

doi: 10.3969/j.issn.1002-0268.2024.08.002

基于 DCGAN 预处理和残差密集注意力网络的路面裂缝识别方法

王强¹, 王浩仰^{*2,3,4}, 高保全⁵, 付渊⁶, 陈煜东⁷

- (1. 山西省交通规划勘察设计院有限公司, 山西 太原 030032; 2. 中公高科养护科技股份有限公司, 北京 100095;
3. 交通运输部公路科学研究院, 北京 100088; 4. 公路与桥梁高效养护及安全耐久国家工程研究中心, 北京 100088;
5. 山西省交通运输厅, 山西 太原 030031; 6. 山西省公路局, 山西 太原 030032;
7. 北京邮电大学 网络空间安全学院, 北京 100876)

摘要: 为了建立一种基于深度学习卷积神经网络 (DCGAN) 的路面裂缝检测模型, 提高特殊路面裂缝 (如模糊裂缝、浅裂缝等) 的识别准确率, 提出了基于 DCGAN 预处理和残差密集注意力网络 (RDAN) 的路面裂缝识别方法。该方法基于对抗学习的思想, 在数据预处理中保留了高分辨率的裂缝特征, 同时抑制背景噪声的影响, 将模糊裂缝增强为清晰裂缝, 将浅色裂缝重建为深色裂缝, 提升模型训练、识别效果; 在模型训练中, 提出了一种全新的适用于路面裂缝识别的 RDAN; 基于传统的残差网络和密集网络的结构进行融合改进, 在元素维度和通道维度同步进行层与层之间特征的传递, 改善深度卷积神经网络训练过程中的退化问题; 通过引入卷积块注意力模块, 在空间与通道 2 个维度实现自适应特征优化, 进一步提升网络在实际复杂路面背景下的裂缝特征提取能力, 抑制背景噪声的干扰。结果表明: 基于 DCGAN 预处理和 RDAN 的路面裂缝识别方法在 CiCS50000 实际工程数据集及多个公开数据集上取得了最优的识别效果, 在 CiCS50000 数据集准确率、精确率、召回率和 Dice 相似系数上分别达到了 97.51%, 87.05%, 83.36% 和 81.02%。

关键词: 智能交通; 裂缝识别方法; 生成对抗网络; 残差密集连接; 注意力机制

中图分类号: U418.1

文献标识码: A

文章编号: 1002-0268 (2024) 08-0011-11

Pavement Crack Recognition Method Based on DCGAN Preprocessing and Residual Dense Attention Network

WANG Qiang¹, WANG Hao-yang^{*2,3,4}, GAO Bao-quan⁵, FU Yuan⁶, CHEN Yu-dong⁷

- (1. Shanxi Traffic Planning Survey and Design Institute Co., Ltd., Taiyuan, Shanxi 030032, China;
2. RoadMainT Co., Ltd., Beijing 100095, China; 3. Research Institute of Highway, Ministry of Transport, Beijing 100088, China;
4. National Engineering Research Center for Efficient Maintenance, Safety and Durability of Highways and Bridges, Beijing 100088, China; 5. Department of Transportation of Shanxi Province, Taiyuan, Shanxi 030031, China;
6. Shanxi Provincial Highway Bureau, Taiyuan, Shanxi 030032, China;
7. School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: To establish a type of pavement crack detection model based on deep learning convolutional neural network for improving the recognition accuracy for special pavement cracks (e.g., fuzzy crack and light-colored crack), the pavement crack detection method based on deep convolution generative adversarial network (DCGAN) preprocessing and residual dense attention network (RDAN) was proposed. In the data

收稿日期: 2024-06-17

基金项目: 中路高科交通科技集团有限公司交通强国试点项目 (JTQG2022-1-3-1)

作者简介: 王强 (1983-), 男, 山西太原人, 硕士, 高级工程师. (kg.83@163.com)

* 通讯作者: 王浩仰 (1989-), 男, 山西长治人, 博士, 副研究员. (wanghaoyang@roadmaint.com)

preprocessing module, guided by the adversarial learning concept, the method preserves high-resolution crack features while suppressing the influence of background noise. The fuzzy cracks can be enhanced to clear cracks. The light-colored cracks can be reconstructed to dark cracks. The training and recognition effect of subsequent models can be improved. In the model training module, a wholly new RDAN for pavement crack recognition was proposed, which was developed through the fusion and improvement of traditional ResNet and DenseNet structures. The layer-to-layer characteristics transfer was synchronized between the element dimension and channel dimension to improve the degradation problem during training of deep convolutional neural networks. The adaptive feature optimization was realized in 2 dimensions of space and channel by introducing the convolutional block attention module to further improve the crack feature extraction ability and suppress the background noise interference in the actual complex road background. The result indicates that the pavement crack recognition method based on DCGAN preprocessing and residual dense attention network has achieved the optimal recognition effect on the CiCS50000 real engineering dataset and multiple public datasets. The accuracy, precision, recall rate and Dice coefficient of the CiCS50000 dataset reach 97.51%, 87.05%, 83.36%, 81.02 respectively.

