E-mail: csa@iscas.ac.cn http://www.c-s-a.org.cn Tel: +86-10-62661041

基于连续注意力机制和卷积金字塔的路面裂缝 检测^①

陈良全, 王彩玲, 刘华军, 蒋国平

(南京邮电大学 自动化学院、人工智能学院, 南京 210023) 通讯作者: 陈良全, E-mail: 1099145164@qq.com

摘 要: 近年来,随着全球汽车保有量的增加和路面的不断扩建,路面裂缝检测受到了广泛的关注.虽然许多裂缝检测器模型已经被提出,但也存在一些问题,例如:一些宽度较细的裂缝可能未被检测而出现裂缝断裂的现象;边缘信息可能会在过滤或池化过程中丢失.本文以 SegNet 为基础框架,编码层设计了一种连续注意力机制,并且在特征图通过解码层之前添加了卷积金字塔结构,以减少裂缝检测中的断裂,获得更完整的边缘信息.与相关方法相比, Precision、Recall和F1-measure 三个指标分别提高了 2.47%、8.21%和 6.87%,对 Crack200、Crack500和 CrackForest 三个开源数据集检测结果的平均交并指标 (MIoU)提高了 14.35%.

关键词:路面裂缝;注意力机制;金字塔结构;编码器-解码器

引用格式: 陈良全,王彩玲,刘华军,蒋国平.基于连续注意力机制和卷积金字塔的路面裂缝检测.计算机系统应用,2021,30(8):249-255. http://www.c-s-a.org.cn/1003-3254/8044.html

Pavement Crack Detection with Continuous Attention Mechanism and Convolution Pyramid Structure

CHEN Liang-Quan, WANG Cai-Ling, LIU Hua-Jun, JIANG Guo-Ping

(College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: With the increase in global vehicles and the expansion of road surfaces, pavement crack detection has received extensive attention in recent years. Many detector models have been proposed, with some problems, though. For example, some narrow cracks may not be detected, leading to discontinuous cracks; the detailed crack edge information may be lost during filtering or pooling. On the basis of SegNet, a continuous attention mechanism is designed in the encoder layer, and a convolutional pyramid structure is added before the feature map passes through the decoder layer to reduce the fracture in crack detection and obtain more complete edge information. The *Precision, Recall*, and *F1-measure* of our approach are 2.47%, 8.21%, and 6.87% higher than those of the related method, respectively, and the Mean Intersection over Union (*MIoU*) of the detection results on the three open datasets, namely, Crack200, Crack500, and CrackForest is improved by 14.35%.

Key words: pavement crack; attention mechanism; pyramid structure; encoder-decoder

路面裂缝是由自然应力引起的路面损坏,如果不及 时处理,很可能导致道路塌陷和交通事故.因此,对路 面进行准确高效的检测尤为重要.最直接的检测方法 是对路面进行目测,并由专家对其进行准确的评估.然

Foundation item: Natural Science Foundation of Nanjing University of Posts and Telecommunications (NY220057) 收稿时间: 2020-11-24; 修改时间: 2020-12-22; 采用时间: 2021-01-07; csa 在线出版时间: 2021-07-31

Software Technique•Algorithm 软件技术•算法 249

① 基金项目:南京邮电大学自然科学基金 (NY220057)

而,这种方法涉及到较高的人工成本,并且可能产生不可靠和不一致的结果,同时检测人员直接在路面上检测是极其危险的且效率也很低^[1].因此有必要设计一种精确、自动化的路面裂缝检测系统.然而路面裂缝的自动检测是一个充满挑战的课题,因为裂纹不仅拓扑结构复杂,尺度不均匀,而且会受低对比度和杂乱背景的影响而降低检测精度^[2].

路面裂缝通常具有以下特点:

(1)裂缝的颜色相对于背景较深;

(2) 裂缝的延伸方向是线性且宽度可变的;

(3) 裂缝边缘的拓扑结构复杂.

