

朱劲松,李欢,王世芳. 基于卷积神经网络和迁移学习的钢桥病害识别[J]. 长安大学学报(自然科学版),2021,41(3):52-63.

ZHU Jin-song, LI Huan, WANG Shi-fang. Defect recognition for steel bridge based on convolutional neural network and transfer learning[J]. Journal of Chang'an University(Natural Science Edition), 2021, 41(3): 52-63.

DOI:10.19721/j.cnki.1671-8879.2021.03.006

# 基于卷积神经网络和迁移学习的钢桥病害识别

朱劲松<sup>1,2</sup>, 李欢<sup>1</sup>, 王世芳<sup>3</sup>

(1. 天津大学 建筑工程学院, 天津 300072; 2. 天津大学 滨海土木工程结构与安全教育部重点实验室, 天津 300072; 3. 安徽工程大学 电气工程学院, 安徽 芜湖 241000)

**摘要:**针对钢桥病害识别效率低、精度不高的现状,提出了一种基于深度学习的钢结构表观病害识别方法。该方法将卷积神经网络 Inception-v4 和迁移学习相结合,分别采用迁移学习中特征提取和微调 2 种训练方式获得 2 种模型,并与全新训练的 Inception-v4 模型进行对比。首先,收集 656 幅钢桥病害图像,包括涂层劣化 176 幅,腐蚀 173 幅,焊缝开裂 151 幅,完好 156 幅,对其进行旋转、翻转、调整对比度等预处理,使数据集扩充至 3 742 幅,按照 8:1:1 的比例,划分训练集、验证集和测试集;然后,采用 3 365 幅钢桥病害图像分别对特征提取模型、微调模型与全新训练模型进行训练和验证,对比了批大小(batch-size)和学习率对模型训练效果的影响,并对这 2 个参数进行了优选;最后,采用 377 幅病害图像进行测试,得到特征提取模型、微调模型和全新训练模型训练一个时期(epoch)的时间分别为 47.2、119.2、121.8 s,测试正确率分别为 89.39%、97.88%与 91.25%。结果表明:迁移学习的 2 种模型较全新训练模型,减少了数据的需求量,提高了运行效率和病害识别准确率;微调模型经历较少的 epoch,便可达到较高的测试准确率,更适于钢桥病害识别的实际应用。

**关键词:**桥梁工程;钢桥;病害识别;卷积神经网络;迁移学习;Inception-v4;图像处理

中图分类号:U446

文献标志码:A

文章编号:1671-8879(2021)03-0052-12

## Defect recognition for steel bridge based on convolutional neural network and transfer learning

ZHU Jin-song<sup>1,2</sup>, LI Huan<sup>2</sup>, WANG Shi-fang<sup>3</sup>

(1. School of Civil Engineering, Tianjin University, Tianjin 300072, China; 2. Key Laboratory of Coast Civil Structure Safety, Ministry of Education, Tianjin University, Tianjin 300072, China; 3. School of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, Anhui, China)

**Abstract:** In view of the low efficiency and accuracy of steel bridge defect recognition, a steel structure apparent defect identification method based on deep learning was proposed. In this method, convolutional neural network Inception-v4 and transfer learning were combined, and two training methods of feature extraction and fine-tuning in transfer learning were used to obtain two

收稿日期:2021-02-19

基金项目:天津市交通运输委员会科技项目(2018-29);国家重点研发计划项目(2018YFB1600300,2018YFB1600301)

作者简介:朱劲松(1975-),男,安徽池州人,教授,博士研究生导师,E-mail:jszhu@tju.edu.cn.

通讯作者:王世芳(1981-),女,安徽芜湖人,讲师,E-mail:307883240@qq.com.

models, and an original Inception-v4 model was trained for comparison. Firstly, 656 images of steel bridge diseases were collected, including 176 images of coating deterioration, 173 images of corrosion, 151 images of weld cracking and 156 images of intact. The data set was expanded to 3 742 images by rotating, flipping and adjusting contrast. According to the ratio of 8 : 1 : 1, the training set, verification set and test set were divided. Secondly, 3 365 steel bridge defect images were used to train and verify the feature extraction model, fine-tuning model and new training model respectively. The effects of batch-size and learning rate on the training effect of the model were compared, and the two parameters were optimized. Thirdly, 377 defect images were used to test these models. The results show that the epoch training time of feature extraction model, fine-tuning model and new training model was 47.2, 119.2 and 121.8 s, and the test accuracy was 89.39%, 97.88% and 91.25%, respectively. Compared with the new training model, the two models of transfer learning reduce the defect of data, improve the operation efficiency and the accuracy of defect recognition, but the fine-tuning model can achieve higher test accuracy with less epoch, which is more suitable for the practical application of steel bridge defect recognition.

**Key words:** bridge engineering; steel bridge; defect recognition; convolutional neural network; transfer learning; Inception-v4; image processing

## 0 引言

钢结构具有强度高、自重轻、跨越能力大等特点,能够满足桥梁对跨度、复杂性和美观性的要求,在桥梁建设中得到广泛应用。但钢桥在使用过程中,长期受到自然侵蚀和人为作用,易出现多种病害<sup>[1-2]</sup>,危及行车安全,因此,对钢桥进行及时有效的病害检测,以便采取合理的维护措施具有重大意义。传统的钢桥病害检测主要依赖于人工检测和无损检测<sup>[3]</sup>,人工检测耗时耗力,且需要检测人员有较高的专业素养,主观性较强,只能做定性的检查;无损检测,虽在一定程度上克服了由于经验差异导致的漏判与误判,但作业成本高,检测技术难以推广,有些技术尚待完善,影响检测结果的准确性。因此,需寻求一种新的智能化识别技术,来减小工作难度,降低工作成本,同时实现高效、精确的病害识别。

计算机视觉和图像处理技术的发展,为病害智能识别提供了动力,尤其是深度学习的出现,其可以对输入数据逐层抽取,实现了从底层信号到高级语义的良好映射,进一步推动了表观病害识别的发展。沙爱民等将卷积神经网络(CNN)技术引入基于图像分析的路面病害识别与测量,建立了病害识别模型,训练裂缝特征提取模型和坑槽特征提取模型,提高了路面病害检测识别的速度和精度,并建立了级联卷积神经网络,实现了自动识别探地雷达图像所反映路基病害的任务<sup>[4-5]</sup>。冯微等构建基于深度置

