

# 用于桥梁底部裂缝无人机检测分类的轻量模型 VGG-A9

虞志远<sup>1</sup>, 代春泉<sup>1,2+</sup>, 王全磊<sup>1,2</sup>

1. 山东科技大学 土木工程与建筑学院, 山东 青岛 266590

2. 山东科技大学 山东省土木工程防灾减灾实验室, 山东 青岛 266590

+ 通信作者 E-mail: dcqwin@sdust.edu.cn

**摘要:** 桥梁底部裂缝因位置隐蔽导致传统人工检测方法效率低、风险高, 现有深度学习检测算法模型参数大、计算复杂度高, 难以在嵌入式设备中实时部署。为此, 提出用于桥梁底部裂缝无人机检测分类的轻量模型 VGG-A9。首先, 通过无人机系统采集桥梁底部裂缝图像, 并采用语义分割驱动的图像二值化处理, 消除光照干扰并保留裂缝形态特征, 构建包含 2720 张裂缝图像的数据集。其次, 简化 VGG 模型, 通过削减网络全连接层参数, 在 7 层卷积层嵌入批量归一化(Batch Normalization, BN)与 Dropout 优化策略, 将 VGG-A9 模型参数压缩至  $0.74 \times 10^6$ 。实验表明, VGG-A9 在自建数据集验证集上的分类准确率达 98.53%, 部署至无人机嵌入式平台后, 单帧分类耗时 21ms, 准确率 98.01%; 在公开道路裂缝数据集 EdmCrack600 的鲁棒性测试中准确率达 93.67%, 显著优于其它模型。该方法在计算资源受限的场景中具备高效性与准确性, 为桥梁智能化巡检提供了可行的技术路径。

**关键词:** 桥梁工程; 机器视觉; 裂缝检测; 无人机; VGG

文献标志码: A 中图分类号: TP391 doi: 10.3778/j.issn.1002-8331.2505-0162

## Lightweight model VGG-A9 for UAV detection and classification of cracks at the bottom of bridges

YU Zhiyuan<sup>1</sup>, DAI Chunquan<sup>1,2+</sup>, WANG Quanlei<sup>1,2</sup>

1. College of Civil Engineering and Architecture, Shandong University of Science and Technology, Qingdao, Shandong 266590, China

2. Shandong Civil Engineering Disaster Prevention and Mitigation Laboratory, Shandong University of Science and Technology, Qingdao, Shandong 266590, China

**Abstract:** Bridge bottom cracks lead to low efficiency and high risks in traditional manual detection methods due to their concealed location, while existing deep learning detection algorithms face difficulties in real-time deployment on embedded devices owing to large model parameters and high computational complexity. To address this, a lightweight model VGG-A9 is proposed for unmanned aerial vehicle (UAV) detection and classification of cracks at the bottom of bridges. Firstly, bridge bottom crack images are collected through a UAV system, and a semantic segmentation-driven image binarization method is applied to eliminate illumination interference while retaining crack morphological features, constructing a dataset containing 2,720 crack images. Secondly, the VGG model is simplified. By reducing parameters in the fully connected layers of the VGG network and embedding Batch Normalization (BN) and

**基金项目:** 山东省住房城乡建设科技计划项目(2022-K7-6)。

**作者简介:** 虞志远(2001—), 男, 硕士研究生, CCF 学生会员, 研究方向为机器视觉; 代春泉, 通信作者, 博士, 副教授; 王全磊, 博士研究生。

Dropout optimization strategies into seven convolutional layers, compressing the parameters of the VGG-A9 model to  $0.74 \times 10^6$ . Experiments show that VGG-A9 achieves a validation accuracy of 98.53% on the self-built dataset. After deployment to the UAV embedded platform, it achieves a per-frame classification time of 21 ms with an accuracy of 98.01%. In robustness testing on the public road crack dataset EdmCrack600, it achieves an accuracy of 93.67%, significantly outperforming other models. The approach exhibits high efficiency and accuracy in computationally constrained scenarios, providing a viable technical solution for intelligent bridge inspection.

**Key words:** bridge engineering; machine vision; crack detection; UAV; VGG

桥梁底部裂缝因位置隐蔽难以及时发现,传统的人工检测方法存在安全性差、效率低且难以覆盖隐蔽区域等问题<sup>[1]</sup>。基于机器视觉的智能检测技术通过自动化图像采集与深度学习算法,可显著提升检测效率和安全性<sup>[2]</sup>,为自动化、智能化桥梁巡检提供了新途径。

基于对自动化解决方案的需求,实现桥梁底部裂缝的智能检测需要采集高质量的裂缝图像作为基础步骤。由于具有便携性、操作简单性和自动飞行控制等优点,无人机已成为桥梁检查任务中的一种常用设备<sup>[3]</sup>。Wang<sup>[4,5]</sup>和 Xu<sup>[6]</sup>等人的研究证实了无人机系统在复杂桥梁环境下捕捉结构细节的可行性,拍摄的图像是后续机器视觉分析的关键输入数据。Lee 等<sup>[7]</sup>使用无人机进行图像采集,并结合基于深度学习的裂缝分类方法与改进的图像处理技术检测量化裂缝。Li 等<sup>[8]</sup>在无人机上安装平行的激光发射器进行裂缝图像采集,并采用校正算法对图像进行几何调整。Shu 等人<sup>[9]</sup>基于 BIM 的箱梁桥几何建模实现全自动轨迹规划,通过坐标-序列映射算法生成最优巡检路径,显著提升导航效率与轨迹精度。Xu 等<sup>[10]</sup>提出增强蛇优化算法,融合混沌映射初始化、减法平均优化器及自适应反向学习策略,有效解决高维环境局部最优与收敛迟滞问题。Shang 等<sup>[11]</sup>基于粒子群优化方法提出了一种同时平衡图像质量与飞行路径的算法,能够达到更好的路径规划效果与更快的优化速度。

