

DOI: 10.19650/j.cnki.cjsi.J2413557

基于特征表征与学习反馈的动态带钢缺陷样本筛选方法*

苑玮琦, 刘文滔, 李绍丽

(沈阳工业大学信息科学与工程学院 沈阳 110870)

摘要:带钢表面缺陷检测是保证钢铁产品质量的关键环节,实现高效准确的缺陷检测对保障产品性能具有重要意义。近年来,深度学习方法在缺陷检测领域进展显著,但在实际应用中仍面临两个问题:一方面,由于工业生产追求高良品率,导致缺陷样本获取受限,且样本标注耗时费力;另一方面,采集的样本中可能存在冗余特征,影响模型训练效率和泛化性能。针对特征冗余问题,提出一种基于特征表征与学习反馈机制的动态样本筛选方法。首先构建包含几何形态、灰度分布及方向特征等多维特征量化模型,系统表征缺陷特征。随后,设计基于特征表征的样本筛选策略,结合特征聚类快速筛选少量具有多样性和代表性的训练样本。最后,设计基于置信度评估的动态优化策略,通过模型的学习反馈获取关键补充样本,提升特征覆盖范围,实现训练样本的自适应优化。NEU-DET数据集的实验结果表明,该方法在将训练样本数量减少52%的情况下,平均检测精度达到76.99%,与完整数据集基本持平。同时,每轮训练迭代时间减少62%,降低了计算开销,验证了方法在样本筛选与检测性能之间的有效平衡。此外,在多种主流目标检测模型上的验证结果表明,该方法在不同检测架构下均能有效提升效率并保持性能,展现出良好的适用性。

关键词:带钢表面缺陷检测;样本筛选;特征表征;形态学特征;深度学习

中图分类号: TP391.41 TH165 **文献标识码:** A **国家标准学科分类代码:** 510.40

A dynamic sample selection method for steel strip defects based on feature representation and learning feedback

Yuan Weiqi, Liu Wentao, Li Shaoli

(College of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China)

Abstract: Surface defect detection of steel strips is crucial for ensuring product quality in steel manufacturing. The achievement of efficient and accurate detection is significant for product performance. While deep learning methods have made significant progress in defect detection, two challenges persist in practical applications: First, due to the pursuit of high yield rates in industrial production, defect samples are limited and sample annotation is time-consuming. Secondly, the collected samples may contain redundant features, affecting model training efficiency and generalization performance. To address feature redundancy, a dynamic sample selection method based on feature representation and learning feedback is proposed. Initially, a multi-dimensional feature quantification model incorporating geometric morphology, grayscale distribution, and directional features is formulated to characterize defect features. Subsequently, a feature representation-based sample selection strategy is designed to select diverse and representative training samples through feature clustering. Finally, a confidence-based dynamic optimization strategy is proposed to obtain supplementary samples through learning feedback, achieving adaptive optimization. Experimental results on the NEU-DET dataset show that the method achieves a mean average precision of 76.99% while reducing the training sample size by 52%. It is comparable with using the complete dataset and decreases training iteration time by 62%. Evaluation of various detection models shows that the method effectively improves efficiency while maintaining performance across different architectures.

Keywords: steel strip surface defect detection; sample selection; feature representation; morphological features; deep learning

0 引言

钢铁产业作为国民经济的基础支柱,在现代化建设进程中发挥着不可替代的作用。带钢作为钢铁生产过程中的关键中间产品,其表面质量直接影响下游产品的性能与附加值^[1]。受生产工艺、设备状态等因素影响,带钢表面可能会出现各种缺陷,不仅会降低产品美观度,更可能引发应力集中、裂纹扩展等严重工程问题^[2]。因此,带钢表面缺陷检测是保证产品质量的关键环节^[3]。

带钢表面缺陷检测具有其特殊性。与普通钢材相比,带钢缺陷因连续轧制工艺的影响,常沿轧制方向表现出显著的延展性和规律性。同时受工艺参数,如温度梯度和压力分布等变化的影响,其灰度分布和形态特征更加复杂多样^[4]。这些特点增加了缺陷检测的难度,对检测方法的实时性和鲁棒性提出了更高要求。

由于传统的人工检测方法存在效率低下和主观性强等局限性,难以满足现代化生产的需求^[5-7]。随着自动化技术的发展,缺陷检测方法不断革新,基于机器视觉的检测方法逐渐成为主流^[8-9]。早期检测方法通常依赖人工设计的特征表征方法,尽管在特征表达的可解释性方面具有优势,但其泛化能力有限,难以适应实际复杂工况的要求。

近年来,深度学习在缺陷检测领域取得了重要进展。卷积神经网络^[10]、Transformer 架构^[11]等深度网络可以自动学习缺陷的层次化特征表示,极大提升了检测性能。一些改进方法如特征融合机制^[12-13]、注意力机制^[14-15]等手段进一步增强了网络的特征表达和检测能力。然而,深度学习算法的有效性和泛化能力通常依赖大规模、高质量的训练样本^[16]。这一特性导致其在工业缺陷检测场景中面临2个问题:1)由于工业生产追求产品的高质量和高良品率,正常生产中出现的缺陷样本数量有限,而专业人员的标注过程又极其耗时。2)受检测环境和设备状态等影响,采集的样本质量存在差异,难以保证它们对模型训练的贡献价值,最终影响模型的训练效率和检测性能。

针对样本数量受限和标注成本高的问题,研究人员提出了多种解决方案。一类方法致力于数据扩充,如文献^[17]提出基于生成对抗网络的缺陷生成网络,有效扩充了训练数据。文献^[18]则利用迁移学习提出基于扩散模型的缺陷图像生成方法,通过缺陷控制和背景迁移,将常见缺陷知识迁移到罕见类别,提升少样本检测性能。然而,这类方法的本质是依赖已有标注数据生成新的缺陷样本,合成数据与真实环境的差异可能影响模型实际表现。另一类方法则着眼于降低标注成本,如文献^[19]通过自监督对比学习,优化特征表征,减少了模型对标注数据的依赖。文献^[20]则结合轻量级网络 and 不确定性

采样策略,提出主动小样本学习方法以解决标注样本获取困难的问题。

尽管上述研究取得了重要进展,但样本仍可能面临质量问题。目前工业缺陷检测领域关于数据质量优化的研究相对较少,但通用机器学习领域已有一些启发性的工作。如文献^[21]提出基于影响函数的筛选方法,通过量化样本对模型性能的影响程度筛选出关键数据。文献^[22-23]则分别基于核心集理论提出了主动学习方法,在维持模型性能的同时降低了训练数据量。文献^[24-25]则提出基于元启发式算法的欠采样方法,以应对样本不平衡等问题。这些通用方法为样本筛选提供了重要思路,但未充分考虑工业缺陷的特殊性。带钢表面缺陷检测场景中,样本常常表现出复杂的特征分布。一方面,缺陷特征在形状、灰度分布等方面存在相似性,导致训练样本中可能存在特征冗余。另一方面,缺陷形态又存在多样化且无规律可循。这种“相似”与“差异”共存特性,使通用方法难以直接应用。因此,本研究从缺陷特征表征理论出发,提出了一种融合特征表征与学习反馈机制的动态样本筛选方法。

