

基于智能型面分析的抛光表面缺陷检测研究

李子豪^{1,2}, 房丰洲^{1,2*}, 任仲贺^{1,2}, 侯高峰^{1,2}

¹天津大学精密测试技术及仪器国家重点实验室, 天津 300072;

²微纳制造实验室, 天津 300072

摘要 工件的表面质量对零件可靠性、质量和使用寿命的影响至关重要。尽管各种基于计算机视觉的目标检测框架已经被广泛应用于工业表面缺陷检测场景,但由于面型的影响以及缺陷之间的混叠性,超精加工工件表面缺陷检测仍然具有挑战性。因此,提出了一种频率嵌入双分支参数预测网络来预测滤波参数,滤除掉型面信息从而使得缺陷特征更加显著。基于智能型面分析的预处理后,提出了一种基于级联区域神经网络感受野增强缺陷检测网络,将可变形卷积间隔地替换到高效网络的卷积模块中,有效地提高了主干网络特征提取的能力,然后重新选择特征图组成新的特征金字塔网络以提高效率,进一步提高网络性能。此外,还构建了具有滤波参数标注信息的滤波参数数据集 UPP-CLS 和具有缺陷类别及位置的缺陷检测数据集 UPP-DET。模型在 UPP-CLS 上达到了 85.36% 的准确性,相较于现有网络提升 3~5 个百分点;在 UPP-DET 上达到了 0.862 的平均精度,相较于现有网络提升 5.3%~7.8%。模型整体性能优于主流网络结构。源代码将在 https://gitee.com/zihadol/detect_app 上开源。

关键词 超精密加工; 计算机视觉; 缺陷检测; 型面分析

中图分类号 O436

文献标志码 A

DOI: 10.3788/LOP230868

Polished Surface Defect Detection Based on Intelligent Surface Analysis

Li Zihao^{1,2}, Fang Fengzhou^{1,2*}, Ren Zhonghe^{1,2}, Hou Gaofeng^{1,2}

¹State Key Laboratory of Precision Measuring Technology and Instrument, Tianjin University, Tianjin 300072, China;

²Laboratory of Micro/Nano Manufacturing Technology (MNMT), Tianjin 300072, China

Abstract Surface quality of the workpiece is critical for part reliability, quality and service life. Although various vision-based target detection frameworks have been widely applied to industrial surface defect detection scenarios, surface defect detection of ultra-precise machining workpieces is still challenging due to the influence of face shape and the confounding nature between defects. Therefore, we propose a frequency-embedded two-branch parametric prediction network to predict the filtering parameters and filter out the profile information to make the defect features more significant. After pre-processing based on intelligent type surface analysis, a cascaded regional neural network-based perceptual field enhancement defect detection network is proposed. It replaces the deformable convolution intervals into the convolution module of the EfficientNet, which effectively improves the feature extraction capability of the backbone network. Then, the feature map is reselected to form a new feature pyramid network to improve the efficiency and further improve the network performance. In addition, the filter parameter dataset ultra precision polishing (UPP-CLS) with filter parameter labelling information and the defect detection dataset UPP-DET with defect category and location are constructed. The model achieves 85.36% accuracy on UPP-CLS, which is 3 to 5 percentage points higher than that of the existing networks, and 0.862 average precision on UPP-DET, which is 5.3%–7.8% higher than that of the existing networks. The overall performance of the model is better than the mainstream network architecture. The source code and dataset will be available at https://gitee.com/zihadol/detect_app.

Key words ultra-precise machining; computer vision; defect detection; profile analysis

1 引言

随着现代科学技术的不断发展,超精密加工技术

在核能、集成电路、激光和航天等领域的应用越来越广,已成为制造领域的主要加工技术。其中超精密抛光(UPP)又是加工金属、光学玻璃、大理石和碳素纤

收稿日期: 2023-03-15; 修回日期: 2023-04-18; 录用日期: 2023-04-23; 网络首发日期: 2023-05-03

基金项目: 国家自然科学基金(52035009)

通信作者: *fzfang@tju.edu.cn

维板等非金属材料以及一些低表面粗糙度零件的关键技术^[1-2]。半导体及光学器件制造领域对加工工件表面的质量也提出了前所未有的高要求,因此对其进行缺陷检测至关重要。

基于图像的传统缺陷检测通常是手工进行的,这种检测方式不仅需要大量的人力物力,而且效率低^[3-5],对缺陷判断时,人的主观因素占很大一部分,无法满足工业智能化的需求^[6]。随着计算机视觉的快速发展,该方法也逐渐应用到缺陷检测领域当中^[7-8]。2012年,Hinton等^[9]提出的 AlexNet 在 ILSVRC (ImageNet Large Scale Visual Recognition Challenge) 以远超第二名的识别率夺冠,从此掀起了深度学习的热潮。基于深度学习的方法解决了手工设计特征算子的局限性,其最大的优点就是可以自动提取图像中的特征及潜在的信息,使用基于学习的方法来进行缺陷检测,避免了大量的特征工程和专家知识。Zhang等^[10]使用卷积神经网络(CNN)来检测大功率圆盘激光焊接过程中3种不同的焊接缺陷,得到了比较好的识别效果。图像分类的方法只能浅显地分出每张照片的缺陷类型,无法区分单张图像中的多种缺陷,也无法得到缺陷的具体位置。目标检测的方法能很好地解决这个问题。郝柏桥等^[11]提出一种迁移学习的钢板缺陷检测方法,建立红外和可见光缺陷数据集,利用迁移学习实现了知识迁移,在钢板脉冲涡流热成像裂纹检测中达到98.6%的检测精度。小目标检测始终是目标检测任务中的难点^[12],为了提高图像中小尺寸缺陷的检测精度,Zeng等^[13]提出了一种新的增强型多尺度特征融合方法(ABFPN),采用具有不同膨胀率的空洞卷积算子来充分利用上下文信息,并且用跳跃连接来实现足够的特征融合,该方法在印制电路板(PCB)表面缺陷检测任务中取得了良好的检测性能。基于该检

测框架不仅能够识别缺陷的种类,而且能够给出缺陷的具体位置信息。

上述工作只针对普通彩色图像进行检测,在超精抛光领域中测量的数据为 μm 甚至 nm 级。微纳尺度下由于数据复杂,检测任务非常具有挑战性,现有的模型不能直接应用于这种数据集的缺陷检测。针对数据中高低频信息混叠问题,本文设计了频率信息提取分支,通过该智能型面分析模型提高缺陷特征显著性,提出了基于数据生成的级联的感受野增强网络来提升对多尺度缺陷的提取性能,实现缺陷的精准检测。

2 数据集

2.1 数据采集

图1为数据采集系统,由白光干涉仪、待测工件、成像上位机和调距上位机组成,其中白光干涉仪能够测量亚 nm 至 mm 的垂直范围,满足采集精度要求。采集的经过化学机械超精抛光的76个工件中包含氧化镓和镍磷合金两种材料,两种材料尺寸相同,直径均为 21 mm ,高均为 12.1 mm 。成像上位机用于接收来自白光干涉仪采集的数据,并将其以伪彩色图像的形式展示。调距上位机显示干涉条纹,通过调整位移台或物镜的 Z 向位移,使调距上位机出现较明显的干涉条纹,此时测量达到要求的结果,其中粗箭头表示位移方向。测量时,每次将工件放置在标记点,调整位移台使其在 X 方向移动 3 mm ,移动3次,然后将 Y 方向向下移动 6 mm ,再向 X 反方向移动 3 mm ,移动3次,每个工件采集6张不同位置的图像。移动过程中如调距上位机条纹消失或者不明显,需要调整使其出现明显干涉条纹。本实验采用 $20\times$ 焦物镜,采集视场尺寸为 $236\ \mu\text{m}\times 315\ \mu\text{m}$,其中数据矩阵大小为 640×480 ,通过算法转换为伪彩色图像,其数据尺寸为 $640\text{ pixel}\times$

