

DOI: 10.11835/j.issn.2096-6717.2021.147



开放科学(资源服务)标识码(OSID):



基于特征金字塔网络的超大尺寸图像裂缝识别检测方法

舒江鹏^{a,b}, 李俊^{a,c}, 马亥波^a, 段元锋^a, 赵唯坚^a

(浙江大学 a. 建筑工程学院; b. 平衡建筑研究中心; c. 浙江大学建筑设计研究院有限公司, 杭州 310058)

摘要:基于图像分析的裂缝自动检测识别一直是桥梁结构健康检测的热点问题之一。深度学习作为裂缝检测的重要解决方法,需要大量数据支持。公开数据集提供的小尺寸裂缝图像不足以解决超大尺寸细长裂缝图像的检测问题。提出一个基于特征金字塔深度学习网络的超大尺寸图像中细长裂缝的检测方法。通过对编码器提取的4个不同层次的特征图分别进行预测,网络能够实现对细小裂缝的高精度分割。试验使用120张大小为 $3\,264 \times 4\,928$ 像素的桥钢箱梁表面裂缝图像对特征金字塔网络进行训练、测试;并将获得的训练模型与通过双线性插值方法缩放图像至 $1\,600 \times 2\,400$ 像素和 $2\,112 \times 3\,168$ 像素两种规格生成的数据集训练后的模型进行对比。结果表明:该方法在对比测试中能够获得最高的裂缝检测交并比(IoU)为0.78,最低的Dice Loss为0.12。测试中,裂缝检测图像显示,缩放图像会导致部分裂缝信息的丢失,该方法能稳定地保留裂缝信息,并实现复杂背景下超大尺寸图像中细长裂缝的高精度自动检测。

关键词:裂缝检测;深度学习;超大尺寸图像;特征金字塔网络

中图分类号:U446.2 文献标志码:A 文章编号:2096-6717(2022)03-0029-08

Crack detection method based on feature pyramid network for super large-scale images

SHU Jiangpeng^{a,b}, LI Jun^{a,c}, MA Haibo^a, DUAN Yuanfeng^a, ZHAO Weijian^a

(a. College of Civil Engineering and Architecture; b. Center for Balance Architecture; c. The Architectural Design & Research Institute of Zhejiang University Co., Ltd., Zhejiang University, Hangzhou 310058, P. R. China)

Abstract: Automatic crack detection based on image analysis is a hot issue in bridge structure health inspection. Crack segmentation based on deep learning is a significant solution, which needs lots of database. The small-scale crack images of open datasets are not sufficient for detection of long and thin cracks of super large-scale image. The study proposes an automatic crack detection method for super large-scale images, based on feature pyramid network. Through four different feature maps in various sizes, the proposed network yields predictions, respectively, which means a highly precise crack segmentation. Experiments are carried on 120 steel box girder crack images in a resolution of $3\,264 \text{ pixels} \times 4\,928 \text{ pixels}$. These images are used to train and test the network. The comparison study is conducted between the

收稿日期:2021-04-28

基金项目:国家自然科学基金(U1709216);国家重点研发计划(2018YFE0125400)

作者简介:舒江鹏(1987-),男,博士,研究员,主要从事结构智能检测和评估,E-mail:jpeshu@zju.edu.cn。

Received:2021-04-28

Foundation items: National Natural Science Foundation of China (No. U1709216); The National Key Research and Development Program of China (No. 2018YFE0125400)

Author brief: SHU Jiangpeng(1987-), PhD, professor, main research interests: inspection and assessment of structures, E-mail: jpeshu@zju.edu.cn.

proposed method and the models trained with crack images resized into $1\ 600\ \text{pixels} \times 2\ 400\ \text{pixels}$ and $2\ 112\ \text{pixels} \times 3\ 168\ \text{pixels}$ with bilinear interpolation algorithm. The results show that our method can achieve the highest crack Intersection over Union (IoU) of 0.78, and the lowest Dice Loss of 0.12 in the comparison study. The predictions of crack images in testing indicate that resizing images sometimes result in the loss of crack information, and our method can maintain the detail of cracks and detect cracks of super large-scale images automatically and precisely.

