

# 改进 U-Net 的无人机航拍路面破损检测方法<sup>\*</sup>

韩建峰<sup>1,2</sup> 张 静<sup>1,2</sup> 宋丽丽<sup>1,2</sup> 陶永昭<sup>1,2</sup>

(1. 内蒙古工业大学信息工程学院 呼和浩特 010080;  
2. 内蒙古自治区感知技术与智能系统重点实验室 呼和浩特 010051)

**摘要:**为了探究无人机航拍视角下细小裂缝难以检测和检测出现断裂的问题,基于 U-Net 架构提出了 ASE-Net 网络。首先,使用改进的 VGG-16 作为编码器,以便提取破损的特征信息;其次在最小尺度网络层引入多尺度特征融合 (MSFF) 模块、通道增强条形池化 (CESP) 模块;最后在解码阶段添加 ECA\_X 注意力模块。实验结果表明,本文模型在自建的无人机航拍路面破损数据集上的 mIoU、mPA、mPrecision 分别达到 0.820 9、0.930 2、0.865 1,相比于基线 U-Net 分别提高 15.97%、12.72%、11.02%。最后,在开源数据集 Crack500 中验证了本文模型相比于其他主流模型具有更强的性能和泛化能力。模型能实现对路面细小裂缝、坑槽、修补的准确检测,有效解决了裂缝检测的断裂问题,可提升大尺寸航拍图像路面破损检测效果。

**关键词:**无人机;破损检测;语义分割;条形池化;多尺度特征;U-Net

中图分类号: TP391; TN911.73 文献标识码: A 国家标准学科分类代码: 520.20

## Improved pavement damage detection method of UAV based on U-Net

Han Jianfeng<sup>1,2</sup> Zhang Jing<sup>1,2</sup> Song Lili<sup>1,2</sup> Tao Yongzhao<sup>1,2</sup>

(1. School of Information Engineering, Inner Mongolia University of Technology, Hohhot 010080, China;  
2. Inner Mongolia Key Laboratory of Perceptual Technology and Intelligent Systems, Hohhot 010051, China)

**Abstract:** In order to explore the difficulties in detecting fine cracks and the occurrence of breaks in the aerial view of UAVs, a network called ASE-Net is proposed based on the U-Net architecture. First, an improved VGG-16 is used as the encoder to extract the broken feature information. Second, multi-scale feature fusion block (MSFF) module and channel enhanced strip pooling (CESP) module are introduced at the minimum scale network layer. Finally, the ECA\_X attention module is added to the decoding stage. The experimental results indicate that the model presented in this paper achieves a mIoU of 0.820 9, a mPA of 0.930 2, and a mPrecision of 0.865 1 on the self-constructed UAV aerial pavement breakage dataset. These results represent improvements of 15.97%, 12.72%, and 11.02% over the baseline U-Net, respectively. Ultimately, the model in this work has been demonstrated to exhibit better performance and generalization ability than other standard models utilizing the open-source dataset Crack500. The model can realize accurate detection of small cracks, potholes, and repairs on the road surface, effectively solving the fracture problem of crack detection, and enhancing the effect of pavement damage detection in large-size aerial images.

**Keywords:** unmanned aerial vehicle; damage detection; semantic segmentation; strip pooling; multi-scale feature; U-Net

## 0 引言

近年来,交通运输行业蓬勃发展,路面裂缝作为病害的常见类型之一,对道路安全具有潜在的威胁<sup>[1]</sup>。路面裂缝会降低道路的使用寿命,影响人们出行安全<sup>[2]</sup>。然而,传统的人工巡检方法主观性强、效率低下、危险系数高。无人机在飞行过程中不会影响正常交通运行,能够完成各种环境

下的路面高分辨率采集任务,从而提高巡检效率。

赵芳等<sup>[3]</sup>将多尺度形态学和双边滤波结合的去噪方法引入 Canny 算法中,不仅加强了降噪效果,还能保留更多的边缘特征。Talab 等<sup>[4]</sup>将自适应阈值和多重滤波相结合,相比于其他传统方法能清晰准确地检测出图像中的裂缝。上述方法对于路面类型、光照条件、复杂背景下的噪声比较敏感,容易出现误检、结果不稳定等问题。

随着深度学习方法的出现,卷积神经网络(convolutional neural network,CNN)逐步应用到路面破损检测中。Liang 等<sup>[5]</sup>提出一种基于双卷积神经网络的裂缝识别方法,利用 CNN 排除干扰因素,通过全卷积网络(fully convolutional network,FCN)以逐层合并的结构进行裂缝分割。Li 等<sup>[6]</sup>将 CNN 架构与 K-means 聚类算法融合,推动了自动化路面裂缝检测技术的进步。Ibragimov 等<sup>[7]</sup>用更快的基于区域的卷积神经网络(faster region-based convolutional neural network,Faster R-CNN)进行破损检测,通过调整滑动窗口的大小,实现大尺寸路面破损图像检测。为了进一步提高路面破损检测性能,张华卫等<sup>[8]</sup>用多尺度卷积和注意力机制融合裂缝场景的特征来提高检测精度。然而,这些方法不能将高低阶的特征信息有效融合,特征流失严重,路面破损检测精度不高。

