

# 基于生成式对抗网络的细小桥梁裂缝分割方法

李良福, 胡敏\*

陕西师范大学计算机科学学院, 陕西 西安 710119

**摘要** 提出了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型桥梁裂缝图像, 再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实现准确分割, 与传统的分割方法相比, 该方法的分割召回率提高了 6%, 平均交并比提高了 10%。

**关键词** 图像处理; 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

中图分类号 TP391.9

文献标识码 A

doi: 10.3788/LOP56.101004

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

Li Liangfu, Hu Min\*

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

**Abstract** For cracks in small bridges, a segmentation method is proposed based on a generative adversarial network. This method introduces a segmental branch into the discriminator structure and combines the generative confrontation network with the semantic segmentation network. In addition, the method is capable of super-resolution image reconstruction and segmentation. To solve the problem of small-bridge-crack segmentation, this method transforms low-resolution small-bridge-crack images into super-resolution coarse-bridge-crack images, which are then segmented. The experimental results show that the proposed method facilitates the identification of small-bridge-crack and its segmentation is accurate. Compared with the traditional segmentation method, the recall rate and mean intersection over union of this method are improved by 6% and 10%, respectively.

**Key words** image processing; 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重要。桥梁作为当今社会交通枢纽的重要组成部分, 不仅承担着运输重任, 还关系着运输人员的安危。然而, 桥梁结构在长期使用中难免产生各种各样的损坏, 这会造成桥梁结构抗力衰减并产生安全隐患, 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, 主要原因是沥青路面的疲劳以及不利的大气条件可能导致材料收缩; 此外, 裂缝也可能由沥青混合物质量差造成的结构路面阻力而引起。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必不可少。现实生活中, 由于裂缝分布杂乱无规律, 很多细小裂缝容易被周边障碍物干扰, 从而造成漏检误检, 这会产生极大的安全隐患。为了避免这类情况的发生, 对细小桥梁裂缝检测的研究十分重要。

收稿日期: 2018-10-31; 修回日期: 2018-12-04; 录用日期: 2018-12-13

基金项目: 国家自然科学基金(61573232, 61401263)

\* E-mail: longford@xjtu.edu.cn

101004-1

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<sup>[4]</sup>提出了一种基于相邻差分直方图的裂缝分割算法,这类基于阈值分割的裂缝识别方法易受光照、噪声和纹理的影响,检测效果往往不稳定。针对这一问题,Landstrom等<sup>[5]</sup>结合形态学处理和逻辑回归算法对裂缝进行检测,利用统计学分类方法滤除噪声,提高了检测精度。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桥梁裂缝的检测中。赵珊珊等<sup>[9]</sup>提出使用卷积神经网络对裂缝进行检测,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,Jégou等<sup>[10]</sup>提出全卷积语义分割网络,该模型可以获得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<sup>[11]</sup>在解码器中使用池化层索引对特征图进行上采样,使得待分割图像可以保持高频细节的完整性。但是,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,以及对位置像素信息与周围接近等原因,上述传统的裂缝检测方法对细小桥梁裂缝的检测效果往往不佳。针对小目标检测问题,通常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<sup>[12-18]</sup>。还可以通过修改网络模型,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方法<sup>[19-21]</sup>。

以上方法均试图通过数据增强或者增加特征维度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,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,再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),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,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,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共同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,对判别器结构中的判别分支的输出部分进行了改进,使得判别器回传的信息更准确;在生成器结构中,引入了像素损失,使得生成器可以生成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的高斯白噪声,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,该过程可表示为

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

式中: $\min_G \max_D V(D, G)$ 为损失函数; $V$ 为损失函数; $E$ 表示均值; $D$ 为判别器模型; $G$ 为生成器模型; $P$ 为向量分布类型; $x$ 为真实图像; $z$ 为输入的随机噪声向量。判别器需要判别输入的是真实图像还是生成器生成的图像,当输入为真实图像时, $D(x)$ 的值趋近于1;当输入为生成器生成的图像时, $D(x)$ 的值趋近于零。生成器则需尽可能调整 $z$ 的分布,使得 $G(z)$ 与 $x$ 的分布差异最小化,即使 $D[G(z)]$ 趋近于1。由于生成式对抗网络具有两个子网络模型,因此采用逐步交叉的训练方式。训练判别器时的损失函数为

