

基于元学习的带钢表面缺陷小样本语义分割

冯 虎, 宋克臣, 崔文琦, 颜云辉

(东北大学 机械工程与自动化学院, 辽宁 沈阳 110819)

摘 要: 由于缺少带钢表面缺陷样本,使得深度神经网络在带钢表面缺陷检测的应用受到了限制,为解决这一实际问题,提出了一种基于元学习思想的小样本语义分割深度学习方法.该方法引入了多尺度解码器和注意力机制.多尺度解码器能够聚合不同尺度的缺陷特征信息,提高网络的分割精度.注意力机制能够有效增强缺陷信息表达,并且抑制背景信息的干扰.此外,构建了一个带钢表面缺陷语义分割数据集,该数据集包含9类带钢表面缺陷.在该数据集上进行了相关实验,结果表明本文方法在平均交并比和前景-背景交并比指标上优于PFENet, SCLNet和HSNet等方法.

关 键 词: 带钢表面缺陷检测;元学习;小样本语义分割;注意力机制;多尺度解码器

中图分类号: TP 391.4 **文献标志码:** A **文章编号:** 1005-3026(2024)03-0354-07

Few-Shot Semantic Segmentation of Strip Steel Surface Defects Based on Meta-Learning

FENG Hu, SONG Ke-chen, CUI Wen-qi, YAN Yun-hui

(School of Mechanical Engineering & Automation, Northeastern University, Shenyang 110819, China.
Corresponding author: SONG Ke-chen, E-mail: songkc@me.neu.edu.cn)

Abstract: Due to the limited availability of strip surface defect samples, the application of deep neural networks in strip surface detection is constrained. To solve this practical issue, a meta-learning-based few-shot semantic segmentation method is proposed. A multi-scale decoder and attention mechanism are used in the proposed method. The multi-scale decoder can aggregate the defect information at different scales and improve segmentation accuracy of the proposed network. The attention mechanism can effectively improve the expression of defect features and suppress the interference of defect background information. In addition, a novel few-shot steel strip surface defect semantic segmentation dataset is constructed including nine classes of strip steel surface defects. Comparison experiments on the proposed dataset show that the proposed method is superior to similar few-shot segmentation methods such as PFENet, SCLNet, and HSNet in terms of evaluation index mean intersection over union and foreground-background intersection over union.

Key words: strip steel surface defect detection; meta-learning; few-shot semantic segmentation; attention mechanism; multi-scale decoder

带钢作为一种重要的工业材料,广泛应用于航空航天、深海探测、桥梁建筑和汽车制造等领域^[1].但是受到制造工艺的影响,带钢表面会产生夹杂、裂纹、划痕等缺陷^[2].为了防止表面缺陷对使用安全的潜在威胁,在带钢生产过程中必须进行表面检测.

传统检测方法需要人工目视对高速移动的

带钢进行缺陷判断^[3],该方法存在效率低下和检测者工作强度大的问题,并且容易受到人为影响导致误检和漏检.因此,基于机器视觉的带钢表面质量自动视觉检测(automated visual inspect, AVI)技术对提高检测速率和检测质量有积极意义^[4-5].传统的自动视觉检测方法依据缺陷的纹理、颜色等特征,将图像处理方法与机器学习方

收稿日期: 2022-10-21

基金项目: 国家自然科学基金资助项目(51805078); 中央高校基本科研业务费专项资金资助项目(N2103011).

作者简介: 冯 虎(1998-),男,陕西汉中,东北大学硕士研究生; 颜云辉(1960-),男,江苏丹阳人,东北大学教授,博士生导师.

法结合对缺陷进行检测,包括基于图像检测和基于频域检测.基于图像检测通常采用边缘检测方法(如Prewitt, Sobel和Canny)对缺陷区域进行定位,而基于频域检测则利用傅里叶变换和小波变换将缺陷特征转换至频域信号进行检测^[6].此外,支持向量机(support vector machine, SVM)和随机森林(random forest, RF)等机器学习方法也被用于缺陷分类.然而,传统方法依赖人工构建缺陷特征,并需要复杂的后处理过程,难以适应复杂的工业生产环境^[7].