在无人机采集图像方法不断进步的同时,研究人员也把基于机器学习的分类算法应用于桥梁裂缝分类。Ding 等<sup>[12]</sup>在 BP 神经网络中引入专家系统对预测结果

进行修正,实现了在复杂工程条件下分类的动态优化。孙朝云等<sup>[13]</sup>通过深度残差网络改进 SSD 模型结构,根据损失函数优化模型的超参数,有效提高了裂缝分类准确率。王纪武等<sup>[14]</sup>针对传统图像处理算法无法解决过饱和图像和多噪声图像有效分类问题,提出了基于改进 Faster R-CNN 的桥梁裂缝分类方法。He 等<sup>[15]</sup>提出了一种结合 YOLOv7 目标检测网络的图像处理方法,对桥梁裂缝修补痕迹和次生裂缝进行准确分类检测。Liang 等<sup>[16]</sup>提出 PCCNet 网络,基于混洗单元与倒残差架构构建轻量化模型,结合小样本训练策略及特征可视化技术,实现高效裂缝分类。Shi 等<sup>[17]</sup>构建了改进 KAN 系列模型,通过 Kolmogorov-Arnold 网络(KAN)替代传统多层感知机(MLP),利用可学习样条函数实现自适应特征变换,在减少了模型参数的同时显著提高了裂缝分类精度。

尽管上述研究优化了桥梁底部裂缝检测流程,但机载设备的有限算力难以满足复杂模型的实时计算需求,制约了智能化检测应用。因此,提出用于桥梁底部裂缝检测分类的轻量模型 VGG-A9,融合无人机图像采集技术与轻量级卷积网络设计,实现对桥梁底部裂缝的智能巡检。

## 1 桥梁底部裂缝数据集构建

图像数据集来源于安徽省某桥梁检测项目,通过设计的无人机图像采集系统进行实地拍摄得到。

### 1.1 无人机图像采集

经调查研究<sup>[18,19]</sup>,选择大疆公司的经纬 M300 飞行

器, 搭配上置云台组件并搭载禅思 H20 云台, 实现对桥梁底部裂缝的垂直拍摄。裂缝图像实际尺寸、传感器尺寸、焦距和相机镜头距底部距离关系如式 1 所示, 其中,  $j$  代表焦距大小,  $c$  为传感器的靶面宽度,  $d$  为镜头距桥梁底部的距离,  $z$  为实际宽度。

$$j = \frac{c \times d}{z} \quad (1)$$

为保证裂缝检测准确性, 选取精度为 0.01mm 获取裂缝图像, 根据式 1 可以计算得出相机距桥梁底部距离为 560mm, 同时考虑到模型计算量等因素, 所拍摄图像均调整为 256×256 像素的正方形规格。对裂缝图像的采集及处理过程如图 1 所示。



图 1 裂缝图像的采集和尺寸调整

Fig.1 Collection and size adjustment of crack images

设置无人机飞行轨迹如图 2 所示, 通过上置云台发现裂缝并按照距离要求对裂缝拍照。

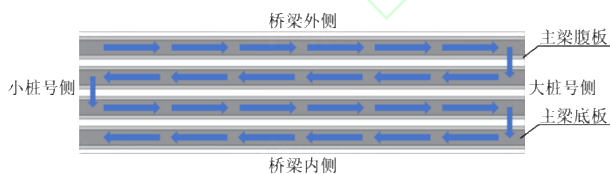


图 2 无人机飞行轨迹示意图

Fig.2 UAV flight trajectory schematic diagram

将无人机自带 GPS 系统与拍摄编号相结合, 可以精确获取裂缝位置。对裂缝图像按照图 3 所示方法进行编号, 其中第一个数字表示第几跨桥梁, 第二个数字表示纵向位置, 第三个数值表示横向位置。如 1-2-3 表示第一跨桥梁、第二纵向的第三处裂缝, 并附带 GPS 坐标。

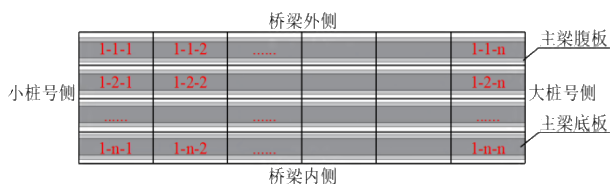


图 3 图像编号示意图

Fig.3 Image numbering diagram

## 1.2 数据集构建

为确保裂缝数据的多样性, 历时两个月完成不同光照条件下、不同位置、不同走向的裂缝拍摄工作, 共获得有效拍摄图像 2720 张。

根据《公路桥梁技术状况评定标准》(JTG/T H21-2011)<sup>[20]</sup>和《公路桥涵养护规范》(JTG 5120-2021)<sup>[21]</sup>对裂缝病害的评定要求, 以及桥梁底部裂缝实际状况, 将桥梁底部裂缝分为横向裂缝、纵向裂缝、斜向裂缝、组合裂缝。将采集的裂缝图像手动分类, 并放置于不同的文件夹中, 建立的数据集如图 4 所示。

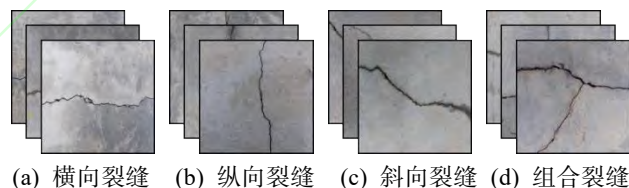


图 4 桥梁底部裂缝图像数据集

Fig.4 Image dataset of cracks at the bottom of bridges

针对传统阈值分割方法对光照敏感、抗噪性不足的缺陷, 对上述分类后图像采用文献<sup>[22]</sup>提出的方法进行二值化。构建适用于桥梁裂缝识别的轻量级改进 DeepLabV3+模型, 将原始图像输入模型进行语义分割, 直接输出裂缝二值化掩码, 部分二值图像如图 5 所示。相较于传统阈值分割方法, 这种语义分割驱动的二值化方法对图像进行像素级分割, 能够较好的处理噪声和边缘信息, 无需进一步的形态学处理。此外, 与原始 RGB 图像相比, 二值化结果有效保留裂缝形态学特征, 消除了色彩冗余信息与光照强度依赖性, 同时抑制了纹

