

FastCrack: 实时路面裂缝分割

乐壮¹, 陈晓冬^{1*}, 汪毅¹, 蔡怀宇¹, 闫卫喜², 侯丽莹³

¹天津大学精密仪器与光电子工程学院光电信息技术教育部重点实验室, 天津 300072;

²天津高速公路集团有限公司科技与信息中心, 天津 300384;

³中交远洲工程咨询有限公司养护部, 河北 石家庄 050030

摘要 在公路的使用过程中,其表面会产生危害结构健康的裂缝,研究高效精确裂缝分割算法已成为交通运输领域的一个重要课题。在现有的基于图像的裂缝分割方法中,由数据驱动的深度学习技术展现出了强大的适用性。但基于神经网络的裂缝分割模型普遍缺乏对模型实时性的关注,为平衡模型精度和速度,选择合适的架构超参数,设计了一套架构超参数挑选框架,提出了实时路面裂缝分割模型(FastCrack-SPOS)。首先,设置不同的待选宽度(16,32,48,64,80)、深度(D1,D2,D3)、下采样倍率(1/4,1/8,1/32),构建出45组不同结构模型,分析每种参数对模型性能的影响;然后,使用神经架构搜索技术为模型每层自动搜索合适类型的卷积块,进而构建出最终模型。实验结果表明:所提架构超参数挑选框架在轻量级裂缝分割模型的设计上十分有效;构建出的FastCrack-SPOS在路面裂缝数据集上的交并比达62.88%;参数量仅 0.29×10^6 ,相比现有模型,减少95%;处理 1024×1024 的图像的速度达147 frame/s,在速度和精度间取得了平衡,具有较高的实际应用价值。

关键词 图像处理; 路面裂缝分割; 轻量级网络; 超参数选择; 神经架构搜索

中图分类号 TP391.9

文献标志码 A

DOI: 10.3788/LOP220754

FastCrack: Real-Time Pavement Crack Segmentation

Yue Zhuang¹, Chen Xiaodong^{1*}, Wang Yi¹, Cai Huaiyu¹, Yan Weixi², Hou Liying³

¹Key Laboratory of Photoelectric Information, Ministry of Education, School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China;

²Technology and Information Center, Tianjin Expressway Group Incorporated Company, Tianjin 300384, China;

³Maintenance Department, China Roads Engineering Consulting Incorporated Company, Shijiazhuang 050030, Hebei, China

Abstract The use of highways can lead to various cracks on their surface, which can harm the structure. Thus, the research on efficient and accurate crack segmentation algorithms in transportation has attracted significant interest in recent times. Data-driven deep learning technology showed the best applicability among the existing image-based crack segmentation methods. However, crack segmentation models based on neural network generally lack attention to real-time performance. Therefore, this study designs a set of structure hyperparameter selection frameworks and proposes a real-time pavement crack segmentation model (FastCrack-SPOS) to balance the accuracy and speed of the model and to select the appropriate structure hyperparameters. First, we constructed 45 groups of different structural models with various widths (16, 32, 48, 64, 80); depths (D1, D2, D3); and down-sampling ratios (1/4, 1/8, 1/32) and analyzed the effects of each parameter on model performance. Then, we used the neural architecture search technology to search for suitable convolution blocks for each layer and constructed the model. Experimental results reveal that the proposed architecture hyperparameter selection method is highly effective for lightweight crack segmentation model design. Our FastCrack-SPOS has an intersection ratio of 62.88% in the pavement crack dataset, and the number of parameters is only 0.29×10^6 , which is a reduction by 95% compared to existing models. For processing images with size of 1024×1024 , the speed attained by the FastCrack-SPOS is 147 frames/s, thereby achieving a balance between speed and accuracy, leading to its high practical application value.

收稿日期: 2022-02-17; 修回日期: 2022-02-28; 录用日期: 2022-03-30; 网络首发日期: 2022-04-12

基金项目: 天津市交通运输科技发展项目(2019-03)

通信作者: *xdchen@tju.edu.cn

Key words image processing; pavement crack segmentation; lightweight network; hyperparameter selection; neural architecture search

1 引言

公路是交通运输领域的重要组成部分。在温度变化、车辆载荷、结构形变等因素的作用下,公路在使用过程中通常会出现结构性损坏,其中最常见的是路面裂缝^[1]。路面裂缝是评估混凝土公路结构健康状况的一个重要依据^[2]。严重的路面缺陷会给交通运输带来很大的安全隐患^[3]。对混凝土路面裂缝进行及时检测是十分重要的^[4]。早期的裂缝检测工作都是通过人工观测来完成的,成本高、工作量大,且由于判别标准不一致,经常出现漏检和误检的现象^[5-6]。

基于图像的裂缝检测方法普遍取代了人工检测,传统的数字图像处理算法由于容易实现、易于移植的特点得到了广泛研究^[7-10]。Zou 等^[7]提出了一种名为 CrackTree 的全自动裂缝检测算法,利用张量投票算法计算裂缝概率图,再由最小生成树细化得到裂缝骨架,该方法可较好抵抗阴影干扰,但张量投票算法速度太慢。为进一步提高检测速度和精度,出现了基于边缘检测^[9]、阈值分割^[10]的系列方法。传统数字图像处理算法在背景简单、噪声较少的情况下具有较好的检测结果,但在实际应用场景中,往往会出现阴影、油漆、坑洞等复杂噪声,传统算法难以精确检测出裂缝区域^[11-12]。

由于传统算法的局限性日渐突出,人工难以设计出可抵抗复杂噪声的数字图像检测算法,机器学习技术被应用到了裂缝检测中。Shi 等^[13]为检测纹理形状更加复杂的裂缝,将积分通道作为裂缝特征,利用基于随机结构森林的裂缝检测器来高效识别复杂裂缝,检测效果明显优于传统方法。基于机器学习的方法虽然取得了更加优异的结果,但特征提取环节仍需手动设计,依赖研究者的先验知识,难以设计出一套通用的特征提取方法^[14]。卷积神经网络能够弥补机器学习方法的不足,其特征提取和分类环节均在模型内部自动完成,无需手动设计特征提取环节。李刚等^[15]将卷积神经网络(CNN)和滑动窗口技术相结合,用于图像分类任务,系统精度达 97.3%,但滑动窗口大小需手动设置,无法根据裂缝尺寸自适应检测。目标检测模型即可弥补分类模型的不足^[16-17]。毛莺池等^[18]将目标检测模型 Faster-RCNN 应用到裂缝检测中,模型可根据裂缝尺寸自适应标记出裂缝所在矩形框,平均交并比达 82.5%,但不能对图像中的裂缝和背景进行像素级分割,模型输出不能直接用于裂缝物理参数的计算。

对于语义分割模型,其输入和输出同尺寸,输出赋予每个像素点一个类别,为只包含 0(代表背景)或者 1(代表裂缝)的二值图^[19-20],可直接用于裂缝物理参数

的计算。目前,语义分割模型在裂缝分割任务上得到了广泛应用^[4,14,21-22]。Liu 等^[4]将 VGG16 作为网络主干,在拼接编码阶段对各层特征图融合多尺度信息,同时使用条件随机场细化裂缝边界,平均交并比达 85.9%,但分割速度只有 9 frame/s。Jiang 等^[14]为提高模型分割速度,提出一种快速二阶段检测方法,首先使用目标检测模型 YOLOv4 生成裂缝所在矩形区域,再对目标区域进行像素级分割,速度提升为单阶段方法的 7.5 倍,但速度仍只有 1.5 frame/s。基于语义分割的裂缝分割模型普遍存在参数量高、计算量大的问题。目前,一些工作注意到了这一问题,致力于设计轻量级裂缝分割模型^[18-20]。Shim 等^[23]将 DenseNet121 作为主干,对编码阶段进行精简,全程使用薄特征图,参数量减少 85.7%。Choi 等^[24]将可分离卷积和稠密连接块结合,在编码器阶段下采样到 1/16,对 1024×512 尺寸图像的处理速度达 36 frame/s。Shim 等^[25]为进一步简化模型,直接删除解码器,仅保留轻量级编码器,全程保留高分辨率薄特征图。

