

DOI: 10.19344/j.cnki.issn1671-5276.2025.02.044

# 自适应光照变化的高频工件图像识别算法

孙成龙<sup>1</sup>, 李节<sup>1</sup>, 李柏林<sup>1</sup>, 王逸涵<sup>1</sup>, 欧阳<sup>2</sup>

(1. 西南交通大学 机械工程学院, 四川 成都 610031; 2. 成都大学 机械工程学院, 四川 成都 610106)

**摘要:**针对工业场景中不同光照条件下高频工件识别准确率低的问题,提出一种自适应光照变化的高频工件图像识别算法。采用通道注意力模块增强图像中对光照变化不敏感的特征,以减少光源照度对识别结果的影响;以 ResNet50 网络为基础,构造了两个分支网络,分别从整体图像中提取全局信息以及从有效的区域中提取局部信息,其中有效的区域由弱监督区域检测模块得到;融合两个分支网络的识别结果,实现了高频工件的图像识别。通过对多类别高频工件的实验结果表明:该算法能够自适应光照的变化,提高了复杂光照条件下的高频工件识别性能,识别准确率达到 94.8%。

**关键词:**光照变化;工件识别;注意力机制;区域检测

**中图分类号:** TP391.41    **文献标志码:** A    **文章编号:** 1671-5276(2025)02-0227-06

## High-frequency Workpiece Image Recognition Algorithm with Self-adaption Illumination Change

SUN Chenglong<sup>1</sup>, LI Jie<sup>1</sup>, LI Bailin<sup>1</sup>, WANG Yihan<sup>1</sup>, OU Yang<sup>2</sup>

(1. School of Mechanical Engineering, Southwest Jiaotong University, Chengdu 610031, China;

2. School of Mechanical Engineering, Chengdu University, Chengdu 610106, China)

**Abstract:** To improve the low recognition accuracy of high-frequency workpieces under different illumination conditions in industrial scenes, an image recognition algorithm of high-frequency workpieces with self-adaptive illumination change is proposed. The channel attention module is used to enhance the features in the image that are insensitive to illumination change, so as to reduce the influence of light source illumination on the recognition results. Based on ResNet50 network, two branch networks are constructed to extract global information from the whole image and local information from the effective region, where the effective region is obtained by the weak supervised region detection module. The recognition results of the two branch networks are fused to achieve the image recognition of high-frequency workpieces. The experimental results of multi-type high-frequency workpieces show that the algorithm can adapt to the change of illumination and improves the recognition performance of high-frequency workpieces under complex illumination conditions, with 94.8% recognition accuracy.

**Keywords:** light change; workpiece recognition; attention mechanism; regional detection

## 0 引言

近年来,随着我国制造业的转型升级并朝着更高质量的方向发展<sup>[1]</sup>,工艺加工自动化水平和工业生产智能化水平也在不断提升<sup>[2-3]</sup>。高频工件作为航天航空设备的重要组成零件,通过在加工流程中引入图像识别技术,不仅可以实现图像处理在工业生产中的实际应用,还可以促进高频工件的加工自动化与制造智能化,进一步服务于中国制造 2025<sup>[4]</sup>。高频工件的识别目前面临以下问题与挑战:1)工件的结构复杂而且类型多样;2)不同种类的工件之间差异较小;3)工件的摆放姿态和光照变化较大。因此,对高频工件加工自

动化的研究仍有较大空间。

图像识别技术在工业领域中有着广泛的应用<sup>[5]</sup>,国内外学者提出了多种相关算法和技术手段。文献[6]提出了一种基于卷积神经网络(convolutional neural networks, CNN)<sup>[7-8]</sup>的机械零件识别算法;文献[9]提出了一种基于分块主成分分析(principal component analysis, PCA)和支持向量机(support vector machine, SVM)<sup>[10]</sup>的零件识别分类系统;文献[11]提出了一种多分支特征融合卷积神经网络(multi-branch feature fusion convolutional neural networks, MFF-CNN)用于主轴轴承零件的自动分类;文献[12]提出了改进的 Inception V3<sup>[13]</sup>与 Xception<sup>[14]</sup>用于螺纹连接件零