信网络的车撞桥墩损伤等级判别模型,在车撞桥墩多种损伤等级判别方面具有良好的准确性和泛化性<sup>[6]</sup>。薛亚东等提出基于深度学习的隧道病害识别方法,针对 GoogLeNet 模型,优化其卷积核,并改进 Inception 模块与网络结构,建立隧道衬砌特征图像分类系统,克服了复杂背景的影响<sup>[7]</sup>。刘新根等提出基于深度学习的隧道衬砌裂缝自动识别算法,并使用 2 个深度卷积神经网络分别实现隧道衬砌图像分类和裂缝识别,提高了隧道衬砌裂缝的识别精度和速度<sup>[8]</sup>。孙朝云等提出基于深度卷积神经网络融合模型的路面裂缝识别方法,其方法将 SSD 网络和 U-Net 网络结合,在识别裂缝类别的同时,可以得到病害的位置<sup>[9]</sup>。李海丰等针对机场道面病害,提出基于注意力机制与特征融合的深度神经网络模型 DetMSPNet,提高了复杂背景下的检测精度<sup>[10]</sup>。Yun 等提出基于深度神经网络的金属表面缺陷检测系统,并在该系统中提出卷积变分自动编码器(CVAE)的数据增强方法,来解决数据不平衡的问题,并基于深层卷积神经网络的缺陷分类算法,实现缺陷的检测分类<sup>[11]</sup>。Cheng 等提出了一种基于快速 R-CNN 的污水管道缺陷自动检测方法,探究了数据集大小、初始化网络及超参数对模型的影响,实现了污水管道缺陷的高精度、快速检测<sup>[12]</sup>。Bastian 等提出了一种基于计算机视觉的水、油、气管道腐蚀检测方法,并采用自主设计卷积神经网络对管道腐蚀图像进行分类,获得了 98.8% 的分类精度<sup>[13]</sup>。

在桥梁病害检测方面, Ali 等提出将红外热像仪(IRT)与深层起始神经网络(DINN)相结合, 来检测钢桁架桥梁亚表面损伤, 该方法提高了检测效率, 降低了计算成本, 但无法区分不同类型的病害<sup>[14]</sup>。Cha 等提出基于快速 R-CNN 的结构视觉检测方法, 其能检测多种损伤, 平均准确率为 87.8%<sup>[15]</sup>。Zhu 等提出一种基于深度学习的水泥混凝土桥面缺陷智能分类系统, 该系统替换 VGG-16 的 Softmax 分类层, 得到改进的 VGG-16 网络, 实现了 7 种缺陷的准确分类<sup>[16]</sup>。沈浩等提出基于深度学习的钢桁架桥螺栓病害智能识别方法, 并利用无人机航拍视频获取图像, 采用迁移学习引入深度学习模型 Inception-v3, 实现了螺栓病害识别的自动化、智能化<sup>[17]</sup>。李雪松等提出基于卷积神经网络的桥梁损伤识别方法, 提高了损伤识别准确率和噪声鲁棒性<sup>[18]</sup>。Xu 等提出了一种改进的融合卷积神经网络结构, 来识别钢箱梁表面疲劳裂纹, 克服了一些干扰信息的影响, 提高了识别精度<sup>[19]</sup>。阮小丽等采用深度学习算法 DeepLabV3+ 模型对桥梁结构表面病害进行识别量化, 提升了对桥梁病害的检测准确度和全面度<sup>[20]</sup>。

基于深度学习的病害检测识别方法, 虽已在钢结构和桥梁中得到应用, 但是其在识别速度和精度上还有待进一步提高, 且目前关于钢桥的病害识别

研究较少。针对以上现状, 本文提出基于卷积神经网络和迁移学习的钢桥病害识别方法, 将在 ImageNet 上预训练的 Inception-v4<sup>[21]</sup> 模型进行迁移, 采用特征提取和微调 2 种方式分别进行训练, 对钢桥常见的涂层劣化、腐蚀、焊缝开裂 3 种病害及完好图像进行识别分类, 并用全新训练的 Inception-v4 模型与之对比, 证明了迁移学习模型在识别速度和精度上有较大的提升, 尤其是微调模型的效果最佳。该方法实现了快速高效的病害检测, 进一步推动了钢结构桥梁智能检测的发展。

## 1 钢桥病害数据集的构建

### 1.1 图像数据的采集

图像数据采用佳能 EOS 7D 数码相机进行拍摄, 分辨率为 5 184 像素×3 456 像素, 拍摄地点为天津市海河上各座钢桥, 拍摄时间为一天中的不同时刻, 以便获取不同光照下的病害图像, 降低光照对训练和测试结果的影响, 对于隐蔽位置的病害, 采用补光灯增加亮度, 以便拍摄。病害图像数据的获取, 除自行拍摄外, 另有少量病害图像由桥梁检测单位提供。本文收集涂层劣化 176 幅, 腐蚀 173 幅, 焊缝开裂 151 幅, 完好 156 幅, 共计 656 幅, 图像数据示例如图 1 所示。

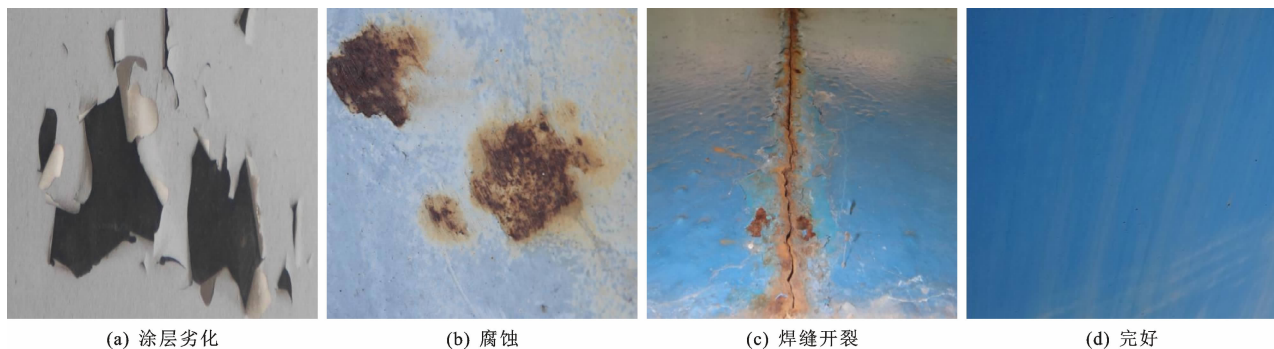


图 1 钢桥图像数据示例

Fig. 1 Examples of steel bridge image data

### 1.2 图像预处理及数据集的建立

为提高模型的鲁棒性, 需对原始数据进行预处理, 先将采集到的图像进行尺寸量纲一化处理, 使其大小一致, 然后对图像进行数据增强, 最终形成本文试验所需的病害数据集。尺寸量纲一化的目的是为了在进行下一步处理时统一图像样本, 便于数据扩增, 本文采用 Python 语言库中的 PIL 库进行尺寸量纲一化处理, 将原始图像缩放至 512 像素×512 像素大小。

图像增强的作用包括: ①增加图像的数量, 缓

解病害样本少的情况; ②丰富数据的多样性, 抵消图像采集中不同亮度和方向的影响, 提高对未知样本的预测能力, 以增强模型的鲁棒性和适应性。常用的数据增强方法有: 平移、旋转、翻转、调整亮度、添加噪声等。为避免对钢桥病害颜色特征产生影响, 本文采用旋转(90°、180°)、翻转、调整亮度(亮度增加、亮度降低)3 种方法, 对数据进行增强。以腐蚀病害为例, 其增强处理后的图像如图 2 所示。