为了将特征信息进行多尺度特征融合,提高分割精度。Ronneberger 等<sup>[9]</sup>基于 CNN 提出了 U-Net 模型,并成功应用于医学分割领域。Liu 等<sup>[10]</sup>将 U-Net 引入裂缝检测,拥有少量数据集的情况下仍能获得较高的精度。Lau 等<sup>[11]</sup>用 ResNet34 作为 U-Net 的编码器,引入残差网络避免梯度消失和网格退化。由于网络层数较少,模型无法提取足够复杂的特征,导致分割结果的精度不高。为了提高分割精度,张伟光等<sup>[12]</sup>利用 VGG-16 作为 U-Net 的编码器,模型具有良好的特征提取和迁移学习能力,但对于狭长细小的裂缝效果欠佳。为了增强对细小裂缝的关注度,王瑞绅等<sup>[13]</sup>在 U-Net 中引入高效通道注意力(efficient channel attention,ECA)机制,改进的网络能自适应提取裂缝特征并抑制背景干扰。夏晓华等<sup>[14]</sup>在空洞空间金字塔池化(atrous spatial pyramid pooling,ASPP)模块中使用条形池化模块代替全局平均池化(global average pooling,GAP),有效捕获裂缝结构的上下文信息,避免无关区域噪声的干扰。Li 等<sup>[15]</sup>利用空间位置自注意力模块和金字塔池化模块获得丰富的上下文信息,改进 PSPNet 算法能准确地提取细小裂缝。张大伟等<sup>[16]</sup>对特征提取层的 4 个不同分辨率表征分别融合卷积注意力模块,提出的 E-HRNet 网络有效避免了漏检、误检等问题。为了缓解样本不均衡的问题,范瑶瑶等<sup>[17]</sup>使用加权 Dice 和二元交叉熵损失(binary cross entropy,BCE)结合的损失函数提高分割精度。然而,上述方法数据集图像以近距采集工具为主,如车载相机、工业相机等,图像采集范围较小,巡检效率低。

近年来,为了提高路面巡检效率,无人机影像裂缝提取领域取得了一些进展。武广臣等<sup>[18]</sup>用 GrabCut 算子提取前景路面信息,结合边缘检测和双阈值轮廓方法探测路面裂缝,排除了大量伪边缘信息和噪声干扰。卢小平等<sup>[19]</sup>提出一种基于无人机影像的迭代二值化道路裂缝检测方法,该方法正确率达到 86.6%,在复杂场景下有较高鲁棒性。He 等<sup>[20]</sup>将目标检测算法与无人机相结合,算法在自建的无人机近远场景图像数据集上具有精度和速度双重优势。

Hong 等<sup>[21]</sup>基于 U-Net 架构引入了卷积块注意力、将改进的编码器和长短跳跃连接相融合,所提方法能有效预测无人机图像中的高速公路裂缝,平均交并比为 77.47%,准确率达到了 68.38%。

上述方法的检测目标均为裂缝,缺少其他常见类型的路面破损研究。此外,由于裂缝像素占比小,细微裂缝通常难以识别,由于检测精度不高导致出现断裂现象。本文基于无人机航拍的方式采集路面破损图像,提高巡检效率,同时,针对路面破损分割算法存在的问题,提出一种改进模型。主要贡献如下:

1) 采用去掉全连接层和第五个池化层的 VGG-16 作为主干网络,使用迁移学习策略,有效提取图像特征。同时,将激活函数改为 GELU。

2) 在网络层最底部引入 ASPP 模块,为了提取更多的特征,采用密集连接,增加网络深度,有效捕获裂缝的上下文信息的同时抑制背景信息。

3) 用 CESP 模块替换 ASPP 模块中的全局平均池化, CESP 模块在条形池化基础上加强通道相关性,有效捕获裂缝细长特征的同时提高检测能力。

4) 解码部分添加 ECA\_X 注意力机制,动态分配权重,内部并联  $3 \times 3$  标准卷积,对输入特征进行二次提取,增强特征提取能力。

## 1 无人机航拍路面破损语义分割 ASE-Net 网络

目前,U-Net 仍是语义分割网络普遍采用的结构<sup>[22-23]</sup>。网络主要由压缩路径、扩展路径和跳跃连接组成,网络结构如图 1 所示。压缩路径包括卷积层和池化层,用于生成小分辨率但压缩的高维语义表示的特征图。扩展路径包含转置卷积层和卷积层,对特征图进行上采样,解码压缩路径特征,逐层还原图像尺寸。由于上采样过程中存在学习参数不稳定和棋盘格效应问题,因此采用双线性插值法替换转置卷积,还原路面破损边缘信息,实现对破损的像素级分割。跳跃连接将同尺度特征图拼接在一起,避免特征流失,像素定位更准确。最后用  $1 \times 1$  卷积核将特征映射转换生

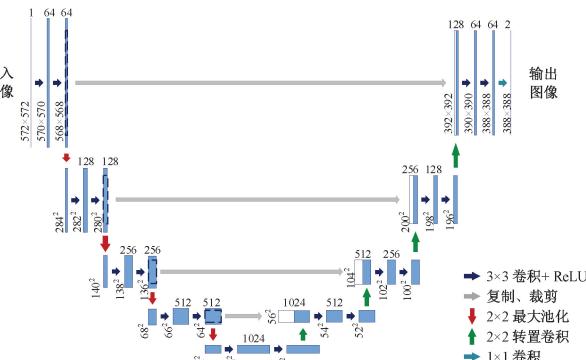


图 1 U-Net 结构

Fig. 1 U-Net structure

成分割映射,实现语义分割。由于编解码器使用的卷积层模式有所不同,网络的通用性受到了限制。

为实现路面破损的高精度检测,本文提出了 ASE-Net

网络,主要由 3 部分构成:基于改进 VGG-16 的编码器;引入 MSFF 和 CESP 模块的最小尺度网络层和结合 ECA\_X 的解码器,网络结构如图 2 所示。