Keywords: crack detection; deep learning; super large-scale images; feature pyramid network

裂缝作为桥梁早期损伤破坏的重要表征之一,及时发现并对其进行检测是桥梁维护工作中的重点^[1]。常规的人工勘测不仅存在检查盲区,同时也需要耗费大量的人力物力,有时还存在安全隐患。随着摄影、遥感、无人机等图像采集技术的快速发展,基于数字图像处理技术的裂缝检测方法得到广泛关注^[2]。通过使用模糊多级中值滤波^[3]、灰度矫正^[4]、非下采样轮廓波变换^[5]等数字图像处理技术去除裂缝图像上的噪声并优化图像质量,裂缝的几何特征和尺寸信息能够被提取出来。在面对复杂光照、背景等外界因素时,传统的数字图像处理技术不能很好地消除干扰。张维峰等^[6]综合应用了不同图像处理算法开发的适用于桥梁缺陷较远距离图像的检测分析软件,在面对细小裂缝的检测时也遇到了瓶颈。

随着计算机领域深度学习的兴起,桥梁裂缝图像处理迎来了新的机遇和挑战。众多学者从实际情况出发,对数据方法、智能算法进行改进,提高了模型的分析能力,推动了人工智能在桥梁检测应用中的提高^[7]。比如,Zhang 等^[8]将深度卷积神经网络(Deep Convolutional Neural Network,DCNN)应用到裂缝信息提取研究中,该算法展现出了比传统方法更高的准确度。Zhang 等^[9]提出网络 CrackNet,能够实现裂缝的自动检测提取。数据集作为数据驱动型深度学习算法的核心组成部分,也得到了广泛的关注。服务于像素级别下桥梁裂缝检测任务的公开数据集^[10]包含了超过一万张的裂缝图像样本,但这些图像的尺寸都只有 256×256 像素。包括其他的公开数据集^[11],服务于裂缝检测任务的裂缝图像尺寸不超过 500×500 像素。在工程应用中,使用高清相机或者无人机拍摄的图像一般都会超过 $1\ 920 \times 1\ 080$ ($1\ 080\ \text{p}$)。一般认为,4k UHD ($3\ 840 \times 2\ 160$)作为新的工业标准,视为高分辨率(High-Resolution);而超过这一尺寸则为超大尺寸,或者是超高分辨率(Very High-Resolution)^[12]。一些超大

尺寸的裂缝图像甚至会超过 $4\ 096 \times 3\ 112$ ^[13]。另外,这些公开数据集中的图像一般都是经过挑选的小尺寸、裂缝特征明显的图像。在实际工程应用中,高清相机拍摄获取的裂缝图像中的裂缝尺寸远小于公开数据集或者试验条件下拍摄的裂缝。

由 Lin 等^[14]首次提出的金字塔型深度学习网络(FPN)表现出了强大的图像分割能力,并赢得了 COCO 2016 挑战。其网络结构突出体现了对小目标检测的优势。受 Lin 等工作的启发,笔者提出一个基于特征金字塔深度学习网络针对超大尺寸裂缝图像的细长裂缝检测方法。在金字塔型的网络架构下,采用特征提取效率更高的 Se_resnext50_32 \times 4d^[15]编码器,区别于传统的图像分割网络,将编码器提取的不同层次(共 4 个)特征图分别进行预测,不同分辨率的特征图得以保留。对于细小裂缝,更大分辨率的特征图能够提供更精确的裂缝几何特征,同时,较小分辨率的特征图提供了更深层的语义信息,能够优化细小目标的预测。不同尺寸裂缝特征图最后通过一包含 3×3 卷积层的 W 运算进行叠加输出统一尺寸的预测,避免产生混叠效应。

该方法包括使用超大尺寸图像的数据集构建方法。针对细长裂缝的检测特征,创新性利用特征金字塔深度学习网络具备的多层识别能力以及识别小目标的优势,通过充分利用裂缝长细几何特征,实现了对超大尺寸裂缝图像进行像素级的裂缝检测。

1 数据集构建方法

数据集是基于深度学习的裂缝自动检测方法的核心之一。为了充分获取裂缝细长的几何特征以及发挥深度学习网络的特征提取能力,提出针对超大尺寸细长裂缝图像的数据集构建方法。该数据集的构建方法包括两个关键部分:训练集与测试集。下文中将分别对训练集和测试集的处理、生成和构建方法做详细介绍。同时,来自 International Project Competition for SHM (IPC-SHM 2020) ANCRiSST^[13]