Key words: intelligent transport; crack recognition method; generative adversarial network; residual dense connection; attention mechanism

0 引言

随着高性能计算机技术的发展,以深度卷积神经网络为代表的人工智能技术逐步在路面裂缝检测领域替代了传统的数字图像处理技术,其在效率、精度等方面均展示了明显的优势。同时,此类技术更适用于“端到端”的路面裂缝检测方法^[1-7]。在数据预处理方面,常用方法包括均值滤波、中值滤波、双边滤波、直方图均衡化和自适应直方图均衡化^[8]等;在网络方面,为了解决训练过程中退化问题,He等^[9]首先将残差学习的思想应用到卷积神经网络,提出了ResNet;不同于ResNet对于网络深度的研究,Huang等^[10]提出了一种网络密集连接方式DenseNet,选择从特征图入手,增强了前后层之间的特征传播与特征服用,相比ResNet减少了参数计算量;Yang等^[11]提出一种特征金字塔和层次增强网络(FPHBN)用于路面裂缝识别,加强不同层次间特征信息的结合,通过对不同层次损失进行加权计算,训练裂纹检测网络;Feng等^[12]提出了一种带有残差连接的多尺度金字塔结构,通过对不同尺度的特征图进行加权融合,改善卷积神经网络在下采样过程中导致的裂缝特征丢失的问题;Nguyen等^[13]针对道路裂缝背景复杂等问题,提出了一种基于灰度图像的深度神经网络DNN的像素级裂缝识别方法,训练数据来源于公开数据集,图片数量仅有几百张;Tong等^[14]使用8 820张图片,提出了基于全卷积网络和一个不确定的框架的像素级裂缝识别方法;Huyan

等^[15]使用3 000张图片提出了基于U-net网络的像素级裂缝识别方法。由于路面背景存在复杂性(如阴影、污渍、标线、修补等),裂缝形态存在多样性,以模糊裂缝和浅色裂缝为代表的特殊裂缝成为制约路面裂缝自动化检测算法总体识别准确率的关键因素。然而,现有方法在实际应用过程中,存在以下不足。

(1) 由于实际路面状况复杂多变,背景噪声众多,传统的图像增强与去噪方法很难在不引入负面效应的条件下完全消除这些影响。同时,传统的路面裂缝预处理算法往往基于像素灰度进行判断,并没有针对裂缝的特征与背景噪声进行准确地区分,由于实际情况下路面裂缝的种类繁多,尤其是对于类似浅色裂缝为代表的特征存在明显差异的特殊裂缝,无法取得有效的预处理效果。

(2) 由于目前裂缝识别领域的论文大多使用分辨率较小、背景干扰较少、裂缝特征较为明显的图像数据集来做相关研究,训练出的模型难以准确拟合实际路面复杂背景下的裂缝特征,往往会造成大量的误识别和漏识别,特别是对以浅色裂缝和模糊裂缝为代表的特殊裂缝类型识别效果较差。

(3) 经典的用于图像分类的卷积神经网络ResNet和DenseNet的核心思想都是通过建立前、后层之间的连接,从而增强特征信息的传播,帮助训练更深层次的卷积神经网络。在此基础上,可尝试基于ResNet和DenseNet的融合改进网络结构,同时保留前后层之间特征图的元素级连接和通道级