基于深度学习的缺陷检测方法能够有效解决传统方法存在的缺陷特征构建困难、方法泛化性不足等问题^[8].此类方法通常基于目标检测或者语义分割^[9].基于目标检测方法能够快速定位和分类缺陷^[10].然而,由于带钢表面缺陷形状的不规则性和类内差异性,目标检测方法难以精确定位缺陷区域,并且容易受到干扰信息的影响,从而导致检测结果存在较大误差.相比之下,语义分割方法是一种高精度的像素级的检测方法,在工业领域常用于进行细粒度的缺陷检测任务^[11].

基于深度学习的语义分割方法在自然图像的分割任务中性能良好.然而,当训练样本稀疏时,这些方法容易出现过拟合问题,导致训练后的网络分割性能难以达到预期效果.在带钢生产领域,由于企业对产品质量控制严格,带钢次品率较低,研究者难以采集充足的带钢表面缺陷数据.此外,高精度的带钢表面缺陷数据集标注工作需要耗费大量人力,导致数据集成本高昂.基于元学习的小样本语义分割方法能够解决因样本稀疏导致的分割性能退化问题.在自然数据集领域中,Zhang等^[12]提出使用支持集辅助原型作为前景原型的补充,以提高网络对目标信息的提取能力.Tian等^[13]提出了一种先验引导特征富集的小样本分割方法.Min等^[14]认为多级特征相关性和4维卷积能够有效提高网络的分割性能.为了解决工业生产中表面缺陷样本不足的问题,Bao等^[9]构建了一个包括金属和非金属表面缺陷的数据集,并在该数据集上验证了小样本分割方法解决表面缺陷问题的可行性.

为了解决生产过程中带钢表面缺陷样本采集困难及样本标注成本高昂等问题,本文提出了一种针对带钢表面缺陷检测的小样本语义分割方法,使用注意力机制和多尺度解码器提高缺陷分割性能;构建了一个新的带钢表面缺陷的数据集,并在该数据集上做了对比实验和消融实验,验证了本文方法在解决带钢表面缺陷样本不足

问题上的有效性.

1 方法描述

1.1 小样本语义分割概述

本文基于元学习(meta-learning)思想,运用小样本语义分割理论解决带钢表面缺陷分割的任务.Dong等^[15]提出了小样本分割范式,包括元训练(meta-train)和元测试(meta-test)两个阶段.首先,将缺陷数据划分为元训练集和元测试集.其中,小样本语义分割将元训练集视为“已知类”,将元测试集视为“未知类”,二者之间不存在交叉的缺陷类别.在元训练阶段,元训练集被用于训练网络参数;在元测试阶段,元测试集被用于评估网络分割性能.与基于全监督学习的语义分割方法不同,训练集和测试集均以小样本任务(few-shot episode)作为基本数据单元,每个任务包含支持集和查询集,其中支持集样本可以为一个或者多个,而查询集样本仅有一个.小样本语义分割训练和测试过程如下:

- 1) 在元训练阶段,使用元训练集中“已知类”缺陷训练网络;

- 2) 网络元训练完成后,将网络的全部可训练参数冻结;

- 3) 在元测试阶段,使用元测试集中的“未知类”样本验证分割性能.

1.2 研究方法

本文提出了网络基于小样本语义分割理论.首先,通过主干网络提取支持集样本和查询集样本的特征图.然后,通过支持特征图的前景特征匹配查询特征图中的相同缺陷类别特征.最后,通过解码器网络对经过匹配的查询集的特征进行特征解码,进而输出对查询集样本的预测结果.本文研究包括如下3个部分:

- 1) 数据集构建.从数据集Surface Defect-4^[9]、X-SDD^[16]和DET-GC10^[17]中筛选含有缺陷的原始样本.通过裁剪、旋转和镜像方式扩充数据集,并使用Labelme对样本进行像素级别标注.该数据集共包括9类缺陷,每类50个样本对,每个样本对中包括缺陷图像及其对应的像素级标签.