本研究构建了一个多维特征量化模型,用于全面表征样本间的视觉差异,为样本筛选提供了理论支持。设计了一种基于特征表征的样本筛选策略。结合聚类分析快速筛选少量具备多样性和代表性的训练样本,有效减少冗余样本引入。提出了一种基于模型置信度评估的动态优化机制,通过学习反馈识别特征覆盖不足的区域,动态补充关键样本,提升特征覆盖范围。最终在维持模型性能的同时有效降低了训练样本数量,平衡了工业场景下样本规模与检测性能之间的矛盾。

1 带钢表面缺陷的形态学特征分析

带钢表面缺陷主要源于轧制成形过程、材料特性以及外部机械损伤等多种因素。与普通钢材相比,带钢表面缺陷在形成机理、视觉特征和检测难度等方面均存在差异。在形成机理上,普通钢材多采用离散加工方式,其表面缺陷分布随机且相对独立,而带钢采用连续轧制工艺,受温度梯度和压力分布的动态影响,缺陷沿轧制方向呈现显著的延展性和规律性。

在视觉特征方面,带钢表面缺陷具有3个显著特点:1)缺陷沿轧制方向的延展性和连续性会随速度和压力变化动态改变;2)轧制设备状态与轧辊磨损的影响,使缺陷具有方向性特征;3)带钢表面特殊的光学反射特性,使缺陷的灰度分布与背景纹理的对比更加复杂。这些特点不仅增加了带钢缺陷在复杂纹理背景下的检测难度,并使得不同缺陷在几何形态、灰度分布和纹理方向上同时表现出相似性与差异性。

其中,相似性在数据集中表现为近似甚至重复的特征信息,容易导致特征冗余问题。这种特征冗余不仅增加了数据标注成本,也影响了深度学习模型的训练效率和泛化能力。为提高深度学习模型的训练效率,需筛选代表性样本,避免模型学习冗余特征。首先,通过形态学

分析,全面了解缺陷的几何特征及其类内分布规律,为后续样本筛选策略提供理论依据。

根据视觉特征,将观察到的带钢表面缺陷归纳为块状、线状和纹理状3种基本类型,其典型样本及特征分布如图1所示。

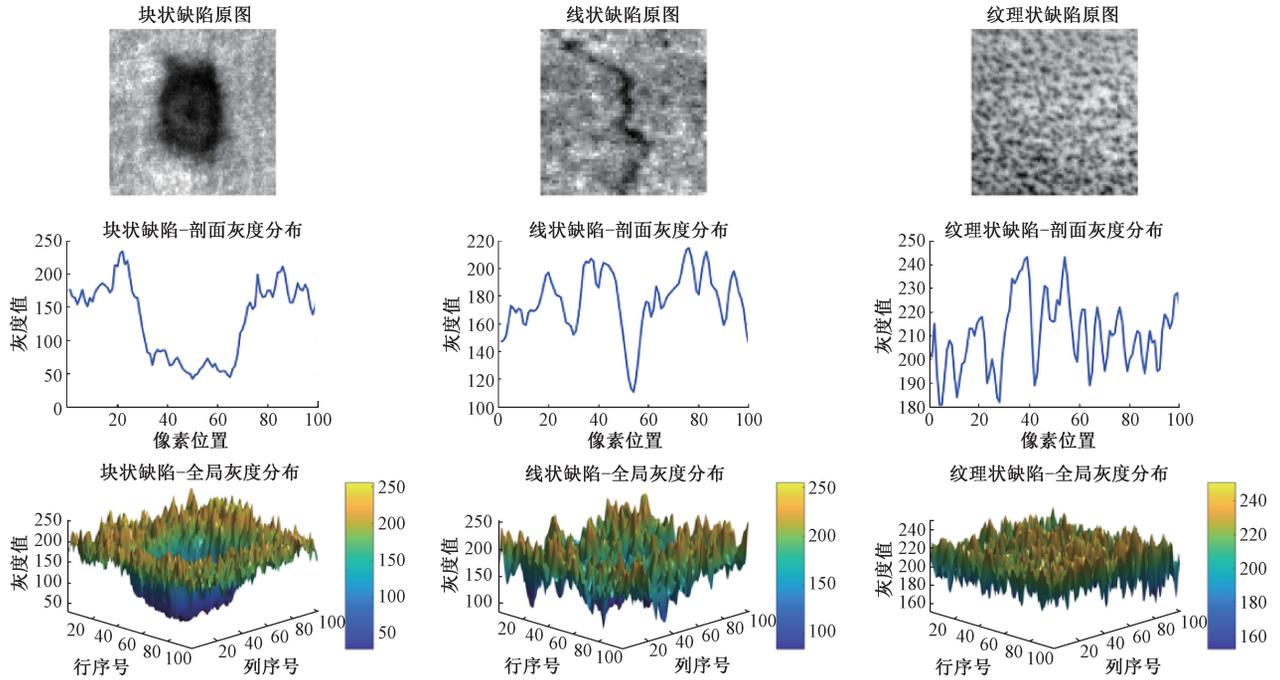


图1 不同形态特征缺陷的代表性图像

Fig. 1 Representative image of defects with different morphological characteristics

块状缺陷表现为长宽比相近的独立区域,其灰度剖面呈现规律的凹陷特征。线状缺陷具有显著的长宽差异,其灰度分布在局部呈现凹陷,但整体上随方向变化形成复杂起伏。纹理状缺陷由密集分布的细小局部点状缺陷构成,其灰度分布表现为高频且不规则的波动特征。为深入分析缺陷的形态相似性,本节对3种基本类型缺

陷的形态特征进行细化分类。

块状缺陷根据其几何形态的复杂度,可分为规则和不规则两类。规则的块状缺陷如图2(a)~(c)所示,其形态可用椭圆、三角形及多边形等基本几何形状来表征。不规则的块状缺陷如图2(d)所示,其形态特征偏离基本几何形状,需要更复杂的轮廓描述方法。

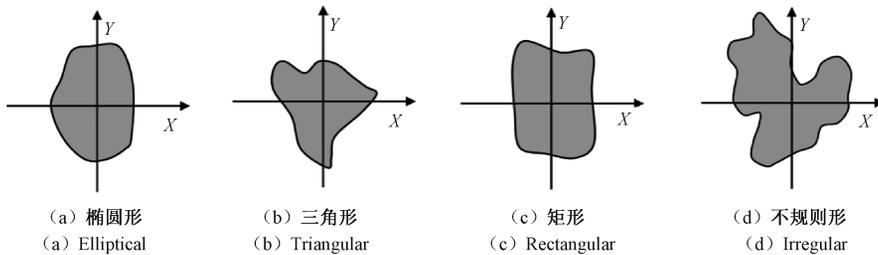


图2 块状缺陷的典型形态

Fig. 2 Typical shape of block defects

线状缺陷根据其特有的明显方向性延伸特征,可分为直线型、曲线型和平行线组合型。如图3所示,直线型缺陷的形态可用沿固定方向延伸的线性结构表征。曲线型缺陷的形态可用具有连续方向变化的线性结构表征。