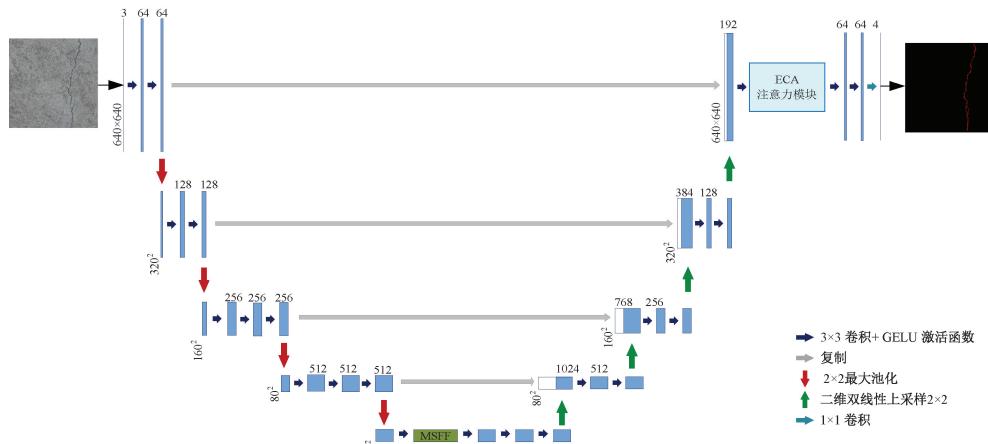


图 2 ASE-Net 结构

Fig. 2 ASE-Net structure

### 1.1 改进 VGG-16 编码器

U-Net 编码器网络层次较浅,特征学习能力有限<sup>[24]</sup>。为了提取更多的有效信息,可以增加网络层数。对于传统的 U-Net,简单的加深网络反而会出现梯度消失问题。同时,经过卷积和池化操作,特征图的分辨率变小,会丢失更多全局信息。VGGNet(visual geometry group)作为一种经典网络模型,能够将预训练好的模型作为特征提取器,利用提取的特征对新任务进行训练,加速模型收敛。同时,VGG-16 部分网络结构与 U-Net 编码阶段的网络结构极为相似,因此将去除全连接层和第五个池化层的 VGG-16 作为主干网络。

VGG-16 模型中卷积神经网络层数多,在反向传播过程中,网络的权重梯度逐渐衰减,较前层的权重几乎无法更新,训练过程中可能会出现梯度消失问题。为了解决该问题,对 VGG-16 进行改进。激活函数由 ReLU 改为 GELU,反向传播时梯度变化更连续,有助于加速收敛。输入为负时,GELU 输出不全为 0,部分神经元不会永远关闭。改进后的网络结构如图 3 所示。

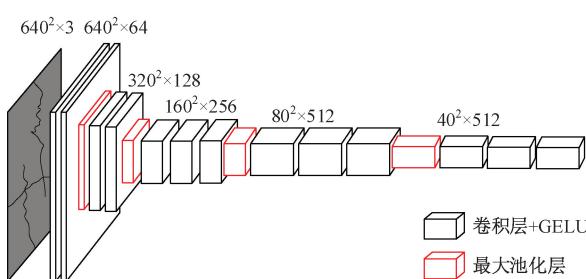


图 3 改进的 VGG-16 结构

Fig. 3 Improved VGG-16 structure

### 1.2 MSFF 模块

为解决裂缝分割过程中细小裂缝不易识别的问题,引入 ASPP<sup>[25]</sup> 进行特征融合。ASPP 是空间金字塔池化(spatial pyramid pooling, SPP)<sup>[26]</sup> 和空洞卷积的结合。后者在标准卷积的基础上引入了膨胀率,相邻像素点之间存在信息冗余,因此,扩大感受野的同时又能捕获多尺度信息。但由于邻近的像素是从相互独立的子集中卷积得到的,远距离相关性不足。

本文基于 ASPP 提出 MSFF 模块,MSFF 位于模型的底部,结构如图 4 所示。由于 CESP 模块擅长捕捉方向性强、细长的特征,因此将 GAP 替换为 CESP;同时,增加网络深度,利用密集连接减少梯度消失。经过空洞卷积后,特征图大小的计算如下:

$$o = \frac{i + 2p - k - (k-1)(d-1)}{s} + 1 \quad (1)$$

式中:  $o$  为输出特征图的长或宽,  $i$  为输入特征图的长或宽,  $p$  为卷积过程中补 0 的个数,  $k$  为卷积核大小,  $d$  为扩张率,  $s$  为步长。

### 1.3 CESP 模块

池化是特征提取的重要步骤,传统池化用方形池化核处理输入特征图,由于核形状单一,结果包含大量背景区域,会造成特征污染。图 5 为池化对比图。如图 5(a)所示,传统池化混淆了裂缝与背景的信息,特征结果不准确。由图 5(b)可知,条形池化<sup>[27]</sup> 沿着水平或垂直方向生成两种池化核,分别执行池化操作,凭借池化核窄、长的特点,既能有效利用长距离关系,又能消除噪声,从而分割出道路表面细小裂缝。水平条形池化和垂直条形池化计算公式如式(2)、(3)。