- 2) 网络训练.将预处理后的数据集划分为元训练集和元测试集.通过元训练集数据集训练网络,根据参加训练的支持集样本数量,训练分为1个支持集训练样本(1-shot)和5个支持集样本(5-shot)两种情况.

- 3) 性能评估.冻结训练后的网络参数,将测

试集输入网络进行预测,然后通过预测结果与真实标签得到评价指标.评价指标包括平均交并比和前景-背景交并比.实验的流程如图 1 所示.

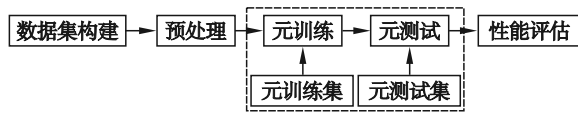


图 1 实验流程图

Fig. 1 Flow chart of the experiment

2 带钢缺陷小样本语义分割方法

2.1 方法概述

本文提出的小样本语义分割方法包括特征

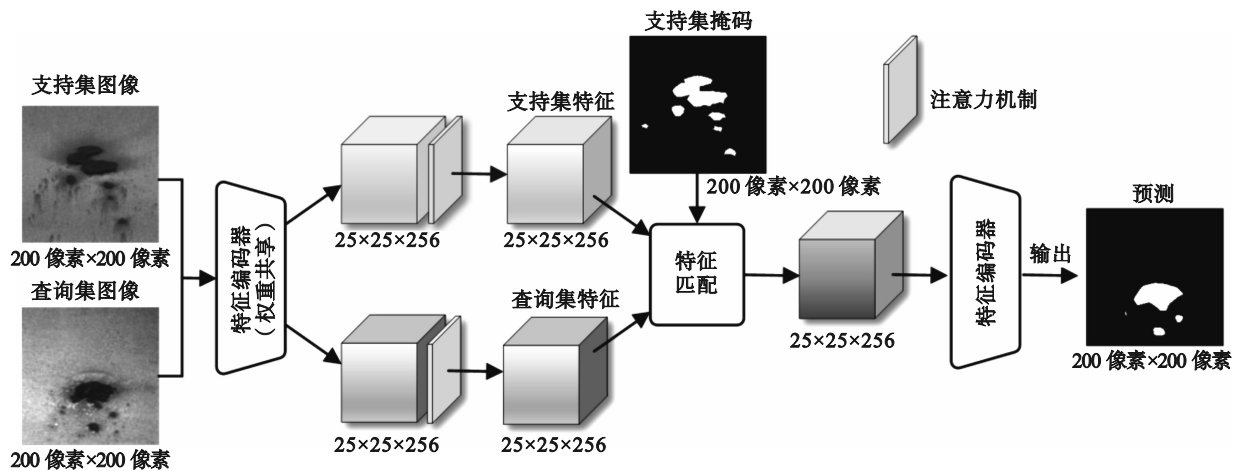


图 2 小样本语义分割网络

Fig. 2 Few-shot semantic segmentation network

2.2 特征提取器

特征提取器用于提取支持集和查询集样本的特征.本文使用在数据集 ImageNet^[18]预训练的主干网络 ResNet-50^[19]作为特征提取器.主干网络输入尺寸为 200 像素×200 像素的单通道支持集和查询集图像,输出维度为 25×25×256 的特征图.需要注意的是,为了保持预训练主干网络良好的特征提取能力,在元训练和元测试过程中均冻结了主干网络的预训练权重.

注意力机制能够抑制与前景无关的背景信息,提高网络对缺陷特征的聚合能力.在特征提取器后嵌入 CBAM^[20]注意力机制模块.该模块沿着两个独立的维度(通道和空间)依次推断注意力图,然后将注意力图与输入特征图矩阵相乘进行自适应特征优化.CBAM 注意力机制模块如图 3 所示.

