

基于生成式数据增强与 Faster-RCNN 改进的发动机打刻面缺陷检测

谭建辉 张峰

中山大学计算机学院 广州 510000

摘要 汽车发动机的打刻面具有承载发动机信息、丢失查找以及防止私自拆改发动机等作用,打刻面的质量将直接决定车辆是否能正常上牌行驶。但是在汽车制造领域,现阶段对打刻面的缺陷检测主要采用人工目视检测的方法,存在漏检的风险。虽然业界已有一些针对表面缺陷检测的研究,但它们无法完全适应发动机打刻面缺陷检测的特点,容易出现错检、漏检情况。为了革新发动机打刻面缺陷检测的方式,提出了一种基于生成式数据增强与 Faster-RCNN 改进的缺陷检测方法。首先,针对发动机打刻面缺陷样本少的小样本问题,提出了一种基于 stable diffusion 的打刻面缺陷图片生成方法。该方法通过两个掩膜图分别控制缺陷生成的位置以及恢复图像的字符特征,从而完成打刻面缺陷样本图像的生成,实现数据集的增强。其次,提出了一种同步双向融合特征金字塔网络替换原模型所使用的特征金字塔网络,增强多尺度特征融合能力,解决打刻面缺陷目标尺度范围广的问题。实验结果表明,所提出的方法在检测发动机打刻面缺陷时,均值平均精度(mAP)达到了 97.52%,相比原始的 Faster-RCNN 模型提高了 34.73%,可以满足发动机打刻面缺陷的检测需求。

关键词: 缺陷检测; 图像生成; 数据增强; Faster-RCNN; 多尺度特征融合

中图分类号 TP391

Defect Detection of Engine Engraved Surface Based on Generative Data Augmentation and Improved Faster-RCNN

TAN Jianhui and ZHANG Feng

School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510000, China

Abstract The engraved surface of automotive engine has the functions of carrying engine information, searching for lost, and preventing unauthorized disassembly and modification of the engine. The quality of the engraved surface will directly determine whether the vehicle can be registered and driven normally. However, in the field of automobile manufacturing, manual visual inspection is mainly used for defect detection of engraved surfaces at present, which poses a risk of missed detection. Although there have been some studies on surface defect detection in the industry, they cannot fully adapt to the defect detection of engine engraved surface, which can easily lead to false positives and false negatives. In order to innovate the method of detecting engine engraved surface defects, this paper proposes a defect detection method based on generative data augmentation and improved Faster-RCNN. Firstly, a method for generating engraved surface defect images based on stable diffusion model is proposed to address the problem of limited samples of engine engraved surface defects. This method controls the location of defect generation and restores the character features of the image through a dual mask image, thereby completing the generation of engraved surface defect images and achieving data augmentation of the dataset. Secondly, a synchronous bidirectional fusion feature pyramid network (SBFFPN) is proposed to replace the feature pyramid network (FPN) used in the original algorithm, enhancing the multi-scale feature fusion capability and solving the problem of wide target scale range of engraved surface defects. The experimental results show that the proposed method achieves mAP of 97.52% in detecting engine engraved surface defects, which is 34.73% higher than the original Faster RCNN model and can meet the detection requirements of engine engraved surface defects.

Keywords Defect detection, Image generation, Data augmentation, Faster-RCNN, Multi-scale feature fusion

1 引言

在汽车生产制造标准中,国标文件 GB 7258 中明确提出车辆需要将型号和出厂编号等信息打刻在发动机上,用以标注发动机的生产厂家、规格、性能、特点、工艺、用途和产品批次等信息,同时也便于丢失查找和防止私自拆改发动机。若发动机打刻面出现质量问题,很可能被车辆管理部门认定为信息不全或私自篡改发动机而导致车辆无法上牌行驶。因此,打刻面表面缺陷的精确检出在发动机生产过程中至关

重要。而现阶段,打刻面的表面缺陷检测完全依赖传统的人工检查,工作人员需要在流动的生产线上目视检查产品的打刻面,快速判断是否存在缺陷,从而挑选出打刻面存在缺陷的产品进行报废或者返修。这种传统的缺陷检查方式效率极其低下,且存在很大的漏检风险,因此本文将探讨采用基于深度学习的图像识别技术来取代人工目视检查的方式。

打刻面的缺陷检测属于金属表面缺陷检测。对于金属表面缺陷的检测,目前已有一些研究工作,主要分为如下两类。

第一类是基于纹理特征或形状特征的传统表面缺陷检测

方法。Yun 等^[1]提出了一种基于 Gabor 滤波器进行纹理特征检测的算法用于钢胚表面缺陷检测,该算法使用了单变量动态编码搜索算法对 Gabor 滤波器进行优化。Liu 等^[2]提出了一种改进的多块局部二值模式算法,通过改变块大小找到合适的尺度来描述缺陷的纹理特征,实现了对钢材表面缺陷 94.3% 的识别准确率。Wang 等^[3]提出了一种磁体表面切割缺陷的检测方法,该方法采用傅里叶变换和霍夫变换对磁体表面图像进行重构,通过比较重构图像与原始图像的灰度差异得到缺陷信息。尽管上述传统的缺陷检测算法随着研究的推进,识别准确率不断提升,但是随着工业自动化不断深化,金属表面缺陷检测的应用场景越来越多样化,需要手工设计缺陷特征描述的传统方法无法满足如此复杂且无规则的缺陷识别。

