

复杂背景下基于图像处理的桥梁裂缝检测算法

李良福¹, 孙瑞赞^{1*}

1. 陕西师范大学计算机科学学院, 西安 710119

摘要: 传统的桥梁裂缝检测算法通常假设裂缝图像的背景简单且无任何障碍物, 但是, 在采集过程中会经常遇到复杂的道路场景, 导致使用现有的桥梁裂缝检测算法不能准确提取裂缝。针对此问题, 本文提出了一种基于图像处理且适用于复杂背景的桥梁裂缝检测算法。首先, 根据深度卷积生成式对抗网络原理, 提出并利用桥梁裂缝图像生成模型对数据集进行扩增; 然后, 针对裂缝特征构建基于语义分割的桥梁裂缝图像分割模型; 最后, 利用桥梁裂缝图像分割模型提取高分辨率裂缝图像中的裂缝。实验结果表明, 与现有算法相比, 本文提出的算法更适用于复杂背景下的桥梁裂缝检测与提取。

关键词: 图像处理; 复杂背景; 桥梁裂缝检测; 深度卷积生成式对抗网络; 语义分割

中图分类号: TP391.9 **文献标识码:** A

Bridge crack detection algorithm based on image processing in complex background

Li Liangfu¹, Sun Ruiyun^{1*}

1. Shaanxi Normal University, School of Computer Science, Xi'an 710119 China

Abstract: The traditional bridge crack detection algorithms usually assumed that the quality of the crack image was perfect, the road surface was clean and there was no interfering material. However, in the actual acquisition process, various complex road scenes will be encountered. And these factors had led to the existed bridge crack detection algorithms could not achieve ideal results. In order to solve the problem, a bridge crack detection algorithm which is based on image processing and suitable for complex scenes was proposed. Firstly, according to the principle of the deep convolutional generative adversarial network, the bridge crack image generative model was proposed and used to amplify the dataset. Secondly, a bridge crack image segmentation model based on semantic segmentation was constructed. Finally, bridge crack image segmentation model was used to extract the bridge crack in the high-resolution images. The experimental results show that the algorithm proposed by this paper has better detection effect and stronger generalization ability in the complex road scenes compared with the existed algorithms.

Key words: Image processing; complex background; bridge crack detection; deep convolutional generative adversarial network; semantic segmentation

OCIS codes: 100.2000; 100.4996; 100.1830; 150.1135

1 引言

交通运输是经济发展的基本需要和先决条件, 现代社会的生存基础和文明标志, 工业发展的基础设施和重要纽带, 关系着国家经济的发展, 承载着社会进步的命脉^[1-2]。据《2017年国民经济和社会发展统计公报》^[3]统计, 在2017年, 我国新改建高速公路里程6796公里, 新建高速铁路投产里程2182公里, 建成了世界上最大的高速公路网、高铁运营网。由于我国重建设而忽养护的公路桥梁管理模式, 使得加快了桥梁与道路老化的速度, 容易造成安全隐患^[4]。已有研究表明, 混凝土桥梁由于直接应力与次生应力而产生桥梁裂缝, 而桥梁裂缝是桥梁损伤状况的外在表现, 是桥梁结构达到承载力极限的标志。因此, 对混凝土桥梁裂缝的检测十分重要。

采取有效手段对桥梁裂缝进行检测^[5]对确保公共交通的安全和正常运行起着十分重要

收稿日期: 年-月-日; 修回日期: 年-月-日; 录用日期: 年-月-日

基金项目: 国家自然科学基金项目(61573232, 61401263)

*E-mail: longford@xjtu.edu.cn; **E-mail: 984789463@qq.com

的作用,长期以来已受到了国内外学术界、工程界的广泛关注,并且取得了一些优秀的研究成果。Oh H 等人提出了迭代的阈值分割算法^[6],该算法的缺点是阈值需要人工设置;孙亮等人提出了基于自适应阈值 Canny 算法的裂缝检测方法^[7],该算法解决了需要人工设置阈值的缺陷,但没有考虑由于光照不均与噪声造成的对裂缝识别的影响,很难取得稳定的效果;Talab A M A 等人提出通过 Otsu 与多重滤波结合的算法提取裂缝^[8],该算法改善了由于光照不均造成的影响,但是单一的阈值使得该算法并不适用于所有的图像。

近几年,机器学习与深度学习成为人工智能方向飞速发展的热点,许多学者成功的将桥梁裂缝检测与之结合。L. Zhang 等人利用深度卷积神经网络实现对桥梁裂缝的提取与检测^[9],但是该算法检测到的裂缝宽度与真实裂缝宽度相差极大;Fu-Chen Chen 等人使用卷积神经网络和朴素贝叶斯数据融合的 NB-CNN 网络进行裂缝检测^[10],该算法的优点是可以测到微小的裂缝,但是却仅仅只能检测到裂缝的位置,并不能实现对裂缝的提取;Yong Shi 等人利用随机结构森林算法对裂缝进行自动提取^[11],但是当该算法应用于复杂背景时,检测效果欠佳。以上算法取得的良好实验效果均建立在裂缝图像的背景简单且不存在任何障碍物的情形之下,低估了桥梁路面的复杂程度。因此,本文提出了一种基于图像处理且适用于复杂背景的桥梁裂缝检测算法。具体做法为:首先,通过桥梁裂缝图像生成模型扩增数据集,然后,利用扩增的数据集训练桥梁裂缝图像分割模型,最后,利用桥梁裂缝图像分割模型检测高分辨率裂缝图像中的裂缝。

2 数据集的扩增

本文提出的复杂背景下桥梁裂缝检测算法属于深度学习领域,众所周知,深度学习中的网络模型一般涉及的权重参数众多,在训练样本不足的情况下很容易出现欠拟合的情况。因此,我们必须建立有充足样本的数据集。但是,到目前为止,全球没有公开的、适用于复杂背景研究的桥梁裂缝数据集,如果直接人工的方式大量采集,不仅工作量极大,而且效率极低。目前,由 Alec Radford 等人提出的深度卷积生成式对抗网络(DCGAN)^[12]是人工智能领域的一个研究热点,已经被成功应用于生成全新的图像。但文献[12]中的 DCGAN 模型适用于生成 64*64 大小、类似 cifar-10、SVHN 数据集的图像,对于 256*256 大小、具有线性拓扑结构的桥梁裂缝图像的生成不能取得满意的效果,甚至无法生成。因此,提出了一种基于深度卷积生成式对抗网络的桥梁裂缝图像生成模型 (Bridge-Crack-Image-Generative-Model),简称 BCIGM。