连接, 增强特征信息传递, 此外将卷积模块注意力机制引入融合改进网络结构, 在通道和空间 2 个维度实现自适应特征优化, 以达到更好的裂缝识别结果。

因此, 本研究主要从特殊裂缝的预处理和裂缝识别网络的构建两方面入手提出改进方法, 着重解决实际工程化应用中由于特殊裂缝识别困难造成的

总体识别准确低的问题, 实现路面裂缝自动化识别能力的提升。

1 本研究方法

本研究方法框架如图 1 所示, 其整体是基于网格的端到端裂缝分割网络; 输入为裂缝图像; 输出为对于该图像每个格子包含裂缝的预测概率矩阵。

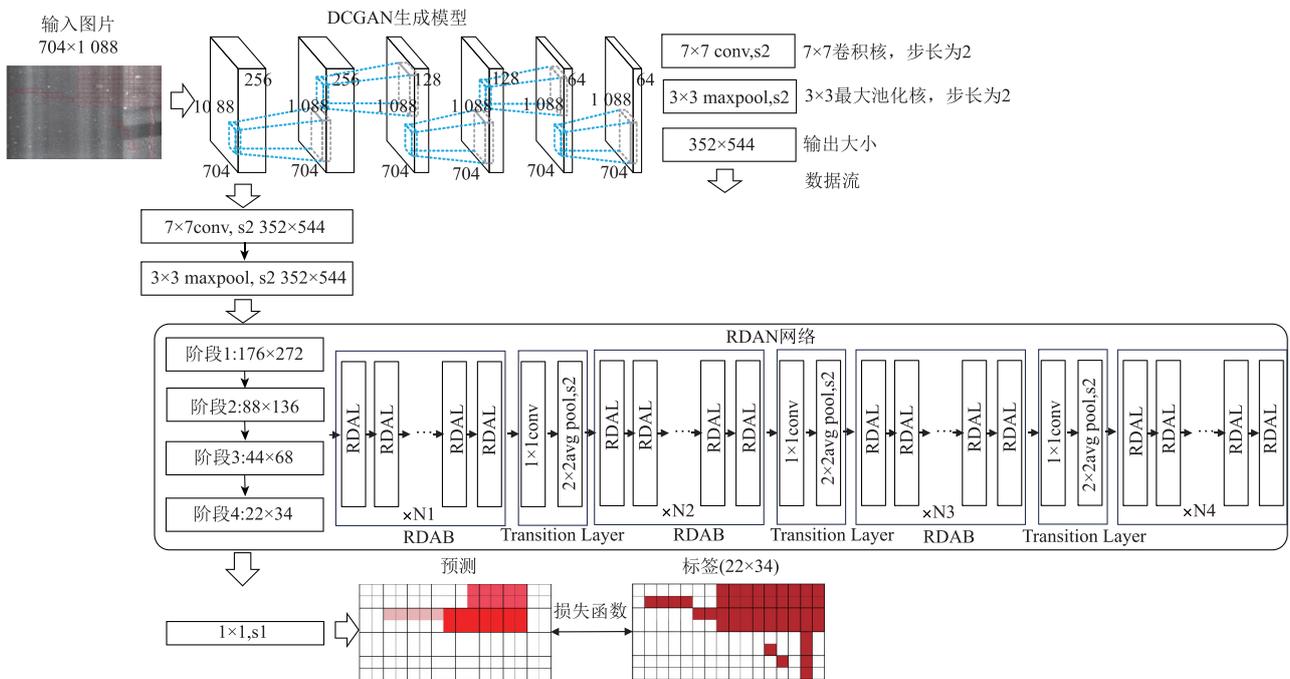


图 1 方法框架

Fig. 1 Method framework

1.1 图片标注

在实际工程化应用过程中, 由于任务量大, 裂缝数量和种类繁多, 采用像素级标记耗费大量的时间和人力, 不符合工程化应用的实效性需求。《多功能路况自动化检测设备》(GB/T 26764—2011)^[16]《公路技术状况评定标准》(JTG 5210—2018)^[17]和配套的《公路技术状况自动化检测规程》(JTG/T E61—2014)^[18]等标准均采用网格标注方法, 采用 10 cm×10 cm 或 100×100 像素的网格, 不同颜色的网格代表不同的路面损坏类型。

本研究实际工程数据集图片, 由多功能路况快速检测设备采集得到, 图像尺寸均为 2 000×3 000 像素左右。为了符合网络的输入输出标准, 首先对图像进行补全 (Padding) 操作, 用 255 灰度值填充原始路面灰度图的下侧与右侧, 将图像补全为 2 200×3 400 像素, 均以 100×100 像素格为单位进行人工标注; 每个格子对应标记为 1 或 0, 其中 1 代表该格子区域内有裂缝, 0 代表无裂缝, 对于每

张图像的裂缝标记矩阵, 即为 22×34。

1.2 图片预处理

采取单通道灰度图作为网络输入, 用于训练的原始路面图像数据尺寸全部被标准化处理为 704×1 088 像素。

1.2.1 深度卷积生成对抗网络 (DCGAN)

DCGAN 生成模型 (G) 和判别模型 (D) 的网络结构构建, 如图 2 所示。其中生成模型的作用是接收浅色裂缝、模糊裂缝样本, 经过全卷积的网络结构, 实现非线性映射, 输出裂缝特征增强或重建后的样本。判别模型的作用是接收清晰裂缝样本和生成模型输出的预处理样本, 在卷积的过程中总共经过 5 次下采样操作, 将原始尺寸为 704×1 088 的输入图像, 输出尺寸为 22×34 的裂缝概率预测矩阵, 用于对裂缝位置的监督以及对于输入数据“真”、“假”的判断。相比原始的 GAN, 本研究在 DCGAN 的设计过程中去掉了全连接层, 采用全卷积的网络结构。同时, 使用带步长的转置卷积层代