轻量级裂缝分割模型正向特征图更薄、分辨率更高、层数更少的方向发展。但现有工作中,难以找到对模型分辨率、层数等架构超参数的调节依据,既没有完善的理论为模型架构精简提供依据,也缺乏实验性的工作来为裂缝分割模型架构超参数的选择提供参考,缺少对实时裂缝分割模型架构超参数研究的工作。

为平衡模型精度和速度,构建轻量化裂缝分割模型,本文针对实时裂缝分割任务,设计了一套架构超参数挑选框架,研究了架构超参数对裂缝分割模型性能的影响,主要考虑了 4 种架构超参数,即宽度(通道数)、深度(层数)、下采样倍率和卷积块类型。首先,为模型宽度、深度、下采样倍率分别设置了多个候选值,在此基础上构建了 45 种不同架构的模型(对应 45 种组合),对这 45 种模型在本文的混合路面裂缝数据集上进行训练,根据实验结果分析架构超参数的变化对模型精度的影响,从这些组合中挑选出模型的最佳宽度、深度和下采样倍率;然后,在已选架构基础上,通过单路径神经架构搜索(SPOS)算法^[26]进一步为模型各层选出合适类型的卷积,在此基础上提出了实时路面裂缝分割模型(FastCrack-SPOS),通过对比实验,验证了 FastCrack-SPOS 的有效性。本文贡献包括:1)针对实时裂缝分割任务,设计了一套架构超参数挑选框架,提出了实时路面裂缝分割模型 FastCrack-SPOS,在 62.8% 的交并比下对 1024×1024 输入的处理速度达 144 frame/s,具有较高的实际应用价值;2)通过对 45 种不同架构模型进行实验,分析了裂缝分割模型宽度、深度、下采样倍率三种架构参数对模型精度和速度

的影响,从中选出了模型的最佳宽度、深度和下采样倍率,为裂缝分割模型的轻量化工作提供了参考;3)首次将神经架构搜索技术应用到裂缝分割任务中,该技术可自动为模型每层搜索出合适类型的卷积块。在本文构建的混合裂缝数据集上,单路径神经架构搜索算法只需一天就可从 1.68×10^8 种模型结构中搜索出高性能架构,大幅缩短了人工挑选卷积块的时间,搜索出的模型在精度上优于由搜索空间中单一卷积块构建的模型,为未来向裂缝分割领域引入更多的自动化模型设计技术带来了先例。

2 所提模型设计方法

2.1 架构超参数挑选框架

模型的架构超参数对性能至关重要,会直接影响模型的精度和实时速度,架构超参数一般包含卷积核尺寸、深度(层数)、宽度(通道数)^[27]。文献[28]表明,大尺寸卷积可由多个小尺寸卷积连续组合来代替,小尺寸卷积的组合在保持感受野的同时,具有更少的参数量和计算量,且能一定程度上提高性能。Zhang 等^[29]对裂缝分割模型进行研究,发现大尺寸卷积核导致模型性能下降。因此,本文不将卷积核尺寸作为架构超参数实验变量。裂缝分割任务中,待分割

裂缝结构细长,在图像中的占比小,细节信息易在多次下采样后丢失,且在已有工作中^[15-17],模型下采样倍率存在多种设置,故将模型的下采样倍率也作为待分析架构超参数。因此,首先考虑模型深度、宽度、下采样倍率 3 个架构超参数,对它们进行多组实验,分析架构参数的变化对模型精度和实时性的影响。另一方面,模型各层卷积类型对性能同样存在重要影响,在模型每层设置不同类型的卷积块,合适的卷积块组合构建出的模型比单一卷积构建的模型具有更好的性能^[30]。Guo 等^[26]通过实验表明,可使用单路径神经架构搜索算法为模型每层搜索出合适类型卷积块,这些卷积块组合构建出的模型在精度上优于由单一卷积块构成的模型。因此,对于模型各层卷积块类型这一超参数,本文以实用化为导向,不进行大量组合实验,而是使用单路径神经架构搜索算法,直接为模型每层自动搜索合适类型的卷积块。在裂缝分割领域,还没有使用神经架构搜索技术来自动设计裂缝分割模型的先例。

设计的架构超参数挑选框架如图 1 所示,主要分为两部分。第一部分,为模型宽度、深度、下采样倍率分别设置多个候选值,在此基础上构建了 45 种不同架构的模型(对应 45 种组合),对这 45 种模型在本文的混

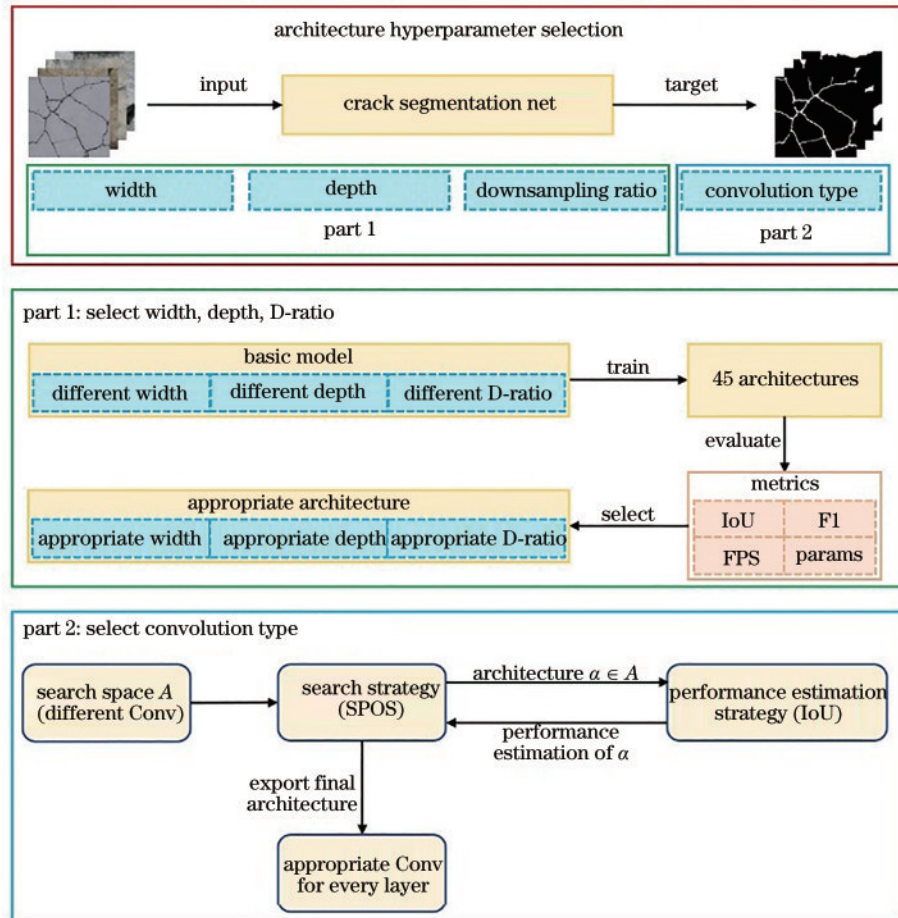


图 1 架构超参数挑选框架

Fig. 1 Architecture hyperparameter selection framework

合路面裂缝数据集上进行训练,根据实验结果分析架构超参数的变化对模型精度和速度的影响,从这些组合中挑选出模型的最佳宽度、深度和下采样倍率;第二部分,在上述架构基础上,使用单路径神经架构搜索算法为模型各层选出合适的卷积块,进而构建出最终模型。最后,对所设计模型和现有模型进行比较,以验证

