

优化 Faster R - CNN 算法的 小样本缺陷检测研究

何军红,温观发,黎长鑫

(西北工业大学 航海学院,陕西 西安 710072)

摘要:随着自动化检测技术的发展,基于深度学习的缺陷检测技术以其高精度、高效率、非接触性的特点正逐渐成为工业和学术领域的研究热点。为解决实际工业生产中由于产品缺陷数据集样本不足、类别不均衡导致的模型过拟合、检测精度低等问题,提出了一种基于 Faster R - CNN 算法框架优化的缺陷检测模型 Faster R - CNN - H - BFC,通过基于多层感知器(multi - layer perceptron, MLP)实现的幻觉网络能够从具有丰富样本的基类中学习类共享特征并为新类生成额外的幻觉样本供模型训练,并且针对 Faster R - CNN 本身存在的识别精度低以及检测效果差等问题,将原始的 VGG16 主干网络替换为具有残差结构的 ResNet50,并引入了特征金字塔网络(Feature Pyramid Networks, FPN)实现多尺度特征融合,添加混合注意力机制(Convolutional Block Attention Module, CBAM)来增强模型的特征提取能力。实验和数据表明:改进后的缺陷检测模型在极少样本场景下具有较好的检测效果,平均检测精度相较于改进前提升了 3.11%。

关键词:小样本;Faster R - CNN;幻觉网络;特征金字塔网络;注意力机制;缺陷检测

中图分类号:TP391.4

文章编号:1000 - 0682(2024)05 - 0094 - 08

文献标识码:A

DOI:10.19950/j.cnki.CN61 - 1121/TH.2024.05.018

Research on small sample defect detection based on optimized Faster R - CNN algorithm

HE Junhong, WEN Guanfa, LI Changxin

(School of Marine Science and Technology, Northwestern Polytechnical University, Shaanxi Xi'an 710072, China)

Abstract: With the development of automatic detection technology, defect detection technology based on deep learning is gradually becoming a research hotspot in industrial and academic fields due to its high precision, high efficiency and non - contact characteristics. In order to solve the problems of model overfitting and low detection accuracy caused by insufficient samples and unbalanced categories of product defect data sets in actual industrial production, a defect detection model Faster R - CNN - H - BFC based on Faster R - CNN algorithm framework optimization is proposed. The illusion network based on multi - layer perceptron (MLP) can learn class sharing features from base classes with rich samples and generate additional illusion samples for new classes for model training. Aiming at the problems of low recognition accuracy and poor detection effect of Faster R - CNN, the original VGG16 backbone network is replaced with ResNet50 with residual structure, and Feature Pyramid Networks (FPN) is introduced to realize multi - scale feature fusion, and the Convolutional Block Attention Module (CBAM) is added to enhance the feature extraction ability of the model. Experiments and data show that the improved defect detection model has better detection effect in very few sample scenarios, and the average detection accuracy is 3.11% higher than that before improvement.

Keywords: small - sample; Faster R - CNN; illusion network; feature pyramid network; attention mechanism; defect detecting

0 引言

在工业质检领域,产品表面缺陷检测是一项至关重要的任务^[1]。传统的缺陷检测方式高度依赖专业人员的手动操作,由于质检人员的主观判断存在差异,产品检测的精度难以保证。此外,这种方式存在检测效率低下和人工成本高等问题^[2]。而随着计算机图像处理技术的快速发展,基于深度学习的产品表面缺陷检测技术以其无损、高效、高精度的特点在工业质检中备受关注。然而,深度学习模型的卓越性能往往依赖于大量可用的训练样本,但在工业生产中,缺陷样本受良品率限制,导致实际应用的检测模型可能因为缺少样本而出现过拟合,检测精度低等问题,因此,如何在训练样本有限的条件下,有效解决过拟合问题,使得神经网络模型在缺陷检测任务中保持高检测精度和效率,成为了当前研究的热点与难点。为了攻克这一难题,研究者们不断尝试新的方法和技术,如数据增强、迁移学习、模型优化等,以提升模型在少量样本下的泛化能力和检测性能。

小样本学习的思想最初由 Li 等人^[3]于 2003 年提出,其核心理念在于,即使在少量样本的情况下,计算机视觉模型也依然能够通过已有的先验知识来有效的学习并识别新的类别。目前,常见的小样本目标检测方案^[4]可分成 4 种方法,即元学习(Meta-learning)、迁移学习(Transfer Learning)、数据增强和度量学习(Metric Learning)。