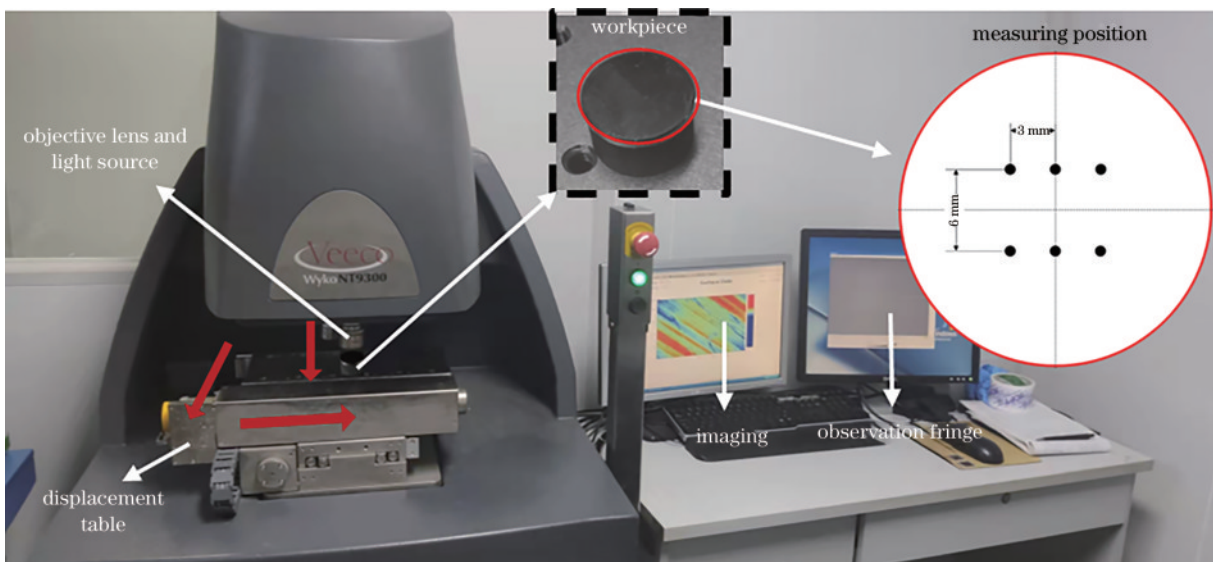


图1 数据采集系统

Fig. 1 Data collection system

480 pixel。最终采集得到 455 个图像数据,但是部分图像在测量时干涉条纹模糊,导致图像中形貌信息要素丢失,通过人工筛选移除 4 张不合格数据,最终得到 451 张图像组成的 UPP 数据集。分别对数据集进行清洗标注,得到了用于预测滤波的 UPP-CLS 数据集和用于检测的 UPP-DET 数据集。

2.2 数据标注

实际加工表面通常包含两种表面形貌信息,可分为低频信息和高频信息,低频信息一般为表面的形貌信息和波纹度信息,高频信息一般为表面的粗糙度信息和缺陷信息。加工时刀具在零件表面产生的不均匀冲击以及外界冲击和材料本身制备等问题,使工件表面产生缺陷损伤,这种损伤会给表面形貌造成轮廓突变(在频域内的表面为高频信息)。低频信息会使缺陷的特征变得不明显,因此会一定程度上降低缺陷检测器的性能。该问题可以通过图像滤波来解决,采用图像滤波的方式将低频的形貌信息去除,从而减少图像中对缺陷检测造成干扰的因素。本研究借鉴 Zeng 等^[14]采用的鲁棒的高斯滤波来消除边界畸变问题,但需要研究如何确定其中的截止波长参数。其中,区域高斯回归滤波器可以通过最小化得到。

$$\int_0^{l_y} \int_0^{l_x} \rho[z(\xi, \eta) - w(x, y)] \cdot s(\xi - x, \eta - y) d\xi d\eta \Rightarrow w(x, y)_{\min}, \quad (1)$$

式中:二值积分函数的范围限制在 $[0, l_x] \times [0, l_y]$; $\rho(\cdot)$ 为误差度量函数; ξ, η, x, y 为空间坐标; $z(\xi, \eta)$ 和 $w(x, y)$ 为测量表面; $s(x, y) = 1 / \{(\alpha^2 \lambda_{cx} \lambda_{cy}) \exp[-\pi(x/\alpha\lambda_{cx})^2 - \pi(y/\alpha\lambda_{cy})^2]\}$, 其中, $\lambda_{cx}, \lambda_{cy}$ 是可人为定义的 x, y 两个方向的截止波长,这两个参数决定了滤除掉型面信息的波长,而为了更加准确和高效地得到这两个参数,搭建了基于频率信息的网络用来回归这两个参数, $\alpha = \sqrt{\ln(2/\pi)}$, 为常量。依据实验,将截止波长划分为间隔 $40 \mu\text{m}$ 的波段,也就是将 $\lambda_{cx}, \lambda_{cy}$ 分别定义为 $40, 80, 120, 160, 200 \mu\text{m}$, 通过两个参数标记每张图像。由于每张图像的型面信息相似,方便起见,将每张图两个方向的截止波长固定为一个值 λ_c 。

3 基本原理

图 2 为数据处理检测的整体流程方案。先将数据输入到滤波参数预测网络进行训练,通过推理阶段得到的滤波参数将数据集进行滤波处理,提高缺陷特征的显著性。为减少过拟合,使训练更加稳定,将滤波后的数据集输入到基于 StyleGan2 的网络^[15]中训练,得到生成数据,并将这些数据和滤波后的数据融合,共同输入到缺陷检测的网络中进行检测,最终可以得到带有缺陷类别信息和边界框信息的图像数据。

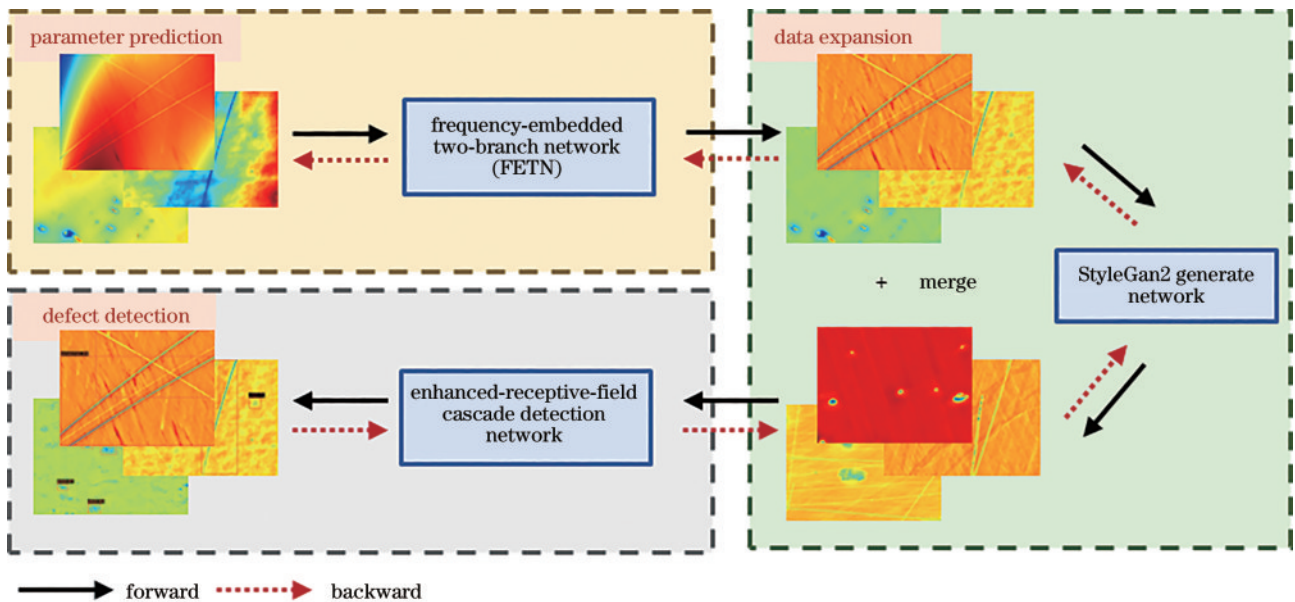


图 2 基于型面分析的缺陷检测整体方案

Fig. 2 Overall defect detection scheme based on surface analysis

3.1 智能型面分析模型原理

采集到的工件表面形貌数据是二维矩阵的形式,为检测工件中产生的缺陷,先将此数据映射为伪彩色图片,再对其进行检测。但是映射后的图像特征混乱,高低频信息混杂在一起,图 3 为图像增强特征的原理。