传统检测方法中,基于最大后验分类器,Oliceria 等^[3] 提出一种利用激光成像系统对图像中的裂缝进行检测, 再利用二值区域和几何特征进行分割.Zou 等^[4]提出 采用地面阴影去除法,通过张量投票法建立裂缝分布 概率图,然后构造随机森林对递归树的边缘进行剪枝 以进行裂缝检测.Oliveria 等^[5]推出一种用于路面裂缝 检测与表征的图像处理算法和技术评价算法工具箱. Peng 等^[6]首次采用 Otsu 阈值分割法去除路面图像中 的道路标志,然后采用自适应迭代阈值分割算法得到 裂缝分割图像.

深度学习方法迅速发展并且已经应用到了路面 裂缝检测任务中.基于深度学习的路面裂缝检测方 法可以分为五类:分割算法,边缘检测算法,分类器 算法,生成对抗方法和无监督方法.文献 [7,8] 将裂缝 检测视为分割任务, 文献 [7] 提出一种分块阈值分割 方法对裂缝进行分割. 文献 [8] 在 SegNet 的基础上通 过融合同一尺度下编码器和解码器的卷积特性来进 行裂缝特征的提取. Yang 等^[9] 将裂缝检测视为边缘 检测任务,以特征金字塔的形式将上下文信息转化 为低层特征进行裂纹检测. Dorafshan 等^[10] 基于混凝 土图像的裂缝检测比较了深度卷积神经网络和边缘 检测器. Zhang 等^[11] 采用基于裂缝修补 (CPO) 的生成 对抗网络模型对路面进行检测并将大尺度裂缝图像 输入到非对称 U 形网络中, 以克服裂缝的产生"全黑" 问题. Li 等^[12] 提出了一种基于无监督多尺度融合裂 纹检测的 (MFCD) 算法. 此外, 文献 [13,14] 将分类算 法应用到路面裂纹检测中. 文献 [13] 提出了深度卷 积神经网络 (DCNN) 实现了自动检测和分类路面裂 纹. 文献 [14] 提出基于概率生成模型 (PGM) 和支持 向量机 (SVM) 的网络来生成和计算裂纹特征的概率

250 软件技术•算法 Software Technique•Algorithm

图,并且提出了一种基于融合概率图的加权展开运 算,增强了算法的有效性.

在一定程度上,无论是传统方法还是深度学习方 法都可以检测到裂缝.然而,保证裂缝检测的连续性和 裂缝边缘的完整性仍需进一步研究.一方面,裂缝沿着 其延展方向是连续分布的,一些比较细的裂缝容易检 测不到.另一方面,裂缝的边缘拓扑结构复杂,检测器 容易丢失边缘信息,导致检测裂缝的边缘过于光滑.如 图 1 所示,图 1(a)中检测结果出现裂缝断裂的现象, 图 1(b)中的检测结果丢失了裂缝的边缘信息.



1 相关方法

1.1 SegNet 的发展

我们将路面裂缝检测视为分割任务,且网络结构 是受 SegNet 的启发. SegNet 是一种编码解码模型,由 Badrinarayanan 等^[15] 首次提出,用于像素级图像分割. 之后,Badrinarayanan 等^[16] 改进了他们之前提出的 SegNet 模型.编码器部分和 VGG16 的 13 层卷积层相 同,且编码层由上采样和卷积构成,最后将每个像素通 过 Softmax 激活函数进行分类.

Schmugge 等^[17] 首先将 SegNet 作为基础框架应用 于裂缝检测任务. 他们提出了一种自动检测方法, 以帮 助对核电站部件的焊接、划痕和研磨进行远程自动检 测,给出了不同光照条件下多帧图像的层次分类置信度. 1.2 基础框架

虽然我们设计的框架是基于 SegNet 的, 但是, 它与原生 SegNet^[16]不一样, 我们没有直接在编码器 部分使用 VGG16 模型. 编码器部分分为 5 层, 前 2 层 是两个卷积, 并且每个卷积层后有 1 个 max-pooling 层, 最后 3 层是 3 个卷积层, 每个卷积层后有 1 个 maxpooling 层.