平行线组合型缺陷则由多条间距小于单线宽度的平行细线构成,整体视觉形态表现为近似单一的线性结构。

纹理状缺陷根据其呈现出的特定的空间分布特征,可分为均匀分布型、聚集分布型和渐变分布型。如图4

所示,均匀分布型的形态可用规则分散的点阵结构表征。聚集分布型缺陷的形态可用局部密集的点簇结构表征。

渐变分布型缺陷的形态表现为点密度在空间上连续变化。

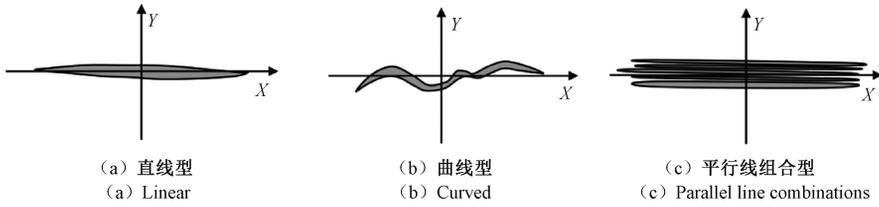


图 3 线状缺陷的典型形态

Fig. 3 Typical shapes of linear defects

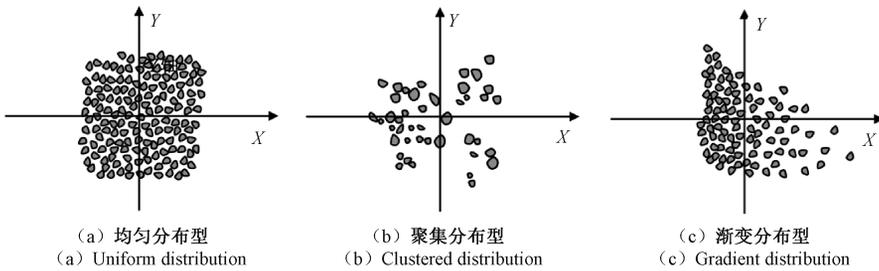


图 4 纹理缺陷的典型形态

Fig. 4 Typical shapes of texture defects

2 缺陷视觉特征的量化模型

通过前述形态特征分析,已明确带钢表面缺陷可基于视觉特征分为块状、线状和纹理状 3 种基本形态类型。由于同一形态类型的缺陷具有明显的视觉相似性,需建立合理的特征量化模型,以评估形态类内样本间的相似程度,从而在各形态类别中筛选代表性样本。本章首先针对 3 种形态类型分别构建形态特征量化指标,随后进一步建立完整的多维特征量化模型。

2.1 形态特征量化模型

本节基于几何学和统计学原理,分别针对块状、线状和纹理状 3 种视觉形态,构建各自对应的形态特征量化指标,以刻画各形态的类内特征差异。每类缺陷仅使用其对应的形态指标进行类内特征量化。

块状缺陷的形态差异主要体现在区域的规则程度上。采用形状规则度 R_s 进行量化,即:

$$R_s = 4\pi \times A_d / P_d^2 \quad (1)$$

其中, A_d 为缺陷区域的面积, P_d 为缺陷轮廓周长。该指标源自圆形度量化理论,以圆形作为最规则的形状基准。经标准化处理, R_s 的取值范围为 $(0, 1]$ 。当缺陷形状为正圆时取得最大值 1,随着形状的不规则程度增加而递减。对于规则的多边形,该值仍保持较高水平,而对于不规则或具有复杂轮廓的缺陷,该值显著降低。

线状缺陷的关键特征体现在其具有明显的线性延展

特性。采用曲直度 C_d 进行量化,即:

$$C_d = L_s / L_e \quad (2)$$

其中, L_s 表示缺陷骨架线累积长度, L_e 表示骨架线两端点间的直线距离。对于标准的线状缺陷, C_d 的典型取值范围为 $[1, +\infty)$; 当缺陷呈直线型时, C_d 接近 1; 当缺陷呈曲线型时, C_d 明显大于 1。对于平行线组合型,由于多条细线的聚集效应,取值通常介于前两者之间。

纹理状缺陷通常由多个局部点状缺陷构成,其空间分布特征是重要的区分特征之一。采用灰度方差比 U_d , 即:

$$U_d = \sigma_d / \mu_d \quad (3)$$

其中,将缺陷区域均分为 3×3 网格的子区域, σ_d 表示各子区域灰度标准差的均值, μ_d 为整个缺陷区域的灰度标准差。当 U_d 接近 1 时,表明缺陷呈现均匀分布特征; 当 U_d 明显小于 1 时,表明缺陷呈现聚集分布特征; 当 U_d 明显大于 1 时,表明缺陷呈现渐变分布特征。

2.2 多维特征量化模型

虽然形态特征是区分缺陷类型的主要依据,但在实际应用中,形状规则度相近的块状缺陷可能因尺寸差异需要不同的处理策略。曲直度相似的线状缺陷可能因方向不同而表现出不同的危害程度。空间分布特征相近的纹理缺陷可能因对比度差异导致检测难度不同。为完整表征缺陷特征,基于 2.1 节构建的形态特征量化指标,本节进一步扩展特征维度,构建包含形态、灰度、尺度和方向等多维的综合特征表征模型。

$$\mathbf{x}_i = [F_i, G_i, S_i, \theta_i, A_i] \quad (4)$$

其中,分量 F_i 表示形态特征,根据缺陷形态类型分别采用形状规则度 R_s 、曲直度 C_d 以及灰度方差比 U_d 。分量 G_i 表示灰度差异指标 G_d ,用于量化缺陷区域与背景的灰度对比度特征,用于准确评估缺陷的识别难度。分量 S_i 表示面积比指标 R_a ,用于量化缺陷的影响范围,描述缺陷的多尺度特性。分量 θ_i 表示方向指标 θ_d ,用于描述缺陷的方向特征。分量 A_i 表示长宽比指标 A_r ,用于描述缺陷的延展性特性。它们可表示为:

$$G_d = (I_d - I_b) / (I_b + \varepsilon) \quad (5)$$

$$R_a = A_d / A_t \quad (6)$$

$$\theta_d = \arctan(d_y / d_x) \quad (7)$$

$$A_r = L_d / W_d \quad (8)$$

其中, I_d 为缺陷区域平均灰度值, I_b 为背景区域平均灰度值, ε 为小正数 1×10^{-6} ,用于避免分母接近0时的计算异常。 A_t 为检测区域总面积。 d_y 和 d_x 分别为缺陷区域在 x 和 y 方向上的投影长度,取值范围为 $[-90^\circ, 90^\circ]$ 。 L_d 表示缺陷区域外接矩形的最大延伸长度, W_d 表示垂直于延伸方向的缺陷区域外接矩形的宽度。