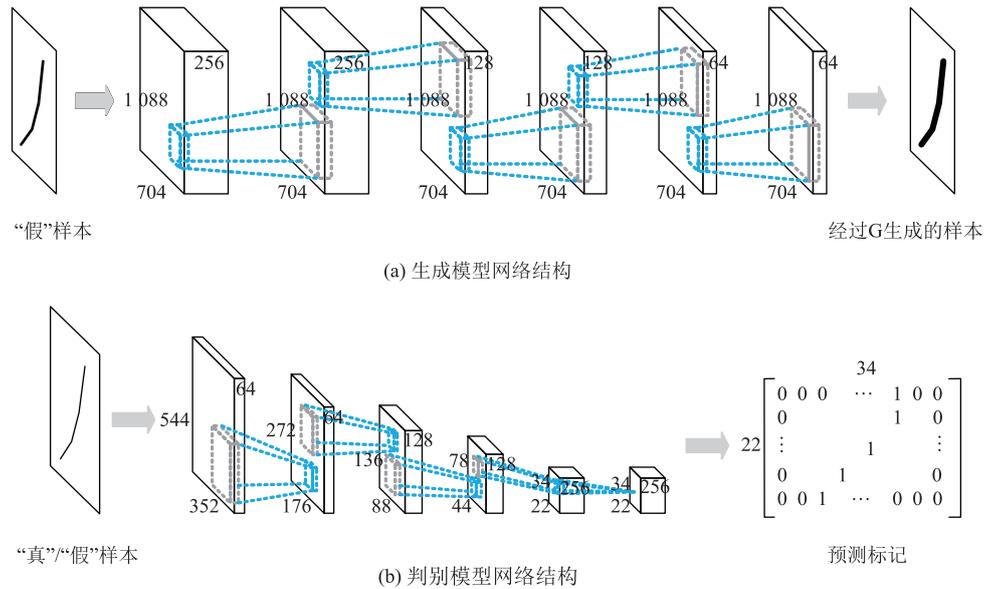


图2 深度卷积生成对抗网络

Fig. 2 Deep convolution generative adversarial network

替上采样和池化层操作,这样做的好处是可以增加对抗网络训练的稳定性。

在具体的训练过程中,首先挑选一定数量的浅色裂缝样本、模糊裂缝样本,和相等数量的清晰裂缝样本作为对抗样本随机配对。其中,区别于原始GAN的输入数据为随机噪声或真实图像样本,本研究使用浅色裂缝或模糊裂缝样本代替原始GAN输入的随机噪声,使用清晰裂缝样本代替原始GAN输入的真实样本,以此使生成模型实现“由图像到图像”的预处理映射。

对抗训练的过程,首先固定生成模型权重,将模型容易识别的清晰裂缝的“真”样本,与模型难以识别的特殊裂缝图像经过生成模型生成的“假”样本,分别输入对抗网络的判别模型进行训练,清晰裂缝样本对应真实裂缝标记矩阵,而特殊裂缝样本对应全“0”标记矩阵,更新判别模型的网络权重;之后,固定判别模型权重,对特殊裂缝图像经过生成模型生成的“假”样本,对应真实裂缝标记,更新生成模型的网络权重。通过对抗训练,最终实现将浅色裂缝经过生成模型后重建为深色裂缝,将模糊裂缝经过生成模型后增强为清晰裂缝。同时,在训练过程中,由于在损失函数的设定过程中加入对于裂缝位置的监督,使得模型重点关注裂缝部分的特征的预处理,因此在生成的图像中,可以起到淡化背景噪声(例如阴影、污渍等)的效果,很大程度上减少后续裂缝识别模型训练、识别过程中的误识别、漏识别情况。

最后,将生成模型用于图片预处理,对输入图像进行预处理增强,提升模型训练效果。同时,该预处理过程也可视为生成模型对于不同类型的裂缝进行特征归一化,统一输出为深色且清晰的裂缝特征,同样有助于加速裂缝识别模型的收敛。

1.2.2 损失函数

对于一张路面裂缝图像,由于图像中真正有裂缝的部分往往占比很少,会导致正负样本分布不均衡,通过设定Dice Loss作为判别模型的损失函数来指导生成对抗网络的训练过程。Dice Loss最早在V-Net^[17]中被提出并广泛用于医学图像分割,根据Dice相似系数计算得到。Dice相似系数是一个用于度量2个集合之间相似程度的指标,它的取值范围为[0, 1],越接近1表示模型预测结果和真实标记之间的相似程度越高,如式(1)所示:

$$S = \frac{2|X \cap Y|}{|X| + |Y|}, \quad (1)$$

式中, $|X \cap Y|$ 为2个集合内正样本的交集; $|X|$ 和 $|Y|$ 为2个集合内各自的正样本的个数。本研究采用拉普拉斯平滑方法,在Dice Loss的分子和分母上同时加上一个较小的平滑系数 α , 取值为0.0001,防止出现分母为0的情况。Dice Loss通过用1减去Dice相似系数来计算,表达式为:

$$\text{Dice_Loss} = 1 - \frac{2|X \cap Y| + \alpha}{|X| + |Y| + \alpha} \quad (2)$$

1.3 残差密集注意力网络

通过结合传统ResNet和DenseNet的基本结构并

加以改进, 引入 CBAM 注意力机制, 分别提出了基于层级的基本连接单元 RDAL, 基于模块级的基本连接单元 RDAB 和完整的裂缝识别网络 RDAN。如图 1 所示, RDAN 由 4 个 RDAB 和 3 个 Transition Layer 交替构成, 每个 RDAB 又由不同数量的 RDAL 构成。在 RDAB 中特征图不断叠加, 使得传递特征维度不断增大, 在 RDAB 后连接过渡层, 通过池化核为 2×2 、步长为 2 的平均池化操作实现一次下采样, 降低维度减少计算量。