**基金项目:**四川省科技厅重点研发项目(2021YFN0020)

**第一作者简介:**孙成龙(1999-),男,山东莘县人,硕士研究生,研究方向为机器视觉、图像处理, sunclwlkq@163.com。

件的识别。上述算法在一定程度上能够克服光照变化对识别结果的影响,但其研究的机械零件结构相对简单、种类较少,不同类型工件之间的差别较大,因此无法有效适用于高频工件的实际加工生产。

为了解决复杂光照下高频工件(采用感光电流使工件局部加热的表面热处理工艺处理后的工件)较难区分的问题,实现加工工序智能化,本文提出一种自适应光照变化的高频工件图像识别算法。该算法基于 ResNet50<sup>[15]</sup>网络,首先使用通道注意力模块提取强光照鲁棒性的全局工件图像特征,再经过主干网络得到一个全局识别结果;然后采用弱监督区域检测模块中的边界搜索机制<sup>[16]</sup>找到工件的重要局部区域,并在裁剪图像后引入分支网络中得到一个局部识别结果;最后在分支融合模块中联合两个识别结果实现工件识别。实验结果表明:与多种工件识别算法相比,该算法对

光照变化有更强的适应性,并且显著提高了识别的准确率。

### 1 本文算法

#### 1.1 总体架构

本文所提出的自适应光照变化的高频工件图像识别算法框架如图 1 所示。算法主要由 3 个模块组成:通道注意力模块、弱监督区域检测模块和分支融合模块。首先将工件全局图像  $I_1$  载入通道注意力模块得到输出特征,接着通过主干网络 ResNet50 生成多层特征图  $M_g$  和识别结果  $P_1$ ;然后在弱监督区域检测模块中根据全局图像  $I_1$  和多层特征图  $M_g$  截取工件局部图像  $I_2$ ,再导入另一个分支的通道注意力模块和卷积神经网络获得识别结果  $P_2$ ;最后利用分支融合模块融合两个分支的识别结果,得到最终的识别结果  $P$ 。

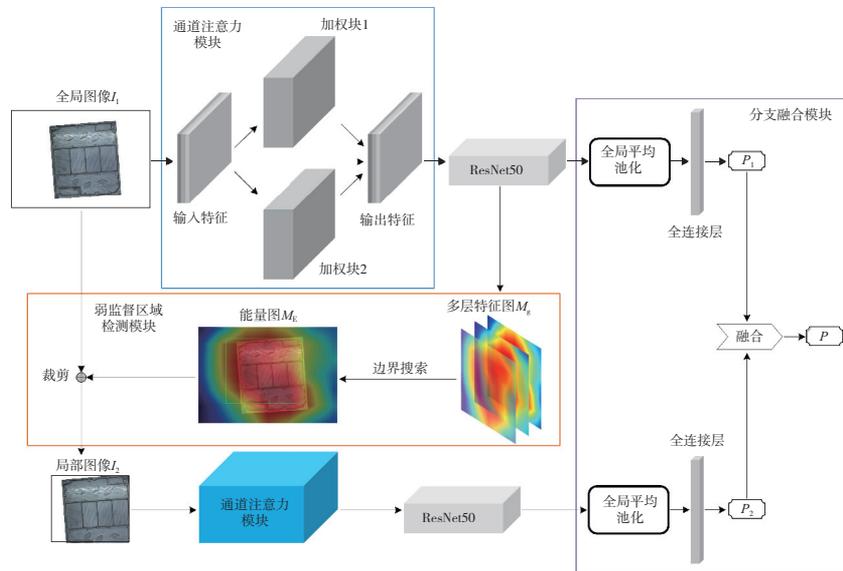


图 1 本文算法框架

#### 1.2 通道注意力模块

工业生产中采集工件图像时受到光照不均和光照变化较大等因素的影响,获取的图片会出现光斑、阴影和光线不足等现象,如图 2 所示。如果直接将工件图像输入网络提取特征,无法获得有

效的图像特征,难以准确识别高频工件的类别。为克服上述干扰信息对工件识别的影响,本文采用通道注意力模块增强工件图像的特征,如图 1 中蓝色区域所示(本刊为黑白印刷,疑问之处请咨询作者)。