2.1 采集数据扩增

使用 BCIGM 需要几千张图像作为训练样本,如果全部人工采集仍然是一个严峻的问题,因此,数据集的扩增分两步完成,首先,对少量的数据通过图像的几何变换、空间滤

波、线性变换这三类图像处理方法进行扩充，然后，使用 BCIGM 生成大量裂缝图像。经过第一步数据集扩充之后，人工手动的挑选出 9362 幅图像作为 BCIGM 的训练集，部分经过图像处理算法扩充之后的图像如图 1 所示，其中第一列为原图，第二至五列依次为经过水平翻转、垂直翻转、线性变换与空间滤波变换后的图像。

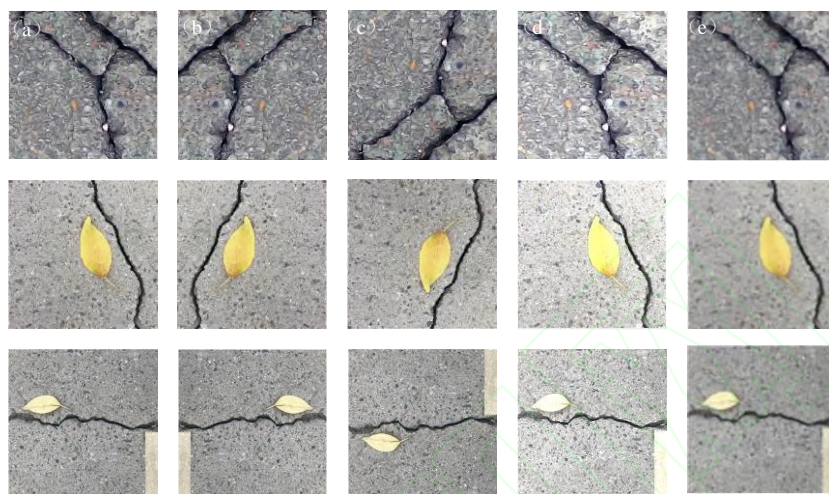


图 1 桥梁裂缝图像数据集扩增示意图 (a) 原图; (b) 水平翻转; (c) 垂直翻转; (d) 线性变换; (e) 空间滤波变换

Fig 1. Amplification diagram of bridge crack dataset (a) Original image; (b) horizontal flip; (c) vertical flip; (d) linear transformation; (e) spatial filtering transformation

2.2 桥梁裂缝图像生成模型

2.2.1 桥梁裂缝图像生成模型的原理

桥梁裂缝图像生成模型的原理与生成式对抗网络的原理一样，由具有对抗关系的生成模型(G)与判别模型(D)组成。其中，生成模型(G)的作用是通过学习已有真实数据样本的概率分布，生成尽可能服从真实样本分布的样本 $G(z)$ ，判别模型(D)的作用是判断输入是来自真实样本还是来自生成样本 $G(z)$ ，实质为一个二分类模型。训练 BCIGM 的过程就是不断交替更新 G 和 D 的过程，当生成网络 G 固定时，优化判别网络，使输入真实数据样本时输出尽可能趋向 1，输入生成数据样本时输出尽可能趋向 0；当判别网络(D)固定时，优化生成网络(G)，使生成样本经过判别网络(D)后输出高概率，当双方达到纳什均衡时，生成模型(G)生成了判别模型(D)无法判断是否是真实样本的生成样本。上述过程表示为如下(1)式：

$$\min_G \max_D V(D, G) = E_{x \sim P_{data}(x)} [\log D(x)] + E_{z \sim P_{z(z)}} [\log (1 - D(G(z)))] \quad (1)$$

其中， z 表示输入生成网络的噪声， $G(z)$ 表示生成样本， x 表示真实数据样本， $D(x)$ 表示 x 通过判别网络后判断为真实样本的概率， $D(G(z))$ 表示生成样本通过判别网络后判断为真实样本的概率。 $P_{data}(x)$ 和 $P_{z(z)}$ 分别表示真实样本数据的概率分布和初始噪音数据的概率分布， $E(.)$ 表示计算期望值。

2.2.2 生成模型

生成模型的作用是将输入的噪声通过上采样生成 $256 \times 256 \times 3$ 的裂缝图像，主要由转置

卷积构成。具体过程如下：首先输入 100 维的噪声，然后，通过全连接层后进行维度转换，转换成 1024 个 8*8 的特征图，最后，通过五个卷积核为 5*5，步幅为 2，卷积核数目依次为 512、256、128、64、3 的转置卷积，且除最后一个转置卷积外，其余转置卷积后均使用 SeLU 激活函数^[13]。模型如图 2 所示：

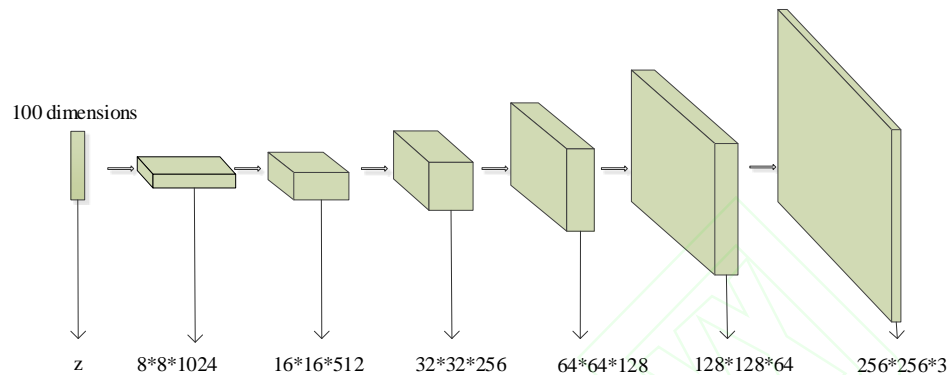


图 2 生成模型

Fig 2. The generative model

一般来说，转置卷积是从深而窄的层次延展为窄而深的层次。由于在生成模型中，输入是由随机噪声转换成的深而窄的特征图，需要通过转置卷积实现上采样，从而生成 256*256 的彩色裂缝图像。之所以选择转置卷积，是因为它相比于紧邻上采样、双样条采样、三次样条采样等手工选择采样方法，我们更需要一种能够自动优化的上采样方法，而且转置卷积已经在 DCGAN 中得到了良好的应用效果。之所以选择 5*5 的卷积核，是因为在深度学习中，常用的卷积核为 3*3、5*5、7*7，且卷积核越小，参数量越低，复杂度越小，但是，如果在桥梁裂缝检测中选择 3*3 大小的卷积核，不仅不利于提取裂缝的结构信息，而且还会受到噪声的严重影响，因此，我们选择相对较大的、5*5 像素分辨率的卷积核。步幅设置为 2 是因为转置卷积的步幅定义了输出层的大小，且在“相同”的填充和步幅为 2 时，输出特征的大小将是输入层的两倍。卷积核的数量与特征图大小的设置综合考虑了模型的复杂度与生成裂缝图像的真实程度。卷积层数目代表了模型的复杂度，层数越多精确度往往更高，但同样消耗计算资源也更高。因此，通过实验得出上述设置是在模型复杂度最低的条件下，生成了更真实的桥梁裂缝图像。