理噪声对分类决策的干扰。将处理后的 2720 张图片按 8:2 的比例划分为训练集和验证集,即训练集 2176 张图片,验证集 544 张图片。构建的桥梁底部裂缝数据集如表 1 所示。

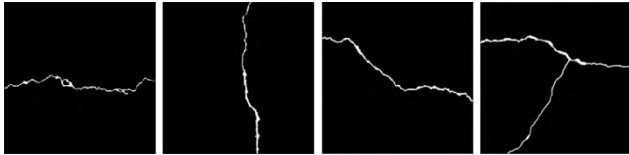


图 5 裂缝二值图像

Fig.5 Binary image of cracks

表 1 桥梁底部裂缝数据集属性

Table 1 Attributes of bridge bottom crack dataset

图像分辨率	256×256
图像格式	单通道二值图像 (背景: 0, 裂缝: 255)
图像总数	2720
裂缝类别	横向、纵向、斜向、组合
类别分布	横向裂缝: 540 (19.9%)
	纵向裂缝: 390 (14.3%)
	斜向裂缝: 830 (30.5%)
	组合裂缝: 960 (35.3%)
训练集/验证集划分比例	8:2

## 2 VGG-A9 模型构建

VGG 模型<sup>[23]</sup>是由牛津大学的视觉几何组(Visual Geometry Group)首先提出的卷积神经网络模型。本研究通过研究卷积神经网络经典的图像分类模型,结合裂缝图像特点和无人机桥梁检测场景中计算资源受限的挑战,改进 VGG 模型。

### 2.1 简化 VGG 模型

本文的裂缝分类输入图像为裂缝二值化图像,相较于原图,特征更明显,更易于分类识别。基于此,简化现有的 VGG 网络模型,简化后的模型共有 9 层,包括

7 层卷积层(Convolution),2 层全连接层(Fully connected),命名为 VGG-A9 模型。该模型中卷积层和池化层迭代组合,提取图像特征,最后汇集于全连接层。每个卷积层后设置 ReLU 函数,经过最大池化(Maxpooling)压缩特征维度,通过 Softmax 函数输出特征向量,得到 4 个类别的识别概率,完成分类任务。模型结构示意图如图 6 所示。

VGG-A9 模型结构解释如下:

#### (1) 输入层

接受单通道二值图像输入,尺寸已在图像预处理阶段统一标准化为 256×256。

#### (2) 卷积层

VGG-A9 模型共设计 7 层卷积层,与原模型一致,其卷积核大小均为 3×3,步长为 1,填充为 SAME Padding,其中第一层至第三层卷积核数量为 16 个,第四层至第六层卷积核数量为 32 个,最后一层卷积层设置 64 个卷积核。卷积层后设置 ReLU 函数,降低计算量,解决梯度饱和问题,使模型具有稀疏性,减少过拟合问题。

#### (3) 池化层

在每一卷积层后均设置池化层,池化选用最大池化,池化核大小为 2×2,数量与卷积核数量保持一致,设置步长为 2。该层对卷积层输出特征图进行下采样操作,压缩特征维度,便于后续计算。

#### (4) 全连接层

VGG-A9 模型设计 2 层全连接层,分别设置 100 个神经元和 4 个神经元,本文模型相比于 VGG 模型在全连接层神经元数量大幅缩减。其原因主要在于该模型输入图像为处理后的二值图像,特征明显,且裂缝图像分类仅为 4 类。调整优化全连接层参数,设置 Softmax 函数,求取各输入图像所属各类别的概率。



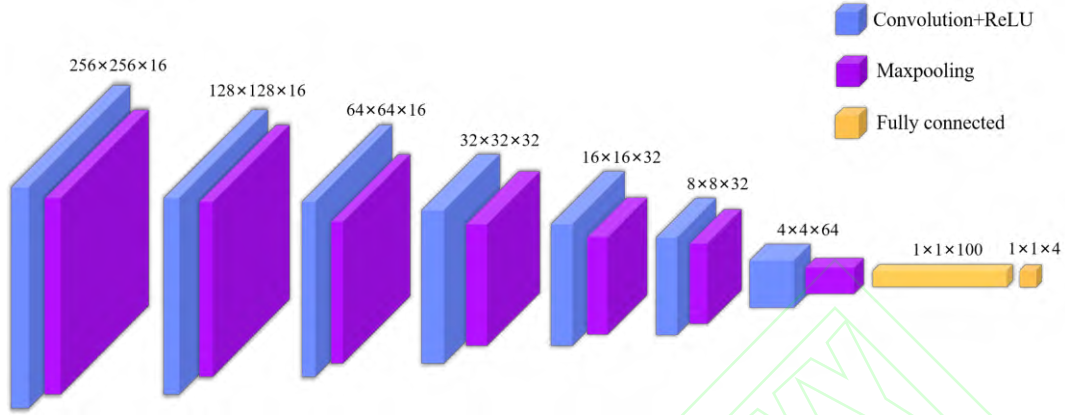


图6 VGG-A9 模型结构

Fig.6 VGG-A9 model structure

## 2.2 VGG-A9 模型优化

通过分析 VGG 模型架构,综合对裂缝分类的要求,优化 VGG-A9 模型,在模型中加入批量归一化层和 Dropout 层,加速模型训练,提高分类准确率。

### (1) 批量归一化

批量归一化(Batch Normalization, BN)<sup>[24]</sup>可使数据分布更为稳定,解决了数据分布不断变化带来的训练缓慢、学习率低等问题。将 BN 插入到卷积层和全连接层后,对数据归一化,批量归一化计算过程如下。