通过数据增强, 图像数量进一步得到了扩充, 剔

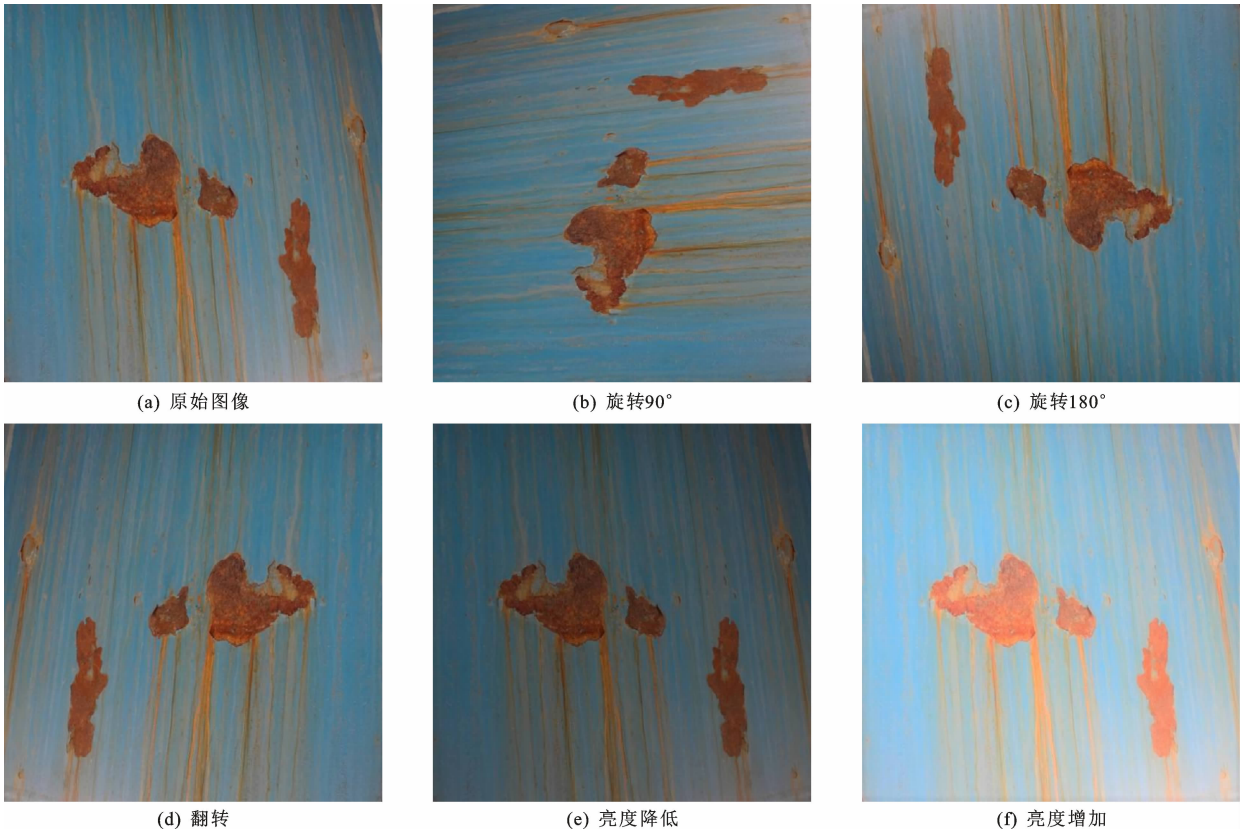


图2 腐蚀图像增强示例

Fig. 2 Examples of corrosion image enhancement

除数据中病害特征不明显的图像后,共获取图像3 742幅,图3为增强后数据在各个病害类别间的分布情况。为方便模型训练,用0、1、2、3作为数据集的标签,0代表涂层劣化、1代表腐蚀、2代表焊缝开裂、3代表完好。依据增强后的图像划分数据集,除训练集和测试集外,增加了验证集对模型进行参数调整,按8:1:1的比例对数据进行随机划分,得到训练集、验证集和测试集。

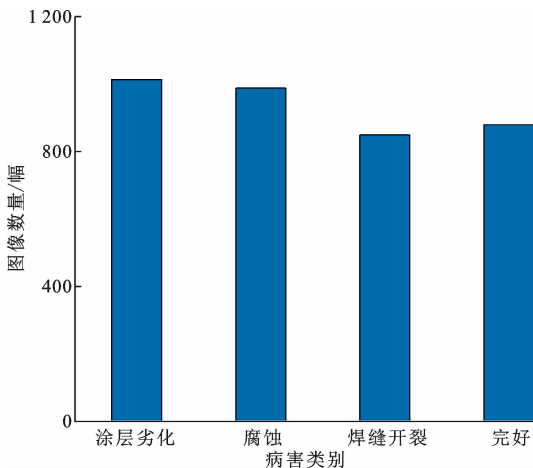


图3 数据增强后图像数量

Fig. 3 Image numbers after data enhancement

## 2 钢桥病害识别模型的搭建

### 2.1 迁移学习

近年来,计算机视觉在一些领域已经取得了较大的成就,但是这些成功的背后是大量数据的支撑,尤其是深度神经网络的训练,若要取得良好的效果,需要几十万甚至上百万的数据,而在图像分类领域,数据较为珍贵,通常只能收集少量的图像,为此可采用迁移学习<sup>[22]</sup>的方法,来解决数据量不足的问题。

迁移学习是将一个预先训练好的模型进行另一种训练,用在新的任务中。其主要目标是利用在源域数据集中训练好的模型参数,来帮助训练目标域中的数据集,这需要给定源域和源域的任务,目标域和目标域的任务,且源域和目标域需存在一定的关联,两者之间相似度越大,迁移效果越好。迁移学习的网络在源域上已经得到训练,在进行目标域的训练时,对数据量的要求不高,且从已有网络的预训练权值开始训练,降低了训练成本,可以显著提高训练效率。迁移学习常见的4种实现方式<sup>[23]</sup>包括:①样本迁移,即在数据集(源领域)中找到与目标领域相似的数据,将该数据放大多倍,与目标领域的数



分为热裂纹、冷裂纹、再热裂纹和层状撕裂,具有尖锐的缺口和大长宽比的特征。由前面卷积层获得较小的感受野,即输出图像的每个像素只利用输入图像很小的一个范围,后面的卷积层感受野逐层加大,用于捕获图像更复杂,更抽象的信息,经过多个卷积层的

运算,得到图像在不同尺度的抽象表示,进行病害特征提取,最后由全连接层完成病害的分类。在运行过程中,预训练参数可以简化模型参数初始化的过程,有助于处理小数据集中出现的训练饱和问题,同时极大提升了训练效率。