的钢箱梁数据集将按提出的数据集构建方法进行处理,以生成后续试验用的数据集。

该钢箱梁数据集中包含 120 张高清相机拍摄的钢箱梁表面细长裂缝图像,这些图像为三通道 RGB 图像,像素为 $3\,264 \times 4\,928$ 大小。除这些裂缝图像外,数据集内还包含每张裂缝图像相对应的标记图(label)。标记图由土木工程领域内专家对裂缝图像中的每一像素点进行标注分类生成,例如将裂缝像素标注为 1,背景像素标注为 0 进行像素级别的区分。

在进行处理前,钢箱梁数据集中的 120 张图像首先被随机分出 100 张作为初始训练集,其余 20 张为初始测试集,命名为 DS0。

将超大尺寸的图像直接放入网络中通常需要超大的计算机 GPU 运算空间。考虑到计算机性能的限制,较为简单直接的方法便是缩小图像尺寸,减小运算空间。传统是使用方法的双线性插值方法是在保证一定质量条件下对图像进行缩放。

为了与数据集构建方法相对比,试验中通过传统的双线性插值对图像进行缩放处理,构建了两个数据集。将原来的高分辨率图像降采样为 $1\,600 \times 2\,400$ 像素和 $2\,112 \times 3\,168$ 像素两种规格。为讨论方便,将图像大小调整为 $1\,600 \times 2\,400$ 像素的训练集命名为 DS1,将图像大小调整为 $2\,112 \times 3\,168$ 像素的训练集命名为 DS2,其具体参数见表 1。对于验证及测试过程,仍然以原规格图像作为验证和测试对象,但为了和训练过程相匹配,测试的裂缝图像均根据双线性插值方法调整为 $1\,600 \times 2\,400$ 像素和 $2\,112 \times 3\,168$ 像素大小。模型输出预测结果($1\,600 \times 2\,400$ 像素和 $2\,112 \times 3\,168$ 像素)后,再利用双线性插值恢复为 $3\,264 \times 4\,928$ 像素,与标注图像进行比较。

表 1 数据集

Table 1 Detail of evaluation datasets

数据集	训练/张	大小/像素	测试/张	大小/像素
DS0	100	$3\,264 \times 4\,928$	20	$3\,264 \times 4\,928$
DS1	100	$1\,600 \times 2\,400$	20	$3\,264 \times 4\,928$
DS2	100	$2\,112 \times 3\,168$	20	$3\,264 \times 4\,928$
DS3	1\,923	480×640	20	$3\,264 \times 4\,928$
DS4	3\,846	480×640	20	$3\,264 \times 4\,928$

由于细小裂缝占裂缝背景极少,双线性插值缩放之后,图像中裂缝像素急剧减小,预测难度增加,

预测准确度降低。考虑到部分裂缝即可提供足够信息进行识别,即可以拆分的几何特性与特征金字塔网络对小目标识别的优势,相应地提出了针对超大尺寸裂缝图像通用的训练集和测试集处理方法。

1.1 训练集

针对实际工程拍摄中裂缝在图像中分布不确定、长细比例较大、裂缝像素占整体像素比例较低的特点,采用多个步骤对超大尺寸图像进行处理,以构建训练集。

步骤 1:获得一张初始的超大尺寸裂缝图像(假设其大小为 $h \times w$,其中, h 为图像的高、 w 为图像的宽),选取一大大小为 $32m \times 32n$ (m, n 为大于 1 的整数)的固定拆分尺寸。该拆分尺寸的取值主要取决于计算资源。将初始图像以数值 0 进行边缘扩增(Padding),将图像统一为尺寸 $32m \times 32n$ 的倍数。扩增后的图像大小为 $H \times W$ 。 H, W 的计算式为

$$H = \left(\left\lceil \frac{h}{32m} \right\rceil + 1 \right) 32m \quad (1)$$

$$W = \left(\left\lceil \frac{w}{32n} \right\rceil + 1 \right) 32n \quad (2)$$

步骤 2:使用一个大小为 $32m \times 32n$ 的滑窗,以行、列步长为 $16m, 16n$ 像素大小对扩充后的图像进行拆分。

步骤 3:保留所有 k 张带裂缝的拆分图像,同时,为保证深度学习能够充分获得背景特征,学习背景信息; k 张只包含背景的图像也被随机选取,进入最终的训练集。最后,一张初始的超大尺寸裂缝图像经过上述操作共可生成 $2k$ 张的 $32m \times 32n$ 拆分子图进入深度学习训练集。对多张超大尺寸裂缝图像进行该构建操作便可构建出相应的训练集。