设计的实时裂缝分割模型的有效性。

2.2 初始模型

DDRNet^[31]为实时语义分割领域的 state of the arts(SOTA)模型,参数量较少,方便在其基础上继续进行轻量化研究。对DDRNet进行精简,将简化得到的模型称为OriginalNet,结构如图2所示。

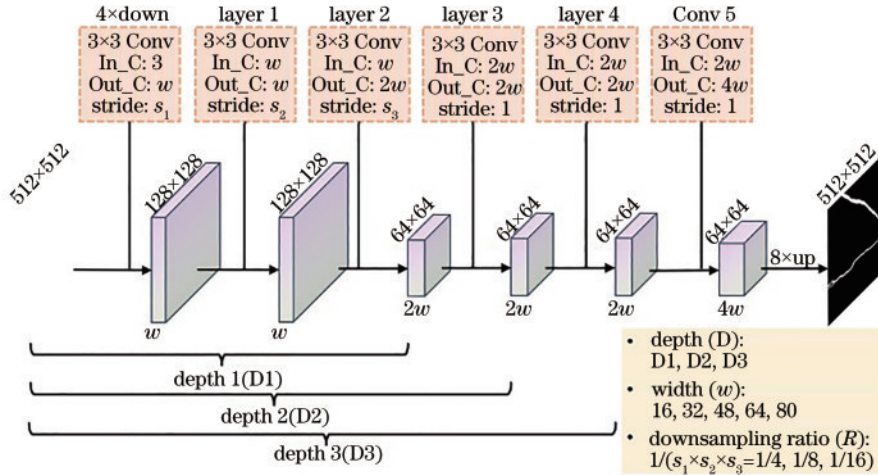


图2 模型宽度、深度和下采样倍率实验

Fig. 2 Width, depth, and downsampling rate experiment of the model

2.3 第一部分:宽度、深度和下采样倍率

模型宽度、深度和下采样倍率3个超参数的实验设置如图2所示,均在初始模型OriginalNet的基础上进行变化。

对于模型深度,为保证模型轻量化,将模型层数逐层删减^[32],设置三种情况,即D1(不包含layer 3和layer 4)、D2(不包含layer 4)和D3(完整层数);对于模型宽度(通道数),令其为w,为保证硬件计算的高效

性,将16设置为最小通道数^[33],在16基础上对通道进行递增,设置待选值为{16, 32, 48, 64, 80};对于下采样倍率,参照裂缝分割工作中的常见值^[15-17],考虑对细节信息和类别信息的关注度,设置待选值为{1/4, 1/8, 1/16}。总共可构建出3×5×3=45种模型架构,在完全相同的实验设置下,对这些模型进行实验。

2.4 第二部分:卷积块类型

模型各层卷积类型选择过程如图3所示。首先,

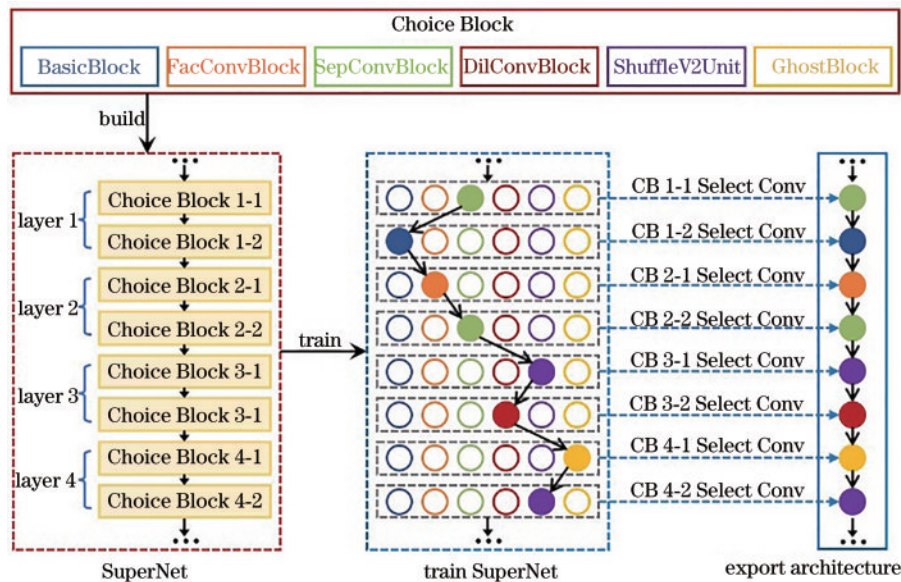


图3 卷积块类型选择

Fig. 3 Convolution block type selection

设置一组待搜索卷积块 Choice Block, 其包含 6 种深度学习领域广泛应用的卷积模块: BasicBlock^[34]、FacConvBlock^[35]、SepConvBlock^[36]、DilConvBlock^[37]、ShuffleV2Unit^[38]、GhostBlock^[39], 再使用 Choice Block 替换原始模型中的卷积块, 构建的超网如图 3 所示, 搜索空间中总共 $6^8 = 1679616$ 种候选架构。然后, 使用 SPOS 算法对超网进行训练, 具体来说, 在每次迭代中, 从超网每个 Choice Block 中随机采样一个卷积块, 将所有采样卷积块的组合作为当前的采样子网, 使用采样子网对训练数据进行正向推理, 反向传播时也仅更新采样子网对应的权值参数, 训练完成后, 挑出超网中精度最高的采样子网作为模型输出, 其每层对应从 Choice Block 中搜索出的卷积块。

算法流程如图 4 所示, 可表示为

$$W^* = \arg \min_w E_{\alpha \in A} \left\{ L_{\text{train}} \left[N(\alpha, W_\alpha) \right] \right\}, \quad (1)$$

式中: W 代表超网权重(所有架构共享); A 为搜索空间; α 为从搜索空间 A 中随机采样的架构; W_α 代表采样架构 α 对应的权重; W^* 代表超网训练完得到的最佳权重; $N(\alpha, W_\alpha)$ 代表架构 α 对应的模型; $L_{\text{train}}[N(\alpha, W_\alpha)]$ 表示架构 α 对应模型在训练集上的损失。式(1)中的期望通过均匀采样近似。

Algorithm 1: convolution type search algorithm

```

1 Input: search space A; SuperNet weight W
2 Output: high performance architecture
3 for  $N_{\text{epoch}}=1, \dots, N_{\text{num\_epochs}}$  do
4   for  $N_{\text{step}}=1, \dots, N_{\text{iterations}}$  do
5     randomly sample an architecture  $\alpha$  from search space A
6     calculate hard loss  $L_{\text{hard}}$  over one step
7     update the weight  $W_\alpha$  of subnet  $N(\alpha, W_\alpha)$ 
8     if  $R_{\text{IoU val}}(\alpha) > R_{\text{IoU val}}(\alpha_{\text{top}})$  then
9        $\alpha_{\text{top}} = \alpha$ 
10    end if
11  end for
12 endfor
13 export  $\alpha_{\text{top}}$  architecture

