

文章编号：1000-4750(2024)00-0001-13

# 基于卷积神经网络的桥梁裂缝两阶段识别算法研究

冯昊龙<sup>1</sup>, 刘宇飞<sup>1</sup>, 刘晓刚<sup>2</sup>, 刘 诚<sup>3</sup>

(1. 清华大学土木工程系, 北京 100084; 2. 北京科技大学土木与资源工程学院土木工程系, 北京 100083;  
3. 中路高科交通检测检验认证有限公司, 北京 100088)

**摘要：**为解决桥梁裂缝智能识别场景下桥梁表面图像条件复杂、对智能识别算法的识别效果产生严重干扰的问题, 提出了“分类筛选-裂缝识别”的两阶段桥梁裂缝识别算法。根据桥检实拍图像的图像特征调整 MobilenetV3 的网络结构, 构建桥梁表面图像的筛选分类模型, 实现对背景图像、干扰图像及低质量图像的自动筛选; 通过迁移学习方法训练框-网格融合式裂缝识别模型, 实现裂缝识别与定位, 为进一步的裂缝精细分割、裂缝测量提供条件。算法开发过程中构建了包含 5000 张图像的图像分类数据集、包含 6400 张图像的裂缝识别数据集; 最终得到的图像分类最高准确率为 97.9%, 裂缝网格识别最高准确率为 65.6%。结果表明: 基于本文提出的桥梁裂缝智能检测方法可实现桥梁裂缝的高效率、自动化检测, 算法具有较强的抗干扰能力。

**关键词:** 桥梁与隧道工程; 裂缝识别算法; 算法研究; 深度学习; 两阶段识别算法; 卷积神经网络

中图分类号: TP391.4; U446 文献标志码: A doi: [10.6052/j.issn.1000-4750.2024.03.0238](https://doi.org/10.6052/j.issn.1000-4750.2024.03.0238)

## RESEARCH ON TWO-STAGE IDENTIFICATION ALGORITHM OF BRIDGE CRACKS BASED ON CONVOLUTIONAL NEURAL NETWORK

FENG Hao-long<sup>1</sup>, LIU Yu-fei<sup>1</sup>, LIU Xiao-gang<sup>2</sup>, LIU Cheng<sup>3</sup>

(1. Department of Civil Engineering, Tsinghua University, Beijing 100084, China;  
2. Department of Civil Engineering, School of Civil and Resource Engineering, University of Science and Technology Beijing, Beijing 100083, China;  
3. China-Road Transportation Verification & Inspection Hi-Tech Co., Ltd., Beijing 100088, China)

**Abstract:** The complex surface image conditions of bridges seriously reduce the crack recognition accuracy of the intelligent identification algorithm. To address this issue, a two-stage bridge crack identification algorithm is proposed. The network structure of MobilenetV3 is adjusted according to the characteristics of bridge surface images, and a classification model for bridge surface images is constructed to achieve automatic filtering of background images, interfering images and low-quality images. A box-grid fusion crack identification model is trained to achieve crack locating, which is used for further crack segmentation and crack measurement. The final image classification accuracy is 97.9%, and the final crack recognition accuracy for grid prediction is 65.6%. The results indicate that the intelligent detection method for bridge cracks proposed in this article can achieve efficient and automated detection of bridge cracks, and the algorithm has strong anti-interference ability.

**Key words:** bridge and tunnel engineering; crack detection algorithm; algorithm research; deep learning; two-stage detection algorithm; convolutional neural network

收稿日期: 2024-03-29; 修改日期: 2024-11-14

基金项目: 国家自然科学基金项目(52192662)

通信作者: 刘宇飞(1988—), 男, 陕西人, 副研究员, 博士, 主要从事工程结构智能诊断与温度效应研究(E-mail: [liuyufei@tsinghua.edu.cn](mailto:liuyufei@tsinghua.edu.cn)).

作者简介: 冯昊龙(2000—), 男, 辽宁人, 硕士生, 主要从事路面与桥梁结构病害识别研究(E-mail: [fhl22@mails.tsinghua.edu.cn](mailto:fhl22@mails.tsinghua.edu.cn));

刘晓刚(1986—), 男, 山东人, 教授, 主要从事纤维增强复合材料及结构、高性能钢结构、装配式钢结构与组合结构、工程结构评估与加固研究(E-mail: [liuxiaogang@ustb.edu.cn](mailto:liuxiaogang@ustb.edu.cn));

刘 诚(1991—), 男, 安徽人, 博士, 主要从事路面与桥梁结构病害识别研究(E-mail: [liucheng1991@gmail.com](mailto:liucheng1991@gmail.com)).

随着我国公路、铁路桥梁建设不断推进、各类桥梁服役年限不断增加，桥梁检测与维护的相关需求也在不断增长。桥梁裂缝是桥梁结构最常见的、对结构安全影响最显著的病害形式之一<sup>[1-3]</sup>，裂缝检测是桥梁病害检测中的一项重要工作内容。桥梁检测行业目前常用的手段为人工检测，其危险性高、检测效率低、检测周期较长且受检测人员主观因素影响较为严重，迫切需要一种自动化的、安全高效的检测方法。

近年来，无人机、爬壁机器人等桥梁表面图像采集设备技术迅速发展，其灵活便捷的操作特性使得其在桥梁检测工作中的应用前景得到研究者的广泛关注。与人工检查手段相比，这些桥梁表面图像采集设备能够覆盖的桥梁表面位置更加全面，检测过程也更加安全高效。将采集得到的桥梁表面影像数据作为病害识别的主要依据，结合数字图像方法或深度学习卷积神经网络算法进行裂缝等桥梁表面病害的识别，可大幅提升桥梁检测工作的效率与安全性，显著降低桥检工作所需的人力成本。

在桥梁裂缝检测算法领域，国内外学者开展了广泛研究，取得了丰富的研究成果。基于计算机视觉的桥梁裂缝识别算法现阶段可分为数字图像处理法和深度学习方法两类。

在工作平台开发与技术方案可行性论证方面，钟新谷等<sup>[4]</sup>以旋翼无人机为工作平台，采用非接触式测量仪验证了无人机悬停状态下机载成像质量的可靠性。

数字图像处理方法主要通过对数字图像应用阈值、滤波等计算机处理实现对裂缝区域的提取，并进一步完成裂缝测量等操作。阮小丽等<sup>[5]</sup>通过裂缝连接、虚假裂缝过滤、裂缝交叉点匹配、通裂缝区域划分等方法提高了裂缝识别与裂缝宽度测量的准确度；谢文高等<sup>[6]</sup>在清水环境实验中对混凝土结构进行表观检测，通过图像处理、高斯滤波以及梯度计算的分权阈值法对存在裂缝的图像进行分析，提取裂缝边缘特征，实现对裂缝宽度的分析计算；常广利等<sup>[7]</sup>通过图像灰度化、直方图均衡化、滤波器去噪、最大类间方差法等手段识别桥梁裂缝，误差率显著降低；杨子涵等<sup>[8]</sup>使用 DIC 计算得出的水平、竖直、剪切应变场计算试件表面的主拉伸应变场，根据应变场提取出开裂位置与裂缝骨架线，并计算裂缝的

发展规律；PRATEEK PRASANNA 等<sup>[9]</sup>提出了基于 STRUM(空间调谐鲁棒多特征) 分类器的裂缝识别算法，并使用机器学习对图像进行分类，代替了手动调整阈值参数的过程。也有学者探究了使用三维数字图像完成裂缝识别的可行性。杨勇等<sup>[10]</sup>将深度图像和彩色图像相结合，通过将二维彩色图像点向三维空间坐标转换，实现了高速铁路桥梁裂缝宽度的检测识别；周浩翔等<sup>[11]</sup>提出了基于水下机器人拍摄平面图像的水下混凝土结构三维重建方法，通过水下折射摄像机成像模型、SIFT 特征提取和匹配算法以及点云重建的方式实现了水下带病害的混凝土结构表面的三维重建。数字图像处理方法虽然算法原理清晰、在特定情况下能够表现出较好的识别效果，但算法的效果部分取决于人工设置的阈值参数，泛化能力较差、抗干扰能力普遍不足，识别效果受图像条件影响较为显著。