第二类是采用基于深度学习自动提取的特征进行金属表面缺陷检测的方法。常见的基于缺陷检测网络主要有两种类型,一种是以 SSD^[4]或 YOLO^[5]为代表的一阶段检测模型,另外一种是以 Faster-RCNN^[6]为代表的两阶段检测模型。Chen^[7]等针对高铁接触网支撑装置上的紧固件的缺陷检测改进了 SSD 网络,级联了 3 个基于深度卷积神经网络的检测阶段,采用不同层的特征图进行目标检测,在复杂环境中具有良好的适应性和鲁棒性,检测率高。为了减少网络计算量,提高网络检测速度,Liu 等^[8]在对热轧钢的缺陷检测场景中对 SSD 网络引入了注意力机制和多特征融合网络,取得了不错的效果。Li 等^[9]采用了由 27 卷积层组成的改进 YOLO 网络对带钢表面缺陷进行识别,通过融合浅层特征提升了对细小缺陷的检测能力。Wang 等^[10]针对金属表面缺陷检测,设计了一种改进的 YOLOv5 检测网络,使用特定的数据增强方法进行正则化,并使用有效的损失函数来解决小尺度对象缺陷引起的数据不平衡问题。改进网络在 GC10-DET 数据集上取得了不错的精度。针对小目标缺陷检测场景,Ding 等^[11]提出了 TDD-Net 网络,通过 K 均值聚类设计合理的锚框并引入了多尺度金字塔网络 FPN^[12],使得网络 Faster-RCNN 适应微小的缺陷检测。Yin 等^[13]在钢表面缺陷检测中,采用多层 RoI 对齐对 Faster-RCNN 网络进行优化,并采用循环 GAN 进行数据增强,最终取得了不错的检测性能。上述的改进算法虽然在不断提升金属表面缺陷检测的准确率,但在发动机打刻面缺陷识别准确率上仍然具有不小的挑战。首先,基于发动机打刻面质量的重要性以及现今工业制造水平不断提升的原因,在最终检查工序上出现打刻面缺陷的产品是罕见的,因此缺陷样本的图像会非常有限,这对于缺陷检测网络的训练是一个严峻的挑战。其次,打刻面上的缺陷目标尺度范围非常广,既有极其微小的缺陷,如脱皮、细微划痕等,也有横跨整个打刻面的缺陷,如污迹、长划痕等,所以如何处理这个问题也是一个挑战。

鉴于两阶段模型在检测准确率上的优势,本文将选择 Faster-RCNN 作为本文研究的基础模型,并针对上述的发动机打刻面缺陷检测的难点提出了以下的解决方案。

1) 针对缺陷样本稀缺的小样本问题,提出了一种基于 stable diffusion 的打刻面缺陷图片生成方法。通过在正常样本图片的特定位置上生成缺陷的方式实现缺陷数据集的扩充,以达到数据增强的目的。

2) 针对打刻面缺陷目标尺寸范围广的问题,提出了一种

双向融合特征金字塔网络(Synchronous Bidirectional Fusion Feature Pyramid Network,SBFFPN),用以替换原模型所使用的特征金字塔网络(Feature Pyramid Network,FPN),实现自顶向下以及自底向上的双向同步特征融合,从而优化多尺度特征融合效果,扩大缺陷目标尺寸的检测范围。

2 相关工作

2.1 Stable diffusion 模型

Stable diffusion 模型是在 Rombach 等^[14]提出的潜在扩散模型的基础上发展而来的,采用了更加稳定、可控和高效的方法来生成高质量图像。Stable diffusion 模型整体是一个端对端的模型,主要由 VAE,U-Net 和 CLIP Text Encoder 这 3 个核心组件构成。VAE 是一种基于编码器-解码器结构的变分自编码器,它的 Encoder(编码器)结构能够将输入的图像转换为低维潜层的特征,Decoder(解码器)结构能将低维潜层特征重建还原回像素级的图像。U-Net 结构能够预测噪声残差,stable diffusion 通过反复调用 U-Net,将预测出来的噪声残差从原来的随机高斯噪声矩阵中去除,这样就能够得到去除噪声后的图像潜层特征,再通过 VAE 中的 Decoder 结构就能重构像素集图像了,这也是 stable diffusion 模型图片生成过程中最核心之处。CLIP Text Encoder 的主要作用是将输入的文本进行编码,转换成文本的语义信息,并输入到 U-Net 结构中作为条件,对模型生成图像的内容进行一定程度的控制和引导。Stable diffusion 模型的架构图如图 1 所示。