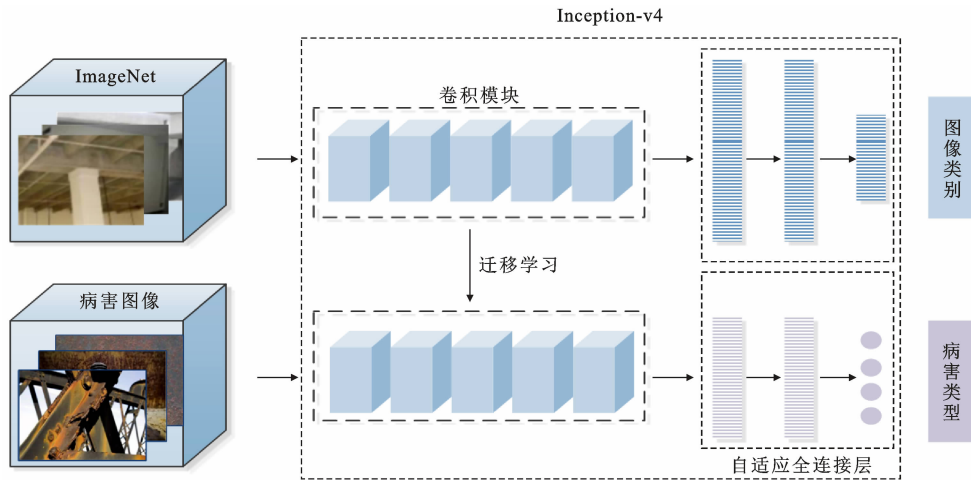


图 5 钢桥病害识别的迁移学习模型框架

Fig. 5 Framework of transfer learning model for steel bridge defect identification

### 3 模型训练和试验分析

#### 3.1 试验环境与设计

表 1 为试验的软件及硬件配置,采用 PyTorch 作为钢桥病害识别模型搭建和训练的平台,其源码易于阅读,设计简洁,可以动态地设计网络,且运行速度快,能更好地完成模型的搭建运行;PyCharm 编辑器由 JetBrains 软件开发公司开发,界面功能强大,合并了多个库(如 Matplotlib 和 NumPy),同时允许创建不同的虚拟环境进行编程,提升了环境管理效率。模型在 GPU 环境下运行,GPU 具有更高的速率和更多的计算单元,通过并行计算处理矩阵运算,能提升运算速度,降低模型训练的时间。

表 1 试验软硬件配置

Tab. 1 Experimental software and hardware configuration

名称	试验软硬件配置
操作系统	Windows10 (64 bit)
编程语言	Python 3.8
深度学习框架	PyTorch 1.7.0
CPU	intel core i5-10300H
GPU	NVIDIA GeForce RTX 2060(6 G)
Cuda	Cuda-11.0
环境管理	Pycharm 2020.2

首先采用特征提取和微调 2 种方式,对迁移模型进行训练。特征提取即将全连接层之前网络中的

卷积模块冻结,提取全连接层之前最后 1 层的张量作为训练数据的瓶颈特征<sup>[26]</sup>,并由分类器进行训练完成分类;微调则会重新训练迁移模型的卷积模块,在每个训练阶段进行一次前向传播。完成迁移模型的训练后,对未经预训练的 Inception-v4 模型,用钢桥病害图像进行全新训练,最后将 3 种模型的训练及测试结果进行对比分析。模型训练次数(epoch)均为 300,采用随机梯度下降(SGD)优化算法和动量法加速<sup>[27]</sup>,动量参数设为 0.9,避免了更新幅度摆动大的问题,同时加快了网络的收敛速度。调整模型的批处理大小(batch-size)和学习率,对模型进行训练,考虑到硬件设备的性能及训练效果,本文只探讨了 batch-size 对采用特征提取训练方式模型的影响。在初始学习率为 0.01 时,将 batch-size 的值设为 16、32、64 进行训练,其余 2 个模型的 batch-size 设置为 8,确定模型的 batch-size 后,将 3 种模型分别在学习率为 0.1、0.01、0.001、0.000 1 下进行训练,探究模型的最佳学习率,得到最终训练模型。用测试集图像在训练好的模型上进行测试,获取测试准确率。

#### 3.2 模型训练结果

##### 3.2.1 batch-size 的影响

当训练集较大时,若一次性将样本输入到神经网络,会导致计算量非常大,增加内存的要求,也会使误差陷入局部极小值;若每次只读取一个样本,在

每个训练样本上得到的目标函数值差别可能较大,模型对样本的泛化能力差。为此需要设置一个参数,将样本按批量输入网络中学习,在简化参数更新的同时避免陷入早期的局部极小值。采用 GPU 对模型训练,在 GPU 中内存以二进制方式存储数据,为训练时能够充分利用 GPU 的矩阵运算特点加速模型收敛。将 batch-size 设置成 2 次幂的形式,分别选择 16、32、64 作为试验的 batch-size 值,分析其对特征提取模型的影响,不同 batch-size 下的准确率曲线和损失曲线如图 6 所示。

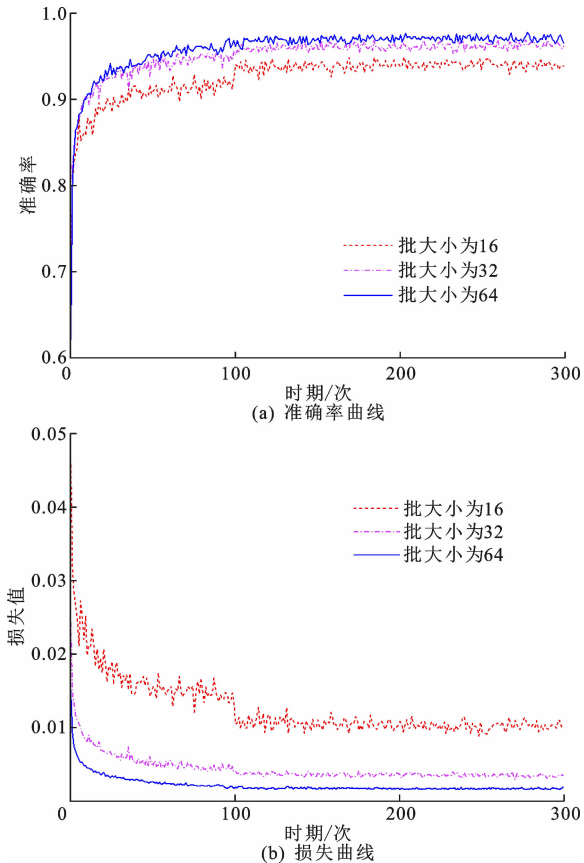


图 6 不同 batch-size 下的准确率曲线和损失曲线

Fig. 6 Accuracy curves and loss curves under different batch-sizes

由图 6 可知:随着 batch-size 数值增大,准确率曲线上升速度加快,所达到的最终准确率增加,损失曲线下降速率加快,所达到的最终损失值减少,且准确率曲线和损失曲线震荡幅度明显降低;但 batch-size 数值由 32 增加至 64 时,准确率曲线和损失曲线间的间距明显小于由 16 增加至 32 时,准确率曲线增加的幅度和损失曲线减少的幅度均在缩小。表 2 为不同 batch-size 下的准确率。由表 2 可知:当 batch-size 数由 16 增加至 32 时,训练准确率增加了 2.34%,测试准确率增加了 2.39%;而由 32 增加至 64 时,训练准确率仅增加了 0.35%,测试准确率增

加了 0.26%,增加量减少。若继续增大 batch-size 值,准确率提升效果可能不明显,甚至不再提升,误差陷入局部极小值,由于显卡内存原因,本文未继续增大 batch-size 值进一步验证,经对比后,选取 64 作为特征提取模型的 batch-size 值。

