

# 基于 U-net 的道路缺陷检测

彭磊 张辉

长沙理工大学电气与信息工程学院 长沙 410000

(1249028596@qq.com)

**摘要** 道路是现代交通运输最主要的途径之一,道路缺陷对于道路安全有着巨大威胁。因此准确检测道路缺陷对道路养护修缮具有重要意义。道路缺陷具有低连续性和低对比度的特点,现阶段多采用人工检测方法,检测效率低,人力成本高,且检测人员的安全可能会遭受威胁。随着深度学习的发展,神经网络方法被广泛应用于工程实践。U-net 是具有编码器-解码器结构的端到端深度学习模型,对微小对象检测能力强,适用于道路裂缝缺陷检测。利用 U-net 深度学习网络对道路缺陷进行检测,能提高检测效率,无需人工干预,保证检测人员安全,降低检测的人工成本。实验结果表明,U-net 网络在数据集 Crack500 上的效果优于 FCN,Segnet 等语义分割网络,在保持较高精度的情况下实现了道路缺陷检测。在此基础上对 U-net 网络层数进行超参数优化,确定该数据集上的最优 U-net 网络结构。

**关键词:** 缺陷检测;U-net;卷积神经网络;深度学习

**中图分类号** TP391

## U-net for Pavement Crack Detection

PENG Lei and ZHANG Hui

School of Electrical & Information Engineering, Changsha University of Science & Technology, Changsha 410000, China

**Abstract** Road is one of the most crucial ways for transportation. Crack on road will cause great danger to transportation if you leave it unchecked, so it is important to detect crack precisely in road maintenance. Road cracks are usually discontinuous and low-contrast which is difficult to detect using traditional methods of image processing. In this paper, we utilize U-net for road crack detection which is an end-to-end with encoder-decoder structure efficient deep learning network on dataset Crack500, while traditional methods are time-consuming and labor-consuming. U-net is appropriate for road crack detection because of its ability to catch fine details in image. Experiment results demonstrate that U-net outperforms other detect methods. Furthermore, we discuss the difference when modifying the number of conv-blocks in U-net. Experiment results show that it achieves best performance when the number of conv-blocks set to be 7.

**Keywords** Defect detection, U-net, Convolutional neural network, Deep learning

### 1 引言

裂缝是经常出现在各种物体表面的一种缺陷,例如铸件表面裂缝<sup>[1]</sup>、桥面裂缝<sup>[2]</sup>、部件缺陷<sup>[3]</sup>。裂缝的出现对物体的质量和安全性会造成巨大的影响。如果忽视裂缝的存在,较小的裂缝会发展成较大的裂缝。例如在高速公路上的裂缝,如果发现的不及时,在雨天气候时会加剧裂缝的扩张,对交通安全的影响是巨大的。所以,及时发现并对裂缝进行修补是至关重要的。全世界有超过七千万公里的高速公路,如果采用人工检测的方法,检测工作人员的安全难以得到保证且检测难度较大,费时费力,成本很高。而利用自动检测方法能在降低成本的同时保证检测效率。同时,早期微小裂缝的修补会较为简单且成本低,具备检测微小裂缝的能力是对检测方法的基本要求之一。

道路裂缝一般呈线状,如果道路缺陷具有较好连续性且具有高对比度,利用传统图像处理方法如边缘检测等能达到好的检测效果。但裂缝的出现一般伴随着间断性和背景噪声,裂缝的连续性会被道路的纹理所影响。在此基础上,道路

环境还具有不确定性。道路环境呈开放式,路边的物体在阳光的照射下会产生阴影,对裂缝存在遮挡的可能性。这种情况下,裂缝与背景对比度降低,缺陷检测的难度进一步提高。因此,具备较好的抗噪声能力也是检测方法的基本要求之一。随着深度学习领域的飞速发展,基于深度学习的裂缝检测方法越来越多。深度学习方法无需人工提取特征,由卷积操作提取特征,简化了操作流程,同时具有更好的泛化性能。同时深度学习方法针对不同的数据集无需重新设计特征,在对神经网络进行微调的基础上即可满足其他任务,并且其具备较好的抗噪声能力,在各领域得到广泛应用。

针对以上问题,本文利用 U-net<sup>[4]</sup> 网络对道路缺陷进行检测。U-net 为语义分割深度学习网络,对图像中每个像素进行分类。其最初用于医学图像分割,在医学领域取得了很好的应用效果,常用于细胞检测、视网膜血管检测等。U-net 对微小物体有很好的检测效果,同时具有较强的抗干扰能力。在实验中发现,卷积模块个数的不同会造成检测精度的不同。因此,本文对卷积层模块个数进行分析实验,整定超参数卷积模块个数,最终利用 U-net 实现对道路裂缝的检测。