$$y_i^h = \frac{1}{W} \sum_{0 \leq j < W} x_{i,j} \quad (2)$$

$$y_j^v = \frac{1}{H} \sum_{0 \leq i < H} x_{i,j} \quad (3)$$

式中:  $W$  和  $H$  分别为宽度和高度;  $y_i^h$  为水平条形池化在

第  $i$  行的结果;  $y_j^v$  为垂直条形池化在第  $j$  列的结果;  $x_{i,j}$  为第  $i$  行  $j$  列元素。

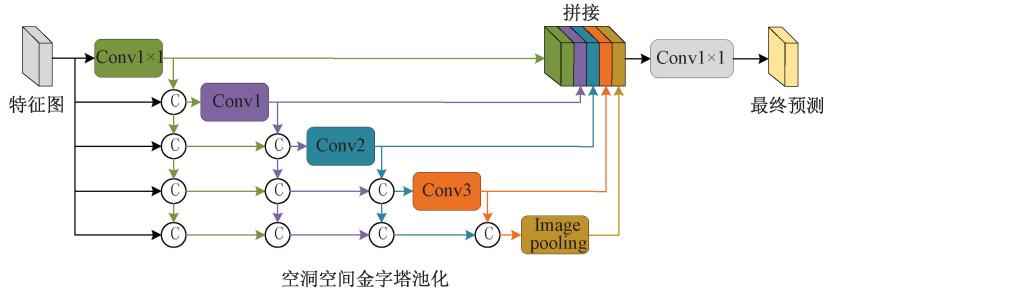


图 4 MSFF 模块

Fig. 4 MSFF module

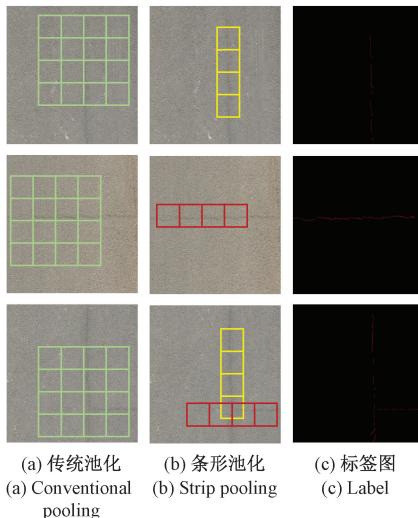


图 5 池化对比图

Fig. 5 Pooling comparison

尽管条形池化擅长处理细小特征,但通道间的联系不紧密,处理具有通道高度相关的特征图时效果欠佳。本文引入 CESP 模块,替换掉 ASPP 模块中 GAP,有效捕获裂缝的上下文信息,防止无关区域干扰预测, CESP 模块如图 6 所示。

CESP 模块通过 GAP 获取通道间的相关性,最后与经过条形池化输出的特征图进行融合,克服局限性的同时提高模型检测能力,裂缝细节保留更完整,识别结果更加符合实际。

#### 1.4 ECA\_X 模块

路面裂缝包括各种不规则形状。随着网络层数加深,每次上采样都会丢失细节信息。跳跃连接也只将特征图进行拼接,并未涉及深层次的融合。本文在最后一次特征图拼接后引入 ECA 模块。ECA<sup>[28]</sup>相比于 SE 机制可避免降维,对跨通道交互的信息比较敏感,增加少数参数量时能带来明显性能提升。

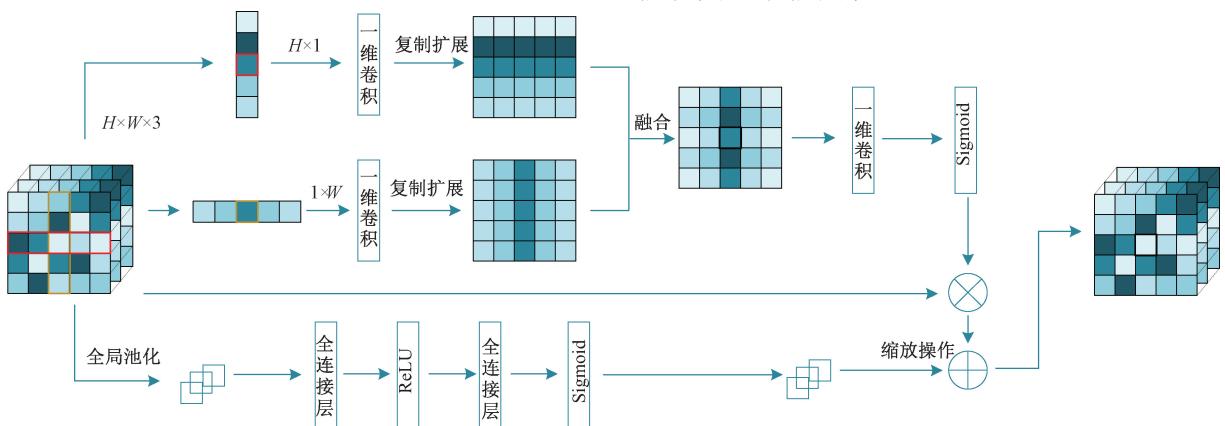


图 6 CESP 模块

Fig. 6 CESP module

路面裂缝大多细小且复杂,分割效果往往不尽人意。为了提取到所有细节特征,提出了 ECA\_X 结构。在 ECA

内部并联  $3 \times 3$  卷积,特征在通道上交互的同时,  $3 \times 3$  卷积对输入特征进行二次提取,最后两个特征结果相乘输出,提

高特征表达能力。ECA\_X 结构如图 7 所示。通过卷积核为  $k$  的快速一维卷积学习通道间的相关性,根据式(4)、(5)动态调整通道权重,保留重要特征。

$$C = 2^{\gamma k - \beta} \quad (4)$$

$$k = \psi(C) = \left( \frac{\log_2 C + \beta}{\gamma} \right)_{odd} \quad (5)$$

式中:  $\left( \frac{\log_2 C + \beta}{\gamma} \right)_{odd}$  为距  $\frac{\log_2 C + \beta}{\gamma}$  最近的奇数  $t$ ,  $k$  为卷积核大小,  $\sigma$  为通道数,  $\gamma, \beta$  为函数指数组线性映射系数与常数项。ECA\_X 的运算表达式如式(6)所示。

$$ECA(x) = \sigma(Conv_{1d}^{1 \times 1}(GAP(x))) \times Conv_{2d}^{3 \times 3}(x) \quad (6)$$

式中:  $\sigma$  为 Sigmoid 激活函数, GAP 为全局平均池化。

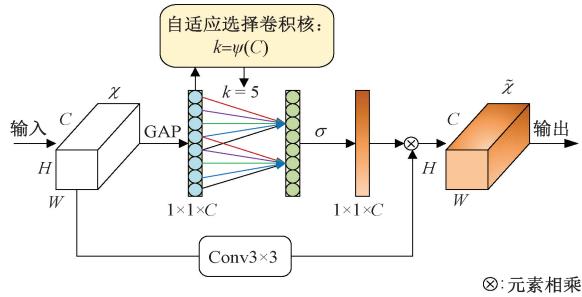


图 7 ECA\_X 结构

Fig. 7 ECA\_X structure

## 1.5 损失函数

现阶段, BCE 损失函数被广泛用于像素分割任务, 计算公式如式(7)。

$$L_{BCE} = -\frac{1}{N} \sum_{i=0}^n [y_i \log_a p_i + (1 - y_i) \log_a (1 - p_i)] \quad (7)$$