计算小批量均值:

$$\mu_{\beta} \leftarrow \frac{1}{m} \sum_{i=1}^m x_i \quad (2)$$

计算小批量方差:

$$\sigma_{\beta}^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_{\beta})^2 \quad (3)$$

归一化:

$$\bar{x}_i \leftarrow \frac{x_i - \mu_{\beta}}{\sqrt{\sigma_{\beta}^2 + \varepsilon}} \quad (4)$$

平移缩放:

$$y_i \leftarrow \gamma \bar{x}_i + \beta \quad (5)$$

其中,  $x_i$  为输入数据,  $y_i$  为输出数据,  $m$  为批次数量。  $\mu$  为均值,  $\sigma$  为方差,  $\gamma$  为拉伸参数,  $\beta$  为偏移

参数,初始均值设置为 0,方差为 1,具体数值通过模型训练时自动学习得到。

在大幅削减 VGG 模型参数后,网络深度变浅,特征提取层对输入分布变化更为敏感。为了稳定关键特征提取层的输出分布,在模型 Conv1(第一层卷积)、Conv4、Conv7 层后插入 BN 层,缓解因层数减少和参数减少导致的内部协变量偏移。此外,二值图像数据分布简单,BN 能更有效地确保中间层输入的均值和方差稳定,从而加速收敛,提升分类精度。

### (2) Dropout

Dropout<sup>[25]</sup>层可以有效缓解过拟合的发生,在一定程度上达到正则化效果。虽然 VGG-A9 模型已大幅简化,但 FC1 层相对于整个模型来说,仍然是参数量较大且易过拟合的部分。在 FC1 层后引入 Dropout 对模型进行优化,通过随机屏蔽神经元强制网络学习鲁棒的特征组合。由于裂缝形态在不同样本间存在显著差异,Dropout 能有效提升模型对不同形态裂缝的泛化能力,减少过拟合风险。

结合桥梁裂缝检测的需求,优化 VGG-A9 模型,以 VGG 模型为基础构建轻量化网络结构,搭建卷积层、池化层、全连接层。模型通过 Conv4、Conv7 层对通道数进行扩展,分别增至 32 和 64 通道,增强对复杂裂缝特征的表达能力。每个最大池化层执行 2 倍下采样操

作,逐步压缩空间维度同时保留显著特征。最终在 FC1 层前通过 Flatten 操作将特征图展平,实现空间特征到分类向量的转换。为增强模型稳定性,增加网络深度,在 Conv1、Conv4、Conv7 层后嵌入 BN 层,在 ReLU

激活函数后对数据归一化处理,缓解二值图像数据分布偏移问题。为防止模型出现过拟合,在 FC1 层后加入 Dropout 层,通过随机失活有效提升泛化能力。优化后的 VGG-A9 模型流程如图 7 所示。

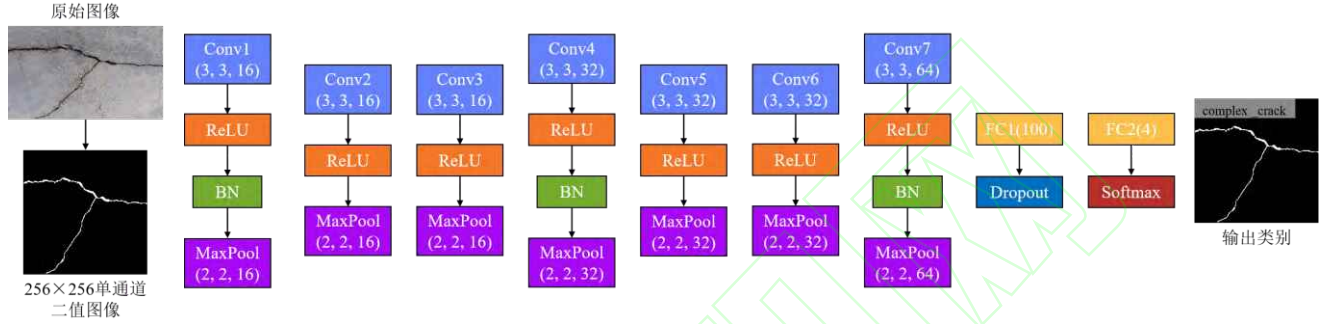


图 7 优化后的 VGG-A9 模型流程图

Fig.7 Optimized VGG-A9 model structure flowchart

### 3 实验结果与分析

#### 3.1 评估指标

本实验中采用三项指标评估模型性能: 准确率 (Accuracy)、精确率(Precision)、召回率(Recall), 计算公式如下所示。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

其中,  $TP$  是指预测为正, 真实为正的样本数量;  $FN$  是指预测为负, 真实为正的样本数量;  $FP$  是指预测为正, 真实为负的样本数量;  $TN$  是指预测为负, 真实为负的样本数量。

#### 3.2 离线训练与性能评估

VGG-A9 模型的训练与验证在高性能工作站完成, 硬件配置为 Intel i5-13600K CPU、NVIDIA RTX 4060 GPU 及 24GB 内存。基于 MATLAB R2021a 深度学习工具箱, 采用交叉熵损失函数与 Adam 优化器 (初始学

习率 0.001, Batch Size=10), 经过 400 次迭代后, VGG-A9 的训练集准确率达 99.17%, 损失值收敛至 0.066, 训练完成后, 在自建验证集上达到 98.53% 的分类准确率。

图 8 分别展示了 VGG-16 和 VGG-A9 在训练过程中损失值和准确率的变化情况。由图 8 可知, 相较于 VGG-16, VGG-A9 的损失值更低, 准确率更高, 且收敛速度显著加快, 训练稳定性更强。VGG-A9 的损失值在第 20 个 epoch 时已降低至 0.2 以下, 在第 67 个 epoch 完全收敛。而 VGG-16 震荡至 300 个 epoch 后才趋于平缓, 第 322 个 epoch 才完全收敛。这是因为 VGG-A9 的网络深度较浅, 避免了过深网络在小数据集上的梯度弥散, 其次嵌入的 BN 层有效抑制了二值图像的分布偏移, 使训练初期梯度传播更稳定。