在 RDAN 前端采用卷积核为 7×7 、步长为 2 的卷积层和池化核为 3×3 、步长为 2 最大池化层, 在实现降维的同时通过较大的感受野保留突出的纹理特征, 同时完成对图像的两次下采样操作; 在 RDAN 后端移除了分类网络最后的全连接层, 采用 1×1 的卷积层将特征图维度降维为 1; 最后经过一个 Sigmoid 激活函数, 将特征图每个格子输出的参数值映射到 $[0, 1]$ 之间, 得到模型对每个格子包含裂缝的预测概率值, 每个格子的预测概率值组合起来得到对于整张图的预测概率矩阵。通过设定阈值进行二值化判断, 在本研究中阈值设为 0.5, 即如果当前格子的预测概率大于 0.5, 则认为这个格子内有裂缝, 标记为 1; 反之认为这个格子内没有裂缝, 标记为 0。

基于模块化的网络结构设计可以根据不同的算力水平和数据量大小, 对网络的深度进行动态调整。本研究设计了 3 种不同深度的 RDAN 网络结构, 如表 1 所示。

表 1 不同深度的 RDAN 网络结构
Tab. 1 RDAN with different depths

层级	输出大小	RDAN-Small	RDAN-Medium	RDAN-Large
RDAB (1)	176×272	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 6$
Transition Layer (1)	176×272	1×1 conv		
Layer (1)	88×136	2×2 average pool, stride 2		
RDAB (2)	88×136	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 12$
Transition Layer (2)	88×136	1×1 conv		
Layer (2)	44×68	2×2 average pool, stride 2		
RDAB (3)	44×68	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 8$	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 24$
Transition Layer (3)	44×68	1×1 conv		
Layer (3)	22×34	2×2 average pool, stride 2		
RDAB (4)	22×34	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 8$	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 8$	$\begin{bmatrix} 1 \times 1 \text{conv} \\ 3 \times 3 \text{conv} \end{bmatrix} \times 16$

1.3.1 残差密集注意力层

本研究提出的残差密集注意力层 (RDAL) 的基本结构如图 3 所示。对于每 1 个 RDAL, 输入特征图 F 分别沿 3 条支路在前馈网络中进行前向传播, 其中黑色实线为中心支路; 中心支路下的虚线为残差支路; 中心支路上的虚线为密集连接支路。

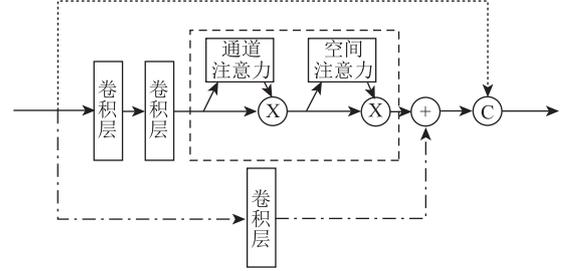


图 3 残差密集注意力层

Fig. 3 Residual dense attention layer

对于中心支路, 输入特征首先经过一组瓶颈层, 瓶颈层指 BN-RELU-Conv (1×1)-BN-RELU-Conv (3×3) 结构, 在减少参数的同时提高了计算效率; 之后, 特征图被发送到 CBAM 模块, 依次通过通道注意力模块和空间注意力模块, 生成注意力权重; 最后 CBAM 将注意力权重与特征图相乘以实现自适应特征优化。本研究将通道注意力模块的非线性映射表示为 $H_c(\cdot)$, 将空间注意力模块的非线性映射表示为 $H_s(\cdot)$ 。则对于第 l 层的输入特征图 F_l 来说, 第 l 层中心支路的输出特征图 $F_{\text{core}(l+1)}$ 为:

$$F_{\text{core}(l+1)} = H_s(H_c(F_l) \otimes F_l) \otimes (H_c(F_l) \otimes F_l) \quad (3)$$

输入经过红色线表示的残差支路, 残差支路的输出与核心支路的输出进行相加。由于残差连接是元素级相加, 因此需要保证特征图具有相同的通道维度。本研究在残差支路中添加一个线性投影 W_{res} 来调整特征图的维度, 这里 W_{res} 使用一个 1×1 的卷积层来实现通道维度匹配, 则对于第 l 层的输入特征图 F_l 来说, 第 l 层残差支路的输出特征图 $F_{\text{res}(l+1)}$ 为:

$$F_{\text{res}(l+1)} = W_{\text{res}}(F_l) \quad (4)$$

最后, 将中心支路与残差支路元素级相加后得到的特征图, 与前面所有的 RDAL 生成的特征图在通道层面进行拼接, 作为所有后续 RDAL 的输入。如图 1 中 RDAN 所示, 本研究使用蓝色线表示在前向传播中, 始终维护一组特征图的集合拼接, 每经过一个 RDAL 都会向该集合中加入当前层输出的特征图, 并且一直将这个集合传递给后续所有层。对于第 l 层, 本研究将密集支路对输入特征的非线性操