滤波有两个内置参数,分别为 x 方向和 y 方向的滤波波长。为后续消耗更少的人力物力,使用神经网络拟合出这两个参数,根据图像中型面波长和频谱之间的关系,融合图像频谱信息来提高网络提取特征的能力。

基于此提出了一种频率嵌入双分支参数预测网络(FETN),输入的图像经过离散傅里叶变换(DFT),从



图 3 智能型面分析原理图

Fig. 3 Schematic diagram of intelligent profile analysis

空间域转换为频域,将从频率特征提取模块得到的高维潜在空间向量和从图像中提取的特征向量进行通道方向的融合,再输入到后续的回归头中,增添的提取频率信息分支会更有效地指导网络提取信息。因为频率参数并非像缺陷信息显而易见,下面将验证该方法的有效性。

图 4 为 FETN 结构图,其中, n 为 FIFE 模块的数量,图像从右边的分支输入,得到包含图像特征的高维空间向量。左边是嵌入频率信息的分支,首先提取图像原始的矩阵信息,格式是二维矩阵,二维矩阵的长宽分别为数据沿 X、Y 方向的采样点数,矩阵中的每个点代表该点 Z 向的高度信息。分别对 X、Y 方向的数据作 DFT,用 M 和 N 来表示数据的行数和列数,以列为例,其信号可以表示为 $\mathbf{x}_n = \{a_0, a_1, a_2, \dots, a_{N-2}, a_{N-1}\}$,其中: a_N 表示每列数据 Z 向的高度; N 为每列数据的采样点个数,也就是数据每列的个数。分别对每列的 N 个样本进行 DFT:

$$\mathbf{X}_k = \sum_{n=0}^{N-1} \mathbf{x}_n \exp\left[(-j2\pi kn)/N\right], k = 0, 1, 2, \dots, N-1, \quad (2)$$

式中, $\mathbf{X}_k \in \mathbb{R}^{N/2}$ 为 DFT 结果,但碍于 DFT 算法复杂度高、计算缓慢,基于其对称性和周期性,使用快速傅里叶变换(FFT)完成计算。基于变换结果,对每条线的频率幅值作平均,得到该方向平均的频率幅值,代表这个方向波长的分布情况:

$$\bar{\mathbf{X}}_k = \left(\sum_{n=0}^{M-1} \mathbf{X}_k \right) / M, \quad (3)$$

式中, $\bar{\mathbf{X}}_k \in \mathbb{R}^{N/2}$ 为平均后的结果,同理可得到每行的 $\bar{\mathbf{Y}}_k \in \mathbb{R}^{M/2}$ 。图 4 中将 $\bar{\mathbf{X}}_k$ 和 $\bar{\mathbf{Y}}_k$ 这两个一维向量在维度上进行扩展合并,将合并的结果输入到频率特征提取(FIFE)模块中。Shao 等^[16]在研究雷达干扰信号分类时使用 1D-CNN 直接对原始信号进行特征提取,并取得了较好的分类效果,本文借鉴该结构构建一维频率信息提取主干网络。

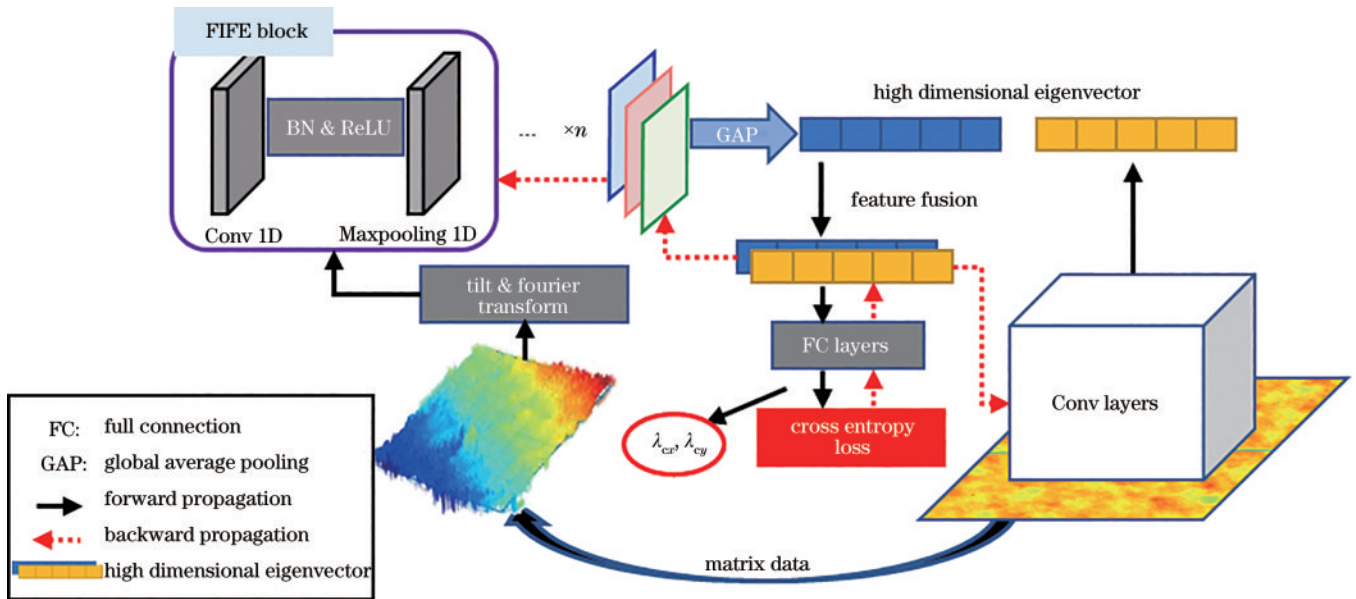


图 4 FETN 结构

Fig. 4 Structure of FETN

图 5 为 FIFE 模块的整体结构图, \mathbf{A}_1 为矩阵数据经过 DFT 后第一个通过卷积得到的特征向量,后面连接了 3 个 base 模块,在每个模块中输入的一维特征向量的形状都不会被改变,也就是特征向量 \mathbf{B} 的维度和特征向量 \mathbf{C} 的维度相同。 \mathbf{B} 和 \mathbf{C} 之间用一维 SE(squeeze and

excitation) 模块连接,特征向量 \mathbf{B} 可以解释为局部特征表述的集合,其中的统计信息可以表达整个图像^[17],通过连接 SE 模块可以更好地筛选一维向量中的有用信息。 \mathbf{A}_4 作为最后一个 base 模块输出的特征向量,后面连接全连接层得到 \mathbf{D} , \mathbf{D} 为抽象频谱信息中高维特征的一