Yan 等^[5]在研究中引入预测模型网络,通过对数据集中的各类目标进行学习,成功获取了各类注意力向量并作为元知识,将其与所需识别的图像通过区域检测网络(Region Proposal Net, RPN)进行通道级融合,实现相应检测图的生成。

Jing 等^[6]将迁移学习应用于布匹缺陷的检测,先将检测模型在 MNIST 数据集上进行预训练,并设计一种基于局部特征检索的检测方法,该方法成

功应用于布匹缺陷的检测,实现了对布匹缺陷的高效、准确识别。

Valente 等^[7]为解决印刷品表面缺陷数据匮乏的问题,基于数据增强思想,将分布先验知识用于合成印刷品表面缺陷数据,并将合成的数据与原始数据进行混合,进一步丰富样本数据的多样性。

Yang 等^[8]提出的 NP-RepMet 通过计算每个候选框的正负样本特征之间的最小距离,将其与学习到的各类别的正负样本最小距离结合,实现对目标的有效检测。

该文针对实际生产中存在的小样本问题,提出了一种基于 Faster R-CNN 算法框架的缺陷检测模型 Faster R-CNN-H-BFC。在原模型的基础上,从数据增强和模型优化两方面进行设计。该模型引入幻觉网络进行数据增强,利用 ResNet 网络提升模型深度,结合特征金字塔网络(Feature Pyramid Networks, FPN)实现多尺度特征的融合,并添加混合域注意力机制(Convolutional Block Attention Module, CBAM)增强模型对于缺陷特征的关注。实验结果表明,改进后的缺陷检测模型在极少样本场景下具有较好的检测效果,模型的检测精度具有明显的提高。

1 Faster R-CNN 算法框架概述

Faster R-CNN 主要由 3 个核心组件构成:特征提取网络使用 VGG16 分类网络用于提取输入图像的特征信息,从而得到共享的特征图;RPN 则基于锚框机制生成一系列覆盖图像中每个目标的尺度不一的候选框;检测网络则通过感兴趣区域(Region of Interest, ROI)池化层进行统一尺度,并通过全连接层实现分类以及回归预测,从而实现对目标的检测。Faster R-CNN 模型基本结构如图 1 所示。Faster R-CNN 网络将 RPN 网络的训练过程与目标检测网络相结合,实现了一种端到端的网络训练方式。这种设计不仅提升了目标检测任务的效率,还进一步增强了检测结果的准确性。

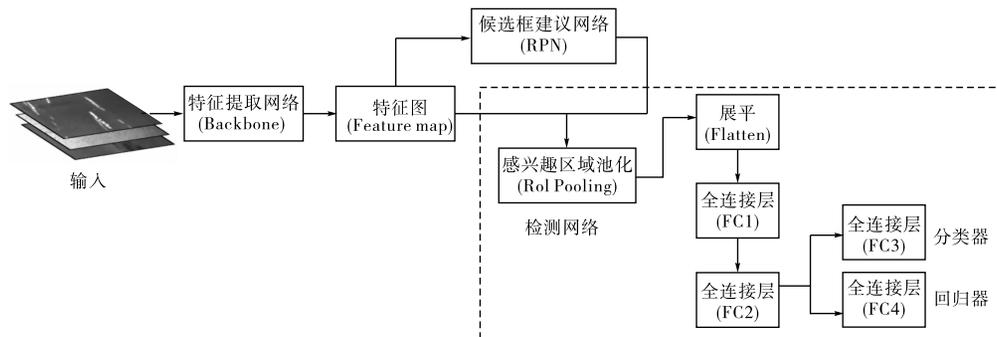


图 1 Faster R-CNN 模型基本结构示意图

2 基于 Faster R - CNN 的缺陷检测模型优化

2.1 基于多层感知器的幻觉网络构建与实现

多层感知器(multi-layer perceptron, MLP)的实质就是一个具备多层结构的前馈神经网络,其中隐藏层的个数可以根据实际的应用场景进行设置,但是最少有一层,整个网络通过每一层对上一层的信号进行处理,并将处理后的输出信号传递给下一层,从而构建出一个逐层连接的网络结构,这样多层感知器就可以通过前向传播以及反向传播不断的更新网络中的权重参数,实现分类、预测以及回归等任务。