式中:  $N$  为类别数,  $y_i$  为第  $i$  个像素的标签值,  $p_i$  为第  $i$  个像素的预测概率值。

本文数据集中各类破损像素分布不均衡, 类型占比如图 8 所示。若只用 BCE 损失函数, 网络只学习样本量大的背景数据, 模型准确率低。因此引入 Dice 损失函数, 弱化预测正确的像素, 强调预测错误的像素, 解决样本失衡问题, 计算公式如式(8)、(9)。

$$D_{Dice} = \frac{1}{n} \sum_{i=1}^n \frac{2a_{ii}}{\sum_{j=1}^n a_{ji} + \sum_{k=1}^n a_{ik}} \quad (8)$$

$$L_{Dice} = 1 - D_{Dice} = 1 - \frac{1}{n} \sum_{i=1}^n \frac{2a_{ii}}{\sum_{j=1}^n a_{ji} + \sum_{k=1}^n a_{ik}} \quad (9)$$

式中:  $D_{Dice}$  为 Dice 系数,  $L_{Dice}$  为 Dice 损失函数,  $n$  为类别数,  $a_{ji}$  是真值为第  $j$  类且预测为第  $i$  类的像素数,  $a_{ii}$ 、 $a_{ik}$  同理。

二者组合使用, 既能保证像素的有效性, 还能解决正负样本不平衡问题, 计算公式如式(10)所示。

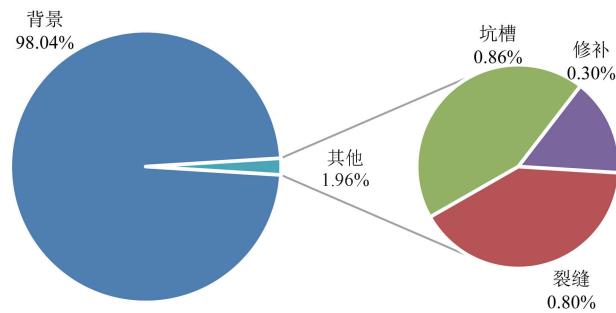


图 8 路面破损像素占比

Fig. 8 Percentage of pixels with broken pavement

$$L = L_{Dice} + \lambda L_{BCE} \quad (10)$$

式中:  $L$  为总损失,  $\lambda$  为 BCE 损失函数占的权重, 设为 0.5。本文的裂缝检测任务属于分割, 像素点分类是辅助, 权重设置较小, 模型才会倾向于预测占比少的裂缝像素, 提高分割精度。

## 2 实验与结果

### 2.1 实验数据与平台

目前关于路面破损的开源数据集大多是单类别的, 没有适用于航拍视角下包含多个破损类别的语义分割数据集。因此由大疆 M300RTK 无人机搭载禅思 P1 相机拍摄路面破损图像, 图像尺寸为  $8192 \times 5460$  像素。为避免训练过程中压缩图像导致信息丢失, 本文的数据集均为  $640 \times 640$  像素的航拍图像切片。为提升模型的鲁棒性, 同时考虑到工程需要, 数据集中包括不同天气状况、拍摄时间和无人机飞行高度等情况, 拍摄对象包括高速公路、国道、省道以及校园内部道路等, 场景包含阴影、杂物、水渍等噪声。最终采集 1500 张图像, 将训练集和测试集按 8:2 比例划分, 并用 Labelme 软件精细标注, 部分破损图像样本如图 9 所示。其中, 路面类型有水泥混凝土和沥青路

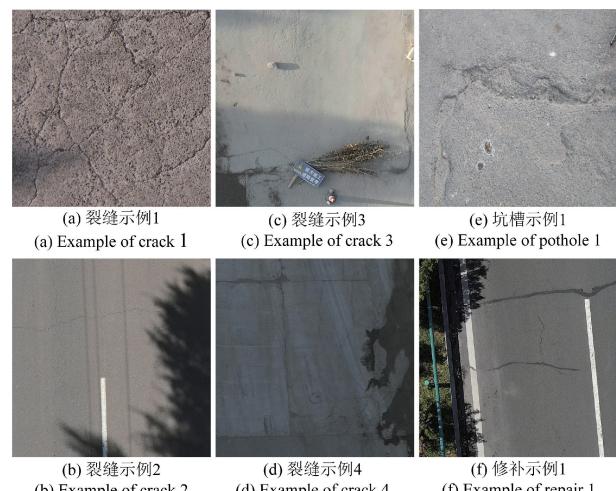


图 9 裂缝图像样本

Fig. 9 Crack image samples

类型包括裂缝、坑槽和修补,不同类别图像数的占比分别为 60%、13% 和 27%。

面,破损本文编程环境为 Python 3.9,模型基于 PyTorch 2.0 深度学习框架。处理器为 Intel i7-10700F,显卡为 NVIDIA GeForce RTX3090,16 GB RAM。采用 Adam 优化方法,批大小为 2,训练 300 轮。

## 2.2 评价指标

为了对比评价所提模型的性能,选择平均交并比(mean intersection over Union, mIoU)、平均精度(mean precision, mPrecision)、平均像素准确率(mean pixel accuracy, mPA)来衡量模型分割效果,计算方法如式(11)~(13)。

$$mIoU = \frac{1}{n} \sum_{i=1}^n \frac{a_{ii}}{\sum_{j=1}^n a_{ji} + \sum_{k=1}^n a_{ik} - a_{ii}} \quad (11)$$

$$mPrecision = \frac{1}{n} \sum_{i=1}^n \frac{a_{ii}}{\sum_{j=1}^n a_{ji}} \quad (12)$$

$$mPA = \frac{1}{n} \sum_{i=1}^n \frac{a_{ii}}{\sum_{j=1}^n a_{ij}} \quad (13)$$

式中:各变量含义同式(8)、(9), $a_{ij}$  同理。

## 2.3 对比实验

为验证改进模型的有效性,选取 UNet、VGG16-UNet、ResNet50-UNet、DeepLabv3+、PSPNet、HRNet 模型与本文模型在航拍路面破损数据集上进行对比实验,如表 1 所示。结果表明,本文提出的模型大小较基线增加了 1.1 倍,但在 mIoU、mPA 和 mPrecision 指标上均优于对比模型,较基线分别提升 15.97%、12.72%、11.02%。