为深入分析分类性能, 图 9 展示了验证集的混淆矩阵, 第五列中前四行的绿色数字为验证集精确率, 第五行中前四列的绿色数字为验证集召回率, 右下角数字 98.53% 为验证集准确率。

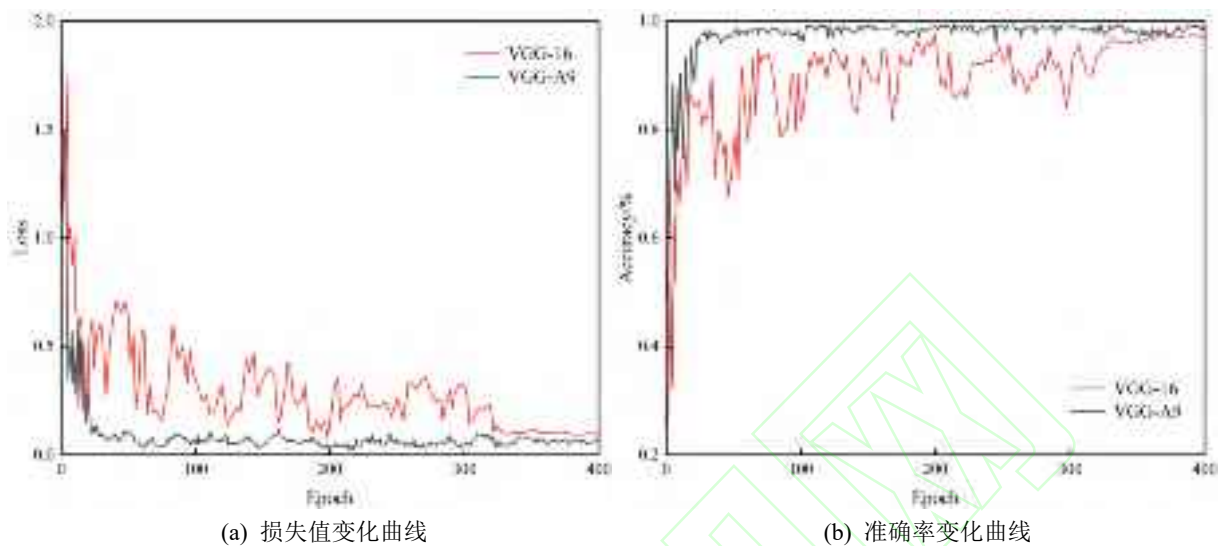


图 8 模型训练过程曲线

Fig.8 Model training process curve

预测类别	横向裂缝	107 19.67%	0 0.00%	0 0.00%	1 0.18%	99.07% 0.93%
	纵向裂缝	0 0.00%	78 14.34%	1 0.18%	0 0.00%	98.73% 1.27%
	斜向裂缝	0 0.00%	0 0.00%	164 30.15%	4 0.74%	97.62% 2.38%
	组合裂缝	1 0.18%	0 0.00%	1 0.18%	187 34.38%	98.94% 1.06%
		99.07% 0.93%	100.00% 0.00%	98.80% 1.20%	97.40% 3.60%	98.53% 1.47%
		实际类别				
		横向裂缝	纵向裂缝	斜向裂缝	组合裂缝	

图 9 验证集混淆矩阵

Fig.9 Verification set confusion matrix

选取经典网络 AlexNet<sup>[26]</sup>、VGG-16、InceptionV3<sup>[27]</sup>、ResNet50<sup>[28]</sup>、MobileNetV2<sup>[29]</sup>与 VGG-A9 模型进行对比，训练中设置超参数与 VGG-A9 模型相同，使用验证集评估性能，对比结果如表 2 所示。从参数量来看，VGG-A9 模型参数大小仅为  $0.74 \times 10^6$ ，显著低于传统模型，并且准确率、精确率和召回率均高于其它模型。结果表明，VGG-A9 通过简化网络结构、引入批量归一化和 Dropout 层优化，在保证高分类精度的同时，压缩模型

大小。这种轻量级设计使其在无人机嵌入式系统等资源受限场景中具有显著优势，能够实现实时、低功耗、高精度的裂缝分类检测。

表 2 模型性能对比

Table 2 Model performance comparison				
模型	准确率/%	精确率/%	召回率/%	参数量/ $10^6$
AlexNet	89.15	88.55	89.88	61.91
VGG-16	96.88	96.59	96.72	138.12
InceptionV3	97.43	97.28	97.31	23.86
ResNet50	97.06	96.95	97.00	25.68
Mo- bileNetV2	95.40	95.12	95.26	3.53
VGG-A9	98.53	98.82	98.59	0.74

3.3 机载实时验证

为验证模型在资源受限场景的实用性，将训练后的 VGG-A9 部署至无人机嵌入式平台 NVIDIA Jetson Xavier NX。输入验证集图像，单帧分类耗时 21ms。在 544 张验证集样本中，机载分类准确率达 98.01%，较工作站环境仅下降 0.52%，峰值内存占用 517MB，平均功耗 9.2W。

为进一步评估 VGG-A9 在边缘计算平台上的性能优势，选取通用轻量基准模型 ShuffleNetV2<sup>[30]</sup>、MobileNetV2 及先进的裂缝检测分类模型 PCCNet、

KAN-MobileViT, 在嵌入式平台 NVIDIA Jetson Xavier NX 上使用相同验证集进行对比实验, 结果如表 3 所示。

表 3 轻量模型在嵌入式平台的性能对比

Table 3 Performance comparison of lightweight models on embedded platforms

模型	准确率 /%	推理延 迟/ms	峰值内 存/MB	平均功 耗/W	参数量 /10 <sup>6</sup>
ShuffleNetV2	97.24	32	593	10.7	5.15
MobileNetV2	95.77	34	610	11.3	3.53
PCCNet	97.06	26	554	10.1	1.02
KAN-MobileViT	93.38	59	758	14.4	7.45
VGG-A9	<b>98.01</b>	<b>21</b>	<b>517</b>	<b>9.2</b>	<b>0.74</b>