2.3 特征匹配器

特征匹配器能够提取支持集中缺陷信息的前景原型向量,通过相似度匹配法将查询集中的

提取器、特征匹配器和特征解码器.使用经过预训练的主干网络作为特征提取器来提取支持集和查询集特征,并在支持集和查询集分支中嵌入注意力机制,以激活缺陷特征并抑制背景信息.在进行特征匹配时,首先使用掩膜平均池化获取支持集的原型向量,然后用其与查询集的特征图进行特征匹配.最后,将特征匹配结果传入多尺度解码器得到预测结果.本文通过 $h \times w \times c$ 来表征特征图的维度,其中: h 和 w 为特征图的几何维度,分别代表了特征图的高和宽,单位为像素; c 为特征图的通道维度.小样本语义分割网络示意图如图 2 所示.

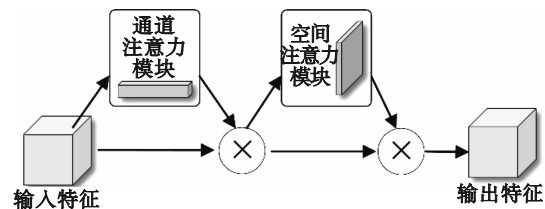


图 3 CBAM 注意力机制模块

Fig. 3 CBAM attention mechanism module

同类别特征与原型向量中的缺陷特征进行对应.

将注意力机制聚合的结果 F_s 和经过下采样的支持集样本的语义标签 M_s 作为特征匹配器的输入.在特征匹配器中,对支持集的特征图进行掩膜平均池化,得到支持集的前景原型向量 P_s .掩膜平均池化表达式为

$$P_s = \text{avgpool}(M_s \odot F_s). \quad (1)$$

其中: $\text{avgpool}(\cdot)$ 表示平均池化; \odot 表示逐像素乘法.掩膜平均池化输出支持集前景原型向量,其维度为 $1 \times 1 \times 256$.

为了匹配支持集特征和查询集的特征,首先

将 P_s 的维度上采样至 $25 \times 25 \times 256$, 然后在通道维度拼接该上采样后的特征图和查询集的特征图. 拼接后的特征图的维度为 $25 \times 25 \times 512$. 将卷积核为 1×1 的卷积层作为度量工具, 对拼接特征图逐像素进行特征匹配, 进而得到匹配特征图, 其维度为 $25 \times 25 \times 256$. 特征匹配模块如图 4 所示.

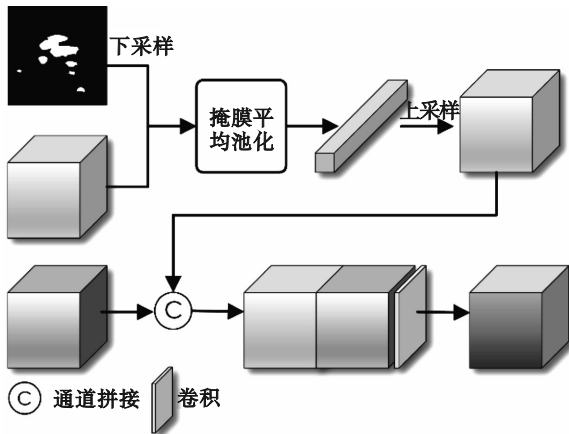


图 4 特征匹配模块
Fig. 4 Feature matching module

2.4 特征解码器

本文构建了一种基于全卷积结构的多尺度解码器, 该解码器能够聚合不同尺度下缺陷特征的上下文信息, 进而提高方法的分割性能. 特征匹配器输出特征图的维度为 $25 \times 25 \times 256$. 首先, 本文使用卷积核为 3×3 , 步距为 2×2 的卷积对该特征图连续 2 次下采样, 分别获得维度为 $12 \times 12 \times 256$ 和 $6 \times 6 \times 256$ 的特征图. 将下采样得到的特征图上采样至 $25 \times 25 \times 256$, 然后沿通道维度拼接原特征图和经过上采样的两个特征图, 拼接后的特征图的维度为 $25 \times 25 \times 768$, 接着用 1×1 卷积层融合拼接后的特征图. 最后, 通过全卷积层对特征图逐像素二分类, 得到预测结果的概率图. 多尺度解码器如图 5 所示.

