

DOI: 10.19650/j.cnki.cjsi.J2513911

基于多尺度特征融合网络的电阻点焊缺陷精确定位与分割方法^{*}

王俞霏, 杜弘志, 胡蕴博, 孙岩标, 郑继贵

(天津大学精密测试技术及仪器全国重点实验室 天津 300072)

摘 要:电阻点焊质量直接影响车身结构的稳定性和安全性,缺陷的像素级分割图对于精确分析缺陷形态和严重程度至关重要。针对传统的目标检测方法在处理细小缺陷时无法进行精确语义分割且分类精度较低的问题,提出了一种基于多尺度特征融合网络的电阻点焊缺陷精确定位与分割方法,通过跨层次特征连接与多尺度特征匹配,使网络在识别整体点焊特征的同时,能够捕捉细小缺陷特征,实现了大场景中缺陷位置的精确语义分割,并提高了电阻点焊区域的分类精度。设计了候选区域生成网络,融合低层次细节特征与高层次语义信息,设计了点焊区域定位损失函数,确保点焊区域的精准定位。随后,提出了缺陷分割与定位网络,结合 ROI Align 与多尺度特征匹配,构建正常焊点特征库,并设计异常评分计算公式,实现点焊区域的像素级异常评分。实验结果表明,方法相比传统目标检测模型,在小目标点焊分类精度上提升了 25.35%, $F1$ 分数提升了 14.81%。此外,所提方法能够生成高精度的像素级缺陷分割图,Pixel AUROC 达到 0.94,展现了优异的缺陷识别能力,同时在开源不同场景下电阻点焊数据集上的测试也取得了良好结果, $F1$ 分数达 0.93,验证了模型的泛化能力。

关键词:电阻点焊;多尺度特征融合;像素级语义分割;异常检测

中图分类号: TH164 TH183

文献标识码: A

国家标准学科分类代码: 510.10

Precise localization and segmentation method for resistance spot welding defects based on multi-scale feature fusion network

Wang Yufei, Du Hongzhi, Hu Yunbo, Sun Yanbiao, Zhu Jigui

(State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China)

Abstract: The quality of resistance spot welding directly affects the structural stability and safety of automobile bodies. Pixel-level segmentation maps of welding defects are crucial for accurately analyzing defect morphology and severity. To address the limitations of traditional object detection methods in precisely segmenting small-scale defects and achieving high classification accuracy, this paper proposes a precise localization and segmentation method for RSW defects based on a multi-scale feature fusion network. By integrating cross-level feature connections and multi-scale feature matching, the network captures both global welding characteristics and fine-grained defect details, enabling accurate semantic segmentation of defects in large scenes and improving classification accuracy in RSW regions. A candidate region generation network is designed to fuse low-level detailed features with high-level semantic information, and a custom localization loss function is introduced to ensure accurate positioning of spot weld regions. Subsequently, a defect segmentation and localization network is proposed, which incorporates ROI Align and multi-scale feature matching to construct a normal feature bank for spot welds and formulates an anomaly scoring function for pixel-level anomaly scoring of weld regions. Experimental results show that, compared with traditional object detection models, the proposed method improves the classification accuracy for small RSW targets by 25.35% and enhances the $F1$ score by 14.81%. Moreover, it produces high-precision pixel-level segmentation maps, achieving a Pixel AUROC of 0.94, demonstrating excellent defect recognition capabilities. The method also achieves good performance on open-source RSW datasets from various industrial scenarios, with an $F1$ score of 0.93, verifying the generalization ability of the model.

Keywords: resistance spot welding; multi-scale feature fusion; pixel-level semantic segmentation; anomaly detection

0 引言

在汽车制造行业中,电阻点焊是关键工艺之一,焊点质量直接影响整车的结构强度、安全性及使用寿命^[1]。汽车车身通常由数千个焊点连接,其强度决定了整车在碰撞等极端工况下的稳定性和乘员安全。由于焊接过程中受到焊接工艺参数、材料特性、设备状态等因素的影响,焊点缺陷不可避免地发生,如气孔、裂纹、未熔合等缺陷可能导致焊接部位的疲劳失效、承载能力下降,甚至影响整车的碰撞安全性能^[2]。因此,汽车焊点缺陷的精准检测对于确保车身结构的完整性和行驶安全至关重要。随着汽车制造向高效、智能化方向发展,如何实现自动化、高精度的焊点缺陷检测成为当前研究的核心问题^[3]。

传统焊点缺陷检测方法主要包括电阻法、视觉检测法和渗透检测法。电阻法通过测量焊点的电阻变化来评估焊接质量,但受限于焊点尺寸和材料特性,难以精准识别微小缺陷^[4]。视觉检测法利用图像处理技术识别焊点表面缺陷,适用于外观缺陷检测,但对复杂背景和光照变化较为敏感,容易产生误检或漏检^[5]。渗透检测法利用荧光或着色渗透剂渗入焊点裂纹,通过紫外光照射显现缺陷,适用于表面微小裂纹的检测,但检测过程繁琐且需要专业操作^[6]。因此,这些传统方法在检测精度、自动化能力及复杂焊点结构适应性方面仍存在诸多局限性^[7]。

近年来,深度学习方法,特别是目标检测技术,逐渐成为焊点缺陷检测的研究热点^[8]。相比传统方法,目标检测能够利用卷积神经网络提取焊点的高层次特征,结合端到端的检测框架,实现对焊点缺陷的高效、精准检测,提高检测的自动化水平和可靠性^[9]。随着目标检测技术的不断发展,其在工业焊接质量检测中的应用已成为重要方向。

目标检测技术已广泛应用于工业异常检测任务,尤其是在焊点缺陷、表面瑕疵检测和制造质量控制领域。YOLO(you only look once)、Faster R-CNN(faster region-based convolutional neural network)等目标检测模型能够在复杂工业环境中高效检测缺陷,并实现精准定位^[10]。Zuo等^[11]提出了HD-YOLO(hyperdefect-YOLO)框架,在YOLO结构中引入超图计算,以增强模型的表示能力,在印刷电路板缺陷检测数据集(high-resolution industrial printed circuit board, HRIPCIB)等数据集上取得了显著提升。此外,Qi等^[12]研究了基于YOLO的产品缺陷检测方法,通过实时分类和定位提高了工业质量控制的效率。Wang等^[13]提出了SSD-Faster Net(single shot multibox detector-faster net),将SSD与Faster R-CNN结合用于工业缺陷检测,并在多个工业组件缺陷检测任务中表现出色。

然而,目标检测方法在工业异常检测中仍存在一定局限性。首先,目标检测算法依赖于大量的训练数据,难以覆盖所有可能的缺陷类型,尤其是在工业应用中,缺陷样本往往较少且具有较大多样性。其次,目标检测算法在小目标检测上存在明显不足,尤其是当焊点尺寸较小时,判断缺陷的精度显著下降。最重要的是,目标检测方法通常只能提供缺陷的边界框信息,而无法提供精确的像素级缺陷分割。这使得它在细粒度缺陷分析中的应用受到限制。

在此背景下,无监督工业异常检测方法逐渐受到关注。无监督方法能够在仅有正常样本的情况下进行训练,并检测异常样本的偏离程度,使其在缺陷样本不足或标注成本较高的场景下表现出较强的适应性。相比于目标检测方法,无监督异常检测方法主要基于数据分布的异常性进行检测,能够在未知缺陷场景中保持较好的泛化能力,并在小样本情况下仍然具备较高的检测准确率。因此,近年来大量研究开始探索基于无监督学习的工业异常检测方法,以弥补目标检测方法的不足。

