

# 基于生成式对抗网络的细小桥梁裂缝分割方法

李良福<sup>1</sup>, 胡敏<sup>1\*</sup>

<sup>1</sup> 陕西师范大学计算机科学学院, 陕西 西安 710119

**摘要** 裂缝具有分布杂乱无规则、粗细不均等特性, 检测准确率往往不高。小目标检测受图像本身分辨率低和噪声的干扰, 使用传统的检测方法处理, 得到的检测结果通常也不理想。针对上述问题, 提出一种基于生成式对抗网络的细小桥梁裂缝分割方法。该方法在判别器结构中引入分割分支, 将生成式对抗网络与语义分割网络合二为一, 兼具超分辨率图像重建功能与分割功能。在处理细小桥梁裂缝分割问题时, 该方法首先将低分辨率的细小桥梁裂缝图像转换为超分辨率的粗大桥梁裂缝图像, 然后对转换后的超分辨率图像进行分割。实验结果表明, 该方法更容易识别出细小桥梁裂缝并进行准确的分割, 与传统的分割方法相比, 该方法的分割召回率较之前增加了 6%, 平均 IU 增加了 10%。

**关键词** 桥梁路面裂缝; 深度学习; 超分辨率图像; 生成式对抗网络; 语义分割

**中图法分类号** TP391.9 **文献标识码** A

## Small Bridge Crack Segmentation Method Based on Generative Adversarial Network

Li Liangfu<sup>1</sup>, Hu Min<sup>1\*</sup>

<sup>1</sup> School of Computer Science, Shaanxi Normal University, Xian Shaanxi 710119

**Abstract** The distribution of cracks is irregular and uneven, which leads to low detection accuracy. Because of the low resolution of the image and the interference of the noise, the detection result of the small target detected by the traditional detection method is not in line with the expectation. In view of the above problems, we propose a segmentation method of cracks for small bridges, which is based on the generative adversarial network. This segmentation method introduces the segmental branch in the discriminator structure, combining the generative adversarial network and semantic segmentation network. Simultaneously, it has the function of image reconstruction and segmentation based on super-resolution. In order to solve the problem of small bridge cracks segmentation, this method transforms the small bridge cracks image with low resolution into the coarse bridge cracks image with super-resolution, and then segments the transformed super-resolution image. Experimental results show that this method makes it easier to identify small bridge cracks, and the segmentation is more accurate. Compared with the traditional segmentation method, the recall rate of this segmentation method is improved by 6% and the mean IU is improved by 10%.

**Key words** bridge pavement crack; deep learning; super-resolution image; generative confrontation network; semantic segmentation

**OCIS codes:** 100.200;100.4996;100.1830;150.1135

## 1 引言

随着交通运输业的发展, 公路的养护工作变得十分的重要。桥梁作为当今社会交通枢纽

收稿日期: 年-月-日; 修回日期: 年-月-日; 录用日期: 年-月-日

作者简介: 李良福(1977—), 男, 博士, 副教授, 从事人工智能、计算机视觉专业; 胡敏(1993—), 女, 硕士研究生, 从事深度学习、计算机视觉专业。

基金项目: 国家自然科学基金项目(61573232, 61401263)

\*E-mail: [longford@xjtu.edu.cn](mailto:longford@xjtu.edu.cn); \*\*E-mail: [550486740@qq.com](mailto:550486740@qq.com);

的重要组成部分,不仅承担着运输重任还关系着运输人员的安危,但由于桥梁结构在长期使用中难免产生各种各样的损坏,造成桥梁结构抗力衰减和安全隐患,因此定期的检查和维护是必不可少的。裂缝是桥梁最常见的缺陷。桥梁裂缝产生的原因多种多样,主要是由于沥青路面的疲劳,再加上不利的大气条件,可能导致材料收缩,或由于沥青混合物质量差造成的结构路面阻力。桥梁路面缺陷的检测是后续维护和修理的前提。因此桥梁裂缝的检测工作必不可少。现实生活中,由于裂缝分布杂乱无规律,往往存在很多细小裂缝,容易被周边障碍物干扰造成漏检误检造成了极大的安全隐患。为了避免这一类情况的发生,对细小桥梁裂缝检测的研究变得十分重要。

随着计算机行业的飞速发展,一系列基于计算机视觉、数字图像处理的方法被应用于裂缝检测<sup>[1-3]</sup>中。Li等人提出了一种基于相邻差分直方图的裂缝分割算法<sup>[4]</sup>;这类基于阈值分割的裂缝识别方法受光照、噪声、纹理的影响,检测效果往往不稳定。针对这一问题,Landstrom等人结合形态学处理和逻辑回归算法对裂缝进行检测<sup>[5]</sup>,利用统计学分类方法滤除噪声,提高检测精度。近几年,在物体检测和识别<sup>[6-8]</sup>方向上取得重要成果的深度学习方法也逐渐被用到了桥梁裂缝的检测中。赵珊珊等人提出使用卷积神经网络对裂缝进行检测<sup>[9]</sup>,对预标记的裂缝图像块进行预处理,将预处理后的图像输入卷积神经网络中完成对裂缝的检测。为了进一步将裂缝检测精确到像素,Evan Shelhamer等人提出的全卷积语义分割网络<sup>[10]</sup>。该模型可以获得的更贴合裂缝本身分布的检测结果,但其将低分辨率分割图上采样至输入图像分辨率机制,花费了大量计算成本。针对这一问题,Badrinarayanan等人在解码器中使用池化层索引对特征图进行上采样<sup>[11]</sup>,使得待分割图像可以保持高频细节的完整性。对于分布杂乱、特征相对不明显的细小桥梁裂缝而言,由于特征信息难提取、裂缝本身分布面积小于周围障碍物以及对对应位置像素信息与周围接近等原因,使得上述传统的裂缝检测方法对细小桥梁裂缝的检测往往不佳。针对小目标检测问题,一个普遍的做法<sup>[12-18]</sup>是增加输入图像的数量以提高小目标的分辨率,从而获得高分辨率的特征图。还有一些<sup>[19-21]</sup>通过修改网络模型,将多个低层特征合成新的高层特征,以生成小目标的多尺度表示。

以上方法均试图通过数据增强或者增加特征维度来提高模型对小目标的检测能力。通过数据增强的方法常常导致训练和测试耗费大量的时间。而由低级特征构造高级特征的方法,则不能保证所构造的特征对于最终的检测有效,且其对检测效果的贡献也仅限于偿还计算成本。对于小目标检测问题,本文认为一种有效的方法是寻找每个类别目标的低分辨率图像和其对应的高分辨率图像内在结构的相关性,将低分辨率图像按照一定的规则转换为与高分辨率图像相似的超分辨率图像,然后使用转换后的超分辨率图像代替低分辨率图像完成检测任