表 2 不同 batch-size 下的准确率

Tab. 2 Accuracies under different batch-sizes

batch-size	16	32	64
训练准确率/%	93.88	96.22	96.57
测试准确率/%	85.94	88.33	88.59

### 3.2.2 学习率的影响

学习率作为深度学习的重要超参数之一,有助于通过损失函数的梯度来调整网络权重,决定了神经网络是否可以收敛到全局最小值。当学习率设置的过大时,会使损失函数直接越过全局最优点,梯度在最小值附近震荡,无法收敛;当学习率设置过小时,损失函数的变化速度较慢,会大大增加网络的收敛复杂度,且易使网络困在局部最小值或者鞍点。

在学习率为 0.1、0.01、0.001、0.000 1 下,3 种模型的训练和测试准确率,如表 3 所示。可知,当学习率为 0.1 时,由于损失函数的梯度震荡幅度大,3 种模型的训练和测试准确率均较低。当学习率缩小 10 倍后,3 种模型的准确率增加,尤其是全新训练模型和微调模型,测试准确率分别增加了 10.87% 和 5.84%,且模型逐渐向最优点靠近。当学习率继续缩小 10 倍后,3 种模型的训练和测试准确率继续增加,但增量在 1% 左右,增幅较少,说明模型已接近最优点。当学习率缩小为 0.000 1 时,3 种模型的训练和测试准确率均出现不同程度的降低,其中全新训练模型的准确率下降量在 2% 之内,降幅最少;特征提取模型训练准确率下降 10.13%,测试准确率下降 6.63%,降幅最大;微调模型的训练准确率虽无明显降低,但测试准确率下降 4.25%,可能出现了过拟合情况。由此可见,当学习率为 0.000 1 时,模型逐渐偏离全局最优点,甚至出现过拟合现象,效果不佳。根据以上对比,当学习率为 0.001 时,3 种模型的训练效果最好,准确率均为最高,故本文学习率选取为 0.001。

### 3.3 模型对比分析

#### 3.3.1 模型运行效率分析

经过第 3.2 节的训练,3 种模型的超参数均已确定,其最终训练、测试准确率及模型训练一个时期(epoch)所用的时间如表 4 所示。观察 epoch 时间可知,特征提取模型因冻结了卷积模块,避免了对

表 3 不同学习率下模型的训练和测试准确率

Tab. 3 Training and testing accuracies of models under different learning rates

模型类型	学习率	训练准确率/%	测试准确率/%
全新训练模型	0.100 0	94.38	80.11
	0.010 0	94.45	90.98
	0.001 0	95.06	91.25
	0.000 1	94.28	89.66
特征提取模型	0.100 0	94.85	87.27
	0.010 0	96.57	88.59
	0.001 0	97.76	89.39
	0.000 1	87.63	82.76
微调模型	0.100 0	99.93	91.51
	0.010 0	100.00	97.35
	0.001 0	100.00	97.88
	0.000 1	99.97	93.63

表 4 不同模型的准确率及 epoch 时间

Tab. 4 Accuracies and epoch time of different models

模型类型	全新训练模型	特征提取模型	微调模型
训练准确率/%	95.06	97.76	100.00
测试准确率/%	91.25	89.39	97.88
epoch 时间/s	121.8	47.2	119.2

卷积层参数的训练,特征提取过程较快,其训练时间最短;微调模型和全新训练模型需对卷积模块进行训练,在每个训练阶段都要经过一次前向传播,耗时长,且微调模型的单个 epoch 时间较全新训练模型相差无几,两者的 epoch 时间约为特征提取模型的 2.5 倍。

虽然微调模型的单个 epoch 时间较全新训练模型没有显著减少,但 3 种模型的准确率曲线和损失曲线,微调模型训练 25 次左右,准确率曲线已经达到稳定,且损失曲线下落较快,准确率曲线和损失曲线较光滑,训练准确率最高,损失最低;而另外 2 种模型,均在 100 个 epoch 后才达到稳定,准确率曲线和损失曲线较微调模型震荡幅度大,见图 7。因此,综合考虑模型一个 epoch 所用的时间和达到稳定时的 epoch 个数,可得迁移学习的 2 种模型均比全新训练模型的运行效率高,且微调模型的效率提升最为显著,鲁棒性更好。

### 3.3.2 精确率、召回率和 $F_1$ 值分析

使用准确率评价模型存在一个问题,即当数据的类别不均衡,特别是有极偏的数据存在情况下,准确率这个评价指标不能客观评价模型的优劣。例如样本中有 99 个负例,1 个正例,模型在预测时,将 100 个样本都预测为负例,准确率即可达到 99%,仅依据准确率进行评价存在一定的不合理性。鉴于此,引入精确率  $P$ 、召回率  $R$  和  $F_1$  值指标,对模型

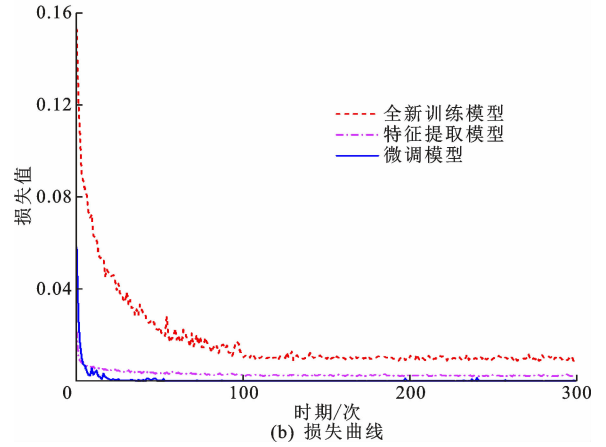
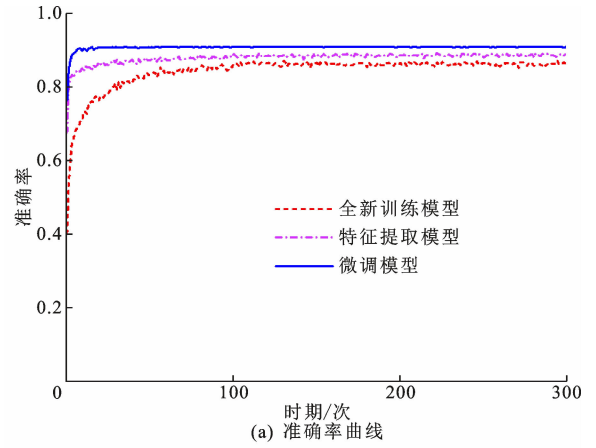


图 7 三种模型的准确率曲线和损失曲线

Fig. 7 Accuracy curves and loss curves of three models

做进一步的分析。精确率是指预测为正例的数据里,预测正确的数据比例;召回率是指真实为正例的数据里,预测正确的数据比例; $F_1$  值又称  $F_1$  分数,同时兼顾了分类模型的精确率和召回率,可看作是模型精确率和召回率的一种加权平均,计算分别为

$$P = \frac{T_p}{T_p + F_p} \quad (1)$$

$$R = \frac{T_p}{T_p + F_n} \quad (2)$$

$$F_1 = \frac{2PR}{P+R} \quad (3)$$

式中: $T_p$  表示真正例,即真实类别为正例,预测类别为正例; $F_p$  表示假正例,即真实类别为负例,预测类别为正例; $F_n$  表示假负例,即真实类别为正例,预测类别为负例。