```

图 4 卷积类型搜索算法

Fig. 4 Convolution type search algorithm

3 实验

3.1 数据集

为增加数据集丰富程度, 考虑到不同的图像分辨率、背景环境、干扰噪声、裂缝几何形状和尺寸, 选取了 6 个公开的路面裂缝数据集: CrackTree^[7]、CFD^[13]、GAPs384^[40]、CrackDataset^[22]、DeepCrack^[4]、Crack500^[40]。在其基础上构建一个混合裂缝数据集, 共包含图片 5514 张, 将训练集和验证集按照 5:1 比例划分, 尺寸从 320×480 变化到 600×800 , 裂缝区域在图像中的占比从 4% 变化到 47%, 裂缝形状各异。背景包含阴影、油斑、水渍、坑洞等多种噪声。

3.2 评价指标

采用裂缝分割任务常见的精确率(precision)、召回率(recall)、F1 分数和交并比(IoU)作为模型评价指标^[4]。精确率表示预测为裂缝的像素点中真实类别为裂缝的比例, 召回率表示真实裂缝点被准确预测的比例, 精确率和召回率有时是相互矛盾的。F1 分数一般作为分类任务的总体指标。交并比表示预测裂缝区域和真实裂缝区域交集和并集的比值, 能够较好反映网络对裂缝的分割性能。

3.3 对比方法

确定模型的最终结构后, 和以下模型进行对比: 经典裂缝分割模型 DeepCrack^[41]、HED^[42]; 常见语义分割模型 UNet^[43]、SegNet^[44]、DeeplabV3+^[45]; 常见实时语义分割模型 ESPNet^[46]、FastSCNN^[47]、DFANet^[48]。

3.4 实验设置

3.4.1 实验平台

实验硬件配置为 NVIDIA GeForce GTX 1080ti GPU (11G), Intel Core i7-7700K CPU @4.2 GHz, 16 GB RAM。系统为 Windows10, 编程语言为 Python3.8, 深度学习框架为 PyTorch1.9.0。所有模型的推理速度均在 1080ti GPU 上以 1024×1024 分辨率进行测试。

3.4.2 数据增强

为进一步提高模型鲁棒性, 使用数据增强来扩大训练集, 采用三种常见的数据增强策略^[31]: 随机水平/垂直翻转图片; 随机调整亮度、饱和度、对比度为初始图像的 0.6~1.4; 随机将图像尺寸缩放为原图的 0.5~1.75。最后从其中裁剪 512×512 的区域作为模型训练时的输入。数据增强策略不会增加数据集图像总数, 而是在每轮读取训练集数据时对图片进行图像处理, 图像处理算法参数在一定范围内随机变化, 可在每轮生成完全不同的训练集, 更好地提高模型鲁棒性。

3.4.3 训练设置

使用 SGD 优化器, 初始学习率为 0.5×10^{-2} , 动量为 0.9, weight decay 为 5×10^{-4} , 采用 poly 学习率衰减策略, poly 中的 power 为 0.9, 预热 10 轮, 所有模型训练 300 个 epochs, batchsize 为 8。针对正负样本不平衡的情况, 采用裂缝分割中广泛使用的 Dice 系数^[14,40-41]作为损失函数。

4 结果与讨论

4.1 第一部分: 宽度、深度和下采样倍率

设置的 45 种结构模型在路面数据集上的实验结果如图 5 所示, 对各模型交并比、F1 分数、精确率、召回率、速度和参数量进行直观比较。图 6 展示了使用网格搜索挑选架构超参数的流程。

1) 在所有模型中, 33 号模型的交并比和 F1 分数最高, 其宽度为 64、深度为 D3、下采样倍率为 1/8; 7 号模型在所有模型中推理速度最高, 其宽度最小、深度最浅、下采样倍率最小。对于所有模型, 交并比从

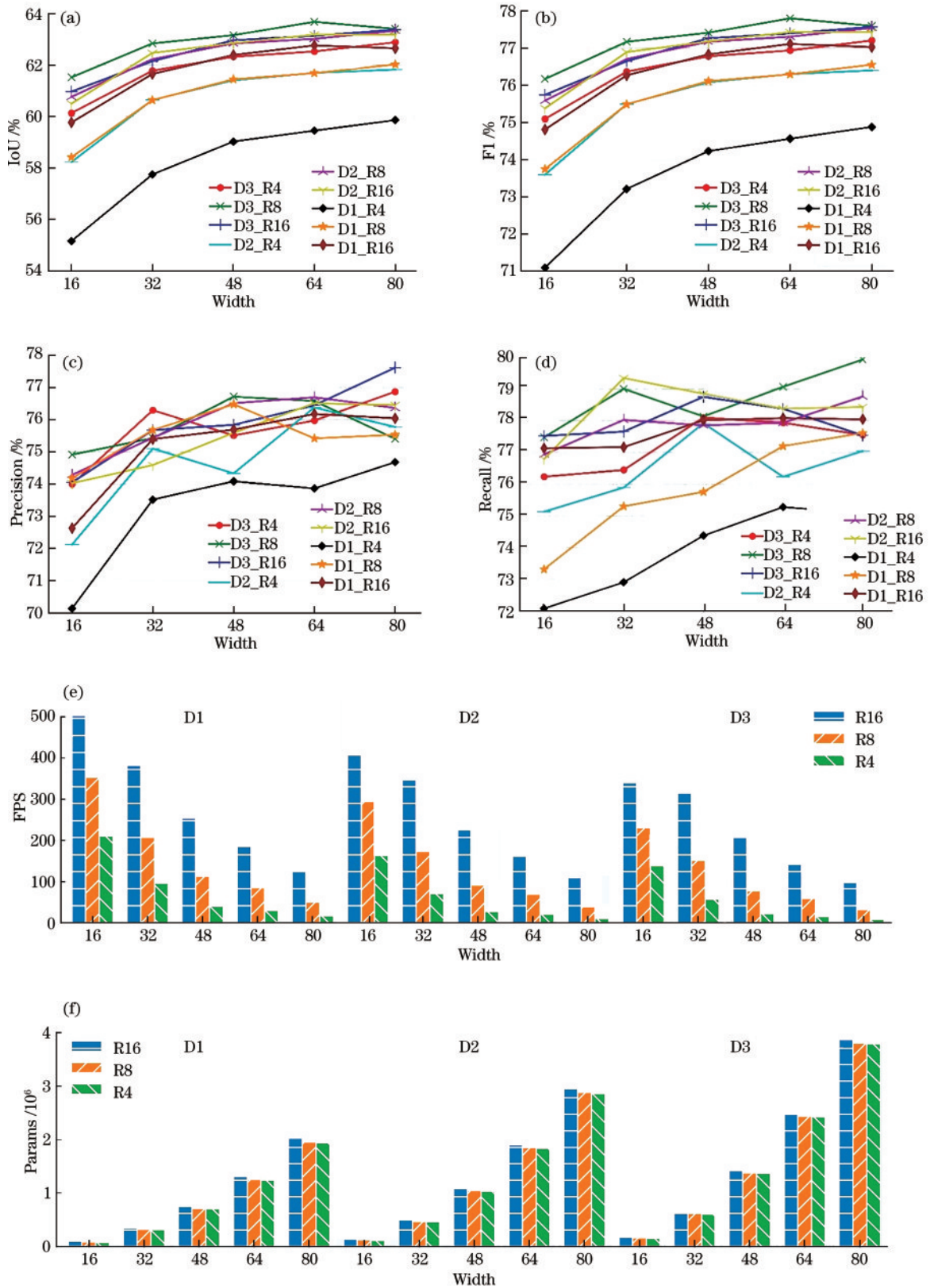


图5 各模型的参数指标比较结果,D为模型深度,R为下采样倍率。(a)交并比;(b)F1分数;(c)精确率;(d)召回率;(e)速度;(f)参数量

Fig. 5 Comparison result of parameter indicators of each model, D represents model depth and R represents downsampling rate.
 (a) IoU; (b) F1 score; (c) precision; (d) recall; (e) FPS; (f) number of parameters