## 2 研究现状

语义分割是图像处理的三大任务之一。分割任务的本质是将图像进行像素级的分类。如今,利用深度学习方法进行分割在各领域都得到广泛应用。比如在医学上可以用于细胞检测,大大减少了医护工作者的工作量;在自动驾驶领域,利用神经网络对出现在汽车视野中的物体进行分割,提高自动驾驶车辆对周围环境的感知能力;对于交通领域来说,国内外很多学者都在逐步将深度学习方法应用到车道线识别、道路缺陷检测等。在语义分割方面,Jonathan 等<sup>[5]</sup>提出利用全卷积神经网络 FCN 对图像进行分割。该方法去除了传统卷积神经网络的全连接层,并用卷积层代替。得益于此,该网络能接收任意尺寸的输入图像。在卷积操作完成之后,利用反卷积将特征恢复到输入图像大小,并以 softmax 作为激活函数,得到图像像素分类概率图,进而得到图像的分割结果。FCN 网络是将深度学习应用于语义分割的开端。在道路缺陷检测方面,如 Cao 等<sup>[6]</sup>提出利用全卷积神经网络作为编码器和解码器对道路缺陷进行检测。该方法采用舍弃全连接层的 VGG16 网络结构对输入图像进行编码,得到高层级的语义特征,将得到的高层特征利用反卷积操作和上采样逐步恢复图像至输入图像分辨率,进而得到图像分割结果。其中,在编码器和解码器之间利用索引池化,即保留编码器中池化操作的最大值索引信息,在上采样过程中利用保留的索引信息,能更好地保留特征的空间信息,提高分割精度。随着深度学习的发展应用,越来越多的高速高精度深度学习网络被应用于道路缺陷检测。

深度学习需要大量的数据量来支撑算法,因此对于数据的数量要求较高。本文采用的数据集为 Crack500<sup>[7]</sup>。该数据集提供 516 张包含道路缺陷的图像,图像大小为  $3\ 264 \times 2\ 448$ ,拍摄地点为美国天普大学(Temple University),经由智能手机采集,每张图像对应已有标注的二值图像标签。该数据集的作者已经将图像分为 3 个部分,即训练集、验证集、测试集,并将过大的原始图像裁剪为  $640 \times 360$  大小的图像。3 个部分分别包括 1 896 张、348 张、1 123 张图像。我们利用训练集对 U-net 进行训练,在训练过程中用验证集评估网络训练情况,并在训练完成后,用测试集来检验网络的泛化能力。图 1 给出了为训练集部分图片及对应标签,验证集部分图片和测试集部分图片。

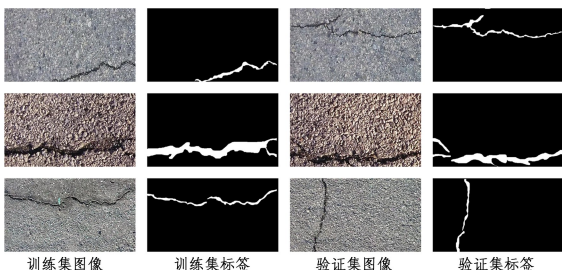


图 1 数据集部分图片

Fig. 1 Samples from dataset

## 3 U-net 网络结构的原理

不同于其他深度学习分割网络,U-net 最大的特点是采用跳跃连接结构(skip-connection),将编码器的到的特征图以

叠加(Concatenate)的方式,与解码器对应的网络层相叠加,在叠加后再进行反卷积。跳跃连接将编码器中获得的特征信息引入到对应解码器中,为上采样或反卷积提供了更多低层次的空间与信息。而这些低层信息通常为轮廓信息、位置信息等,为后期图像分割提供多尺度多层次的信息,由此可以得到更精细的分割效果。U-net 网络结构如图 2 所示。