对于试验用钢箱梁数据集,考虑计算机的计算性能,选取 $m=15, n=20$ 。首先将尺寸为 $3\,264 \times 4\,928$ 的图像(如图 1(a))填充为 $3\,360 \times 5\,120$,并置零,保证后续拆分时能够覆盖到图像的边缘,见图 1(b),根据重叠分割的步骤 2,设定滑窗大小为 480×640 对图像以行/列 $240, 320$ 像素的步长,得到拆分子图,如图 1(c)所示。一些将子图像包含 0 填充部分。利用所有的子图像作为输入需要大量的训练时间。因此,根据对应的子标签,自动筛选出包含裂纹损伤像素的 1 923 个子图像,如图 2(a)所示,即 $k=1\,923$ 。这 1 923 张裂缝子图像组成了一个训练集,命名为 DS3,作为缺少背景图像信息的数据集,与本方法进行对比。其余的 13 939 个子图像不包

含属于裂纹损伤的像素,只表示背景信息,如图2(b)所示。在 13 939 个子图像中随机选取 1 923 张子图像。1 923 张裂缝子图像和 1 923 张背景子图像构成一个训练集,命名为 DS4,见表 1,即为本方法提出的数据集构建方法的最终结果。

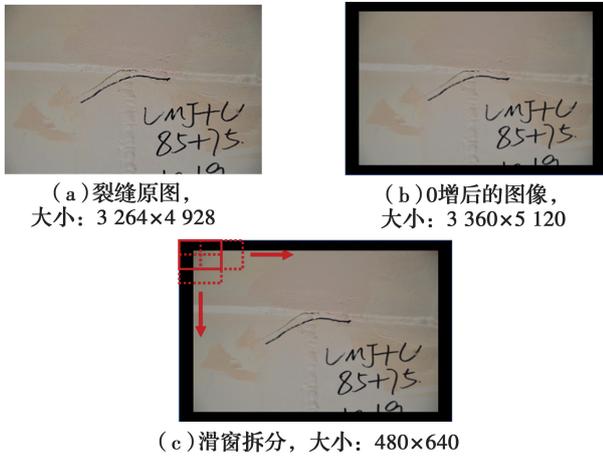


图 1 训练图像处理过程
Fig. 1 Treating process of train image

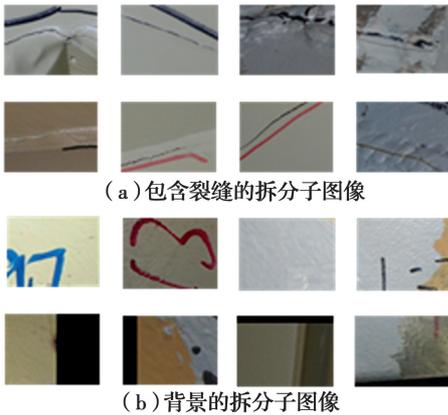


图 2 代表性的拆分子图像
Fig. 2 Representative sub-images

1.2 测试集

在图像测试方法上,同样针对细长裂缝的位置和几何特征提出了新的超大尺寸细长裂缝图像的测试处理方法。该方法包含 3 个步骤:

1)输入测试图像(假设为 $h \times w$),将图像以 0 进行扩增,不同于训练集,为了保证边缘像素能够被后续拆分覆盖,扩增后尺寸 $H \times W$ 的计算式为

$$H = \left(\left\lceil \frac{h}{32m} \right\rceil + 2 \right) 32m \quad (3)$$

$$W = \left(\left\lceil \frac{w}{32n} \right\rceil + 2 \right) 32n \quad (4)$$

式中: m 、 n 的取值同训练集中的取值。

2)使用 $32m \times 32n$ 的滑窗以行、列步长为 $32m$ 、 $32n$ 对测试图像进行第 1 次拆分,随后将分割起点定位于图内的第 $16m$ 行、第 $16n$ 列像素,进行第 2 次拆分。通过此方法,第 1 次拆分的 4 张相邻图像交点将成为第 2 次分割图像的中心。针对裂缝的位置不确定性对分割图像边缘进行多次覆盖。将步骤 2 生成的一系列拆分图一同输入经训练集训练后的特征金字塔深度学习网络,获得预测结果。

3)最后将该预测结果按照滑窗顺序重新拼装成为带 0 增的图像。去除 0 增区域后,便获得与原测试图像大小相同的裂缝预测图。同时,由于预测图中的每一个像素经过二次分割拥有两个预测值。本方法从较为安全的角度设定其中只要有一个预测判断该像素为裂缝,该像素即可认为裂缝像素,保证尽可能多的裂缝像素被检测识别。