务。相比低分辨率图像，超分辨率图像具有更多可以被分类模型检测到的细节特征，改善了网络的欠拟合问题，从而提高了检测准确率。生成式对抗网络<sup>[22]</sup>具有学习图像间潜在关联并依此生成以假乱真的图像的特性。因此本文提出一种分割生成对抗网络SE-GAN (Segmental Generative Adversarial Network)，将具有低分辨率的细小桥梁裂缝图像输入到网络，得到超分辨率图像以及对应的裂缝分割结果。在SE-GAN模型中，本文对判别器结构和生成器结构分别进行了改进。在判别器结构中，本文将原本识别输入图像真伪的传统判别器更名为判别分支，同时为判别器结构引入了分割分支。并通过分割分支反馈的分割损失和判别分支反馈的对抗损失一起引导生成器生成更为真实的超分辨率图。与此同时，为了进一步挖掘真实的高分辨率图像与生成器生成的图像之间的潜在关联，本文对判别器结构中的判别分支的输出部分进行了改进，使得判别器回传更准确的信息；在生成器结构中，本文引入了像素损失，使得生成器可以生成从视觉上与真实图像更相似的超分辨率图像。

## 2 相关工作

### 2.1 生成式对抗网络原理

生成式对抗网络由生成器和判别器两个部分组成。判别器是一个简单地卷积神经网络模型，其以真实图像和生成器构造的虚假图像为输入，通过一系列卷积层、激励层、归一化层、池化层对输入数据进行特征提取，最终输出 $[0, 1]$ 区间的概率值；生成器则是一个反向的卷积神经网络模型，通过一系列反卷积层进行上采样，结合激励层，将低维向量转化为与真实图像维度相同的向量输出。其中生成器的输入是随机生成高斯白噪声，经过生成器网络的解码，最终输出一个与真实图像大小相同的向量，然后通过判别器预测的置信值计算与真实类别对应标签之间的差距，直接将此误差作为反向传播的误差来更新参数以及最初的输入向量。生成式对抗网络采用交叉熵损失函数进行参数优化，公式如（1）所示，

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

其中  $D$  表示判别器 (discriminator)， $G$  表示生成器 (generator)， $z$  是输入的随机噪声向量。判别器需要判别输入是真实图像还是生成器生成的图像，当输入是真实图像时， $D(x)$  值趋近于一；当输入是生成器生成的图像时， $D(x)$  值趋近于零。生成器则是尽可能的调整  $z$  的分布，使得  $G(z)$  与  $x$  的分布差异最小化，即  $D(G(z))$  趋近于一。由于生成式对抗网络具有两个子网络模型，因此采用逐步交叉训练方式。训练判别器时的损失函数为公式（2）所示，此时默认生成器性能已达到最优，故只需训练判别器的分类性能。

$$\max_D V(D, G) = E_{x \sim P_{data}(x)} [\log(D(x))] + E_{z \sim P_G(z)} [\log(1 - D(G(z)))] \quad (2)$$

而训练生成器时的损失函数时，默认判别器具有最优的判别能力，即  $E_{x \sim P_{data}(x)} [\log(D(x))]$  部分始终为固定常数，不需要加入训练，故训练生成器时的损失函数如公式（3）所示。

$$\min_G V(D, G) = E_{z \sim P_G(z)} [\log(1 - D(G(z)))] \quad (3)$$

由于生成式对抗网络具有模拟生成图像的功能，本文将生成式对抗网络运用于细小桥梁裂缝图像的超分辨率桥梁裂缝图像重建中，并在生成式对抗网络的判别器中引入差异判别分支，增加像素损失，为细小桥梁裂缝图像补充更多的细节信息，将细小桥梁裂缝图像转化为易于被语义分割模型分割的超分辨率粗大桥梁裂缝图像。

## 2.2 语义分割网络原理

语义分割网络与目标分类网络结构类似，同样由卷积层、池化层、激励层构成。不同的是，目标分类网络在卷积层特征提取之后得到一个长度等于分类类别数目的列向量，然后通过激活函数将列向量中的值映射到  $[0, 1]$  区间，得到目标属于每个类的概率，概率值最大的类即为模型预测的类别。而语义分割网络在卷积层特征提取之后进行反卷积和分阶段特征融合操作，得到个数等于分类的类别数目，大小等于网络输入尺寸的热图，并根据输入图像对应的分割标签按局部区域像素点分类准确率计算损失值，实现梯度回传，计算公式如公式（4）所示。

$$L(x; \theta) = \sum_{ij} l(x_{ij}; \theta) \quad (4)$$

$x_{ij}$  为图像  $i$  行  $j$  列对应的像素值， $\theta$  为模型对应的超参数， $l(x_{ij}; \theta)$  为局部像素的分类损失函数， $L(x; \theta)$  为分割损失函数。

针对粗大型、中型桥梁裂缝图像，使用传统的语义分割网络可很好的对裂缝信息进行像素级别的分割。但针对形状细小、纹理特征不清晰的细小型桥梁裂缝图像，由于在整幅图像中裂缝像素信息所占比例少，供分割模型学习的显著特征不足，同时受噪声的干扰，仅使用单一语义分割网络对其进行分割，容易产生误检、漏检等问题。因此，本文将语义分割网络引入生成式对抗网络的判别器结构中，建立针对细小桥梁裂缝图像的语义分割生成式对抗网络 SE-GAN。首先将低分辨率细小型桥梁裂缝图像输入生成式对抗网络，生成对应的超分辨率粗大型桥梁裂缝图像，再使用语义分割网络对超分辨率粗大型桥梁裂缝图像进行分割。该方法将小目标检测问题转换为正常尺度目标检测问题，避免了因细小桥梁裂缝图像中裂缝像素信息不足、噪声干扰导致的检测不准确情况，提升了分割模型针对细小桥梁裂缝图像的分割准确率，本文方法的具体阐述如下。



### 3 细小桥梁裂缝语义分割生成式对抗网络

本文将生成式对抗网络突破性地运用于小目标图像语义分割任务中,建立细小桥梁裂缝语义分割生成式对抗网络 SE-GAN, 并针对细小桥梁裂缝图像特征分别对生成式对抗网络的判别器结构和生成器结构进行调整。具体的改进如下, 在判别器结构中, 新增加分割分支, 将传统的判别器保留并以判别分支的形式与分割分支共同构成新的判别器。使得判别网络兼具判别图像真伪以及裂缝图像语义分割的功能, 与此同时, 传统的损失函数忽略了真实图像与生成图像之间的差异性, 不能够学习到二者间更多的潜在关系, 为了更深层次地挖掘两者间的内在关联, 同时增强判别器非线性拟合能力, 本文对判别分支的损失函数进行调整, 将原本的二分类问题转换为差异程度分析问题。