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

此时,默认生成器性能已达到最优,只需训练判别器的分类性能。训练生成器的损失函数时,默认判别器具有最优的判别能力,即  $E_{x \sim P_{\text{data}}(x)} \{\text{lb}[D(x)]\}$  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为

$$\min_G V(D, G) = E_{z \sim P_{G(z)}} \{\text{lb}\{1 - D[G(z)]\}\}. \quad (3)$$

由于生成式对抗网络具有模拟生成图像的功能,这里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型桥梁裂缝图像。

## 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,接着,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,此时分割损失函数可表示为

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

式中: $x_{ij}$ 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。

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地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,同时易受噪声干扰,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,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 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。

## 3 细小桥梁裂缝语义分割生成式对抗网络

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。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,无法学习到更多两者之间的潜在关系,为了更深层次地挖掘两者之间的内在关联,同时增强判别器非线性拟合能力,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不规则,且其对应的像素分布占图像整体像素的比例少,这里将细小桥梁裂缝图像进行四等分,得到 16 个低分辨率子图,低分辨率子图切分示意图如图 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,并对重建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。

### 3.1 SE-GAN 判别器

不同于传统的判别器,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。判别分支与传统的判别器功能相同,用于判别输入是真实图像还是生成器生成的图像。语义分割分支的作用有两个,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,判别器结构示意图如图 2 所示。判别器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输入,经过特征提取分类,最终输出。

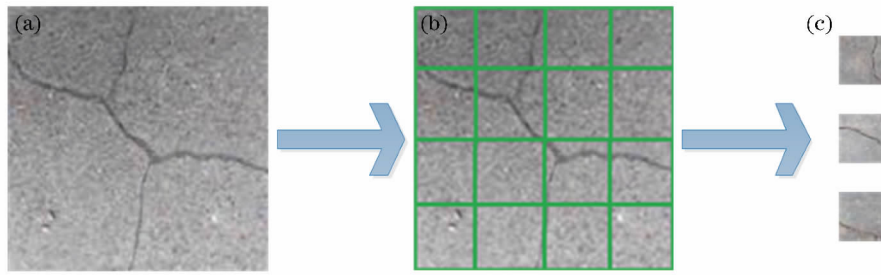


图1 低分辨率子图切分示意图。(a)原始图像;(b)切分示意图;(c)切分成的低分辨率子图

Fig. 1 Schematic of low-resolution subgraph segmentation. (a) Original image; (b) schematic of segmentation; (c) low-resolution subgraphs after segmentation

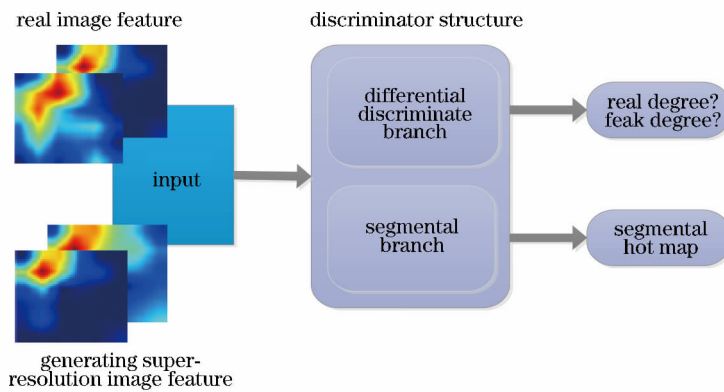


图2 判别器结构示意图

Fig. 2 Structural diagram of 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之间的内在联系,无法将判别器功能最大化。为了更好地挖掘真实图像分布与生成图像分布之间的潜在相关性,引导网络生成的图像更完美地拟合真实图像,引入差异判别分支,使判别器不再分裂化地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之间的相对差异而言的,两者之间的差异可以表示为

