文基于实际应用的实时性需求，针对 RHNet 模型不同阶段的特点，设计了 3 种不同结构的 SDBM（SDBM1、SDBM2 和 SDBM3），分别用于 RHNet 模型第三至第五阶段。SDBM1、SDBM2 和 SDBM3 的上下文分支结构如图 2 所示。

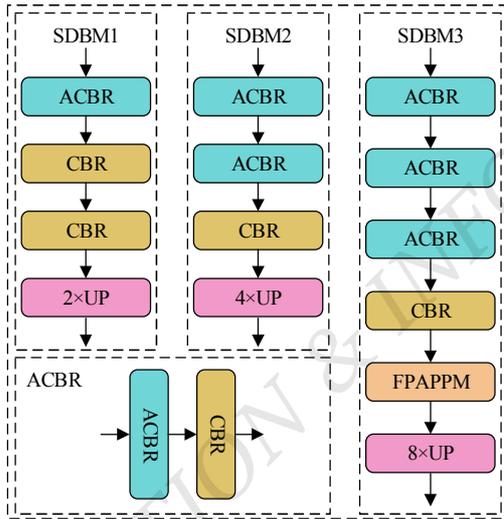


图 2 3 种 SDBM 的上下文分支结构图

上下文分支采用平均池化层（AvgP）与步距为 1 的卷积层相结合的 ACBR 模块进行下采样。为了更好地提取深层次的特征图，3 种 SDBM 均采用了逐步下采样的策略，即 SDBM1 将输入特征图下采样至原始分辨率的 1/2，SDBM2 将输入特征图下采样至原始分辨率的 1/4，SDBM3 将输入特征图下采样至原始分辨率的 1/8。

在融合方式上，3 种 SDBM 均采用了简单的相加

操作，不仅能够在网络中长时间保持高分辨率特征图，还能提升推理速度，从而更适应实时性需求。

为了提升推理速度，SDBM 尽可能简化了空间分支和融合模块。具体而言，SDBM1 和 SDBM2 的空间分支未进行额外操作；SDBM3 的空间分支仅添加了一个 1×1 卷积层，用于调节通道数，以便与上下文分支的输出特征图进行融合。

此外，SDBM3 的上下文分支还引入了更多的卷积层，并加入了 FPAPP，旨在加速高分辨率特征图中多尺度信息的提取和融合，同时保证检测性能不受过多影响。

1.3 FPAPP

为了提高缺陷检测的准确性，需充分挖掘深层信息和多尺度上下文信息。虽然金字塔池化模块（pyramid pooling module, PPM）及其改进版本深度聚合金字塔池化模块（deep aggregation PPM, DAPP）^[8]和并行聚合金字塔池化模块（parallel aggregation PPM, PAPP）^[9]在多项任务中表现较好，但其结构复杂，若直接将其应用于 RHNet 模型，可能会降低推理速度。为此，本文提出一种 FPAPP，旨在提高推理速度的同时，保持 PAPP 在高分辨率特征图下的性能。FPAPP 主要由一系列串联的 BRC 模块与池化层构成，这种串联结构在提取多尺度特征的过程中有效降低了计算开销，其结构如图 3 所示。

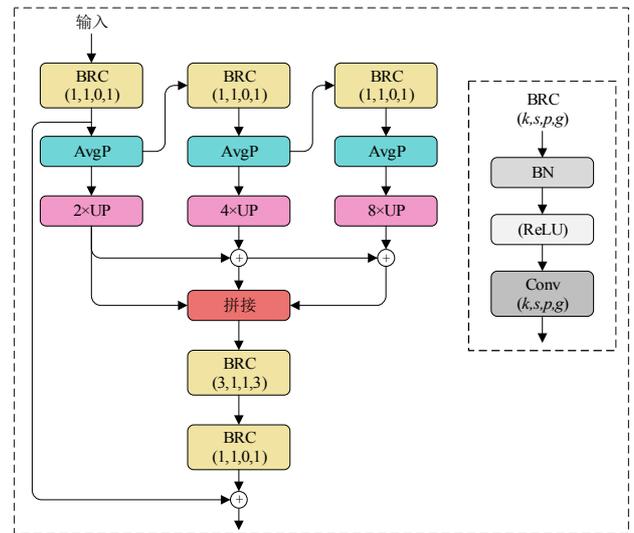


图 3 FPAPP 结构图

BRC 模块中(k, s, p, g)分别代表卷积核大小、步距、填充和卷积组数。首先, 输入特征图依次通过 3 个 BRC 模块与平均池化层 AvgP 的组合, 提取 3 个不同层次的特征图; 然后, 通过插值操作将这些不同层次的特征图调整为相同尺寸, 并进行拼接操作; 最后, 将拼接后的特征图输入到 2 个 BRC 模块中进行信息融合, 以进一步提高模型的特征表达能力。