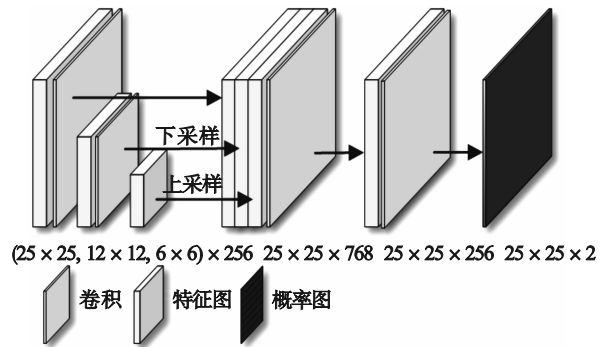


图 5 多尺度解码器
Fig. 5 Multi-scale decoder

2.5 损失函数

本文选择交叉熵损失函数用于分类和分割任务. 带钢表面缺陷语义分割实际上是对缺陷逐像素地分类任务, 每个类预测概率为 p_i 和 $1 - p_i$, 交叉熵损失表达式为

$$L = \frac{1}{N} \sum_i - [y_i \times \ln(p_i) + (1 - y_i) \times \ln(1 - p_i)]. \quad (2)$$

其中: y_i 为正样本 i 的标签, 正样本为 1, 负样本为 0; p_i 为预测样本 i 为正的的概率.

3 实 验

3.1 数据集构建

本文从带钢表面缺陷数据集 Surface Defect-4^[9], X-SDD^[16] 和 DET-GC10^[17] 中整合 9 类常见带钢表面缺陷图像, 包括夹杂、划痕、孔洞、红铁皮、铁皮灰、板系氧化皮、轧辊印、水污和油污. 为了网络批次化训练, 将缺陷图像尺寸均裁剪为 200 像素 × 200 像素. 为了减少运算所占用的内存, 将缺陷图像由三通道转换为单通道. 通过标注软件 Labelme 对缺陷图像进行像素掩膜标注, 生成缺陷图像的语义标签. 经过标注后的缺陷数据集包含 9 种表面缺陷, 每类缺陷包括 50 个缺陷样本, 缺陷样本由缺陷图像和对应的缺陷语义标签组成, 如图 6 所示.

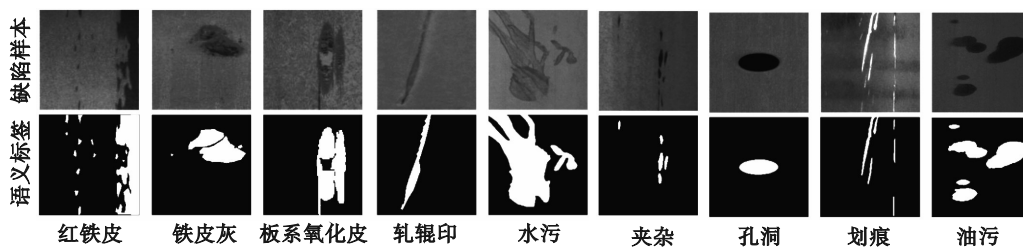


图 6 缺陷及对应语义标签
Fig. 6 Defects and corresponding semantic labels

根据小样本分割的设定, 本文将上述数据集随机划分为 3 个集合: Fold-1, Fold-2 和 Fold-3.

每个集合都包含 3 种不同类别的缺陷. 数据集的划分如表 1 所示.

表 1 数据集的划分
Table 1 Division of dataset

集合	缺陷		
Fold-1	红铁皮	铁皮灰	板系氧化皮
Fold-2	轧辊印	水污	夹杂
Fold-3	油污	孔洞	划痕

3.2 评价指标

本文使用小样本语义分割方法中最常用的两种指标:平均交并比和前景-背景交并比作为评价指标。

1) 平均交并比. 交并比是预测结果的前景区域与标签前景交集面积除以两者的并集面积. 该指标的取值范围为 0~1, 其中 0 表示两者无重合, 1 表示完全重合. 在小样本任务的交叉验证中, 需要计算每一个类的交并比, 求和后取平均值作为平均交并比. 通常将平均交并比作为评价方法精度的主要指标。