随着近几年深度学习技术的快速发展，基于深度学习方法的桥梁裂缝识别研究得到了学界的广泛关注，不断有学者尝试将各类深度学习模型应用于桥梁裂缝识别场景。刘洪公等<sup>[12]</sup>改进了裂缝分类的卷积神经网络模型，最终提出一种智能裂缝检测方案；李良福等<sup>[13]</sup>提出一种基于卷积神经网络的分类模型，用于桥梁背景面元和桥梁裂缝面元的识别；XU 等<sup>[14]</sup>提出了一种基于 atrous 卷积、ASPP 模块和深度可分离卷积的端到端裂纹检测模型；梁雪慧等<sup>[15]</sup>提出了一种基于改进 GoogleNet 的桥梁裂缝检测算法，通过图块裁切与子图块分类定位裂缝位置，完成裂缝宽度的计算；李雯雯等<sup>[16]</sup>提出了一种基于全卷积神经网络和朴素贝叶斯数据融合模型的桥梁裂缝识别算法用于裂缝分割；余加勇等<sup>[17]</sup>使用 Mask R-CNN 深度学习算法训练和建立裂缝识别模型，并构建了图像后处理方法以获取裂缝形态及宽度信息；刘森森<sup>[18]</sup>基于迁移学习的思想，建立了基于 VGG16 的迁移学习裂缝识别网络。此外，也有学者尝试将 YOLO、ResNet、RS-Unet 等近年新出现的各类深度学习模型应用于桥梁裂缝识别场景，实现了较高的识别准确度<sup>[19-23]</sup>。

综合领域内研究现状与应用场景中的实际需求，现有相关研究主要存在以下不足：

1) 算法的研究范围局限于规整、无干扰物的混凝土表面图像，算法的抗干扰能力不足。与其

他结构类型相比，桥梁结构的结构构件直接外露的面积占比较大，也因此更适合进行基于视觉方法的病害检测；然而，即使是对于桥梁结构，在实际检测工作中混凝土结构构件的表面也会附有大量排水管道、线缆、防护网等附属设施，对裂缝识别算法产生严重干扰。此外，混凝土构件表面本身也会带有浇筑接缝，极易与结构裂缝之间产生混淆；为实现无人机拍摄对结构表面的完全覆盖，一些图像会部分拍摄到环境背景，这些干扰因素都会显著降低裂缝自动识别的准确性，需要妥善处理。

2) 在实际的桥梁检测工作中，桥梁表面严重污染、仅通过图像信息无法准确识别裂缝的情况常有出现，易导致自动识别算法的大面积误判。桥梁结构表面大部分与外界环境直接接触，在长时间的服役过程中可能会产生大面积的水渍、油漆等污染，污染痕迹与裂缝之间彼此混杂，使得针对此类图像的裂缝识别较为困难，现阶段只能将图像交由人工判断或使用非视觉手段进行裂缝判别。

为解决以上问题，本文提出了一种“分类筛选+裂缝识别”的两阶段桥梁裂缝识别算法。首先，训练一个图像分类模型，实现对背景图像、干扰图像及低质量图像的自动筛选；将筛选后确认适合进行智能裂缝识别的图像输入训练好的裂缝识别模型，实现裂缝定位。通过将两阶段的深度学习算法相结合，保证桥梁裂缝智能识别算法的可靠性，确保智能识别算法能够代替部分人工工作，提高生产效率。

## 1 研究框架及算法

### 1.1 研究框架

本文研究框架与技术路线如图1所示。

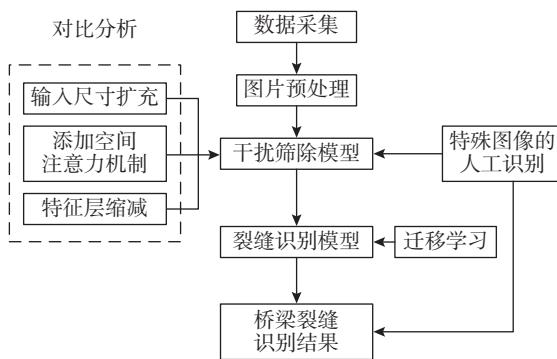


图1 技术路线图  
Fig. 1 Technical roadmap

本文主要完成了以下研究工作：1) 开发基于图像分类模型的干扰图像筛选算法，通过对比试验分析输入尺寸扩充、添加空间注意力机制等模型结构改造方法对分类准确度的影响；2) 通过迁移学习方法开发基于框/网格双重格式目标检测模型的裂缝识别算法；3) 通过将两个阶段的深度学习算法与人工处理相结合，提出桥梁表面裂缝识别的工作流，实现裂缝智能识别算法对人工的部分替代，并保证裂缝识别结果的可靠性。无人机采集得到的照片进行预处理后输入图像分类模型，将背景图像、干扰图像等转交人工处理；将筛查后确认适合应用裂缝自动识别算法的图像输入目标检测模型，完成该部分图像中裂缝的自动识别；人工识别结果与自动识别结果合并输出，完成基于桥梁表面图像的裂缝检测工作。

### 1.2 图像分类深度学习算法

#### 1.2.1 MobileNetV3 基本结构

与其他参数量较大的重量级神经网络相比，轻量级神经网络的模型参数量更小，处理效率也更高，更加适用于需要进行快速图像处理的应用场景。MobileNetV3 网络是当前图像分类领域应用较为广泛的轻量级深度学习网络之一，其在实际应用中表现出了较强的准确度与速度优势<sup>[24]</sup>。在图像分类算法的开发过程中，本文以该网络为基础，首先根据桥梁表面裂缝图像的特征对网络结构进行调整，随后尝试向网络中添加空间注意力机制，使得改造后的深度学习模型在桥梁表面图像分类场景下具有更高的准确度与适用性。

MobileNetV3 网络主要由主干网络与检测头两部分构成，如图2 所示。

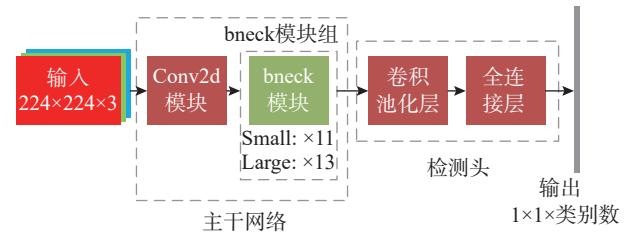


图2 MobileNetV3 网络结构图  
Fig. 2 Network structure of mobilenetv3 model

主干网络为模型的特征提取部分，其主要由一系列 bneck 模块顺次链接构成。bneck 模块主要实现了通道可分离卷积、通道注意力机制、残差连接三类网络改进机制，具有较强的特征提取能力。检测头为模型的特征解析部分，通过综合分析各类图像特征完成图像分类。

### 1.2.2 输入尺寸调整

MobileNetV3 模型的输入尺寸为  $224 \times 224$  像素, 尺寸超出限制的图像需要在前处理过程中降采样至该尺寸后才能输入模型进行处理。若图像尺寸过大, 前处理降采样倍数过高, 会导致大量图像细节的丢失, 严重影响算法整体的分类准确度。与图像分类模型相比, 裂缝识别模型的输入尺寸相对更大(图像尺寸一般在 640 像素左右);而在本研究中, 需要将相同尺寸的图像先后输入分类模型与识别模型进行处理, 因此需要将分类模型的图像输入尺寸扩充至与目标识别模型相近, 保证二者协同工作。综合考虑裂缝识别模型的输入尺寸需求与无人机实采图像的尺寸信息, 计划将分类模型输入尺寸扩充为原来的 4 倍, 即  $896 \times 896$  像素。

为修改网络结构、调整输入尺寸, 向主干网络中添加了两个额外的 bNeck 模块, 并对部分网络结构参数进行了调整。调整前后的网络结构如表 1、表 2 所示。

表 1 MobileNetV3 模型结构参数

Table 1 Network structure parameters of mobilenetv3 model

编号	网络模块	输出特征层数	卷积核尺寸	步长
1	conv2d	16	$3 \times 3$	2
2	bneck	16	$3 \times 3$	2
3	bneck	24	$3 \times 3$	2
4	bneck	24	$3 \times 3$	1
5	bneck	40	$5 \times 5$	2
6	bneck	40	$5 \times 5$	1
7	bneck	40	$5 \times 5$	1
8	bneck	48	$5 \times 5$	1
9	bneck	48	$5 \times 5$	1
10	bneck	96	$5 \times 5$	2
11	bneck	96	$5 \times 5$	1
12	bneck	96	$5 \times 5$	1
13	conv2d	576	$1 \times 1$	1
14	pool	576	$7 \times 7$	1
15	FN	1024	—	—
16	FN	$k$	—	—