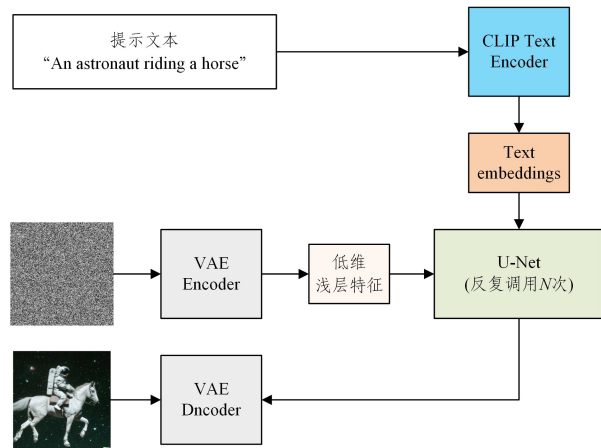


图 1 Stable diffusion 模型架构图

Fig. 1 Architecture of Stable diffusion model

2.2 Faster-RCNN 模型

Faster-RCNN 模型是在 R-CNN 基础上发展起来的目标检测模型,它的结构由特征提取网络 Backbone、区域建议网络 RPN、ROI 池化以及分类器 4 部分组成。特征提取网络使用多组基础的卷积层、池化层以及激活层提取出输入图像的特征图,用于后续的区域建议网络以及分类。常见的用于 Faster-RCNN 的特征提取网络有 AlexNet, Inception, VGG16, ResNet, MobileNet 等。区域建议网络的作用是在 Backbone 提取的特征图上生成一系列的候选框,用于最终的目标检测。在候选框获取的过程中,网络首先在特征图每个点上生成 9 个不同尺寸比例的 anchors 锚点框,然后使用 softmax 分类函数预测锚点框内的图像是前景还是背景,并同时预测属于前景的锚点框的位置坐标的偏移量,最后使用非极大抑制 NMS 算法筛选输出用于目标检测的候选框。

ROI 池化将遍历 PRN 层产生的候选框并映射到特征图上确定具体的区域,将这个区域划分为 7×7 的区域,通过使用最大池化的方式将不同候选框对应的特征图区域形成固定的 7×7 的维度,方便送到后面的网络进行目标检测。最后的分

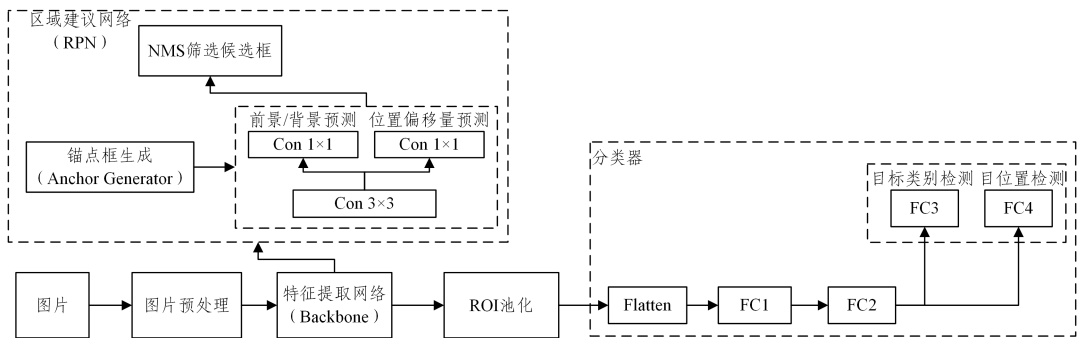


图 2 Faster RCNN 网络架构图

Fig. 2 Architecture of Faster RCNN

3 方法描述

3.1 基于 stable diffusion 的缺陷图像生成

如前所述,打刻面缺陷检测的最大挑战之一是缺陷样本极其稀少。表 1 统计了某整车制造厂发动机装配线上发动机打刻面缺陷样本的发生频率。如此低的发生频率,采用自然收集缺陷样本的方法,根本无法在有限的时间内完成模型的训练工作。而传统的数据增强方法因存在局限性,无法有效提升打刻面缺陷检测模型训练的性能,如表 2 所列。

表 1 发动机装配线上打刻面缺陷样本统计

Table 1 Sample statistics of engraved surface defects on the engine assembly line

No	Defect type	Frequency/%	Capacity	Defects per month
1	脱皮	0.03	1000	22
2	划痕	0.01	1000	2
3	撞痕	0.01	1000	2
4	污迹	0.1	1000	6

表 2 传统数据增强方法及其局限性

Table 2 Traditional data augmentation methods and their limitations line

No	Methods	Limitation
1	几何变换	没有考虑到数据集的特性和分布,对原缺陷样本的数量有要求
2	噪声添加	难以控制噪声质量,且可能导致训练标签的错误
3	颜色变换	难以大规模扩展,只能在已有特征上进行操作
4	复制粘贴	生成的缺陷样本数据与原始数据过于相似,导致模型泛化能力不足
5	GAN 或 VAE 图像生成	训练过程非常不稳定,难以调试;生成的样本图像质量较低,无法满足需求