Wang 等人^[9]基于元学习的思想以及借鉴人类的视觉系统,提出了通过幻觉器生成额外的训练样本,并将其结合元学习器联合训练,得到一个应用于小样本场景的分类模型,并在 ImageNet 数据集上取得不错的检测效果。幻觉网络的应用如图 2 所示。原始图像输入网络中后,经过 RPN 网络得到一些候选框,这些候选区域经过幻觉网络后生成幻觉样本。

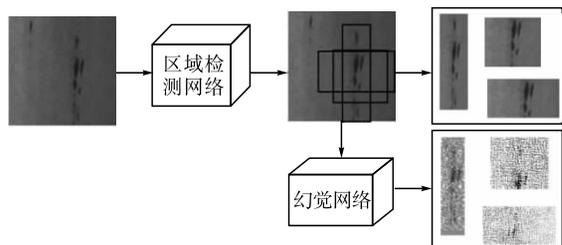


图 2 幻觉网络应用目标检测框架示意图

对于幻觉网络的结构设计,相关研究使用循环神经网络^[10]、记忆增强网络^[11]或者多层感知器。该文采用的幻觉网络参考 Wang 等人^[9]提出的幻觉器设计,使用一个以 ReLU 为激活函数的三层 MLP

构成幻觉网络,其中 ReLU 激活函数作为一种常用的激活函数,只有在输入值大于 0 时才能输出,否则只能输出 0,因此并不存在饱和区域。这一特性有助于在一定程度上缓解梯度消失的问题,并确保幻觉的非负性,使得生成的样本和真实的例子一样。

幻觉网络 G 本质是通过学习基类的共享内部特征变化为新类生成额外的训练样本,样本生成公式如式(1)所示:

$$\tilde{x} = G(x, z; W_G) \tag{1}$$

其中, x 为输入样本; z 为随机噪声向量; W_G 为幻觉网络参数; \tilde{x} 为新生成的幻觉样本输出。幻觉样本生成的过程如图 3 所示,其可以概括为如下具体流程:首先是从训练样本集 S_{train} 中随机采样一个真实的样本种子 (x, y) 并采样一个噪声向量 z ,然后将 x 以及 z 传递给 G ,最后生成一个幻觉样本 (\tilde{x}, y) ,对每个类别都进行该操作最后得到生成的样本集 S_{train}^G ,将其和原始样本集结合得到一个增广后的训练数据集 $S_{train}^{aug} = S_{train} \cup S_{train}^G$ 。 y 表示样本的类别,由上述幻觉生成公式可知,生成的幻觉样本不会改变原始类别。

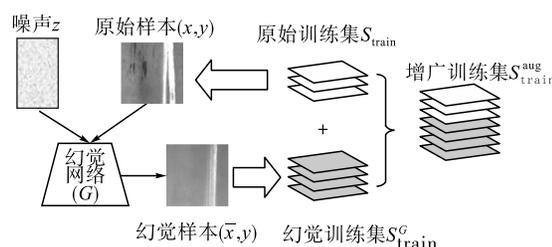


图 3 幻觉样本生成示意图

基于图 1 所示的 Faster R - CNN 缺陷检测模型,本节将幻觉网络集成到检测网络中并将其命名为 Faster R - CNN - H,其结构示意图如图 4 所示。