在生成器模型中, 为了提高模型生成图像的能力, 更好地刻画细小桥梁裂缝图像细节, 首先引入更深层的残差结构, 在保证不会造成梯度消失问题发生的情况下, 提高网络的非线性拟合能力。为了缩小生成图像与真实图像在视觉上的差异, 引入像素损失, 保证两种图像从潜在分布到视觉分布均具有相似性。考虑到细小桥梁裂缝的形状大多细长无规律, 且其对应的像素分布占图像整体像素的比例少的特性, 本文将细小桥梁裂缝图像进行四等分, 得到十六个低分辨率子图, 切分方法如图 1 所示。使用生成式对抗网络对子图依次进行超分辨率重建, 并对重新后的裂缝图像进行语义分割, 得到相应的检测结果图, 将结果图按原切分顺序进行拼接, 从而得到最终的检测结果。接下来, 本章将对上述方法进行具体的阐述。

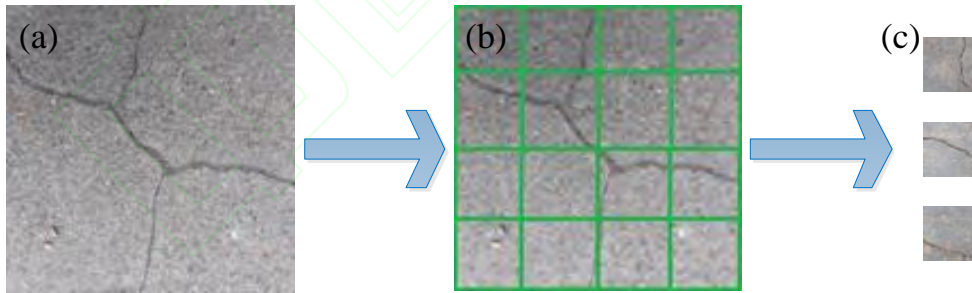


图 1 低分辨率子图切分示意图。(a) 原始图像 (b) 切分示意图 (c) 切分成的低分辨率子图

Figure 1 Schematic diagram of the low resolution subgraph segmentation. (a) original image (b) segmentation diagram (c) split into low resolution subgraphs

#### 3.1 SE-GAN 判别器

不同于传统的判别器, 本文中判别器模型由两个分支构成, 分别是判别分支和语义分割分支, 判别分支用传统的判别器功能相同, 来判别输入是真实图像还是生成器生成的图像。语义分割分支作用有两个, 一是生成裂缝的语义分割结果完成网络对裂缝检测, 二是获得分割损失, 与判别器得到的对抗损失共同调整生成器生成超分辨率图像的分布, 判别器结构如

图 2 所示,将真实的高分辨率桥梁裂缝图像对应的特征图和生成器生成的超分辨率特征图作为判别器结构的输入,经过特征提取分类,最终输出。

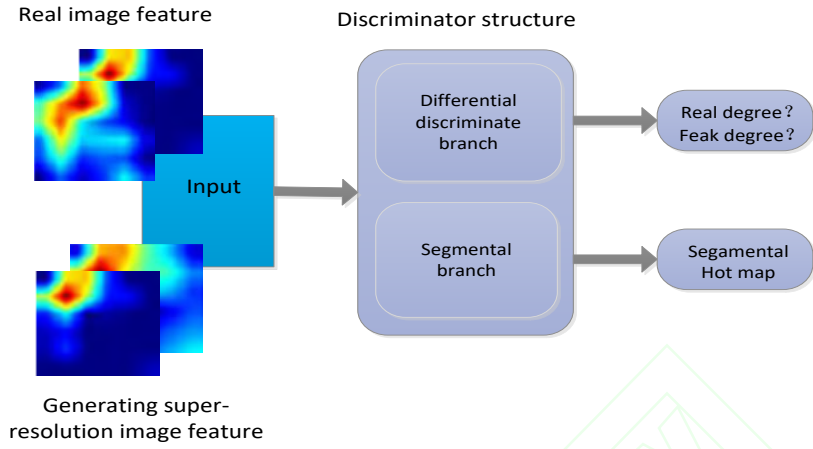


图 2 判别器结构示意图

Figure 2 Schematic diagram of the discriminator

### 3.1.1 差异判别分支

传统的判别器结构中,判别器模型分别对真实的裂缝图像与生成的裂缝图像进行学习,并按照相应的标签分别对两种图像的学习结果进行修正,这样做忽略了真实的裂缝图像与生成的裂缝图像间的内在联系,无法将判别器功能最大化。为了更好的挖掘真实图像分布与生成图像分布之间的潜在相关性,引导生成网络生成的图像可以更完美的拟合真实图像,本文引入差异判别分支,判别器不再分裂化的对两种图像进行判别,而是对两者之间的差异进行判别。考虑到高分辨率粗大型桥梁裂缝图像与生成器生成的超分辨率细小桥梁裂缝图像之间具有一定的相关性,并且真实的高分辨率图像的真实程度和生成器生成的超分辨率图像的虚假程度都是相对于判别器学习到的两者间的相对差异而言的,所以这两者之间的差异可以表示为:

$$dis(x, G(z)) = \|F(x) - E(F(G(z)))\|_{L_1} \quad (5)$$

传统的判别器输出经激活函数  $D = \sigma(F(\bullet))$  映射到  $[0, 1]$  区间的概率值,其中  $F$  表示判别网络特征提取过程,  $\sigma$  表示激活函数。因此,公式 (5) 中  $F(x)$  表示真实高分辨率桥梁裂缝图像经过判别网络后提取到的特征,  $F(G(z))$  表示生成的超分辨率桥梁裂缝图像经过判别网络后提取到的特征,  $E(\bullet)$  表示均值。对于真实的高分辨率图像判别器输出为  $D(dis(x, G(z)))$ , 对应生成的超分辨率图像判别器输出为  $D(dis(G(z), x))$ , 经过激活函数映射,此时的判别器输出仍为  $[0, 1]$  的概率值,与传统的判别器一样,真实图像相对生成图像的真实程度对应标签为

一，生成器生成图像相对真实图像的虚假程度对应标签为零。此时生成式对抗网络的损失具有如下定义：

$$L_{adv} = E_{x \sim P_{data}} [\log(D(dis(x, G(z))))] + E_{x \sim P_{G(z)}} [\log(1 - D(dis(G(z), x)))] \quad (6)$$

其中  $L_{adv}$  表示对抗损失。在训练判别器时，损失函数为：

$$\max_D V(D, G) = E_{x \sim P_{data}} [\log(D(dis(x, G(z))))] + E_{x \sim P_{G(z)}} [\log(1 - D(dis(G(z), x)))] \quad (7)$$

训练生成器时，损失函数为：

$$\min_G V(D, G) = E_{x \sim P_{data}} [\log(1 - D(dis(G(z)), x))] \quad (8)$$

本文采用的差异判别分支结构如图 3 所示。分支的输入由两部分组成， $G(z)$  表示生成器生成的超分辨率桥梁裂缝图像，HR 表示真实的高分辨率桥梁裂缝图像。分支的结构主要由卷积层 (Conv)、激活层 (Leaky Relu)、归一化层 (Batch Normalization) 构成。