3 基于特征表征与学习反馈的样本筛选方法

由于带钢表面缺陷样本存在着表现相似、特征冗余,可能影响模型对缺陷特征的有效学习。为此,本节设计了一种融合特征表征与学习反馈的动态样本筛选方法。总体框架如图5所示,主要包含特征表征和动态优化两个模块。

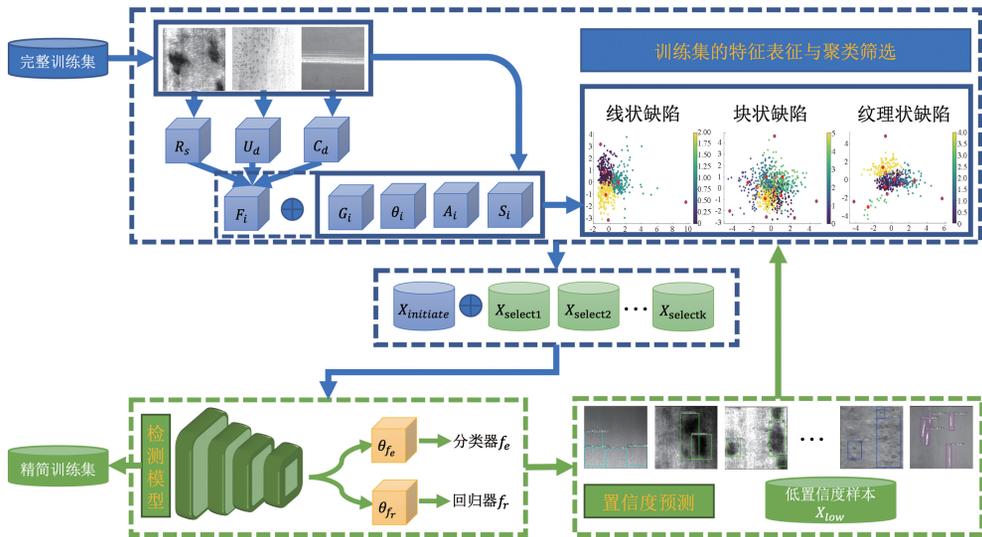


图5 动态样本筛选方法框架

Fig. 5 Framework of the dynamic sample selection method

特征表征模块,基于视觉特征构建了包含形态特征 F_i 、灰度差异 G_i 、面积比 S_i 、方向 θ_i 以及长宽比 A_i 的五维量化模型,结合特征聚类分析筛选初始训练集 $X_{initiate}$ 。动态优化模块在每轮迭代中,首先利用当前训练集训练模型,随后对剩余训练样本进行置信度预测。将置信度低于动态阈值 $Tc(k)$ 的样本归入低置信度样本集 X_{low} ,对其再次进行特征聚类分析,筛选其中的关键补充样本构成 X_{select} 。这些补充样本被加入训练集进行新一轮训练,形成特征表征与学习反馈的动态优化循环。通过这种方式,在保持类内特征多样性的同时,最小化冗余样本数量,以构建具有代表性的精简训练集。

3.1 基于特征表征的样本筛选策略

为高效获取各类缺陷的差异性代表样本,本节设计了基于特征表征的样本筛选策略。首先,根据缺陷的形态类型获取每个样本对应的特征向量,为确保各特征分

量的量纲一致性,对特征向量进行标准化处理,即:

$$\mathbf{X}'_i = (\mathbf{X}_i - \mathbf{X}_{\min}) / (\mathbf{X}_{\max} - \mathbf{X}_{\min}) \quad (9)$$

其中, \mathbf{X}_{\min} 和 \mathbf{X}_{\max} 分别表示各特征维度的最小值和最大值。

随后,对不同形态类型的缺陷分别进行 K-means 聚类分析。为确定最优聚类数 K ,引入轮廓系数 $S(k)$ 作为评价指标,即:

$$S(k) = (b(i) - a(i)) / \max(a(i), b(i)) \quad (10)$$

其中, $a(i)$ 为样本 i 与同类其他样本的平均距离, $b(i)$ 为样本 i 与最近邻类样本的平均距离。通过遍历可能的 K 值,选择使 $S(k)$ 最大的 K 值作为最优聚类数。

获得特征聚类结果后,为快速获取少量具有差异性的代表性样本,在每个聚类簇中选取 3 个关键位置样本构成初始训练集 $X_{initiate}$ 。第 1 个是簇中心样本,用于表征该类别的典型特征。第 2 个是与其他类别交界处的边

界样本,用于刻画类别间的区分特征。第3个是类自身的外围边界样本,用于刻画类别的分布范围。

3.2 基于学习反馈的动态优化机制

由于初始训练集 $\mathbf{X}_{initiate}$ 的少量样本仅能覆盖特征空间中的显著差异区域,同时受统计特征表征的局限性,其覆盖范围在复杂缺陷特征的抽象表征方面仍存在不足。为弥补这一不足,本节设计了基于模型学习反馈的动态优化机制。该机制通过分析模型预测过程中对样本的置信度,识别特征表达能力薄弱的区域,并有针对性地补充样本,从而实现训练集的自适应优化。具体来说,该机制首先获取检测模型输出的置信度得分 $score(x_i)$ 。同时,为使不同训练阶段的置信度具有可比性,计算归一化置信度评分 $Conf(i)$ 。随后,结合动态阈值获取低置信度样本集 \mathbf{X}_{low} , 即:

$$Conf(i) = (score(x_i) - score_{min}) / (score_{max} - score_{min}) \quad (11)$$

$$Tc(k) = T_0 + \alpha \times k \quad (12)$$

$$\mathbf{X}_{low} = \{x_i \mid Conf(i) < Tc(k)\} \quad (13)$$

其中, $score_{max}$ 和 $score_{min}$ 分别为当前批次样本的最大和最小置信度得分。该归一化处理使得不同批次的置信度评估具有可比性。 k 为迭代轮次, T_0 为初始阈值, α 为步长系数。该动态阈值机制能够随着训练进程逐步提高筛选标准,确保在模型性能的情况下仍能筛选出真正具有价值的补充样本。

这些低置信度样本表明当前模型对其特征表达能力不足,反映了训练集中特征覆盖的薄弱环节。然而,这些低置信度样本之间可能存在特征相似的冗余样本,直接将所有样本加入训练集会降低训练效率。为此,采用3.1节的样本筛选策略对低置信度样本集 \mathbf{X}_{low} 进行筛选,从中选取关键补充样本,并填入关键补充样本集 \mathbf{X}_{select} , 即:

$$\mathbf{X}_{select} = \{x_{c_k}, x_{b_k}, x_{m_k} \mid k = 1, 2, \dots, K\} \quad (14)$$