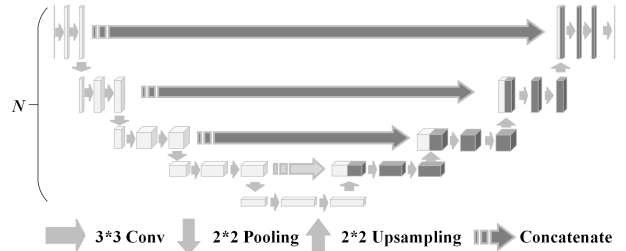


图 2 U-net 网络结构图

Fig. 2 Structure of U-net

U-net 为编码器-解码器结构。在编码器部分,采用卷积模块进行特征提取。图 2 中  $N$  表示卷积模块个数,并设定基准 U-net 卷积模块个数为 5。每个卷积模块包括两个卷积层,在每个卷积层后跟随 BN 层。BN 层的作用是防止神经网络在训练过程中出现梯度弥散或梯度爆炸,加快收敛速度,减少训练时间。每个卷积模块的卷积层中卷积核个数分别为 32,64,128,256,512。每经过一个卷积模块,有池化层对得到的特征进行降维,在维持通道数不变的情况下,将尺寸缩小为上一个卷积模块的一半。在经过 5 个卷积模块后,得到输入图像的高层级特征向量,将特征向量传入解码器部分。编码器部分的激活函数采用 ReLU 函数。

在解码器部分,采用上采样和卷积组合的方法实现反卷积,增大特征图像分辨率。在解码器部分采用 4 个解码模块对提取特征进行解码。解码模块由反卷积和卷积模块组成。每个反卷积跟随一个卷积模块,解码部分的卷积模块与编码器中相同层级的卷积模块一致。解码模块的输入不止有上一层级的输出,还有来自相对应编码器得到的特征图。在通道维度上相连接,将编码器得到的低等级特征进行融合(Concatenate),保证对低等级特征的充分利用,提高网络对微小目标的分割效果。在编码器中仍采用 ReLU 作为激活函数。

在将特征还原为输入图像大小后,将与输入图像大小相同的特征图送入 softmax 层,得到每个像素属于各个类别的概率,确定输入各个类别的概率阈值,若高于该阈值则认为该像素属于指定类别。最终,得到图像分割结果。

## 4 实验结果与分析

### 4.1 实验环境

本文研究的硬件环境为 i9-9900k intel 处理器,内存为 32GB,GPU 为 GeForce GTX 2080 显卡(8GB 显存);操作系统为 Windows10,软件环境为 Keras 2.2.4,Python 3.6。

### 4.2 实验结果

在得到训练完成的模型后,在有 1 123 张图片的测试集上进行测试,得到最终测试结果。

部分实验结果如图 3 所示,其中由上至下算法模型依次为 FCN,Segnet<sup>[8]</sup>,Enet<sup>[9]</sup>,U-net 以及 Ground Truth。由图 3 可知,在经过网络训练后,U-net 具备较强的缺陷检测能力,并且检测性能优于其他网络模型,尤其是对细微缺

陷的检测效果更好。

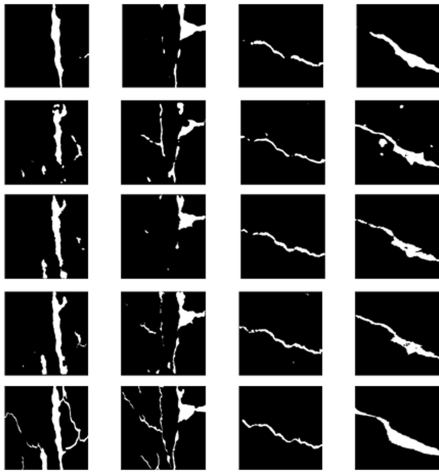


图3 部分实验结果图

Fig. 3 Partial results of detection

#### 4.3 实验测试指标及分析

语义分割任务通常采用准确率(Precision)、召回率(Recall)以及 F1-score 来对模型效果进行评价。各指标的表达式如下:

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN} \quad (1)$$

$$F_1 = \frac{2RP}{R + P}$$

其中,  $TP$  为正确预测的正样本数,  $FP$  为错误预测的正样本数,  $FN$  为错误预测的负样本数。F1-score 用来衡量准确率和召回率, 为二者的调和均值。利用以上指标对模型进行评估, 结果如表 1 所列。

表 1 不同分割算法的效果对比

Table 1 Performance comparison of different methods

Method	Precision	Recall	F1-score
FCN	0.55	0.44	0.52
Segnet	0.53	0.51	0.53
Enet	<b>0.59</b>	0.45	0.55
U-net	0.55	<b>0.55</b>	<b>0.55</b>

在使用相同数据集训练和测试后, 结果表明 U-net 在召回率和 F1-score 上的得分高于其他模型, 而 Enet 在准确率上的得分是最高的。得益于 U-net 的跳跃连接(skip-connection)结构, 在上采样阶段重复利用低层级提取的特征, 因此细节分割能力更为突出。从对结果图的观察来看, Enet 的准确率较高但检测效果并不好, 原因可能是其对图像像素进行二分类时正负样本不平衡所致。