表 1 不同模型在航拍数据集上的结果对比

Table 1 Comparison of results of different models on aerial photography dataset

模型	mIoU	mPA	mPrecision	Params/ KB
U-Net	0.661 2	0.803 0	0.754 9	76 555
VGG16-UNet	0.735 0	0.839 2	0.827 0	97 248
ResNet50-UNet	0.697 3	0.793 9	0.818 0	171 938
DeepLabV3+	0.732 9	0.830 4	0.829 9	214 744
PSPNet	0.636 2	0.787 2	0.738 4	182 806
HRNet	0.625 6	0.727 2	0.766 1	258 248
本文模型	<b>0.820 9</b>	<b>0.930 2</b>	<b>0.865 1</b>	162 573

图 10 为在航拍数据集上使用不同模型的结果对比,第 1、2 列为裂缝,第 3、4 列为坑槽,第 5、6 列为修补。第 1 列中,图 10(c)、(g)、(h) 模型出现不同程度的漏检,图 10(d)、(e) 模型将道路表面的痕迹分割为了裂缝;第 2 列中,图 10(c)、(e)、(f)、(g) 模型对块状裂缝有多处漏检,图 10(c)、(d)、(f)、(g)、(h) 模型将裂缝最右端错误分类为修补;第 3 列中,本文方法实现了对坑槽的近似准确分割,仅在边界上不精确,其他模型除边界分割不精确外,均存在不同程度的漏检、错检、分类错误现象;第 4 列中,只有图 10(d) 模型和本文方法的分割结果与标签近似,其中本文方法在最下方分割出的边界较为准确;第 5 列中,图 10(f)、(g) 模型没有分割出裂缝,图(c) 模型存在漏检和裂缝断裂现象,图 10(d)、(e)、(h) 模型有多检现象,本文模型分割结果与标签最接近;第 6 列中,图(c)、(d)、(f)、(g)、(h) 模型对裂缝存在不同程度的漏检,图(e) 模型与本文模型的检测结果与标签近似。综上,本文模型在各类路面破损上的分割效果均优于对比模型,尤其在裂缝和修补类型上表现最佳。

## 2.4 消融实验

为验证改进的 3 个模块对分割性能的影响,分别将 ECA\_X、MSFF、CESP 模块添加到基线网络上进行训练,结果如表 2 所示。实验 1 表示修改主干网络的结果。实验 2 引入 MSFF 模块后 mIoU、mPrecision 较实验 1 分别提高 0.011 6、0.013 5, mPA 降低 0.001 4, 说明扩大模型感受野、特征图密集连接能保留更多特征信息,但是增加了参数量。实验 4 在实验 2 基础上引入 ECA\_X 模块,模型能动态分配权重并进行二次特征提取,提升了网络的局部和全局特征提取能力,各指标结果虽然只有小幅提升,但是参数量只增加了 1.298 M。针对裂缝狭长且细小的结构,实验 6 在实验 4 基础上引入 CESP 模块, mIoU、mPA、mPrecision 分别达到 0.820 9、0.930 2、0.865 1, 说明模型对裂缝尾端和纤细部分的特征处理较好。综上,通过航拍数据集的消融实验,验证了 4 个改进部分的结合可有效提升分割性能。

## 2.5 模型泛化性实验

由于本文的实验都是在自建的无人机航拍数据集上完成的,为了更好说明网络与模型性能,将上述 7 种模型与文献[29]提出的算法在 Crack500 开源数据集上进行比较,结果如表 3 所示。本文模型在 mIoU、mPA、mPrecision 指标上优于其他模型,证明本文模型在处理不同路面状况时具备更强的鲁棒性。