其中, x_{c_k} 表示第 k 个聚类簇的中心样本, x_{b_k} 表示与相邻类别交界处的边界样本, x_{m_k} 表示类自身的边界样本, K 为最优聚类数。

此动态优化机制充分考虑了特征空间分布特性,既保证了样本的代表性,又有效避免了冗余样本的引入。筛选得到的关键补充样本集 \mathbf{X}_{select} 可直接加入初始训练集 $\mathbf{X}_{initiate}$ 进行下一轮训练。

4 实验与分析

4.1 实验环境与数据集

1) 数据集

本研究选用带钢表面缺陷检测领域广泛认可的标准

数据集 NEU-DET^[26]。该数据集的所有图像样本均来源于真实热轧带钢连续生产线现场采集的图像数据,所涵盖的缺陷类型能够真实体现带钢轧制生产过程中常见缺陷的典型特征,如裂纹的延展性、氧化皮的复合分布特征、划痕的方向性规律等。数据集共包含 1 800 幅图像,涵盖热轧带钢表面 6 种典型缺陷类型。如图 6 所示,包括裂纹(crazing, Cr)、夹杂(inclusion, In)、斑块(patch, Pa)、点蚀(pitted surface, Ps)、氧化皮(rolled in scale, Rs)和划痕(scratches, Sc)。每类缺陷均含有 300 幅尺寸为 200 pixels×200 pixels 的灰度图像。

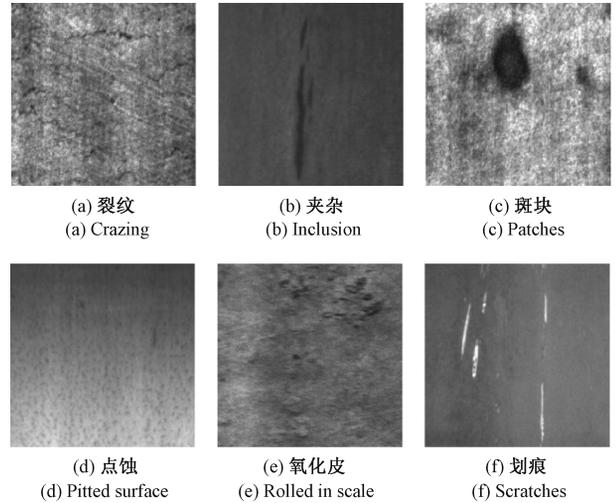


图6 NEU-DET中的6类缺陷

Fig. 6 Six types of defects in NEU-DET

2) 参数设置与环境配置

实验框架使用 PyTorch^[27], 采用 Python 3.8 编程语言和 PyTorch 1.7.1 深度学习框架。实验在配备 24 GB 显存的 Nvidia GeForce RTX 3090 GPU 上运行,操作系统为 Ubuntu 18.04。

为确保实验的可靠性和类别均衡性,数据集采用分层抽样方法进行划分,保持每类缺陷的训练集与测试集比例为 8:2,即每类缺陷随机选取 240 张图像作为待筛选训练集,60 张图像作为测试集。为避免采样偏差,设置固定随机种子为 42 进行数据划分,以确保实验的可重复性。

目标检测模型选用快速区域卷积神经网络(faster region-based convolutional neural network, Faster R-CNN),主干网络采用残差网络(residual network, ResNet)。训练采用自适应矩估计优化器进行参数优化,初始学习率设置为 0.001,批大小为 16,训练过程共计 260 轮。动态优化过程中,通过网格搜索法确定初始置信度阈值 T_0 为 0.6,步长系数 α 设置为 0.02。模型训练采用早停策略,当连续 3 轮的平均精度均值提升幅度 < 1% 时停止迭代,以平衡计算效率和模型性能。

3) 模型评价指标

对于模型的整体检测效果,选择目标检测中最常使用平均精度均值(mean average precision, mAP)来评价网络模型的整体性能。并采用精确率-召回率(precision-recall, PR)曲线评估模型在不同置信度阈值下的检测性能。PR 曲线展示了不同置信度阈值下的精确率(Precision, P)和召回率(Recall, R)。它们定义为:

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

$$Recall = \frac{TP}{TP + FN} \tag{16}$$

$$AP = \frac{1}{11} \sum_{R \in \{0, 0.1, \dots, 1\}} \max_{R' \geq R} P(R') \tag{17}$$

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \tag{18}$$

其中, AP 表示平均精度, TP 为成功预测的正例, FP 为被模型误判为正例的负例, FN 表示被模型错误预测为负例的正例。n 表示模型的预测类别总数。

4.2 可行性验证

表 1 列出了各类型缺陷在样本筛选过程中的统计结果。相比原始训练集,本研究方法最终保留了约 48% 的训练样本。结果显示,不同类型缺陷的筛选比例存在差异,这反映了各类缺陷的特征复杂度和冗余程度不同。其中,斑块缺陷的保留率最低,仅为原始训练集的 39.2%,说明该类缺陷样本中存在较多冗余特征。而裂纹缺陷的保留率最高,达到了 56.7%,表明其样本特征分布较为分散,需要更多样本来保证特征空间的覆盖。

表 1 数据集筛选结果

Table 1 Data set filtering results

标记方式	整体	裂纹	夹杂	斑块	点蚀	氧化皮	划痕
原始训练集	1 440	240	240	240	240	240	240
筛选完训练集	690	136	118	94	113	128	101
测试集	360	60	60	60	60	60	60

基于上述筛选结果,设计了 3 组对照实验以验证本研究方法的可行性。第 1 组为基准组,使用完整训练集。第 2 组为随机组,即随机选择相同数量的训练样本。第 3 组为本研究所提方法获得的精简数据集。表 2 展示了各组在不同缺陷类型上的检测性能。

实验结果表明,采用本研究所提方法筛选的精简数据集在减少约 52% 训练样本的情况下,保持接近完整数据集的检测水平。观察具体结果可以发现,本研究所提方法在总体 mAP 上达到 76.99%,与完整数据集的 77.89% 相比仅下降 0.9 个百分点。特别是在裂纹(Cr)

表 2 可行性验证

Table 2 Feasibility verification (%)

数据集	整体	裂纹	夹杂	斑块	点蚀	氧化皮	划痕
完整数据	77.89	52.49	76.51	89.04	84.68	74.39	90.25
随机选择	70.83	43.66	62.27	89.47	88.53	61.63	79.43
本研究方法	76.99	60.52	75.65	87.91	80.93	70.18	86.76

缺陷上,本研究所提方法取得了 60.52% 的性能,相比完整数据集的 52.49% 有明显提升,这一现象验证了本研究提出的样本筛选策略有效地保留了具有代表性的样本,通过筛选过程去除了一些质量较差或有噪声的样本。相比之下,随机选择方法的 mAP 为 70.83%,各类缺陷检测指标均出现明显下降,其中夹杂(In)缺陷降幅达 14.24 个百分点。这表明简单的随机抽样无法保证样本代表性,容易导致关键特征信息缺失。