无监督工业异常检测主要包括基于重构的方法、自监督学习方法以及深度特征嵌入方法。基于重构的方法通过在正常样本上训练模型,使其能够重构正常数据,并假设异常样本难以被模型重构。然而,基于重构的方法往往容易过拟合,使得异常样本的重构误差减小,从而降低检测效果。典型方法包括自编码器^[14](Autoencoder)、生成对抗网络^[15](generative adversarial network, GAN)、Transformer^[16]和扩散模型^[17](diffusion model)。其中,Autoencoder通过编码和解码学习数据的低维表示,但可能学习到异常样本的分布特征,导致检测能力下降。生成对抗网络依赖生成器与判别器的对抗训练以捕获数据分布,但训练过程容易出现模式崩溃,影响异常检测的稳定性。Transformer通过自注意力机制建模全局特征,在异常检测中表现出较好的特征表达能力,但通常需要大规模数据训练,计算资源需求较高。扩散模型则通过逐步添加噪声并学习去噪过程实现数据生成,具有较强的异常检测能力,但推理速度较慢,影响工业检测的实时性。

相比之下,自监督学习方法通过设计预训练任务,使模型在无标注数据上学习有助于异常检测的特征。例如,CutPaste^[18]通过在正常图像上剪切并粘贴区域生成伪异常样本,使模型学习区分正常与伪异常,但伪异常可能与真实异常存在较大差异,影响检测泛化能力。RealNet^[19]利用扩散模型生成更自然的异常样本,但合成异常的多样性有限,可能导致泛化性能下降。去噪残差与异常编码器——解码器^[20](denoising residual and anomaly encoder-decoder, DRAEM)通过在正常图像上添加噪声并训练模型重构正常图像,以此检测异常区域,但对噪声类型和强度较为敏感,在复杂场景下可能出现误检。

深度特征嵌入方法则利用预训练模型提取图像特征,并在特征空间中检测异常。其中图像块分布建模^[21](patch distribution modeling, PaDiM),通过构建多元高斯分布模型并计算马氏距离进行异常检测,能够较好地刻画正常样本分布,但存储大量特征带来较高的内存开销,并增加计算复杂度。PatchCore^[22]通过构建记忆库来提升异常检测性能,减少模型对异常样本的依赖,但在高分辨率图像上的计算资源需求仍然较大。尽管这些方法在无监督工业异常检测领域取得了显著进展,但仍然存在各自的局限性,如对异常样本建模能力不足、计算资源需求较高、泛化能力受限等问题。因此,如何进一步优化这些方法,提高检测性能和计算效率,仍然是当前研究的关键问题。

1 方 法

1.1 概述

针对电阻点焊缺陷检测中存在的焊点密集微小、缺陷形态细碎多样、样本覆盖不足等特性,传统方法面临语义分割精度不足及异常分类性能下降的双重挑战,提出多尺度特征融合网络,网络框架如图1所示,网络架构由候选区域生成网络和缺陷分割与定位网络两部分组成,候选区域生成网络通过跨层级特征连接实现对多尺度电阻点焊区域的精准识别,缺陷分割与定位网络通过多尺度特征匹配实现像素级的异常评分与精细缺陷分割定位,有效提升大场景焊接质量检测可靠性。

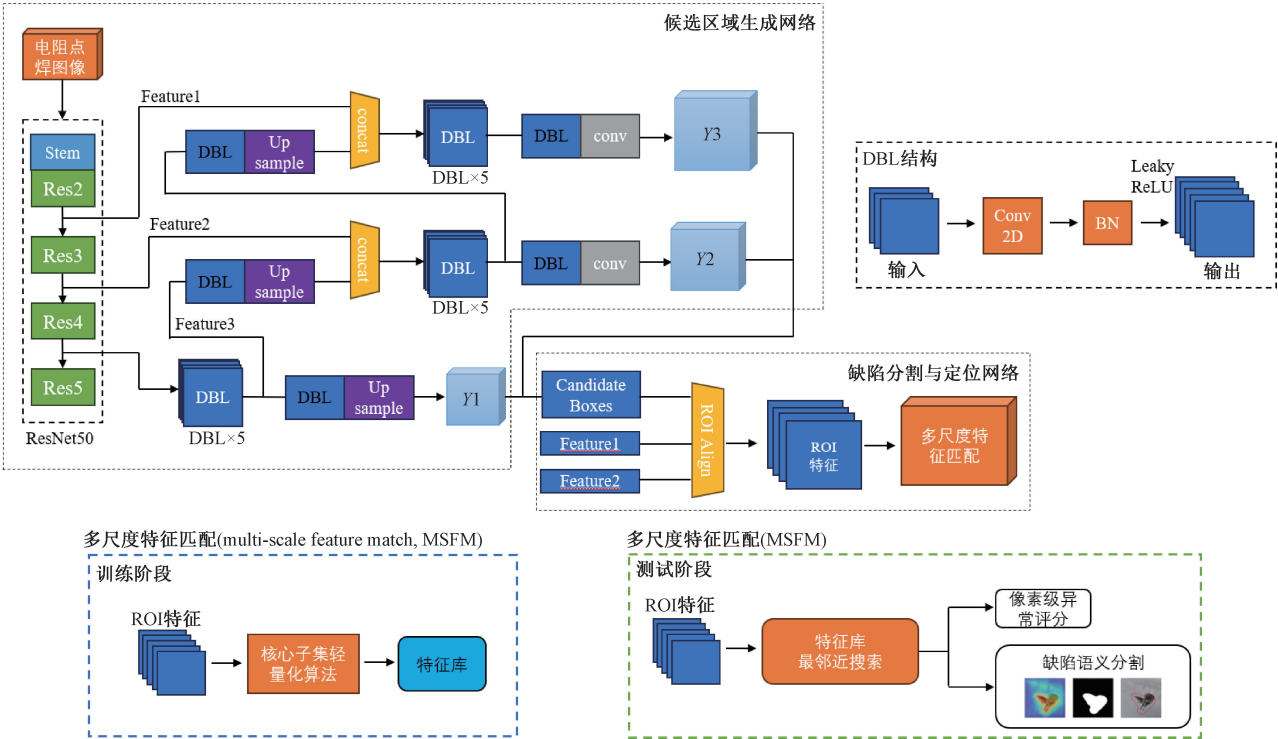


图1 电阻点焊缺陷检测网络框架

Fig. 1 Weld point defect detection network framework

候选区域生成网络首先利用预训练的 ResNet50 骨干网络提取输入图像的多尺度特征,这些特征有助于捕捉不同尺度的焊点缺陷信息。网络通过跨层特征连接与上采样操作,使得低层次的细节特征和高层次的语义信息能够有效融合。最终,网络根据融合后的特征生成候选框,并对每个候选区域进行精细预测,精确定位可能包含缺陷的区域。

缺陷分割与定位网络利用感兴趣区域对齐(region of interest align, ROI Align)对每个候选框内的区域进行高精度特征提取,并将其映射为固定维度的特征表示。通

过特征匹配方法,该网络将提取的特征与正常焊点特征库中的特征进行比较,计算像素级的异常评分,从而精确确定缺陷的位置和大小。此外,网络还会生成精确的缺陷分割图,通过像素级分割结果,提供焊点缺陷的详细位置和健康状态评估。

在训练阶段,候选区域生成网络通过多尺度特征提取与融合在标注数据集上进行训练,采用回归分析任务学习精确定位焊点区域,从而生成电阻点焊区域的候选框。在此阶段,网络不训练分类任务,专注于定位目标的准确性。与此同时,缺陷分割与定位网络利用数据集中

的标注区域与低层次特征经过 ROI Align 操作,从正常焊点区域中精确提取特征,并基于这些特征构建正常焊点的特征库。在特征库的构建过程中,通过核心子集选择策略对特征库进行优化,筛选出具有代表性的特征,确保在后续测试阶段能够高效且准确地匹配和检测焊点缺陷。