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

式中: $F$ 表示判别网络特征提取过程,因此 $F(\mathbf{x})$ 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; $F[G(\mathbf{z})]$ 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。传统

的判别器输出经激活函数 $D = \sigma[F(\cdot)]$ 映射为 $[0, 1]$ 区间的概率值,其中 $\sigma$ 为激活函数。对于真实的高分辨率图像判别器,其输出为 $D\{\text{dis}[\mathbf{x}, G(\mathbf{z})]\}$ ,对应生成的超分辨率图像判别器输出为 $D\{\text{dis}[G(\mathbf{z}), \mathbf{x}]\}$ ,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1,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定义为

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

式中: $L_{\text{adv}}$ 为对抗损失。在训练判别器时,损失函数为

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

训练生成器时,损失函数为

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

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, $G(\mathbf{z})$ 为生成器生成的超分辨率

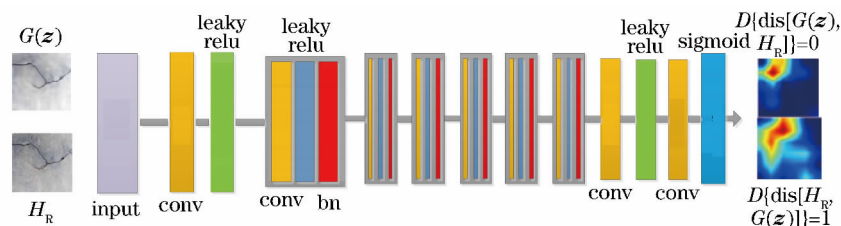


图3 差异判别分支结构

Fig. 3 Structural diagram of Discrimination branch

桥梁裂缝图像,  $H_R$  为真实的高分辨率桥梁裂缝图像。分支的结构主要由卷积层、激活层、归一化层构成。

### 3.1.2 分割分支

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,分割分支结构图如图4所示,第一行表示网络的大致结构,由 dense block 结构、transition down 结构、卷积层、反卷积层构成。其中 dense block、layer 以及 transition down 结构的组成如图4第二行所示,箭头表示将箭头起点所在层的输出与箭头终点所在层的输出进行内容拼接。拼接后形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,将超分辨率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的过程可表示为

$$L_{\text{seg}} = \sum_{ij} [y_{ij} \text{lb } x_{ij} + (1 - y_{ij}) \text{lb}(1 - x_{ij})], \quad (9)$$

式中:  $L_{\text{seg}}$  为分割损失;  $x_{ij}$  为网络预测热图对应  $i, j$  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。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并对生成的超分辨率图像进行细化,促使生成器生成更易于实现分割分支检测的细节信息。

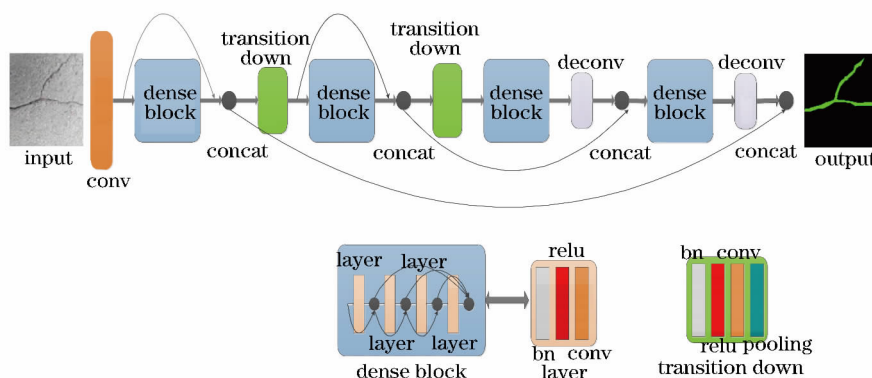


图4 分割分支结构图

Fig. 4 Structural diagram of segmentation branch

## 3.2 SE-GAN 生成器

### 3.2.1 SE-GAN 生成器网络结构

在处理超分辨率生成问题中,往往通过在生成网络中加入残差结构增加模型的复杂性,同时又避免了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,从而可生成与高分辨率图像分布类似的超分辨率图像。超分辨率生成式对抗网络(SR-GAN)<sup>[23]</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,增添了非线性特征,从而使模型的学习能力