表 3 量化了不同样本数量在计算资源消耗上的差异。结果显示,由于筛选后的训练集规模降至原始数据集的 48%,每轮迭代时间从完整数据集的 73 s 降至 28 s,降幅达 62%。证明了本研究所提方法在保持检测性能的同时,有效节省了计算开销。

表 3 训练效率对比

Table 3 Training efficiency comparison

评估指标	完整标签	本研究方法	节省比例
每轮迭代时间/s	73	28	62%
标注样本量/张	1 440	690	52%

图 7 展示的 PR 曲线进一步支持了上述结论。本研究方法的曲线与完整数据方法走势接近,随着召回率增加至 0.7,开始持续逐渐出现性能差距。随机选择方法的曲线则在整个召回率范围内均明显低于其他 2 种方法。验证了本研究所提筛选策略能在有效降低训练时间

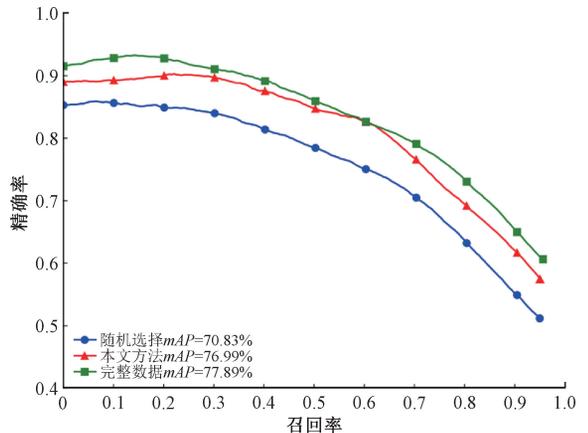


图 7 不同训练集的 PR 曲线

Fig. 7 PR curves of different training sets

的同时,能够维持接近完整数据集的检测性能。

4.3 不同筛选方法的对比实验

将本研究所提的筛选方法与基于核心集的 Core-set 方法^[28]、基于不确定性的 Uncertainty 方法^[29]以及基于密

度的 Density 方法^[30]进行性能对比。为保证实验的公平性,所有方法均采用 48% 的训练集筛选比例,并在相同的实验环境和参数设置下进行训练与测试,得到的检测结果如表 4 所示。

表 4 不同方法的性能对比

Table 4 Performance comparison of different methods

(%)

参数	标签	整体	裂纹	夹杂	斑块	点蚀	氧化皮	划痕
P	Core-set	74.99	70.33	71.82	80.42	84.29	67.21	75.89
	Uncertainty	71.94	68.77	69.50	74.62	83.02	61.54	74.19
	Density	75.36	72.32	73.10	79.38	85.12	65.18	77.04
	本研究方法	77.16	75.02	74.22	78.03	86.34	70.11	79.22
R	Core-set	85.90	83.95	91.27	95.56	77.04	80.23	87.33
	Uncertainty	84.38	83.83	89.76	94.24	76.98	70.02	91.46
	Density	85.13	81.24	92.92	92.08	76.03	71.37	97.16
	本研究方法	87.45	84.47	91.32	94.71	80.11	81.92	92.15
AP	Core-set	74.15	58.16	74.39	89.93	79.89	69.81	72.73
	Uncertainty	74.35	58.76	75.06	87.14	79.13	64.21	81.81
	Density	73.98	57.07	67.32	86.91	78.85	65.49	88.25
	本研究方法	76.99	60.52	75.65	87.91	80.93	70.18	86.76

由表 4 可知,从整体性能来看,本研究所提方法获得了 76.99% 的 mAP,相比 Core-set、Uncertainty 和 Density 方法分别提升了 2.84%、2.64% 和 3.01%。在处理不同类型缺陷时,本研究所提方法表现出了一定的类内特征捕捉能力。对于具有复杂类内形态的裂纹(Cr)缺陷,本研究所提方法相较其他方法在检测精度上有所提升,这得益于所提出的筛选策略在识别和保留类内关键代表性样本方面的作用。对于具有多样化线性结构的夹杂(In)和划痕(Sc)这两类缺陷,本研究所提方法分别达到 75.65% 和 86.76% 的检测率,其中划痕类缺陷的性能仅次于 Density 方法。在表面形态变化丰富、区域范围相对

混乱的氧化皮(Rs)和点蚀(Ps)这两类缺陷上分别取得最高的 70.18% 和 80.93% 的性能。对于形态特征相对稳定的斑块(Pa)类缺陷,本方法的检测性能仅次于 Core-set 方法,达到 87.91%。这些结果表明所提筛选策略在识别和保留类内代表性样本方面具有一定效果。

为进一步说明本研究所提方法的实际应用效果,图 8 展示了本研究所提方法在不同类型缺陷上的检测效果示例。

结果显示,本研究所提方法不仅能够准确定位和识别各类缺陷,而且在复杂背景和多个目标的场景下依然保持稳定的性能表现。图 8(a)展示了检测难度相对较

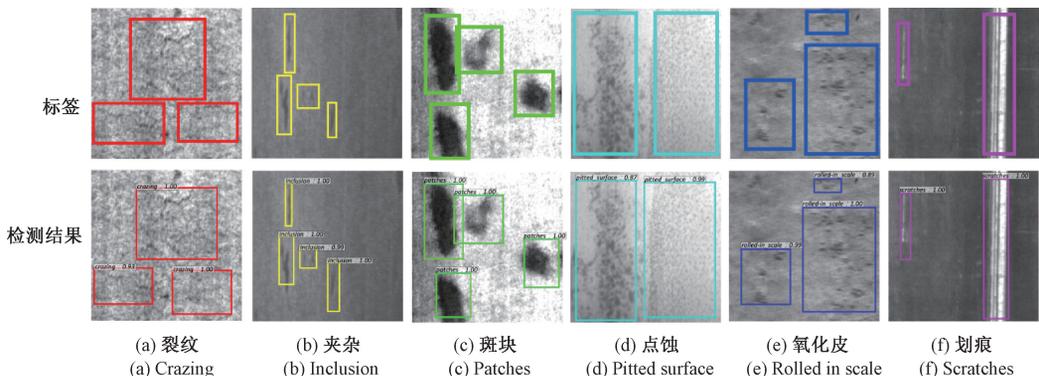


图 8 可视化检测结果

Fig. 8 Visualize the test results

高的裂纹(Cr)缺陷,尽管其形态细长、分布情况混乱且部分区域对比度较低,但本研究提方法仍能准确检测出缺陷边界范围,且置信度达到1。图8(b)、(c)、(f)展示了夹杂(In)、斑块(Pa)以及划痕(Sc)等局部特征明显的缺陷,这些缺陷的检测框基本都能完整包围目标区域,且框选精准。特别是对于夹杂(In)以及划痕(Sc)中的一些尺寸偏小的缺陷,也能实现准确检测,这验证了所提出的样本筛选策略在保留例如形状、尺度等关键特征方面的有效性。同时,图8(b)显示在多个缺陷相互靠近的情况下,算法仍能准确区分各个独立的缺陷目标,展现出良好的独立目标区分能力。图8(d)、(e)展示了对于面积较大的点蚀缺陷和氧化皮缺陷,本方法能够准确框选缺陷区域。值得注意的是,即使在缺陷边界模糊、背景纹理复杂的情况下,检测框也能较好地贴合实际缺陷轮廓,这