55.15% 变化到 63.67%, 参数量从 0.0797×10^6 变化到 3.869×10^6 , 速度从 8 frame/s 变化到 503 frame/s。

2) 深度。深度对性能的影响如图 5(a) 和图 5(b) 所示, 呈现出单调趋势, 随着深度(层数)增加, 性能也

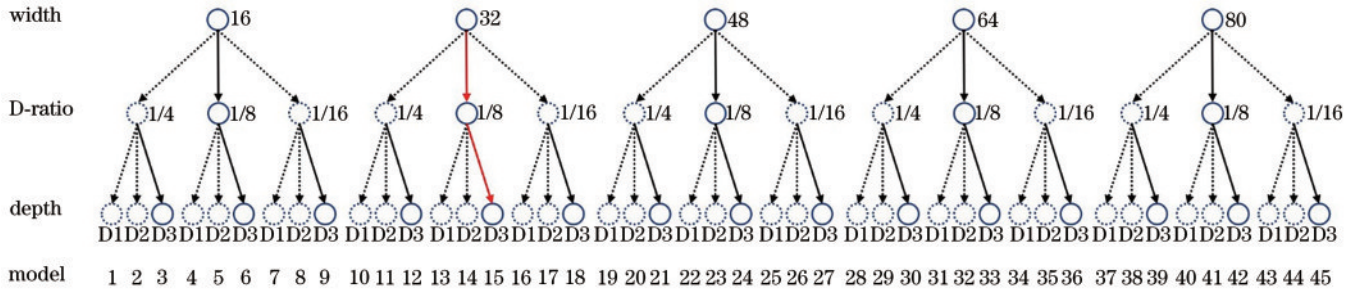


图 6 网格搜索挑选架构超参数

Fig. 6 Hyperparameters selected by grid search

随之上升,但层数越多时,性能的提升就越慢;模型深度越大,宽度增加对模型性能的提升越慢。深度对性能的影响也可从模型本身参数量来理解,模型层数越多,通道数越高,模型参数量越大,对数据集的拟合情况也越好。深度对参数量和推理速度的影响如图 5(e)和图 5(f)所示,深度越大,模型参数量越多,推理速度越低。本次实验最高深度仅为 4 层,参数量保持在个位数以下,较为轻量,且深度越深,精度越高。因此,当宽度和下采样倍率固定时,选择实验中的 D3 作为最终深度,相比其他深度值,得到的精度值最高,分割性能最好。

3) 下采样倍率。下采样倍率对模型交并比呈现非单调的影响趋势。如图 5(a)所示,当模型深度达 D2 时,下采样倍率先从 1/4 变到 1/8,再变到 1/16,交并比呈现出先升高后降低的趋势,此时 1/8 倍率下精度最高。实际上,倍率为 1/4 时,特征图尺寸更大,网络能关注到更多细节信息,但相对地缺失了高级语义信息,对区域类别的判断会存在偏差;而倍率为 1/16 时,特征图尺寸更小,模型能关注到更多语义类别信息,但相应地特征图中就会缺失更多的细节信息^[32];在 1/8 的倍率下,网络则能在细节信息和语义信息间取得平衡。如图 5(e)和图 5(f)所示,下采样倍率对参数量没有影响,对速度影响较大,倍率越高,推理速度越低;宽度越大时,下采样倍率加倍导致的速度损失越大。在深度为 D3、宽度值固定时,1/8 的倍率下能够达到最大的精度指标,且下采样倍率的变化对模型参数量几乎没有影响。显然,选择同等参数下精度更高的 1/8 下采样倍率。

4) 宽度。宽度对性能的影响同样如图 5(a)和图 5(b)所示,呈现出和深度类似的单调趋势,当宽度增加时,无论深度和下采样倍率为多少,模型的交并比都会逐渐上升,但随着宽度继续增大,性能的上升幅度越来越小。这正好验证了现有结论,更宽的网络能捕获更加细粒度的特征,模型的分割性能就会上升,但随着宽度继续增大,模型中缺乏高层特征,性能的上升幅度也就越来越小^[32]。图 5(e)和图 5(f)表明,宽度增加后,模型计算量增加,推理速度明显降低,且宽度相比深度对参数量和计算量的影响更大。挑选宽度时,综合考虑模型精度和推理速度,计算速度的降低能否带来与其匹配的精度增益。采取如下方法:确定深度为 D3、下采样倍率为 1/8 后,还有 5 种待选模型,宽度为 {16, 32, 48, 64, 80}, 设置一个宽度为 8 的基线模型,以该模型交并比和 FPS 为基准,计算 5 个待选模型相对于基线模型的交并比变化量 ΔR_{IoU} 和速度变化量 ΔS_{FPS} ,将 $\Delta R_{IoU} / \Delta S_{FPS}$ 作为衡量指标,此值越大,表明模型精度提高时推理速度的降低程度越小,或者说速度降低时模型精度提升越大。

ΔR_{IoU} 和 ΔS_{FPS} 的计算结果如表 1 所示。宽度为 16 的模型即为本文使用的初始模型 OriginalNet,其交并比为 61.51%。而宽度为 32 的模型的 $\Delta R_{IoU} / \Delta S_{FPS}$ 值最大,说明该模型降低单位速度时给精度指标带来的增幅最大,性能提升更划算。因此,从 45 种模型中选择架构超参数为 $w=32$ 、 $D=D3$ 、 $R=1/8$ 的 15 号模型作为最终的轻量级模型结构,将其称为 FastCrack。

4.2 第二部分:卷积类型

对确定深度、宽度和下采样倍率的模型结构

表 1 各模型相对基线模型的 ΔR_{IoU} 和 ΔS_{FPS}
Table 1 ΔR_{IoU} and ΔS_{FPS} of each model relative to baseline model

w	D	R	IoU / %	$S_{FPS} / (\text{frame} \cdot \text{s}^{-1})$	$\Delta R_{IoU} / \%$	$\Delta S_{FPS} / \%$	$\frac{\Delta R_{IoU}}{\Delta S_{FPS}}$
80	D3	1/8	63.40	32	3.84	322	0.01193
64	D3	1/8	63.67	59	4.11	295	0.01393
48	D3	1/8	63.15	78	3.59	276	0.01301
32	D3	1/8	62.83	152	3.27	202	0.01619
16	D3	1/8	61.51	231	1.95	123	0.01585
8	D3	1/8	59.56	354		322	

FastCrack, 使用 SPOS 算法搜索模型各层的卷积类型。算法在单个 1080ti GPU 上花费 1 天, 最终从包含 1.68×10^6 种架构的超网中为模型各层搜索出的卷积块如图 7 所示, 将此卷积组合构建的模型称为 FastCrack-SPOS。在本文构建的混合数据集上对 FastCrack-SPOS 和搜索空间中分别用 6 种单一卷积块构建的模型进行训练和测试, 结果如表 2 所示。FastCrack-SPOS 的交并比、精