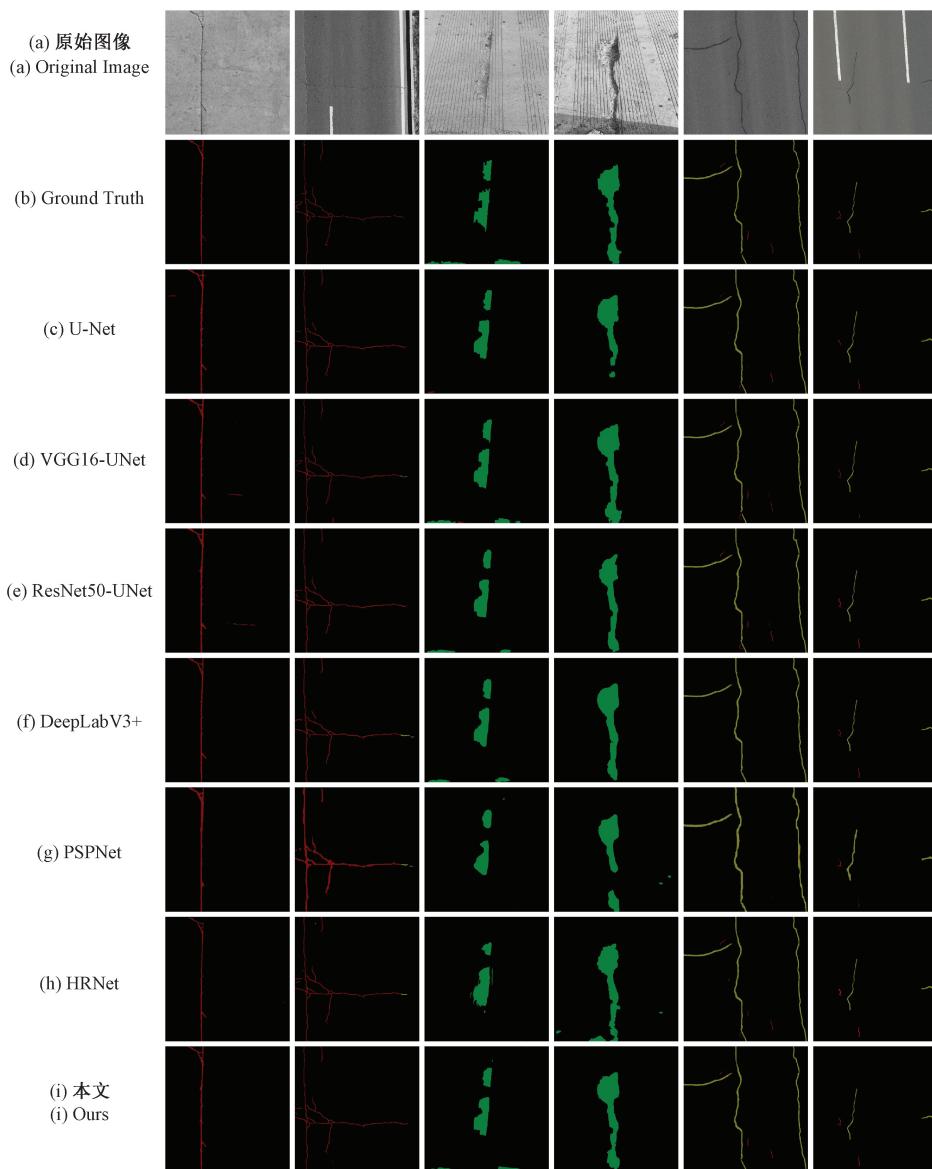


图 10 分割结果可视化

Fig. 10 Visualization of segmentation results

表 2 模型消融实验

Table 2 Model ablation experiments

Plan	Network	MSFF	ECA	CESP	mIoU	mPA	mPrecision	Params/KB
1	✓	—	—	—	0.784 4	0.920 1	0.831 0	97 248
2	✓	✓	—	—	0.796 0	0.918 7	0.844 5	159 795
3	✓	—	✓	—	0.795 9	0.921 7	0.843 1	98 546
4	✓	✓	✓	—	0.803 0	0.927 3	0.847 7	161 093
5	✓	✓	—	✓	0.814 7	0.925 1	0.862 4	161 275
6	✓	✓	✓	✓	<b>0.820 9</b>	<b>0.930 2</b>	<b>0.865 1</b>	162 573

表 3 Crack500 数据集泛化实验

Table 3 Generalisation experiments on the Crack500 dataset

模型	mIoU	mPA	mPrecision
U-Net	0.729 5	0.854 8	0.797 6
VGG16-UNet	0.772 0	0.886 6	0.828 9
ResNet50-UNet	0.768 8	0.890 3	0.826 8
DeepLabV3+	0.761 4	0.898 8	0.812 8
PSPNet	0.744 5	0.891 2	0.796 6
HRNet	0.756 2	0.896 6	0.807 8
文献[29]	0.791 0	0.893 0	0.731 0
本文模型	<b>0.809 1</b>	<b>0.910 8</b>	<b>0.852 2</b>

### 3 结 论

本文提出一种用于航拍路面破损分割的模型。模型基于 U-Net, 使用改进的 VGG-16 增强特征提取; 在最小尺度网络层引入 MSFF、CESP 模块, MSFF 扩大了模型感受野、增加网络深度的同时利用密集连接减少特征流失; CESP 在条形池化基础上进行通道增强, 使模型聚焦于不同通道间的关键信息; 解码阶段增加 ECA\_X 模块来注重通道关联性。结果表明, 本文模型可有效提升航拍图像中细小裂缝的检测精度, 解决裂缝检测的断裂问题, 且具有更好的泛化能力。但本文模型参数量大, 难以实现路面破损实时分割, 后续将平衡模型的检测精度和处理速度, 探究轻量化网络在该领域中的应用。