进一步验证了本研究所提出的样本筛选策略对于复杂特征亦具备有效性。

综上所述,通过与不同的样本筛选方法的全面对比,实验结果充分证明了本研究提出的样本筛选策略的有效性。并且,该方法在有效降低训练时间的同时,实现了接近全量数据训练的检测性能。

4.4 不同检测模型的适应性验证

为验证所提方法在不同目标检测模型上的适用性,本节选取文献[31-33]的方法进行对比实验。这些模型分别代表单阶段检测器、多阶段检测器和基于Transformer的检测器,覆盖了主干网络优化、检测器架构优化和框架设计优化的不同研究方向。实验设置与4.1节相同,完整训练集和筛选后训练集在不同检测模型上的性能表现如表5所示。

表5 不同检测模型的性能对比

Table 5 Performance comparison of different detection models

检测模型	主干网络	训练数据	每轮迭代时间/s	效率节省/%	mAP/%	性能损失/%
文献[31]	CSPNet-L	完整数据	17	41.2	81.12	0.73
		筛选数据	9		80.39	
文献[32]	ResNet50	完整数据	46	50.0	79.95	1.81
		筛选数据	23		78.14	
文献[33]	ResNet18	完整数据	28	35.7	82.56	1.26
		筛选数据	18		81.30	

实验结果表明,本研究提出的样本筛选方法在各类检测架构上均有效提升了训练效率。其中,文献[32]的多阶段检测器获得了最显著的效率提升,训练时间减少了50%,同时性能损失为1.81%。这是因为多阶段检测器在训练过程中需要对每个候选区域进行多次特征提取和优化,样本数量的减少能大幅降低计算开销。文献[31]的单阶段检测器展现出最佳的性能保持能力,在训练时间减少41.2%的同时,性能损失仅为0.73%,这主要是由于其本身就采用了高效的单次推理机制。文献[33]基于Transformer的检测器获得了35.7%的效率提升,性能损失为1.26%,效率提升幅度最小,这可能是因为其注意力机制的计算复杂度主要受特征维度影响,对样本数量的依赖相对较小。

上述分析证明了本研究所提方法在保留数据集关键特征信息方面的有效性。特别是在计算开销较大的多阶段检测器上,能够在保持较高检测精度的同时实现显著的效率提升。同时,本研究提出的样本筛选方法具有良好的通用性,能够适应不同类型的检测架构,在提升训练效率的同时保持模型性能。

5 结 论

本研究针对带钢表面缺陷样本存在着表现相似、特征冗余,影响模型对缺陷特征的学习效率和检测性能的问题。提出了一种融合特征表征与学习反馈的动态样本筛选方法。该方法首先从形态学视角构建了一个多维特征量化模型,实现了对不同形态类型缺陷特征的全面表征。其次,设计了基于特征表征的样本筛选机制,通过聚类分析特征空间分布,快速筛选出具有差异性的少量代表性样本,在保证样本代表性的同时有效降低了数据冗余。最后,提出了基于置信度评估的动态优化机制,实现了训练样本的自适应扩充和质量优化。通过实验部分的结果,表明本研究所提方法在将训练样本量降低52%的情况下,实现了76.99%的检测准确率。实验证明,本研究所提方法不仅保持了与完整数据集相当的检测性能,同时有效降低了训练时间。

此外,通过进一步在涵盖单阶段、多阶段及Transformer架构的多种检测模型上的验证结果表明,该方法在实现训练效率提升的同时,也能保持相对稳定的

检测精度,验证了方法在不同检测架构上的适应性和有效性,为提升带钢表面缺陷检测模型的训练效率提供了可行的研究思路。

参考文献

- [1] ZHAO B T, CHEN Y R, JIA X F, et al. Steel surface defect detection algorithm in complex background scenarios[J]. *Measurement*, 2024, 237:115189.
- [2] TANG B, CHEN L, SUN W. Review of surface defect detection of steel products based on machine vision[J]. *IET Image Processing*, 2023, 17(2):303-322.
- [3] 李丹,王慢慢,刘俊德,等. 基于轻量级卷积神经网络的带钢表面缺陷识别[J]. *仪器仪表学报*, 2022, 43(3):240-248.
- LI D, WANG M M, LIU J D, et al. Steel strip surface defect recognition based on lightweight convolutional neural network [J]. *Chinese Journal of Scientific Instrument*, 2022, 43(3):240-248.
- [4] LI ZH G, WEI X M, HASSABALLAH M, et al. A deep learning model for steel surface defect detection [J]. *Complex & Intelligent Systems*, 2024, 10(1):885-897.
- [5] LI SH L, YUAN W Q. Detection of contamination defect on ice cream bar based on fuzzy rule and absolute neighborhood[J]. *Instrumentation*, 2017, 4(3):24-34.
- [6] 伍麟,郝鸿宇,宋友. 基于计算机视觉的工业金属表面缺陷检测综述[J]. *自动化学报*, 2024, 50(7):1261-1283.
- WU L, HAO H Y, SONG Y. A survey on industrial metal surface defect detection based on computer vision[J]. *Acta Automatica Sinica*, 2024, 50(7):1261-1283.
- [7] 苑玮琦,齐健婷. 冰勺表面缺陷检测方法研究[J]. *仪器仪表学报*, 2022, 43(6):239-250.
- YUAN W Q, QI J T. Research on the detection method of surface defects of ice spoons[J]. *Chinese Journal of Scientific Instrument*, 2022, 43(6):239-250.
- [8] ZHAO ZH Q, ZHENG P, XU SH T, et al. Object detection with deep learning: A review [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 30(11):3212-3232.
- [9] 李德健,李绍丽,苑玮琦,等. 基于自适应方向模板搜索的FPC走线薄浆检测[J]. *电子测量与仪器学报*, 2022, 36(7):177-188.
- LI D J, LI SH L, YUAN W Q, et al. FPC trace paste attenuation detection based on adaptive directional template[J]. *Journal of Electronic Measurement and Instrumentation*, 2022, 36(7):177-188.
- [10] SONG CH H, CHEN J X, LU ZH, et al. Steel surface defect detection via deformable convolution and background suppression [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72:5017709.
- [11] 陈俊英,李朝阳,黄汉涛,等. 并行特征提取和渐进特征融合的计算机主板装配缺陷检测[J]. *光学精密工程*, 2024, 32(10):1622-1637.
- CHEN J Y, LI CH Y, HUANG H T, et al. Computer motherboard assembly defect detection based on parallel feature extraction and progressive feature fusion [J]. *Optics and Precision Engineering*, 2024, 32(10):1622-1637.
- [12] ZHAO Y H, LIU Q, SU H, et al. Attention-based multiscale feature fusion for efficient surface defect detection[J]. *IEEE Transactions on Instrumentation and Measurement*, 2024, 73:5013310.
- [13] 赵洋,王军凯,林志毅,等. 基于改进YOLOv8n的钢材表面缺陷检测[J]. *电子测量技术*, 2024, 47(13):191-198.
- ZHAO Y, WANG J K, LIN ZH Y, et al. Steel surface defect detection algorithm based on improved YOLOv8n[J]. *Electronic Measurement Technology*, 2024, 47(13):191-198.
- [14] 张航,周毅,邱宇峰. 融合HGnetv2和注意力机制的钢材表面缺陷检测方法[J]. *电子测量与仪器学报*, 2025, 39(1):36-49.
- ZHANG H, ZHOU Y, QIU Y F. Detection method of steel surface defects with fusion of HGnetv2 and attention mechanism[J]. *Journal of Electronic Measurement and Instrumentation*, 2025, 39(1):36-49.
- [15] ZHANG H, FU W, WANG X M, et al. An efficient model for metal surface defect detection based on attention mechanism and multi-scale feature [J]. *The Journal of Supercomputing*, 2025, 81(1):1-25.
- [16] SHEN L, SUN Y, YU ZH Y, et al. On efficient training of large-scale deep learning models[J]. *ACM Computing Surveys*, 2024, 2025(3):3700439.
- [17] NIU SH L, LI B, WANG X G, et al. Defect image sample generation with GAN for improving defect recognition [J]. *IEEE Transactions on Automation Science and Engineering*, 2020, 17(3):1611-1622.
- [18] 岳忠牧,张喆,吕武,等. De-DDPM:可控、可迁移的缺陷图像生成方法[J]. *自动化学报*, 2024, 50(8):1539-1549.
- YUE ZH M, ZHANG ZH, LYU W, et al. De-DDPM: A controllable and transferable defect image generation method [J]. *Acta Automatica Sinica*, 2024, 50(8):1539-1549.
- [19] HU X J, YANG J, JIANG F L, et al. Steel surface defect detection based on self-supervised contrastive representation learning with matching metric[J]. *Applied Soft Computing*, 2023, 145:110578.
- [20] 刘金海,赵真,付明芮,等. 基于主动小样本学习的管