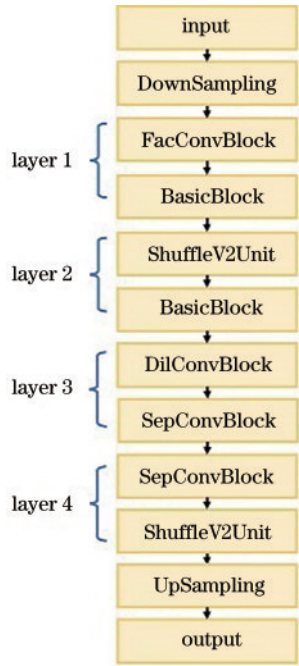


图 7 FastCrack-SPOS 模型结构

Fig. 7 Network structure of FastCrack-SPOS

确率和 F1 分数明显高于其他模型, 仅以 0.29×10^6 的参数数量达到 147 frame/s 的速度。

实验结果表明, 单路径神经架构搜索算法大幅减少了人工选择卷积类型所需的时间, 对包含 1.68×10^6 种架构的超网搜索只需 1 天, 若单独训练这些子网需要 1.68×10^6 GPU days, 显然难以实现。另一方面, FastCrack-SPOS 和由单一卷积块构建的模型的对比实验表明, SPOS 算法在裂缝分割任务中十分有效, 通过训练包含多种卷积块的超网, 能够自动为模型每层挑选出合适卷积块, 在目标数据集上构建出高性能轻量级架构, 说明将神经架构搜索与裂缝分割相结合是完全可行的, 为未来向裂缝分割领域引入更多的自动机器学习技术带来了先例。

4.3 和其他模型对比

确定模型宽度、深度、下采样倍率和各层卷积块类型共 4 种超参数后, 对 FastCrack-SPOS 模型和其他系列模型进行对比。实验结果如表 3 所示: FastCrack-SPOS 取得了最大的交并比和 F1 分数, 且具有最快的推理速度和最少的参数量; 在交并比、精确率、召回率、F1 分数指标上显著高于 DeepCrack、HED 等现有检测模型, 参数量与它们相比减少 95% 左右; 而 UNet、SegNet、DeeplabV3+ 等常见的语义分割模型虽拥有较大的参数量, 但交并比和 F1 分数仍不及 FastCrack-SPOS, 且速度仅为 FastCrack 的 1/15; DFANet、ESPNet、FastSCNN 等实时语义分割模型的速度均超过 30 frame/s, 且 ESPNet 参数量仅为 0.364×10^6 , 但在精度上和 FastCrack-SPOS 相比仍存在一定差距。

表 2 FastCrack-SPOS 模型的有效性分析

Table 2 Validity analysis of FastCrack-SPOS model

Method	IoU / %	Precision / %	Recall / %	F1 / %	FLOPs / 10^9	$S_{FPS} / (\text{frame} \cdot \text{s}^{-1})$	Params / 10^6
FastCrack-Basic	62.83	75.42	79.00	77.17	12.64	152	0.61558
FastCrack-Fac	61.64	74.91	77.67	76.27	9.19	163	0.438
FastCrack-Sep	61.26	74.77	77.22	75.97	4.42	125	0.208
FastCrack-Dil	61.44	75.50	76.74	76.12	4.42	126	0.208
FastCrack-Shuffle	59.81	74.20	75.51	74.85	3.62	175	0.169
FastCrack-Ghost	61.62	73.91	78.76	76.25	3.98	122	0.1855
FastCrack-SPOS	62.88	76.07	78.38	77.21	7.08	147	0.291

表 3 FastCrack-SPOS 和常见语义分割模型的对比

Table 3 Comparison between FastCrack-SPOS and common semantic segmentation models

Model	IoU / %	Precision / %	Recall / %	F1 / %	FLOPs / 10^9	$S_{FPS} / (\text{frame} \cdot \text{s}^{-1})$	Params / 10^6
DeepCrack	23.23	31.24	54.41	39.69	2190.8	1.6	30.91
HED	59.91	73.04	76.92	74.93	321.5	10.2	14.72
UNet	58.99	74.49	73.93	74.21	669.75	5.1	31.04
SegNet	62.46	77.22	76.57	76.89	1205.34	4.46	16.31
DeeplabV3+	62.56	73.41	80.89	76.97	360.25	7.7	59.34
ESPNet	61.89	75.22	77.73	76.46	88	99.2	0.364
DFANet	48.21	62.77	67.52	65.06	7.18	47	2.177
FastSCNN	61.74	76.86	75.83	76.34	3.48	132	1.136
FastCrack-SPOS	62.88	76.07	78.38	77.21	7.08	147	0.291

各模型对裂缝的分割结果如图 8 所示。从第 1 列和第 5 列可看出：在面对大范围的油漆干扰时，各模型对细小裂缝的分割结果普遍不高；DeepCrack 对图像细节噪声很敏感，存在诸多误检；而 HED、DFANet 等模型不同程度地受到油漆干扰；UNet 和 SegNet 对细小裂缝检测结果稍好，但仍存在诸多漏检，且检出的裂缝不连续；所提模型漏检、误检情况少，对细小裂缝的

检出程度更高。从第 2 列和第 3 列可以看出：在背景颜色干扰较大时，DeepCrack 漏检严重；HED、DFANet 存在大面积误检；SegNet、ESPNet、FastSCNN 对细小裂缝的分割同样存在断裂、不连续的情况；FastCrack-SPOS 分割较为精准，可较好检出裂缝细节信息。从第 4 列可以看出，对于背景简单没有噪声的图像，所有模型普遍可达到较好的分割效果。从第 6 列可看出，