作统称为 $H_l(\cdot)$ ，则第 l 层的输出 F_{l+1} 为：

$$F_{l+1} = H_l[F_{\text{core}(0)} + F_{\text{res}(0)}, \dots, F_{\text{core}(l)} + F_{\text{res}(l)}]。 \quad (5)$$

在 RDAN 中，底层结构还是沿用 DenseNet 中的密集连接，这样做的好处是，可以使用更少的参数计算量实现更深层次的特征复用。而在 RDAL 内部连接中，依靠引入注意力机制，实现了前馈传播过程中拼接特征图的信息增强，同时依靠 ResNet 中的短路连接，保证最差学习情况也是原始特征图的恒等映射，进而改善训练退化问题。

1.3.2 残差密集注意力模块

如图 1 中 RDAB 中结构所示，RDAB 由一系列 RDAL 相连所组成，每经过一个 RDAL，前馈网络中的特征图维度增加 $growthrate$ ，本研究中 $growthrate$ 设置为 32。设进入 RDAB 的特征图维度为 D_{in} ，当前 RDAB 由 N 层 RDAL 构成，则这个 RDAB 的输出特征图维度 D_{out} 为：

$$D_{out} = D_{in} \times growthrate。 \quad (6)$$

1.3.3 过渡层

随着网络层数的增加，经过 RDAB 连接之后的

特征图维度会不断增大，如式 (6) 所示。参数过多会导致计算效率低下，因此本研究采用 DenseNet 网络中的过渡层，进行参数压缩。如图 1 中 RDAN 中结构所示，过渡层由 BN-RELU-Conv(1×1)-AvgPool(2×2, stride 2) 的组合构成，通过卷积核为 1×1 的卷积层，将通道数压缩为一半，减少参数计算量。通过步长为 2 的平均池化操作，将每一个输入特征图的长和宽分别缩减为原来的 1/2，下采样后的特征图大小为原始的 1/4。

1.3.4 损失函数

对于裂缝识别网络，同样使用 Dice Loss 作为损失函数指导训练。优化方法选用 Adam 优化法对网络参数自适应地优化更新。1 阶矩衰减率 β_1 取值选择 0.9；2 阶矩衰减率 β_2 取值选择 0.999；学习率 α 取值选择 10^{-3} ；平滑因子 ϵ 取值选择 10^{-8} 。

2 试验设计

2.1 试验数据

(1) 实际工程数据集包括 CiCS50000 裂缝精修数据集，如表 2 所示。

表 2 工程数据集

Tab. 2 Engineering dataset

数据集名	图片数/张	图像像素	采集方式	说明
CiCS50000 裂缝精修	50 000	约 2 000×3 000	同一多功能路况快速采集设备，采集角度与路面垂直且相机放置于车顶，镜头到路面的距离保持不变	有裂缝和无裂缝的图像数量比例为 1:1，包含各种类型的裂缝（40 138 张清晰裂缝、2 132 张模糊裂缝、7 730 张浅色裂缝等）及各种类型的噪声干扰（油漆、污渍、标线、修补等）

(2) 公开数据集包括 CFD 数据集^[18]、AigleRN 数据集^[19]和 CRACK500 数据集^[11]，如表 3 所示。为

了符合本研究的网络输入与评价标准，对 AigleRN 和 CRACK500 数据集中图片尺寸进行了裁剪。

表 3 公开数据集

Tab. 3 Public dataset

数据集名	图片数/张	图像像素	采集方式	说明
CFD	118	320×480	手机	—
AigleRN	38	462×991 或 462×311	Aigle-RN 系统	将图像尺寸裁剪为 448×960 像素或 448×288 像素
CRACK500	500	约 2 000×1 500	手机	将图像裁剪为 352×640 像素

2.2 硬件和软件环境

本研究试验的硬件环境为 1 台包括 4 个 NVIDIA TITAN X 图形处理器的深度学习服务器。每个 GPU 的单精度浮点计算能力约为 11TFlops，显存大小为 12 G。服务器内存为 128G，处理器为 2 个 Intel Xeon E5-2637v4，共有 8 个内核，最大频率为 3.7 GHz。本研究中的算法通过 Python3.6 进行编程与实现。

2.3 裂缝识别网络试验

本试验旨在验证本研究所提出的 RDAN 裂缝识别网络的效果，初步构建 3 种不同深度的 RDAN 结构，如表 1 所示，在 CiCS50000 实际工程数据集和 CFD，AigleRN，CRACK500 这 3 个公开数据集上分别进行训练、测试，并且将裂缝识别领域内其他主流裂缝识别算法在上述数据集上进行复现，在测试

集上通过计算各项评价指标对比各算法的优劣。

2.4 生成对抗网络预处理试验

为了探究本研究所提出基于生成对抗网络的路面裂缝图像预处理算法在实际工程应用过程中的有效性, 本研究分别设计了对比试验: 分别对比了 DCGAN 预处理后对于各种网络模型训练效果的提升情况, 和 DCGAN 预处理方法与传统预处理方法分别对于 RDAN 模型训练效果的提升情况。