在测试阶段,输入的电点焊图像首先通过候选区域生成网络生成候选框。随后,缺陷分割与定位网络利用 ROI Align 操作精确提取每个候选框内的特征,确保从图像中获取高质量的局部特征表示。接着,提取的特征与正常焊点特征库中的特征进行匹配,通过计算候选区域特征与特征库中代表性特征之间的相似度,进行像素级异常评分,从而判断焊点是否存在缺陷。这一过程通过最近邻匹配算法完成,确保每个焊点区域的异常得分准确反映其缺陷程度。最后,网络生成精确的缺陷分割图,明确标定缺陷的位置、形态及其严重性,从而为焊点质量评估提供详尽的决策支持。

1.2 候选区域生成网络

候选区域生成网络采用端到端单阶段目标检测框架,通过将输入图像划分为多个网格,每个网格基于多个预定义的 anchor 框,直接预测目标区域的位置、大小、置信度及类别概率。网络架构主要包括特征提取、特征融合和检测头 3 个部分。

鉴于电阻点焊缺陷检测对特征表达能力的高要求,在特征提取方面,采用了 ResNet50 作为骨干网络。ResNet50 采用深度残差结构,通过残差连接有效缓解深层网络训练中的梯度消失问题,从而提高训练稳定性和收敛速度。ResNet50 由 Stem(初始卷积与池化层)以及 Res2、Res3、Res4、Res5 这 4 个阶段组成,所提方法选取 Res2、Res3、Res4 作为特征融合模块的输入,以更好地适应电阻点焊检测任务对多尺度特征的需求。

在特征融合阶段,所提方法主要采取跨层级特征连接与密集块层(densely blocked layer, DBL)结构的组合来增强模型的多尺度感知能力。具体而言,深层低分辨率特征图通过上采样逐步放大,并与浅层高分辨率特征图进行拼接(concat),以便充分结合低层的细节信息和高层的语义信息,从而提高模型对不同尺度目标的识别能力,尤其在检测复杂背景下的小目标电阻点焊时,能够展现出更强的表现。

跨层级特征连接通过将不同层次的特征图进行结合,使得网络能够同时利用低层次的细节特征和高层次的语义特征。低层特征图通常能够捕捉到电阻点焊的边缘、纹理等局部细节,而高层特征图则更多地聚焦于图像的全局语义,如电阻点焊的形状等。通过跨层连接,能够有效地将这两者结合,从而提升网络在多尺度目标检测中的能力。

在具体实现时,所提方法使用上采样操作将低分辨率的深层特征图恢复到与高分辨率浅层特征图相同的尺寸,使用双线性插值进行上采样,即:

$$\hat{F}(x, y) = \sum_{i=0}^1 \sum_{j=0}^1 F(x_i, y_j) \cdot \omega(i, x) \cdot \omega(j, y) \quad (1)$$

式中: $\hat{F}(x, y)$ 是上采样后的位置 (x, y) 处的特征值; $F(x_i, y_j)$ 是原特征图中的 4 个邻近像素值; $\omega(i, x)$ 和 $\omega(j, y)$ 是双线性插值的权重函数,用于计算目标位置的值。

然后通过拼接操作将这 2 种不同层次的特征进行融合。假设第 l 层的特征图为 F_L , 第 $l+1$ 层的特征图为 F_H , 跨层级连接后的融合特征图 F_{fused} 可以通过式(2)表示,即:

$$F_{fused} = \text{concat}(F_L, F_H) \quad (2)$$

这一过程将细节信息和语义信息结合在一起,为模型提供更加丰富的特征表示,特别是在对小目标和细小缺陷的检测中,能更好地保持空间信息和语义信息的平衡。

DBL 结构由卷积层、批归一化(batch normalization, BN)层和 Leaky ReLU 激活函数组成,具体设置为:

1) 卷积层

卷积核大小:采用 3×3 的卷积核大小,能够有效捕捉图像的局部特征。相较于 1×1 的小卷积核, 3×3 的卷积核可以提取更丰富的上下文信息,能够在空间上更好地融合周围的特征信息,对于处理不同尺寸的缺陷有更强的适应性。其较大的感受野能够帮助网络更好地理解局部和全局的空间关系,提高了对各种尺度缺陷的检测能力。

步长设置为 1,确保特征图的空间尺寸保持不变,并且有助于提取更多的细节信息。

填充方式:采用同填充(Padding = 1),即在输入图像的边缘填充适量的零值,保证卷积操作后输出特征图的空间尺寸不变。

2) 批归一化

批归一化层用于对卷积层输出进行标准化,使得每层的输入数据均值为 0,方差为单位值。批归一化的公式为:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (3)$$

式中: x_i 是输入数据; μ_B 是当前批次的均值; σ_B^2 是当前批次的方差; ϵ 是一个小常数,用于避免除零错误。通过批归一化,数据的分布得以标准化,这使得网络训练更加稳定,加速了收敛过程。

3) Leaky ReLU 激活函数

Leaky ReLU 能够引入非线性并增强模型的学习能力。其数学公式为:

$$f(x) = \max(0, 1x, x) \quad (4)$$

Leaky ReLU 在负半轴也提供了一个小的梯度,避免了“死神经元”问题。通过给负值区域提供一个小的斜率(0.1),Leaky ReLU 激活函数确保了网络不会在训练过程中因为负值输入而“失活”,从而提高了训练效率和模型的表达能力。这种激活函数对于细节信息的提取尤其重要,尤其是在小尺寸电阻点焊检测任务中,能够使网络更好地学习和保留微弱的缺陷特征。

检测头负责在融合后的特征图上进行目标区域的定位与预测,生成每个候选区域的边界框位置、目标类别以及置信度。在这一部分,网络通过多尺度特征图进行预测,以适应不同大小的目标。在每个检测尺度上的特征图经过 1×1 卷积和 3×3 卷积提取局部信息,并最终基于预定义的 anchor 框进行目标区域的预测,即每个网格单元预测多个 anchor 框的边界框位置、类别概率和目标置信度。为了减少冗余候选框并提高检测精度,最终检测结果通过置信度阈值过滤和非极大值抑制(non-maximum suppression, NMS)进行筛选,得到最优目标框。

在所提方法中,候选区域生成网络的总损失函数主要包括两个部分:边界框回归损失、目标置信度损失,其表达式如式(5)所示。

$$L_{total} = \lambda_{coord} L_{bbox} + \lambda_{conf} L_{conf} \quad (5)$$

式中: L_{bbox} 是边界框回归损失,优化框的位置和大小; L_{conf} 是目标置信度损失,用于判断候选框是否包含电阻点焊目标。

边界框回归损失的目标是优化中心点坐标 (x, y) 和宽度、高度 (w, h) 的回归误差,具体而言,目标框的坐标回归损失和尺寸回归损失分别计算,并在总损失函数中加权求和。通过最小化这些损失,模型能够精确回归目标的位置和大小,即:

$$L_{bbox} = \lambda_{coord1} L_{coord1} + \lambda_{coord2} L_{coord2} \quad (6)$$

式中: L_{coord1} 为中心点坐标 (x, y) 的回归损失; L_{coord2} 为宽度、高度 (w, h) 的回归损失,其表达式如式(7)~(8)所示。

$$L_{coord1} = \sum_i \text{smooth}_{L1}(\hat{x}_i, x_i) + \text{smooth}_{L1}(\hat{y}_i, y_i) \quad (7)$$

$$L_{coord2} = \sum_i \text{smooth}_{L1}(\hat{w}_i, w_i) + \text{smooth}_{L1}(\hat{h}_i, h_i) \quad (8)$$