交并比计算公式为

$$R_i = \frac{TP_i}{TP_i + FP_i + FN_i}, \quad (3)$$

$$\bar{R} = \frac{1}{N} \sum_{i=1}^N R_i. \quad (4)$$

其中: N 为缺陷类别总数; R_i 为交并比; \bar{R} 为平均交并比; TP 为被模型预测为正类的正样本; FP 为被模型预测为正类的负样本; FN 为被模型预测为负类的正样本。

2) 前景-背景交并比. 该评价指标忽略类别信息, 计算前景交并比和背景交并比. 通常将前景-背景交并比作为评价模型精度的辅助指标。

3.3 环境配置

本文的全部对比实验和消融实验采用 ubuntu 20.04 操作系统, pytorch 1.7.0 的深度学习框架, Python 3.8 编程语言, Pycharm 2022 编译环境, NVIDIA GeForce RTX3060 图像处理器, Intel

Core-i5 11400F@2.60GHz 处理器, 运行内存 16 GB, 使用了 opencv-python 跨平台开源计算机视觉库。

3.4 参数配置

本文方法是一种端到端神经网络, 依据文献[13], 使用随机梯度下降(stochastic gradient descent, SGD)优化器, 优化器的动量系数设置为 0.9, 初始学习率为 0.025, 权重衰减系数为 1×10^{-3} . 训练迭代次数为 100, 每次传入小样本任务数为 4. 为了减小网络训练时产生的随机性偏差, 依据文献[13], 选择 5 个随机种子进行多组实验, 取其平均值作为结果. 实验验证了 1 个支持集样本(1-shot)和 5 个支持集样本(5-shot)两种设定下的指标. 在元训练阶段冻结主干网络的参数, 在训练后冻结网络的全部可训练参数。

3.5 对比实验

如表 2 所示, 在本文构建的带钢表面缺陷数据集上对比了本文方法和现有的小样本分割方法. 各方法均采用 ResNet-50 作为主干网络, 使用相同的图像预处理方法和参数配置. 在 1-shot 时, 本文方法的平均交并比在 Fold-1 比 PFENet 高 2.0%, 在 Fold-2 比 PFENet 低 0.5%, 在 Fold-3 比 HSNet 高 1.3%, 比 PFENet 高 4.3%. SCLNet 的前景-背景交并比超过了其他方法. 在 5-shot 时, 本文方法的平均交并比在 Fold-1 比 PFENet 高 1.1%, 在 Fold-2 比 HSNet 低 2.9%, 在 Fold-3 比 HSNet 低 0.9%, 平均交并比均值比 PFENet 和 HSNet 分别高 0.2% 和 0.3%. 而前景-背景交并比则比 SCLNet 低 1.7%。

分析表 2 结果可知, 相比其他对比方法, 本文方法在最极端的 1-shot 情况分割性能更为突出, 更适应缺陷样本极度稀疏的场景. 虽然 SCLNet 的前景-背景交并比高于其他方法, 但是在缺陷样本不充足时, SCLNet 对缺陷的分割能力受到抑制。

表 2 1 个支持集样本和 5 个支持集样本方法对比实验
Table 2 Comparative experiment of the methods for 1-shot and 5-shot

方法	1 个支持集样本					5 个支持集样本				
	平均交并比/%				前景-背景 交并比/%	平均交并比/%				前景-背景 交并比/%
	Fold-1	Fold-2	Fold-3	均值		Fold-1	Fold-2	Fold-3	均值	
SCLNet	53.3	38.0	46.4	45.9	75.9	55.6	39.2	47.5	47.4	76.5
PFENet	66.1	44.2	55.7	55.3	72.9	67.2	45.9	56.1	56.4	74.9
HSNet	61.4	43.1	58.7	54.4	73.3	63.8	46.3	58.9	56.3	74.4
本文方法	68.1	43.7	60.0	57.3	73.9	68.3	43.4	58.0	56.6	74.8

注: 表中黑体数字表示所对比的指标达到最优。

3.6 消融实验

对多尺度解码器和注意力机制在 1-shot 和

5-shot 设定下进行消融实验, 如表 3 所示. 评价指标为 3 个集合平均交并比的均值和 3 个集合前

景-背景交并比的均值。

1) 基准:同时移除多尺度解码器和注意力机制.基准的平均交并比和前景-背景交并比均最低,在1-shot时,平均交并比和前景-背景交并比分别为51.8%和71.3%;在5-shot时,平均交并比和前景-背景交并比分别为53.8%和72.0%。