为了解决上述的问题,实现良好的打刻面缺陷检测模型的训练效果,提出了基于 stable diffusion 的打刻面缺陷样本图像生成方法。使用工业相机对发动机打刻面区域的图像进行采集,获取每种缺陷的少量样本图片,同时收集正常的打刻面样本图片至少 1000 张。接下来的处理主要分 3 个方面:一是对缺陷样本图像的处理;二是对正常样本图片的处理;三是在前面两个处理步骤的基础上使用 stable diffusion 的图像修复功能,在正常样本图片的特定区域生成缺陷。

1) 针对缺陷样本图片的处理。首先将获取的缺陷样本图

类器将实现最终的目标检测分类以及目标区域坐标确定。它首先将从 ROI 池化输出的特征图进行展平处理,再接上两个全连接层,最后通过两个全连接层分别预测目标分类以及目标位置。Faster-RCNN 的网络架构图如图 2 所示。

片中的缺陷区域裁剪出来,形成缺陷区域图库,如图 3 所示。然后使用缺陷区域图库对 stable diffusion 的基础模型进行微调,生成打刻面缺陷区域图片。模型微调时,向 CLIP Text Encode 模型中添加打刻面缺陷类别的描述词汇,同时将基础模型中的 VAE 以及 U-Net 模型参数进行冻结,从而将该特定的描述词汇与对应的打刻面缺陷区域图像相关联。使用微调后的模型生成打刻面缺陷图像如图 4 所示。

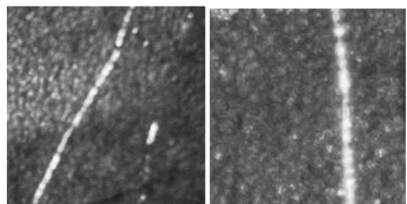


图 3 裁剪后的打刻面缺陷区域图片

Fig. 3 Image of the defect area on the engraved surface

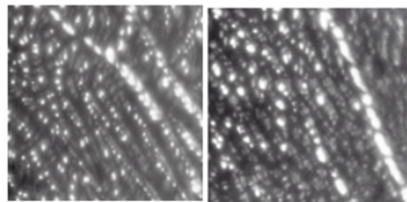


图 4 使用微调模型生成的缺陷图片

Fig. 4 Defect images generated using fine-tuning models

2) 针对正常样本图片的处理。采用轮廓检测算法提取正常打刻面上的字符轮廓,并基于字符轮廓形成对应样本图片的字符掩模图。该掩模图是字符轮廓区域为黑色,背景区域为白色的反向掩模图,最终得到正常样本库以及对应的字符掩模图库。

3) 缺陷图片的生成。基于前述准备条件的缺陷图片生成过程如图 5 所示。首先随机选取正常样本图片库中的图片,通过 VAE 的 Encoder 转成潜层特征 Z_{ini} ,并加入一定的高斯噪声得到 Z_t ;然后将通过 CLIP Text Encoder 转成语义信息的缺陷类别的描述词汇以及 Z_t 送入 U-Net 中,生成去噪后的潜层特征 $Z_{t-1, f_{\text{g}}}$ 。接着,在迭代去噪过程的每一步,利用区域生成掩模图将 $Z_{t-1, f_{\text{g}}}$ 中不需要重建的部分都替换成原图最初的特征,只在区域生成掩模图中白色区域进行特征的重建与

优化。最后利用字符掩模图将迭代完成后的潜层特征 Z_{0-0} 中的字符区域的特征替换成原图最初的特征输出最终的潜层特征 Z_{0-1} ，从而纠正图像在缺陷生成过程中对打刻面字符造成

的影响。最后，通过 VAE 的 Decoder 将最终的潜层特征 Z_{0-1} 转成像素级图像。生成缺陷图片与真实缺陷图片的对比如表 3 所列。

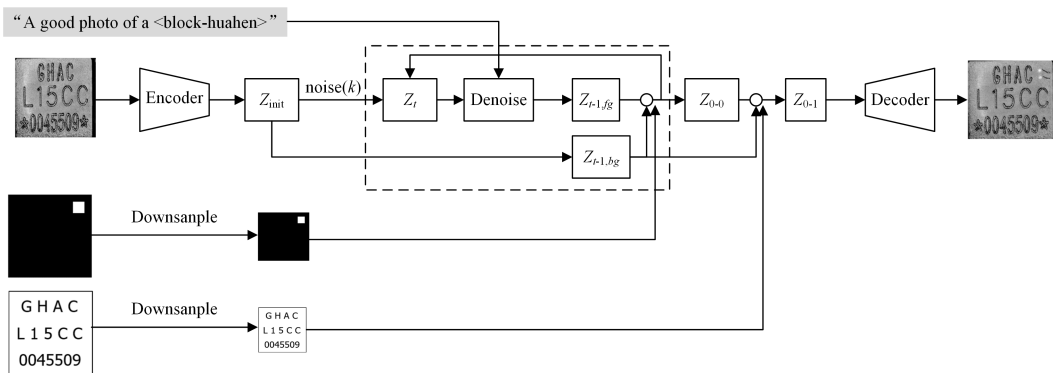


图 5 打刻面缺陷图片的生成过程

Fig. 5 Process of generating images of engraved surface defects

表 3 真实缺陷与生成缺陷图片

Table 3 Real defects and generated defect

Type	划痕	撞痕	脱皮
Real			
Generated			

3.2 Faster-RCNN 的改进

如前所述，发动机打刻面缺陷检测的另外一项挑战是缺陷目标尺度问题。为解决这个问题，提出了一种同步双向融合特征金字塔网络。通过将原模型中的 FPN 替换为该网络，实现多尺度特征融合和优化，进而完成 Faster-RCNN 模型的优化。