FPAPPM 通过并行处理不同尺度的特征图来减少计算量, 同时保留重要的上下文信息, 从而在高分辨率特征图下提高推理速度。

1.4 损失函数

本文选取交叉熵损失函数作为 RHNet 模型的损失函数, 其数学表达式为

$$L_{ce}(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K \left[y_{ik} \log \hat{y}_{ik} + (1 - y_{ik}) \log(1 - \hat{y}_{ik}) \right] \quad (1)$$

式中: N 为像素总数, K 为类别总数, y 和 \hat{y} 分别为缺陷标签和预测结果。

为了更好地处理困难样本, 本文采用在线困难样本挖掘 (online hard example mining, OHEM) 策略^[10]来训练 RHNet。OHEM 策略主要包括 2 个步骤: 1) 模型使用标准的交叉熵损失函数进行训练; 2) 在每个训练周期结束后, 根据损失值对样本进行排序, 并对损失较大的样本进一步训练。损失较大的样本通常代表模型预测困难的样本, 强化训练这些困难样本有助于提高模型对复杂情况的辨识能力, 从而提升模型的整体性能。在训练过程中, 额外的监督信号可以优化模型参数。此外, 将辅助分割头放置在较浅的卷积层, 能够加速浅层卷积层参数的优化。因此, 在 SDBMI 后引入了辅助分割头。RHNet 模型的损失 L_{total} 为

$$L_{total} = L_d + 0.4L_a \quad (2)$$

式中: L_d 为主分割头的损失, L_a 为辅助分割头的损失。

2 实验验证

2.1 实验环境设置

本文采用不同的训练平台和测试平台, 验证 RHNet 模型在边缘端设备上的性能。

训练平台: 深度学习框架 PyTorch 1.13.1、语义分割工具包 mmsegmentation 1.2.2、计算架构 CUDA 11.7 以及操作系统 Ubuntu 22.04.6、英特尔至强 Gold 5320 CPU 以及英伟达 RTX 3090 GPU。

测试平台: ONNX Runtime-GPU 1.14.1、CUDA 11.6 以及操作系统 Windows 11、酷睿 i7-8750H CPU 以及英伟达 GTX 1060 GPU。

2.2 数据集构建与预处理

本文利用 3 个公开的表面缺陷数据集 (单晶硅电池表面缺陷数据集 (SCD)、SD-saliency-900 数据集 (SDD)^[11]和道路裂缝数据集 (CFD)^[12]) 进行 RHNet 模型的性能验证。

SCD 包含 190 幅缺陷图像, 图像分辨率为 448×448 像素, 且这些图像仅包含一种缺陷。

SDD 包含热轧带钢的 3 种典型表面缺陷 (夹杂物、斑块和划痕)。每种缺陷类别包含 300 幅图像, 图像分辨率为 200×200 像素。

CFD 专门针对细小裂缝检测任务构建, 包含 118 幅带有标签的图像, 图像分辨率为 480×320 像素。考虑到道路裂缝与工业产品表面缺陷在图像上具有高度的相似性, 且细微裂缝能够更好地考验模型的检测能力, 因此选择该数据集进行测试。

由于上述数据集相对较小, 模型在训练过程中可能面临过拟合的问题。因此, 本文采用水平翻转、垂直翻转、逆时针旋转 90°、180°等数据增强方法来扩展数据集, 并将 3 个数据集按 8:2 的比例随机划分为训练集和测试集。

2.3 评价指标

本实验采用平均交并比 (mean intersection over union, mIoU), mF1score 和平均准确率 (mean accuracy,

mAcc) 来评价模型的检测精度; 采用每秒检测帧数 (FPS) 来评价模型的推理速度。

mIoU 衡量了模型的预测分割结果与真实分割结果之间的重叠程度, 计算公式为

$$m_{IoU} = \frac{1}{N} \sum_{i=1}^N \frac{P_{ii}}{\sum_{j=1}^N P_{ij} + \sum_{j=1}^N (P_{ji} - P_{ii})} \quad (3)$$