#### 5 网络参数整定

卷积层个数对于神经网络性能的影响一直是深度学习领域的研究难题。不同的卷积模块个数对于特征提取的程度不同, 会影响图像最后的分割结果。故对网络的神经网络的卷积模块个数进行整定, 比较不同卷积模块数下 U-net 对缺陷的检测能力。由于输入图像尺寸的限制, 实验设置的卷积模块数  $N$  范围限制为 2~7。实验结果如图 4 所示。

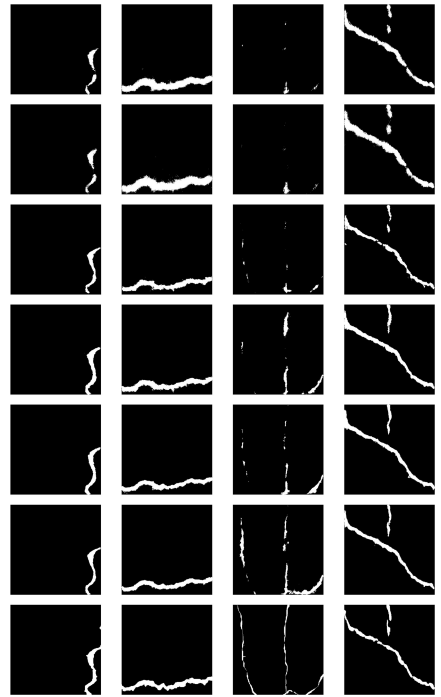


图 4 不同卷积模块数下 U-net 的结果

Fig. 4 Results of U-net under different number of convolution modules

图 4 中结果从上至下所用卷积模块数  $N=2, 3, 4, 5, 6, 7$ , 最后一行为 Ground truth。由图 4 可知当  $N=4$  时, 模型已经具备较强的对简单缺陷进行检测的能力, 但对困难样本的检测仍不满足任务要求。随着网络层数的加深, 对困难样本的处理能力也相应增强, 同时保留对简单样本的检测能力。当  $N=7$  时, U-net 的检测能力最强。其对应的评价指标对比如表 2 所列。

表 2 不同卷积模块数下 U-net 算法的效果对比

Table 2 Performance comparison of U-net under different number of convolution

$N$	Precision	Recall	F1-score
$N=2$	0.56	0.48	0.54
$N=3$	0.54	0.52	0.54
$N=4$	0.55	0.55	0.55
$N=5$	0.57	<b>0.59</b>	0.58
$N=6$	<b>0.62</b>	0.53	0.60
$N=7$	<b>0.62</b>	<b>0.59</b>	<b>0.61</b>

由表 2 可知, 随着网络层数的加深, U-net 对于缺陷的检测能力也逐步提高。随着卷积模块的增加, 网络的特征提取能力越来越强, 对于缺陷的语义特征获取更为准确, 同时在上采样时利用了编码器的低层级特征, 因此在加深网络的同时并不会导致浅层特征的丢失。  $N=7$  时模型的各项指标都高于或等于其他网络模型, 在各测试模型中性能最好。

**结束语** 本文针对道路裂缝缺陷数据集利用 U-net 神经网络对图像中缺陷进行语义分割。该网络在经过训练后具备对道路缺陷进行检测的能力, 得到了稳定的模型参数, 并且通过和其他深度学习方法进行比较, 验证了 U-net 网络在缺陷检测任务中优于其他深度学习方法, 在检测任务中具有较高准确率, 对道路缺陷的准确检测, 能够保证交通运输的安全进行, 降低交通运输风险, 减少财产损失。但是目前的工作仍然有很多需要改进之处, 细微缺陷的检测与全局缺陷的检测存

在矛盾,加深网络的同时检测微小缺陷的能力也减弱。后续工作考虑改进网络结构,在保证全局缺陷检测精度的情况下,提高对细微缺陷的检测能力。此外,数据集中存在路面情况较为恶劣的情况,但道路并未出现缺陷,易让神经网络产生错误识别,将非缺陷区域识别为缺陷。因此后续工作还将提高网络对图像噪声的处理能力,降低错误识别率,并加强对检测结果的评价。