以上公式主要用来评估二分类模型,为了测量多分类任务中不同模型之间的性能差异,采用宏平均值,即所有类别的每一个统计指标值的算数平均值进行评价。首先,计算每个混淆矩阵的精确度  $P_i$  和召回率  $R_i$ ,将其表示为  $(P_1, R_1), (P_2, R_2), \dots, (P_n, R_n)$ 。然后,通过计算各精确度和召回率的平

均值,得到宏精确率  $M_p$ 、宏召回率  $M_r$  和宏  $F_1$  值  $M_{F_1}$ ,计算如下所示<sup>[28]</sup>

$$M_p = \frac{1}{n} \sum_{i=1}^n P_i \quad (4)$$

$$M_r = \frac{1}{n} \sum_{i=1}^n R_i \quad (5)$$

$$M_{F_1} = \frac{2M_p M_r}{M_p + M_r} \quad (6)$$

根据以上公式,计算得到3种模型的宏精确率、宏召回率和宏  $F_1$  值如表5所示。

表5 不同模型的宏精确率、宏召回率和宏  $F_1$  值

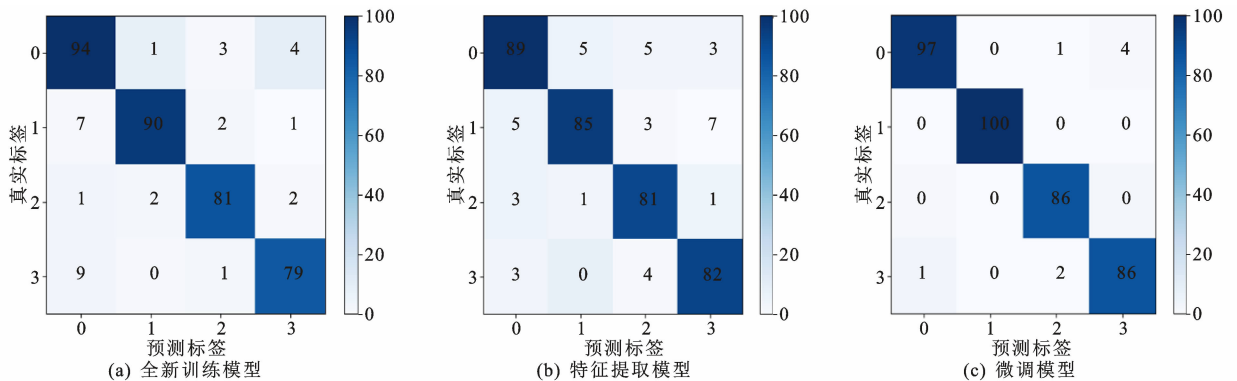
Tab.5 Macro accuracy, macro recall rate and macro  $F_1$

values of different models

%

模型类型	全新训练模型	特征提取模型	微调模型
宏精确率	91.60	89.42	97.79
宏召回率	91.28	89.64	97.93
宏 $F_1$ 值	91.44	89.53	97.86

宏平均值平等对待每一个类别,它的值主要受到稀有类别的影响,能较客观评价模型的识别效果。



单位:幅

图8 三种模型的混淆矩阵

Fig.8 Confusion matrix of three models

### 3.4 特征图可视化与识别结果可视化

#### 3.4.1 特征图可视化

神经网络对特征的学习是一个抽象的过程,为更直观观察这个抽象提取过程,以便了解模型对钢桥病害样本特征的学习情况,可将卷积过程中的特征映射到像素空间进行可视化。选取钢桥腐蚀病害样本,对其前4层卷积层的特征图进行可视化,依据可视化结果,进而观察不同卷积层的学习特点,可视化特征图如图9所示。

由图9可知,随着卷积层的加深,滤波器提取到的特征由低层次逐渐转变为高层次特征,CNN特征提取具有层次化的特点。卷积过程可以看作计算被卷积对象和卷积核的相关性,卷积核和输入的对像中对应部分越相似,则卷积得到的结果就会越大,前

由表5可知,3种模型的宏精确率、宏召回率和宏  $F_1$  值均较稳定,其中微调模型达到了97%以上,高出另外2个模型6%以上,说明全新训练模型和特征提取模型受光照强度、图像位置变换等影响偏大,微调模型则具有更好的几何(镜像、旋转、缩放)不变性和光度(亮度、饱和度)不变性,能够克服外界复杂环境的影响,鲁棒性更强。各模型的混淆矩阵如图8所示。可见不同模型下类别之间相互误分的情况如下:全新训练模型对腐蚀和完好类别的误测数量最多,为10幅,误测数量最少是焊缝开裂类,为5幅;而特征提取模型对完好类别的识别效果较全新训练模型稍好,对涂层劣化和腐蚀类识别能力偏弱,分别误测13幅和15幅;微调模型共误测8幅图像,其中涂层劣化误测5幅,完好误测3幅,腐蚀和焊缝开裂类无错误预测,对4种类别的识别能力均比前2个模型强。可见,无论是宏平均值还是对单个类别的识别能力,微调模型都优于全新训练模型和特征提取模型。

2层卷积后得到的特征图,较好提取到了腐蚀病害的边缘轮廓,且在图像中病害位置的颜色较为突出,说明滤波器对腐蚀病害进行了有效提取,并主要学习了图像中纹理特征和颜色特征等底层特征。第3、4层卷积特征图随着卷积核数量的增加,得到的特征图也在增多,特征图轮廓信息不再明显,逐渐变得抽象,可见随着网络不断加深,滤波器也变得更为复杂,提取图像特征更加深入,卷积层对图像提取逐渐向高层次发展。同时,也说明了底层特征需要少量的迭代训练就能收敛,而高层特征则需要更多次的迭代才能收敛。