### 1.2.3 特征层缩减

与其他更加泛化的图像分类应用场景不同, 桥梁表面图像的场景比较单一, 可供提取的图像特征种类也相对较少; 需要识别的类别数量也大幅减少, 不需要提取种类过多的图像特征用于特征分析。因此, 对于桥梁表面图像分类这一应用场景, MobileNetV3 网络在特征层数量上存在明显

表 2 扩充输入尺寸后模型结构参数

Table 2 Network structure parameters after input size expanding

编号	网络模块	输出特征层数	卷积核尺寸	步长
1	conv2d	16	$3 \times 3$	2
2	bneck	16	$3 \times 3$	2
3	bneck	24	$3 \times 3$	2
4	bneck	24	$3 \times 3$	1
5	bneck	40	$5 \times 5$	2
6	bneck	40	$5 \times 5$	1
7	bneck	40	$5 \times 5$	1
8	bneck	48	$5 \times 5$	1
9	bneck	48	$5 \times 5$	1
10	bneck	96	$5 \times 5$	2
11	bneck	96	$5 \times 5$	1
12	bneck	96	$5 \times 5$	2
13	bneck	96	$5 \times 5$	1
14	bneck	96	$5 \times 5$	1
15	conv2d	576	$1 \times 1$	1
16	pool	576	$7 \times 7$	1
17	FN	1024	—	—
18	FN	$k$	—	—

冗余, 可以适当缩减特征层的数量以降低模型的参数量, 提高模型训练与推测过程的效率。在 1.2.2 节输入尺寸扩充后网络的基础上修改了部分模块中的输入、输出特征层数量, 进一步修改后的网络结构如表 3 所示。

表 3 特征层缩减后的模型结构参数

Table 3 Network structure parameters after reducing feature layers

编号	网络模块	输出特征层数	卷积核尺寸	步长
1	conv2d	16	$3 \times 3$	2
2	bneck	16	$3 \times 3$	2
3	bneck	24	$3 \times 3$	2
4	bneck	24	$3 \times 3$	1
5	bneck	40	$5 \times 5$	2
6	bneck	40	$5 \times 5$	1
7	bneck	40	$5 \times 5$	1
8	bneck	48	$5 \times 5$	2
9	bneck	48	$5 \times 5$	1
10	bneck	64	$5 \times 5$	2
11	bneck	64	$5 \times 5$	1
12	bneck	96	$5 \times 5$	2
13	bneck	96	$5 \times 5$	1
14	bneck	96	$5 \times 5$	1
15	conv2d	120	$1 \times 1$	1
16	pool	120	$7 \times 7$	1
17	FN	256	—	—
18	FN	$k$	—	—

#### 1.2.4 注意力机制

注意力机制是近几年深度学习领域应用较为广泛、应用效果较好的模型改进手段之一。通过模仿人类视觉在数据处理过程中将注意力集中于某些特定位置的过程，注意力机制可以引导模型重点关注对预测准确性影响更显著的位置或图像特征，从而提升模型的性能。注意力机制主要分为通道注意力机制与空间注意力机制两类。通道注意力机制通过对每个通道的重要性施加权重，在训练过程中调整这一权重的方式促使模型专注于更加重要的图像特征，而空间注意力机制则通过对同一特征图中不同位置施加权重、对这一权重进行训练的方式促使模型专注于更加重要的图像位置<sup>[26]</sup>。

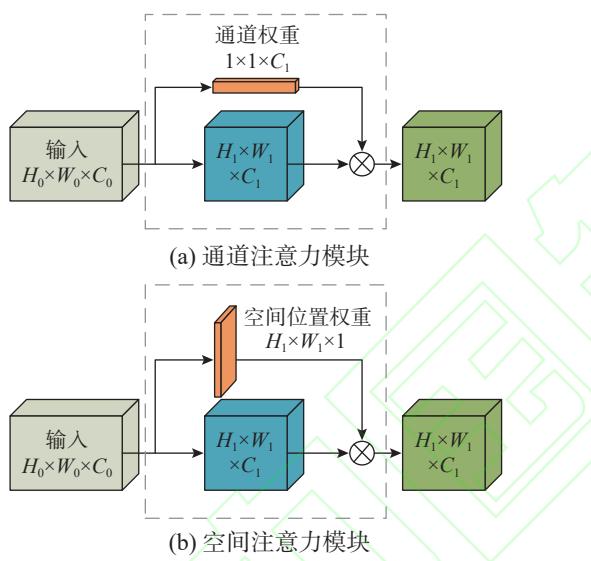


图 3 通道注意力模块与空间注意力模块

Fig. 3 Channel attention module and spatial attention module

特征层注意力机制已经应用于 MobileNetV3 中的 bneck 模块，但空间注意力机制在该模型中尚未得到应用。在桥梁检测实际工作场景下，图像采集设备的位置、拍摄时的取景范围以及各类图像特征的位置分布情况会表现出显著的不规则性，与算法测试场景或其他应用场景下拍摄得到的图像数据集之间有较大差距。在一般的图像分类数据集中，目标对象一般位于图像中心或占据图像中较大的面积比例；而在本研究的目标场景下，需要分类筛选的虚化背景、附属设施、表面污染等的分布情况相对而言更为随机，可能位于图像边缘处或只占据图像中的一小部分面积，因此更加需要引入空间注意力机制，使模型注重于分析更有价值的局部图像。

为引入空间注意力机制，对 MobileNetV3 图像分类模型中的 bneck 模块进行了代码重构。在该模块中添加了一个空间权重层，该权重层高度、宽度与模块输出尺寸相同，深度为 1。在深度学习模型的前向传播过程中，将该权重层和模块的原输出按位相乘，作为该模块的新输出，在反向传播过程中同样对该权重层的内部参数进行迭代优化。

#### 1.2.5 图像分类模型的损失函数

在进行神经网络训练时，需要定义损失函数用来量化模型预测与真实结果之间的差距，从而为神经网络提供正确的反馈。在输入图像进行神经网络的正向传播时，网络的最终输出为一个一维向量，其长度与分类的类别数相等，每一位上的值代表模型判断下对应类别的置信度。在计算图像分类模型的损失函数时，本研究直接使用了图像分类任务中常用的交叉熵损失函数，其计算式为

$$L = 1/N \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (1)$$

式中：N 为样本总量；M 为类别总数； $y_{ic}$  为符号函数，样本  $i$  真实类别等于  $c$  时取 1，否则取 0； $p_{ic}$  为模型预测第  $i$  个样本属于第  $c$  类的概率。

### 1.3 裂缝识别深度学习算法

#### 1.3.1 融合式裂缝识别模型

在裂缝识别算法方面，本研究使用了 LI 等<sup>[25]</sup>的研究成果：基于框/网格双重格式的目标检测深度学习网络。该模型开发了框与网格两种模型输出格式；其中，框 (box) 格式通过给出裂缝等对象的边界框判定识别对象的位置与尺寸，网格 (grid) 格式通过将图像划分为网格、判定每个网格所属的对象类别的方式完成对各类识别对象的定位。

#### 1.3.2 迁移学习方法

在训练用数据集难以获取、数据规模较小时，深度学习算法的训练效果一般较差，且由于样本较少极易出现过拟合现象；针对这一问题，可使用迁移学习方法，在数据集规模较小的情况下尽可能提升模型性能。

在一般的深度学习模型训练过程中，首先随机初始化权重，此时模型预测也是随机的；在反向传播过程中，权重数值得到调整，使得模型预测与人类预测之间更为接近；因此可以认为，模型内部权重即代表了模型在训练过程中对训练数据集的“理解”。基于这一原理，研究者提出了迁移学习方法，通过内部权重在模型之间的传递