式中: p_{ii} 为被正确分类为第 i 类的像素数量, p_{ji} 为被错误分类为第 i 类的像素中实际属于第 j 类的像素数量。

mF1 为查准率 (P) 和查全率 (R) 的调和平均数, mAcc 为所有类别的平均准确率, 二者的计算公式分别为

$$P_i = \frac{TP_i}{TP_i + FP_i} \quad (4)$$

$$R_i = \frac{TP_i}{TP_i + FN_i} \quad (5)$$

$$m_{F1} = \frac{1}{N} \sum_{i=1}^N \frac{2 \times P_i \times R_i}{P_i + R_i} \quad (6)$$

$$m_{Acc} = \frac{1}{N} \sum_{i=1}^N \frac{TP_i + TN_i}{TP_i + FP_i + TN_i + FN_i} \quad (7)$$

式中: TP_i 、 FP_i 、 TN_i 和 FN_i 分别为第 i 类的真阳性、假阳性、真阴性和假阴性的数量, P_i 、 R_i 分别为第 i 类的查准率和查全率。

2.4 实验结果

本实验采用了 SGD 优化器和 Poly 学习率下降策略来训练模型, 共训练 150 个 epoch。其中, 动量、权重衰减和初始学习率分别设置为 0.9、0.0005 和 0.01。在训练过程中, 未对图像进行任何处理, 直接将原始图像作为输入。3 个数据集的批量大小设置如下: SCD 为 24、SDD 为 32、CFD 为 8。

为验证本文 (RHNet) 模型的有效性与优越性, 在相同的实验数据集和测试环境下, 将 RHNet 模型与 FCN8s (Res18)^[13]、PSPNet (Res18)^[14]、DeepLabv3+ (Res18)、BiSeNetV1^[16]、BiSeNetV2^[17]、STDCV1^[18]、STDCV2^[18]、DDRNet^[19]、PIDNet^[20] 模型进行对比实验, 结果如表 1 所示。其中, (Res18) 表示该模型使用 ResNet18 作为骨干网络。

表 1 10 种模型对比实验结果

模型	SCD				SDD				CFD			
	mIoU/ %	mF1/ %	mAcc/ %	FPS/ (f/s)	mIoU/ %	mF1/ %	mAcc/ %	FPS/ (f/s)	mIoU/ %	mF1/ %	mAcc/ %	FPS/ (f/s)
FCN8s(Res18)	77.2	85.7	82.8	33.0	82.3	90.0	89.2	128.3	56.8	62.7	58.2	44.3
PSPNet(Res18)	77.5	85.9	83.3	32.2	82.5	90.0	89.2	126.7	61.6	69.3	64.0	43.7
DeepLabv3+(Res18)	79.1	87.1	85.7	30.1	84.9	91.6	91.2	119.2	70.5	79.4	75.1	39.5
BiSeNetV1	77.8	86.1	83.4	80.7	82.9	90.3	89.4	182.5	63.0	71.2	66.0	102.2
BiSeNetV2	78.6	86.8	84.9	62.1	84.0	91.1	90.4	130.0	66.2	74.9	70.4	89.5
STDCV1	75.5	84.3	80.1	76.6	82.3	90.0	88.9	161.4	58.6	65.4	60.5	111.6
STDCV2	74.7	83.6	78.9	58.0	82.1	89.8	88.6	110.8	57.8	64.3	59.5	82.1
DDRNet	78.3	86.6	84.3	74.3	85.8	92.2	91.5	153.0	68.8	77.7	74.2	114.3
PIDNet	77.7	86.1	83.4	58.4	85.5	92.0	91.6	110.0	67.2	76.0	71.8	89.2
RHNet	80.3	88.1	87.1	100.8	88.0	93.5	93.2	257.8	76.1	84.5	83.3	156.3

由表 1 可知, RHNet 模型在 3 个数据集上的各项指标均达到了最佳水平, 尤其在 CFD 中, RHNet 模型的 mIoU、mF1 和 mAcc 分别比排名第二的