大幅提升。

为了进一步提高生成器生成超分辨率图像的能力,增强网络的非线性拟合功能,以使深层卷积可以拟合得到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,这里将 SR-GAN 中生成网络的 16 个残差块扩展成 32 个,生成的网络结构对比图如图5所示。网络由卷积层、激活层、归一化层、反卷积层构成,图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。



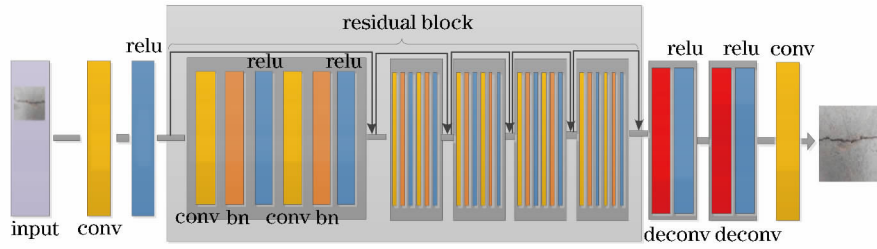


图5 生成网络结构对比图

Fig. 5 Comparison chart of generated network structures

### 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,往往会出现虽然网络损失很低,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超分辨率图像的优化梯度信息,而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,为了弥补像素分布差异而造成的视觉缺憾,在对生成器生成的超分辨率图像进行优化时,加入了像素损失项,可表示为

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

式中: $L_p$  为像素损失。像素损失与对抗损失、分割损失结合,共同控制超分辨率图像的分布,得到 SE-GAN 模型的损失函数为

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

式中: $L_{adv}$  为对抗损失; $L_{seg}$  为分割损失; $\lambda$  为分割参数; $\mu$  为像素参数。

## 4 实验与结果分析

具体的运行操作环境分为两部分:硬件部分和软件环境。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 3.31 GHz,主机内存为 8 G,硬盘大小为 1024 G。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,采用手工拍摄的 10350 张 128 pixel×128 pixel 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进行 4 倍下采样得到 10350 张 32 pixel×32 pixel 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 3 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,桥梁裂缝图像与手工标注的语义分割标签图如图 6 所示。图 6 (b) 中的绿色区域对应裂缝位置,黑色部分对应背景区域。

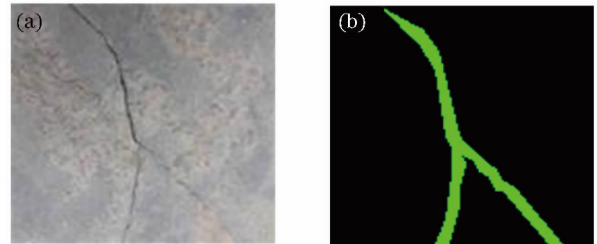


图6 桥梁裂缝图像与手工标注的语义分割标签图。

(a) 桥梁裂缝图像; (b) 手工标注的分割图像

Fig. 6 Bridge-crack image and manually labelled semantic segmentation image. (a) Bridge-crack image; (b) manually labelled 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。实验中生成器结构较深,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,初始学习率为 0.0001,学习率衰减为 0.1,衰减步长为 500000,批次训练图像张数为 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,这里使用 adam 梯度下降方式寻求最优解,初始学习率为 0.0005,动量项取值 0.5,迭代 100000 次后保存训练结果;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模型  $G$  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 1。实现步骤为:

- 1) 使用真实样本训练判别分支  $D_{adv}$ 。
- 2) 保持生成器模型  $G$  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) 重复步骤 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 3 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,分别使用不同的算法对这 3 种桥梁裂缝进行分割。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。对不同类型的裂缝的各算法分割效果对比如图 7 所示,第 1 行为粗大型桥梁裂缝,第 2 行为中型桥梁裂缝,第 3 行为重点研究的细小桥梁裂缝。图 7 (a) 为原始图像;图 7 (b) 为标签图;图 7 (c) 为 AdaptNet (Adapt Structured Output Space for Semantic Segmentation)<sup>[24]</sup> 方法对应的分割效果;图 7 (d) 为 FC-DenseNet (Fully convolutional networks for semantic segmentation)<sup>[25]</sup> 方法对应的分割效果;图 7 (e) 为金字塔场景解析网络 (PSPNet)<sup>[26]</sup> 方法对应的分割效果;图 7 (f) 为本文方法对应的分割效果。从图 7 中可以看出,4 种算法对于粗大型桥梁裂缝都有较好的表现,对于中型桥梁裂缝效果稍差,而对于细小桥梁裂缝,传统的算法出现了漏检以及误检的情况。由此可见,仅使用传统的方法进行细小桥梁裂缝的语义分割无法得到很好的检验效果。

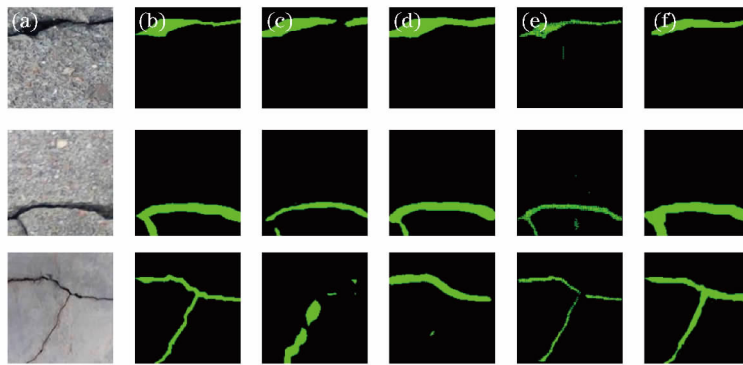


图 7 算法分割效果对比图。(a)原始图像;(b)标签图;(c) AdaptNet;(d) FC-DenseNet;(e) PSPNet;(f) SE-GAN

Fig. 7 Comparison of segmentation results using different algorithms. (a) Original images; (b) labelled images; (c) AdaptNet; (d) FC-DenseNet; (e) PSPNet; (f) SE-GAN

#### 4.3.1 分割结果

将所提算法与分割效果比较好且主流的分割算法进行对比,得到细小桥梁裂缝分割结果对比图如图 8 所示。由图 8 可以看出,所提方法对细小桥梁裂缝分割具有明显的优势。图 8 中 3 行分别对应 3 张不同形态的细小桥梁裂缝图。图 8(a)为原始的

高分辨率图像;图 8(b)为原始图像对应的标签图;图 8(c)为 AdaptNet 方法对 3 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;图 8(d)为 FC-DenseNet 方法对 3 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但对细小桥梁裂缝几乎检测

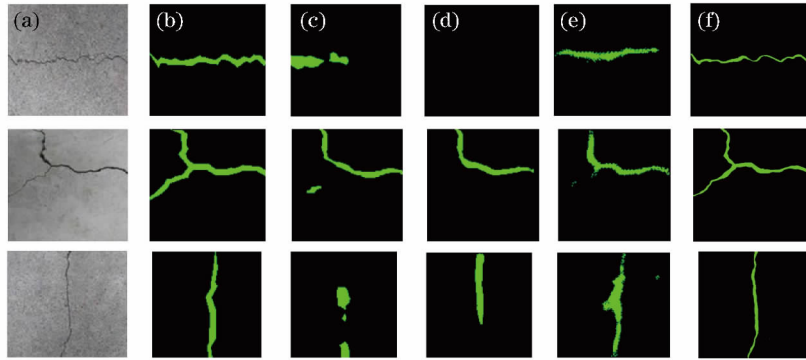


图8 细小桥梁裂缝分割结果对比图。(a)原始图像;(b)标签图;(c) AdaptNet;(d) FC-DenseNet;(e) PSPNet;(f) SE-GAN