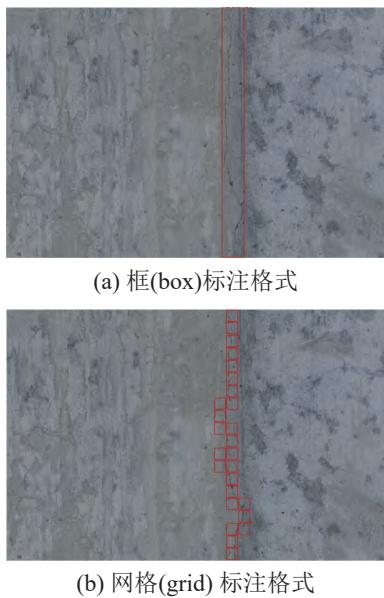
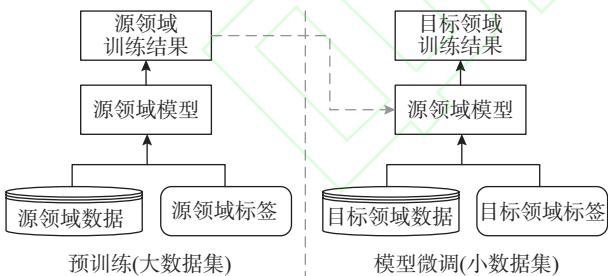


图 4 融合识别模型的两种标注数据格式

Fig. 4 Two label formats for fusion recognition model

实现模型训练过程中对先前已有数据的充分利用。

迁移学习方法主要分为预训练与参数微调两部分。在预训练过程中，使用的数据集规模应较大且应与参数微调中使用的图像较为相似；通过设置训练任务完成预训练，得到与预训练数据集图像特征较为匹配的权重文件。在第二步参数微调过程中，将预训练得到的权重文件导入目标模型，此时目标模型便拥有了与目标场景相似的部分前置知识。在此基础上进行第二轮训练，即可充分利用两个数据集的信息，引导模型关注正确的对象特征，尽量避免过拟合。

图 5 迁移学习技术路线示意  
Fig. 5 Principle of transfer learning

### 1.3.3 裂缝识别模型的损失函数

本研究使用的裂缝识别模型包含两种输出格式，并对每种格式设置了对应的损失函数。对于框选格式的识别结果，通过位置损失函数、类别损失函数、置信度损失函数量化预测框与真实框之间的位置相似度和对象分类的准确性；对于网格格式的预测结果，设置了类别损失函数与置信度损失函数。损失函数的详细计算方法参见文献 [25]。

## 2 数据采集与图像处理

### 2.1 图像来源

本研究使用的桥梁表面图像采集于广州黄龙带特大桥，为实际桥检工作中使用无人机拍摄到的桥梁表面图像。在图像拍摄过程中，通过无人机路径规划与扫描式拍摄遍历了桥梁的各索塔表面、梁底、梁侧位置，得到了桥梁表面裂缝全面检测所需的完整图像数据集。采集得到的原始图像尺寸为  $8192 \times 5460$  像素，综合考虑深度学习模型的输入尺寸需求，将图像分割为  $1024 \times 682$  像素的子图像，从中选择部分图像构建数据集。

### 2.2 图像分类数据集的建立

采集得到的数据集中大部分为正对桥梁表面拍摄的规整无污染的混凝土表面图像，同时也包含部分环境背景、附属结构、污染等干扰图像。通过对整个数据集进行观察分析，将不适用于裂缝自动识别、需要剔除或交由人工识别的图像分为四类：背景干扰、附属结构干扰、缆索干扰以及混凝土表面严重污染。各类图像示例如图 6。

图 6 图像分类数据集 各类别图像示例  
Fig. 6 Image examples of various categories in image classification dataset

从 2.1 节分割得到的  $1024 \times 682$  像素子图像中人工筛选四类干扰图像与正常混凝土表面图像每类各 1000 张，构成共包含 5000 张图像的图像分类数据集。由于  $1024 \times 682$  像素子图像与图像分类模型要求的输入尺寸 ( $896 \times 896$  像素) 未对齐，在

人工分类标记前对分类数据集中的图像进行了缩放与裁切。将子图像等比例放大为  $1345 \times 892$  像素，而后取左侧  $892 \times 892$  像素的部分作为深度学习模型的输入。其中，正常混凝土表面图像仅指代无严重污染或干扰的混凝土表面图像，不论是否含有裂缝，只要图像干净规整、符合自动识别条件即算作此类。图像分类数据集中这一类别也包含了裂缝图像、无裂缝图像两类，从而保证数据集具有较强的代表性。

### 2.3 裂缝识别数据集的建立

从正常混凝土表面图像中筛选出带有裂缝的 6400 张图像，构建裂缝识别数据集，图像尺寸同样为  $1024 \times 682$  像素。

由于本数据集中桥梁结构裂缝多数为水平或竖直方向，为增强模型的泛化性，选择 50% 的图像进行了  $-30^\circ \sim +30^\circ$  之间的随机角度旋转。由于深度学习模型具有固定的输入尺寸（640 像素高），因此在前处理阶段又进行了图像缩放。输入裂缝识别模型的图像尺寸共两类：对于未旋转的图像，缩放为  $960 \times 640$  像素；对于需要旋转的图像，在旋转、边角裁切、缩放后输入模型的图像尺寸为  $640 \times 640$  像素。两类图像每类各 3200 张。

正常混凝土表面图像虽图像质量较高、适用于裂缝自动识别方法，但仍然包含混凝土接缝、线缆两类干扰性杂物（如图 7），这些干扰对象易与裂缝之间产生混淆，需要在识别模型中新建类别并针对性训练加以解决，在数据集构建过程中也需要针对这两类对象进行标注。

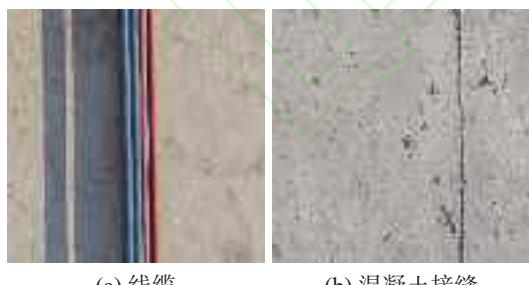


图 7 需要标注的混凝土表面干扰物

Fig. 7 Concrete surface interferences that need to be labeled

对以上三种对象类别，均进行了标注框（box）、标注网格（grid）两种格式的人工标注。标注网格的划分规则为：沿高度方向将网格划分为 20 行，即每个网格高度为 32 个像素；按照图像的长宽比例计算宽度方向的网格划分列数，即  $960 \times 640$  像素的图像划分为 30 列， $640 \times 640$  像素的图

像划分为 20 列。

整个裂缝识别数据集中各类识别目标的 box/grid 标签数量如表 4 所示。

表 4 裂缝识别数据集 各类 box/grid 标签数量

Table 4 Number of various box/grid labels in crack identification dataset

类别	box 标签数量	grid 标签数量
裂缝	7182	143634
接缝	1347	38411
线缆	133	6564

### 2.4 数据集划分方法

对于图像分类数据集和裂缝识别数据集，均按照 8 : 2 的比例划分训练集和验证/测试集。在裂缝识别数据集中，对于进行了旋转处理的图像和未进行旋转处理的图像采用了相同的数据集划分比例。

## 3 桥梁表面图像分类筛选方法

### 3.1 模型训练的编程环境

本研究中深度学习分类模型的程序使用 python3.8 进行编写，采用的代码框架为 pytorch1.12。模型训练使用的计算机配置为 Intel Xeon Gold 4362 处理器和 NVIDIA A100 Tensor Core 显卡。使用 GPU 驱动模式加速模型训练，CUDA 版本为 11.3。

### 3.2 试验对比与讨论

图像 batchsize 取 64，使用 adam 参数优化器，初始学习率为 0.003，进行模型训练。对原始 MobileNetV3 模型、输入尺寸扩充模型、特征层缩减模型、空间注意力机制模型（按顺序编号为分类模型①~分类模型④）分别进行训练，训练过程中记录损失函数与准确度随迭代次数增加的变化情况，并记录训练过程中模型在验证集上达到的最高准确度，输出对应的分类混淆矩阵用于结果分析。