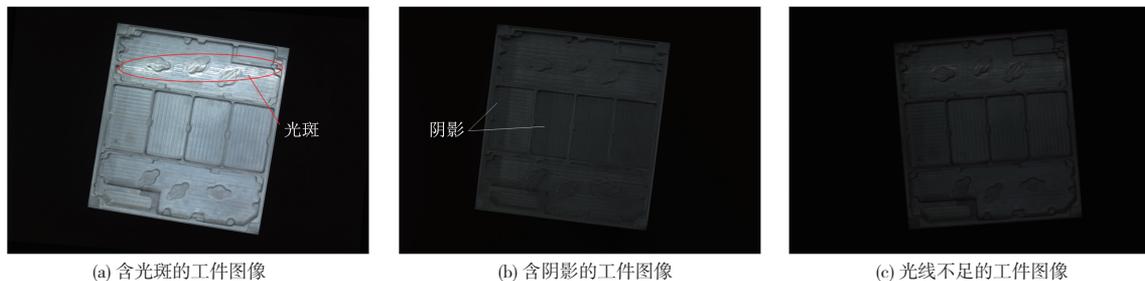


图 2 不同光照条件下的高频工件图像

采集到的工件图像为 RGB 三通道彩色图像,有的颜色通道含有受光照变化影响较大的信息,而有的颜色通道含有对光照变化不敏感的信息。通道注意力模块的实质是在原始图像中显式地建立颜色特征通道之间的相互依赖关系,然后通过学习的方式自动获取每个颜色特征通道的特征信息,并依照其重要程度增加对有用颜色特征通道的关注,抑制对当前图像识别任务不利的颜色特征通道,使光照变化的影响最小化。本文利用加权块使网络专注于有效通道的学习,如图 3 所示。

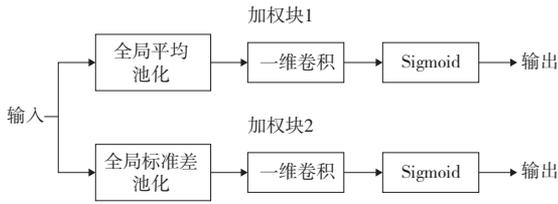


图 3 加权块

本文对工件的全局图像  $I_1$  分别采用全局平均池化和全局标准差池化,将每个二维颜色特征通道压缩成通道表示符  $z^m$  和  $z^s$ 。这两个通道表示符在一定程度上代表着特征通道上的全局信息,其计算公式为:

$$z^m = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W x \quad (1)$$

$$z^s = \sqrt{\frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W (x - z^m)^2} \quad (2)$$

式中:  $x \in R^{C \times H \times W}$  为输入特征图;  $C$ 、 $H$  和  $W$  分别为图像的特征通道数、高度和宽度。

$z^m$  和  $z^s$  以不同的方式捕获全局信息,提高了全局信息的表征能力。首先将二者分别载入不同的一维卷积,通过给每个特征通道添加能够不断学习的注意力权重  $g^m \in R^{C \times 1 \times 1}$  和  $g^s \in R^{C \times 1 \times 1}$ ,建立通道间有效的联系以获得较为准确的通道相关性。

$$g^m = \sigma(F_m(z^m)) \quad (3)$$

$$g^s = \sigma(F_s(z^s)) \quad (4)$$

式中:  $F_m(\cdot)$  和  $F_s(\cdot)$  为一维卷积操作;  $\sigma(\cdot)$  为 sigmoid 函数。

将注意力权重加权到输入特征图上,重新定义在通道维度上的原始特征,得到增强的输出特征  $y \in R^{C \times H \times W}$ :

$$y = x \otimes g^m \otimes g^s \quad (5)$$

式中  $\otimes$  为逐通道乘法。

### 1.3 弱监督区域检测模块

由于高频工件种类繁多,类间差异较小,相似

工件之间的差异经常出现在特定的局部区域。因此,本文采用弱监督区域检测模块定位工件图像中有显著性差别的区域,以提升识别性能,如图 1 中橙色区域所示,该模块包括边界搜索和裁剪两种机制。