原始的 FPN 通过对特征提取网络中得到的特征图进行自上而下的融合，得到各个层级的融合特征图，如图 6(a) 所示，这样就可以实现预测的特征图上既保留深层特征的语义信息，又保留了浅层特征的空间信息。尽管如此，传统的 FPN 依然存在改善点。首先，它的特征融合是单向的，不同特征层之间存在语义上的差异，单向融合会降低在多尺度上的表达能力。其次，它自上而下地采用上采样的方式进行特征融合，会更多地保留深层语义信息，而相对缺少浅层的空间信息。基于这些不足，研究人员提出了各种改进的 FPN 结构，主要分为两种改进结构。一种是简单双向融合结构，如图 6(b) 所示，如 Liu 等^[15] 提出的 PANet；另外一种复杂双向融合结构，如图 6(c) 所示，如 Tan 等^[16] 提出的 BiFPN 以及 Qiao 等^[17] 提出的 RFP。该网络的结构示意图如图 6(d) 所示。

各种改进的 FPN 尽管在不同的目标检测场景中取得了不错的改进效果，但是在解决打刻面缺陷目标尺度非常广的问题上仍有改进空间。无论是简单双向融合特征金字塔，还是复杂双向融合特征金字塔，它们的双向融合都是基于自上向下融合而成的特征图开展的融合，这样最终形成的特征图在保留浅层的空间信息的能力上仍有所欠缺。为了改善这个问题，本文提出了同步双向融合金字塔网络架构，架构共包含

3 个融合路径，分别是自底向上的融合路径、自顶向下的融合路径以及平行融合路径。以 ResNet50 为特征提取网络的网络架构如图 7 所示，通过对 ResNet50 提取的特征图 $\{C_2, C_3, C_4, C_5\}$ 进行自底向上融合得到特征图 $\{D_2, D_3, D_4, D_5\}$ ，进行自顶向下融合得到特征图 $\{T_2, T_3, T_4, T_5\}$ ，最后通过平行融合得到特征图 $\{P_2, P_3, P_4, P_5\}$ 。

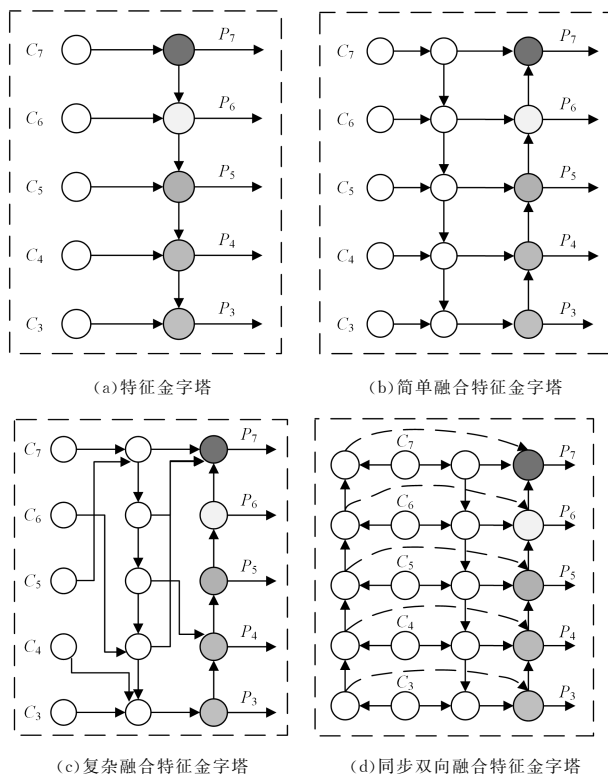


图 6 不同结构的特征金字塔网络

Fig. 6 Feature pyramid networks with different structures

1) 自底向上融合路径：首先， D_2 就是 C_2 ，并不需要进行特别的处理。从 D_3 开始，每个融合特征图都是经过如图 8(a) 所示的融合构建块形成的，即将 D_i 经过平均池化操作下采样到与 C_{i+1} 同样大小的特征图，并与 C_{i+1} 进行特征图相加操作，得到 D_{i+1} 。

2) 自顶向下融合路径：首先， T_5 就是 C_5 ，无需进行任何的操作。从 T_4 开始，每个融合特征图都是经过如图 8(b) 所示的

融合构建块形成的,即将 T_{i+1} 通过最近邻插值的方法上采样到与 C_i 同样大小的特征图,并与 C_i 进行特征图相加操作,得到 T_i 。

3) 平行融合路径:在通过上述两个路径的融合得到特征图 $\{D_2, D_3, D_4, D_5\}$ 与 $\{T_2, T_3, T_4, T_5\}$ 后,将同等大小的特征图进行相加操作,即 D_i 与 T_i 相加,最终输出 $\{P_2, P_3, P_4, P_5\}$ 。