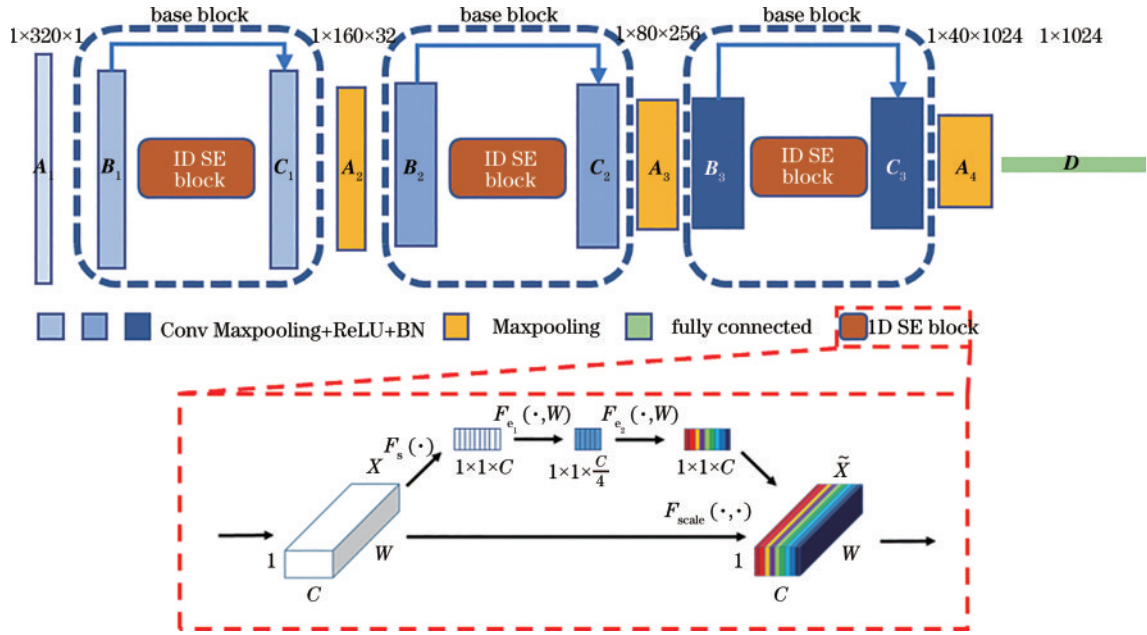


图5 FIFE模块结构

Fig. 5 Structure of FIFE block

维向量与右边分支图像特征一维向量的融合,输入到后面进行滤波参数预测。一维SE模块的工作原理为

$$m_k = F_s(X) = \frac{1}{W} \sum_{i=1}^W x_k, \quad (4)$$

式中: m_k 为 X 缩小数量为 W 的空间维度后的第 k 个元素; $F_s(\cdot)$ 为全局平均池化函数。使用最简单的聚合方法将具有全局信息的向量统计到局部描述中,其统计量可以代表整个图像的响应信息。信息的扩展层为

$$n = F_{e_2}[F_{e_1}(m, W_1), W_2] = \sigma[W_2 \delta(W_1 \cdot m)], \quad (5)$$

式中: n 为激活向量; $F_{e_1}(\cdot, \cdot)$ 、 $F_{e_2}(\cdot, \cdot)$ 为Sigmoid函数; $W_1 \in \mathbb{R}^{C \times (C/4)}$ 、 $W_2 \in \mathbb{R}^{C \times (C/4)}$, C 为通道维数; $\sigma(\cdot)$ 为ReLU函数。先将特征向量通道维度降维,旨在降低模型的复杂性,ReLU激活函数使该模块具有一定的非线性,最终使用激活向量 n 重新缩放 X :

$$\tilde{X} = F_{\text{scale}}(X_k, n_k) = X_k n_k, \quad (6)$$

式中: \tilde{X} 的维度和 X 的维度相同; $F_{\text{scale}}(\cdot, \cdot)$ 是 X_k 和 n_k 在通道维度上逐通道进行点积运算。把特征图中的全部信息聚合为一个向量,该向量可以看作带有特征的权重向量,通过可学习的方式进行参数更新,在通道方向进行自动特征整合。

在base模块中, B 与 C 之间添加跳跃连接,增强了 B 与 C 特征图的特征交互,并且在训练时可以稳定训练,整个base模块工作原理为

$$B(x) = S(x) + x, \quad (7)$$

式中: $S(x)$ 为一维SE模块; x 为输入。FIFE模块即为3个base模块串联得到:

$$o = D\{A_1[B_1(x)] \oplus A_2[B_2(x)] \oplus A_3[B_3(x)]\}, \quad (8)$$

式中: $o \in \mathbb{R}^{1 \times R}$ 为频率提取分支输出的高维向量, R 为一维张量的长度; \oplus 为卷积操作符。频率向量每经过

一个base模块,都通过池化层进行特征聚合和计算参数量降低。最终输出和图像分支相同形状的特征向量,将两向量进行组合后,传入全连接层进行特征融合,经过网络训练后即可得出所需的滤波参数。

3.2 缺陷检测模型原理

通过智能滤波算法,对输入的数据进行表面形貌去除,使图像中的缺陷特征更加显著。本节提出了一种基于增强的级联双阶段检测网络,将滤波后的图像输入主干网络,加入可变形卷积以实现精准的缺陷区域特征提取,这种稳健的工件抛光表面缺陷检测方法可以实现精准的缺陷定位和缺陷分类。

图6为级联缺陷检测模型,主体结构采用两阶段的检测网络。级联区域搜索网络是基于Faster-RCNN改进的一个网络^[18],该网络检测头部分设置多阶段的检测器,每个阶段的IoU(intersection over union)设置值逐渐增加,后一阶段检测器的输入使用前一阶段检测器的输出,这样可以逐步提高样本质量,从而提高整体检测器的检测性能。

为提升网络的整体检测性能,将网络的骨干部分换为EfficientNet^[19],该网络具有更加高效和精准的性能。EfficientNet由数个MBCConv(building convolutional modules for EfficientNet)模块组成,将图7(a)中第2、4、6、8个模块的卷积替换为可变形卷积^[20],可以提高形状各异缺陷的识别率以及精准性。传统的卷积模块只能以固定尺寸对图像中的特征进行建模学习,图7(b)中固定尺寸的感受野并不能完全覆盖缺陷特征,导致缺陷特征学习不准确。可变形卷积在没有额外监督信号的情况下,可以通过在图像中自主学习偏移量来改变卷积的空间采样位置,从而将感受野更加贴近地覆盖