首先,由主干网络 ResNet50 生成全局输入图像的多层特征图  $M_g$ ,然后把所有通道的特征图相加得到能量图  $M_E$ 。为消除负元素的干扰,将  $M_E$  的所有元素归一化到  $[0, 1]$ 。

$$\hat{M}_E(i) = \frac{M_E(i) - M_{E,\min}}{M_{E,\max} - M_{E,\min}} \quad (6)$$

式中:  $M_{E,\max}$  和  $M_{E,\min}$  分别为  $M_E(i)$  中最大和最小元素的值;  $\hat{M}_E$  为缩放能量图。

为了更精确地定位,采用双线性插值将  $\hat{M}_E$  上采样到  $25 \times 25$  大小的能量图  $\bar{M}_E$ 。

$$\bar{M}_E = \text{bilinear}(\hat{M}_E) \quad (7)$$

在二维空间实现搜索较为复杂,所以将  $\bar{M}_E$  聚集成两个一维结构化能量向量:

$$\begin{cases} V_w = \sum_{i=0}^H \bar{M}_E(i, W) \\ V_h = \sum_{j=0}^W \bar{M}_E(H, j) \end{cases} \quad (8)$$

式中  $V_w$  和  $V_h$  分别为沿着空间宽度和高度方向的一维结构化能量向量。

以  $V_w$  为例,提出了包含宽度向量中不同元素能量的概念:

$$\begin{cases} E_{[0:W]} = \sum_{i=0}^W V(i) \\ E_{[x_1:x_2]} = \sum_{j=x_1}^{x_2} V(j) \end{cases} \quad (9)$$

式中:  $E[0:W]$  为宽度向量中所有元素的能量和;  $E[x_1:x_2]$  为沿空间宽度从  $x_1$  到  $x_2$  的区域能量。

设置一个阈值  $\gamma$  并定义全局图像中的关键区域为占据面积最小且满足如下条件:

$$\begin{cases} E[x_1:x_2]/E[0:W] > \gamma \\ E[y_1:y_2]/E[0:H] > \gamma \end{cases} \quad (10)$$

然后,利用边界搜索机制自动寻找该区域的宽度边界坐标  $[x_1 : x_2]$  和高度边界坐标  $[y_1 : y_2]$ 。最后,采用裁剪机制根据整个边界坐标  $[x_1, x_2, y_1, y_2]$  在原始图像中的位置截取有效的工件信息和有价值的背景信息,同时去除其他的噪声区域,得到工件局部图像  $I_2$ 。

### 1.4 分支融合模块

通过弱监督区域检测模块能够获得有价值的

局部图像,但忽略了全局图像在识别分类任务中的作用。为避免网络只关注图像的局部信息,本文采用分支融合模块,如图1中紫色区域所示。

本文所提算法中,两个通道注意力模块不共享参数,以提取不同尺度的工件特征。由于不同分支网络对同一类工件有不同的识别效果,本文采用分支融合模块兼顾两个分支的识别结果以进一步提高识别的准确率,融合的识别分数  $P$  的计算公式如下:

$$P = \mu P_1 + \lambda P_2 \quad (11)$$

式中:  $P_1$  为全局图像输入网络的识别结果;  $P_2$  为局部图像输入网络的识别结果;  $\mu$  和  $\lambda$  为权衡不同分支识别结果的平衡因子。

## 2 实验结果与分析

### 2.1 实验数据集

本文研究的实验数据来自某军工院所生产的高频工件图像,从中挑选了 20 000 张高频工件图像组成实验数据集,该数据集含有 20 个类别的工件。其中,每个类别的工件有 1 000 张图像,每张图像的尺寸均为  $3\ 822 \times 2\ 702$  像素;将数据集按 7:3 的比例随机分成由 14 000 张和 6 000 张的高频工件图像组成的训练集和验证集。

### 2.2 实验设置

实验所用的计算机配置如下: CPU 为 Intel(R) Core(TM) i5-10400F, GPU 为 NVIDIA GeForce GTX 1660 SUPER, 运行内存为 16 GB, Windows10 系统。本文采用 Python3.6, 并选用 PyTorch1.2 深度学习框架, CUDA 10.2 深度学习网络加速库。网络参数设置中本文使用 ImageNet 上的预训练权重初始化网络,选择 Adam 作为优化器,采用交叉熵损失作为损失函数, batch size 设定为 32, 迭代次数设置为 50 次, 迭代的学习率初始化为  $10^{-4}$ , 每迭代 20 次除以 10。