分析表 3 可知, VGG-A9 在保持最低参数量的同时, 实现了最高的分类准确率和最低的推理延迟, 在内存占用和功耗方面表现也是最优。相较于裂缝专用的次优模型 PCCNet, VGG-A9 的准确率提高 0.95%, 推理速度加快 5ms, 内存占用减少 37MB, 功耗降低 0.9W。KAN-MobileViT 由于 Transformer 架构的边缘适配缺陷, 推理延迟达 59ms, 平均功耗 14.4W, 综合表现最差。实验表明, VGG-A9 凭借针对裂缝检测分类任务的轻量化设计, 实现了精度、速度和功耗之间的平衡, 在无人机桥梁底部裂缝检测分类任务中具有显著优势。

随机选取 50 张验证集样本对 VGG-A9 进行实时测试, 部分分类结果如图 10 所示。图中, horizontal\_crack 为横向裂缝、vertical\_crack 为纵向裂缝、diagonal\_crack 为斜向裂缝、complex\_crack 为组合裂缝。

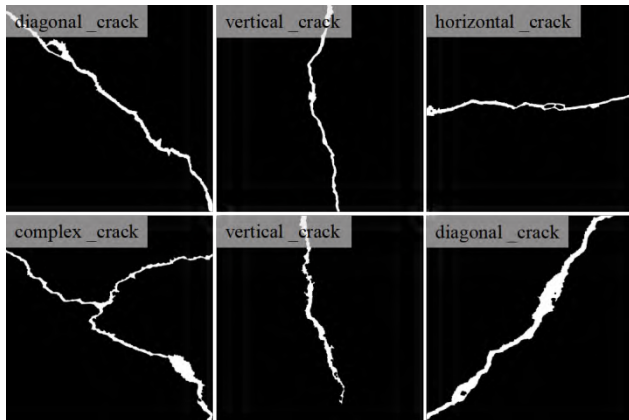


图 10 模型对裂缝分类结果

Fig.10 Model classification results of cracks

### 3.4 消融实验

为验证 BN 层和 Dropout 组件的有效性, 将无 BN 和 Dropout 层的 9 层 VGG 模型定义为 Baseline, 在离线工作站和嵌入式平台的环境下进行消融实验, 实验结果如表 4 所示。

表 4 消融实验结果

Table 4 Results of ablation experiment

模型	离线工作站			嵌入式平台		
	准确率/%	训练收敛 epoch	泛化间隙 /%	准确率/%	部署准确率损失/%	推理延迟 /ms
Baseline	94.12	142	2.31	92.28	1.84	<b>19</b>
Baseline+BN	97.43	72	2.02	95.40	2.03	21
Baseline+Dropout	95.04	125	0.82	94.30	<b>0.74</b>	20
VGG-A9	<b>98.53</b>	<b>67</b>	<b>0.64</b>	<b>98.01</b>	<b>0.74</b>	21

如表 4 所示, 泛化间隙为训练集和验证集准确率差值, 用于衡量模型过拟合程度; 部署准确率损失为离线工作站环境与嵌入式平台环境准确率差值, 用于衡量模型部署鲁棒性。实验结果表明: 插入 BN 层后, 模型离线准确率提升 3.31%、部署准确率提升 3.12%, 且训练收敛加速 49%, 证明 BN 层通过稳定特征分布显著增强学习效率。Dropout 降低泛化间隙 1.49%, 并实现最低部署准确率损失, 有效抑制参数过拟合并增强在边缘场景的泛化能力。相较于 Baseline, 融合 BN 和 Dropout 的 VGG-A9 模型仅增加 2ms 推理延迟, 达成最高离线准确率与部署准确率, 同时泛化间隙与部署损失均显著降低, 充分验证了 BN 与 Dropout 组件在提升模型精度、加速收敛及增强部署鲁棒性方面的有效性, 二者协同作用实现了轻量化设计的最优综合性能。

### 3.5 鲁棒性测试

为了验证 VGG-A9 具有较强的鲁棒性, 在高性能工作站环境下对 VGG-A9 进行跨数据集鲁棒性测试。选用公开道路裂缝数据集 EdmCrack600<sup>[31]</sup>, 该数据集是用安装在车上的摄像头在汽车移动时连续拍摄的, 裂缝形态相较于本文构建的数据集更加杂乱无序, 对分类



任务来说是一个很大的挑战。EdmCrack600 包含 600 张道路裂缝图像,对裂缝图像进行手动分类,得到横向裂缝 164 张,纵向裂缝 55 张,斜向裂缝 113 张,组合裂缝 268 张。

由于 EdmCrack600 已给出裂缝相应标注完成的二值图像,因此无需再对裂缝图像进行二值化处理。将二值图像送入模型完成分类,不同模型的部分分类结果对比如图 11 所示,与 AlexNet、VGG-16、InceptionV3、ResNet50、ShuffleNetV2、MobileNetV2、PCCNet、KAN-MobileViT 分类精度的对比结果如表 5 所示。