BCIGM 选择 SeLU 激活函数代替 Relu 激活函数和 Batch Normalization 的原因有：首先，SeLU 引入了自归一化的属性，使神经元激励值可以自动地收敛到零均值和单位方差，并且即使是存在噪声和扰动的环境下，通过许多层的前向传播后还是将收敛到零均值和单位方差。此外，对于不逼近单位方差的激励值，其方差存在上确界和下确界，因此梯度消失和梯度爆炸不可能出现，这样大大增加了 BCIGM 的稳定性。

2.2.3 判别模型

判别模型的作用是通过特征提取判断输入样本是否是真实样本，由全卷积网络构成。具

体过程如下：首先，输入 256*256*3 的样本图像，包括真实样本与生成样本两类，然后，经过 6 个卷积核为 5*5，步幅为 2、卷积核数目依次为 64、128、256、512、1024、2048 的卷积，其次，通过 1 个卷积核为 1*1 的卷积，最后，通过 Sigmoid 激活函数，映射出输入样本概率。之所以加入 1*1 的卷积核是因为它可以在不改变特征图大小的情况下进行降维，减少参数的个数，从而降低计算时间。具体模型如图 3 所示：

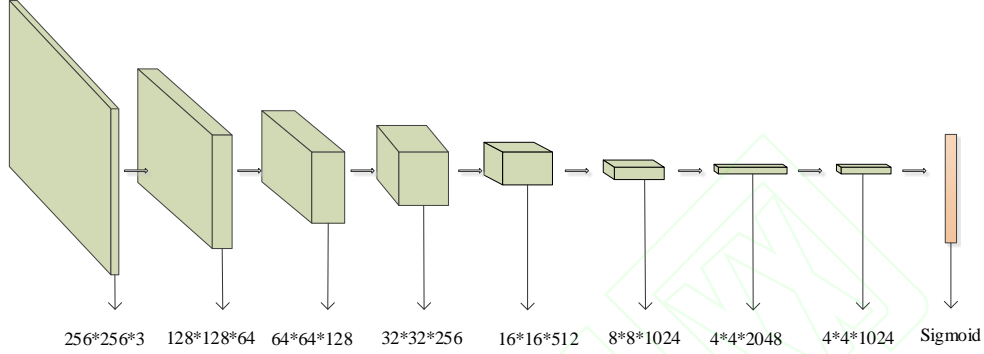


图 3 判别模型

Fig 3. The discriminant model

3 基于语义分割的桥梁裂缝图像分割模型

图像语义分割^[14-15]作为人工智能领域的重要分支，是计算机视觉中关于图像理解的重要一环，SegNet^[16]、FCN^[17]、DeepLab^[18]等网络模型被证明在多目标分类中取得了先进的效果。众所周知，在一定条件下，网络的深度越深，提取到的特征越精确，检测效果也随之越好，但是在实际应用中发现网络越深时越容易出现梯度扩散的问题。在 2017 年，由 Gao Huang 等人提出的密集卷积网络 Dense Convolutional Network^[19]解决了这个问题。因此，本文提出一种基于 Dense Convolutional Network 的裂缝分割模型。

3.1 密集卷积网络

密集卷积网络(Dense Convolutional Network , DenseNet)是一种具有密集连接的卷积神经网络。在该网络中，任何两层之间都有直接的连接，也就是说，网络每一层的输入都是前面所有层输出的并集，而该层所学习的特征图也会被直接传给其后面所有层作为输入。DenseNet 不仅可以高效利用特征图，并且可以在网络很深时有效解决梯度消失问题。可表示为如下(2)式：

$$X_l = H_l([X_0, X_1, \dots, X_{l-1}])。 \quad (2)$$

式中 l 表示层数， X_l 代表 l 层的输出， $[X_0, X_1, \dots, X_{l-1}]$ 表示将 0 到 $l-1$ 层的输出特征图按深度链接，即 Filter Concatenation，使用 Filter Concatenation 的好处在于防止了由于层数增多带来的计算资源的爆炸性需求，从而使网络的宽度和深度均可扩大。 $H_l(.)$ 由 Batch Normalization，ReLU 激活函数和卷积操作组成。由于 Filter Concatenation 要求特征图

X_0, X_1, \dots, X_{i-1} 的尺寸相同, 而池化操作会改变特征图尺寸且不可或缺, 因此, 提出 DenseBlock, 使得 DenseBlock 内的特征图尺寸相同, 每经过 DenseBlock 的一层, 特征图数量增加 k , 通过 k 可以控制网络的宽度。如图 4 所示为包括 4 层的 DenseBlock。

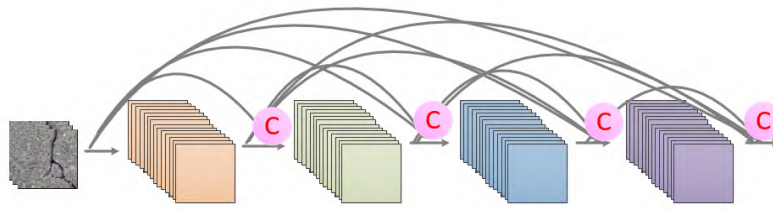


图 4 4 层 DenseBlock 示意图

Fig 4. The diagram of 4-layer DenseBlock

3. 2 桥梁裂缝图像分割模型的构建

文献[20]提出用于语义分割的 FC-DenseNet103 模型, 并在 CamVid 数据集中取得了满意的效果, 但是当应用于复杂背景下的桥梁裂缝提取时效果不佳, 且参数众多、训练模型时间较长。因此, 针对 FC-DenseNet103 模型进行了改进, 提出了适用于复杂背景的桥梁裂缝图像分割模型(Bridge-Crack-Image-Segmentation-Model), 简称 BCISM。