式中: (\hat{x}_i, \hat{y}_i) 表示网络预测的第 i 个框的预测中心点坐标; (x_i, y_i) 表示第 i 个框的真实中心点坐标; (\hat{w}_i, \hat{h}_i) 表示网络预测的第 i 个框的预测宽度和高度; (w_i, h_i) 表示第 i 个框的真实宽度和高度。平滑 L1 损失的定义为:

$$\text{smooth}_{L1}(a, b) = \begin{cases} 0.5(a - b)^2, & |a - b| < 1 \\ |a - b| - 0.5, & |a - b| \geq 1 \end{cases} \quad (9)$$

目标置信度损失被用来优化候选框的预测置信度,即每个框是否包含目标物体。具体地,置信度损失通过计算网络预测的置信度与真实标签之间的差异,来引导网络更准确地定位包含目标的框,并提高预测结果的可

靠性,置信度损失定义为:

$$L_{conf} = - \sum_i \sum_j I_{obj}^{ij} [\hat{c}_i^j \log(c_i^j) + (1 - \hat{c}_i^j) \log(1 - c_i^j)] - \lambda_{noobj} \sum_i \sum_j I_{noobj}^{ij} [\hat{c}_i^j \log(c_i^j) + (1 - \hat{c}_i^j) \log(1 - c_i^j)] \quad (10)$$

式中: I_{obj}^{ij} 表示第 i 个网格的第 j 个框是否包含目标物体; \hat{c}_i^j 是网络预测的第 j 个候选框包含目标的置信度; c_i^j 是实际情况中框 j 是否包含目标的真实标签; λ_{noobj} 为权重系数,用于调节框内没有目标对训练的影响; I_{noobj}^{ij} 表示框 j 中是否不包含目标物体。

在得到候选框后,网络通过 NMS 进一步处理预测结果,以减少冗余的候选框。NMS 的过程为:

1) 计算 IoU: 对每对候选框 B_1 和 B_2 , 计算它们之间的交并比 IoU, 即:

$$IoU(B_1, B_2) = \frac{\text{Area of Intersection}(B_1, B_2)}{\text{Area of Union}(B_1, B_2)} \quad (11)$$

2) 筛选框: 如果候选框 B_2 与框 B_1 的 IoU 超过设定的阈值 θ , 并且 B_1 的置信度较低, 则去除 B_1 , 保留 B_2 。

3) 最终框选择: 经过 NMS 处理后, 网络仅保留最终检测框中置信度最高的候选框, 去除冗余框, 得到最优的目标框。

1.3 缺陷分割与定位网络

本研究提出了一种基于多尺度特征库构建的缺陷分割与定位网络。在训练阶段构建正常电阻点焊图像的多尺度特征库, 推理阶段则通过将测试图像的多尺度特征与特征库进行匹配, 计算异常分数, 从而实现电阻点焊缺陷的精准检测与像素级语义分割。

多尺度特征的选择: 虽然更深层的特征(Res4、Res5)具备更强的全局表达能力, 但由于经过较多的卷积和池化操作, 其空间分辨率较低, 导致局部细节信息的损失, 特别是在检测小型缺陷(如气孔、裂纹等)时, 这种信息丢失会降低检测的精确性。此外, 深层特征往往更加偏向于 ImageNet 预训练任务中的高层语义特征, 而这些特征主要服务于自然图像分类任务, 与工业电阻点焊异常检测任务的匹配度较低。相反, Res2、Res3 层的特征保留了更多局部结构信息, 能够更好地捕获小尺度焊点的局部异常模式, 从而提升异常检测的表现。因此选取 ResNet50 提取出的 Res2、Res3 层特征进行后续的特征库的构建与多尺度特征匹配。

首先, 利用候选区域生成网络预测的候选框与高分辨率多尺度特征图(Res2、Res3 层的特征)作为输入, 通过 ROI Align 精确提取候选区域内的特征, 并将其映射为固定尺寸的特征表示。

假设预测的候选框在原图坐标系中的坐标为 (x_1, y_1, x_2, y_2) , 原图像大小为 (H, W) , 对应特征图的大小为 (H', W') 。首先将坐标映射至特征图尺度上, 即:

$$\begin{cases} x'_1 = x_1 \times \frac{W'}{W} \\ y'_1 = y_1 \times \frac{H'}{H} \\ x'_2 = x_2 \times \frac{W'}{W} \\ y'_2 = y_2 \times \frac{H'}{H} \end{cases} \quad (12)$$

将映射后的 ROI 区域划分为固定尺寸的 $K \times K$ 网格,每个网格的宽度和高度为:

$$\begin{aligned} \text{bin_width} &= \frac{x'_2 - x'_1}{K} \\ \text{bin_height} &= \frac{y'_2 - y'_1}{K} \end{aligned} \quad (13)$$

对于每个网格,再进行双线性插值采样,设每个网格内采样点位置为 (x, y) ,与其相邻的特征图像素点位置分别为: (x_l, y_l) 、 (x_r, y_r) 、 (x_l, y_b) 和 (x_r, y_b) ,对应特征值分别为 $f(x_l, y_l)$ 、 $f(x_r, y_r)$ 、 $f(x_l, y_b)$ 和 $f(x_r, y_b)$ 则该采样点 (x, y) 处的插值特征值由双线性插值公式计算,即:

$$\begin{aligned} f(x, y) &= \frac{f(x_l, y_l)(x_r - x)(y_b - y)}{(x_r - x_l)(y_b - y_l)} + \\ &\frac{f(x_r, y_l)(x - x_l)(y_b - y)}{(x_r - x_l)(y_b - y_l)} + \frac{f(x_l, y_b)(x_r - x)(y - y_l)}{(x_r - x_l)(y_b - y_l)} + \\ &\frac{f(x_r, y_b)(x - x_l)(y - y_l)}{(x_r - x_l)(y_b - y_l)} \end{aligned} \quad (14)$$

式中: (x_l, y_l) 、 (x_r, y_r) 、 (x_l, y_b) 和 (x_r, y_b) 为特征图上邻近的4个点坐标; $f(\cdot, \cdot)$ 表示对应位置处的特征值。最后,每个网格的输出特征通过对其中所有采样点的特征值取平均,得到固定尺寸的特征输出。

在完成 RoI Align 后,进一步对预测框区域的特征进行平均池化处理,以确保存入特征库的特征具有更大的视野范围,提高对于小范围偏差的鲁棒性。因此,定义操作 $f_{u,v}[c]$,用于对 RoI Align 提取的特征进行平均池化。

$$f_{u,v}[c] = \frac{1}{p^2} \sum_{a=0}^{p-1} \sum_{b=0}^{p-1} \mathbf{X}_c(u+a, v+b) \quad (15)$$

式中: $\mathbf{X}_c(i, j)$ 是特征图 \mathbf{X} 在通道 c 处的像素值; (u, v) 是特征图中的像素坐标; p 是平均池化窗口大小,设置为3。对于 RoI Align 提取特征假设为深度 C^* 、高度 H^* 、宽度 W^* 的三维张量集合进行上述操作,得到预测框对应的局部特征集 $\mathbf{P}_p(\mathbf{F}_{ROI})$,即:

$$\mathbf{P}_p(\mathbf{F}_{ROI}) = \left\{ f_{u,v}[c] \mid \begin{cases} 0 \leq u \leq H^* \\ 0 \leq v \leq W^* \\ 0 \leq c \leq C^* \end{cases} \right\} \quad (16)$$

在训练阶段,仅使用训练集中正常焊点的标注框通过 ROI Align 提取区域特征,形成固定维度的特征向量。然而,直接存储所有焊点区域的特征可能会导致存储开销过大,并显著增加后续最近邻搜索的计算复杂度。因此,采用核心子集选择策略,在保证特征空间覆盖范围的

同时,减少存储需求并提高计算效率。具体而言,从大量正常焊点特征中选取一个代表性子集,以保证该子集能够最大程度地近似原始特征分布。该核心子集选择问题可以形式化表示为:

$$\mathbf{M}_C = \underset{\mathbf{M}_C \subseteq \mathbf{M}_B}{\text{argmin}} \max_{\mathbf{m} \in \mathbf{M}_B} \min_{\mathbf{n} \in \mathbf{M}_C} \|\mathbf{m} - \mathbf{n}\|_2 \quad (17)$$