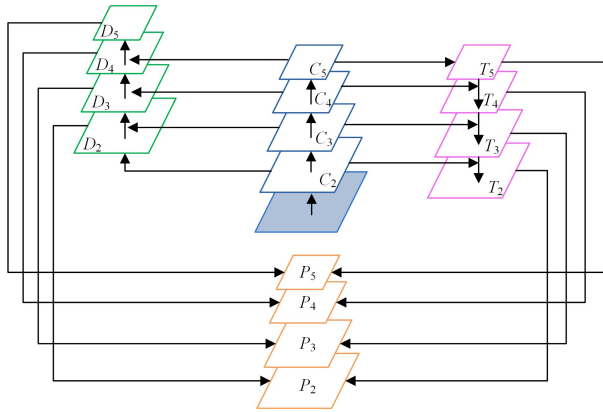


图7 同步双向融合特征金字塔网络框架

Fig. 7 Framework of synchronous bidirectional fusion feature

pyramid network

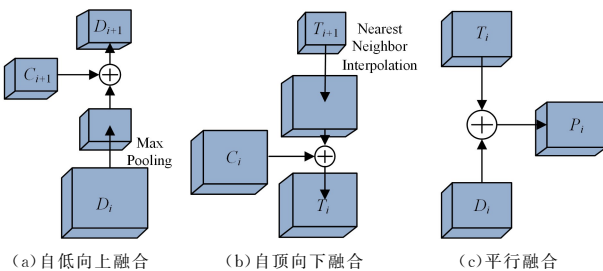


图8 3个融合路径的构建块

Fig. 8 Building blocks of three fusion paths

4 实验与结果分析

4.1 实验设计

本文的实验数据主要来自某汽车整车制造公司发动机装配生产线上,使用1200彩色工业相机配合85mm长焦镜头在流动的生产线上对打刻面区域进行取景。通过剔除模糊等质量不高的图像并经生产质量专家的确认,形成正常图片库3000张,真实缺陷图片库145张,其中缺陷类型包含划痕、撞痕、脱皮、污迹4种类型。对缺陷图片库按照8:2的比例划分训练集与测试集,然后通过本文提出的缺陷图片生成方法将训练集补充至1200张。

在Faster-RCNN模型的参数设置方面,优化器选择SGD,初始学习率 lr 设置为0.01,衰减系数 $weight\ decay$ 设置为0.0001,动量参数 $momentum$ 设置为0.9,训练批大小 $batch\ size$ 设置为4。学习率调整策略选择StepLR,调整步长为3,调整倍数为0.33。

针对生产线上的发动机打刻面缺陷检测,本文首先研究了一种基于stable diffusion的打刻面缺陷图片生成方法对缺陷训练集进行补充,然后提出同步双向特征金字塔网络对Faster-RCNN模型进行了改进。为了验证所提出方法的有效性,设计了以下3组对比实验。

1) 数据增强的对比实验:使用原Faster-RCNN模型在不

同的特征提取网络下对扩充数据集前后的检测效果进行对比,同时在同一条件下使用不同的数据增强方法进行对比,以确认本文采用的缺陷图片生成方法的有效性。

2) 特征金字塔网络改进前后的对比实验:原Faster-RCNN模型使用ResNet50为特征提取网络添加不同的特征金字塔网络进行检测效果的对比,使用的数据集为扩充后的数据集。该实验的目的,是确认在打刻面缺陷检测场景下,改进的同步双向融合特征金字塔网络的有效性。

3) 本文模型与其他目标检测模型之间的对比实验:将改进后的Faster-RCNN模型与原Faster-RCNN模型、SSD、YOLO进行实验对比。该实验主要是为了验证改进的有效性。

衡量缺陷检测模型的检测效果需要使用一些评价指标,比如精确率(Precision)、召回率(Recall)、平均精度(AP)、均值平均精度(mAP)等。在这些评价指标中,均值平均精度是最为综合,最能体现模型的精度。

1) 均值平均精度(mAP)

均值平均精度是各类别缺陷的平均精度的均值,mAP的值越大,代表模型的检测性能越好。

$$mAP = \frac{\sum AP_{\text{defect}}}{N_{\text{defect}}} \quad (1)$$

其中, AP_{defect} 是某一类缺陷的平均精度, $\sum AP_{\text{defect}}$ 表示所有类别缺陷的平均精度之和, N_{defect} 是模型检测缺陷的类别数。

2) 平均精度(AP)

平均精度是模型对某一缺陷类别检测的精确率(Precision)数值以及召回率(Recall)数值所组成曲线下的面积,表征单个缺陷类别的检测精度。精确率的计算如式(2)所示,召回率的计算如式(3)所示。