本文选择 FC-DenseNet103 模型作为基础模型, 原因有以下几点: 第一, 桥梁裂缝为线性拓扑结构且图像背景纹理复杂, 语义分割作为一种针对像素级别的精确分割, 可以对每个像素进行预测, 用于桥梁裂缝的检测有充足的理论支撑; 第二, 采用 DenseNet 的网络结构不仅能够使得模型的参数使用效率提高, 而且由于到每个特征图都有短路径连接使其实现了隐含的深层监督, 此外, 特殊的设计使所有层都可以很容易的访问到它们前面的所有层, 这就使得前面计算的特征图信息可以很容易的被重用。第三, FC-DenseNet103 网络模型中 103 层的模型结构可以使得精确度大幅提高, 并且有效避免了梯度消失, 这是大部分语义分割模型无法实现的。

BCISM 一共由 74 个卷积层构成, 网络模型具体包括由 DenseBlock 和 Transition Down 组成的下采样路径, 以及由 DenseBlock 和 Transition Up 组成的上采样路径, 以及 Softmax 函数。其中, DenseBlock 依次由 4、5、7、10、12、10、7、5、4 个层构成, 每个层由 Batch Normalization、ReLU 激活函数、3*3 卷积和 Dropout 构成, Dropout 是指在深度学习网络的训练过程中, 对于神经网络单元, 按照一定的概率将其暂时从网络中丢弃, 以使每一个 batch 都在训练不同的网络, 网络中 Dropout=0.2。Transition Down 的作用是减少特征图的空间维度, 由 Batch Normalization、ReLU 激活函数、1*1 卷积与 2*2 池化操作组成, 其中 1*1 的卷积用于保存特征图的数量, 2*2 的池化操作用于降低特征图的分辨率, 以此来补偿由于网络层数大幅增加造成的特征图数量的线性增长。Transition Up 由一个转置卷积构成, 作用是恢复输入图像的空间分辨率, 转置卷积仅对最后一个 DenseBlock 的特征图使用, 原因是由于最后一个 DenseBlock 综合了所有之前 DenseBlock 的信息。Softmax 函数的作用是输出裂缝

与非裂缝的概率。BCISM 的网络结构参数如表 1 所示：

表 1 BCISM 的网络结构参数
Tab.1 Network structure parameters of BCISM

Name of the layer	Size of the kernel	Stride	Size of the output feature map	Number of the feature map
Input Layer	---	---	256*256	3
Convolution	Convolution 5*5	1	256*256	48
DenseBlock	Convolution [3*3]*4	1	256*256	96
TransitionDown	Convolution 1*1	1	256*256	96
	Max pooling 2*2	2	128*128	96
DenseBlock	Convolution [3*3]*5	1	128*128	156
TransitionDown	Convolution 1*1	1	128*128	156
	Max pooling 2*2	2	64*64	156
DenseBlock	Convolution [3*3]*7	1	64*64	240
TransitionDown	Convolution 1*1	1	64*64	240
	Max pooling 2*2	2	32*32	240
DenseBlock	Convolution [3*3]*10	1	32*32	360
TransitionDown	Convolution 1*1	1	32*32	360
	Max pooling 2*2	2	16*16	360
DenseBlock	Convolution [3*3]*12	1	16*16	504
TransitionUp	Deconvolution 3*3	2	32*32	504
DenseBlock	Convolution [3*3]*10	1	32*32	624
TransitionUp	Deconvolution 3*3	2	64*64	624
DenseBlock	Convolution [3*3]*7	1	64*64	444
TransitionUp	Deconvolution 3*3	2	128*128	444
DenseBlock	Convolution [3*3]*5	1	128*128	300
TransitionUp	Deconvolution 3*3	2	256*256	300
DenseBlock	Convolution [3*3]*4	1	256*256	204
Convolution	Convolution 1*1	1	256*256	2
Softmax	---	---	---	---

3.3 高分辨率图像裂缝检测

在实际应用中桥梁裂缝图像分辨率都比较大，常用为 512*512、1024*1024、甚至 2048*2048 分辨率的图像，但是，深度学习网络中一般使用的输入图像分辨率都较小，因为如果将网络设计为高分辨率输入，会由于数据量庞大造成训练时间很久，甚至由于网络模型太深、设计不当而造成模型无法收敛，导致无法得到理想的输出，因此，本文提出了针对高分辨率裂缝图像的检测算法。具体做法为：首先，用滑动窗口算法对高分辨率图像顺次进行裁剪，并对裁剪图像块按序标号；然后，将图像块依次输入 BCISM 进行裂缝检测并按原序依次输出；最后，将输出图像按原顺序拼接。如图 5 所示为对 512*512 的裂缝图像进行检测的示意图，将原图顺次裁剪为 (a)(b)(c)(d)，依次输入 BCISM 后按原序输出为 (a1)(b1)(c1)(d1)，最后拼接为原图的检测结果。

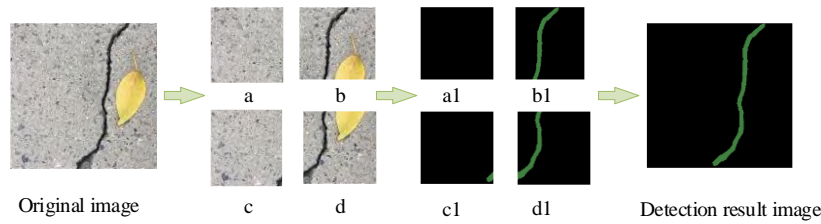


图 5 高分辨率图像检测示意图

Fig 5.Detection diagram of high-resolution image

4 实验结果与分析

4.1 数据集

本文一共采集了 1183 张桥梁裂缝图像，分为背景简单的图像、背景中含有障碍物的图像、背景中含有大面积污渍的图像三类。其中，背景简单的图像具体指裂缝图像中仅含有宽度较大的桥梁裂缝，其背景无任何干扰物；含有障碍物的图像具体指在裂缝图像的背景中包含各种各样的障碍物，本文实验中包括的障碍物类型有落叶、车道线、阴影、钉子、油漆、污渍等；背景中含有大面积污渍的图像具体指背景中存在大面积的污渍，其与第二类中含有污渍障碍物的区别为污渍所占面积比例的不同，此类中强调的是大面积。采集的裂缝图像原始分辨率为 2448×3264 ，为了便于在后续算法中使用，对原始图像进行了一系列处理：首先，将采集来的图像短边缩放到 2048 大小，裁出中央 2048×2048 的区域；然后，将 2048×2048 分辨率的图像下采样为 1024×1024 分辨率的图像；其次，将 1024×1024 的图像采用滑动窗口算法，不重叠裁剪为 16 幅 256×256 大小的图像；最后，通过第 2 章所述进行数据集扩充，最终数据集中的各种类型图像所占的比例如表 2 所示。

表 2 各类型图像在数据集中的占比

Tab.2 The proportion of each type of image in the dataset

Type of the picture	Simple background	Background with obstacles	Background with large areas of stains
Number of the picture	10449	13275	5710
The proportion (%)	35.5	45.1	19.4