2 本文提出的模型

CACPNet 的整体设计结构如图 2 所示, 左边虚线 框出的部分为编码器, 右边为解码器, 编码器上半部分 结构为多层卷积金字塔结构, 下半部分为连续注意力 机制.

2.1 连续注意力机制

路面裂缝通常是连续的,因此裂缝之间通常不会 出现断开的现象.另外,在检测过程中经常会混淆背景 和裂缝,这就导致了断裂和裂纹信息丢失问题.受注意 力门控^[18]的启发,我们设计了一个连续特征注意机制 模块,以保证路面裂缝的连续性.

连续特征注意机制结构如图 3 所示. 输入图像或

特征经过的上一层为 Encoder 模块,下一层为连续注 意力机制模块.连续注意力机制模块中,输入的原始图 片或者上层特征图依次经过 1×1×1 的卷积,批量归一 化处理和 ReLU 函数激活,接着通过 Softmax 函数激 活和最大值池化处理生成注意力参数与上层特征进行 加权,得到本层的特征输出.通常在 Encoder-Decoder 框架中添加注意力机制是和 Attention UNet^[18] 类似,在 编码层和解码层的跳跃连接中添加注意力机制,而本 文则是在编码层之间添加注意力机制,因为我们只是 关注裂缝特征在特征提取过程中的细节特征丢失,所 以采用了在编码层之间连续使用符合各个解码层特征 的注意力门控机制来对各层特征进行加权.



$$F' = \alpha \times F + F \tag{1}$$

注意力参数的生成过程为:

$$F_{\text{att}}^{l} = \sigma_1 \left(B N_{\gamma,\beta} \left(W_i^{\mathrm{T}} x_i^l + b_i^l \right) \right) \tag{2}$$

$$F^{l} = F_{\text{att}}^{L} \left(x_{i}^{l}; \Theta_{i}^{l} \right) \tag{3}$$

$$\alpha = P(\sigma_2(F^l)) \tag{4}$$

在式(2)中, x_i^l 是第l层的输入, W_i^T 和 b_i^l 分别为第

Software Technique•Algorithm 软件技术•算法 251



Attention module

图 3 连续注意力机制结构

l 层的权重和偏量,如果特征的数量级过大,在通过激活函数时会提前进入饱和区,因此需要采取批量归一 化 $BN_{\gamma\beta}, \sigma_1$ 是 ReLU 激活函数, F_{att}^l 是注意力特征生成 公式, F^l 为转换后注意参数的特征,最后通过 Softmax 激活函数增加非线性和 max-pooling 层归一化尺度得 到注意力参数.

2.2 多层卷积金字塔结构

传统的特征金字塔结构^[19,20] 主要从不同尺度的图 像中提取特征,但这增加了训练数据量,降低了效率. 目前用于路面裂缝检测的特征金字塔网络^[9,21] 大多是 利用网络最后一层的特征映射来构造特征金字塔,从 而保证了裂缝特征对路面裂缝的完整性.但是在一定 程度上缺乏解决裂纹边缘信息丢失的能力.因此,我们 提出了一种多层卷积特征金字塔结构,利用各编码层 的卷积特征进行特征融合.多层卷积特征金字塔结构 可以充分利用各层图像特征,既能保证裂纹特征的连 续性,同时也减少了裂纹边缘信息的丢失.

多层卷积特征金字塔如图 4 所示. 每层的输入为 编码器每一层的输出特征. 从上到下每层的特征图尺 寸为 64×240×160、128×120×80、256×60×40、 512×30×20、512×15×10.