Fig. 8 Comparison of small-bridge-crack segmentation results. (a) Original images ; (b) labelled images ; (c) AdaptNet ; (d) FC-DenseNet ; (e) PSPNet ; (f) SE-GAN

不出;图8(e)为PSPNet方法的分割结果,与前两种算法相比,PSPNet可以检测到部分细小桥梁裂缝,但是检测效果不稳定且有误检的情况发生;图8(f)为所提算法的分割结果,从图中可以看出,相较于其他算法,所提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,使用精确度、召回率、 $F_1$ 分数对分割结果进行了更为直观的评估,不同算法分割效果对比如表1所示。由表1可见,对于不同形态的细小桥梁裂缝,本文算法的召回率、 $F_1$ 分数、平均交并比(IoU)3项指标均明显高于其他3种传统方法。

表1 不同算法分割效果对比

Table 1 Comparison of segmentation results using different algorithms %

Algorithm	Precision	Recall	$F_1$ score	Mean IOU
Adapt Net	82.6	50.9	63.0	70.5
FC-DenseNet	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

使用峰值信噪比(PSNR)对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,得到不同损失对生成图像峰值信噪比的影响,如图9所示。图9中方框曲线表示仅含对抗损失时的PSNR曲线,三角曲线表示包含对抗损失和像素损失时的PSNR曲线,圆点曲线表示包含对抗损失、像素损失以及分割损失的PSNR曲线。可以看出,原点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;三角曲线对应的网络模型生成效

果略优于方框曲线对应的网络模型,但从曲线的走向可以明显发现,加入像素损失后模型的稳定性较强,而仅包含对抗损失的模型在Epoch为190附近生成图像的质量急剧下降,生成图像的质量不高,后期分割得到的有用特征变少,因此仅使用对抗损失的对抗网络无法得到十分准确的分割结果。

损失函数效果对比图如图10所示,图10(a)为原高分辨率图像,图10(b)为仅使用对抗损失对网络进行优化后生成的超分辨率图像,图10(c)为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,图10(d)为使用对抗损失、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10中可以明显看出,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,这为后期裂缝的分割增加了许多噪声,极大地影响了分割质量;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,又较高程度地还原了裂缝的特征。

为了防止单个评价指标对生成图像评价不准确的情况,使用PSNR、均方误差(MSE)和结构相似性(SSIM)指标共同进行评价。不同损失函数对生成图像的影响如表2所示。从表2中可以看出,所提算法生成的超分辨率图像的3项指标均优于另外两种方法。综上所述,使用3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。