2.5 评价指标

本研究中的问题场景结合了计算机视觉领域中的物体定位和图像分类两类问题, 若关注单个检测窗口则可以看作针对窗口小图像的二分类问题; 若关注整张路面图像则可以看作在图像中寻找裂缝区域的位置。针对某 1 个 100×100 像素的识别窗口, 根据识别结果的二分类与实际标注的二分类可以将单个窗口的识别结果分为 4 种情况:

(1) 真阳性 (True Positive, TP), 即识别结果为真, 实际标注为真;

(2) 真阴性 (True Negative, TN), 即识别结果为假, 实际标注为假;

(3) 假阳性 (False Positive, FP), 即识别结果为真, 实际标注为假;

(4) 假阴性 (False Negative, FN), 即识别结果为假, 实际标注为真。

由以上 4 个基本参数可以推导出以下几个应用最为广泛的评价指标。

准确率 (Acc): 对于一张路面裂缝图像, 准确率表示预测正确的格子数占总格子数的比例, 其中预测正确的情况包括正确的识别为有裂缝 (TP) 和正确的识别为无裂缝 (TN), 表示为:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

精确率 (P_r): 对于一张路面裂缝图像, 精确率表示模型识别为有裂缝的格子数中, 有多少确实预测正确, 因此精确率亦被称作查准率。

$$P_r = \frac{TP}{TP + FP} \quad (8)$$

召回率 (R_e): 对于一张路面裂缝图像, 召回率表示图像中实际为有裂缝的格子中, 有多少被模型正确地识别出来, 因此召回率亦被称作查全率。

$$R_e = \frac{TP}{TP + FN} \quad (9)$$

Dice 相似系数: 本研究引入 VNet 中提出的用于分割任务的 Dice 相似系数, 用于评价模型预测裂缝

矩阵与实际标记矩阵的正样本相似程度。在本问题中 Dice 相似度指标的计算方式等价于分类问题中的 F1 指数, 它是一种综合考虑精确率和召回率的参数, 可以用 TP , FP , FN 表示为:

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (10)$$

3 结果与讨论

3.1 裂缝识别网络试验

3.1.1 在 CiCS50000 数据集上进行对比试验

本研究将 CiCS50000 数据集按照 8 : 1 : 1 的比例随机划分为训练集、验证集和测试集。初始学习率设置为 0.0001, 每训练 50 轮衰减 10 倍。BatchSize 设置为 16, 每个模型共训练 200 轮, 取验证集指标最优轮次的模型在测试集上进行测试和自动化评价指标计算。各模型的参数量和测试集上评价指标结果如表 4 所示。

表 4 不同模型在 CiCS50000 数据集上的识别效果指标

Tab. 4 Recognition effect indicators of different models on CiCS50000 dataset

模型	参数量/ M	准确率	精确率	召回率	Dice 系数
ResNet34 ^[9]	81.33	0.971 0	0.824 8	0.807 6	0.748 3
ResNet50 ^[9]	90.14	0.969 9	0.806 8	0.815 8	0.743 6
ResNet101 ^[9]	162.85	0.970 3	0.823 6	0.809 1	0.750 0
U-Net ^[20]	18.08	0.969 3	0.746 0	0.823 8	0.705 8
Pyramid Residual Network ^[12]	7.69	0.965 0	0.668 1	0.805 1	0.624 5
ConvNet ^[21]	6.40	0.874 0	0.467 0	0.794 1	0.424 1
CrackForest ^[18]	—	0.766 7	0.157 5	0.665 3	0.101 1
DenseNet-Small ^[10]	8.83	0.970 1	0.825 3	0.784 3	0.729 3
DenseNet-Res-Small	9.93	0.971 2	0.842 5	0.795 3	0.749 9
本研究提出的 RDAN- Small	10.01	0.971 2	0.853 8	0.801 8	0.763 8
DenseNet-Medium ^[10]	11.43	0.971 0	0.825 2	0.806 5	0.748 7
DenseNet-Res-Medium	12.93	0.971 5	0.840 1	0.813 2	0.766 0
本研究提出的 RDAN- Medium	13.02	0.973 2	0.856 1	0.821 8	0.781 5
DenseNet-Large ^[10]	27.12	0.973 0	0.844 1	0.829 6	0.782 9
DenseNet-Res-Large	31.57	0.972 1	0.854 0	0.821 1	0.780 1
本研究提出的 RDAN- Large	31.72	0.973 5	0.868 1	0.827 1	0.793 0

从试验结果来看, 本研究所提出的 RDAN-Large 模型在 CiCS50000 实际工程数据集上, 获得了最高的准确率 97.35%, 最高的精确率 86.81% 以及最高

的 Dice 相似系数 79.30%，同时也保证了较高的召回率。通过对比同一深度下 RDAN 添加残差密集连接和 CBAM 模块前后的试验指标可以看出，对于相对浅层次的网络（RDAN-Small 和 RDAN-Medium），在添加了残差密集连接与 CBAM 模块之后，在精确率和召回率上都取得了明显的提升，综合指标 Dice 相似系数分别提升了 3.45% 和 3.28%；而对于较深层次的网络（RDAN-Large），精确率也由 84.41% 提升到 86.81%，综合指标 Dice 相似系数提升了 1.01%。