这一优化目标的含义是:在给定的原始特征集合 \mathbf{M}_B 中,选择一个子集 \mathbf{M}_C ,使得 \mathbf{M}_C 能够尽可能地覆盖整个特征空间,即保证每个原始特征点 \mathbf{m} 到最近的子集特征点 \mathbf{n} 之间的欧式距离最小化。由于该优化问题是一个非确定性多项式时间难 (non-deterministic polynomial-time hard, NP-hard) 问题,在 Sener 等^[23]提出的核心集选择算法的基础上进行改进,根据 Johnson-Lindenstrauss 定理:对于 $\mathbf{P}_p(\mathbf{F}_{ROI})$ 特征集中的每一个 d 维特征,存在线性映射 $f: \mathbb{R}^d \rightarrow \mathbb{R}^k$,其中 $d > k$,对于任意的两个特征 $\mathbf{m}, \mathbf{n} \in \mathbf{P}_p(\mathbf{F}_{ROI})$,式(18)~(19)成立。

$$k = \frac{c \log \dim(\mathbf{m})}{\varepsilon^2} \quad (18)$$

$$(1 + \varepsilon) \|\mathbf{m} - \mathbf{n}\|_2^2 \leq \|f(\mathbf{m}) - f(\mathbf{n})\|_2^2 \leq (1 - \varepsilon) \|\mathbf{m} - \mathbf{n}\|_2^2 \quad (19)$$

式中: $\dim(\mathbf{m})$ 是原始特征点 \mathbf{m} 的维度; ε 是允许的误差容忍度,设置为0.1; $f(\cdot)$ 在所提方法中使用随机线性投影。将高维特征降维到低维空间时,特征之间的相对距离变化不会超过设定的误差容忍度,从而保持数据的几何结构的同时进一步缩减训练时间。

该算法将核心集选择问题转化为 k-Center 问题,并通过贪心算法提供近似解,在贪心算法的复杂度分析中,该算法的时间复杂度为 $O(n \cdot b)$,其中 n 为数据点的数量, b 为选择的样本数。这表明,算法的计算量与数据规模呈线性关系,适合处理大规模数据集;在收敛性分析方面, Sener 等理论证明该贪心算法能够在一定条件下收敛到一个接近最优解的近似值,且通过迭代过程接近 2-OPT (2-Optimization) 解,从而保证算法的有效性和高效性。算法的伪代码如算法1所示。

算法1:特征库核心子集轻量化算法

Require: ROI Align 提取的特征 \mathbf{F}_{ROI} , 核心子集数量 l , 随机线性投影 ψ

Ensure: 轻量化特征库 \mathbf{M}_B

```

1:  $\mathbf{M}_B \leftarrow \emptyset$ 
2: for each image  $x_i$  do
3:    $\mathbf{M}_B \leftarrow \mathbf{M}_B \cup \mathbf{P}_p(\mathbf{F}_{ROI}(x_i))$ 
4: end for
5: /* 应用贪心算法进行核心子集选择 */
6:  $\mathbf{M}_C \leftarrow \emptyset$ 
7: for  $i = 0$  to  $l - 1$  do
8:    $m_i \leftarrow \arg \max_{\mathbf{m} \in \mathbf{M}_B - \mathbf{M}_C} \min_{\mathbf{n} \in \mathbf{M}_C} \|\psi(\mathbf{m}) - \psi(\mathbf{n})\|_2$ 
9:    $\mathbf{M}_C \leftarrow \mathbf{M}_C \cup \{m_i\}$ 
10: end for
11:  $\mathbf{M}_B \leftarrow \mathbf{M}_C$ 
```

利用核心子集采样构建的正常特征库,确保在推理阶段能够高效地进行异常检测,并提高系统的实时性和鲁棒性。

1.4 异常评分计算

在推理阶段,提出一种基于多尺度特征匹配的像素级异常评分计算方法,利用构建的特征库 M_B 对测试样本进行像素级异常评分计算。具体步骤为:

1) 候选框提取:利用训练好的候选区域生成网络获取预测框。

2) ROI Align:在预测框区域内提取 Res2、Res3 特征,并进行平均池化后拼接 (concat) 得到最终特征表示。

3) 最近邻匹配:计算提取的多尺度特征与相匹配的特征库特征的最邻近距离,以确定异常程度。

对于给定的测试图像 x_i , 经过训练好的目标检测模块得到预测框,然后通过 RoI Align 精确截取候选框对应区域内的 Res2、Res3 特征 $F_{ROI}(x_i)$, 对该特征进行平均池化得到 $P_p(F_{ROI})$ 然后通过对多尺度特征进行拼接 (concat) 得到测试样本特征 P^{test} 。

对于测试样本特征 P^{test} , 在特征库 M 中找到与之匹配最邻近的特征 P^* , 即:

$$P^{test*}, P^* = \underset{P^{test} \in P_p(F_{ROI})}{\operatorname{argmax}} \underset{P \in M_B}{\operatorname{argmin}} \|P^{test} - P\|_2 \quad (20)$$

式中: argmin 是寻找特征库中与 P^{test} 最相似的特征 P^* ; argmax 是在 P^{test} 中寻找与特征库中最相似特征中距离最大的特征 P^{test*} ; 异常评分定义为测试特征与其最邻近记忆库特征的欧式距离,即:

$$s^* = \|P^{test*} - P^*\|_2 \quad (21)$$

为了更稳定地反映异常程度,进一步引入了局部邻域归一化策略。考虑到某些测试特征的最近邻特征可能具有较大距离,而这些特征在特征库中的分布较为稀疏,因此采取了该策略以增强模型对稀有特征的鲁棒性,因此对 s^* 进行加权修正,即:

$$s = \left(1 - \frac{\log(1 + \|P^{test*} - P^*\|_2)}{\sum_{P \in N(P^*)} \log(1 + \|P^{test*} - P\|_2)} \right) \cdot s^* \quad (22)$$

式中: $N(P^*)$ 为特征库 M_B 中的 n 个最近的特征 (所提方法中设置为 9), 由此得到每个特征的异常评分后,再通过同式 (1) 中的双线性插值将特征异常分数图的分辨率放大至与原图像相同,得到图像的像素级异常评分。为了将像素级的异常评分标准化到统一的范围,使用最大最小归一化对异常分数进行处理。最大最小归一化的公式为:

$$S_{i,j}^* = \frac{S_{i,j} - \min(S)}{\max(S) - \min(S)} \quad (23)$$

式中: $S_{i,j}$ 是原始的异常分数,在位置 (i,j) 处的异常值; $\min(S)$ 和 $\max(S)$ 分别表示图像异常分数的最小值和最

大值; $S_{i,j}^*$ 是归一化后的异常分数,归一化后的分数范围在 $[0, 1]$ 之间。

在得到归一化后的异常分数后,通过设定一个阈值 τ 来进行缺陷分割。具体来说,对于每个像素点,如果其归一化后的异常分数 $S_{i,j}^*$ 大于设定的阈值 τ ,则认为该位置为缺陷区域,否则认为是正常区域,所提方法中 τ 设定为 0.5。分割公式为:

$$D_{i,j} = \begin{cases} 1, & S_{i,j}^* \geq \tau \\ 0, & S_{i,j}^* < \tau \end{cases} \quad (24)$$

式中: $D_{i,j}$ 是分割结果,表示在位置 (i,j) 处是否属于缺陷区域; $D_{i,j} = 1$ 表示该位置是缺陷区域; $D_{i,j} = 0$ 表示该位置是正常区域,将异常分数转化为二值图像,实现焊点缺陷的判断的同时,实现缺陷区域的提取和分割。