准确度定义方式如下：将各类对象中模型预测置信度最高的一类作为模型的预测结果，预测值和真实结果相同的图像判定为预测准确与真实结果不同则为预测错误；预测准确的图像数量占图像总数的比例为模型准确度。由于本模型的主要作用为筛选剔除不适合输入自动识别算法的图像，正常混凝土表面图像与其他图像之间的二分类准确度也是评价模型性能的重要参考指标之一。

分类模型①在训练集和验证集上的损失函数、分类准确率随迭代次数增加的变化情况如图 8

所示。

随迭代次数增加，模型在训练、验证集上的损失值逐渐降低并最终趋于稳定，验证集预测准确度逐渐上升，最终稳定于 0.93 左右，准确度最高达到 0.953。模型充分收敛、各项数值基本稳定后，模型在验证集中的准确度略低于训练集，但验证集准确度没有出现明显的降低，可以认为训练后期基本没有出现过拟合现象。

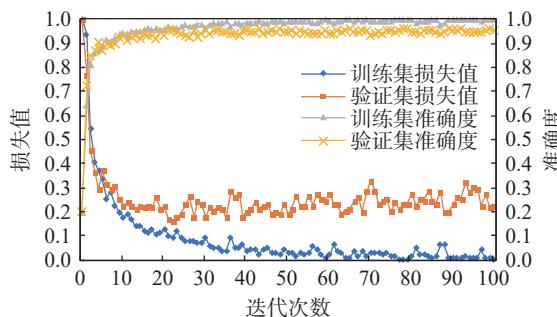


图 8 分类模型①指标变化曲线  
Fig. 8 Index curve for classification model I

训练结果表明，MobileNetV3 模型虽然属于轻量化模型、参数量相对较小，但基本能够满足图像分类场景的应用需求，表现出了较强的模型性能；训练过程中未出现明显的过拟合现象，模型具有较强的泛化性能。

取训练过程中效果最佳的模型进行准确度测试，模型在测试集上的分类混淆矩阵如图 9 所示。



图 9 模型①在测试集上的分类混淆矩阵  
Fig. 9 Classification confusion matrix of model I  
on the test set

分析混淆矩阵可知，该模型准确度的主要影响因素为正常混凝土表面与严重污染混凝土表面之间的混淆、环境背景与附属设施之间的混淆。其中，影响正常混凝土表面图像与干扰图像之间二分类准确度的因素主要为正常混凝土表面图像与严重污染混凝土表面图像之间的混淆。推测产

生混淆的主要原因为：正常图像与污染图像之间的主要区别在于图像的细节纹理，而模型的输入尺寸过小、降采样倍数过高导致前处理过程中细节信息部分丢失，影响分类准确率。

分类模型②在训练集和验证集上的损失函数、分类准确率随迭代次数增加的变化情况如图 10 所示。

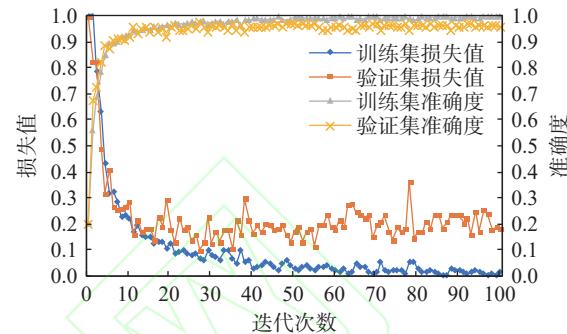


图 10 分类模型②指标变化曲线  
Fig. 10 Index curve for classification model II

分类模型②的收敛性、泛化性与分类模型①相似，但模型收敛时的分类准确率有所提高，最高准确率提升至 97.2%。

取训练过程中效果最佳的模型进行准确度测试，模型在测试集上的分类混淆矩阵如图 11 所示。

		附属结构	缆索干扰	严重污染	环境背景	正常图像	
真实值	附属结构	0.005	0.020	0.000	0.000	0.975	
	缆索干扰	0.006	0.000	0.000	0.988	0.006	
严重污染	严重污染	0.029	0.000	0.971	0.000	0.000	
	环境背景	0.004	0.971	0.000	0.000	0.025	
正常图像	正常图像	0.965	0.000	0.015	0.000	0.020	
	环境背景	0.000	0.000	0.000	0.000	0.000	
		正常	环境	严重	缆索	附属	
		图像	背景	污染	干扰	结构	预测值

图 11 模型②在测试集上的分类混淆矩阵  
Fig. 11 Classification confusion matrix of model II  
on the test set

分析对比混淆矩阵可知，与模型①相比，正常混凝土表面与严重污染混凝土表面之间的混淆现象有所缓解，说明输入尺寸扩充后模型的图像细节处理能力有所增强。此外，在添加了两个 bNeck 模块、模型的网络深度增加后，模型并未表现出明显的性能退化，验证了残差连接机制有效解决了深度增加导致的模型性能退化问题。

分类模型③在训练集和验证集上的损失函

数、分类准确率随迭代次数增加的变化情况、测试集混淆矩阵如图 12、图 13 所示。

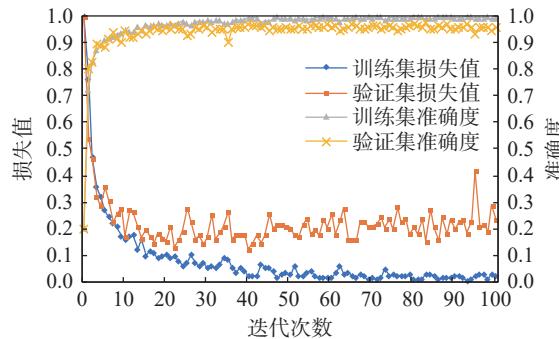


图 12 分类模型③指标变化曲线

Fig. 12 Index curve for classification model III

真实值 真实值	预测值				
	正常	环境	严重	缆索	附属
图像	0.000	0.015	0.000	0.000	0.985
背景	0.000	0.000	0.000	0.994	0.006
污染	0.044	0.000	0.956	0.000	0.000
干扰	0.004	0.971	0.000	0.000	0.025
结构	0.980	0.000	0.005	0.000	0.015

图 13 模型③在测试集上的分类混淆矩阵  
Fig. 13 Classification confusion matrix of model III on the test set

与分类模型②相比，分类模型③虽然减少了模型参数量，但分类准确率与模型②基本相同（最高准确率为 97.4%），说明特征层缩减能够有效地精简网络结构，对本任务场景下的分类准确度没有明显影响。

分类模型④在训练集和验证集上的损失函数、分类准确率随迭代次数增加的变化情况如图 14 所示。

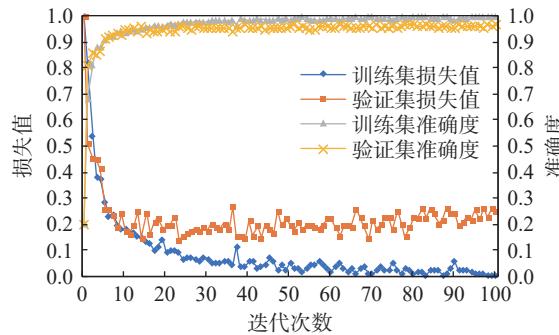


图 14 分类模型④指标变化曲线

Fig. 14 Index curve for classification model IV

与其他模型相比，模型④在迭代过程中的收敛速度更快，收敛时的最高准确率（97.9%）也高于

其他三个分类模型，说明添加空间注意力机制对模型性能的提升有积极作用。

真实值 真实值	预测值				
	正常	环境	严重	缆索	附属
图像	0.000	0.010	0.000	0.000	0.990
背景	0.006	0.000	0.000	0.988	0.006
污染	0.015	0.000	0.985	0.000	0.000
干扰	0.004	0.975	0.000	0.000	0.021
结构	0.980	0.000	0.010	0.000	0.010

图 15 模型④在测试集上的分类混淆矩阵

Fig. 15 Classification confusion matrix of model IV on the test set

为进一步验证如前所述几个模型结构修改方式的有效性，选择扩充输入尺寸、未缩减特征层、应用了空间注意力机制的模型作为对照组进行消融试验。该对照模型应用的改进措施最多、结构最复杂，内部参数量也最大，理论上应具有最高的识别准确率。分别取消输入尺寸扩充措施、取消空间注意力机制、取消通道注意力机制、缩减特征层数量，与消融试验对照组作对比，探究几类改进措施对准确率的影响。除各组准确率外，也测算了个模型的内部参数量以及对于单图像的推理用时，结果整理如表 5。