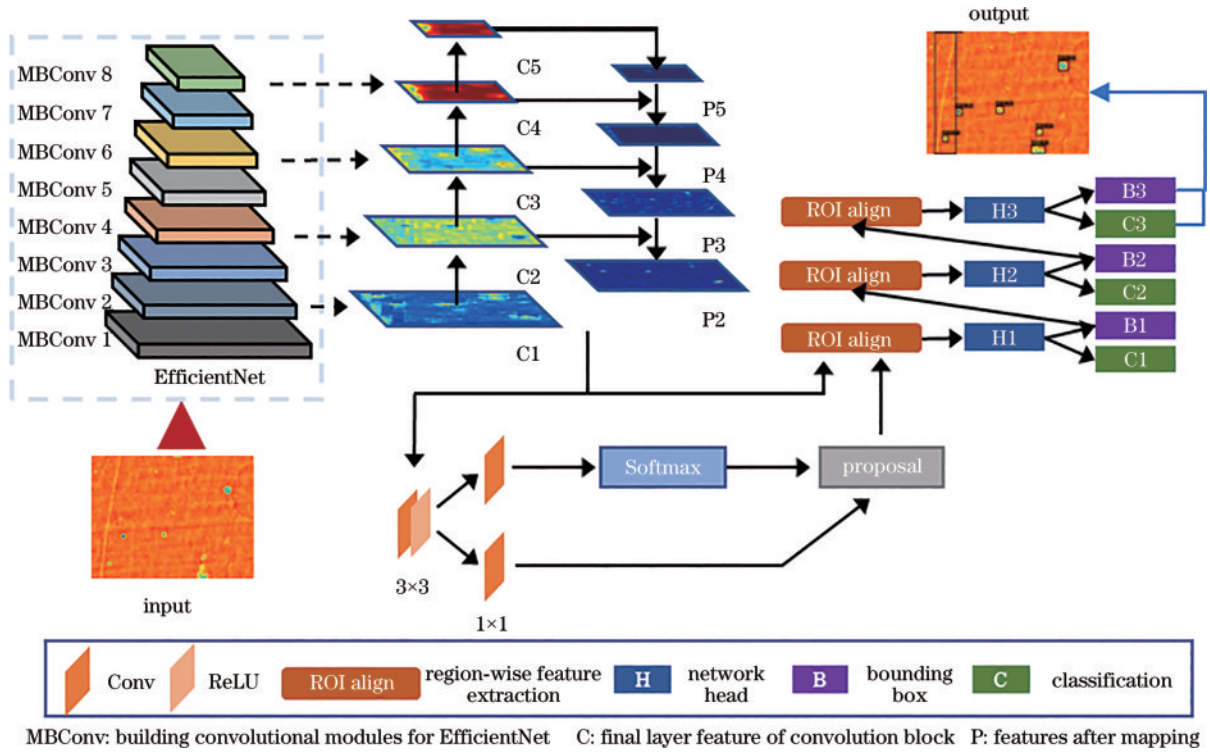


图 6 级联的感受野增强缺陷检测模型结构

Fig. 6 Structure of cascaded receptive field enhancement defect detection model

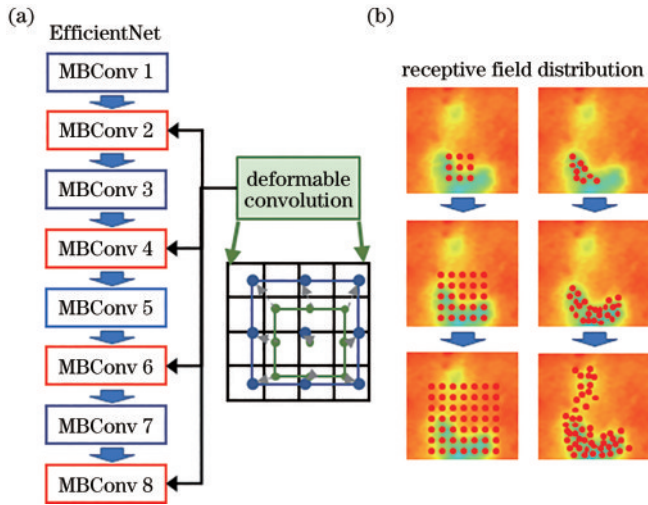


图 7 缺陷检测模型主干网络。(a)主干网络结构;(b)缺陷特征感受野对比

Fig. 7 Backbone network structure of defect detection model. (a) Backbone network structure; (b) comparison of defect characteristic receptive field

到缺陷尺寸上,有助于提高特征提取的性能。

4 结果及分析

4.1 实验设备及实验参数

训练以及测试实验是在 Ubuntu 20.04.4 LTS 操作系统上进行的,该设备配备了 Intel Core i9-10900k CPU,一块 GeForce RTX 3090 显卡,显存为 24 GB,系统内存为 32 GB,借助 CUDA 11.3 运算平台加速训

练。表 1 中列出了两个数据集实验的详细信息。

表 1 实验基本参数

Table 1 Basic parameters of the experiment

Parameter	UPP-CLS	UPP-DET
Batch size	8	4
Epoch	50	15
Initial learning rate	0.001	0.0002
Weight decay	-	0.05
Optimizer	Adam	AdamW
Regularization	L2 weight decay	-

4.2 智能型面分析实验结果

预处理阶段使用 UPP-CLS 数据集,数据集包含 451 张图像,其滤波参数由专家完成标注,选择正确的滤波参数进行滤波后可以显著地去除图像中的型面信息。

图 8 为 UPP-CLS 数据集标签分布,由图 8(a)可以看出数据集存在类别不均衡的问题,最少的类别只有 42 张图像,而最多的类别有 170 张图像。直接将数据集进行训练会导致类别数据少的部分出现过拟合情况,这样会直接影响分类器的性能,使用随机裁剪、水平翻转进行图像扩增,图 8(b)为数据平衡后的标签分布,不同类别的数据标签分布均衡。

经过数据扩增后数据集的图像数量为 761 张。其中 75% 的数据用于训练,25% 的数据用于测试,每次实验均对数据进行重新打乱,取 3 次实验的平均结果。在输入图像之前,对数据集的整体灰度作了统计,并进

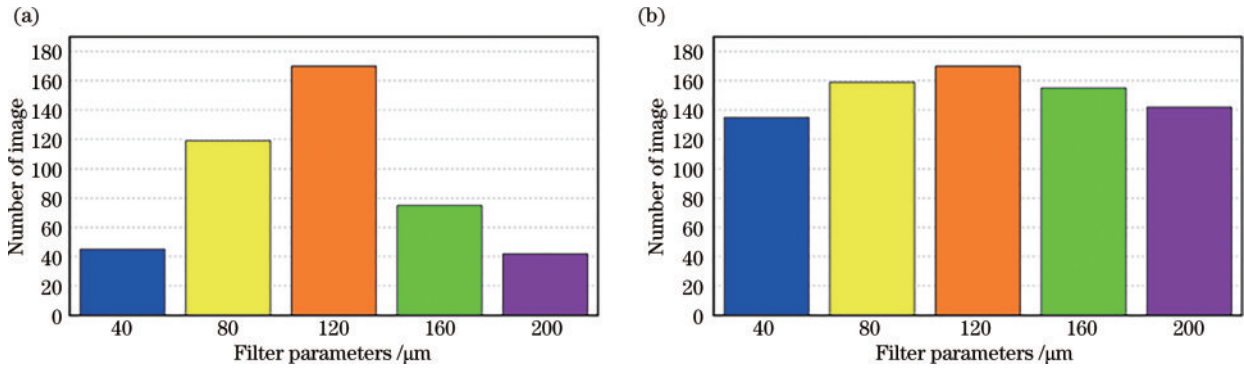


图 8 UPP-CLS 数据集标签分布。(a)数据平衡前标签分布;(b)数据平衡后标签分布

Fig. 8 UPP-CLS dataset label distribution. (a) Label distribution before data balancing; (b) label distribution after data balancing

行归一化处理,从而更易选择合适的学习率,并且移除掉共同的部分更能够凸显出个体差异。经统计,UPP-CLS 数据集的平均值 $M_{\text{mean}}=[0.672, 0.726, 0.324]$ 、标准差 $S_{\text{std}}=[0.382, 0.331, 0.385]$ 。

为提高 FIFE 分支模块的特征提取性能,对以下 3 个改进点进行实验:

1) 特征向量的组合方式。图像分支中图像经过主干网络进行特征提取,得到抽象特征的高维向量;频率信息分支中型面数据经过 FIFE 模块进行特征提取后,得到抽象频率特征的高维向量。两个向量在分类头处有两种组合方式,一种是沿长度方向的“L”组合方式,一种是沿通道方向的“C”组合方式。

2) 跳跃连接。参考 ResNet^[21],在设计特征提取模块时,加入跳跃连接以避免梯度消失和梯度爆炸问题。此外,跳跃连接将上一层的信息输入到处理后的下一层,使得网络学习的时候只需学习残差即可,加快了网络训练并提高了网络训练的稳定性。

3) 自注意力机制。加入一维的 SE 模块来提高特征提取的能力,SE 模块可以通过建模通道之间的关系来自适应地选择特征表达,使用全局信息加强有用特征,淡化无用特征。

实验使用 ResNet-50 进行测试,其他的实验参数同表 1,通过表 2 可以看出:在仅使用图像输入的单分支网络中,其准确率为 75.18%;而加入双分支输入,在长度方向上进行向量组合,准确率降低了 0.79 个百分点,长度方向的组合方式导致特征混淆,准确率反而出现下降;而将组合方式改为通道方向,借鉴 U-Net 的特征组合方式^[22],先在通道方向组合再进行加权更有助于特征融合,其准确率提高了 2.89 个百分点;在此基础上添加了跳跃连接,从数据中可以看出添加跳跃连接后准确率仅增长了 0.17 个百分点;每个 base 模块的一条分支中加入一维 SE 模块,其中 SE 模块为一个瓶颈结构,这种结构可以在不增加太多参数量的情况下,自适应地调整特征图各通道的权重,使准确率提高 1.23 个百分点,最终的准确率为 79.47%。该实验证明了各个模块对整体分支的增益,也表明模块之间的

表 2 FIFE 模块中模块性能分析

Table 2 Module performance analysis in the FIFE Block

Vector ordering		SC	SE	Accuracy / %
L	C			
-	-	-	-	75.18
✓	-	-	-	74.39
-	✓	-	-	78.07
-	✓	✓	-	78.24
-	✓	✓	✓	79.47

Note: “Vector ordering” indicates the way of vector combination, where “L” means the vectors are combined in the length direction, “C” means the vectors are combined in the channel direction. “SC” is skip connection, which means jump connection is added to the network.