2 实 验

2.1 数据集构建

在国家重点研发项目“国产机器人汽车焊装生产线关键技术与应用示范”的支持下,在汽车装配流水线上采集汽车白车身不同位置的电阻点焊数据图像,电阻点焊缺陷检测数据集采集现场如图 2 所示。



(a) 白车身样本
(a) Body-in-white sample



(b) 工业相机采集装置
(b) Industrial camera setup

图 2 电阻点焊缺陷检测数据集采集现场

Fig. 2 On-site data acquisition setup for resistance spot welding defect detection

使用工业相机拍摄,单张图像分辨率为 4 560 pixels×3 024 pixels,工业相机提供较高的分辨率与清晰度,能够详细捕捉焊点区域的细节。拍摄过程中,采用稳定光源并调整相机角度,确保覆盖焊点的不同视角,数据采集过程覆盖了多个工作站和生产流程,旨在模拟实际生产环境中的焊点检测任务,并且通过变换拍摄角度和位置,增强了数据的多样性,减少了相同视角对模型性能的依赖,从而提高了模型在实际应用中的鲁棒性。

实验共采集了 2 000 张图像,电阻点焊数据集示意图如图 3 所示,涵盖多种焊点类型和缺陷类别,如虚焊、连焊、少锡、气孔等。数据集中所有焊点区域均由人工标注,包含位置框及缺陷的语义分割。其中,80% 图像用于训练,20% 用于测试,其中训练集包含有正常焊点与多种缺陷焊点,在特征库的构建过程中仅选取其中正常焊点参与构建。



图 3 电阻点焊数据集示意图

Fig. 3 Schematic diagram of the weld point dataset

2.2 评价指标

为了全面评估所提方法与传统目标检测方法在小目标焊点缺陷检测中的性能,对比了正负样本分类的精度,即模型对正常焊点和缺陷焊点的正确分类能力,并且对于电阻点焊缺陷分割进行评估。为了准确反映分类精度与缺陷分割精度,因此,采用了以下 3 个主要的分类评价指标与 2 个分割评价指标。

精确度 (Precision): 用于评估模型预测为缺陷焊点 (正样本) 中,实际为缺陷焊点的比例。其计算公式为:

$$Precision = \frac{TP}{TP + FP} \quad (25)$$

式中: TP 为模型正确预测为缺陷焊点的数量; FP 为模型错误预测为缺陷焊点但实际上是正常焊点的数量,精确度越高,表示模型预测的缺陷焊点更准确,误报 (即将正常焊点误分类为缺陷) 越少。

召回率 (Recall): 用于评估模型对所有实际缺陷焊点的识别能力。其计算公式为:

$$Recall = \frac{TP}{TP + FN} \quad (26)$$

式中: 假负例 (false negative, FN) 为模型未能检测到的缺陷焊点,实际为缺陷但被模型误分类为正常,召回率越高,表示模型漏检的缺陷焊点越少,检测能力更强。

$F1$ 值 ($F1$ -Score): $F1$ 值是精确度和召回率的调和平均值,用于综合评价模型的检测能力。计算公式为:

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (27)$$

像素级 AUROC (pixel-level area under the receiver operating characteristic curve, Pixel AUROC): Pixel AUROC 衡量模型区分缺陷与非缺陷区域的能力,尤其适用于像素级别的分类任务。通过计算模型在不同决策阈值下的假阳性率 (false positive rate, FPR) 与真正率 (true positive rate, TPR) 的关系,得出一个综合评分。Pixel AUROC 的值越接近 1,表示模型的区分能力越强。其计算公式为:

$$Pix AUROC = \int_0^1 TPR(t) d(FPR(t)) \quad (28)$$

式中: $TPR(t)$ 为真正率,即正确预测为缺陷的像素占有所有实际缺陷像素的比例; $FPR(t)$ 为假阳性率,即错误预测为缺陷的像素占有所有实际非缺陷像素的比例。积分过程考虑了模型在所有可能的决策阈值下的表现。

Pixel $F1$ Score 用于衡量模型在缺陷区域分割任务中的精度与召回率之间的平衡,专注于像素级别的分割效果。该指标结合了模型在精度和召回率方面的表现,提供了一个综合的评估标准。Pixel $F1$ Score 值越高,表示模型在缺陷区域分割的精度和召回率之间达到了更好的平衡,分割效果越优。

2.3 实验结果

实验在配备 NVIDIA RTX 3090 显卡的计算机上进行,利用 PyTorch 框架完成深度学习模型的训练和测试。实验中的数据集被划分为训练集 (1 600 张图像)、验证集 (200 张图像) 和测试集 (200 张图像),确保了模型的评估与优化能够覆盖多种实际场景,为了增强模型的泛化能力,在训练过程中对训练集应用了数据增强技术,包括平移、旋转、镜像、拼接和增加噪声等方式。

为了更直观地展示 DBL 各参数的作用,仅在候选区域生成网络中进行测试,并通过调整不同参数设置,测试候选区域生成网络的定位精度,探讨 DBL 各参数对模型性能的影响。评估了 3 种配置:配置 1 使用 3×3 卷积核、Leaky ReLU 激活函数和批归一化 (默认配置);配置 2 使用 1×1 卷积核、Leaky ReLU 激活函数和批归一化;配置 3 使用 3×3 卷积核、ReLU 激活函数和批归一化。实验结果如表 1 所示,默认配置 (配置 1) 在精度、召回率和 $F1$ 分数上表现最好,表明使用 3×3 卷积核和 Leaky ReLU 激活函数能够显著提升模型性能。相比之下,配置 2 虽然使用 1×1 卷积核减少了计算开销,但在精度和 $F1$ 分数

上有所下降;配置 3 虽然在召回率上与配置 1 相差不大,但精度略低。因此,默认配置下的性能最佳。

表 1 不同 DBL 结构配置对模型性能的影响
Table 1 Impact of different DBL configuration on model performance

配置	卷积核	激活函数	精确度	召回率	F1 分数
1	3×3	Leaky ReLU	0.95	0.97	0.96
2	1×1	Leaky ReLU	0.93	0.96	0.94
3	3×3	ReLU	0.92	0.97	0.94

在 DBL 配置 1 的条件下,在汽车现场电阻点焊数据集上进行了实验,评估了所提方法在电阻点焊分类精度、缺陷区域分割精度的表现。数据量为 700 张训练图像、150 张验证图像和 150 张测试图像。为了验证所提方法的有效性,对不同模型在焊点缺陷检测任务中的表现进行了对比。对于每个模型,记录了其在测试集上的精确度、召回率和 F1 分数。同时,还记录了每个模型的推理时间,以便在速度上进行对比。实验结果如表 2 所示。所提方法在自制焊点数据集上的检测效果如图 4 所示,bad 表示缺陷的电阻点焊,good 表示正常的电阻点焊,所有点焊区域都被准确框选并给出了像素级的异常评分。

表 2 所提方法与目标检测典型模型比较结果
Table 2 Comparison results of proposed model with typical models

模型	推理速度/s	精确度	召回率	F1 分数
Tiny-YOLOv3	0.424	0.60	0.81	0.69
YOLOv3	0.759	0.65	0.91	0.76
YOLOv5	0.730	0.66	0.94	0.78
YOLOv8	0.712	0.71	0.95	0.81
本文方法	0.785	0.89	0.97	0.93

实验结果表明,所提方法在精确度召回率和 F1 分数上优于目前主流的目标检测算法,尤其在精确度和召回率之间实现了较好的平衡。模型在精确度上达到 0.89,在召回率上达到 0.97,F1 分数为 0.93,相较于 YOLOv8 模型,精确度提升 25.35%,F1 分数提升 14.81%,表现出了较为明显的优势。