因为各层的特征通道数和图像尺度不同,无法直接进行特征融合.使用 1×1×1 卷积对通道数进行变换,然后通过最大池改变图像特征尺寸的大小,最后,前四层金字塔结构融合生成 512 通道特征映射,再与编码器的最后一个特征层进行特征融合,最后得到解码层的输入特征图.

3 实验分析

3.1 训练细节

整个网络是由 PyTorch 实现的, 训练环境为 3.5 GHz,

252 软件技术•算法 Software Technique•Algorithm

8 GB RAM, NVIDIA 2080Ti GPU, 且在 Ubuntu 环境下运行.

3.2 数据集

比较实验采用 3 个公共数据集,包括 Crack500 数据集^[9], Crack200 数据集^[4]以 CrackForest 数据集 (CFD)^[22].因为 3 种数据集图片的大小是不同的,我们 有所有的图像和地面真相调整到 480×320 分辨率.同时,由于 CFD 和 crack200 数据集中的数据量太小,我 们随机选取翻转输入图像以增强数据以防止过度拟合.

3.3 评价指标

为了比较生成的分割图像与真实分割图像的像素级 相似度,分割任务中主要使用 4 个指标,分别是 Precision、 Recall、F1-measure 和 MIoU.

Precision 是指所有样本中真正样本占所有正样本的百分比, Recall 是正样本被模型判定为正样本的百分比, F1-measure 是衡量二分类精度的一个指标. 它同时考虑了分类器的准确性和召回率.

因此, Precision 可以定义为:

$$Precision = \frac{TP}{TP + FP}$$
(5)

Recall 可以被定义为:

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$F1$$
-measure 可以定义为:
 $F1 = \frac{2*Precision*Recall}{Precision+Recall}$
(7)

MIoU可以被定义为:

$$MIoU = \frac{TP}{TP + FN + FP}$$
(8)

式(6)-式(9),要计算这4个指标,我们需要计算 3个指标,TP是正样本预测为正样本的个数,FP是预 测为负样本的正样本个数,FN是预测为负样本的负样 本个数.

3.4 参数设置

进行训练时我们设置了两个停止训练的条件,一个是训练步数达到了设定值,第二个是学习率下降到 1e-9. 训练的 batch 设置为 6.

由于交叉熵损失不适用分布不均匀的数据集,因此我们采用 Tversky 损失^[23]和 L1 正则损失的加权作为此次实验的损失函数. Tversky 系数是 Dice 系数和 Jaccard 系数加权之后的广义系数. 不平衡分布路面裂 缝可能导致某些数据出现异常值,但 L1 正则损失可以

处理数据中的异常值.因此,我们采用 L1 损失来提高 系统的鲁棒性.

因此损失函数可以定义为:

$$loss(t,p) = \alpha * T(t,p) + (1-\alpha)(L1(t,p))$$
(9)

式 (11) 中, *t* 为真实分割图像, *p* 为预测分割图像, *T* 为 Tversky 损失函数, *L*1 为 *L*1 正则损失函数, *α* 为权重系 数, 设置为 0.7.

3.5 实验结果

如表1和表2所示,为本文定量实验结果.图5至 图7为3种数据集上可视化检测结果.从图5可以看出, 与图5(a)-图5(e)裂缝检测器的结果相比,CACPNet 检测的裂缝边缘信息更丰富,而不是平滑的.从图6可 以看出,裂缝在原始图像和真实检测图像中是连续分 布的,没有断开,但是除CACPNet的检测结果外,其他 的裂缝检测器检测出的裂缝发生了断裂.图7是CFD 数据上的检测结果,可以看出检测的效果也是要优于 其他的检测器.图8为3种不同数据集训练时的损失 和 *MIoU* 指标曲线图.