- 道焊缝缺陷检测方法[J]. 仪器仪表学报, 2022, 43(11): 252-261.
- LIU J H, ZHAO ZH, FU M R, et al. Pipeline weld defect detection method based on active few-shot learning[J]. Chinese Journal of Scientific Instrument, 2022, 43(11): 252-261.
- [21] WANG Z F, ZHU H, DONG ZH H, et al. Less is better: Unweighted data subsampling via influence function[C]. 34th AAAI Conference on Artificial Intelligence, 2020, 34(4): 6340-6347.
- [22] YANG SH, XIE Z K, PENG H Y, et al. Dataset pruning: Reducing training data by examining generalization influence [J]. ArXiv preprint arXiv: 2205.09329, 2022.
- [23] SHAO J Y, WANG Q, LIU F B. Learning to sample: An active learning framework[C]. 2019 IEEE International Conference on Data Mining, 2019: 538-547.
- [24] SOLTANZADEH P, FEIZI-DERAKHSHI M R, HASHEMZADEH M. Addressing the class-imbalance and class-overlap problems by a metaheuristic-based under-sampling approach[J]. Pattern Recognition, 2023, 143: 109721.
- [25] 仲林林, 吴奇, 叶俊杰, 等. 基于元学习的变电设备小样本缺陷图像检测 [J]. 仪器仪表学报, 2024, 45(10): 154-167.
- ZHONG L L, WU Q, YE J J, et al. Meta-learning-based few-shot image detection of defects in substation equipment[J]. Chinese Journal of Scientific Instrument, 2024, 45(10): 154-167.
- [26] BAO Y Q, SONG K CH, LIU J, et al. Triplet-graph reasoning network for few-shot metal generic surface defect segmentation[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 3083561.
- [27] PASZKE A, LERER A, KILLEEN T, et al. PyTorch: An imperative style, high-performance deep learning library[J]. ArXiv preprint arXiv:1912.01703, 2019.
- [28] SENNER O, SAVARESE S. Active learning for convolutional neural networks: A core-set approach[J]. ArXiv preprint arXiv:1708.00489, 2017.
- [29] ELLEUCH J F, MEHDI M Z, KAMMOUN H M, et al. New possibilistic formalism based on confidence levels for metallic surface defect classification[J]. IEEE Access, 2024, 12: 93010-93019.
- [30] GUI X Q, LU X D, YU G X. Cost-effective batch-mode multi-label active learning[J]. Neurocomputing, 2021, 463: 355-367.
- [31] VARGHESE R, SAMBATH M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness [C]. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems, 2024: 1-6.
- [32] CAI ZH W, VASCONCELOS N. Cascade R-CNN: High quality object detection and instance segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(5): 1483-1498.
- [33] ZHAO Y, LYU W Y, XU SH L, et al. DETRs beat YOLOs on real-time object detection [C]. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

作者简介



苑玮琦, 1982年于湖南大学获得学士学位, 1988年于东北大学获得硕士学位, 1997年于东北大学获得博士学位, 现为沈阳工业大学教授, 主要研究方向为机器视觉和图像识别。

E-mail: yuan60@126.com

Yuan Weiqi received his B. Sc. degree from Hunan University in 1982, M. Sc. degree from Northeastern University in 1988, and Ph. D. degree from Northeastern University in 1997. He is currently a professor at Shenyang University of Technology. His main research interests include machine vision inspection and image recognition.



刘文滔 (通信作者), 2017年于吉林财经大学获得学士学位, 2020年于吉林大学获得硕士学位, 现为沈阳工业大学博士研究生, 主要研究方向为机器视觉检测。

E-mail: 972151125@qq.com

Liu Wentao (Corresponding author) received her B. Sc. degree from Jilin University of Finance and Economics in 2017, and received her M. Sc. degree from Jilin University in 2020. She is currently a Ph. D. candidate at Shenyang University of Technology. Her main research interest includes machine vision inspection.



李绍丽, 2013年于沈阳工业大学获得学士学位, 2015年于沈阳工业大学获得硕士学位, 2018年于沈阳工业大学获得博士学位, 现为沈阳工业大学副教授, 主要研究方向为机器视觉检测、数字图像处理及模式识别。

E-mail: lishaoli@sut.edu.cn

Li Shaoli received her B. Sc. degree from Shenyang University of Technology in 2013, M. Sc. degree from Shenyang University of Technology in 2015 and Ph. D. degree from Shenyang University of Technology in 2018. She is currently an associate professor at Shenyang University of Technology. Her main research interests include machine vision inspection, digital image processing and pattern recognition.