DeepLabv3+ 高出 5.6%、5.1% 和 8.2%; RHNet 模型的 FPS 比排名第二的 DDRNet 高出 42 f/s。

为了更直观地比较各模型的性能, 对实验结果进

行可视化处理，结果如图 4 所示。其中，第四行展示了形状较大且轮廓信息复杂的缺陷。

由图 4 可知：RHNet 模型不仅能够完整检测出所有缺陷，还能够较好地还原缺陷的轮廓细节；且能够

准确检测出 CFD 中大部分细小且狭长的裂缝，而其他模型的检测结果或者不连续，或者未完全检测出缺陷，表明 RHNet 模型在保持空间细节信息方面具有明显优势。

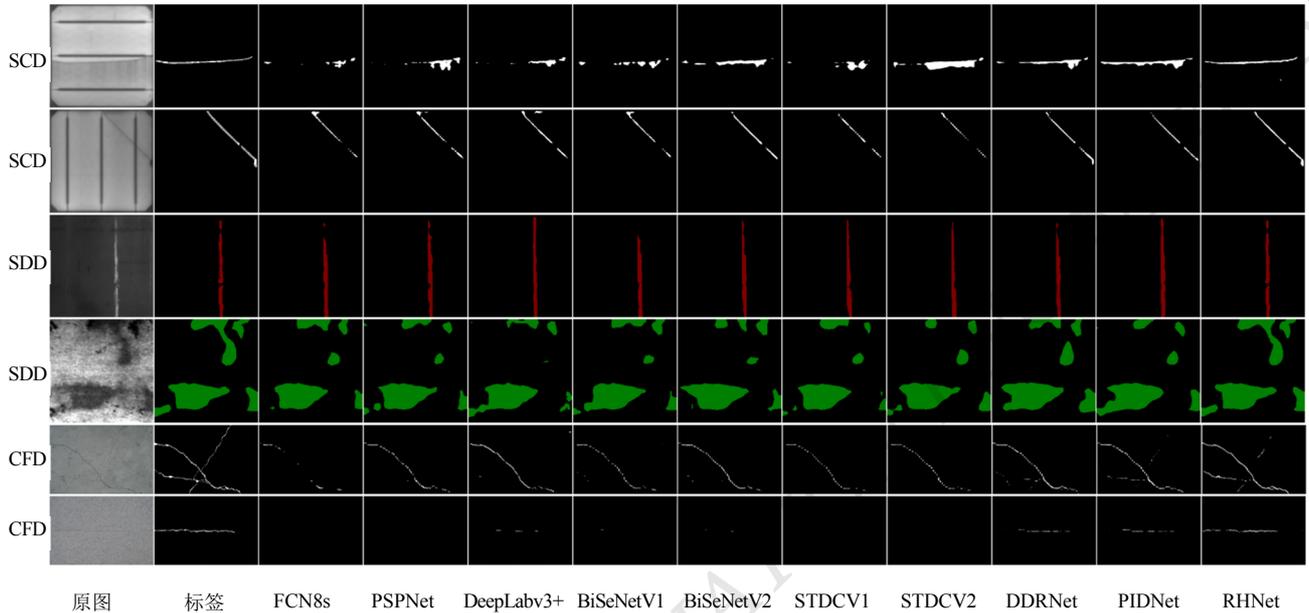


图 4 10 种模型检测结果可视化对比图

为了验证本文模型的有效性，在 SCD 上进行消融实验，比较 PAPPm 和 FPAPPm 在 DDRNet、PIDNet 和 RHNet 模型上的性能表现，结果如表 2 所示。

表 2 消融实验结果

模型	mIoU/ %	mF1/ %	mAcc/ %	FPS/ (f/s)
DDRNet+PAPPm	78.17	86.41	84.33	85.27
DDRNet+FPAPPm	78.15	86.38	84.42	90.68
PIDNet+PAPPm	77.65	86.06	83.39	58.39
PIDNet+FPAPPm	77.61	86.01	83.27	65.19
RHNet+PAPPm	80.42	88.12	87.37	89.10
RHNet+FPAPPm	80.31	88.06	87.13	100.82

由表 2 可知，虽然 FPAPPm 的检测精度稍微降低，但其推理速度明显提升，且推理速度差距随着输入图像尺寸的增大而进一步加大，表明 FPAPPm 有效缓解了 RHNet 模型中因过大特征图所导致的推理速度下降问题。