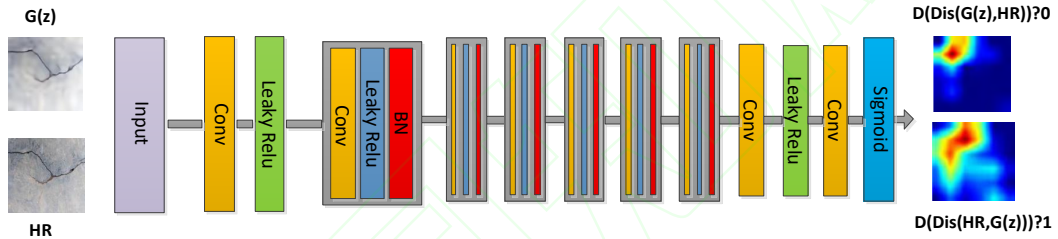


图 3 差异判别分支结构

Figure 3 Discriminate branch structure

### 3.1.2 分割分支

本文分割分支使用独立的分割结构模型，直接使用 FC-DenseNet103 中的网络模型对桥梁裂缝图像进行语义分割，具体的模型结构如图 4 所示，上排表示网络的大致结构，是由 Dense Bloc 结构、Transition Down 结构、卷积层 (conv)、反卷积层 (Deconv) 构成。其中 Dense Block 由 layer 结构构成，Dense Block、layer 以及 TransitionDown 结构的组成如图 4 下排所示，箭头表示由箭头起点所在层的输出与箭头终点所在层的输出内容拼接，组成具有更厚结构的特征图，并将组合后的特征图作为下一层的输入。先由生成模型生成细小桥梁裂缝的超分辨率图像，将超分辨图像及其对应的语义分割标签图一起送入分割分支，随后生成对应热图，通过给定的语义分割标签图计算损失如公式 (9)，将损失回传给生成器，生成器根据返回的梯度信息调整模型超参数以及对生成的超分辨图像进行细化，促使生成器生成更易于分割分支检测的细节信息。

$$L_{seg} = \sum_{ij} (y_{ij} \log x_{ij} + (1 - y_{ij}) \log(1 - x_{ij})) \quad (9)$$

其中  $L_{seg}$  表示分割损失， $x_{ij}$  表示网络预测热图对应  $i$ 、 $j$  位置的像素所属类别， $y_{ij}$  为标签图对应  $i$ 、 $j$  位置的像素所属类别。

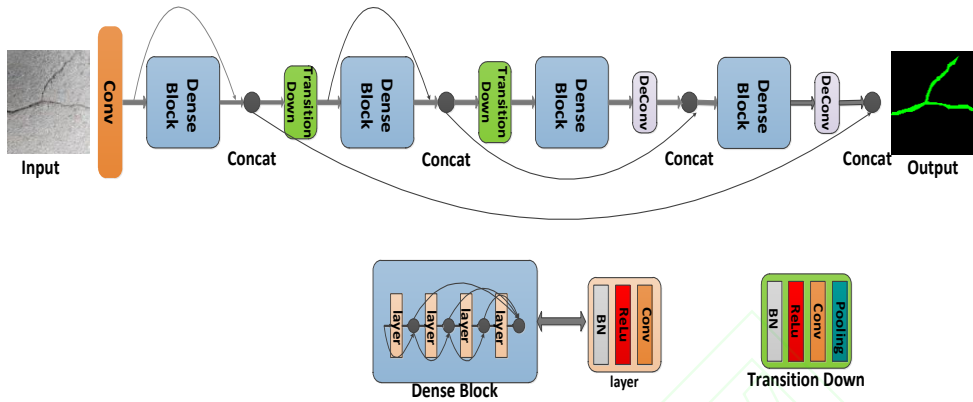


图 4 分割分支结构图

Figure 4 Split branch structure diagram

### 3.2 SE-GAN 生成器

#### 3.2.1SE-GAN 生成器网络结构

在处理超分辨率生成问题中，往往在生成网络中加入残差结构来增加模型的复杂性，同时又防止因过拟合造成的泛化能力差的问题，提高了网络的生成能力和泛化能力，从而生成与高分辨率图像分布类似的超分辨率图像。SRGAN<sup>[26]</sup>是典型在生成器中运用残差结构的方法，其在卷积神经网络的卷积层后增加残差结构，使得网络具有更深的层，增添了非线性性，从而使模型的学习性能大幅的提升。

为了进一步提高生成器生成超分辨率图像的能力，同时增加更多的非线性特征，使得深层卷积可以拟合浅层卷积输出的特征，并在防止网络出现梯度消失问题的同时将低级特征和高级特征融合从而形成多层残差融合模型，本文将 SRGAN 中生成网络的 16 个残差块扩展成 32 个，具体结构如图 5 所示。网络由卷积层（conv）、激活层（Relu）、归一化层（Batch Normalization）、反卷积层（Deconv）构成，其中灰色结构为残差块部分，深灰色方块对应单个残差块的内部结构。

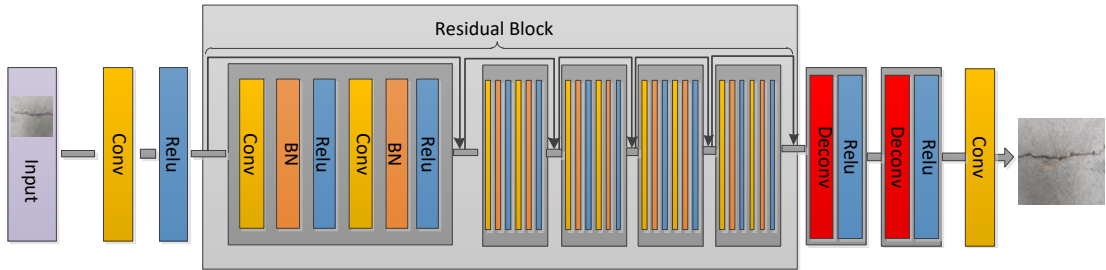


图 5 生成网络结构对比图



Figure 5 shows the network structure comparison chart

### 3.2.2 像素损失

仅使用对抗损失和分割损失对 SE-GAN 进行调整, 往往会出现网络最终的损失很低, 但肉眼看到的结果却并不好的问题。这是因为对抗损失和分割损失均是从图像的特征分析角度来获得用于生成的超分辨率图像的优化梯度信息, 对于模型而言可以得到等价的像素分布, 由于不同位置的取值不同, 图像呈现的效果差距很大。因此, 为了降低像素分布差异而造成的视觉缺憾, 本文在对生成器生成的超分辨率图像进行优化时, 加入了像素损失  $L_p$ ,

$$L_p = |G(z) - x| \quad (10)$$

结合对抗损失、分割损失共同控制超分辨率图像的分布。

$$L = L_{adv} + L_{seg} + L_p \quad (11)$$

其中,  $L$  为 SE-GAN 模型的损失函数, 为分割参数, 为像素参数。

## 4 实验与结果分析

本文具体的运行操作环境分为两部分: 硬件部分和软件环境。(1) 硬件部分: 图像采集装置采用日本 Panasonic 公司的 wv-BP330 相机。计算机硬件采用 Inter(R) Core(TM) i5-6600 中央处理器, 主频为 3.31GHZ, 主机内存为 8G, 硬盘大小为 1024G。(2) 软件环境: 软件开发环境为 Ubuntu16.04, 深度学习框架采用 Tensorflow1.0, python2.7。