2) 基准+多尺度解码器:在基准方法上加入多尺度解码器.与基准方法相比,在1-shot时,平均交并比和前景-背景交并比分别提高4.4%和2.0%.在5-shot时,平均交并比和前景-背景交并比分别提高2.7%和1.7%。

3) 基准+注意力机制:在基准方法中嵌入注意力机制.与基准方法相比,在1-shot时,平均交并比和前景-背景交并比分别提高2.1%和0.7%。

在5-shot时,平均交并比和前景-背景交并比分别提高1.0%和0.8%。

4) 基准+多尺度解码器+注意力机制:将多尺度解码器和注意力机制同时加入基准方法,其平均交并比和前景-背景交并比均最优.在1-shot时,平均交并比和前景-背景交并比分别提高5.5%和2.6%.在5-shot时,平均交并比和前景-背景交并比分别提高2.8%和1.8%。

分析表3结果可知,多尺度解码器和注意力机制均能提升方法性能.多尺度解码器能够建立不同尺度下的特征联系,聚合不同尺度特征图的缺陷信息,有效提升了方法的分割性能.注意力机制提高本文方法对局部前景特征的代表能力,抑制了背景信息的干扰。

表3 1个支持集样本和5个支持样本模块消融实验
Table 3 Ablation experiment of the modules for 1-shot and 5-shot

方法	1个支持集样本		5个支持集样本	
	平均交并比/%	前景-背景交并比/%	平均交并比/%	前景-背景交并比/%
基准	51.8	71.3	53.8	72.0
基准+多尺度解码器	56.2	73.3	56.5	73.7
基准+注意力机制	53.9	72.0	54.8	72.8
基准+多尺度解码器+注意力机制	57.3	73.9	56.6	73.8

注:表中黑体数字表示所对比的指标达到最优。

3.7 定性结果

可视化9类带钢表面缺陷的分割结果,如图

7所示.可视化结果表明,本文所提方法能够准确地分割出缺陷的位置和类别。

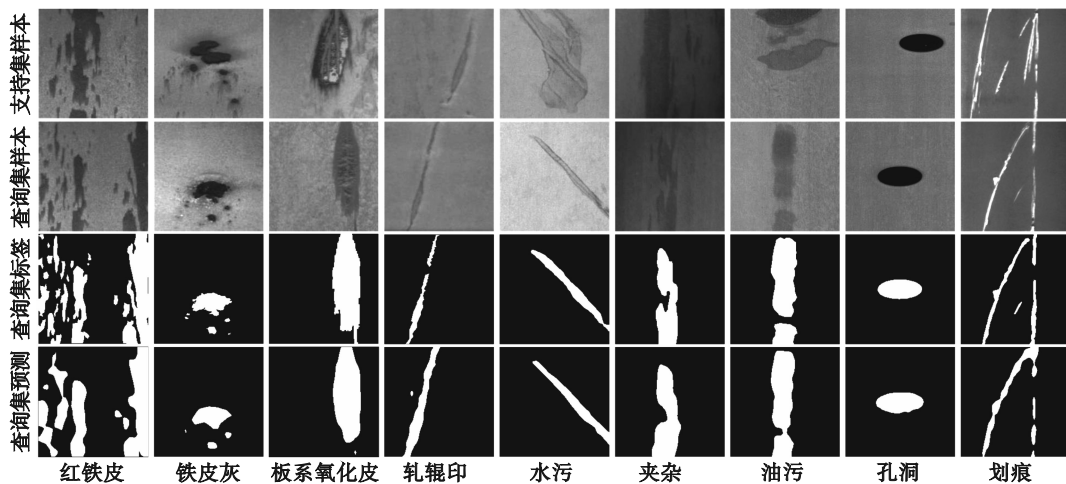


图7 带钢表面缺陷小样本语义分割定性结果

Fig. 7 Qualitative results of strip steel surface defects' few-shot semantic segmentation