### 2.3 识别性能对比实验

为了验证本文提出的高频工件图像识别算法的识别效果,将其与其他多种算法进行对比实验。对比的算法包括: ResNet50、基于卷积神经网络的机械零件识别算法(记为 WorkNet-2)、基于深度学习的主轴承盖零件识别算法(记为 MFF-CNN)、基于改进卷积神经网络的零件识别算法(记为 AP-Xception)。实验结果如表 1 所示,图 4 所示为本文算法识别结果的混淆矩阵。

表 1 不同算法的工件识别结果

算法	图像总量/张	正确识别/张	准确率/%
ResNet50	6 000	4 818	80.3
WorkNet-2	6 000	5 370	89.5
MFF-CNN	6 000	5 424	90.4
AP-Xception	6 000	5 472	91.2
本文算法	6 000	5 688	94.8

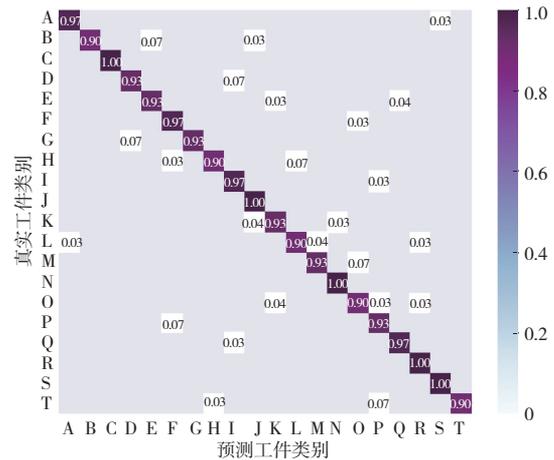


图 4 本文算法的测试集混淆矩阵

从表 1 的实验结果中可以看出,本文算法的图像识别性能明显优于其他对比算法。相比于 ResNet50、WorkNet-2、MFF-CNN 和 AP-Xception,本文算法分别将准确率提高了 14.5 个百分点、5.3 个百分点、4.4 个百分点和 3.6 个百分点。从图 4 的实验结果中可以看出,本文算法对每一类别的高频工件识别准确率都在 90% 以上。

### 2.4 模型参数选择实验

本文所提算法的裁剪范围阈值  $\gamma$ , 决定了从工件全局图像中提取有效区域的大小, 进一步影响着工件识别的准确率。此外, 式(11)中的参数  $\mu$  和  $\lambda$  较大程度影响着网络对不同分支的侧重。

如果裁剪范围阈值太小, 会导致工件特征信息丢失过多; 如果裁剪范围阈值太大, 将会使得网络无法聚焦于重要的局部特征。因此, 截取的区域应该被限制在一个合理的范围之内。为了确定最佳的阈值, 实验中先将平衡因子  $\mu$  和  $\lambda$  分别设置为 0.6 和 0.4, 再测试  $\gamma \in \{0.60, 0.65, 0.70, 0.75, 0.80\}$  时高频工件的识别准确率, 实验结果如表 2 所示。

表 2 阈值  $\gamma$  的大小对识别结果的影响

$\gamma$ 的取值	准确率/%
0.60	92.4
0.65	94.1
0.70	94.8
0.75	93.5
0.80	92.6

从表 2 中可以看出,随着裁剪范围阈值的增大,准确率先不断提升然后不断下降。当阈值  $\gamma$  为 0.70 时,本文算法达到了最佳性能。因此,本文选取  $\gamma=0.70$  作为最终的裁剪范围阈值。

为了评估平衡因子值  $\mu$  和  $\lambda$  对识别结果的影响,实验中分别测试了  $\gamma$  为 0.70,  $\mu$  和  $\lambda$  取不同值时算法对高频工件的识别准确率,实验结果如表 3 所示。