兼容性和整体结构的有效性。

表 3 为各模型对比结果,分别使用几种先进的分类器进行消融实验,其中灵敏度 Sensitivity、漏报率 (FNR)、特异度 Specificity 和误报率 (FPR) 这些评价指标均是每类加和的平均结果。将几种先进的特征提取网络作为图像分支的主干网络,打勾代表加入频率信息提取分支 FIFE 模块。准确率随着不断更换图像分支主干网络而提高。其中 EfficientNet-b4 作为性能较强的特征提取网络嵌入频率信息分支可以达到最高 85.36% 的准确率,相较于 ResNet-50 提升了 5.89 个百分点,灵敏度也达到了最高的 81.49%。特异度最低为 12.28%,表明对于预测每一类来说,其他类预测错误的概率很低。添加 FIFE 模块的不同主干网络的准确率可以提升 3~5 个百分点,多种评价指标充分证明了 FIFE 模块的有效性,实现了深度学习对滤波参数的预测。

图 9 是使用 FETN 滤波后得到的一些结果图,其中:“before filtering”是未经滤波的原始输入图;“after filtering”是经过滤波后的图;顶部字体是详细的处理结果。可以很清楚地看到,智能滤波的操作有效地去除了型面信息,一定程度地增强了高频的缺陷特征。

表 3 各模型嵌入 FIFE 模块对比分析

Table 3 Comparative analysis of each model embedded in FIFE block

Image branch backbone	FIFE	Accuracy / %	Sensitivity / %	FNR / %	TNR / %	Specificity / %
InceptionV3 ^[23]	–	70.47	68.83	31.17	73.17	26.83
	✓	74.44	71.23	28.77	77.70	22.30
ResNet-50 ^[21]	–	75.18	73.64	26.36	77.88	22.12
	✓	79.47	76.12	23.88	76.27	23.73
ResNet-101 ^[21]	–	78.63	76.03	23.97	80.00	20.00
	✓	83.98	79.68	20.32	83.21	16.79
MobileNetV3 ^[24]	–	70.09	72.11	27.89	71.67	28.33
	✓	75.69	72.99	27.01	76.35	23.65
ResNext-50 ^[25]	–	78.87	77.69	22.31	80.04	19.96
	✓	82.21	77.15	22.85	84.76	15.24
EfficientNet-b2 ^[19]	–	80.41	78.25	21.75	82.67	17.33
	✓	84.86	80.89	19.11	86.38	13.62
EfficientNet-b4 ^[19]	–	82.11	79.97	20.03	84.64	15.36
	✓	85.36	81.49	18.51	87.72	12.28

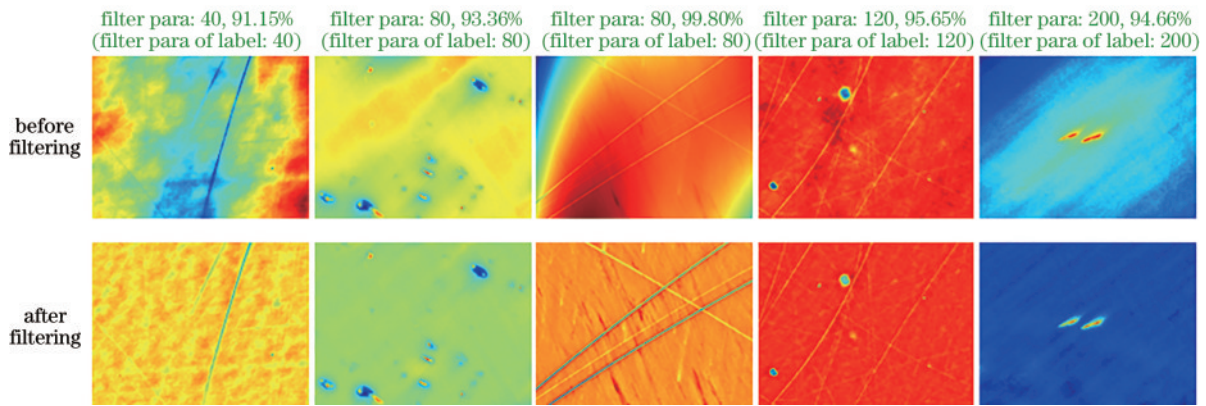


图 9 FETN 智能型面分析滤波结果图

Fig. 9 Results of FETN intelligent profile analysis and filtering

4.3 缺陷检测实验结果

使用提出的 UPP-DET 数据集进行实验,该数据集包含原始数据 442 张,每张图像的分辨率为 640 pixel × 480 pixel。为了更好地评估检测器的性能,尽量减少因为数据不足带来的过拟合现象,使用 StyleGan2 网络对数据集进行扩增,生成了 200 张拟真图像,最终的 UPP-DET 数据集包含 642 张图像。在用白光干涉仪进行数据采集后,根据图像将缺陷分为划痕和凹坑两个基础类,通过手工的方法对其进行标注,UPP-DET 的注释包括两部分,一个是每个缺陷的类别,一个是每个缺陷的边界框坐标信息。

初始化对于一个优化任务来说非常重要,它影响着模型迭代的稳定性以及迭代次数,并且在一定程度上会影响模型精度。图 10 根据所提出的数据集,统计了标注框的高宽比信息,从图 10 中可以看出标注框的高宽比大部分都分布在 6 之前,其中绝大多数分布在 2 之前,因此对锚框的初始化比例划分为 $R_{aspect} = [0.25, 0.50, 1.00, 2.00, 3.00, 5.00]$ 。

为了提升模型缺陷检测性能,分别从数据预处理、

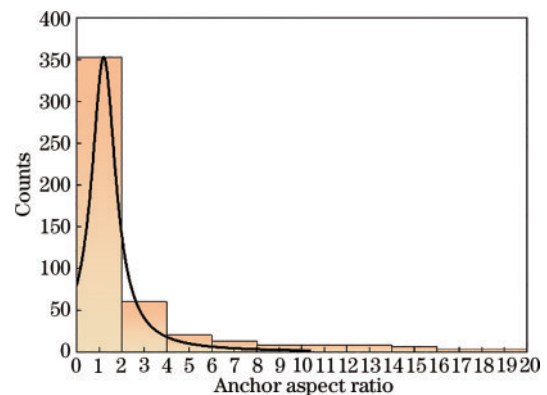


图 10 标注框在数据集中高宽比统计

Fig. 10 Height-to-width ratio statistics of dimension box in dataset

模型改进等 3 个模块进行实验:

- 1) 使用扩增数据。为了减少网络训练的过拟合,增加网络训练的稳定性,使用 StyleGan2 进行数据扩充,输出 200 张生成图像,将生成的拟真数据也进行人工标注,作为检测数据集的一部分。

2) Mixup 预处理。为了增加数据的多样性,在网络的预处理阶段使用 Mixup^[26]数据增强算法,该算法将两个图像按一定的比例混合生成新图像,并且还会生成新图像对应的新标签,使用新的图像和标签参与训练。

3) 嵌入可变形卷积。由于卷积网络层中卷积单元在固定位置对输入特征图进行采样,对于图像中不同位置具有不同尺寸与形状的对象,需要定义一种自适应感受野大小的卷积操作。这种将可变形卷积嵌入主干网络来提高网络的特征提取能力的方法是有效的^[26]。

实验使用 Faster R-CNN 进行测试,主干网络使用 ResNet-50,其他实验参数同表 1。表 4 中第一行没添加任何模块,其平均精度(mAP)为 0.697;使用 StyleGan2 扩增数据,此时 mAP 提高了 0.024,生成数据也增加了数据的多样性;在加入 Mixup 数据增强算法后,mAP 又提高了 0.008,略微增加了检测器性能;在主干网络中嵌入可变形卷积,其 mAP 再次增加了 0.011,可变形卷积增加了网络特征提取的性能。该实验证明了各模块对检测器性能提升的有效性。

表 5 为分别使用不同的先进检测器进行消融实验的结果,可以看出在一阶段的检测器中,Yolov5 检测器

表 4 缺陷检测模型模块性能分析

Table 4 Performance analysis of each model embedded in FIFE block

	AD	Mixup	DC	mAP
	—	—	—	0.697
	✓	—	—	0.721
	✓	✓	—	0.729
	✓	✓	✓	0.740

Note: “AD” is “amplified data”, which means using amplified data; “Mixup” means using this preprocessing operation; “DC” is deformable convolution, which means embedding deformable convolution in the network.