此外,所提方法在像素级缺陷分割方面也展现了优异的性能。实验结果表明,所提方法的 Pixel AUROC 达到了 0.94,显示了模型在区分缺陷与非缺陷区域的优越能力。与此同时,Pixel F1 Score 达到 0.79,表明在缺陷分割的精度和召回率之间达到了良好的平衡,确保了高

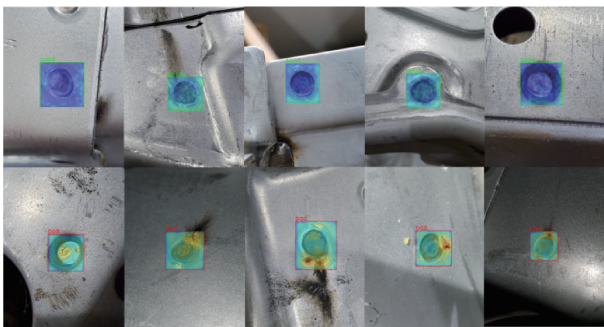


图 4 所提方法检测结果
Fig. 4 Detection results of the proposed model

精度的缺陷定位和分割。同时,该方法也保有较快的推理速度,能够满足工业的实时性检测。

如图 5 所示,图 5(a)和(b)分别为 YOLOv5、YOLOv8 的检测结果的局部放大图,图 5(c)为所提方法检测结果的局部放大图,图 5(d)为相应的缺陷的真值图。由于电阻点焊缺陷占比往往较小,传统的目标检测难以捕捉细小缺陷,YOLOv5 和 YOLOv8 在检测小型缺陷时常出现误判,将其错误地识别为正常焊点,而所提方法成功检测出这些小型缺陷,表现出了更强的检测能力。真值图用于验证缺陷分割结果的准确性,局部放大图展示了缺陷的细节,进一步证明了该方法在缺陷定位和分割上的高精度。

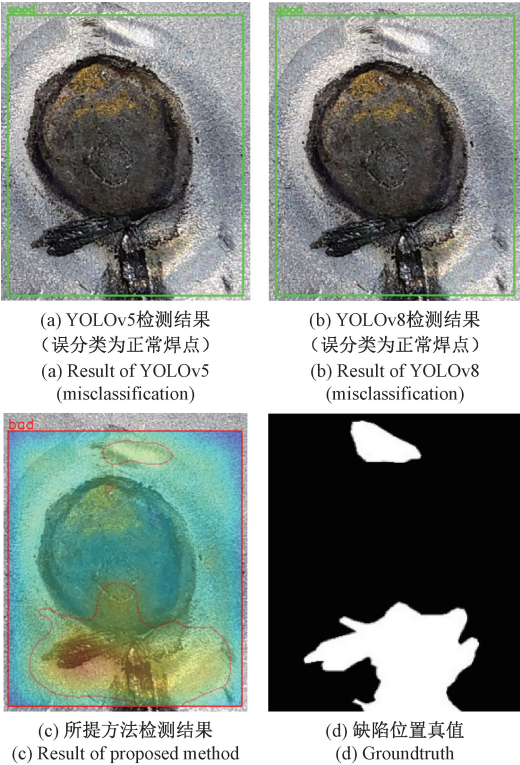


图 5 检测结果局部放大对比
Fig. 5 Comparison of detection results

为了评估模型性能与数据量之间的关系,通过逐步增加训练数据量进行多次实验。在实验中,从少数数据量(300 张图像)逐步增加到高数据量(1 600 张图像),并评估数据量变化对目标区域生成精度、缺陷区域分割精度以及分类准确度的影响。实验结果如表 3 所示,随着训练数据量的增加,模型在分类精度和召回率上逐步提高,并且在数据量达到 700 张图像后,各项指标基本趋于饱和。同时,Pixel AUROC 和 Pixel F1 Score 这两个分割精度指标也显著提升,表明随着训练样本量的增加,缺陷分割与定位网络在焊点区域分割能力上的表现也在数据量达到 700 后趋于平稳。更大规模的数据集使得模型在处理复杂背景和细微差异时更加精准和鲁棒,但在数据量达到一定程度后,进一步增加数据量对性能的提升效果逐渐减小。然而,更大规模的数据集也会导致训练时间的显著增加,给模型的训练和优化带来更高的计算成本。

表 3 不同训练数据量对模型性能的影响

Table 3 Impact of different training data sizes on model performance

训练数据 量/张	像素级 AUROC	像素级 F1 分数	精确度	召回率	F1 分数
300	0. 87	0. 74	0. 84	0. 90	0. 86
500	0. 91	0. 76	0. 88	0. 95	0. 92
700	0. 94	0. 79	0. 89	0. 97	0. 93
900	0. 94	0. 80	0. 89	0. 97	0. 93
1 200	0. 95	0. 80	0. 91	0. 98	0. 94
1 600	0. 96	0. 81	0. 92	0. 98	0. 95

为了进一步验证模型的泛化能力,在汽车现场电阻点焊 700 张图像数据集上训练模型的基础上,在一份包含不同工业场景下的开源电阻点焊数据集(resistance spot welding dataset, RSW Dataset)上进行测试。该数据集图像均包含多个焊点,涵盖了多种工业场景,包括:不同工业场景目标的点焊图像(如重型设备框架、运输工具部件、钢结构支架、机械装配部件等)、多种类型的点焊缺陷(如孔洞、裂纹等)、复杂背景下的点焊图像(如车间环境背景)、以及多种光照条件下的点焊图像(如强光、低光等),旨在模拟真实工业环境中的多样化挑战,数据集涵盖不同工业场景示意图如图 6 所示。

从该数据集中选取了 200 张图像进行测试,以评估所提出的模型在复杂工业环境中的表现,并验证其在不同场景下的泛化能力和鲁棒性。实验结果如表 4 所示,传统的目标检测模型在新数据集上的性能普遍下降,主要由于这些模型是基于特定数据集训练的,因而对训练集之外的缺陷分布缺乏有效的识别能力。与之相比,所



图 6 不同工业场景的电阻点焊

Fig. 6 Resistance spot welding images from different industrial scenarios

提方法通过基于多尺度特征匹配的方法对电阻点焊进行检测,有效提升了模型的泛化能力。同时,所提方法在缺陷语义分割任务中也表现出色。具体而言,模型在 Pixel AUROC 和 Pixel F1 Score 上分别达到了 0. 95 和 0. 83,进一步验证了该方法在精确缺陷定位与分割方面的优越性能及其良好的泛化能力。该方法在开源电阻点焊数据集

上的测试结果如图 7 所示,进一步证明了其在不同工业场景中的有效性。

表 4 开源数据集模型性能对比结果

Table 4 Comparison of model performance on the open-source dataset			
模型	精确度	召回率	F1 分数
Tiny-YOLOv3	0.57	0.76	0.65
YOLOv3	0.61	0.82	0.70
YOLOv5	0.64	0.89	0.74
YOLOv8	0.69	0.93	0.79
本文方法	0.90	0.96	0.93