表 5 消融试验各组模型结构与模型性能

Table 5 Model Performance of the models in the ablation test

序号	相较于对照组的修改	准确度/ (%)	模型参数量/M	单张图像平均推理用时/ms
1	无	97.5	2.31	28
2	取消尺寸扩充	95.3	1.59	22
3	取消空间注意力机制	97.2	2.14	26
4	取消通道注意力机制	95.2	2.28	28
5	缩减特征层数量	97.9	1.72	22

由试验结果可知，输入尺寸扩充措施和两种注意力机制的应用均对提升模型性能有一定作用，而特征层数量缩减并未对模型准确度产生明显影响。综合模型准确率、参数量集单张图像平均推理用时，可知试验 5 中的模型准确率、处理效率均较高，最适用于实际桥检工作。

## 4 桥梁表面图像裂缝识别方法

### 4.1 模型训练的编程环境

深度学习裂缝识别模型的编程环境、训练过

程中的硬件配置与第三节中分类模型的编程环境相同。

## 4.2 模型评价指标

### 1) 框选预测准确性评价指标

由于识别模型中框选格式的预测框与真实框之间无法完全重合，需要通过设置 IOU 阈值判断框选预测的准确性。对于有部分重叠的预测框和真实框，二者之间的 IOU 计算公式如下：

$$\text{IOU} = S_I / S_U \quad (2)$$

式中： $S_I$  为两框交集的面积； $S_U$  为两框并集的面积，如图 16 所示。

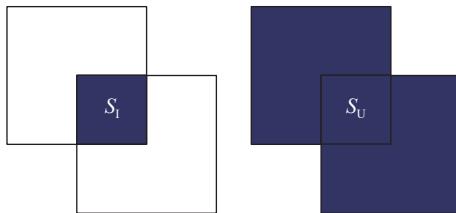


图 16  $S_I$  与  $S_U$  计算方式示意

Fig. 16 Calculation method of  $S_I$  and  $S_U$

当计算得到的 IOU 值高于预先统一设置的 IOU 阈值时，预测框与真实框之间匹配，认为该预测框为真阳性 (TP)；若预测框没有与之匹配的真实框，则记为假阳性 (FP)；若真实框没有与之匹配的预测框，则记为假阴性 (FN)。在本研究中取 IOU 阈值为 0.6，略高于一般场景下的常规取值，以保证裂缝定位算法具有足够的识别准确度与定位精确度。

根据以上三个指标计算精确率 (P) 与召回率 (R)：

$$P = \frac{TP}{TP+FP} \quad (3)$$

$$R = \frac{TP}{TP+FN} \quad (4)$$

式中， $TP$ 、 $FP$ 、 $FN$  代表验证集中真阳性、假阳性、假阴性出现的总次数。

精确率、召回率不仅受 IOU 阈值影响，也受到预测置信度阈值的影响。随置信度阈值增大，预测结果的召回率降低、精确率升高；由于实际应用场景需要模型在同一置信度阈值下同时保持较高的精确率和召回率、尽可能减少漏判与误判情况的发生，因此一般将模型的平均精度 (AP) 作为模型性能的评价指标。AP 值通过计算 P-R 曲线下包围的面积得出，该指标仅与模型的性能相关联，不受后处理过程中置信度阈值参数的影响。对于同一模型，每种识别对象对应一个 AP 值，由于本研究重点关注裂缝识别的准确性，因此将裂缝类别的 AP 值作为框选预测准确性的评价指标。

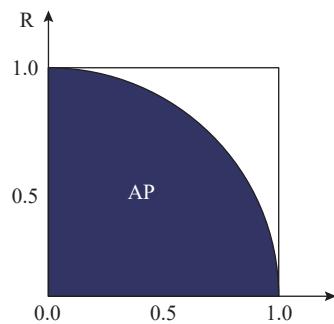


图 17 P-R 曲线与 AP 计算方式

Fig. 17 P-R curve and AP calculation

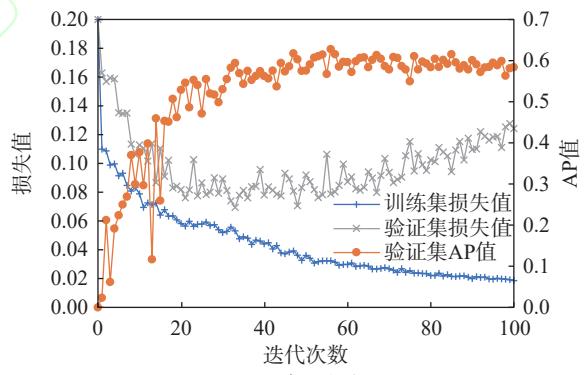
### 2) 网格预测准确性评价指标

由于网格数据格式的真实值和预测值均以预先划分好的网格为基础，在对二者进行匹配时不需要设置 IOU 阈值，只要网格某一位置的模型预测值与真实值相同即视为匹配，否则为不匹配。采用与框预测相同的计算方式统计 TP、FP、FN 指标，计算 P-R 曲线与 AP 值，同样将裂缝类别的 AP 值作为网格预测准确性的评价指标。

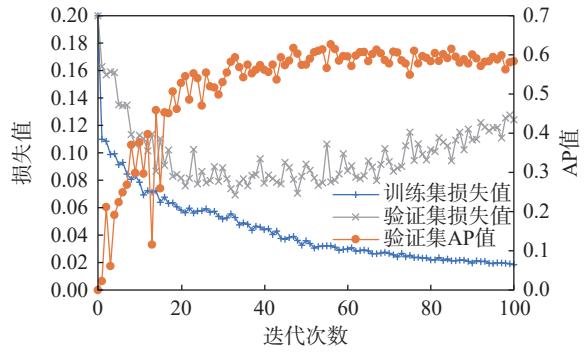
## 4.3 试验对比与讨论

### 4.3.1 非迁移学习模型的模型性能

首先只使用桥梁裂缝识别数据集进行模型训练，迭代共 100 轮，框预测、网格预测在训练集、测试集上的损失值和 AP 值随迭代次数增加的变化情况如图 18 所示。



(a) 框预测



(b) 网格预测

图 18 非迁移学习模型指标曲线

Fig. 18 Index curve of non transfer learning model

观察验证集中网格预测损失值变化情况可以发现，在训练约 40 轮后验证集损失值开始上升，网格预测在验证集上的识别准确度开始出现下降趋势，判断此时模型开始出现过拟合。

挑选训练过程中裂缝识别框预测 AP 值最高的模型进行模型测试，框预测和网格预测的各类别 P-R 曲线如图 19。

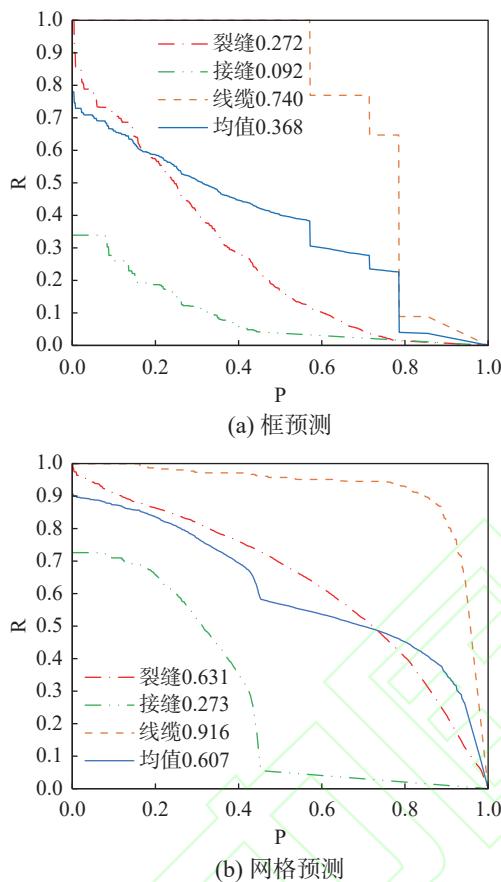


图 19 非迁移学习模型在测试集上的 P-R 曲线及 AP 值  
Fig. 19 P-R curve and AP value of the non transfer learning model on the test set