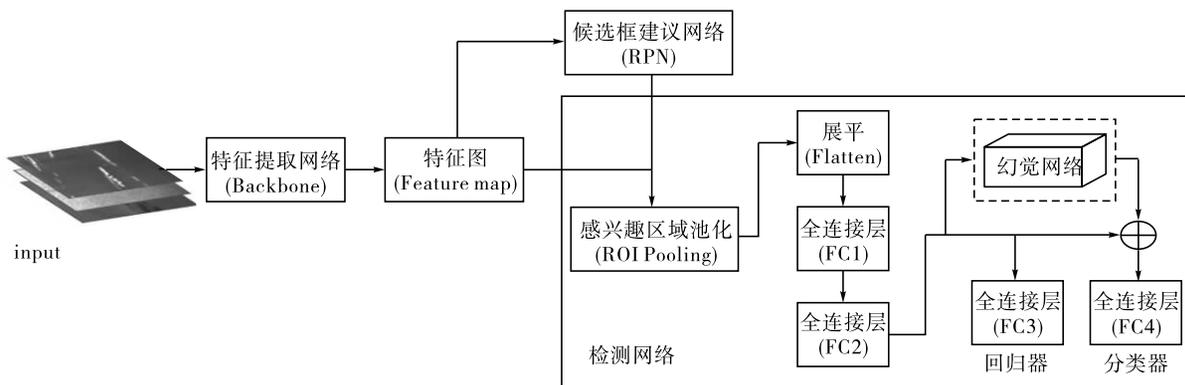


图 4 集成幻觉网络的缺陷检测网络

幻觉网络集成在 RoI head 中,具体插入全连接层之后,分类器之前。其中幻觉网络将可用的训练

样本作为输入并生成幻觉样本,生成的幻觉样本数据集将用于新类缺陷训练分类器,幻觉样本将不对

回归器产生任何影响。

2.2 Faster R – CNN 的 Backbone 改进

Faster R – CNN 原模型中使用 VGG16 作为主干网络,如图 5 所示就是 VGG16 网络的结构图,该网络使用 3 × 3 这种小卷积核替换 5 × 5 等大卷积核,从而减少模型训练参数,加快模型训练速度。

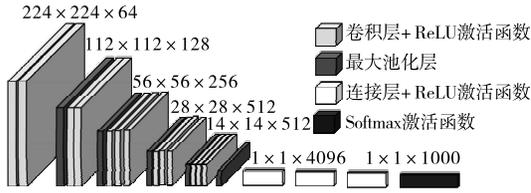


图 5 VGG16 结构图

尽管增加网络的深度通常有望提升训练效果,但随着层数的增加,梯度在反向传播的过程中可能逐层衰减(梯度消失)或放大(梯度爆炸),模型的训练往往达不到理想的效果。而残差网络通过引入残差连接,使得输入的信号可以直接传播到高层,从而避免了梯度消失或爆炸的问题。如图 6 所示就是深

层网络常用的残差单元结构,该结构将输入分成两条分支,主分支依然是一个卷积模型操作,并在头尾插入一个 1 × 1 的小卷积用于降维和升维,确保尺度不变的情况下减少模型参数;辅助分支则不进行任何卷积操作,只是保留原本的输入,最后将两条分支进行求和并用 ReLU 激活函数处理。

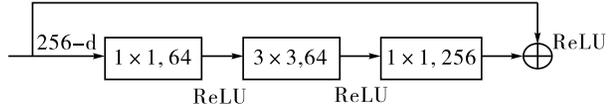


图 6 残差单元结构图

ResNet 网络嵌入了残差单元用于解决深度网络退化的问题,考虑工业生产环境中硬件条件限制,为了在实际的生产作业中更好的部署网络模型,该文将以 ResNet50 作为 backbone 替换 Faster R – CNN 中的原始主干网络 VGG16。ResNet50 的结构如图 7 所示,该网络主要由 conv1、conv2_x、conv3_x、conv4_x 以及 conv5_x 等五部分组成。