### 4.1 数据集

为了提高分割准确率, 本文采用手工拍摄的 10350 张 128\*128 大小的 RGB 彩色桥梁裂缝图像作为高分辨率输入图像。对采集的高分辨率图像作四倍下采样得到 10350 张 32\*32 大小的 RGB 彩色桥梁裂缝图像作为生成网络的低分辨率输入图像。同时, 在高分辨率图像中挑选 3000 张包含各种形态的高分辨率桥梁裂缝图像, 制作对应的语义分割标签, 将挑选的 3000 张桥梁裂缝图像及其对应的标签作为判别器的分割分支输入, 辅助生成器生成超分辨率桥梁裂缝图像同时完成细小桥梁裂缝的语义分割。在分割分支中, 数据集分为三个部分, 包括训练集验证集和测试集, 其中训练集 2400 张, 占图片总数的 80%, 验证集和测试集分别 300 张, 各占图片总数的 10%。每一张图像都对应一张手工标注的语义分割标签, 图 6 中绿色区域对应裂缝位置, 黑色部分对应背景区域。

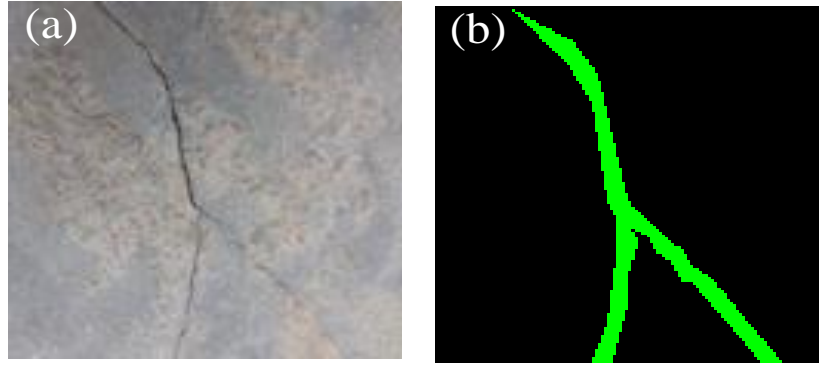


图6 桥梁裂缝图像与手工标注的语义分割标签图。(a) 桥梁裂缝图像 (b) 手工标注的分割图像

Figure 6 Semantic segmentation label image of bridge crack image and manual labeling. (a) Bridge crack image (b)

Manually labeled segmentation image

#### 4.2 分割生成式对抗网络的训练

SE-GAN 结构由两个部分构成，分别是生成器和判别器。本文中生成器结构较深，因此采用分步训练的策略。第一步单独训练生成器网络，以下采样得到的低分辨率桥梁裂缝图像作为网络的输入，对应的高分辨率桥梁裂缝图像作为标签，计算像素损失  $L_p$ ， $L_p$  值越高说明生成器生成的超分辨率图像与真实图像之间的差距越大，网络的训练过程即最小化像素损失的过程。在训练中，初始学习率  $\text{learning-rate}=0.0001$ ，学习率衰减  $\text{decay-rate}=0.1$ ，衰减步长  $\text{decay-step}=500000$ ， $\text{batch-size}=16$ ，迭代 1000000 次后将训练结果保存；第二步单独训练判别器的分割分支，将挑选出的 3000 张桥梁裂缝图作为分割分支的输入，对应的由人工标注的语义分割结果作为分割分支的标签，计算分割损失  $L_{seg}$ ， $L_{seg}$  值越高说明分割分支对输入裂缝图像的分割越不准确，本文使用 *adam* 梯度下降方式来寻求最优解，初始学习率  $\text{learning-rate}=0.0005$ ，动量项取值 0.5，迭代 100000 次后将训练结果保存；第三步分别读取前两步的训练参数，在此基础上共同训练生成器和判别器的判别分支，使用人工采集的高分辨率桥梁裂缝图像和生成器  $G$  生成的超分辨率桥梁裂缝图像作为判别器判别分支的输入，其中生成样本的标签为零，真实样本的标签为一：

- 1) 使用真实样本训练判别分支  $D_{adv}$ ；
- 2) 保持生成器  $G$  的参数不变，使用  $G$  生成的样本训练判别分支  $D_{adv}$ ；
- 3) 保持判别分支  $D_{adv}$  的参数不变，使用低分辨率图像作为输入生成超分辨率图像，将生成图像输入到  $D_{adv}$  中，计算  $D_{adv}$  的损失值，由此获得  $G$  网络更新超参数以及调整生成图像分布的梯度信息，并将梯度信息用来完成  $G$  网络的更新；
- 4) 保持分割分支  $D_{seg}$  的参数不变，使用低分辨率图像作为输入生成超分辨率图像，将

生成图像输入到  $D_{seg}$  中，计算  $D_{seg}$  的损失值后将值返回给  $G$ ，完成  $G$  网络的更新；

5) 重复 1) ~4)，迭代训练 100000 次后完成对抗网络的更新。

### 4.3 结果比较与分析

区别于传统的分割网络，本文将裂缝的类型分为三种，分别为粗大型桥梁裂缝、中型桥梁裂缝以及细小桥梁裂缝，并分别使用不同的算法对这三中桥梁裂缝进行分割。细小桥梁裂缝不同于一般的小目标，由于裂缝本身具有细长不规则的特性，且当背景较为昏暗时，模型很难提取到明显的裂缝特征，而细小桥梁裂缝作为几种裂缝中特征最不明显的一类，检测与分割工作更是难上加难。各算法对不同类型的裂缝的分割结果如图 7 所示，第一行为粗大桥梁裂缝，第二行为中型桥梁裂缝，第三行为本文重点研究的细小桥梁裂缝，图中每一列分别对应一种算法的分割效果。从图中可以看出，四种算法对于粗大型桥梁裂缝都具有较好的表现，对于中型桥梁裂缝效果稍差，对于细小桥梁裂缝，传统的算法出现了漏检以及误检的情况，由此可见，对于细小桥梁裂缝的语义分割仅使用传统的方法进行分割无法得到很好地检验效果。