3 结论

本文针对工业产品表面缺陷检测任务存在的空间信息丢失、缺陷轮廓不完整和漏检细小缺陷等问题，提出了 RHNet 模型。该模型通过保持较高分辨率的特征图来保留更多的空间细节信息，提升了检测精度，能够检测出更多缺陷的轮廓细节，对细小缺陷更为敏感。为了提高推理速度，RHNet 模型采用了 SDBM 和 FPAPPm，以减少计算量。实验结果验证，RHNet 模型能够有效平衡检测精度和推理速度，满足工业产品表面缺陷检测任务的要求。未来研究将尝试进一步优化 RHNet 模型的检测性能，包括提升大尺寸输入下的推理速度，以及提高模型对复杂缺陷模式的预测能力。

©The author(s) 2024. This is an open access article under the CC BY-NC-ND 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

参考文献

- [1] QI Qinghua, XU Xiaowei, ZHAN Liu, et al. Research on YOLOX-based tire defect detection method[C]. 2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI). IEEE, 2022:1-6.
- [2] PAN J, ZENG D, TAN Q, et al. EU-Net: A novel semantic segmentation architecture for surface defect detection of mobile phone screens[J]. IET Image Processing, 2022,16(10):2568-2576.
- [3] ZHANG J, DING R, BAN M, et al. FDSNeT: An accurate real-time surface defect segmentation network[C]//ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022:3803-3807.
- [4] ZHU J, HE G, ZHOU P. MFNet: A Novel Multilevel Feature Fusion Network with Multibranch Structure for Surface Defect Detection[J]. IEEE Transactions on Instrumentation and Measurement, 2023,72:1-11.
- [5] CAO J, YANG G, YANG X. A pixel-level segmentation convolutional neural network based on deep feature fusion for surface defect detection[J]. IEEE Transactions on Instrumentation and Measurement, 2020,70:1-12.
- [6] ZHANG Z, WANG W, TIAN X. Semantic segmentation of metal surface defects and corresponding strategies[J]. IEEE Transactions on Instrumentation and Measurement, 2023,72:1-13.
- [7] 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.
- [8] PAN H, HONG Y, SUN W, et al. Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes[J]. IEEE Transactions on Intelligent Transportation Systems, 2022,24(3):3448-3460.
- [9] XU J, XIONG Z, BHATTACHARYYA S P. PIDNet: A real-time semantic segmentation network inspired by PID controllers[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:19529-19539.
- [10] SHRIVASTAVA A, GUPTA A, GIRSHICK R. Training region-based object detectors with online hard example mining [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:761-769.
- [11] SONG G, SONG K, YAN Y. Saliency detection for strip steel surface defects using multiple constraints and improved texture features[J]. Optics and Lasers in Engineering, 2020,128:106000.
- [12] SHI Y, CUI L, QI Z, et al. Automatic road crack detection using random structured forests[J]. IEEE Transactions on Intelligent Transportation Systems, 2016,17(12):3434-3445.
- [13] SUN W, WANG R. Fully convolutional networks for semantic segmentation of very high resolution remotely sensed images combined with DSM[J]. IEEE Geoscience and Remote Sensing Letters, 2018,15(3):474-478.
- [14] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017:2881-2890.
- [15] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation[J]. arXiv preprint arXiv:1706.05587, 2017.
- [16] YU C, WANG J, PENG C, et al. Bisenet: Bilateral segmentation network for real-time semantic segmentation[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018:325-341.
- [17] YU C, GAO C, WANG J, et al. Bisenet v2: Bilateral network with guided aggregation for real-time semantic segmentation [J]. International Journal of Computer Vision, 2021,129:3051-3068.
- [18] FAN M, LAI S, HUANG J, et al. Rethinking bisenet for real-time semantic segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:9716-9725.
- [19] PAN H, HONG Y, SUN W, et al. Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes[J]. IEEE Transactions on Intelligent Transportation Systems, 2022,24(3):3448-3460.
- [20] XU J, XIONG Z, BHATTACHARYYA S P. PIDNet: A real-time semantic segmentation network inspired by PID controllers[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:19529-19539.

作者简介:

王伊杰, 男, 1999 年生, 硕士研究生, 主要研究方向: 图像处理、语义分割。E-mail: wangyijie_2023@163.com

蔡建扑, 男, 2000 年生, 硕士研究生, 主要研究方向: 工业自动化、大模型。E-mail: 13798295109@163.com

任志刚 (通信作者), 男, 1987 年生, 博士研究生, 副教授, 主要研究方向: 复杂工业过程优化与控制、工业智能。

E-mail: renzhigang@gdut.edu.cn