4.2 实验环境

本文算法的程序基于主流的深度学习开源框架 TensorFlow,使用 Python 语言开发；实验的硬件环境为 Intel i7 处理器，NVIDIA GeForce GTX 1070 显卡；软件环境为 Ubuntu 16.04 LTS 操作系统。

4.3 对比试验

为了验证本文提出算法的有效性和准确性，本文分别设计了五组对比实验进行验证。其中，第一组实验用于验证 DCGAN 模型与 BCIGM 对于生成 256×256 的裂缝图像效果的影响。该实验包含三个小对比实验，实验 1 为 DCGAN 模型与 BCIGM 生成裂缝图像的可视化对比；实验 2 为 BCIGM 采用 SeLU 激活函数与采用 ReLU 激活函数生成裂缝图像的可视化对比。这两组实验选取前 4 张生成图像作为代表。实验 3 为无 1×1 卷积核的 BCIGM 与有 1×1 卷积核的 BCIGM 训练网络时每个 batch 的使用时间对比，使用时间通过对每个 batch 的运行时间进行求和取平均值计算得到。

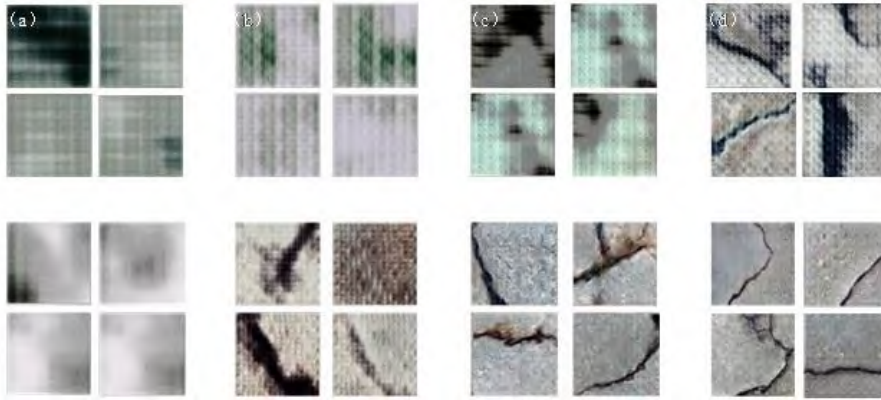


图 6 DCGAN 与 BCIGM 生成裂缝的可视化对比 (a)Epoch=01;(b)Epoch=03;(c)Epoch=16;(d)Epoch=25

Fig 6. Visualization of generated cracks by DCGAN and BCIGM (a)Epoch=01 ;(b) Epoch=03; (c)Epoch=16; (d)Epoch=25

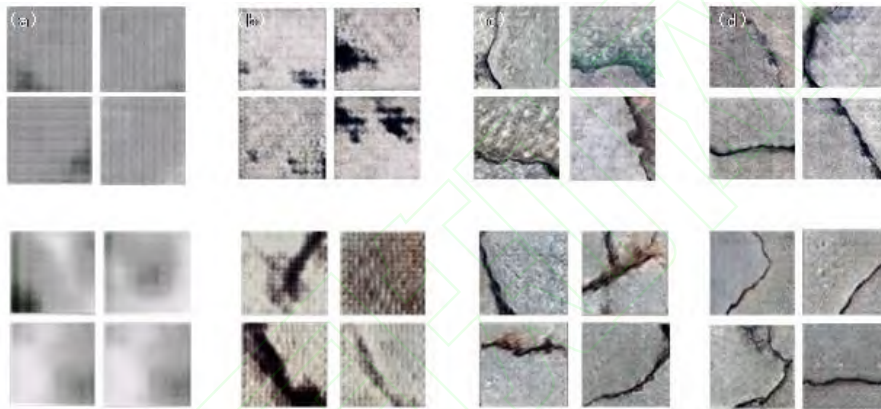


图 7 ReLU 激活函数与 SeLU 激活函数生成裂缝图像的可视化对比 (a)Epoch=01;(b) Epoch=03;(c)Epoch=16;(d)Epoch=25

Fig 7. Visualization of generated cracks by ReLU and SeLU (a)Epoch=01;(b) Epoch=03;(c)Epoch=16;(d)Epoch=25

表 3 有无 1*1 卷积核的 BCIGM 对训练速度的影响

Tab.3 The influence of BCIGM with or without 1*1 convolution kernel on training speed

with or without 1*1 convolution kernel	Time/s
without 1*1 convolution kernel	5.4801
with 1*1 convolution kernel	5.4312

图 6 为实验 1 的结果，其中，第一、二行为通过 DCGAN 生成的裂缝图像，第三、四行为通过 BCIGM 生成的裂缝图像，第一、二列为当 Epoch=01 时的结果图，第三、四列为 Epoch=03 时的结果图，第五、六列为 Epoch=16 时的结果图，第七、八列为 Epoch=25 时的结果图。图 7 为实验 2 的结果，其中，第一、二行为使用 ReLU 激活函数生成的裂缝图像，第三、四行为使用 SeLU 激活函数生成的裂缝图像，第一、二列为当 Epoch=01 时的结果图，第三、四列为 Epoch=03 时的结果图，第五、六列为 Epoch=16 时的结果图，第七、八列为 Epoch=25 时的结果图。表 3 为实验 3 的结果。通过观察实验 1 的两组图可以看出，原 DCGAN 模型生成的裂缝图像网格现象严重，且学习到第 16 个 Epoch 时才初步出现条状裂缝特征，而本文提出的 BCIGM 最后生成的裂缝图像清晰，基本无网格现象，与真实采集到的裂缝图像相似度极高，并且第 3 个 Epoch 就出现了连接的线性裂缝特征。通过观察实验 2 可以看出，采用 ReLU 激活函数的 BCIGM 在第 3 个 Epoch 时仅出现了不连续的黑色裂缝点，而采用 SeLU 激活函数的

BCIGM 出现了连续的裂缝雏形，除此之外，最后生成的裂缝图像也更加清晰，更符合实际场景下拍摄的裂缝图像，更重要的是提高了模型的稳定性，在测试期间没有出现任何一次模型崩塌现象。通过观察表 3 可以看出增加了 1*1 卷积核以后的 BCIGM 训练速度变快，虽然每个 batch 提升速度不多，但当训练数据集庞大，Epoch 增加时，仍然可以减少很大一笔时间开支。