#### 3.4.2 识别结果可视化

通过第3.3节的对比分析可知,微调模型的识别性能最佳,故本文试验最终将微调模型作为钢桥

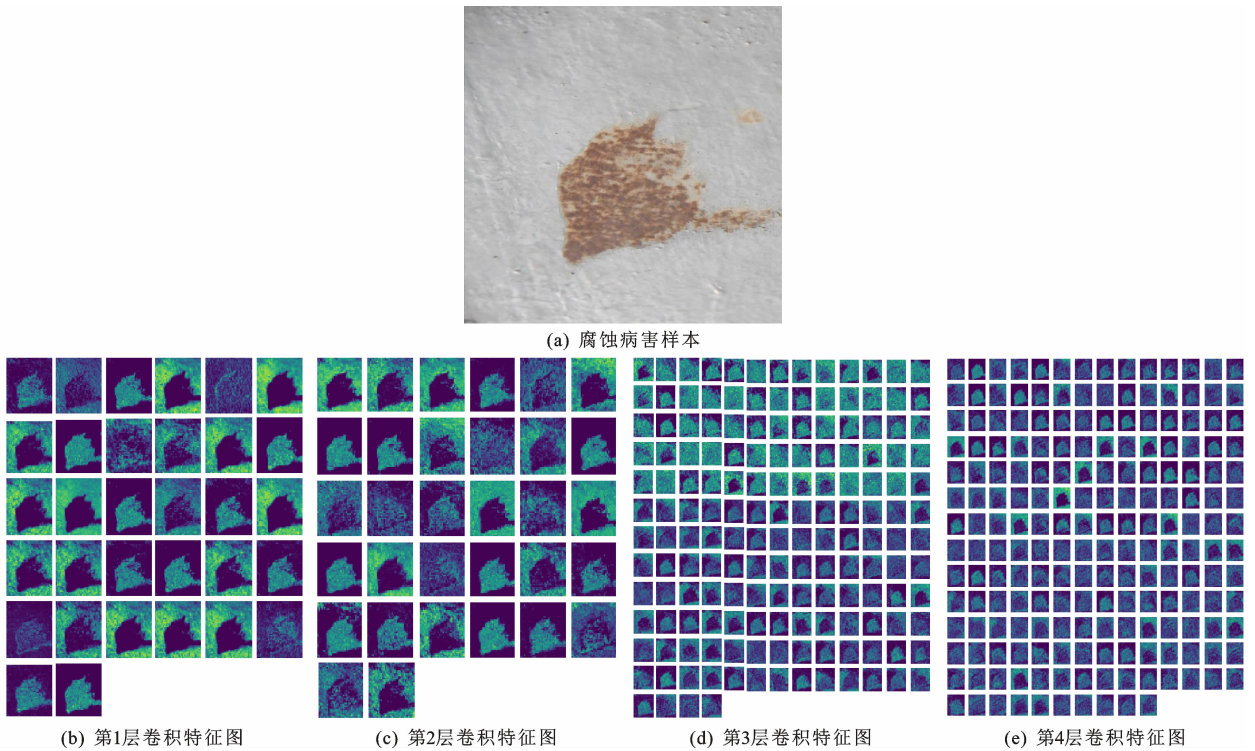


图 9 腐蚀病害样本卷积特征图

Fig. 9 Convolution characteristics of corrosion defect samples

病害识别的模型。为更清楚地知道每幅病害图像的识别情况,在钢桥病害识别模型中,加入了展示识别结果的程序。将待识别的病害图像输入模型中,模型在遍历图像后,可以把每幅图像的识别结果进行可视化,识别结果以窗口的形式展现,每个窗口可同时显示 4 幅图像,具体效果如图 10 所示。为更直观明确预测错误的图像,当预测结果与真实标签不同时,纵横轴标签颜色输出变为红色,如图 10 展示的示例窗口,前 3 幅图像的预测结果均与真实标签相同,标签颜色统一为黑色,而第 4 幅的预测标签与真实标签不一致,预测错误,标签颜色为红色。进一步观察该图像,其涂层劣化的形状与焊缝形状相似,对于模型识别有一定误导,模型将其误识为焊缝开裂。经过对识别结果的可视化,在更加直观看到模型识别效果的同时,也可以进一步分析模型出现误测的原因,增强了钢桥病害检测模型的实际应用性。

## 4 结 语

(1)当学习率过大时,会导致全新训练模型中损失函数值振荡或出现梯度爆炸,迁移学习模型中的参数更新过快,破坏了原本训练好的权重信息;学习率过小时,模型偏离最优点甚至出现过拟合现象,在学习率为 0.001 时,模型取得了最佳识别效果。

(2)特征提取模型单个 epoch 时间为 47.2 s,在训练 100 次左右达到稳定;微调模型单个 epoch 时间为 119.2 s,在训练 25 次左右达到稳定。与全新训练模型相较,迁移学习极大缩短了模型训练时间,提升了运行效率,在识别性能方面,微调模型各类评价指标均优于全新训练模型,具有较好的鲁棒性。

(3)预训练 Inception-v4 模型的源域与钢桥病害图像存在一定的差异,迁移学习后采用特征提取和微调 2 种训练方式进行训练,分别得到 89.39% 和 97.88% 的准确率,且微调模型借助于对中高级特征的数据扩充和调整,减少了过拟合,识别准确率较特征提取模型高,选取微调模型作为钢桥病害识别模型,实现了快速、准确识别。

(4)目前,本文模型仅实现了对病害的识别分类,尚无法确定每幅图像中病害的具体位置,下一步可研究在对钢桥病害分类的同时,实现病害的定位,以便于采取相应的维护措施,更好地应用于实际工程中。

## 参考文献:

### References:

- [1] CJJ/T 233—2015, 城市桥梁检测与评定技术规范[S].  
CJJ/T 233—2015, Technical code for test and evaluation of city bridges[S].