表 1	3种数据集上的 MIoU 指标比较			
方法	Crack200	Crack500	CFD	
CrackIT	0.397	0.434	0.454	
CrackForest	0.501	0.563	0.506	
FPHBN	0.420	0.600	0.405	
SegNet	0.507	0.736	0.557	
文献[21]	0.528	0.736	0.620	
CACPNet	0.657	0.753	0.721	

表 2 CFD 数据集上 Precision, Recall, F1-measure 指标 (%)

方法	Precision	Recall	F1-measure
CrackIT	67.23	76.69	71.64
CrackTree	73.22	76.45	70.80
CrackForest	82.28	89.44	85.71
SegNet	83.46	87.50	85.16
MFCD	89.90	89.47	88.04
CACPNet	92.37	97.68	94.91





3.6 消融实验

我们通过消融实验验证了 CACPNet 各个部分的 重要性,在 CFD 数据集上设计了与原生 SegNet 的对

比实验,所有参数设置,包括学习率和损失函数都保持 一致.消融实验一共为3组,第1组为原生 SegNet,第 2组为原生 SegNet 加上连续注意力机制,第3组为第

Software Technique•Algorithm 软件技术•算法 253

二组的基础上加上多层卷积金字塔结构.具体结果见表 3. 此外,我们还绘制了 P-R 曲线来验证添加注意机 制和金字塔模型的有效性.一般来说,精确度降低,召 回率提高. 从图 9 中可以看出, 融合了注意机制和金字 塔模型的曲线高于其他曲线, 进一步证明了我们提出 的方法的必要性.





表 3 消融实验结果							
Attention	Pyramid	Precision (%) <i>Recall</i> (%)	F1-measure (%)			
Ν	Ν	83.46	87.50	85.16			
Y	Ν	91.66	92.54	92.03			
Y	Y	92.37	97.68	94.91			
Precision		SegNet SegNet wi CACPNet	th attention	0.8 10			
	0	0.2 0.4	0.0	0.8 1.0			
кесан							
图 9 消融实验 P-R 曲线							

4 结论与展望

本文提出了一种基于 SegNet 的路面裂缝检测模型. 我们基于 SegNet 融合了连续特征注意机制与多层卷积金字塔结构. 连续特征注意力机制可以保证裂纹之间的连续性, 在某种程度上, 减少了检测裂缝的断裂. 多层卷积金字塔结构采用对每层特征输出进行再利用, 保证全局信息的完整性, 包括可以对细微裂缝检测, 同

254 软件技术•算法 Software Technique•Algorithm

时可以保证裂缝边缘信息的完整性.此外, CACPNet 已 经在3个公共数据集上进行了实验验证,包括 Crack500 数据集, Crack200 数据集和 Crack Forest 数据集.与现 有的方法相比,经过实验验证,我们的网络在各种指标 上都得到了一定的提升.此外,还进行了消融实验以验 证其每个模块的有效性.

参考文献

- 1 Nejad FM, Zakeri H. An optimum feature extraction method based on wavelet-radon transform and dynamic neural network for pavement distress classification. Expert Systems with Applications, 2011, 38(8): 9442–9460. [doi: 10.1016/ j.eswa.2011.01.089]
- 2 Hu Y, Zhao CX, Wang HN. Automatic pavement crack detection using texture and shape descriptors. IETE Technical Review, 2010, 27(5): 398–405. [doi: 10.4103/0256-4602.62225]
- 3 Oliveira H, Correia PL. Automatic crack detection on road imagery using anisotropic diffusion and region linkage. Proceedings of the 18th European Signal Processing Conference. Aalborg, Denmark. 2010. 274–278.
- 4 Zou Q, Cao Y, Li QQ, *et al.* Cracktree: Automatic crack detection from pavement images. Pattern Recognition Letters, 2012, 33(3): 227–238. [doi: 10.1016/j.patrec.2011. 11.004]
- 5 Oliveira H, Correia PL. CrackIT An image processing

toolbox for crack detection and characterization. Proceedings of 2014 IEEE International Conference on Image Processing (ICIP). Paris, France. 2014. 798–802.