表 3 平衡因子  $\mu$  和  $\lambda$  的大小对识别结果的影响

$\mu, \lambda$ 的取值	准确率/%
$\mu=1.0, \lambda=0$	92.2
$\mu=0.9, \lambda=0.1$	92.5
$\mu=0.8, \lambda=0.2$	93.0
$\mu=0.7, \lambda=0.3$	93.6
$\mu=0.6, \lambda=0.4$	94.8
$\mu=0.5, \lambda=0.5$	93.1
$\mu=0.4, \lambda=0.6$	92.7
$\mu=0.3, \lambda=0.7$	92.3
$\mu=0.2, \lambda=0.8$	91.9
$\mu=0.1, \lambda=0.9$	91.4
$\mu=0, \lambda=1.0$	90.7

从表 3 中可以看出,当  $\mu=0.6, \lambda=0.4$  时,高频工件识别取得了最高的准确率。随着  $\mu$  的减小和  $\lambda$  的增大,工件识别的效果逐渐降低。这是因为裁剪的局部图像包含的信息较少,过于依赖该部分会造成网络无法获得较好结果。当  $\mu$  较小  $\lambda$  较大时,容易忽视截取的关键区域从而导致准确率降低。因此,本文选取  $\mu=0.6, \lambda=0.4$  作为平衡不同分支网络的参数。

### 2.5 消融实验

为了验证算法中的各个模块对网络性能的影响,在保留 ResNet50 网络的基础上,通过控制变量去掉各项改进设计消融实验,实验结果如表 4 所示。

表 4 不同模块对网络性能的影响

通道注意力模块	弱监督区域检测模块	分支融合模块	准确率/%
×	×	×	80.3
✓	×	×	92.2
✓	✓	×	93.7
✓	✓	✓	94.8

注:×表示没有采用该模块;✓表示采用该模块。

从表 4 的实验结果中可以看出,首先在主干网络 ResNet50 的基础上加入通道注意力模块后,准确率提升了 11.9 个百分点,这表明网络具有对不同颜色通道特征信息的感知能力,克服了光照变化的影响,更有效地提取到了高频工件图像的特征;然后加入弱监督区域检测模块,准确率进一步提升了 1.5 个百分点,这表明网络的特征学习专注于有区别的有效区域,能够提升识别的准确率;最后采用分支融合模块,准确率再进一步提升了 1.1 个百分点。这表明结合全局信息和局部信息能够提升高频工件的识别性能。

### 2.6 可视化实验

为了验证所提算法的有效性,在同一测试集上随机选取一些图像进行可视化展示。图 5 展示了不同类别工件的细微差异和算法改进前后注意力特征的可视化图。

从图 5 中可以看出,与改进前的算法相比,改进后的算法使网络更加关注于凸台与其他工件有差异的区域。因此,所提算法能够提取到区分性更强的工件图像特征,显著提高了网络的识别性能。

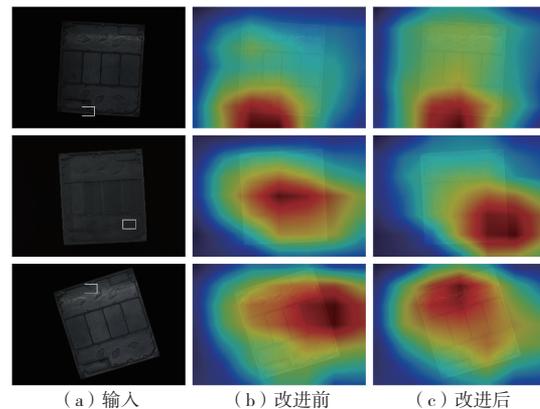


图 5 注意力特征可视化

## 3 结语

本文提出了一种自适应光照变化的高频工件图像识别算法。首先通过引入通道注意力模块增强了网络的特征提取能力,降低了因光照变化对

高频工件识别效果的影响;然后针对工件类间差异小的特点,采用弱监督区域检测模块搜索和定位有区别的局部图像;最后结合分支融合模块使网络有效地权衡工件的全局特征和局部特征。实验结果表明:与原始的 ResNet50 网络相比,本文算法对高频工件识别准确率提高了 14.5 个百分点,具有更强的特征提取能力。此外,与多种图像识别算法相比,本文算法对高频工件识别准确率均有一定程度的提升。实验结果证明了本文算法的有效性。未来可进一步探索更高效的网络结构和识别算法。