第二组实验用于验证数据集扩增方法对于 BCISM 的影响。该实验的具体操作步骤如下：首先，直接使用人工采集的 1183 张桥梁裂缝图像在不经数据集扩增方法的情况下训练 BCISM；然后，使用本文第 2 章所描述的数据集扩增方法进行扩增，使用扩增之后的数据集训练 BCISM。最后，从数据集的测试集中随机选取 156 张桥梁裂缝图像分别对训练好的 BCISM 进行测试。测试结果从两个角度进行评价，第一个角度为裂缝检测结果的可视化对比，如图 8 所示，其中第一列为原图，第二列为标签图，第三列为不扩充数据集检测结果，第四列为扩充数据集后的检测结果。第二个角度采用量化的指标精确率 Pre 与召回率 Rec 进行评价。在本文中，TP 代表被正确检测出来的裂缝区域像素的数量，FP 代表被误判为裂缝区域像素的数量，FN 代表属于裂缝区域的像素但是却没有被检测出来的像素的数量，精确度 Pre 与召回率 Rec 的具体计算公式如(3)式、(4)式所示，实验结果如表 4 所示。

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

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

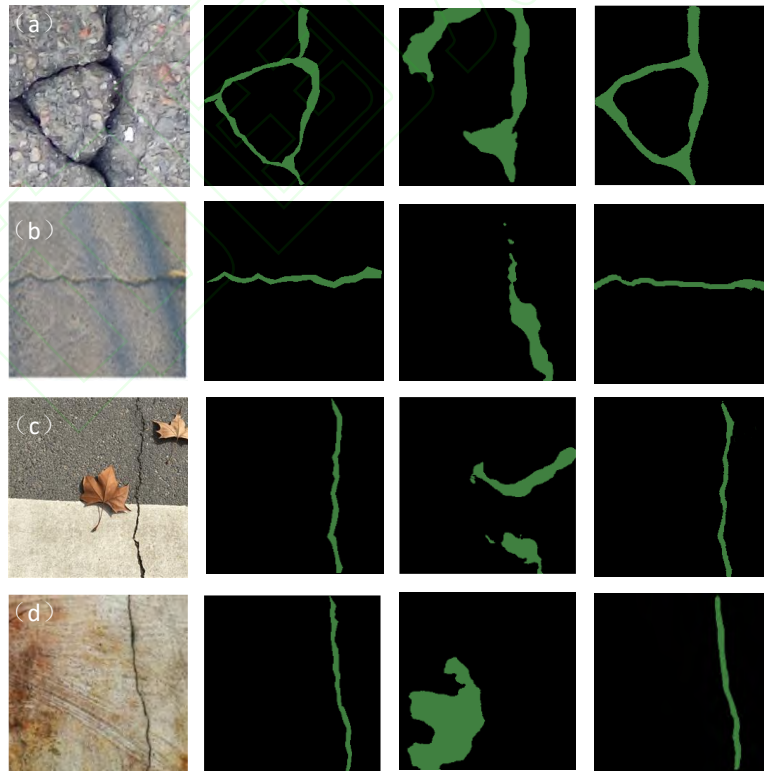


图 8 有无数据集扩增实验结果的可视化对比 (a)原图;(b)标签;(c)无数据集扩增;(d)有数据集扩增

Fig 8. Comparison of experimental results with or without dataset amplification (a) Original image (b) Label (c) without dataset amplification (d) with dataset amplification

表 4 有无数据集扩增对 BCISM 的影响

Tab.4 Comparison of experimental results with or without dataset amplification

Number of training samples	With or without dataset amplification	Number of verification samples	Pre	Rec
1183	Without dataset amplification	156	13.5%	17.9%
29436	With dataset amplification	156	92.9%	92.6 %

通过观察图 8 可知，通过未经过数据扩增的数据集训练的 BCISM 进行测试时，裂缝部分与原图中裂缝匹配度很低，检测效果极其不好。当裂缝背景相对简单时，未经数据集扩充模型的检测结果与标签图有相似之处，如第一行所示。但是当背景中存在障碍物或大面积污渍时，未经数据集扩充的模型的检测结果几乎与标签图像无任何相似度，如第二行至四行所示。观察表 4 可知，第一组实验由于没有对采集来的数据进行扩充，导致了网络模型的训练样本严重不足，从而产生了欠拟合现象，其直接表现就是精确率与召回率极低，而通过数据集扩充之后的数据集训练的 BCISM 进行测试时，可以看到精确率与召回率显著提高，网络模型的欠拟合现象基本消除。综上所述，充足的训练样本对于模型的训练成败至关重要，这也证明了本文提出的数据集扩增方法是极其必要的。

第三组实验用于验证本文提出的 BCISM 与当前主流的语义分割模型的对比。本组实验选用了 SegNet、FCN、DeepLab、以及 FC-DenseNet56、FC-DenseNet67、FC-DenseNet103 模型作为 BCISM 的对比模型，对是否预训练、参数大小、精确率 Pre、召回率 Rec、F1_Score 以及训练时每张图所用时间几方面进行对比。对比结果如表 5 所示，其中，精确率 Pre 与召回率 Rec 与第 4 章第二组对比实验所述一致，F1_score 可以看作是模型精确率和召回率的一种加权平均，兼顾了精确率和召回率的查全和查准的作用。F1_Score 的具体计算公式如(5)式所示。

$$F_1 = 2 * \frac{Pre * Rec}{Pre + Rec} \quad (5)$$

表 5 BCISM 与主流语义分割模型对比

Tab.5 Comparison of the exiting semantic segmentation models and BCISM

Model	Pre-training	Parameter (M)	Pre	Rec	F1_Score	Time(s)
SegNet	T	29.5	74.0%	78.5%	76.2%	0.5823
FCN8	T	134.5	86.9%	83.4%	85.1%	0.3739
DeepLab	T	37.3	82.6%	80.9%	81.7%	0.9751
FC-DenseNet56(k=12)	F	1.5	89.8%	87.6%	88.7%	0.1685
FC-DenseNet67(k=16)	F	3.5	89%	88.8%	88.9%	0.2635
FC-DenseNet103(k=16)	F	9.4	93.0%	92.1%	92.5%	0.2795
BCISM(k=12)	F	2.8	92.9%	92.6%	92.8%	0.1998

通过观察表 5 可知：BCISM 与 SegNet、FCN8、DeepLab 三个模型相比，在没有预训练的情况下，参数量与时间均较低，且 Pre、Rec、F1_Score 均有大幅提高。BCISM 与 FC-DenseNet56 相比，虽参数量与时间稍高，但 Pre、Rec、F1_Score 均提高 4%左右；BCISM 与 FC-DenseNet67 相比，参数量与时间均低，且 Pre、Rec、F1_Score 均提高；BCISM 与 FC-DenseNet103 相比，参数量不及三分之一且每张图的时间减少 0.1 秒左右，但是二者的 Pre、Rec、F1_Score 结果相近，均在 92%以上；综上所述，BCISM 在没有预