在测试试验中, m 、 n 的取值同训练中的取值,分别为 15、20。其余处理按上述 3 个步骤进行,与训练集虽然稍有差别,但较为相似。

2 特征金字塔深度学习网络

特征金字塔网络 (Feature Pyramid Network) 的结构大致可以分为编码器 (Encoder)、解码器 (Decoder) 和组装 (Assembling) 3 个部分,如图 3 所示。编码器是卷积神经网络的前馈过程,用来提取不同阶段的特征,并不断降低特征图分辨率。解码器是自下而上的特征图放大过程。通过上采样,顶层特征图和底层特征图实现合并,丰富网络获取的语义信息。在融合完成后,网络使用了 3×3 的卷积核对每一个融合结果并进行卷积,目的是消除上采样中的混叠效应 (Aliasing Effect)。最后,在组装部分添加各个阶段的特征图获取各层信息,这使得特

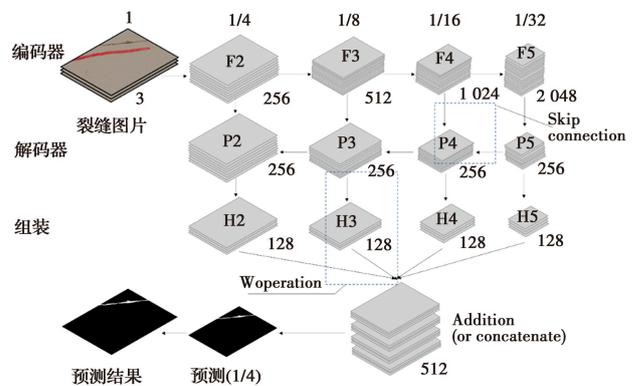


图 3 特征金字塔网络的结构
Fig. 3 Structure of feature pyramid network

征金字塔网络拥有较强的语义信息获取能力,同时也能够满足速度和内存的要求。特征金字塔网络的特别之处在于预测是在不同的特征层中独立进行的,有助于检测不同尺寸的裂缝目标,契合工程获取裂缝图像中裂缝大小不确定性的特点。

2.1 编码器

编码器是一个特征提取网络,一般采用 VGG^[16]和 ResNet^[17]作为骨干。在本方法中,为了提高对裂缝的特征提取效果,采用 Se_ResNeXt50_32×4d^[15]为编码器。它包括3个部分:ResNet,SE(Squeeze-and-Excitation)模块和X模块,使网络更加深入、更快收敛和更易优化。同时,该模型的参数少,复杂度低,适合裂缝检测任务。SE模块采用的是 SENet^[15]的一个计算单元。

压缩(Squeeze)采用了一个全局平均池化层,获取具有全局视野的特征图。激励(Excitation)利用了一个全连接的神经网络结构对压缩后的结果进行非线性变换,然后将其作为一个权重乘以输入特征。SENet 主要对通道之间的相关性进行学习,减弱了对通道本身的关注,虽然稍微增加了计算量,但能够实现更好的分割效果。X模块来自 ResNeXt^[18],一个更优版本的 ResNet。ResNeXt 的核心创新在于采用了聚合转换(Aggregated Transformations),用相同拓扑结构的并行堆叠取代原始的 ResNet 的3层卷积块,提高了模型的准确度,而不显著增加参数量。同时,由于拓扑结构相同,超参数减少,模型也较原始的 ResNet 更易移植。ResNeXt50_32×4d 从有50层网络深度的 ResNet50 改进而来。其中,32×4d 代表的是32条路径,每条路径的通道数为4个。SE模块嵌入到 ReNext50_32×4d 中,最终获得 Se_resnext50_32×4d。其预训练参数通过 imagenet1000 得到。

总的来说,编码器是一个以 Se_resnext50_32×4d 为骨干的自下而上裂缝特征提取网络。在编码器入口,将大小为裂缝图像输入(32m×32n)。后续中,每一个阶段便是特征金字塔的一个层次。选取 conv2、conv3、conv4 和 conv5 层提取的特征为{C2, C3, C4, C5},这是 FPN 网络的4个层次。N 是 batch size,特征向量则分别为 $F2=(N, 256, 8m, 8n)$, $F3=(N, 512, 4m, 4n)$, $F4=(N, 1024, 2m, 2n)$, $F5=(N, 2048, m, n)$ 。需要注意的是,由于 $F5$ 是原图的 1/32 大小,所以,拆分滑窗的长度和宽度应是 32 的倍数。