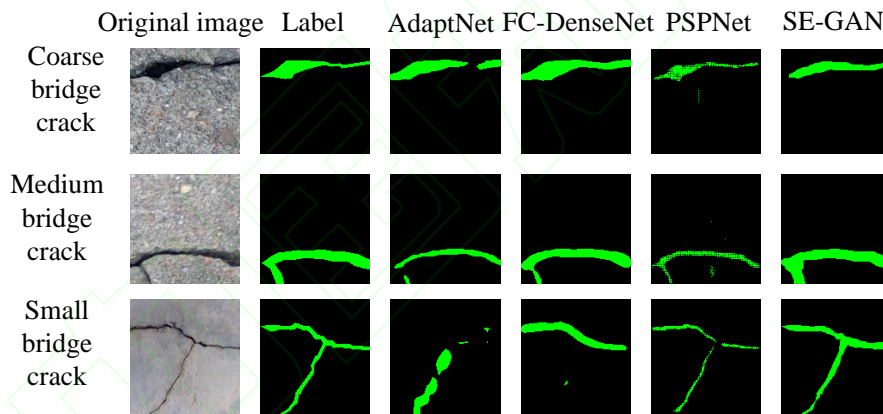


图 7 算法分割效果对比图

Figure 7 Comparison of algorithm segmentation effects

#### 4.3.1 分割结果

本文将当前分割效果比较好且主流的分割算法进行对比实验，可以看出本文方法对细小桥梁裂缝分割具有明显的优势。图 8 中三行分别对应三张不同形态的细小桥梁裂缝图，第一列为原始的高分辨率图像，第二列为原始图像对应的标签图，第三列为 AdaptSegNet 方法对三张图的分割结果，从图中可以看出对于细小桥梁裂缝，该方法存在严重的漏检现象，第四列为 FC-DenseNet103 方法对三种图的分割结果，从图中可以看出 FC-DenseNet 对中型桥梁裂缝检测结果较好细小桥梁裂缝几乎检测不出，第五列对应 PSPNet 方法的分割结果，与前两种算法相比 PSPNet 可以检测出部分细小桥梁裂缝，但是检测效果不稳定且有误检的情况

发生，最后一列为本文算法的分割结果，从图中可以看出，较其他算法本文算法对细小桥梁裂缝检测效果更好，几乎没有误检以及漏检的情况发生。同时，本文使用精确度(precision)、召回率 (recall)、F1 分数对分割结果进行了更为直观的评估，评估结果如表 1 所示。对于不同形态的细小桥梁裂缝，对比结果如图所示，在召回率、F1 分数、平均 IU 三项指标中本文算法明显高于其他三种传统方法。

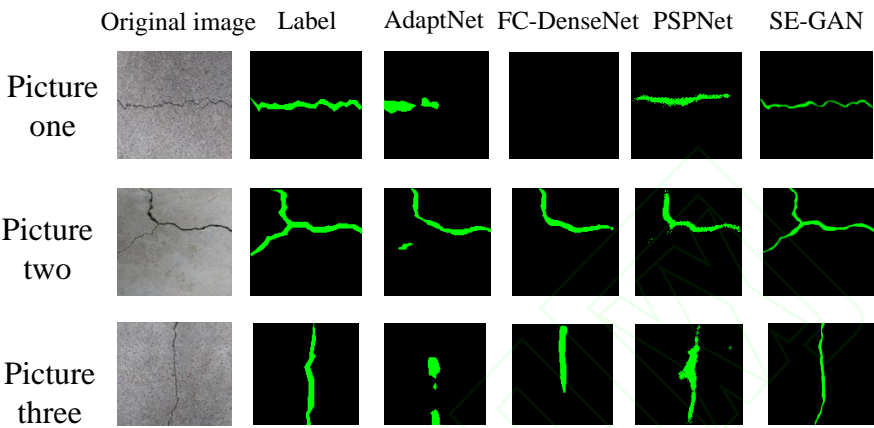


图 8 细小桥梁裂缝分割效果对比图

Figure 8 Comparison of small bridge crack splitting effects

表 1 不同算法分割效果对比表

Table 1 Comparison of different algorithm segmentation effects

Algorithm name	Precision	Recall	F1 score	Mean IU
AdaptSegNet	82.6%	50.9%	63.0%	70.5%
FC-DenseNet103	89.3%	52.8%	66.4%	72.9%
PSPNet	95.1%	56.9%	71.2%	76.0%
SE-GAN	95.5%	70.4%	81.1%	82.2%

4. 3. 2 损失函数

本文使用峰值信噪 (PSNR) 比对比仅包含对抗损失的生成式对抗网络，包含对抗损失和像素损失的生成式对抗网络以及本文中包含对抗损失、像素损失以及分割损失的生成式对抗网络生成的超分辨率图像进行对比评估，如图 9 所示橙色的曲线表示仅含对抗损失时的 PSNR 曲线，紫色的曲线表示包含对抗损失和像素损失时的 PSNR 曲线，绿色的曲线表示包含对抗损失、像素损失以及分割损失的 PSNR 曲线。从图中可以看出，绿色曲线对应的网络模型生成的超分辨率图像质量远高于其他两种模型；紫色曲线对应的网络模型生成效果略优于橙色曲线对应的网络模型，但从曲线的走向可以明显发现，加入像素损失后的模型的稳定性较强，而仅包含对抗损失的模型在 epoch=190 附近生成图像的质量急剧下跌，生成图像的质量不高，后期分割得到的有用特征变少，因此仅使用对抗网络无法得到十分准确的分割结果。

不同损失生成图片效果对比图如图 10 所示，第一列表示原高分辨率图像，第二列表示

仅使用对抗损失对网络进行优化后生成的超分辨率图像，第三列表示使用对抗损失和像素损失对网络进行优化后生成的超分辨率图像，第四列表示本文使用中使用对抗损失、像素损失以及分割损失对网络进行联合优化后生成的超分辨率图像。从图中可以明显的看出，仅使用对抗损失训练网络会出现较为明显的马赛克现象，即细节部分有许多多余的纹理特征，为后期裂缝的分割增添了许多噪声，因此大大影响了分割质量；而使用了对抗损失和像素损失训练的网络因为所有像素点具有相同的权重，因此更加的关注生成图像整体与原高分辨率图像的视觉一致性，改善了有对抗网络自动生成的纹理噪声，但使裂缝的细节信息有所丢失；本文采用的损失函数既改善了生成网络生成纹理噪声的现象，又较高程度的还原了裂缝的特征。

为了防止一个评价指标不准确的情况发生，本文还针对 SSIM、MSE 和 PSNR 指标共同给出具体的对比，对比情况如下表 2 所示。从表中可以看出，使用这三项指标对本文算法生成的超分辨率图像进行评估得到的结果均优于另外两种方法。在综上所述，使用三种损失函数同时对生成式对抗网络的超参数和生成向量提供梯度信息可以得到生成能力更强、生成的超分辨率图像质量更高的网络模型。