具有最好的性能(mAP 为 0.799);在二阶段的检测器中,Faster R-CNN 具有最快的速度(28.3 frame/s);Cascade R-CNN 在使用 ResNet-50 作为主干网络时,mAP 为 0.819,相较于 Faster R-CNN 提高 0.079。将主干网络替换为 EfficientNet-b4,并且将其中 MBConv 2、4、6、8 模块中的卷积替换为可变形卷积,为确保检测速度没有过多损失,重新设计特征金字塔网络(FPN)结构,间隔提取特征图,mAP 达到 0.854,相对于 Cascade R-CNN 提升 0.035。以精度为首要目标,在不损失太多速度的同时多阶段检测器获得了最好的性能。

表 5 各检测模型与所提模型对比结果

Table 5 Comparison results between proposed model and other detection models

Stage	Method	Backbone	MST	Scratch	Pit	mAP	FPS / (frame · s ⁻¹)
One-stage	SSD300 ^[27]	VGG-16	—	0.363	0.335	0.349	—
	SSD512 ^[27]	VGG-16	—	0.581	0.593	0.587	—
	RetinaNet ^[28]	ResNet-50	—	0.775	0.749	0.762	—
	Yolov3 ^[29]	DarkNet-53	—	0.632	0.594	0.613	—
	Yolov5	DarkNet-53	—	0.816	0.782	0.799	—
Two-stage	Faster R-CNN ^[30]	ResNet-50	✓	0.729	0.751	0.740	28.3
	Dynamic R-CNN ^[31]	ResNet-50	✓	0.816	0.798	0.807	24.9
	Cascade R-CNN ^[18]	ResNet-50	✓	0.815	0.822	0.819	22.7
	Proposed	EfficientNet-b4	✓	0.867	0.840	0.854	21.1

Note: “MST” is “multi-size test”, indicating the usage of multi-scale test; “Scratch” and “Pit” are the categories of detection; “FPS” is “frames per second”, indicating the number of frames per second processed.

网络级联的 3 个检测头中的 IoU 阈值(T_{IoU})对模型的影响如表 6 所示,其他实验参数使用最好结果的参数。设置级联检测头结构能够逐渐精准地回归出检测框,并且能够克服训练和测试中输入候选框质量不匹配的问题。表 6 中第一行是测试基线,在使用原始参数时,网络测试的 mAP 为 0.854。过小的 IoU 阈值并不能起到筛选候选框的作用,因此修改第一个检测头的参数 $T_{IoU_1} = [0.2, 0.3, 0.4]$ 。从 $T_{IoU_1} = 0.2$ 开始测试,在固定后两个检测头参数的情况下,网络检测的性能随着 T_{IoU_1} 的提高而提升,在 $T_{IoU_1} = 0.4$ 时 mAP 达到 0.856。继续调整 T_{IoU_2} 和 T_{IoU_3} 的参数,最终在 $T_{IoU} = [0.4, 0.5, 0.7]$ 时达到最好的效果(mAP 为 0.862),

表 6 缺陷检测模型检测头各阶段 IoU 阈值分析结果

Table 6 IoU threshold analysis results at each stage of defect detection model detection head

No.	T_{IoU}	mAP
Baseline	[0.5, 0.6, 0.7]	0.854
1	[0.2, 0.6, 0.7]	0.841
2	[0.3, 0.6, 0.7]	0.853
3	[0.4, 0.6, 0.7]	0.856
4	[0.4, 0.5, 0.6]	0.851
5	[0.4, 0.5, 0.7]	0.862
6	[0.4, 0.5, 0.8]	0.839

相较于基线提升了 0.9%,相较于 Yolov5 提升了 7.8%,相较于 Cascade R-CNN 提升了 5.3%。

图 11 是所提检测模型同其他检测模型结果对比图,图 11(a)是检测的总体结果图,每行为检测的数据样本。图 11 中有些未标注的缺陷,在实际预测时可以检测出来,由图 11 所示,所提模型检测得更加全面。图 11(b)为图 11(a)部分区域放大图,可以看出,所提方法相较于 Faster R-CNN 和 Cascade R-CNN 两种方法能够检测出较为全面的缺陷特征。箭头标出的均为对比区域,前两种方法均没有检测出下面的划痕缺陷,

所提方法以 0.92 的置信度检测到划痕缺陷。Faster R-CNN 模型也未能将箭头标识的划痕检测出来。并且对比检测缺陷的置信度,所提方法在确定性上也相对较高,相比之下可以证明所提方法具有较好的性能。

通过所提算法,得到了一个高性能的缺陷检测器。图 12 是使用该检测器得到的结果,上面一行为标注信息,下面一行是对应的检测结果。

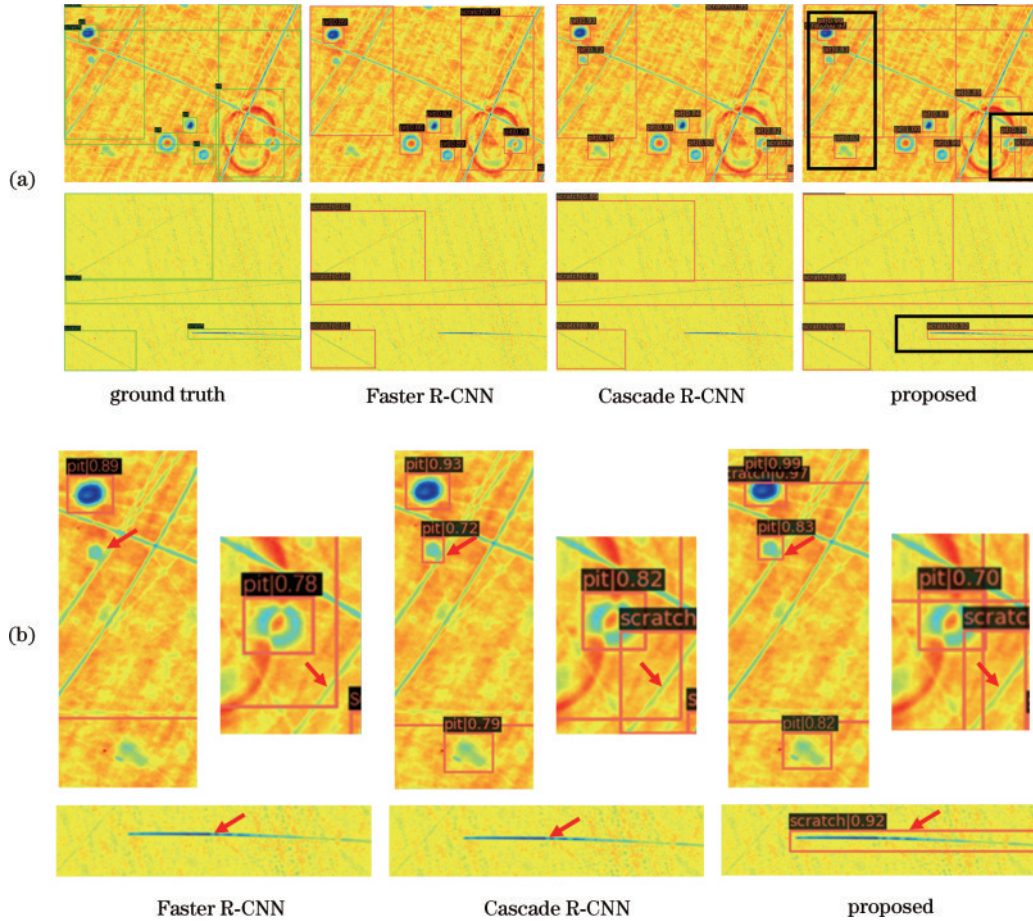


图 11 所提检测模型和其他主流检测模型在 UPP-DET 数据集上的结果对比。(a)整体对比图;(b)局部对比图
Fig. 11 Comparison of the results between the proposed detection model and other mainstream detection models on the UPP-DET dataset. (a) Overall comparison; (b) partial comparison