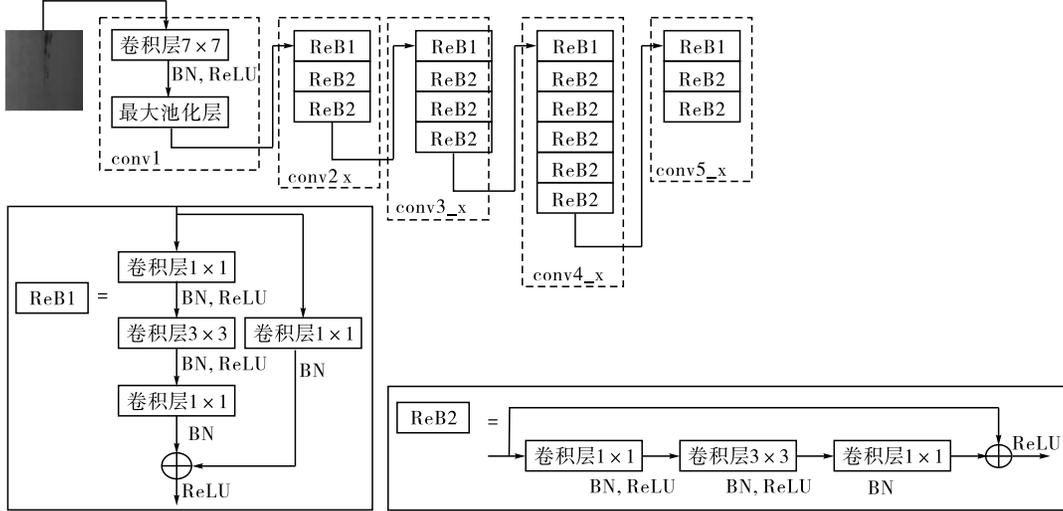


图 7 ResNet50 结构图

其中 conv1 部分只由一个卷积操作和最大池化层组成,可以看成该网络的预处理部分,其余 4 个部分基本类似,都是由 ReB1 和 ReB2 这种残差结构重复组成。

2.3 特征金字塔网络的引入

在卷积神经网络中,随着网络的深度不同,高层特征图和浅层特征图关注重点不同。高层特征图主要包含语义特征信息利于分类任务,但是缺乏空间信息,浅层特征图主要包含空间位置信息利于回归任务。特征金字塔则通过不同的连接方式将高层丰富的语义特征以及浅层丰富的空间信息进行融合,使得生成的特征表示更加丰富以及准确。其进行特

征融合主要可以分为如下 3 个部分:自底向上、自顶向下以及横向连接。

(1) 自底向上的结构本质就是一个卷积神经网络的前向传播过程,根据输入的特征图尺度大小逐层缩小一半,每层输出的最后一个特征图构成整个特征金字塔结构;

(2) 自顶向下的结构通过将顶层的特征图进行上采样(通常是使用最近邻插值的方法),使其尺寸与下一层的特征图相同,为后续的融合操作打下基础,提高了特征表示能力,为后续的目标检测任务提供了更强大的特征支持;

(3) 横向连接的结构如图 8 所示,为了实现特

征图之间的融合,将底层的特征图通过一个 1×1 的普通卷积操作使得输出的特征图和在自顶向下的结构中通过上采样输出的特征图通道数一致。

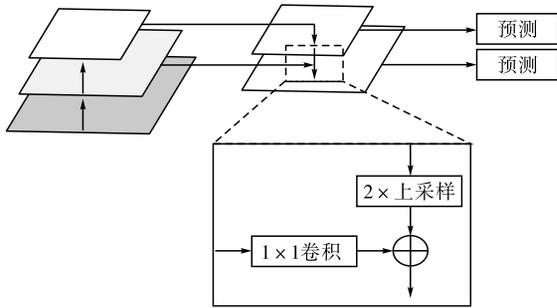


图 8 FPN 特征图融合示意图

为了提升缺陷检测模型的特征提取能力,故将 FPN 融合到特征提取网络中实现多尺度特征融合。如图 9 所示即为 ResNet50 融合 FPN 示意图。

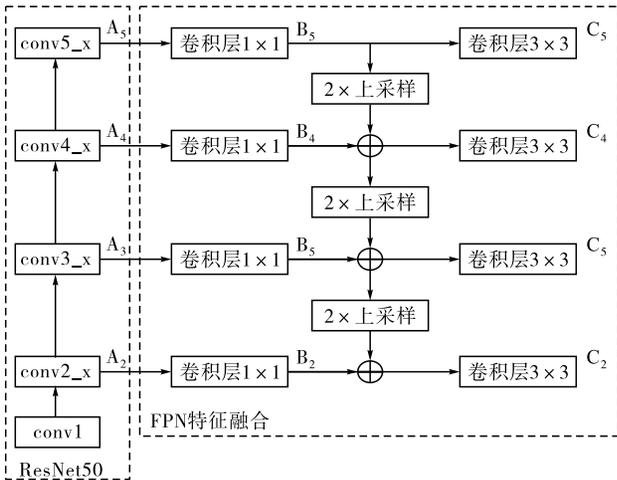


图 9 ResNet50 融合 FPN 示意图

ResNet50 网络由 conv1、conv2_x、conv3_x、conv4_x、conv5_x 这 5 个部分组成,采用特征金字塔进行多尺度融合的过程主要可以分为如下几个步骤:首先是 ResNet50 自底向上的卷积操作,得到尺度以及通道数不同的特征 $\{A_2, A_3, A_4, A_5\}$; 然后这些特征自上而下进行特征融合,先通过一个 1×1 的卷积操作得到特征 $\{B_5, B_4, B_3, B_2\}$,从 B_4 特征开始,和上一层的经过 2 倍上采样后的特征进行融合,再通过 3×3 的卷积操作消除不同层的混叠效果,最后将 FPN 得到的所有特征图输入到 RPN 网络。从而完成整个高层特征和低层特征融合的操作,提升模型的特征提取能力。