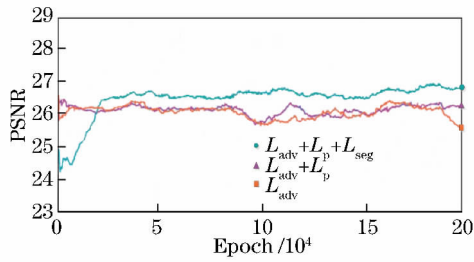


图9 不同损失对生成图像 PSNR 的影响

Fig. 9 Effect of loss on PSNR of generated image

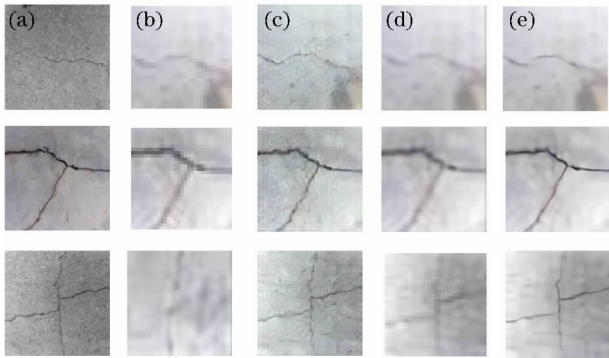
图10 损失函数效果对比图。(a)原始图像;(b)低分辨率图像;(c)  $L_{adv}$ ;(d)  $L_{adv} + L_p$ ;(e)  $L_{adv} + L_p + L_{seg}$ Fig. 10 Comparison of effect of loss function. (a) Original images; (b) low-resolution images; (c)  $L_{adv}$ ; (d)  $L_{adv} + L_p$ ; (e)  $L_{adv} + L_p + L_{seg}$ 

表2 不同损失函数对生成图像的影响

Table 2 Effect of loss function on generated image

Loss function for training	PSNR	MSE	SSIM
$L_{adv}$	25.0076	165.2917	0.8210
$L_{adv} + L_p$	26.0750	160.5380	0.8546
$L_{adv} + L_p + L_{seg}$	27.6830	110.8601	0.8702

#### 4.3.3 判别分支结构

为了进一步改善生成的超分辨率图像的质量,对判别器中的判别分支结构进行修改,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,为了进一步衡量两者的相似性,在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。不同判别器生成的超分辨率图像如图11所示。每行对应一张桥梁裂缝图像,图11(a)为原高分辨率图像,图11(b)为将原图进行4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,图11(c)为使用传统的判别器结构生成的桥梁裂缝图像,图11(d)为使用所提判别器结构生成的桥梁裂缝图像。由图11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,而使用所提出的判别器可以改善这一问题,从而得到更为清晰且真实的图片。与此同时,使用PSNR、MSE、SSIM 3个图像质量评价指标对3种情况下得到的图像进行评估,不同判别器生成的超分辨率图像质量如表3所示。由表3可见,相较于传统判别器生成的超分辨率图像,使用所提的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。

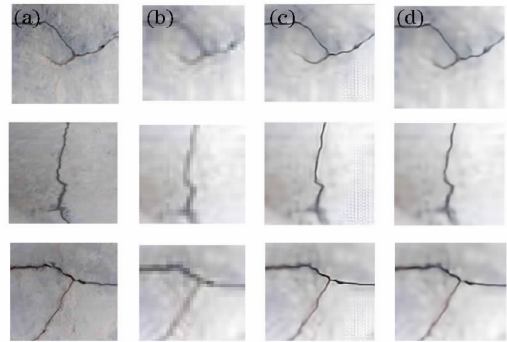


图11 不同判别器生成的超分辨率图像。(a)原始图像;(b)低分辨率图像;(c)传统判别器;(d)差异判别器

Fig. 11 Super-resolution images generated by different discriminators. (a) Original images; (b) low-resolution images; (c) traditional discriminator; (d) discrimination branch

表3 不同判别器生成的超分辨率图像质量

Table 3 Quality of super-resolution images generated by different discriminators

Discriminator	PSNR	MSE	SSIM
Low-resolution image	22.2210	389.9000	0.3152
Traditional discriminator	26.7781	136.5404	0.8620
Discrimination branch	27.6830	110.8601	0.8702

#### 4.4 小结

在4.3.1节中,首先使用Adapt Net、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,提出的SE-GAN方法具有较高的分割准确率,由于细小桥梁裂缝本身提供的细节信息较少,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用以对模型进行联合优化。在4.3.2节中,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,进一步提高了模型生成超分辨率图像能力,这一结论在 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逼真。

## 5 结 论

将生成式对抗网络运用到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 SE-GAN。相比传统的超分辨率图像生成算法,利用所提算法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,所提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地补充完整,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,所提算法具有更高的分割准确率。