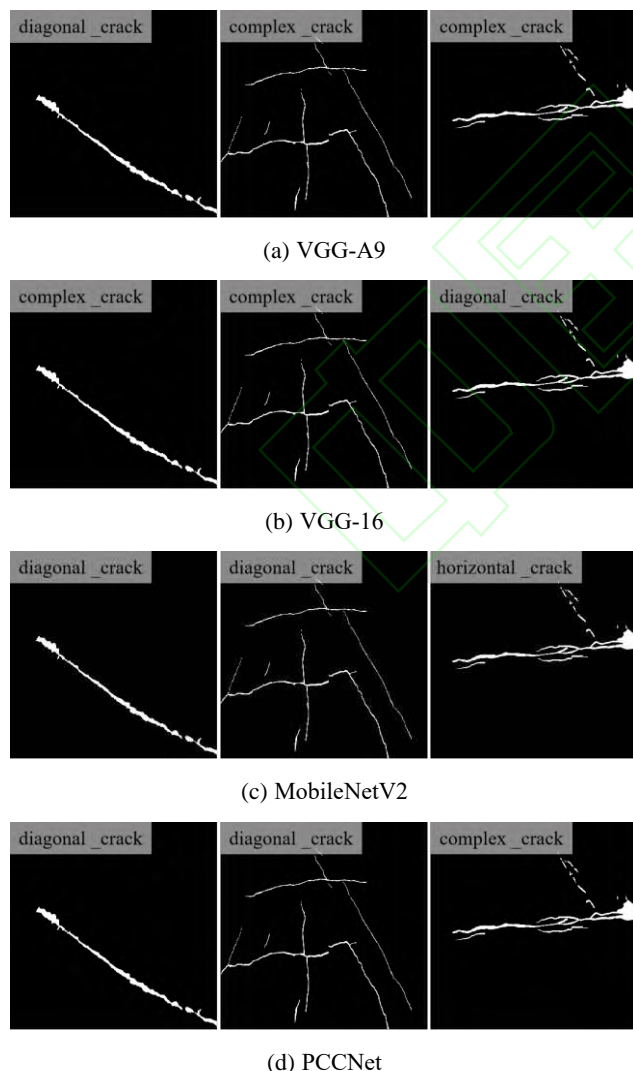


图 11 鲁棒性实验中不同模型分类结果对比

Fig.11 Comparison of classification results of different models in robustness experiments

图 11 中展示的一张裂缝图像分别为斜向裂缝、组合裂缝、组合裂缝, VGG-A9 对三张图像的分类结果正确, VGG-16 将第一和第三张图像误分类为组合裂缝和斜向裂缝, MobileNetV2 将第二和第三张图像误分类为斜向裂缝和横向裂缝, PCCNet 将第二张图像误分类为斜向裂缝。

表 5 鲁棒性实验模型性能对比

Table 5 Comparison of model performance in robustness experiments

模型	准确率/%	精确率/%	召回率/%
AlexNet	83.00	80.78	82.36
VGG-16	89.83	90.16	90.51
InceptionV3	91.83	92.23	92.01
ResNet50	91.50	91.07	90.60
ShuffleNetV2	91.83	91.33	90.67
MobileNetV2	90.50	90.31	89.93
PCCNet	93.17	<b>93.00</b>	92.83
KAN-MobileViT	86.34	85.67	85.84
VGG-A9	<b>93.67</b>	92.88	<b>93.18</b>

通过分析表 5 得知,在公开数据集 EdmCrack600 的鲁棒性测试中, VGG-A9 以 93.67% 的准确率和 93.18% 的召回率领先所有对比模型,精确率 92.88% 仅略低于裂缝检测分类模型 PCCNet 0.12 个百分点。实验证明, VGG-A9 在保持轻量化的同时,展现出优异的跨场景泛化能力,具有较强的鲁棒性。

## 4 结束语

针对桥梁底部裂缝检测的隐蔽性挑战与无人机平台的计算资源限制,提出一种结合无人机拍摄系统与轻量级 VGG-A9 模型的协同解决方案。通过无人机图像采集系统与语义分割驱动的二值化策略,构建了包含 2720 张裂缝图像的数据集。提出的轻量级模型 VGG-A9 仅有 9 层,通过削减全连接层参数与嵌入 BN、Dropout 优化策略,将参数模型压缩至  $0.74 \times 10^6$ 。实验验证表明:模型在自建数据集上分类准确率达 98.53%,机载部署实现了实时(21ms/帧)、低功耗(9.2W)、

高精度 (98.01%) 的分类检测; 在公开数据集 EdmCrack600 的鲁棒性测试中准确率达 93.67%, 优于其它对比模型。该方法显著降低了计算复杂度并具备较强的泛化能力, 为桥梁智能化巡检提供了可行的技术路径。

本研究模型的输入图像为单通道二值图像, 识别效果依赖二值化预处理质量, 在极端光照或复杂背景场景下性能可能受限。未来工作将基于本架构直接处理原始 RGB 图像并融合多尺度形态特征以提升环境适应性, 同时推进边缘计算硬件的算子编译优化与分布式协同检测框架开发, 构建适应复杂场景、低延迟且可扩展的智能裂缝检测系统。