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

其中, TP 为模型正确检出缺陷的检测框数量, FP 为模型检出缺陷多余的检测框数量, FN 为模型没有检测到缺陷的数量。

4.2 实验结果分析

打刻面缺陷识别的效果评价指标将使用均值平均精度(mAP),交并比IOU(Intersection over Union)均设置为0.5,将从第4.1节的3个方面进行实验并对比分析,从而验证本文提出的方法的有效性。

1) 数据增强的对比实验

首先,在原Faster-RCNN模型上分别使用MobileNet, VGG16以及ResNet50作为特征提取网络对数据增强前后的训练集进行训练,并使用同一验证集进行验证。实验结果如表4所列。

表4 使用本文数据增强方法的对比实验

Table 4 Comparative experiment using the proposed data augmentation method

No.	Backbone	Before enhancement	After enhancement (%)
1	MobileNet	20.98	32.36
2	VGG16	39.14	55.03
3	ResNet50	62.79	79.63

从表 4 的结果可知,在对打刻面的缺陷数据集采用本文提出的缺陷图片生成方法进行增强后,模型检测的效果有了明显的提升。

然后,在使用 ResNet50 作为特征提取网络的原 Faster-RCNN 模型上采用不同的数据增强方法进行实验对比,结果如表 5 所列。

表 5 不同数据增强方法的对比实验

Table 5 Comparative experiment of different data augmentation

methods		
No.	Methods	mAP / %
1	几何变换	64.58
2	噪声添加	63.14
3	颜色变换	63.25
4	缺陷复制粘贴	72.14
5	Our method	79.63

从表 5 的结果可知,采用本文提出的缺陷图片生成方法进行数据增强相对已有的数据增强方法,在模型检测效果方面取得了最好的提升效果。由于打刻面缺陷图片数据集的数量有限,采用 GAN 以及 VAE 方法训练生成的打刻面缺陷图片质量无法满足模型训练的要求,故未进行实验效果的对比。

2) 特征金字塔网络改进前后的对比实验

本文提出的同步双向融合特征金字塔网络在原 FPN 的结构上增加了基于特征图的自底向上的融合分支以及平行融合分支,在本文扩充后的打刻面缺陷数据集上使用现有的各种 FPN 结构进行对比实验,结果如表 6 所列。

表 6 各种结构的 FPN 对比实验

Table 6 Experiments of FPNs with various structures

No.	FPNtype	mAP / %
1	FPN	79.63
2	CARAFE-FPN ^[18]	79.70
3	RFP	80.10
4	BFP ^[19]	84.80
5	BiFPN	87.10
6	PANet	89.40
7	DyFPN ^[20]	92.30
8	FPG ^[21]	95.60
9	Our method	97.52

从表 6 中可以看出,使用同步双向融合特征金字塔网络的 Faster-RCNN 模型相比其他改进的 FPN 取得了最好的效果:相比第二名的 FPG 结构,mAP 提升了约 2%;相对原始特征金字塔网络的 Faster-RCNN 模型,mAP 更是提升了约 17%。图 9 为改进后的 Faster-RCNN 模型与原模型的 mAP 在数据集上随着迭代次数变化的曲线图。可以看出,改进后的模型的精度更高,准确率提升的速度也更快。

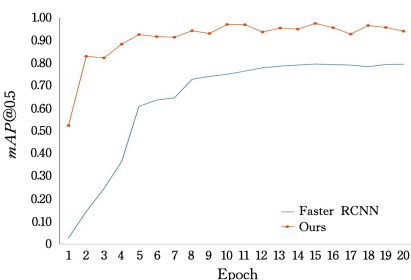


图 9 改进前后模型 mAP 曲线的对比

Fig. 9 Comparison of mAP curves of the model before and after improvement

3) 本文模型与其他目标检测模型的对比实验

为了验证本文提出的改进的 Faster-RCNN 模型在打刻面缺陷检测场景中的优越性,将其与 SSD 以及 YOLO 系列模型进行比较实验,其中 SSD 和 YOLO 系列模型均使用原始的特征提取网络,采用的训练集以及验证集均为本文所提出的扩充后的打刻面缺陷数据集。实验对比结果如表 7 所列。

表 7 主流检测模型的检测结果

Table 7 Detected results of mainstream model

No.	Model	mAP / %
1	Faster-RCNN	79.63
2	SSD	84.00
3	YOLOv3 ^[22]	90.10
4	YOLOx ^[23]	93.21
5	YOLOv8	95.62
6	YOLOv10 ^[24]	95.23
7	Our method	97.52

从表 7 中可以看出,本文改进后的 Faster-RCNN 模型在打刻面缺陷数据集上取得了最佳的性能,相对其他主流的目标检测模型取得了约 1.9% 的 mAP 提升。