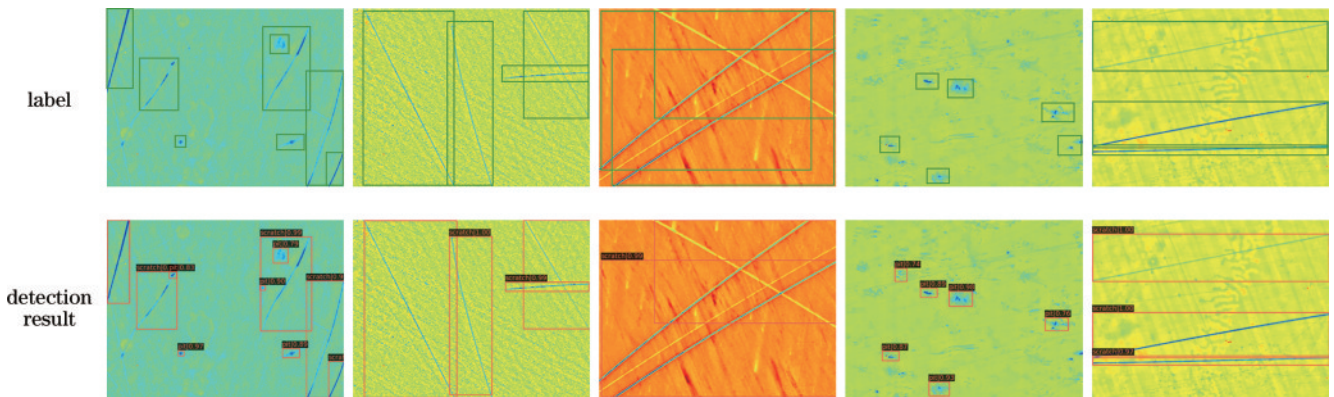


图 12 缺陷检测模型检测结果图
Fig. 12 Test results of defect detection model

5 结 论

提出了一种智能型面分析方法——FETN,该方法可以批量处理高精度的工件表面图像,滤除掉表面低频的型面信息,保留高频的缺陷信息,从而提高缺陷检测效率。由于其可以滤除型面信息,因此可以用于处理平面、球面甚至自由曲面的工件。该方法将图像本身和图像的频率信息同时输入到网络,引导网络学习图像和滤波参数之间的关联,通过该网络得到滤波参数。对图像滤波可以有效滤除型面信息,提高缺陷特征的显著性。在此智能预处理的基础上,提出了一种基于级联区域网络的感受野增强缺陷检测模型,将主干网络替换为高性能的 EfficientNet,为了保证其高效性重新设计了 FPN 结构,将可变形卷积间隔嵌入 EfficientNet 的 MBConv 中让感受野能够自适应检测物体,提高了检测性能。为此构建了两个数据集:UPP-CLS 包含 451 张标注图像,用于智能滤波处理;UPP-DET 包含 642 张标注图像,用于缺陷检测。实验证明,所提方法在 UPP-CLS 上实现了 85.36% 的准确率,在 UPP-DET 上实现了 0.862 的 mAP,准确性优于现有网络。

参 考 文 献

- [1] 袁巨龙, 张飞虎, 戴一帆, 等. 超精密加工领域科学技术发展研究[J]. 机械工程学报, 2010, 46(15): 161-177.
Yuan J L, Zhang F H, Dai Y F, et al. Development research of science and technologies in ultra-precision machining field[J]. Journal of Mechanical Engineering, 2010, 46(15): 161-177.
- [2] Kong L B, Cheung C F. Design, fabrication and measurement of ultra-precision micro-structured freeform surfaces[J]. Computers & Industrial Engineering, 2011, 61(1): 216-225.
- [3] Tulbure A A, Tulbure A A, Dulf E H. A review on modern defect detection models using DCNNs-deep convolutional neural networks[J]. Journal of Advanced Research, 2021, 35: 33-48.
- [4] 刘兰, 叶芸, 郭太良. 基于拓展型特征金字塔的 OLED 像素缺陷检测方法[J]. 光学学报, 2023, 43(2): 0212005.
Liu L, Ye Y, Guo T L. Detection method for OLED pixel defects based on extended feature pyramid[J]. Acta Optica Sinica, 2023, 43(2): 0212005.
- [5] Ren Z H, Fang F Z, Yan N, et al. State of the art in defect detection based on machine vision[J]. International Journal of Precision Engineering and Manufacturing-Green Technology, 2022, 9(2): 661-691.
- [6] Gao Y P, Li X Y, Wang X V, et al. A review on recent advances in vision-based defect recognition towards industrial intelligence[J]. Journal of Manufacturing Systems, 2022, 62: 753-766.
- [7] Wang J L, Xu C Q, Zhang J, et al. Big data analytics for intelligent manufacturing systems: a review[J]. Journal of Manufacturing Systems, 2022, 62: 738-752.
- [8] Gao R X, Wang L H, Helu M, et al. Big data analytics for smart factories of the future[J]. CIRP Annals, 2020, 69(2): 668-692.
- [9] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [10] Zhang Y X, You D Y, Gao X D, et al. Welding defects detection based on deep learning with multiple optical sensors during disk laser welding of thick plates[J]. Journal of Manufacturing Systems, 2019, 51: 87-94.
- [11] 郝柏桥, 范玉刚, 宋执环. 基于深度迁移学习的脉冲涡流热成像裂纹缺陷检测[J]. 光学学报, 2023, 43(4): 0415002.
Hao B Q, Fan Y G, Song Z H. Deep transfer learning-based pulsed eddy current thermography for crack defect detection[J]. Acta Optica Sinica, 2023, 43(4): 0415002.
- [12] 程松, 杨洪刚, 徐学谦, 等. 基于 YOLOv5 的改进轻量型 X 射线铝合金焊缝缺陷检测算法[J]. 中国激光, 2022, 49(21): 2104005.
Cheng S, Yang H G, Xu X Q, et al. Improved lightweight X-ray aluminum alloy weld defects detection algorithm based on YOLOv5[J]. Chinese Journal of Lasers, 2022, 49(21): 2104005.
- [13] Zeng N Y, Wu P S, Wang Z D, et al. A small-sized object detection oriented multi-scale feature fusion approach with application to defect detection[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 3507014.
- [14] Zeng W, Jiang X, Scott P J. Fast algorithm of the robust Gaussian regression filter for areal surface analysis[J]. Measurement Science and Technology, 2010, 21(5): 055108.
- [15] Karras T, Laine S, Aittala M, et al. Analyzing and improving the image quality of StyleGAN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 8107-8116.
- [16] Shao G Q, Chen Y S, Wei Y S. Deep fusion for radar jamming signal classification based on CNN[J]. IEEE Access, 2020, 8: 117236-117244.
- [17] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [18] Cai Z W, Vasconcelos N. Cascade R-CNN: delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 6154-6162.
- [19] Tan M, Le Q. Efficientnet: rethinking model scaling for convolutional neural networks[J]. Proceedings of Machine Learning Research, 2019, 97: 6105-6114.
- [20] Dai J F, Qi H Z, Xiong Y W, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 764-773.

- [21] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [22] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [23] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2818-2826.
- [24] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2020: 1314-1324.
- [25] Xie S N, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5987-5995.
- [26] Zhang H, Cisse M, Dauphin Y N, et al. Mixup: beyond empirical risk minimization[EB/OL]. (2018-04-27)[2023-04-14]. <https://arxiv.org/abs/1710.09412>.
- [27] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Cham: Springer International Publishing, 2016, 9905: 21-37.
- [28] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2999-3007.
- [29] Farhadi A, Redmon J. Yolov3: an incremental improvement[EB/OL]. (2018-04-08)[2023-01-02]. <https://arxiv.org/abs/1804.02767>.
- [30] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [31] Zhang H K, Chang H, Ma B P, et al. Dynamic R-CNN: towards high quality object detection via dynamic training [M]//Vedaldi A, Bischof H, Brox T, et al. Computer vision-ECCV 2020. Lecture notes in computer science. Cham: Springer, 2020, 12360: 260-275.