### 参 考 文 献

- [1] FERGUSON M, RONAY A, LEE Y T, et al. Detection and Segmentation of Manufacturing Defects with Convolutional Neural Networks and Transfer Learning[C]// Smart Sustain Manuf Syst. 2018;2-10.
- [2] PRATEEK P, KRISTIN J D, NENAD G, et al. Automated Crack Detection on Concrete Bridges[J]. IEEE Transactions on Automation Science and Engineering, 2016, 13(2): 591-599.
- [3] CHEN F, 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.
- [4] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C]// Medical Image Computing and Computer-assisted Intervention, 2015: 234-241.
- [5] JONATHAN L, EVAN S, TREVOR D. Fully Convolutional Networks for Semantic Segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015;3431-3440.
- [6] CAO V D, LE DUC A. Autonomous concrete crack detection using deep fully convolutional[J]. Automation in Construction, 2019, 99: 52-58.
- [7] ZHANG L, YANG F, ZHANG D Y M. Road crack detection using deep convolutional neural network[C]// IEEE International Conference on Image Processing. 2016;3708-3712.
- [8] 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.
- [9] ADAM P, ABHISHEK C, SANGPIL K, et al. Enet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation[J]. arXiv:1606.02147.
- [10] QIN Z, ZHENG Z, LI Q Q, et al. DeepCrack: Learning Hierarchical Convolutional Features for Crack Detection [J]. IEEE Transactions on Image Processing, 2019, 28(3): 1498-1512.
- [11] LIU Y H, YAO J, LU X H, et al. DeepCrack: A deep hierarchical feature learning architecture for crack segmentation[J]. Neurocomputing, 2019, 338(21): 139-153.
- [12] ZHANG K G, ZHANG Y T, CHENG H D. CrackGAN: Pavement Crack Detection Using Partially Accurate Ground Truths Based on Generative Adversarial Learning [J]. arXiv: 1909.08216v2.
- [13] WU H, ZHANG J, HUANG K, et al. FastFCN: Rethinking Dilated Convolution in the Backbone for Semantic Segmentation [J]. arXiv:1903.11816v1.
- [14] LIU W J, HUANG Y C, LI Y, et al. FPCNet: Fast Pavement Crack Detection Network Based on Encoder-Decoder Architecture[J]. arXiv:190248v1.
- [15] CHEN T Y, CAI Z H, ZHAO X, et al. Pavement crack detection and recognition using the architecture of segNet[J]. Journal of Industrial Information Integration. 2020, 18.
- [16] ZHAO H S, SHI J P, QI X J, et al. Pyramid Scene Parsing Network[J]. arXiv:1612.01105v2.
- [17] RON L, CAGKAN Y, GITTA K, et al. RadioUNet: Fast Radio Map Estimation with Convolutional Neural Networks [J]. arXiv:1911.09002.
- [18] ZHU H G, MIAO Y, ZHANG X, et al. Semantic Image Segmentation with Improved Position Attention and Feature Fusion[J]. Neural Processing Letters, 2020, 52: 329-351.
- [20] ZHOU Z W, MD M R S, NIMA T, et al. UNet++: A Nested U-Net Architecture for Medical Image Segmentation[J]. arXiv: 1807.10165v1.
- [21] DOMINGO M. Aluminum Casting Inspection Using Deep Learning: A Method Based on Convolutional Neural Networks [J]. Journal of Nondestructive Evaluation, 2020, 39: 12.
- [22] LIN J H, YAO Y, MA L, et al. Detection of a casting defect tracked by deep convolution neural network[J]. The International Journal of Advanced Manufacturing Technology, 2018, 97: 573-581.
- [23] YU F, VLADLEN K. Multi-Scale Context Aggregation by Dilated Convolutions[J]. arXiv:1511.07122.
- [24] HAO M, LU C F, WANG G Q, et al. An improved Neural Segmentation Model for Crack Detection-Image Segmentation Model[J]. Bulgarian Academy Sciences, 2017, 17(2): 119-133.
- [25] GANG S, LI S, SUN G, et al. Squeeze-and-Excitation Networks [J]. arXiv:1709.01507v4.
- [26] XU H Y, SU X, WANG Y, et al. Automatic Bridge Crack Detection Using a Convolutional Neural Network [J]. Applied Science, 2019, 9(14), 2867.
- [27] YANG F, ZHANG L, YU S J, et al. Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection[J]. IEEE Transaction on Intelligent Transportation Systems, 2020, 21(4): 1525-1535.



**PENG Lei**, born in 1996, postgraduate. His main research interests include image processing and deep learning.



**ZHANG Hui**, Ph.D, assistant professor, visiting scholar. His main research interests include machine vision, sparse representation, visual tracking.