训练的情况下，用较少的参数，较短的训练时间，达到了最精确的检测结果。

第四组实验用于验证本文提出的算法与现有裂缝检测算法的检测效果对比。实验结果如图 9 所示，其中第一列为原图，第二列为标签，第三列为采用阈值分割算法的检测结果，第四列为采用 Canny 算法的检测结果，第五列为采用 NB-CNN 算法的检测结果，第六列为采用随机结构森林算法的检测结果、第七列为采用本文算法的检测结果。

第四组对比实验选择两幅背景简单的裂缝图像、两幅含有障碍物的裂缝图像，一幅含有大面积污渍的裂缝图像作为裂缝图像的典型代表。选择阈值分割算法、Canny 算法作为传统裂缝检测算法代表，选择 NB-CNN 算法、随机结构森林算法作为深度学习算法代表进行对比。通过观察实验结果可知，通过阈值分割算法进行检测的结果会产生大量的噪声，并且当背景复杂时，会受到障碍物和大面积污渍的影响，使得检测效果更差，结果如图 9 的第三列所示。Canny 算法的检测结果同样不理想，提取出的裂缝中非裂缝部分占据了极大的比例，结果如图 9 的第四列所示。NB-CNN 算法可以检测到图像中所有的裂缝，但是仅仅可以标注裂缝所在位置，不能实现精确提取，而且把车道线、大片污渍等全部纳入裂缝提取框中，结果如图 9 的第五列所示。通过随机结构森林算法检测的图像，噪声量减少，但是当存在障碍物钉子与落叶时，会把钉子与部分落叶边缘错分为裂缝，而且当背景含有大面积污渍时，检测效果极差，如图 9 第六列所示。但是，本文提出的复杂背景下基于图像处理的桥梁裂缝检测算法在以上几种情况下均取得理想效果，准确提取出了裂缝部分，如图 9 第七列所示。因此可以证明：与现有算法相比，本文提出的算法更适用于复杂背景下的桥梁裂缝检测。

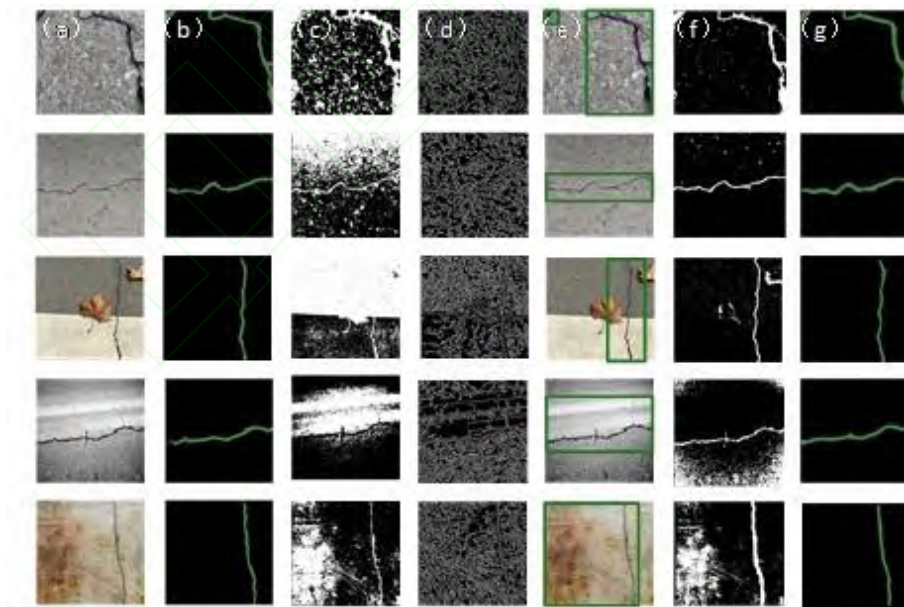


图 9 主流裂缝检测算法与本文算法的对比图 (a)原图;(b)标签;(c)阈值分割算法;(d)Canny 算法;(e)NB-CNN 算法;(f)随机结构森林算法;(g)本文算法

Fig 9. Comparison of the exiting algorithms and the algorithm proposed by this paper (a) Original image;(b)Label;(c)Threshold segmentation algorithm;(d)Canny algorithm;(e)NB-CNN algorithm;(f) Random structure forest algorithm;(g) algorithm proposed by this paper

第五组对比实验用于验证本文提出的算法在背景简单的裂缝图像、背景含有障碍物的裂缝图像以及背景含有大面积污渍的裂缝图像三种场景中进行裂缝检测的有效性。本文最终对测试集中不同背景的桥梁裂缝图像进行了检测，实验结果如图 10 所示。其中第一列、第四列以及第七列为原图、第二列、第五列、第八列为标签图、第三列、第六列、第九列为采用本文提出算法进行裂缝检测的结果图。

第五组实验为了证明本文提出的算法适用于复杂背景的桥梁裂缝检测，故挑选了典型的九幅图像进行说明。按照由上至下，由左及右的顺序对图 10 的图像依次进行编号。第一幅、第二幅与第三幅为不同类型但背景均简单的裂缝图像，第四幅为背景中含有阴影的裂缝图像，第五幅为背景中含有车道线的裂缝图像，第六幅为背景中含有落叶与车道线的裂缝图像，第七幅为背景中含有油漆箭头与小块污渍的裂缝图像，第八幅为光照不均且含有车道线、钉子等障碍物的裂缝图像，第九幅为含有大面积污渍的裂缝图像。通过观察上图实验结果可以看出，本文的算法不仅适用于背景简单的桥梁裂缝图像检测，而且适用于背景复杂的桥梁裂缝检测，从而可以证明本文算法的有效性。

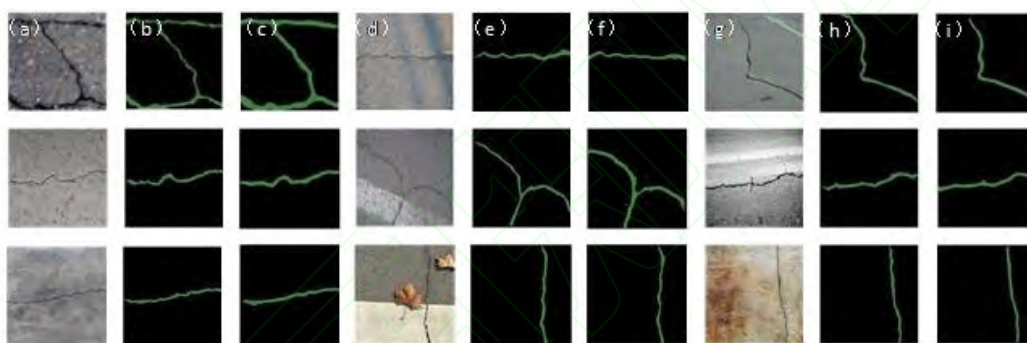


图 10 本文算法部分检测结果(a)原图;(b)标签;(c)检测结果;(d)原图;(e)标签;(f)检测结果;(g)原图;(h)标签;(i)检测结果

Fig 10. Partial test results of the algorithm proposed by this paper (a)Original image;(b)label;(c)the result;(d)original image;(e)label;(f)the result;(g)original image;(h)label;(i)the result