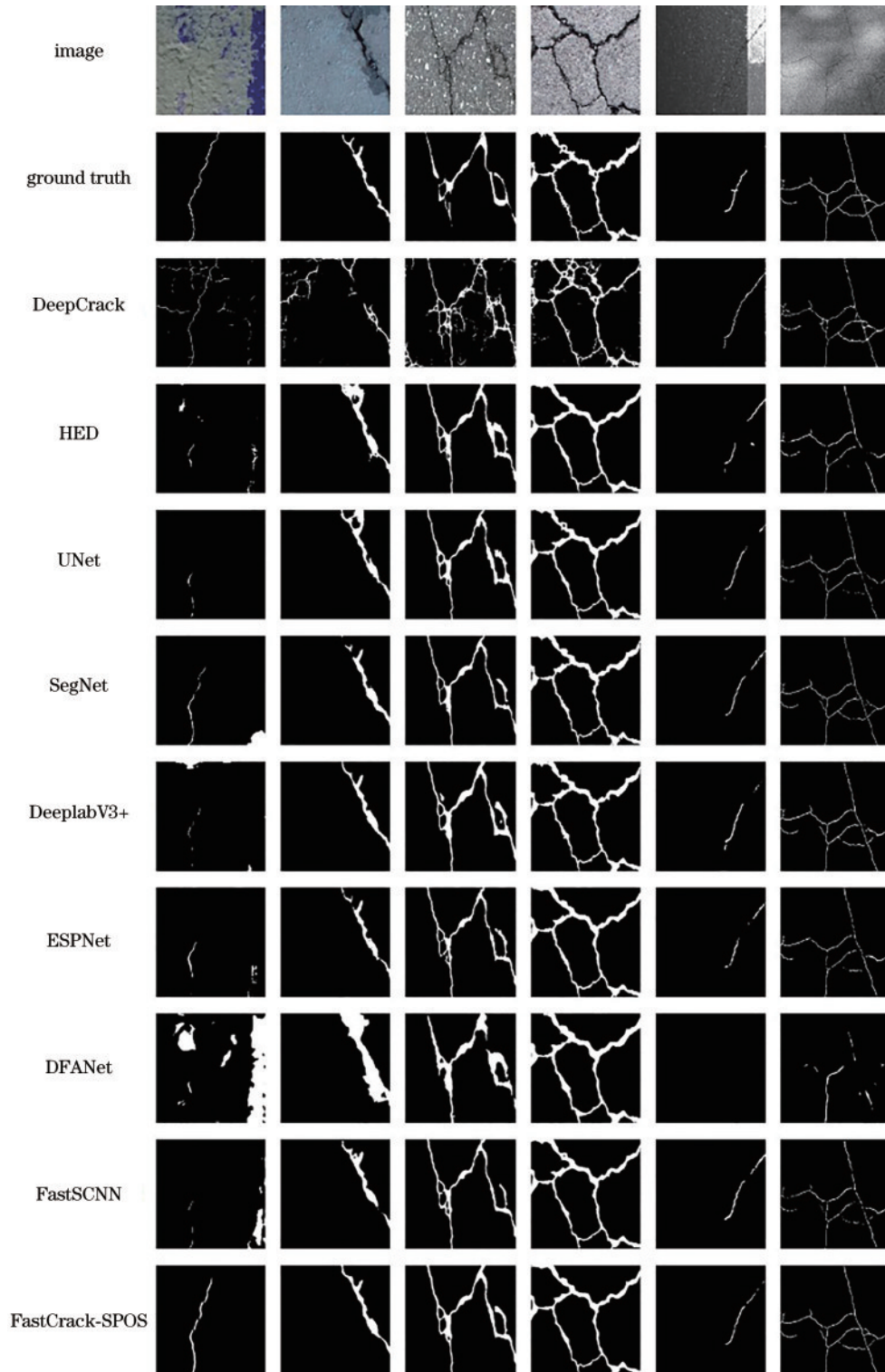


图 8 裂缝分割结果

Fig. 8 Crack segmentation results

FastCrack-SPOS 能够较好地抵抗阴影干扰。

实验结果表明:相比于常见的裂缝分割、边缘检测及语义分割模型,设计的架构超参数挑选框架所构建出的 FastCrack-SPOS 不仅具有更好的裂缝分割性能,且推理速度远快于其他模型,对 1024×1024 输入的处理速度达 147 frame/s,在速度和精度之间取得了平衡;同时参数量只有 0.29×10^6 ,十分轻量,易于后续将模型部署到移动设备上,具备较高的实际应用价值。

4.4 模型泛化性探究

为了进一步验证 FastCrack-SPOS 模型的泛化性,增加了额外的消融实验。对在路面裂缝数据集上训练好的所有模型在开源的隧道裂缝数据集上进行测试,以验证所提模型的泛化能力。隧道裂缝数据集^[49]中包含了更为复杂的油漆、污渍干扰及不同的背景。实验结果如表 4 所示。可以看到,在隧道裂缝数据集的评测上, FastCrack-SPOS 仍然能够保持最高的交并比和 F1 分数,综合性能最好,说明利用所设计的架构超参数挑选框架设计的 FastCrack-SPOS 不仅在路面裂缝数据集上具有最佳的性能,而且在隧道裂缝的分割上同样优于其他模型,具有一定的泛化性。

表 4 FastCrack-SPOS 在隧道裂缝数据集上的泛化性实验
Table 4 Generalization experiment of FastCrack-SPOS on tunnel crack dataset

Method	IoU / %	Precision / %	Recall / %	F1 / %
DeepCrack	12.87	20.07	32.57	21.21
HED	20.27	25.24	50.77	33.72
UNet	19.88	41.20	27.76	33.17
SegNet	20.18	28.92	40.04	33.59
DeeplabV3+	21.91	28.01	50.15	35.94
ESPNet	20.38	26.43	47.10	33.86
DFANet	9.53	11.14	39.80	17.41
FastSCNN	20.68	28.89	42.13	34.28
FastCrack-SPOS	23.64	38.52	43.21	40.73

5 结 论

针对实时裂缝分割任务,为平衡模型精度和速度,设计了一套架构超参数挑选框架,提出了实时路面裂缝分割模型 FastCrack-SPOS。首先,对模型宽度、深度、下采样倍率 3 个超参数进行了一系列组合实验,发现 $1/8$ 下采样倍率更适合深层模型;深度和宽度的增加会提高模型精度,但增加幅度越高,精度提升越慢,过大的宽度还会导致精度的下降。实际选择时需考虑参数量增加带来的额外计算成本,为裂缝分割模型的轻量化工作提供参考。然后,对于卷积块类型这一超参数,首次向裂缝分割任务中引入神经架构搜索技术,自动为模型每层挑选合适类型的卷积,大幅缩短人工设计裂缝分割网络所需时间,针对性构建出了高

性能轻量级裂缝分割模型。实验结果表明, FastCrack-SPOS 模型的交并比达 62.88%,参数量仅有 0.29×10^6 ,相比现有模型减少 95% 左右,推理速度达到 147 frame/s,具备较高的实用价值,便于部署到无人机、手机等便携式计算设备中。

参 考 文 献

- [1] Mei Q P, Gül M. A cost effective solution for pavement crack inspection using cameras and deep neural networks [J]. Construction and Building Materials, 2020, 256: 119397.
- [2] Miao P Y, Srimahachota T. Cost-effective system for detection and quantification of concrete surface cracks by combination of convolutional neural network and image processing techniques[J]. Construction and Building Materials, 2021, 293: 123549.
- [3] 陈晓冬, 艾大航, 张佳琛, 等. Gabor 滤波融合卷积神经网络的路面裂缝检测方法[J]. 中国光学, 2020, 13(6): 1293-1301.
Chen X D, Ai D H, Zhang J C, et al. Gabor filter fusion network for pavement crack detection[J]. Chinese Optics, 2020, 13(6): 1293-1301.
- [4] Liu Y H, Yao J, Lu X H, et al. DeepCrack: a deep hierarchical feature learning architecture for crack segmentation[J]. Neurocomputing, 2019, 338: 139-153.
- [5] Oliveira H, Correia P L. Automatic road crack detection and characterization[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(1): 155-168.
- [6] Faghih-Roohi S, Hajizadeh S, Núñez A, et al. Deep convolutional neural networks for detection of rail surface defects[C]//2016 International Joint Conference on Neural Networks (IJCNN), July 24-29, 2016, Vancouver, BC, Canada. New York: IEEE Press, 2016: 2584-2589.
- [7] Zou Q, Cao Y, Li Q Q, et al. CrackTree: automatic crack detection from pavement images[J]. Pattern Recognition Letters, 2012, 33(3): 227-238.
- [8] Guan H Y, Li J, Yu Y T, et al. Iterative tensor voting for pavement crack extraction using mobile laser scanning data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(3): 1527-1537.
- [9] Zhao H L, Qin G F, Wang X J. Improvement of Canny algorithm based on pavement edge detection[C]//2010 3rd International Congress on Image and Signal Processing, October 16-18, 2010, Yantai, China. New York: IEEE Press, 2010: 964-967.
- [10] Li P, Wang C, Li S M, et al. Research on crack detection method of airport runway based on twice-threshold segmentation[C]//2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), September 18-20, 2015, Qinhuangdao, China. New York: IEEE Press, 2015: 1716-1720.
- [11] 李良福, 武彪, 王楠. 基于多分辨率网络的桥梁裂缝检测方法研究[J]. 激光与光电子学进展, 2021, 58(12): 1210004.
Li L F, Wu B, Wang N. Method for bridge crack