2.4 CBAM 混合注意力机制的添加

CBAM 注意力机制^[12]通过通道注意力模块学习图像中目标的类型标签特征信息,然后利用空间

注意力模块学习输入图像中目标的位置信息,再将这 2 个特征信息进行融合从而增强网络的特征提取能力,其示意图如图 10 所示。由于这 2 个模块是相对独立的,故可以根据需要选择各自的权重。基于该文使用的带钢缺陷数据集中:缺陷背景灰暗,容易和缺陷混淆。因此选择 CBAM 注意力机制引入到该文使用的缺陷检测模型,提升缺陷检测模型对图像中特征提取能力。

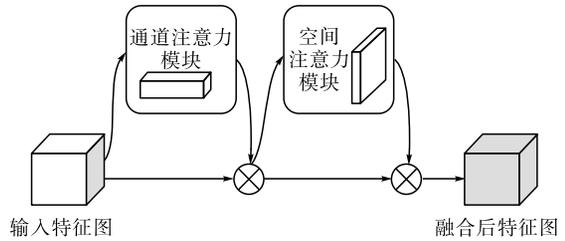


图 10 CBAM 注意力机制

CBAM 注意力模块可以实现即插即用的效果,故可以集成到该文使用的缺陷检测网络中。为了降低背景和噪声等干扰,提升模型的特征提取能力,故选择将 CBAM 注意力模块加入 backbone 中的 conv1 以及 conv5_x 末尾,增强模型对于高层和底层的特征提取能力。主干网络集成 CBAM 模块结构图如 11 所示。

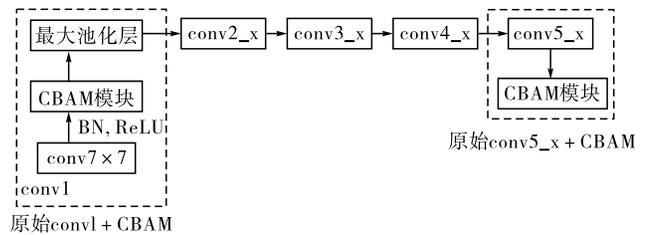


图 11 ResNet50 网络集成 CBAM 模块结构图

将 CBAM 模块分别置于底层和高层,一方面可以通过设置不同的权重使得 CBAM 中的通道注意力模块重点关注高层的语义信息,空间注意力模块重点关注底层的位置信息。另一方面在 FPN 进行多尺度特征融合前加入注意力机制,可以先对特征图进行特征增强,再实现特征融合,这样可以进一步的提升模型的特征提取能力。

3 实验设计与结果分析

3.1 数据集与实验环境

3.1.1 小样本数据集及数据处理

根据小样本学习任务的定义,小样本数据集以 N -way K -shot 任务形式来构建,其中 N 代表有多

少缺陷类别, K 表示每个缺陷类别的样本数量。该文采用的小样本数据集是 Severstal 带钢缺陷数据集, 该数据集包含 4 种类型的带钢缺陷, 分别是: 裂纹、划痕、凹坑以及氧化铁皮。部分缺陷样本如图 12 所示。