4 结 论

本文提出了一种小样本语义分割深度学习方法,用于解决在实际工业生产中因缺陷样本难以采集和样本细粒度标注成本高昂所带来的

问题.本文结合生产实际,构建了新的带钢表面缺陷分割数据集,并在该数据集上进行了对比实验,以验证本文方法的性能优势.通过消融实验,验证了注意力机制和多尺度解码器在解决缺陷问题时的积极作用.将本文方法应用在生

产现场,可以实时分割轧线上可能出现的表面缺陷。

参考文献:

- [1] Song G R, Song K C, Yan Y H. 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.
- [2] Song K C, Yan Y H. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects [J]. *Applied Surface Science*, 2013, 285: 858-864.
- [3] Zhou D H, Wei M H, Si X S. A survey on anomaly detection, life prediction and maintenance decision for industrial processes [J]. *Acta Automatica Sinica*, 2013, 39(6): 711-722.
- [4] Luo Q W, Fang X, Su J, et al. Automated visual defect classification for flat steel surface: a survey [J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(12): 9329-9349.
- [5] 周鹏,徐科,杨朝霖.基于SIFT的中厚板表面缺陷识别方法[J].清华大学学报(自然科学版),2018,58(10): 881-887.
(Zhou Peng, Xu Ke, Yang Chao-lin. Surface defect recognition for moderately thick plates based on a SIFT operator [J]. *Journal of Tsinghua University (Science and Technology)*, 2018, 58(10): 881-887.)
- [6] Chu M X, Gong R F, Gao S, et al. Steel surface defects recognition based on multi-type statistical features and enhanced twin support vector machine [J]. *Chemometrics and Intelligent Laboratory Systems*, 2017, 171: 140-150.
- [7] 张德富,宋克臣,牛孟辉,等.基于一维卷积的生产线冷态重轨表面缺陷快速检测[J].东北大学学报(自然科学版),2021,42(2):276-281.
(Zhang De-fu, Song Ke-chen, Niu Meng-hui, et al. Rapid detection of cold heavy rail surface defects of production line based on one-dimensional convolution network [J]. *Journal of Northeastern University (Natural Science)*, 2021, 42(2): 276-281.)
- [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方法[J].东北大学学报(自然科学版),2021,42(5):687-692.
(He Yu, Song Ke-chen, Zhang De-fu, 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] Bao Y Q, Song K C, 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: 1-11.
- [10] Cha Y J, Choi W, Suh G, et al. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types [J]. *Computer Aided Civil and Infrastructure Engineering*, 2018, 33(9): 731-747.
- [11] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [C]//Conference on Computer Vision and Pattern Recognition (CVPR). Boston, 2015: 3431-3440.
- [12] Zhang B F, Xiao J M, Qin T. Self-guided and cross-guided learning for few-shot segmentation [C]//Conference on Computer Vision and Pattern Recognition (CVPR). Kuala Lumpur, 2021: 8312-8321.
- [13] Tian Z T, Zhao H S, Shu M, et al. Prior guided feature enrichment network for few-shot segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(2): 1050-1065.
- [14] Min J H, Kang D, Cho M. Hypercorrelation squeeze for few-shot segmentation [C]//International Conference on Computer Vision (ICCV). Montreal, 2021: 6941-6952.
- [15] Dong N Q, Xing E P. Few-shot semantic segmentation with prototype learning [C]//British Machine Vision Conference (BMVC). Newcastle, 2018: 1-13.
- [16] Feng X L, Gao X W, Luo L. X-SDD: a new benchmark for hot rolled steel strip surface defects detection [J]. *Symmetry*, 2021, 13(4): 706-723.
- [17] Lyu X, Duan F, Jiang J, et al. Deep metallic surface defect detection: The new benchmark and detection network [J]. *Sensors*, 2020, 20(6): 1562-1578.
- [18] Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge [J]. *International Journal of Computer Vision*, 2015, 115(3): 211-252.
- [19] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, 2016: 770-778.
- [20] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision (ECCV). Munich, 2018: 3-19.