### 参考文献:

- [1] ZHOU JZ, WEN X. Research on influencing factors and multiple driving paths of intelligent transformation in China's manufacturing industry [J]. Journal of Computational Methods in Sciences and Engineering, 2021, 21(5): 1561-1573.
- [2] 李昌明,李东年,陈成军,等. 基于卷积神经网络与虚拟训练集的零件识别[J]. 组合机床与自动化加工技术, 2021(8): 40-43.
- [3] 张科研,孙继旋,庞启硕. 工件形状的激光雷达三维扫描识别方法[J]. 激光杂志, 2021, 42(5): 133-137.
- [4] MARCATO M. The Made in China 2025 amid hyperglobalization: upgrading, intangible assets, and internationalization strategies [J]. Economia e Sociedade, 2022, 31: 355-384.
- [5] 杨涛,欧阳,苏欣,等. 联合损失监督的高频工件深度学习识别算法[J]. 机械制造与自动化, 2023, 52(1): 30-33, 47.
- [6] DUAN S L, YIN C C, LIU M M. Recognition algorithm based on convolution neural network for the mechanical parts [C]//International Workshop of Advanced Manufacturing and Automation. Singapore: Springer,

2019:337-347.

- [7] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [8] GAI R L, CAI J R, WANG S Y. Research review on image recognition based on deep learning[J]. Journal of Chinese Computer Systems, 2021, 42(9): 1980-1984.
- [9] 李奇,王宇钢. 基于分块 PCA 和 SVM 的零件识别分类系统[J]. 机械工程与自动化, 2021(4): 21-23, 26.
- [10] CORTES C, VAPNIK V. Support-vector networks [J]. Machine Learning, 1995, 20(3): 273-297.
- [11] 张鹏飞,石志良,李晓垚,等. 基于深度学习的主轴轴承分类识别算法[J]. 图学学报, 2021, 42(4): 572-580.
- [12] 杨乐,甘忠,李义龙,等. 基于改进卷积神经网络的零件识别[J]. 仪表技术与传感器, 2022(5): 82-87.
- [13] 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). Las Vegas, NV, USA: IEEE, 2016: 2818-2826.
- [14] CHOLLET F. Xception: deep learning with depthwise separable convolutions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 1800-1807.
- [15] 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). Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [16] WANG Q, HUANG W, XIONG Z T, et al. Looking closer at the scene: multiscale representation learning for remote sensing image scene classification [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(4): 1414-1428.

收稿日期: 2023-07-10

(上接第 200 页)

### 参考文献:

- [1] 王佐才,丁雅杰,戈壁,等. 桥梁结构非线性模型修正研究综述[J]. 交通运输工程学报, 2022, 22(2): 59-75.
- [2] 方圣恩,林友勤,夏樟华. 考虑结构参数不确定性的随机模型修正方法[J]. 振动·测试与诊断, 2014, 34(5): 832-837, 973.
- [3] 秦仙蓉,詹彦明,赵书振,等. 基于替代模型的岸桥随机有限元模型修正[J]. 振动与冲击, 2020, 39(1): 43-48.
- [4] 林光伟,张熠. 不同结构复杂度下结合集成学习的模型修正方法[J]. 工程力学, 2022, 39(增刊1): 153-157.
- [5] 方圣恩,张秋虎,林友勤,等. 参数不确定性估计的随机响应面模型修正方法[J]. 振动工程学报, 2016,

29(4): 594-602.

- [6] WU X, KOZLOWSKI T. Inverse uncertainty quantification of reactor simulations under the Bayesian framework using surrogate models constructed by polynomial chaos expansion [J]. Nuclear Engineering and Design, 2017, 313: 29-52.
- [7] 韩忠华. Kriging 模型及代理优化算法研究进展[J]. 航空学报, 2016, 37(11): 3197-3225.
- [8] 平梦浩,张文华,唐亮. 不确定性结构的非高斯模态参数模拟方法[J]. 华南理工大学学报(自然科学版), 2024, 52(9): 81-92.
- [9] 陈承滨,余岭,潘楚东,等. 基于蚁狮优化算法与迹稀疏正则化的结构损伤识别[J]. 振动与冲击, 2019, 38(16): 71-76, 99.

收稿日期: 2024-10-16