对于裂缝识别任务，预测框和预测网格的最高 AP 值分别为 0.272、0.631，裂缝网格识别的效果相对更好；综合以上实验结果，可以发现对于裂缝类别和接缝类别，框选格式的识别准确度明显不足。推测影响准确度的主要原因为：裂缝与接缝的几何形态过于细长，在边界框中对裂缝识别提供参考的有效像素占比过低，使得裂缝与接缝的框选识别任务本身较为困难，准确提取图像特征所需的样本量更大；数据集中的样本量相对不足，导致模型在训练过程中没有有效提取到裂缝的图像特征，峰值准确度较低且较早出现了过拟合现象。

#### 4.3.2 迁移学习模型的模型性能

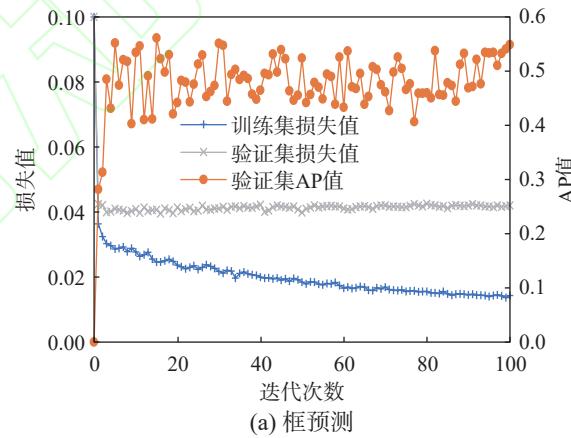
在迁移学习过程中，选择了一般混凝土表面裂缝识别作为预训练任务。预训练使用的数据集现阶段共包括 18 000 张各类混凝土结构表面图像，包含接缝干扰类别和标注框、标注网格两种标注格式，各类目标的 box/grid 标签数量如表 6 所示。

表 6 预训练裂缝识别数据集 各类 box/grid 标签数量

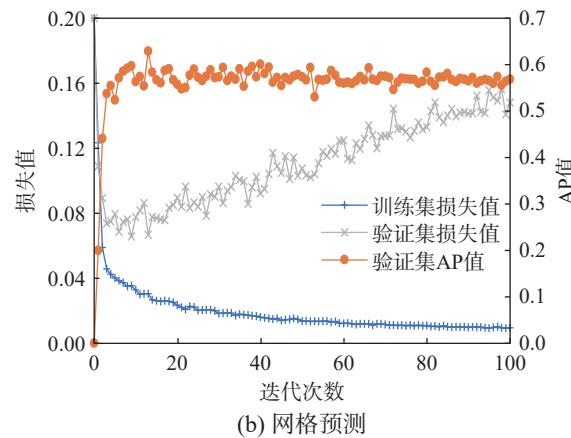
Table 6 Number of various box/grid labels in the pre-training crack identification dataset

类别	box 标签数量	grid 标签数量
裂缝	23 158	425 366
接缝	13 268	267 695

完成预训练后，将得到的权重数据导入桥梁裂缝识别模型进行迁移学习。在迁移学习的前 20 个迭代轮次中，冻结主干网络的模型权重，仅对检测头进行训练；迭代 20 轮后解除权重冻结继续训练，共迭代 100 轮。框预测、网格预测在训练集、测试集上的损失值和 AP 值随迭代次数增加的变化情况如图 20。



(a) 框预测



(b) 网格预测

图 20 迁移学习模型指标曲线  
Fig. 20 Index curve of transfer learning model

挑选训练过程中裂缝识别框预测 AP 值最高的模型进行模型测试，框预测和网格预测的各类别

P-R 曲线如图 21:

与未使用迁移学习的模型相比, 迁移学习模型在迭代过程中的收敛速度明显更快, 两种格式的识别准确度也有提升, 预测框和预测网格的最高 AP 值分别提升至 0.352、0.656。将两次实验结果进行对比, 可知迁移学习方法有效发挥了水泥路面裂缝数据集的类比参考作用, 有效提升了模型性能。在训练中后期, 迁移学习模型出现了与无迁移学习模型相似的过拟合现象, 且本次训练过程中准确度降低的趋势表现得更加明显。分析这一现象的原因主要为: 模型为进一步提升训练集上的识别准确度, 在训练中后期的迭代过程中重新开始关注与裂缝无关的图像背景细节, 逐步舍弃了预训练过程中获取的水泥路面裂缝信息。

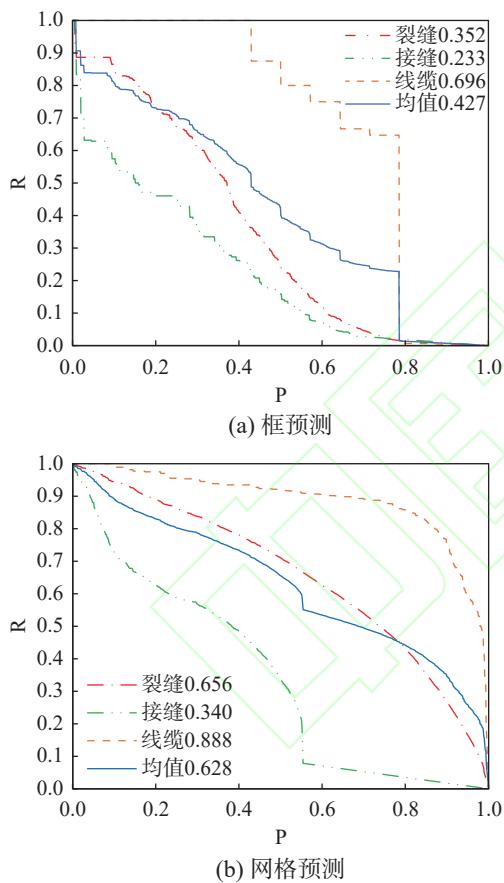


图 21 迁移学习模型在测试集上的 P-R 曲线及 AP 值  
Fig. 21 P-R curve and AP value of the transfer learning model on the test set

## 5 总结与展望

本文通过分别开发桥梁表面图像分类筛选模型与桥梁裂缝识别模型, 提出了桥梁表面裂缝的两阶段识别方法, 并结合实际工程数据对该方法的可行性进行了验证。得出了以下结论:

(1) 通过模型结构调整、添加空间注意力机制, 图像分类模型对图像细节的处理能力有所增强, 分类准确度有所提升。

(2) 由于桥梁裂缝的图像特征较为细长、裂缝形态比较复杂, 在现有数据集条件下裂缝的框选预测准确度较低, 模型的泛化性能也不强; 迁移学习方法的引入略微改善了框选识别方法的性能, 但在训练后期模型过拟合较为明显, 为进一步提升框选识别方法的准确度需要继续扩充数据集。与框选识别方法相比, 网格裂缝识别方法的准确度相对更高, 基本满足裂缝识别的应用需求。

## 参考文献:

- [1] 黄军生. 钢筋混凝土桥梁裂缝成因综述[J]. 世界桥梁, 2002, 30(2): 59—63.  
HUANG Junsheng. A summary of the causes of cracking in reinforced concrete bridges [J]. World Bridges, 2002, 30(2): 59—63. (in Chinese)
- [2] 段翔远. 基于模糊理论的混凝土公路桥梁技术状况评定方法研究[J]. 市政技术, 2023, 41(7): 61—64, 105.  
DUAN Xiangyuan. Study on evaluation method for technical condition of concrete highway bridges based on fuzzy theory [J]. Journal of Municipal Technology, 2023, 41(7): 61—64, 105. (in Chinese)
- [3] 盛宇静, 王春. 某简支空心斜板梁桥结构性病害成因分析[J]. 市政技术, 2023, 41(1): 172—177.  
SHENG Yujing, WANG Chun. Analysis of structural damage cause of a simply supported hollow skew beam bridge [J]. Journal of Municipal Technology, 2023, 41(1): 172—177. (in Chinese)
- [4] 钟新谷, 彭雄, 沈明燕. 基于无人飞机成像的桥梁裂缝宽度识别可行性研究[J]. 土木工程学报, 2019, 52(4): 52—61.  
ZHONG Xingu, PENG Xiong, SHEN Mingyan. Study on the feasibility of identifying concrete crack width with images acquired by unmanned aerial vehicles [J]. China Civil Engineering Journal, 2019, 52(4): 52—61. (in Chinese)
- [5] 阮小丽, 王波, 荆国强, 等. 桥梁混凝土结构表面裂缝自动识别技术研究[J]. 世界桥梁, 2017, 45(6): 55—59.  
RUAN Xiaoli, WANG Bo, JING Guoqiang, et al. Study of automatic identification technology for surface cracks in bridge concrete structures [J]. World Bridges, 2017, 45(6): 55—59. (in Chinese)
- [6] 谢文高, 张怡孝, 刘爱荣, 等. 基于水下机器人与数字图像技术的混凝土结构表面裂缝检测方法[J]. 工程力学, 2022, 39(增刊 1): 64—70.  
XIE Wengao, ZHANG Yixiao, LIU Airong, et al. Method for concrete surface cracking detection based on Rov and digital image technology [J]. Engineering Mechanics, 2022, 39(Suppl 1): 64—70. (in Chinese)
- [7] 常广利, 冷冰, 杨凯银. 基于多旋翼无人机桥梁裂缝智能化识别分析[J]. 市政技术, 2021, 39(10): 46—48.  
CHANG Guangli, LENG Bing, YANG Kaiyin. Intelligent identification of bridge cracks based on multi-rotor UAV [J]. Journal of Municipal Technology, 2021, 39(10): 46—48. (in Chinese)