## 参 考 文 献

- [1] Pang X R, Wang D P. Oblique crack monitoring of concrete beam with macro FBG strain sensor [J]. *Laser & Optoelectronics Progress*, 2017, 54 (12): 120603.  
庞香润, 王大鹏. FBG 宏应变传感器的混凝土梁斜裂缝监测[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.* Pavement crack recognition based on aerial image [J]. *Acta Optica Sinica*, 2017, 37(8): 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(8): 0810004.
- [4] Li Q Q, Liu X L. Novel approach to pavement image segmentation based on neighboring difference histogram method [C]// 2008 Congress on Image and Signal Processing, May 27-30, 2008, Sanya, Hainan, China. New York: IEEE, 2008: 792-796.
- [5] Landstrom A, Thurley M J. Morphology-based crack detection for steelslabs [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]// 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2363-2371.
- [7] Lin T Y, Goyal P, Girshick R, *et al.* Focal loss for dense object detection [C]// 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 2999-3007.
- [8] Huang C, Chen P, Yang X, *et al.* REDBEE: A visual-inertial drone system for real-time moving object detection [C]// 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), September 24-28, 2017, Vancouver, BC, Canada. New York: IEEE, 2017: 1725-1731.
- [9] Zhao S S, He N. Road surface crack detection based on CNN [J]. *Transducer and Microsystem Technologies*, 2017, 36(11): 135-138.  
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(11): 135-138.
- [10] Jégou 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 (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1175-1183.
- [11] 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.
- [12] Bell S, Zitnick C L, Bala K, *et al.* Inside-outside net: detecting objects in context with skip pooling and recurrent neural networks [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 2874-2883.
- [13] Benenson R, Omran M, Hosang J, *et al.* Ten years of pedestrian detection, what have we learned? [C]// European Conference on Computer Vision,

- September 6-7 and 12, 2014, Zurich, Switzerland. Switzerland AG; Springer International Publishing, 2014; 613-627.
- [14] Zhang C, Chen Y. Object detection based on hard examples mining using 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 Z, Kundu K, Zhu Y K, *et al.* 3d object proposals for accurate object class detection[C]// *Proceedings of Advances in Neural Information Processing Systems 28 (NIPS 2015)*, December 7-12, 2015, Montreal, Canada. San Francisco: Margan Kaufmann, 2015; 424-432.
- [16] Zhu J W, Liu W H, Yin J F, *et al.* Infrared small target regions detection based on improved image 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 multibox detector[C]// *European conference on computer vision*. Springer, October 11-14, 2016, Amsterdam, The Netherlands. Switzerland AG; Springer International Publishing, 2016; 21-37.
- [18] Yang F, Choi W, Lin Y Q. Exploit all the layers: fast and accurate CNN object detector with scale dependent pooling and cascaded rejection classifiers[C]// *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016; 2129-2137.
- [19] He K M, Zhang X Y, 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 X, Lin Z, Shen X H, *et al.* A convolutional neural network cascade for face detection[C]// *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015; 5325-5334.
- [21] Wu Y H, Liu Y L, Li J M, *et al.* Traffic sign detection based on convolutional neural networks[C]// *The 2013 International Joint Conference on Neural Networks (IJCNN)*, August 4-9, 2013, Dallas, TX, USA. New York: IEEE, 2013; 1-7.
- [22] Goodfellow L J, Pouget A J, Mirza M, *et al.* Generative adversarial nets [C]// *Proceedings of Advances in Neural Information Processing Systems 27 (NIPS 2014)*, December 8-13, 2014, Montreal, Canada. San Francisco: Margan Kaufmann, 2014; 2672-2680.
- [23] Ledig C, Theis L, Huszár F, *et al.* Photo-realistic single image super-resolution using a generative adversarial network[C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017; 105-114.
- [24] Tsai Y H, Hung W C, Schuster S, *et al.* Learning to adapt structured output space for semantic segmentation[C]// *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018; 7472-7481.
- [25] Simon J, Michal D, David V, *et al.* The one hundred layers tiramisu: fully convolutional denseNets for semantic segmentation[C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017; 1175-1183.
- [26] Zhao H S, Shi J P, Qi X J, *et al.* Pyramid scene parsing network [C]// *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017; 6230-6239.