2.2 解码器

解码器是一个自上而下的放大裂缝特征图的过程。 $P5=(N, 256, m, n)$ 是通过 1×1 卷积层直接从 $F5$ 得到的。在接下来的操作中,通过插值算法将 $P5$ 放大两倍为 $(N, 256, 2m, 2n)$ 。 $F4$ 经过 1×1 卷积层,变成 $(N, 256, 2m, 2n)$ 。将上述两个特征向量相加,得到 $P4=(N, 256, 2m, 2n)$ 。这些操作被称为跳跃连接(Skip Connection),如图4所示,继续使用这些操作来获取 $P4$ 、 $P3$ 和 $P2$ 。跳跃连接的优势在于它既可以利用顶层的高级语义特征(有助于裂缝特征分类),又可以利用底层的高分辨率信息(有助于裂缝特征定位)。

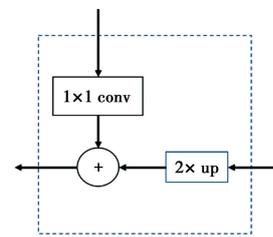


图4 跳跃连接

Fig. 4 Skip connection

2.3 组装

组装部分要求金字塔每一层输出的特征图具有相同尺寸大小,为了实现该目的,首先选择包含 3×3 卷积层的 W 运算,采用双线性插值放大,如图5所示。

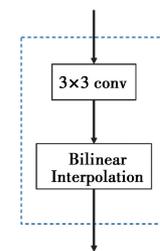


图5 W操作

Fig. 5 W operation

$P5$ 执行 3 次 W 操作后, $H5=(N, 128, 8m, 8n)$ 。以此类推, $H4$ 和 $H3$ 分别进行两次和一次 W 操作。此外, $H2$ 不需要放大。然后直接将 H_i ($i=2\sim 5$) 相加,得到一个向量 $(N, 256, 8m, 8n)$ 。该向量经过 3×3 的卷积层和双线性插值放大到原始裂缝图像尺寸 $(N, 1, 32m, 32n)$ 。为了便于预测过程,通过将激活函数 sigmoid 的值改为 $0\sim 1$ 得到掩码 $(N, 1, 480, 640)$ 。如果某点的值大于某一固定阈值,则预测该点为裂纹。在试验中,选择 0.5 作为阈值。

3 试验

3.1 试验细节

所有试验均在 Intel(R) Xeon(R) E5-2678 v3 @ 2.50 GHz, 64.0 GB RAM 和 NVIDIA RTX2080TI, 11.0 GB RAM 的计算平台上进行。基于 4 个不同的训练集训练了 4 个基于特征金字塔网络结构的预测模型。为了讨论方便,用 DS1 和 DS2 训练获得的特征金字塔网络称为 Model1 和 Model2,用 DS3 和 DS4 训练获得的网络称为 Model3 和 Model4。

3.2 试验指标

假设在标签和预测中均属于裂缝的像素数为 I ,在标签或预测中表示裂缝的像素数为 U ,交并比 (Intersection over Union IoU) 是图像分割任务常用的评价指标,通过式(5)计算。

$$\text{IoU} = \frac{I}{U} \quad (5)$$

另一评价指标 Dice Loss 不仅可以作为深度学习训练中的损失函数,也可以在一定程度上评价图像分割的准确度。假设在裂纹分割任务中 X 为图像对应的标记, Y 是裂缝预测,Dice Loss 通过式(6)计算。

$$\text{Dice Loss} = 1 - \frac{2 |X \cap Y|}{|X| + |Y|} \quad (6)$$

4 结果和讨论

图 6 为 4 个模型在一些典型测试图像上的裂缝检测结果,图 7 为基于 4 个不同训练集测试中的评价指标 IoU 和 Dice Loss 的对比。图 6 中展示的 IoU 值通过单张预测图片与对应标记计算得到,图 7 中展示的 IoU 和 Dice Loss 测试指标值为测试中所有图片的预测同对应标记计算获得的平均 IoU 和 Dice Loss 的值。从图 7 可以看出,在 IoU 和 Dice Loss 方面,在测试图像上,Model2 比 Model1 出性能更好,Model2 的 IoU 高于 Model1,同时 Model2 的 Dice Loss 小于 Model1。在分析了 Model1 和 Model2 的预测之后,发现一个值得注意的现象,如图 6 所示,Model1 和 Model2 的预测都忽略了测试图像中的部分裂纹,导致预测中的裂纹长度比标签中的短,在预测中丢失了部分裂纹。造成这种现象的原因之一是用于调整图像形状大小的双线性插值算法,双线性插值考虑已知像素值围绕未知像素计算位置的最接近的 2×2 邻域,然后取这 4 个像素的加权平均值来得到最终的插值值。调整大小的方法