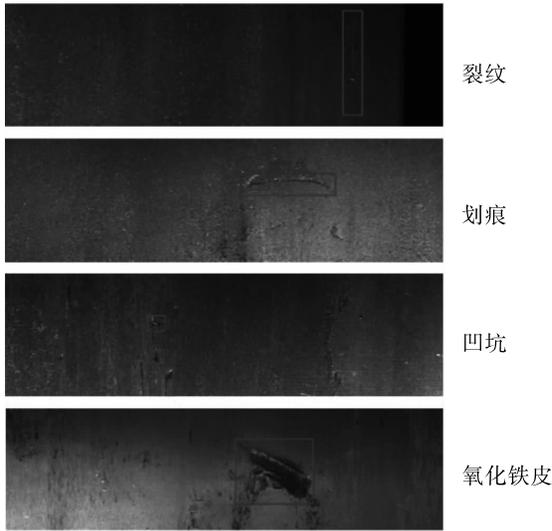


图 12 带钢表面部分缺陷示意图

该数据集中训练集包含了 12568 张缺陷图片, 但是可用于模型训练的已经标注好的缺陷样本图像只有 7095 张, 测试集包含了 5506 张缺陷图片。对该数据集进行分析可以了解到数据集中每种缺陷样本数量分布不平衡, 其中划痕缺陷样本占比达到了约 73%, 在所有标注样本中有 5150 张, 其占据了所有缺陷样本中的大部分比重, 而裂纹缺陷最少只有 247 张, 占比只有约 4%。带钢缺陷数据集上不同缺陷样本类别的数量占比如图 13 所示。

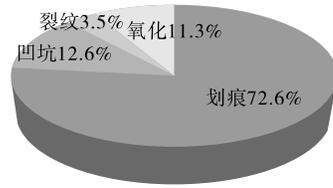


图 13 带钢缺陷不同类别占比分布

这种样本数量分布不均的情况恰好契合小样本学习任务中将训练集划分为丰富的基类以及稀缺的新类设定, 为此将样本数量最多的前 2 种 (划痕以及凹坑) 作为具有丰富样本的基类数据集 D_{base} , 将数量相对较少的裂纹以及氧化铁皮作为新类数据集 D_{novel} 。根据 $N - way K - shot$ 任务定义实验中的 N 为 2, 而 K 则按照 1、2、3、5、10 进行设置实现微调训练, 从而构成一个带钢缺陷小样本数据集。

3.1.2 实验环境设置

实验环境的具体配置信息如下表 1 所示。

表 1 实验环境配置信息

名称	型号或版本
操作系统	Ubuntu18.04
CPU	Intel(R) Core(TM) i7-10870H
GPU	NVIDIA GeForce RTX 2070
内存	32G
Pytorch	1.8
Python	3.8
CUDA	10.2

3.2 模型训练与测试过程

3.2.1 训练流程设置

实验基于 Fast R-CNN 检测模型通过两阶段微调的方式训练检测模型。本实验使用的两阶段微调训练模型流程如图 14 所示。

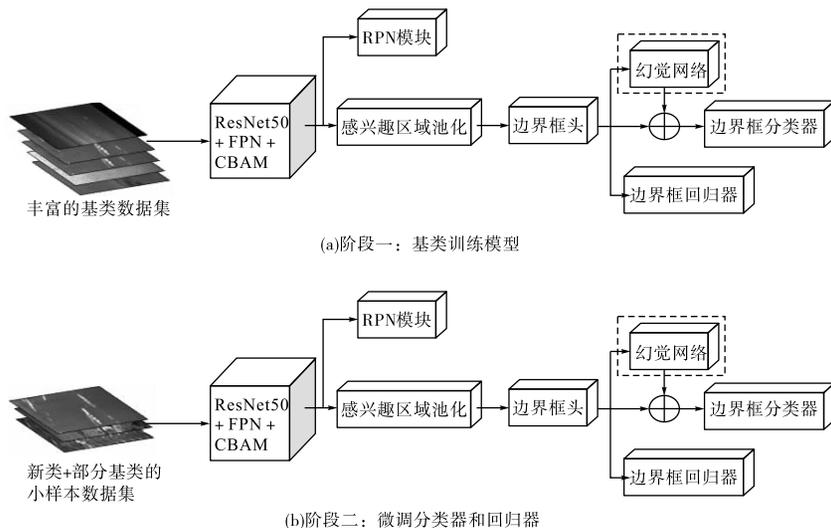


图 14 模型训练流程示意图

基于小样本学习的两阶段微调将模型的训练分成两个阶段:第一个阶段主要是通过丰富的基类对检测模型进行训练,该阶段又分为 2 个步骤,先是冻结幻觉网络,对没有幻觉网络的检测模型进行训练,然后是冻结除幻觉网络的其余模块,利用已经在基类上训练好的分类器对幻觉网络进行训练;第二阶段则是冻结处理分类器以及回归器的其余模块(上图中阶段二的灰色模块),通过新类数据部分基类数据构成的小样本数据集对分类器以及回归器进行微调。