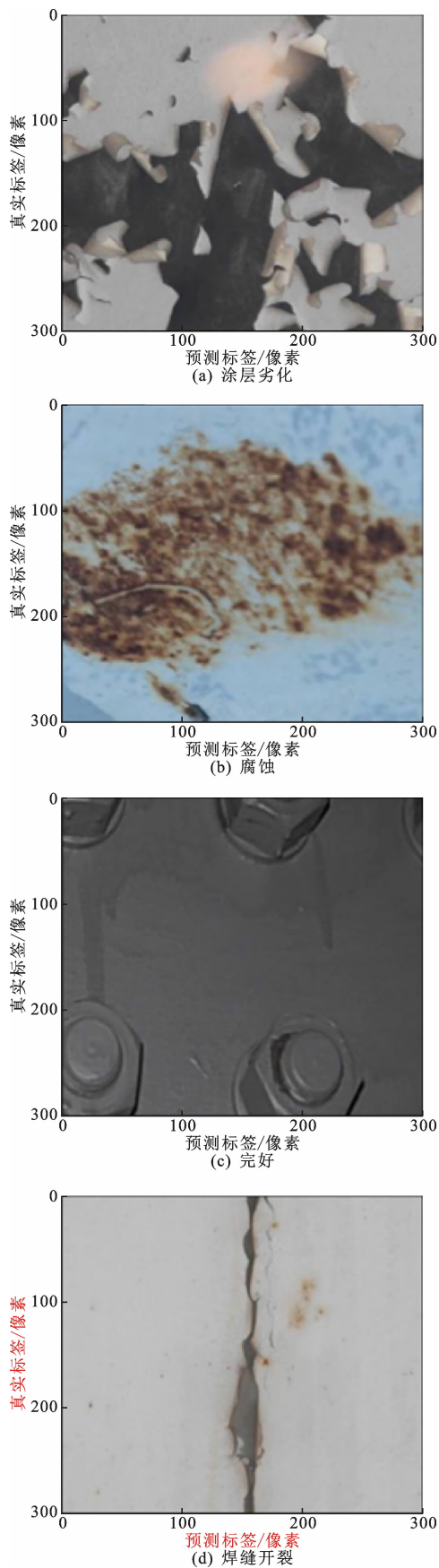


图10 可视化识别结果

Fig. 10 Visual recognition results

- [2] JTG/T H21—2011,公路桥梁技术状况评定标准[S]. JTG/T H21—2011, Standard for technical condition evaluation of highway bridges[S].
- [3] 马 晔. 混凝土结构缺陷的红外热成像检测识别技术[J]. 公路交通科技, 2017, 34(12): 59-65. MA Ye. Infrared thermal imaging detection and recognition technology for defects in concrete structure [J]. Journal of Highway and Transportation Research and Development, 2017, 34(12): 59-65.
- [4] 沙爱民, 童 峥, 高 杰. 基于卷积神经网络的路表病害识别与测量[J]. 中国公路学报, 2018, 31(1): 1-10. SHA Ai-min, TONG Zheng, GAO Jie. Recognition and measurement of pavement disasters based on convolutional neural networks [J]. China Journal of Highway and Transport, 2018, 31(1): 1-10.
- [5] 沙爱民, 蔡若楠, 高 杰, 等. 基于级联卷积神经网络的公路路基病害识别[J]. 长安大学学报(自然科学版), 2019, 39(2): 1-9. SHA Ai-min, CAI Ruo-nan, GAO Jie, et al. Subgrade distresses recognition based on convolutional neural network[J]. Journal of Chang'an University (Natural Science Edition), 2019, 39(2): 1-9.
- [6] 冯 微, 张兆津, 邵海鹏. 一种基于深度置信网络的车撞桥墩损伤等级判别方法[J]. 长安大学学报(自然科学版), 2018, 38(6): 135-145. FENG Wei, ZHANG Zhao-jin, SHAO Hai-peng. A method for discriminating damage level in vehicle-bridge pier collision based on deep belief network [J]. Journal of Chang'an University (Natural Science Edition), 2018, 38(6): 135-145.
- [7] 薛亚东, 李宜城. 基于深度学习的盾构隧道衬砌病害识别方法 [J]. 湖南大学学报(自然科学版), 2018, 45(3): 100-109. XUE Ya-dong, LI Yi-cheng. A method of disease recognition for shield tunnel lining based on deep learning[J]. Journal of Hunan University (Natural Sciences), 2018, 45(3): 100-109.
- [8] 刘新根, 陈莹莹, 朱爱玺, 等. 基于深度学习的隧道裂缝识别方法 [J]. 广西大学学报(自然科学版), 2018, 43(6): 2243-2251. LIU Xin-gen, CHEN Ying-ying, ZHU Ai-xi, et al. Tunnel crack identification based on deep learning [J]. Journal of Guangxi University (Natural Science Edition), 2018, 43(6): 2243-2251.
- [9] 孙朝云, 马志丹, 李 伟, 等. 基于深度卷积神经网络融合模型的路面裂缝识别方法[J]. 长安大学学报(自然科学版), 2020, 40(4): 1-13. SUN Zhao-yun, MA Zhi-dan, LI Wei, et al. Pavement crack identification method based on deep convolu-

- tional neural network fusion model[J]. Journal of Chang'an University (Natural Science Edition), 2020,40(4):1-13.
- [10] 李海丰,韩红阳.复杂背景下机场道面细带状结构病害检测算法[J/OL].北京航空航天大学学报,2021,5:1-9[2021-05-18].DOI:10.13700/j.bh.1001-5965.2020.0512.  
LI Hai-feng,HAN Hong-yang. Algorithm for the detection of thin strip-shaped structural diseases on airport pavement under complex background[J/OL]. Journal of Beijing University of Aeronautics and Astronautics,2021,5:1-9[2021-05-18].DOI:10.13700/j.bh.1001-5965.2020.0512.
- [11] YUN J P,SHIN W C,KOO G, et al. Automated defect inspection system for metal surfaces based on deep learning and data augmentation[J]. Journal of Manufacturing Systems,2020,55:317-324.
- [12] CHENG J C P,WANG M Z. Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques[J]. Automation in Construction,2018,95:155-171.
- [13] BASTIAN B T,JASPREETH N,RANJITH S K, et al. Visual inspection and characterization of external corrosion in pipelines using deep neural network[J]. NDT & E International,2019,107:102134.
- [14] ALI R,CHA Y J. Subsurface damage detection of a steel bridge using deep learning and uncooled microbolometer[J]. Construction and Building Materials, 2019,226:376-387.
- [15] 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.
- [16] ZHU J S,SONG J B. An intelligent classification model for surface defects on cement concrete bridges[J]. Applied Sciences,2020,10(3):972.
- [17] 沈浩,江臣,陈宇文,等.基于深度学习的钢桁架桥螺栓病害智能识别方法[J].南京工业大学学报(自然科学版),2020,42(5):608-615.  
SHEN Hao,JIANG Chen,CHEN Yu-wen, et al. Intelligent identification method for bolt diseases of steel truss bridge based on deep learning model[J]. Journal of Nanjing Tech University (Natural Science Edition),2020,42(5):608-615.
- [18] 李雪松,林逸洲,马宏伟,等.基于卷积神经网络的桥梁损伤识别方法应用研究[J].青海大学学报,2018,36(2):41-46.  
LI Xue-song,LIN Yi-zhou,MA Hong-wei, et al. Application of bridge damage identification based on convolutional neural network[J]. Journal of Qinghai University,2018,36(2):41-46.
- [19] XU Y,BAO Y Q,CHEN J H, et al. Surface fatigue crack identification in steel box girder of bridges by a deep fusion convolutional neural network based on consumer-grade camera images[J]. Structural Health Monitoring,2019,18(3):653-674.
- [20] 阮小丽,王波,吴巨峰,等.基于深度学习的钢筋混凝土桥梁掉块露筋病害识别[J].世界桥梁,2020,48(6):88-92.  
RUAN Xiao-li,WANG Bo,WU Ju-feng, et al. Identification of spalled concrete and exposed reinforcement in reinforced concrete bridge based on deep learning[J]. World Bridges,2020,48(6):88-92.
- [21] SZEGEDY C,IOFFE S,VAN HOUCKE V, et al. Inception-v4,inception-ResNet and the impact of residual connections on learning[EB/OL].(2016-08-23)[2021-05-18].<https://arxiv.org/abs/1602.07261>.
- [22] PAN S J,YANG Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering,2010,22(10):1345-1359.
- [23] XIONG P,ZHU Y,SUN Z R, et al. Application of transfer learning in continuous time series for anomaly detection in commercial aircraft flight data[C]//IEEE. Proceedings of 2018 IEEE International Conference on Smart Cloud (SmartCloud). New York: IEEE,2018:13-18.
- [24] DENG J,DONG W,SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]//IEEE. Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2009:248-255.
- [25] 王卫东. 钢箱梁桥常见病害及其检测[J].城市道桥与防洪,2014(12):131-135,15.  
WANG Wei-dong. Common faults and detection of steel box girder bridge[J]. Urban Roads Bridges & Flood Control,2014(12):131-135,15.
- [26] SONG Y,MCLOUGHLIN I,DAI L R. Deep bottleneck feature for image classification[C]//ACM. Proceedings of the 5th ACM on International Conference on Multimedia Retrieval. New York: ACM,2015:491-494.
- [27] LEI Y W,HU T,LI G Y, et al. Stochastic gradient descent for nonconvex learning without bounded gradient assumptions[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(10): 4394-4400.
- [28] ZHU J S,ZHANG C,QI H D, et al. Vision-based defects detection for bridges using transfer learning and convolutional neural networks[J]. Structure and Infrastructure Engineering,2020,16(7):1037-1049.