- 6 Peng L, Chao W, Li SM, *et al.* Research on crack detection method of airport runway based on twice-threshold segmentation. Proceedings of the 5th International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC). Qinhuangdao, China. 2015. 1716–1720.
- 7 Zhang KG, Cheng HD, Zhang BY. Unified approach to pavement crack and sealed crack detection using preclassification based on transfer learning. Journal of Computing in Civil Engineering, 2018, 32(2): 04018001. [doi: 10.1061/(ASCE)CP.1943-5487.0000736]
- 8 Zou Q, Zhang Z, Li QQ, *et al.* Deepcrack: Learning hierarchical convolutional features for crack detection. IEEE Transactions on Image Processing, 2019, 28(3): 1498–1512. [doi: 10.1109/TIP.2018.2878966]
- 9 Yang F, Zhang L, Yu SJ, et al. Feature pyramid and hierarchical boosting network for pavement crack detection. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(4): 1525–1535. [doi: 10.1109/TITS.2019.2910595]
- 10 Dorafshan S, Thomas RJ, Maguire M. Comparison of deep convolutional neural networks and edge detectors for 33-8044-based crack detection in concrete. Construction and Building Materials, 2018, 186: 1031–1045. [doi: 10.1016/j. conbuildmat.2018.08.011]
- 11 Zhang KG, Zhang YT, Cheng HD. CrackGAN: Pavement crack detection using partially accurate ground truths based on generative adversarial learning. arXiv: 1909.08216, 2019.
- 12 Li HF, Song DZ, Liu Y, *et al.* Automatic pavement crack detection by multi-scale image fusion. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(6): 2025–2036. [doi: 10.1109/TITS.2018.2856928]
- 13 Yusof NAM, Ibrahim A, Noor MHM, *et al.* Deep convolution neural network for crack detection on asphalt pavement. Journal of Physics: Conference Series, 2019, 1349(1): 012020.
- 14 Ai DH, Jiang GY, Kei LS, et al. Automatic Pixel-level pavement crack detection using information of multi-scale

neighborhoods. IEEE Access, 2018, 6: 24452-24463. [doi: 10.1109/ACCESS.2018.2829347]

- 15 Badrinarayanan V, Handa A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. arXiv: 1505.07293, 2015.
- 16 Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481–2495. [doi: 10.1109/ TPAMI.2016.2644615]
- 17 Schmugge SJ, Rice L, Lindberg J, et al. Crack segmentation by leveraging multiple frames of varying illumination. Proceedings of 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). Santa Rosa, CA, USA. 2017. 1045–1053.
- 18 Oktay O, Schlemper J, Le Folgoc L, *et al.* Attention U-Net: Learning where to look for the pancreas. arXiv: 1804.03999, 2018.
- 19 Lazebnik S, Schmid C, Ponce J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. Proceedings of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). New York, NY, USA. 2006. 2169–2178.
- 20 Young J. Spatial pyramid match kernels for brain image classification. Proceedings of 2016 International Workshop on Pattern Recognition in Neuroimaging (PRNI). Trento, Italy. 2016. 1–4.
- 21 Xiang XZ, Zhang YQ, El Saddik A. Pavement crack detection network based on pyramid structure and attention mechanism. IET Image Processing, 2020, 14(8): 1580–1586. [doi: 10.1049/iet-ipr.2019.0973]
- 22 Shi Y, Cui LM, Qi ZQ, *et al.* Automatic road crack detection using random structured forests. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(12): 3434 – 3445. [doi: 10.1109/TITS.2016.2552248]
- 23 Salehi SSM, Erdogmus D, Gholipour A. Tversky loss function for image segmentation using 3D fully convolutional deep networks. Proceedings of the 8th International Workshop on Machine Learning in Medical Imaging. Quebec City, QC, Canada. 2017. 379–387.

Software Technique•Algorithm 软件技术•算法 255