## 参考文献:

- [1] 张劲泉, 晋杰, 汪云峰, 等. 公路桥梁智能检测技术与装备研究进展[J]. 公路交通科技, 2023, 40(1): 1-27+58.
- [2] ZHANG J Q, JIN J, WANG Y F, et al. Study progress of intelligent inspection technology and equipment for highway bridge[J]. Journal of Highway and Transportation Research and Development, 2023, 40(1): 1-27+58.
- [3] LI G, REN X L, QIAO W T, et al. Automatic bridge crack identification from concrete surface using ResNeXt with postprocessing[J]. Structural Control & Health Monitoring, 2020, 27(11): e2620.
- [4] AI D H, JIANG G Y J, LAM S K, et al. Computer vision framework for crack detection of civil infrastructure-A review[J]. Engineering Applications of Artificial Intelligence, 2023, 117(PA): 105478.
- [5] WANG F, ZOU Y, CASTILLO E D, et al. Automated UAV path-planning for high-quality photogrammetric 3D bridge reconstruction[J]. Structure and Infrastructure Engineering, 2024, 20(10): 1595-614.
- [6] WANG F, ZOU Y, CHEN X Y, et al. Rapid in-flight image quality check for UAV-enabled bridge inspection[J]. Isprs Journal of Photogrammetry and Remote Sensing, 2024, 212: 230-50.
- [7] XU T, WU Y P, QIN Y, et al. Automatic steel girder inspection system for high-speed railway bridge using hybrid learning framework[J]. Computer-Aided Civil and Infrastructure Engineering, 2025, 40(11): 1508-27.
- [8] LEE J H, YOON S S, KIM I H, et al. Diagnosis of crack damage on structures based on image processing techniques and R-CNN using unmanned aerial vehicle (UAV)[C]//Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018, USA, March 5-8, 2018. Bellingham: SPIE, 2018: UNSP 1059811.
- [9] LI J P, LI X D, LIU K, et al. Crack identification for bridge structures using an unmanned aerial vehicle (UAV) incorporating image geometric correction[J]. Buildings, 2022, 12(11): 1869.
- [10] SHU J P, XIA Z, GAO Y F. BIM-Based Trajectory Planning for Unmanned Aerial Vehicle-Enabled Box Girder Bridge Inspection[J]. Remote Sensing, 2025, 17(4): 682.
- [11] XU W Y, CUI C, JI Y C, et al. A UAV path planning algorithm for bridge construction safety inspection in complex terrain[J]. Scientific Reports, 2025, 15(1): 13564.
- [12] SHANG Z X, BRADLEY J, SHEN Z G. A co-optimal coverage path planning method for aerial scanning of complex structures[J]. Expert Systems With Applications, 2020, 158: 113535.
- [13] DING H, JIANG X H, LI K, et al. Intelligent classification method for tunnel lining cracks based on PFC-BP neural network[J]. Mathematical Problems in Engineering, 2020, 2020: 8838216.
- [14] 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.
- [15] SUN Z Y, MA Z D, LI W, et al. Pavement crack identification method based on deep convolutional neural network fusion model[J]. Journal of Chang'an University(Natural Science Edition), 2020, 40(4): 1-13.
- [16] 王纪武, 鱼鹏飞, 罗海保. 基于改进 Faster R-CNN+ZF 模型的铁路桥梁裂缝分类方法[J]. 北京交通大学学报, 2020, 44(1): 106-112.
- [17] WANG J W, YU P F, LUO H B. Classification method of railway bridge cracks based on improved Faster R-CNN+ZF model[J]. Journal of Beijing Jiaotong University, 2020, 44(1): 106-112.
- [18] HE Z Y, XU W. Deep learning and image preprocessing-based crack repair trace and secondary crack classification detection method for concrete bridges[J/OL]. Structure and Infrastructure Engineering. <http://doi.org/10.1080/15732479.2024.2330702>.
- [19] LIANG J, ZHANG Q P, GU X Y. Small-sample data-driven lightweight convolutional neural network for asphalt pavement defect identification[J]. Case Studies in Construction Materials, 2024, 21: e03643.
- [20] SHI T S, LUO H. Deep learning for automated detection and classification of crack severity level in concrete structures[J]. Construction and Building Materials, 2025, 472: 140793.
- [21] 刘宇飞, 冯楚乔, 陈伟乐, 等. 基于机器视觉法的桥梁表观病害检测研究综述[J]. 中国公路学报, 2024, 37(02): 1-15.
- [22] LIU Y F, FENG C Q, CHEN W L, et al. Review of bridge apparent defect inspection based on machine vision[J]. China Journal of Highway and Transport, 2024, 37(02): 1-15.
- [23] 江晨轶, 潘家伟, 张丽军, 等. UAV SfM 技术在活动构造研究中的应用——以青藏高原西北部龙木错断裂为例[J]. 地质力学学报, 2024, 30(2): 332-347.
- [24] JIANG C Y, PAN J W, ZHANG L J, et al. Application of UAV SfM technology in active tectonic research: A case study of the Longmu Co Fault, Northwestern Qinghai-Tibet Plateau[J]. Journal of Geomechanics, 2024, 30(2): 332-347.
- [25] 公路桥梁技术状况评定标准: JTG/T H21-2011[S]. 北京: 人民交通出版社, 2012.
- [26] Highway performance assessment standard: JTG/T H21-2011[S]. Beijing: China Communications Press, 2012.

- [21] 公路桥涵养护规范: JTG 5120 – 2021[S].北京: 人民交通出版社, 2021.  
Specifications for maintenance of highway bridges and culverts: JTG 5120 – 2021[S].Beijing: China Communications Press, 2021.
- [22] YU Z Y, DAI C Q, ZENG X M, et al. A lightweight semantic segmentation method for concrete bridge surface diseases based on improved DeeplabV3+[J]. Scientific Reports, 2025, 15(1): 10348.
- [23] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv:1409.1556v6, 2015.
- [24] 杨寒雨,赵晓永,王磊.数据归一化方法综述[J].计算机工程与应用, 2023, 59(3): 13-22.  
YANG H Y, ZHAO X Y, WANG L. Review of Data Normalization Methods[J].Computer Engineering and Applications, 2023, 59(3): 13-22.
- [25] ZHAO X, WANG L M, ZHANG Y F, et al. A review of convolutional neural networks in computer vision[J]. Artificial Intelligence Review, 2024, 57(4): 99.
- [26] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the 26th International Conference on Neural Information Processing Systems (NIPS 2012), USA, December 3-6, 2012. New York: Curran Associates, 2012: 1097-1105.
- [27] 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), USA, June 27-30, 2016. Piscataway, NJ: IEEE, 2016: 2818-2826.
- [28] HE K M, ZHANG X Y, REN S Q, et al. Deep Residual Learning for Image Recognition[J].arXiv:1512.03385, 2015.
- [29] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[J]. arXiv:1801.04381v4, 2019.
- [30] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design[C]//Proceedings of the 15th European Conference on Computer Vision(ECCV), GERMANY, September 08-14, 2018. Berlin: Springer, 2018, 11218: 122-138.
- [31] MEI Q, GUL M, AZIM R M. Densely connected deep neural network considering connectivity of pixels for automatic crack detection[J]. Automation in Construction, 2020, 110: 103018-103018.