5 结 语

本文提出了复杂背景下基于图像处理的桥梁裂缝检测算法。主要介绍了基于深度卷积生成式对抗网络的桥梁裂缝图像生成模型，提出了基于语义分割的桥梁裂缝图像分割模型，并实现了对高分辨率图像的裂缝提取。实验结果表明：数据集扩充有效缓解了由于数据不充足引起的欠拟合现象；与现有的语义分割模型相比，本文提出的桥梁裂缝图像分割模型参数量减少、训练时间缩短，但是检测效果提高；与现有桥梁裂缝检测算法相比，本文提出的算法在复杂背景下有更理想的检测效果。

参考文献(References)

- [1] Wang W Q. Development and expectation of bridge engineering technology[J]. Construction Technology, 2018, 47(06): 103-108.
王武勤. 桥梁工程技术发展与展望[J]. 施工技术, 2018, 47(06): 103-108.

- [2] Qu L, Wang K R, Chen L L, *et al.* Fast road detection based on RGBD images and convolutional neural network[J]. *Acta Optica Sinica*, 2017, 37(10):1010003.
曲磊, 王康如, 陈利利, 等. 基于RGBD图像和卷积神经网络的快速道路检测[J]. *光学学报*, 2017, 37(10): 1010003.
- [3] National Bureau of Statistics of People's Republic of China. The statistics communique on national economy and social development of China 2017 [N]. *China Information News*, 2018-03-01(003).
中华人民共和国国家统计局. 中华人民共和国2017年国民经济和社会发展统计公报[N]. *中国信息报*, 2018-03-01(003).
- [4] Deng X L, Tian S Z. Bridge deformation detection and data processing based on 3D laser scanning[J]. *Laser & Optoelectronics Progress*, 2018, 55(7): 071201.
邓晓隆, 田石柱. 基于三维激光扫描的桥梁变形监测及数据处理[J]. *激光与光电子学进展*, 2018, 55(7): 071201.
- [5] Wang B, Wang X, Chen F, *et al.* Pavement crack recognition based on aerial image. *Acta Optica Sinica*, 2017, 37(08): 0810004.
王博, 王霞, 陈飞, 等. 航拍图像的路面裂缝识别[J]. *光学学报*, 2017, 37(08): 0810004.
- [6] Oh H, Garrick N W, Achenie L. Segmentation algorithm using iterative clipping for processing noisy pavement images[C].//*Imaging Technologies@s Techniques and Applications in Civil Engineering*.1998: 138-147.
- [7] Sun L, Xing J C, Xie L Q, *et al.* An adaptive threshold-based Canny algorithm for crack detection[J]. *Microcomputer&its Applications*, 2017, 36(05): 35-37+41.
孙亮, 邢建春, 谢立强, 等. 基于自适应阈值Canny算法的裂缝检测方法研究[J]. *微型机与应用*, 2017, 36(05): 35-37+41.
- [8] Talab A M A, Huang Z, Xi F, *et al.* Detection crack in image using Otsu method and multiple filtering in image processing techniques[J]. *Optik - International Journal for Light and Electron Optics*, 2016, 127(3):1030-1033.
- [9] Zhang L, Yang F, Zhang Y D, *et al.* Road crack detection using deep convolutional neural network[C].// *IEEE International Conference on Image Processing*, 2016: 3708-3712.
- [10] Chen F C, Jahanshahi M R. NB-CNN: deep learning-based crack detection using convolutional neural network and naïve bayes data fusion[J]. *IEEE Transactions on Industrial Electronics*, 2018, 65(5): 4392-4400.
- [11] Shi Y, Cui L, Qi Z, *et al.* Automatic road crack detection using random structured forest[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 17(12): 3434-3445.
- [12] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. *Computer Science*, 2015: 2333-9721.
- [13] Klambauer G, Unterthiner T, Mayr A, *et al.* Self-Normalizing neural networks[C].// *31st Conference on Neural Information Processing System*, 2017.
- [14] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2014, 39(4): 640-651.
- [15] Guo C C, Yu F Q, Chen Y. Image semantic segmentation based on convolutional neural network feature and improved superpixel matching[J]. *Laser & Optoelectronics Progress*, 2018,55(8): 081005.
郭呈呈, 于凤芹, 陈莹. 基于卷积神经网络特征和改进超像素匹配的图像语义分割[J]. *激光与光电子学进展*, 2018, 55(8): 081005.
- [16] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for scene segmentation[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 39(12): 2481-2495.
- [17] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2014, 39(4): 640-651.
- [18] 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 & Machine Intelligence*, 2018, 40(4): 834-848.
- [19] Huang G, Liu Z, Maaten L V D, *et al.* Densely connected convolutional networks[C].//*IEEE Conference on Computer Vision and Pattern Recognition*. *IEEE Computer Society*, 2017: 2261-2269.
- [20] Jegou S, Drozdal M, Vazquez D, *et al.* The one hundred layers tiramisu: fully convolutional DenseNets for semantic segmentation[C].//*2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017: 1175-1183.

网络首发:

标题: 复杂背景下桥梁路面裂缝检测与分割算法研究

作者: 李良福, 孙瑞赞

收稿日期: 2018-09-12

录用日期: 2018-09-30

DOI: 10.3788/lop56.061002

引用格式:

李良福, 孙瑞赞. 复杂背景下桥梁路面裂缝检测与分割算法研究[J]. 激光与光电子学进展, 2019, 56(06): 061002.

网络首发文章内容与正式出版的有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

基于卷积神经网络特征和改进超像素匹配的图像语义分割

郭呈呈 于凤芹 陈莹

江南大学物联网工程学院, 江苏 无锡 214122

激光与光电子学进展, 2018, 55(8): 081005

复杂地物背景下的车辆目标激光主动偏振成像研究

徐飞飞 曾朝阳 陈杭

装备学院研究生管理大队, 北京 101416

激光与光电子学进展, 2016, 53(5): 051103