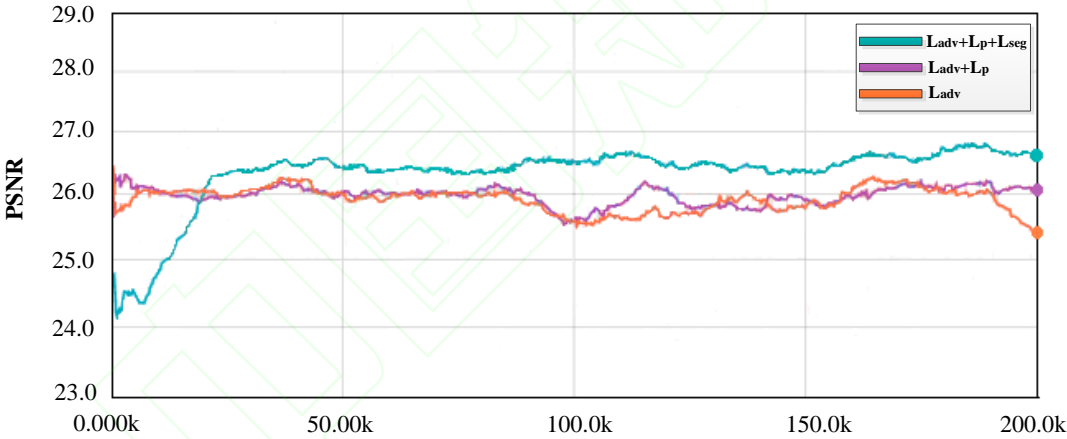


图 9 不同损失对生成图像峰值信噪比的影响曲线

Figure 9 Effect of different losses on the peak signal-to-noise ratio of the generated image



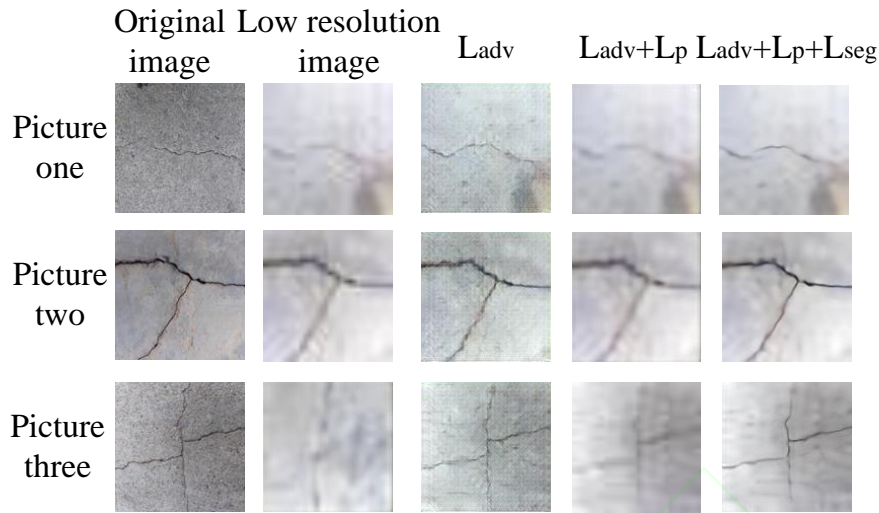


图 10 损失函数效果对比图

Figure 10 Comparison of the effect of the loss function

表 2 不同损失函数对生成图像的影响对比表

Table 2 Comparison of the effects of different loss functions on generated images

Loss function for Training	PSRN	MSE	SSIM
Ladv	25.0076	165.2917	0.8210
Ladv+Lp	26.0750	160.5380	0.85462
Ladv+Lp+Lseg	27.6830	110.8601	0.8702

#### 4.3.3 判别分支结构

为了进一步的改善生成超分辨率图像的质量，本文对判别器中的判别分支结构进行了修改，考虑到原高分辨率图像与由低分辨率图像生成的超分辨率图像之间存在一定的关联，为了更近一步衡量的两者相似性，首先计算两者之间的特征差异后，将差异输入到判别分支进行判别。实验证明，改进的判别分支可以促使生成网络生成质量更高的超分辨率图像。实验对比结果如图 11 所示，每一行对应一张桥梁裂缝图像，第一列为原高分辨率图像，第二列为将原图进行四倍下采样后用于训练的低分辨率图像使用双线性插值上采样后得到的图像，第三列为使用传统的判别器结构生成的桥梁裂缝图像，第四列为使用本文判别器结构生成的桥梁裂缝图像，从图中可以看出，使用传统的判别器生成的桥梁裂缝图像带有纵向格子，而本文使用的判别器改善了这一问题，得到了更为清晰且真实的图片。与此同时，本文使用了 PSRN、MSE、SSIM 三个图像质量评价指标对三种情况得到下的图像进行了评估，从数值中可以发现较传统判别器生成的超分辨率图像而言，使用本文判别分支结构可以得到质量更高且与原高分辨率图像更为相似的超分辨率图像。

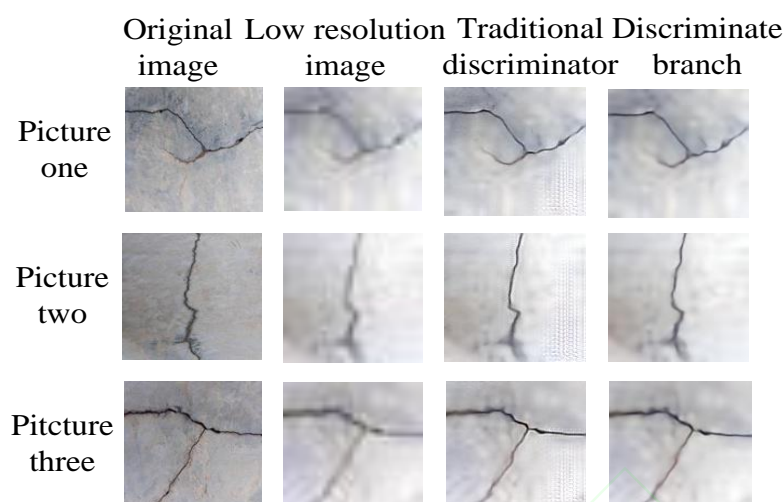


图 11 不同判别器生成超分辨图像对比图

Figure 11 Comparison of different discriminators to generate super-resolution images

表 3 不同判别器生成超分辨图像质量对比表

Table 3 Comparison of different discriminators to generate super-resolution image quality

Discriminant branch structure	PSRN	MSE	SSIM
Low resolution image	22.2210	389.9	0.3152
Traditional discriminator	26.7781	136.5404	0.8620
Discriminating branch	27.6830	110.8601	0.8702

#### 4.4 本文方法总结

在 4.3.1 节中，本文首先使用 AdaptSegNet、FC-DenseNet、PSPNet 以及 SE-GAN 分别对粗大型桥梁裂缝、中型桥梁裂缝和细小型桥梁裂缝进行分割，验证了研究细小桥梁裂缝分割方法的必要性。随后使用不同算法分别对不同形态的细小桥梁裂缝进行分割，实验结果表明，针对细小桥梁裂缝分割问题，本文提出的 SE-GAN 具有较高的分割准确率，由细小型桥梁裂缝本身提供的细节信息较少和噪声干扰导致的模型漏检情况也得到了明显的改善。SE-GAN 突破性地引入分割分支，将生成式对抗网络的超分辨率图像重建和语义分割联系起来，并增加由分割分支回传的分割损失以及由高分辨率细小桥梁裂缝图像与超分辨率细小桥梁裂缝图像相减得到的像素损失对模型进行联合优化，在 4.3.2 节中，本文对新增损失的充分性进行验证，实验证明使用 SE-GAN 中损失函数生成的图像质量更高，模型对细小桥梁裂缝细节的补充更完整。最后，SE-GAN 创新性地将判别分支的真伪判别功能改进为差异程度判别功能，更进一步提高模型生成超分辨率图像能力，并在 4.3.3 节中得到了验证。综上所述，SE-GAN 针对细小桥梁裂缝分割问题具有绝对性优势，且生成的超分辨率桥梁裂缝图像更加的逼真。