### 参考文献

- [1] XIANG X ZH, ZHANG Y Q, EL SADDIK A. Pavement crack detection network based on pyramid structure and attention mechanism [J]. IET Image Processing, 2020, 14(8): 1580-1586.
- [2] FAN ZH, LI C, CHEN Y, et al. Automatic crack detection on road pavements using encoder-decoder architecture [J]. Materials, 2020, 13(13): 2960.
- [3] 赵芳, 周旺辉, 陈岳涛, 等. 改进的 Canny 算子在裂缝检测中的应用 [J]. 电子测量技术, 2018, 41(20): 107-111.
- ZHAO F, ZHOU W H, CHEN Y T, et al. Application of improved canny operator in crack detection [J]. Electronic Measurement Technology, 2018, 41(20): 107-111.
- [4] TALAB A M A, HUANG ZH C, XI F, et al. Detection crack in image using Otsu method and multiple filtering in image processing techniques [J]. Optik, 2016, 127(3): 1030-1033.
- [5] LIANG D, ZHOU X F, WANG S, et al. Research on concrete cracks recognition based on dual convolutional neural network [J]. KSCE Journal of Civil Engineering, 2019, 23(7): 3066-3074.
- [6] LI W, HUYAN J, GAO R, et al. Unsupervised deep learning for road crack classification by fusing convolutional neural network and k-means clustering [J]. Journal of Transportation Engineering Part B: Pavements, 2021, 147(4): 04021066.
- [7] IBRAGIMOV E, LEE H J, LEE J J, et al. Automated pavement distress detection using region based convolutional neural networks [J]. International Journal of Pavement Engineering, 2022, 23(6): 1981-1992.
- [8] 张华卫, 贾文娟, 张金龙, 等. 并联提取与特征融合注意力网络下的裂缝检测 [J]. 电子测量技术, 2022, 45(10): 102-111.
- ZHANG H W, JIA W J, ZHANG J L, et al. Crack detection under parallel extraction and feature fusion attention network [J]. Electronic Measurement Technology, 2022, 45(10): 102-111.
- [9] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]. 2015 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015: 234-241.
- [10] LIU ZH Q, CAO Y W, WANG Y Z, et al. Computer vision-based concrete crack detection using U-net fully convolutional networks [J]. Automation in Construction, 2019, 104: 129-139.
- [11] LAU S L H, CHONG E K P, YANG X, et al. Automated pavement crack segmentation using U-net-based convolutional neural network [J]. IEEE Access, 2020, 8: 114892-114899.
- [12] 张伟光, 钟靖涛, 呼延菊, 等. 基于 VGG16-UNet 语义分割模型的路面龟裂形态提取与量化 [J]. 交通运输工程学报, 2023, 23(2): 166-182.
- ZHANG W G, ZHONG J T, HU Y J, et al. Extraction and quantification of pavement alligator crack morphology based on VGG16-UNet semantic segmentation model [J]. Journal of Traffic and Transportation Engineering, 2023, 23(2): 166-182.
- [13] 王瑞绅, 宋公飞, 王明. 引入 ECA 注意力机制的 U-Net 语义分割 [J]. 电光与控制, 2023, 30(1): 92-96, 102.
- WANG R SH, SONG G F, WANG M. U-Net semantic segmentation with ECA attention mechanism [J]. Electronics Optics & Control, 2023, 30(1): 92-96, 102.
- [14] 夏晓华, 苏建功, 王耀耀, 等. 基于 DeepLabV3+ 的轻量化路面裂缝检测模型 [J]. 激光与光电子学进展, 2024, 61(8): 182-191.

- XIA X H, SU J G, WANG Y Y, et al. Lightweight pavement crack detection model based on Deeplabv3[J]. *Laser & Optoelectronics Progress*, 2024, 61(8): 182-191.
- [15] LI L, WANG N, WU B, et al. Segmentation algorithm of bridge crack image based on modified PSPNet[J]. *Laser & Optoelectronics Progress*, 2021, 58(22): 2210001.
- [16] 张大伟,田抑阳,徐培娟,等.基于 E-HRNet 的路面破损区域识别方法[J].北京交通大学学报,2023,47(4): 110-119.
- ZHANG D W, TIAN Y Y, XU P J, et al. Road surface damage area identification method based on E-HRNet[J]. *Journal of Beijing Jiaotong University*, 2023, 47(4): 110-119.
- [17] 范瑶瑶,王兴芬,刘亚辉.改进 DeepLabv3+网络的钢板表面缺陷检测研究[J].计算机工程与应用,2023,59(16): 150-158.
- FAN Y Y, WANG X F, LIU Y H. Improved DeepLabv3+ model for surface defect detection on steel plates [J]. *Computer Engineering and Applications*, 2023, 59(16): 150-158.
- [18] 武广臣,刘艳.无人机影像 GrabCut 路面裂缝识别[J].测绘通报,2024(8):90-95.
- WU G CH, LIU Y. Pavement cracks recognition by using UAV image based on GrabCut method [J]. *Bulletin of Surveying and Mapping*, 2024(8): 90-95.
- [19] 卢小平,张航,张冬梅,等.一种基于无人机影像的迭代二值化道路裂缝检测方法[J].河南理工大学学报(自然科学版),2019,38(6):55-59.
- LU X P, ZHANG H, ZHANG D M, et al. An iterative binary road crack detection method based on UAV image [J]. *Journal of Henan Polytechnic University(Natural Science)*, 2019, 38(6): 55-59.
- [20] HE X Y, TANG ZH W, DENG Y B, et al. UAV-based road crack object-detection algorithm [J]. *Automation in Construction*, 2023, 154: 105014.
- [21] HONG ZH H, YANG F, PAN H Y, et al. Highway crack segmentation from unmanned aerial vehicle images using deep learning[J]. *IEEE Geoscience and Remote Sensing Letters*, 2022, 19: 1-5.
- [22] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [23] ZOU Q, ZHANG ZH, LI Q Q, et al. DeepCrack: Learning hierarchical convolutional features for crack detection [J]. *IEEE Transactions on Image Processing*, 2019, 28(3): 1498-1512.
- [24] 惠冰,李远见.基于改进 U 型神经网络的路面裂缝检测方法[J].*交通信息与安全*,2023,41(1): 105-114,131.
- HUI B, LI Y J. A detection method for pavement cracks based on an improved U-shaped network[J]. *Journal of Transport Information and Safety*, 2023, 41(1): 105-114,131.
- [25] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [26] HE K, ZHANG X Y, REN SH Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [27] HOU Q B, ZHANG L, CHEN M M, et al. Strip pooling: Rethinking spatial pooling for scene parsing[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 4002-4011.
- [28] 赵志宏,郝子晔,何朋.融合注意力机制与 GhostUNet 的路面裂缝检测方法[J].*电子测量技术*,2023,46(24):164-171.
- ZHAO ZH H, HAO Z Y, HE P. Combining attention mechanism with GhostUNet method for pavement crack detection[J]. *Electronic Measurement Technology*, 2023, 46(24): 164-171.
- [29] ZHOU Q, QU ZH, LI Y X, et al. Tunnel crack detection with linear seam based on mixed attention and multiscale feature fusion[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-11.

## 作者简介

韩建峰,教授,硕士生导师,主要研究方向为图像处理与图像识别、信号处理、计算机视觉。

E-mail:hanjianfeng@imut.edu.cn

张静,硕士研究生,主要研究方向为信号与信息处理。

E-mail:jingz0915@163.com

宋丽丽(通信作者),副教授,硕士生导师,主要研究方向为图像处理与图像识别、信号处理。

E-mail:songlili@imut.edu.cn