- detection based on multiresolution network[J]. *Laser & Optoelectronics Progress*, 2021, 58(12): 1210004.
- [12] 闫河, 赵其峰, 谢敏, 等. 自适应 Canny 算子 PCBA 元器件边缘检测及修复[J]. *光学学报*, 2021, 41(5): 0515003.
- Yan H, Zhao Q F, Xie M, et al. Edge detection and repair of PCBA components based on adaptive Canny operator[J]. *Acta Optica Sinica*, 2021, 41(5): 0515003.
- [13] Shi Y, Cui L M, Qi Z Q, et al. Automatic Road crack detection using random structured forests[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 17(12): 3434-3445.
- [14] Jiang W B, Liu M, Peng Y N, et al. HDCB-net: a neural network with the hybrid dilated convolution for pixel-level crack detection on concrete bridges[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(8): 5485-5494.
- [15] 李刚, 刘强伟, 万健, 等. 复杂背景下交错低秩组卷积混合深度网络的路面裂缝检测算法研究[J]. *激光与光电子学进展*, 2020, 57(14): 141031.
- Li G, Liu Q W, Wan J, et al. A novel pavement crack detection algorithm using interlaced low-rank group convolution hybrid deep network under a complex background[J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141031.
- [16] 刘荻, 张焱, 赵琰, 等. 基于特征重聚焦网络的多尺度近岸舰船检测[J]. *光学学报*, 2021, 41(22): 2215001.
- Liu D, Zhang Y, Zhao Y, et al. Multi-scale inshore ship detection based on feature re-focusing network[J]. *Acta Optica Sinica*, 2021, 41(22): 2215001.
- [17] 宋子壮, 杨嘉伟, 张东方, 等. 基于无监督域适应的低空海面红外目标检测[J]. *光学学报*, 2022, 42(4): 0415001.
- Song Z Z, Yang J W, Zhang D F, et al. Low-altitude sea surface infrared object detection based on unsupervised domain adaptation[J]. *Acta Optica Sinica*, 2022, 42(4): 0415001.
- [18] 毛莺池, 唐江红, 王静, 等. 基于 Faster R-CNN 的多任务增强裂缝图像检测方法[J]. *智能系统学报*, 2021, 16(2): 286-293.
- Mao Y C, Tang J H, Wang J, et al. Multi-task enhanced dam crack image detection based on Faster R-CNN[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(2): 286-293.
- [19] 赵亮, 胡杰, 刘汉, 等. 基于语义分割的深度学习激光点云三维目标检测[J]. *中国激光*, 2021, 48(17): 1710004.
- Zhao L, Hu J, Liu H, et al. Deep learning based on semantic segmentation for three-dimensional object detection from point clouds[J]. *Chinese Journal of Lasers*, 2021, 48(17): 1710004.
- [20] 张世宽, 吴清潇, 林智远. 焊缝图像中结构光条纹的检测与分割[J]. *光学学报*, 2021, 41(5): 0515002.
- Zhang S K, Wu Q X, Lin Z Y. Detection and segmentation of structured light stripe in weld image[J]. *Acta Optica Sinica*, 2021, 41(5): 0515002.
- [21] Zhang A, Wang K C P, Li B X, et al. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2017, 32(10): 805-819.
- [22] Yang X C, Li H, Yu Y T, et al. Automatic pixel-level crack detection and measurement using fully convolutional network[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2018, 33(12): 1090-1109.
- [23] Shim S, Cho G C. Lightweight semantic segmentation for road-surface damage recognition based on multiscale learning[J]. *IEEE Access*, 2020, 8: 102680-102690.
- [24] Choi W, Cha Y J. SDDNet: real-time crack segmentation [J]. *IEEE Transactions on Industrial Electronics*, 2020, 67(9): 8016-8025.
- [25] Shim S, Kim J, Lee S W, et al. Road surface damage detection based on hierarchical architecture using lightweight auto-encoder network[J]. *Automation in Construction*, 2021, 130: 103833.
- [26] Guo Z C, Zhang X Y, Mu H Y, et al. Single path one-shot neural architecture search with uniform sampling [M]//Vedaldi A, Bischof H, Brox T, et al. *Computer vision-ECCV 2020. Lecture notes in computer science*. Cham: Springer, 2020, 12361: 544-560.
- [27] Cai H, Gan C, Wang T Z, et al. Once for all: train one network and specialize it for efficient deployment[EB/OL]. (2019-08-26)[2021-05-04]. <https://arxiv.org/abs/1908.09791>.
- [28] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04)[2021-05-04]. <https://arxiv.org/abs/1409.1556>.
- [29] Zhang A, Wang K C P, Fei Y, et al. Deep learning-based fully automated pavement crack detection on 3D asphalt surfaces with an improved CrackNet[J]. *Journal of Computing in Civil Engineering*, 2018, 32(5): 4018041.
- [30] Chen X, Xie L X, Wu J, et al. Progressive differentiable architecture search: bridging the depth gap between search and evaluation[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 1294-1303.
- [31] Hong Y D, Pan H H, Sun W C, et al. Deep dual-resolution networks for real-time and accurate semantic segmentation of road scenes[EB/OL]. (2021-01-15)[2021-05-06]. <https://arxiv.org/abs/2101.06085>.
- [32] Tan M X, Le Q V. EfficientNet: rethinking model scaling for convolutional neural networks[EB/OL]. (2019-05-28)[2021-05-06]. <https://arxiv.org/abs/1905.11946>.
- [33] Zagoruyko S, Komodakis N. Wide residual networks [C]//Proceedings of the British Machine Vision Conference 2016, September 19-22, 2016, York, UK. London: British Machine Vision Association, 2016: 87.1-87.12.
- [34] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [35] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on

- Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1800-1807.
- [36] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2021-05-06]. <https://arxiv.org/abs/1704.04861>.
- [37] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[EB/OL]. (2015-11-23)[2021-05-06]. <https://arxiv.org/abs/1511.07122>.
- [38] Ma N N, Zhang X Y, Zheng H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design [M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11218: 122-138.
- [39] Han K, Wang Y H, Tian Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1577-1586.
- [40] Yang F, Zhang L, Yu S J, et al. Feature pyramid and hierarchical boosting network for pavement crack detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(4): 1525-1535.
- [41] Zou Q, Zhang Z, Li Q Q, et al. DeepCrack: learning hierarchical convolutional features for crack detection[J]. IEEE Transactions on Image Processing, 2019, 28(3): 1498-1512.
- [42] Xie S N, Tu Z W. Holistically-nested edge detection [C]//2015 IEEE International Conference on Computer Vision, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 1395-1403.
- [43] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [44] 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.
- [45] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 833-851.
- [46] Mehta S, Rastegari M, Caspi A, et al. ESPNet: efficient spatial pyramid of dilated convolutions for semantic segmentation[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11214: 561-580.
- [47] Poudel R P K, Liwicki S, Cipolla R. Fast-SCNN: fast semantic segmentation network[C]//30th British Machine Vision Conference 2019, September 9-12, 2019, Cardiff, UK. London: British Machine Vision Association, 2019.
- [48] Li H C, Xiong P F, Fan H Q, et al. DFANet: deep feature aggregation for real-time semantic segmentation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 9514-9523.
- [49] Ren Y P, Huang J S, Hong Z Y, et al. Image-based concrete crack detection in tunnels using deep fully convolutional networks[J]. Construction and Building Materials, 2020, 234: 117367.