## 5 结语

本文将生成式对抗网络首次运用到了对小目标进行语义分割的领域，将生成式对抗网络的超分辨率图像重建和语义分割联系起来设计了一个新的分割生成式对抗网络（SE-GAN）。相比传统的超分辨率图像生成算法，本文生成的超分辨率细小桥梁裂缝图像质量更高，且与原高分辨率图像更为相似。与此同时，由于分割分支的引入，本文方法更偏向于生成易于被

分割模型检测到并进行正确分割的细节特征,使得在生成的超分辨率图像中,原低分辨率图像的细节特征能够被更加精准的补充完整,从而使分割分支在处理分割任务时得到更多的有用信息,实验结果证明,对于细小桥梁裂缝分割问题,较传统的分割算法,本文具有更高的分割准确率。

## 参考文献

- [1] Pang X R, Wang D P. Study on the Oblique Crack of Concrete Beam with Macro FBG Strain sensor[J]. Laser & Optoelectronics Progress, 2014, 54(12): 120603.  
庞香润, 王大鹏. 基于FBG宏应变传感器的混凝土梁斜裂缝检测研究[J]. 激光与光电子学进展, 2017, 54(12): 120603.
- [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] WANG B, WANG X, CHEN F, et al. Identification of pavement cracks in aerial imagery[J]. 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.
- [4] Li Q, Liu X. Novel approach to pavement image segmentation based on neighboring difference histogram method[J]. IEEE Computer Society, 2008, 2:792-796.
- [5] Landstrom A, Thurley M J. Morphology-based crack detection for steel slabs[J]. IEEE Journal of selected topics in signal processing, 2012, 6(7):866-875.
- [6] Lu Y Y, Lu C W, Tang C K. Online Video Object Detection Using Association LSTM[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2017: 2363-2371.
- [7] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2017: 2999-3007.
- [8] Huang C, Chen P, Yang X. REDBEE: A visual-inertial drone system for real-time moving object detection[C]//Proceedings of the IEEE Computer Society Conference on IEEE, 2017: 1725-1731.
- [9] Zhao S S, He N. Pavement Crack Detection Based on Convolutional Neural Network[J]. Journal of Sensors and Microsystems, 2017, 36(10): 147-153.  
赵珊珊, 何宁. 基于卷积神经网络的路面裂缝检测[J]. 传感器与微系统, 2017, 36(10): 147-153.
- [10] Jegou S, Drozdal M, Vazquez D, et al. The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2017: 1175-1183.
- [11] 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.
- [12] Bell S, Zitnick C L, Bala K. Inside-outside net: Detecting objects in context with skip pooling and recurrent neural networks[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016: 2874-2883.
- [13] Benenson R, Omran M, Hosang J. Ten years of pedestrian detection, what have we learned?[J]. Springer International Publishing, 2014, 8926: 613-627.
- [14] Zhang C, Chen Y. Target Detection Based on Difficult Sample Mining under Residual Network[J]. Laser & Optoelectronics Progress, 2018, 55(10): 101003.  
张超, 陈莹. 残差网络下基于困难样本挖掘的目标检测[J].激光与光电子学进展, 2018, 55(10):101003.
- [15] Chen X, Du K K, Zhu Y. 3d object proposals for accurate object class detection[C]//Proceedings of Neural Information Processing Systems, 2015: 424-432.
- [16] Zhu J W, Liu W H, Yin J F, et al. Infrared weak target region detection algorithm based on improved complexity[J]. Laser & Optoelectronics Progress, 2018, 55(10):101006.  
朱婧文, 刘文好, 印剑飞, 等. 基于改进复杂度的红外弱小目标区域检测算法[J].激光与光电子学进展, 2018, 55(10): 101006.
- [17] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multi-box detector[C]//Proceedings of European Conference on Computer Vision, 2016: 21-37.
- [18] Yang F, Choi W, Lin Y. Exploit all the layers: Fast and accurate cnn object detector with scale dependent pooling and cascaded rejection classifiers[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016: 2129-2137.
- [19] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(9):1904-1916.
- [20] Li H, Lin Z, Shen X. A convolutional neural network cascade for face detection[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015: 5325-5334.
- [21] Wu Y, Liu Y, Li J. Traffic sign detection based on convolutional neural networks[C]//Proceedings of International Joint Conference on Neural Networks, 2014: 1-7.
- [22] Goodfellow L J, Pouget A J, Mirza M, et al. Generative adversarial nets[C]//Proceedings of Neural Information Processing Systems, 2014: 2672-2680.
- [23] Evan S, Jonathan L, Trevor D. Fully convolutional networks for semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2015, 79(10): 3431-3440.
- [24] Yi H S, Tsai I, Hung W C, et al. Learning to Adapt Structured Output Space for Semantic Segmentation//Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7472-7481.
- [25] Zhao H S, Shi J P, Qi X J, et al. Pyramid Scene Parsing Network[C]//Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 6230-6239.
- [26] Ledig C, Theis L, Huszar F. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[C]//Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, 2017:105-114.

网络首发:

**标题:** 基于生成式对抗网络的细小裂缝分割方法

**作者:** 李良福, 胡敏

**收稿日期:** 2018-10-31

**录用日期:** 2018-12-13

**DOI:** 10.3788/lop56.101004

**引用格式:**

李良福, 胡敏. 基于生成式对抗网络的细小裂缝分割方法[J]. 激光与光电子学进展, 2019, 56(10): 101004.

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

**基于模糊不变卷积神经网络的遥感飞机识别**

刘坤 苏彤 王典

上海海事大学信息工程学院, 上海 200135

激光与光电子学进展, 2018, 55(8): 082001

**深度学习在视觉定位与三维结构恢复中的研究进展**

鲍振强 李艾华 崔智高 袁梦

火箭军工程大学, 陕西 西安 710025

激光与光电子学进展, 2018, 55(5): 050007

**深度学习与激光诱导荧光在假酒识别中的应用**

来文豪 周孟然 王亚 胡锋 李大同 赵舜

安徽理工大学电气与信息工程学院, 安徽 淮南 232000

激光与光电子学进展, 2018, 55(4): 043001

**基于区域全卷积网络结合残差网络的火焰检测方法**

洪伟 李朝锋

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(4): 041011