结束语 为了实现发动机打刻面缺陷检测方式的革新,实现自动检测,本文提出了基于生成式数据增强与 Faster-RCNN 改进的缺陷检测。针对发动机生产线最终检查工序中打刻面缺陷样本罕见导致模型训练数据不足的问题,提出了一种基于 stable diffusion 的打刻面缺陷图片生成方法,实现对模型训练集的数据增强。针对发动机打刻面缺陷目标尺度范围广的问题,采用了同步双向融合特征金字塔网络改进原来的特征融合机制,实现模型在深层特征图语义信息以及浅层特征空间信息的融合,从而提升模型在多目标尺度检测的准确性。第 4 节的实验数据分析表明,采用本文方法进行数据增强后,缺陷检测模型的性能有明显的提升;采用本文改进的 Faster-RCNN 模型在缺陷识别的准确率上也取得了很大的提升,且改进模型的正确识别率也都超过了其他的主流算法。最终,本文提出的检测方法实现了在发动机生产线上的实际检测应用。但是,本文所提出的算法仍然有很多方面需要进一步改进。首先,在识别速度上还有很多的提升空间;其次,在保证本方法效果的基础上,算法参数量有待改善。这些都将是下一步的研究方向。

参考文献

- [1] YUN J P, CHOI S H, KIM J W, et al. Automatic detection of cracks in raw steel block using Gabor filter optimized by univariate dynamic encoding algorithm for searches(uDEAS)[J]. NDT & E International, 2009, 42(5): 389-397.
- [2] LIU Y, XU K, XU J. An Improved MB-LBP Defect Recognition Approach for the Surface of Steel Plates[J]. Applied Sciences, 2019, 9(20): 4222.
- [3] WANG J, FU P, GAO R X. Machine vision intelligence for product defect inspection based on deep learning and Hough transform[J]. Journal of Manufacturing Systems, 2019, 51: 52-60.
- [4] WEI L, DRAGOMIR A, DUMITRU E, et al. SSD: Single Shot MultiBox Detector[J]. arXiv: 1512. 02325, 2015.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C] // 2016 IEEE/Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 779-788.

- [6] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [7] CHEN J W, LIU Z G, WANG H R, et al. Automatic Defect Detection of Fasteners on the Catenary Support Device Using Deep Convolutional Neural Network [J]. IEEE Transactions on Instrumentation & Measurement, 2017, 67(2): 257-269.
- [8] LIU X Y, GAO J. Surface Defect Detection Method of Hot Rolling Strip Based on Improved SSD Model [C] // International Conference on Database Systems for Advanced Applications. 2021.
- [9] LI J, SU Z, GENG J, et al. Real-time Detection of Steel Strip Surface Defects Based on Improved YOLO Detection Network Science Direct [J]. IFAC-Papers OnLine, 2018, 51(21): 76-81.
- [10] WANG K, TENG Z, ZOU T. Metal Defect Detection Based on Yolov5 [J]. Journal of Physics: Conference Series, 2022, 2218(1): 012050.
- [11] DING R, DAI L, LI G, et al. TDD-Net: A Tiny Defect Detection Network for Printed Circuit Boards [J]. CAAI Transactions on Intelligence Technology, 2019, 4(2): 7.
- [12] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection [C] // 2017 IEEE/Conference on Computer Vision and Pattern Recognition (CVPR). 2017: 936-944.
- [13] YIN T P, YANG J. Detection of Steel Surface Defect Based on Faster R-CNN and FPN [C] // ICCAI'21. 2021.
- [14] ROMBACH R, BLATTMANN A, LORENZ D, et al. High-Resolution Image Synthesis with Latent Diffusion Models [C] // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022: 10674-10685.
- [15] LIU S, QI L, QIN H, et al. Path Aggregation Network for Instance Segmentation [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018: 8759-8768.
- [16] TAN M, PANG R, LE Q V. EfficientDet: Scalable and Efficient Object Detection [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020: 10778-10787.
- [17] QIAO S, CHEN L C, YUILLE A. DetectoRS: Detecting Objects with Recursive Feature Pyramid and Switchable Atrous Convolution [C] // 2021 IEEE/Computer Vision and Pattern Recognition. 2021.
- [18] WANG J, CHEN K, XU R, et al. CARAFE: Content-Aware Re-Assembly of Features [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). 2019: 3007-3016.
- [19] PANG J, CHEN K, SHI J, et al. Libra R-CNN: Towards Balanced Learning for Object Detection [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019: 821-830.
- [20] ZHU M J, HAN K, YU C B, et al. Dynamic Feature Pyramid Networks for Object Detection [C] // Fifteenth International Conference on Signal Processing Systems (ICSPS 2023). 2024: 504-511.
- [21] CHEN K, CAO Y, LOY C C, et al. Feature Pyramid Grids [J]. arXiv:2004.03580, 2020.
- [22] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement [J]. arXiv:1804.02767, 2018.
- [23] GE Z, LIU S, WANG F, et al. YOLOX: Exceeding YOLO Series in 2021 [J]. arXiv:2107.08430, 2021.
- [24] WANG A, CHEN H, LIU L, et al. YOLOv10: Real-Time End-to-End Object Detection [C] // NIPS'24. 2025: 107984-108011.



TAN Jianhui, born in 1990, postgraduate, intermediate engineer. His main research interests include industrial automation, industrial vision and artificial intelligence.