3.2.2 训练策略设置

训练策略的相关参数如下所示:优化器选择带有动量方法的随机梯度下降算法(Stochastic Gradient Descent, SGD);基类训练的学习率(Learning rate)为 0.01,新类微调的学习率为 0.001;动量(Momentum)为 0.915;权重衰减(Weight decay)系数为 0.0005;批次大小为 4。

3.3 实验结果对比分析

3.3.1 幻觉网络的有效性验证

针对小样本微调方案对于带钢缺陷小样本场景的应用进行了对比实验,实验设置分别在 1-shot、2-shot、3-shot、5-shot 以及 10-shot 上进行了验证,其实验结果如表 2 所示,表中的数值表示的是 *mAP*,实验结果对比如图 15 所示。

表 2 幻觉网络对于小样本学习的有效性实验结果表

方案 K-shot	1	2	3	5	10
Faster R-CNN	31.61	36.32	52.36	56.14	63.47
Faster R-CNN-H	37.57	41.73	52.41	55.61	63.35

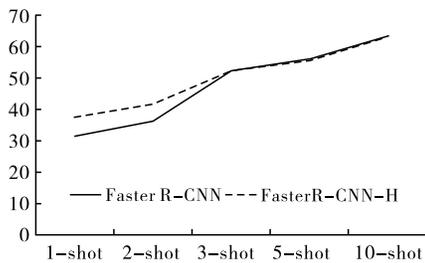


图 15 添加幻觉网络实验结果对比图

根据实验结果分析,幻觉网络的加入在 1-shot 下对于网络的 *mAP* 提升最大,提高了 5.96%,随着 shot 的增加,幻觉网络的效果有所下降,尤其是到 5-shot 开始逐渐有了抑制效果,添加了幻觉网络的检测模型反而降低了 *mAP*。这说明了幻觉网络在训练样本很少的情况下对检测性能的增益比较大,但是随着训练样本数量的增加,比如从每类缺陷中只有 1 个样本扩增到有 10 个样本,性能增益会不断

的减少,甚至不再是有益的。

3.3.2 不同改进模块的有效性验证

针对不同改进方案,在 Faster R-CNN-H 检测模型的基础上通过带钢缺陷数据集进行消融实验设计,分别增加一种改进方法进行对照实验,验证改进方法对于模型的改进效果。具体的实验结果如表 3 所示。其中“—”表示没有添加该模块,“√”表示添加了该模块。

表 3 不同改进方案的消融实验结果表

方案	ResNet50	FPN	注意力机制	<i>mAP</i>
Faster R-CNN-H	—	—	—	63.35
改进 1	√	—	—	64.26
改进 2	√	√	—	65.80
改进 3	√	√	√	66.58

从实验结果可知:

改进 1 使用 ResNet50 主干网络替换原始的 VGG16,残差结构使得改进后的 backbone 拥有更深的卷积层,使得算法的 *mAP* 提高了 0.91%;

改进 2 使用 FPN 进行多尺度的特征融合,使得特征图融合了高层语义信息以及底层的位置信息,从而增强模型的特征提取能力,使得算法的 *mAP* 相对改进 1 提高了 1.54%;

改进 3 引入注意力机制 CBAM 模块,使得模型从空间以及通道两个方面充分关注图像上的缺陷目标特征,使得算法的 *mAP* 相比于改进 2 提高了 0.78%。

综合整个消融实验可以看到这些改进都对提高带钢缺陷的检测精度有一定的积极作用,相比于原始的 Faster R-CNN-H 模型的 *mAP* 一共提高了 3.23%。

4 结论

针对小样本学习场景下的缺陷检测任务,该文提出 Faster R-CNN-H-BFC 模型,通过引入幻觉网络扩增样本,同时替换 ResNet50 为主干网络、结合特征金字塔实现多尺度特征融合以及嵌入 CBAM 注意力机制模块等方式,解决了样本不足、样本分布不均和背景干扰等问题,显著提升了模型的检测精度。实验结果验证了各项改进的有效性,但考虑到实际工业生产中硬件性能受限的情况,为了使缺陷检测模型可以更好的落地于实际生产,需要对检测模型进行轻量化设计。故后续的研究可以着重于如何减少模型参数,提升模型的训练速度以及模型的推理速度等。