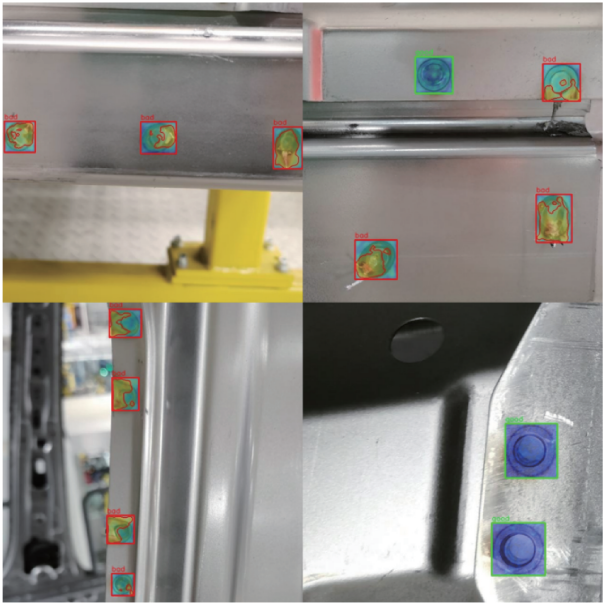


图 7 电阻点焊数据集测试结果

Fig. 7 Detection results of the RSW dataset

3 结 论

本研究提出了一种基于多尺度特征融合网络的电阻点焊缺陷精确定位方法,所提方法针对电阻点焊设计多尺度特征的跨层次连接以及候选区域训练的损失函数,实现电阻点焊区域的精准定位。然后通过基于多尺度特征匹配的缺陷分割与定位网络,构造正常电阻点焊特征库,生成电阻点焊区域的像素级异常评分,有效克服了传统目标检测中无法对细小缺陷区域语义分割以及分类精度低的问题。实验结果表明,所提方法在保证较快的推理速度的同时,不仅实现了小目标电阻点焊缺陷的语义分割,在电阻点焊的分类检测精度方面明显优于传统的目标检测网络,同时在不同工业场景下也表现出较高的

鲁棒性与泛化能力。
特征库的构建和最近邻匹配计算可能导致存储和计算开销增加,未来可通过优化特征存储和匹配算法提高效率。此外,尽管方法具有良好的实时性,处理高分辨率图像时仍可能面临计算瓶颈,未来可以通过优化网络结构或采用更高效的计算框架进一步提升性能。

参考文献

[1] 刘传洋,吴一全,刘景景. 基于视觉的输电线路金具锈蚀缺陷检测方法研究进展[J]. 仪器仪表学报, 2024, 45(3): 286-305.
LIU CH Y, WU Y Q, LIU J J. Research progress of vision-based rust defect detection methods for metal fittings in transmission lines [J]. Chinese Journal of Scientific Instrument, 2024, 45(3):286-305.

[2] 郭玥,于希明,王少军,等. 遥感图像云检测的多尺度融合分割网络方法[J]. 仪器仪表学报, 2019, 40(6): 31-38.
GUO Y, YU X M, WANG SH J, et al. Cloud detection in remote sensing images with mutilevel scale fused network[J]. Chinese Journal of Scientific Instrument, 2019,40(6):31-38.

[3] ALGHANNAM E, LU H, MA J Y, et al. A novel method of using vision system and fuzzy logic for quality estimation of resistance spot welding [J]. Symmetry, 2019, 11(8): 990.

[4] 张瀚丹,吴一全. 基于视觉的汽车装配件缺陷检测研究进展[J]. 仪器仪表学报, 2023, 44(8): 1-20.
ZHANG H D, WU Y Q. Research progress of vehicle assembly defect detection methods based on vision[J]. Chinese Journal of Scientific Instrument, 2023, 44(8): 1-20.

[5] 赵朗月,吴一全. 基于机器视觉的表面缺陷检测方法研究进展[J]. 仪器仪表学报, 2022, 43(1): 198-219.
ZHAO L Y, WU Y Q. Research progress of surface defect detection methods based on machine vision [J]. Chinese Journal of Scientific Instrument, 2022, 43(1): 198-219.

[6] LIU J Q, XIE G Y, WANG J B, et al. Deep industrial image anomaly detection: A survey [J]. Machine Intelligence Research, 2024, 21(1): 104-135.

[7] 张小俊,王志鹏,马秋实,等. 一种基于深度学习的图像语义分割方法[J]. 机械设计, 2021, 38(3): 16-23.
ZHANG X J, WANG ZH P, MA Q SH, et al. Method of image semantic segmentation based on deep learning[J]. Journal of Machine Design, 2021, 38(3): 16-23.

- [8] TAO X, GONG X Y, ZHANG X, et al. Deep learning for unsupervised anomaly localization in industrial images: A survey [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 3196436.
- [9] 赵恩玄,何云勇,沈宽,等. 基于深度学习的铸件 CT 图像分割算法[J]. 仪器仪表学报, 2023, 44(11): 176-184.
- ZHAO EN X, HE Y Y, SHEN K, et al. Casting CT image segmentation algorithm based on deep learning[J]. Chinese Journal of Scientific Instrument, 2023, 44(11): 176-184.
- [10] CUI Y J, LIU ZH X, LIAN SH G. A survey on unsupervised anomaly detection algorithms for industrial images[J]. IEEE Access, 2023, 11: 55297-55315.
- [11] ZUO Z, DONG J H, GAO Y, et al. HyperDefect-YOLO: Enhance YOLO with hyper graph computation for industrial defect detection [J]. ArXiv preprint arXiv: 2412.03969, 2024.
- [12] QI ZH, DING L W, LI X T, et al. Detecting and classifying defective products in images using YOLO[J]. ArXiv preprint arXiv:2412.16935, 2024.
- [13] WANG J Y, YU N G. SSD-faster net: A hybrid network for industrial defect inspection[J]. ArXiv preprint arXiv: 2207.00589, 2022.
- [14] BAUR C, WIESTLER B, ALBARQOUNI S, et al. Deep autoencoding models for unsupervised anomaly segmentation in brain MR images [C]. Brainlesion: Part I, 2019: 161-169.
- [15] AKÇAY S, ATAPOUR-ABARGHOUEI A, BRECKON T P. Skip-ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection [C]. 2019 International Joint Conference on Neural Networks, 2019: 1-8.
- [16] PIRNAY J, CHAI K. Inpainting transformer for anomaly detection[C]. International Conference on Image Analysis and Processing, 2022: 394-406.
- [17] LU F B, YAO X F, FU CH W, et al. Removing anomalies as noises for industrial defect localization[C]. 2023 IEEE/CVF International Conference on Computer Vision, 2023: 16120-16129.
- [18] LI CH L, SOHN K, YOON J, et al. Cutpaste: Self-supervised learning for anomaly detection and localization[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 9659-9669.
- [19] ZHANG X M, XU M, ZHOU X ZH. Realnet: A feature selection network with realistic synthetic anomaly for anomaly detection[C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16699-16708.
- [20] ZAVRTANIK V, KRISTAN M, SKOCAJ D. Draem-a discriminatively trained reconstruction embedding for surface anomaly detection[C]. 2021 IEEE/CVF International Conference on Computer Vision, 2021: 8330-8339.
- [21] DEFARD T, SETKOV A, LOESCH A, et al. Padim: A patch distribution modeling framework for anomaly detection and localization[C]. International Conference on Pattern Recognition, 2021: 475-489.
- [22] ROTH K, PEMULA L, ZEPEDA J, et al. Towards total recall in industrial anomaly detection[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 14298-14308.
- [23] SENER O, SAVARESE S. Active learning for convolutional neural networks: A core-set approach[J]. ArXiv preprint arXiv:1708.00489, 2017.

作者简介



王俞霏,2023 年于天津大学获得学士学位,现为天津大学硕士生,主要研究方向为计算机视觉和深度学习。

E-mail: yufei_w@tju.edu.cn



Wang Yufei received his B.Sc. degree from Tianjin University in 2023. Now he is currently a M.Sc. candidate at Tianjin University. His main research interests include computer vision and deep learning.

孙岩标(通信作者),2010 年于武汉大学获得学士学位,2015 年于北京大学获得博士学位,现为天津大学副教授,主要研究方向为大尺寸视觉精密坐标测量。

E-mail: yanbiao.sun@tju.edu.cn

Sun Yanbiao (Corresponding author) received his B.Sc. degree from Wuhan University in 2010, received his Ph.D. degree from Peking University in 2015. Now he is an associate professor at Tianjin University. His main research interest includes large-scale visual precision coordinate measurement.