- [8] 杨子涵, 舒江鹏, 杨静滢, 等. 基于 DIC 技术的钢筋混凝土土梁剪切裂缝自动提取与量化方法[J]. *工程力学*, 2024, 41(增刊 1): 187—196.  
YANG Zihan, SHU Jiangpeng, YANG Jingying, et al. Automation extraction and quantification of shear cracks in reinforced concrete beams based on DIC technology [J]. *Engineering Mechanics*, 2024, 41(Suppl 1): 187—196. (in Chinese)
- [9] PRASANNA P, DANA K J, GUCUNSKI N, et al. Automated crack detection on concrete bridges [J]. *IEEE Transactions on Automation Science and Engineering*, 2016, 13(2): 591—599.
- [10] 杨勇, 赵维刚. 基于深度图像和彩色图像相结合的高速铁路桥梁裂缝宽度检测识别算法研究[J]. 国防交通工程与技术, 2015, 13(6): 12—16.  
YANG Yong, ZHAO Weigang. On the depth-image-and-color-image-combination-based calculation method for the crack-width detection of high-speed railway bridges [J]. *Traffic Engineering and Technology for National Defense*, 2015, 13(6): 12—16. (in Chinese)
- [11] 周浩翔, 刘爱荣, 李伟财, 等. 基于水下机器人的水下混凝土结构表观病害三维重建方法[J]. *工程力学*, 2024, 41(增刊 1): 129—135.  
ZHOU Haoxiang, LIU Airong, LI Weicai, et al. Method for 3D reconstruction of surface diseases of underwater concrete structures based on Rov [J]. *Engineering Mechanics*, 2024, 41(Suppl 1): 129—135. (in Chinese)
- [12] 刘洪公, 王学军, 李冰莹, 等. 基于卷积神经网络的桥梁裂缝检测与识别[J]. 河北科技大学学报, 2016, 37(5): 485—490.  
LIU Honggong, WANG Xuejun, LI Bingying, et al. Detection and recognition of bridge crack based on convolutional neural network [J]. *Journal of Hebei University of Science and Technology*, 2016, 37(5): 485—490. (in Chinese)
- [13] 李良福, 马卫飞, 李丽, 等. 基于深度学习的桥梁裂缝检测算法研究[J]. 自动化学报, 2019, 45(9): 1727—1742.  
LI Liangfu, MA Weifei, LI Li, et al. Research on detection algorithm for bridge cracks based on deep learning [J]. *Acta Automatica Sinica*, 2019, 45(9): 1727—1742. (in Chinese)
- [14] XU H Y, SU X, WANG Y, et al. Automatic bridge crack detection using a convolutional neural network [J]. *Applied Sciences*, 2019, 9(14): 2867.
- [15] 梁雪慧, 程云泽, 张瑞杰, 等. 基于卷积神经网络的桥梁裂缝识别和测量方法[J]. 计算机应用, 2020, 40(4): 1056—1061.  
LIANG Xuehui, CHENG Yunze, ZHANG Ruijie, et al. Bridge crack classification and measurement method based on deep convolutional neural network [J]. *Journal of Computer Applications*, 2020, 40(4): 1056—1061. (in Chinese)
- [16] 李雯雯, 李喜媛, 周健, 等. 基于全卷积神经网络和朴素贝叶斯数据融合的桥梁裂缝识别算法[J]. 公路交通科技, 2023, 40(2): 44—52.  
LI Wenwen, LI Xiyuan, ZHOU Jian, et al. An algorithm for recognizing bridge cracks based on full convolution neural network and naive Bayesian data fusion [J]. *Journal of Highway and Transportation Research and Development*, 2023, 40(2): 44—52. (in Chinese)
- [17] 余加勇, 李锋, 薛现凯, 等. 基于无人机及 Mask R-CNN 的桥梁结构裂缝智能识别[J]. 中国公路学报, 2021, 34(12): 80—90.  
YU Jiayong, LI Feng, XUE Xiankai, et al. Intelligent identification of bridge structural cracks based on unmanned aerial vehicle and mask R-CNN [J]. *China Journal of Highway and Transport*, 2021, 34(12): 80—90. (in Chinese)
- [18] 刘森森. 基于迁移学习的桥梁裂缝智能检测与识别[J]. 信息技术与信息化, 2021(7): 233—235.  
LIU Sensen. Intelligent detection and recognition of bridge cracks based on transfer learning [J]. *Information Technology and Informatization*, 2021(7): 233—235. (in Chinese)
- [19] 张振海, 季坤, 党建武. 基于桥梁裂缝识别模型的桥梁裂缝病害识别方法[J]. 吉林大学学报(工学版), 2023, 53(5): 1418—1426.  
ZHANG Zhenhai, JI Kun, DANG Jianwu. Crack identification method for bridge based on BCEM model [J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2023, 53(5): 1418—1426. (in Chinese)
- [20] 余加勇, 刘宝麟, 尹东, 等. 基于 YOLOv5 和 U-Net3+ 的桥梁裂缝智能识别与测量[J]. 湖南大学学报(自然科学版), 2023, 50(5): 65—73.  
YU Jiayong, LIU Baolin, YIN Dong, et al. Intelligent identification and measurement of bridge cracks based on YOLOv5 and U-Net3+ [J]. *Journal of Hunan University (Natural Sciences)*, 2023, 50(5): 65—73. (in Chinese)
- [21] 梁栋, 李英俊, 张少杰. 融合改进 ResNet-14 和 RS-Unet 模型的混凝土桥梁裂缝识别[J]. 北京交通大学学报, 2023, 47(3): 10—18.  
LIANG Dong, LI Yingjun, ZHANG Shaojie. Identification of cracks in concrete bridges through fusing improved ResNet-14 and RS-Unet models [J]. *Journal of Beijing Jiaotong University*, 2023, 47(3): 10—18. (in Chinese)
- [22] WANG J K, HE X H, SHAO F M, et al. A real-time bridge crack detection method based on an improved inception-resnet-v2 structure [J]. *IEEE Access*, 2021, 9: 93209—93223.
- [23] 高庆飞, 王宇, 刘晨光, 等. 基于卷积神经网络算法的混凝土桥梁裂缝识别与定位技术[J]. 公路, 2020, 65(9): 268—274.  
GAO Qingfei, WANG Yu, LIU Chenguang, et al. Identifying and positioning technologies of concrete bridge crack based on convolutional neural network [J]. *Highway*, 2020, 65(9): 268—274. (in Chinese)
- [24] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE, 2019: 1314—1324.
- [25] LI B L, QI Y, FAN J S, et al. A grid-based classification and box-based detection fusion model for asphalt pavement crack [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2023, 38(16): 2279—2299.
- [26] ZHU X Z, CHENG D Z, ZHANG Z, et al. An empirical study of spatial attention mechanisms in deep networks [C]// IEEE/CVF International Conference on Computer Vision (ICCV) Seoul: IEEE, 2019: 6687—6696.