图 6 不同模型的预测结果(预测 1 代表 Model1 对裂缝的预测,其余类似)

Fig. 6 Predictions of four models (Predict 1 represent the prediction of Model1)

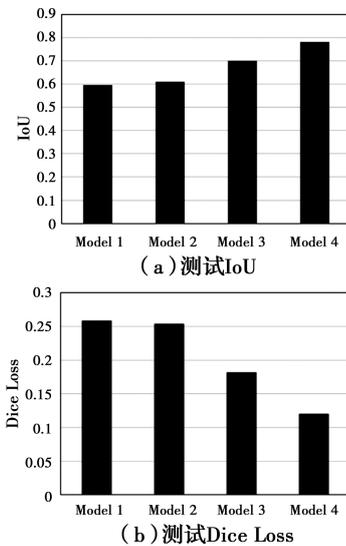


图 7 不同模型的测试指标值

Fig. 7 Evaluation index of four models

不是对测试图像中的每个像素进行预测,而是对像素的统计值进行预测。一方面,Model1 和 Model2 的预测不能在像素级上进行精确预测,但另一方面,模型仍然可以指出裂缝的位置。根据 Model1 和 Model2 预测的裂纹位置,可以采用一些方法改进裂纹预测,如利用扩展阈值算法可以扩展裂缝长度。从图 7(a)可以看出,Model1 的 IoU 仅比 Model2 的 IoU 小 0.02,但 DS1 中的图像大小几乎是 DS2 的

1/3。这表明,更大的图像尺寸并不能在很大程度上改善预测,而作为输入的大尺寸图像却需要大量的存储和计算空间。当需要缩放图像以获得一个新的训练集时,建议首先考虑计算效率,而不是图像大小。

使用该方法构建数据集时,测试中评价指标 IoU 和 Dice Loss 都有了明显提升。Model3 和 Model4 的 IoU 分别为 0.70 和 0.78。由于 Model3 只在裂纹子图像(DS3)上进行训练,因此忽略了一些背景信息。从图 6 第 3、第 4 行的预测结果可以看出,Model3 错误地将边缘上的一些像素点预测为裂纹,钢结构的一些凹槽特征也被预测为裂纹。尽管 Model3 的预测不如 Model4 准确,但 Model3 在结构健康监测中可以视为提供了较为安全的估计。从图 6 第 5 行可以看出,使用 Model3 对结构的主裂纹进行了预测,并将一些可疑特征视为裂纹,保证不遗漏裂纹损伤。Model3 的 IoU 虽然不是最好的,但在工程上具有理想的特性。经过 DS4 的训练,Model4 变得比 Model3 更加优秀和准确。Model4 同时学习了背景信息和裂纹信息。在图 6 中的预测 4, Model4 的性能最好,很好地处理了背景和沟槽的干扰。较高的精度意味着 Model4 可以为下一步的研究或测量提供精确的裂纹信息。在裂纹分割任务中,Model3 和 Model4 的性能优于 Model1 和 Model2。对于基本训练集的图像,该方法被证明是比直接调整超大尺寸图像大小方法更好的解决方案。部分裂缝便可以提供足够的检测信息。因此,该方法是一种合适的针对超大尺寸细长裂缝的处理方法。此外,该方法对图像的分辨率没有任何要求,适用于任意的高分辨率和低分辨率图像。

5 结论

针对钢箱梁内部包含复杂背景和裂缝的超大尺寸图像,提出了一种基于特征金字塔深度学习网络的裂缝自动检测方法,并进行了一系列对比试验。主要结论如下:

1)基于本方法构建的训练集,特征金字塔网络能够在测试中对 $3\ 264 \times 4\ 928$ 像素的桥钢箱梁表面裂缝图像实现最大交并比(IoU)为 0.78,最小 Dice Loss 为 0.12,表现优于将图像进行简单缩放构建数据集训练的模型。

2)采用双线性插值方法将超大尺寸图像($3\ 264 \times$

$4\ 928$)缩放至 $1\ 600 \times 2\ 400$ 像素或 $2\ 112 \times 3\ 168$ 像素,将会导致部分裂缝信息丢失,降低裂缝检测准确度。建议采用本方法对超大尺寸裂缝图像进行拆分,拆分尺寸大小可设定为 480×640 。

3)部分裂缝可以为深度学习网络提供足够的特征信息。该方法适合于裂缝检测任务,并且该方法对于裂缝图像的具体分辨率没有严格的要求,为今后的裂缝自动检测提供了一个较有潜力的解决方案。

参考文献:

- [1] 李艳霞. 基于深度学习的桥梁裂缝图像信息自动提取方法研究[D]. 北京: 北京交通大学, 2020.
LI Y X. Research on automatic bridge crack extraction based on image deep learning [D]. Beijing: Beijing Jiaotong University, 2020. (in Chinese)
- [2] 彭玲丽, 黄少旭, 张申申, 等. 浅谈无人机在桥梁检测中的应用与发展[J]. 交通科技, 2015(6): 42-44.
PENG L L, HUANG S X, ZHANG S S, et al. Discussion on UAV application and development in bridge inspection [J]. Transportation Science & Technology, 2015(6): 42-44. (in Chinese)
- [3] 陈艳君. 基于特征空间的路面裂缝检测与识别算法研究[D]. 武汉: 武汉工程大学, 2012.
CHEN Y J. Research on detection and recognition method for pavement crack based on feature space [D]. Wuhan: Wuhan Institute of Technology, 2012. (in Chinese)
- [4] 孙朝云, 褚燕利, 樊瑶, 等. 基于 VC++ 路面裂缝图像处理系统研究[J]. 计算机应用与软件, 2009, 26(8): 82-85.
SUN Z Y, CHU Y L, FAN Y, et al. Pavement crack image processing system research based on VC++ [J]. Computer Applications and Software, 2009, 26(8): 82-85. (in Chinese)
- [5] 马常霞, 赵春霞, 胡勇, 等. 结合 NSCT 和图像形态学的路面裂缝检测[J]. 计算机辅助设计与图形学学报, 2009, 21(12): 1761-1767.
MA C X, ZHAO C X, HU Y, et al. Pavement cracks detection based on NSCT and morphology [J]. Journal of Computer-Aided Design & Computer Graphics, 2009, 21(12): 1761-1767. (in Chinese)
- [6] 张维峰, 刘萌, 杨明慧. 基于数字图像处理的桥梁裂缝检测技术[J]. 现代交通技术, 2008, 5(5): 34-36.

- ZHANG W F, LIU M, YANG M H. Bridge crack detection technology based on digital image processing [J]. *Modern Transportation Technology*, 2008, 5(5): 34-36. (in Chinese)
- [7] 勾红叶, 杨彪, 华辉, 等. 桥梁信息化及智能桥梁 2019 年度研究进展[J]. *土木与环境工程学报(中英文)*, 2020, 42(5): 14-27.
- GOU H Y, YANG B, HUA H, et al. State-of-the-art review of bridge informatization and intelligent bridge in 2019 [J]. *Journal of Civil and Environmental Engineering*, 2020, 42(5): 14-27. (in Chinese)
- [8] ZHANG L, YANG F, ZHANG Y D, et al. Road crack detection using deep convolutional neural network [C]//2016 IEEE International Conference on Image Processing (ICIP). September 25-28, 2016, Phoenix, AZ, USA. IEEE, 2016: 3708-3712.
- [9] 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.
- [10] DORAFSHAN S, THOMAS R J, MAGUIRE M. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete [J]. *Construction and Building Materials*, 2018, 186: 1031-1045.
- [11] AZIMI M, ESLAMLOU A D, PEKCAN G. Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review [J]. *Sensors (Basel, Switzerland)*, 2020, 20(10): 2778.
- [12] CHENG H K, CHUNG J, TAI Y W, et al. CascadePSP: Toward class-agnostic and very high-resolution segmentation via global and local refinement [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13-19, 2020, Seattle, WA, USA. IEEE, 2020: 8887-8896.
- [13] BAO Y Q, LI H. Machine learning paradigm for structural health monitoring [J]. *Structural Health Monitoring*, 2020: 147592172097241.
- [14] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 936-944.
- [15] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. 2014. <https://arxiv.org/abs/1409.1556>.
- [17] 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 (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [18] XIE S N, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 5987-5995.

(编辑 胡玲)