通过对比不同深度网络的试验指标，本研究发发现原始 Small 深度的 DenseNet 网络，在加入本研究所提出的残差密集连接和 CBAM 模块之后，精确度和 Dice 相似系数已经超过了原始 Medium 深度的 DenseNet 网络；同样的对于 Medium 深度的 DenseNet 网络，在加入本研究提出的残差密集连接和 CBAM 模块之后，精确度和 Dice 相似系数也达到接近原始 Large 深度的 DenseNet 网络的水平。综上所述，本研究的 RDAN 网络的改进使得原始 DenseNet 网络在较浅的层次即可达到或超过原始更深层次网络的识别效果，同时模型的参数量并没有太多的增加。

3.1.2 在 CFD 公开数据集上进行对比试验

本研究在 CFD 公共数据集上训练和测试各个模型。由于 CFD 数据集只有 118 张图像，因此本研究采用 5 折交叉验证训练方法。将数据集等分成 5 份，每次选取 4 份作训练集和验证集，1 份作测试集，训练集、验证集和测试集的比例为 7:1:2，图片数量分别为 82, 12, 24，最终对每折测试集上结果取平均。在训练过程中，BatchSize 设置为 16，模型共训练 200 轮。初始学习率设置为 0.000 1，每 50 轮衰减 10 倍。最后，挑选出测试集上最优效果进行对比分析，如表 5 所示。

表 5 不同模型在 CFD 数据集上的识别效果指标

Tab. 5 Recognition effect indicators of different models on CFD dataset

模型	准确率	精确率	召回率	Dice 系数
CrackForest	0.941 8	0.853 8	0.871 1	0.838 7
ConvNet	0.922 8	0.854 7	0.733 3	0.783 0
FPHBN	0.708 6	0.382 6	0.912 6	0.525 3
Pyramid Residual Network	0.961 9	0.941 4	0.865 5	0.894 8
本研究提出的 RDAN-Small	0.951 1	0.907 5	0.847 6	0.869 4
本研究提出的 RDAN-Medium	0.956 2	0.925 4	0.857 0	0.884 4
本研究提出的 RDAN-Large	0.961 4	0.932 2	0.871 4	0.894 6

在 CFD 公开数据集上，本研究所提出的 RDAN-Large 网络在精确率上略低于 Feng 等^[8]提出的

Pyramid Residual Network，而在召回率上略高于该网络，准确率和 Dice 相似系数基本持平。这是由于 CFD 数据集的图像路面背景比较干净清晰，没有太多噪声干扰，裂缝特征相对明显，因此采用结构相对简单的网络也能够取得不错的效果。

3.1.3 在 AigleRN 公开数据集上进行对比试验

本研究在 AigleRN 公开数据集上，对各个模型进行训练预测。由于该数据集图片数量同样较少，总共只有 38 张，因此本研究采用 5 折交叉验证的训练、测试方式，将数据集等分成 5 份，每次选取 4 份作训练集和验证集，1 份作测试集，训练集、验证集和测试集的比例为 7:1:2，图片数量分别为 26, 4, 8，最终对每折测试集上结果取平均。训练过程中，由于每张图像尺寸不同，本研究将 BatchSize 设为 1，初始学习率设为 0.000 1，训练 200 个 epoch，每 50 轮学习率衰减 10 倍，挑选出测试集上最优效果进行分析，如表 6 所示。

表 6 不同模型在 AigleRN 数据集上的识别效果指标

Tab. 6 Recognition effect indicators of different models on AIGLERN dataset

模型	准确率	精确率	召回率	Dice 系数
CrackForest	0.720 7	0.457 5	0.911 0	0.549 7
ConvNet	0.843 9	0.561 7	0.525 6	0.534 9
FPHBN	0.874 1	0.440 4	0.530 8	0.453 5
Pyramid Residual Network	0.920 4	0.742 7	0.848 6	0.736 9
本研究提出的 RDAN-Small	0.939 8	0.807 6	0.741 3	0.728 0
本研究提出的 RDAN-Medium	0.955 1	0.875 7	0.748 0	0.765 4
本研究提出的 RDAN-Large	0.955 9	0.894 5	0.743 3	0.779 9

本研究提出的 RDAN-Large 模型获得了最高的准确率 95.59%，最高的精确率 89.45% 及最高的 Dice 相似系数 77.99%。随着网络深度的加深，精确度提升效果越明显，进而在 AigleRN 数据集上取得了最好的效果。

3.1.4 在 CRACK500 公开数据集上进行对比试验

本研究在提出的 CRACK500 公开数据集上对各个模型进行训练预测。按照 Yang 等^[11]的做法，对 500 张图像进行裁剪，每张图像裁剪成 16 个不重叠的图像区域，最终只保留包含 1 000 像素以上裂缝的区域。训练集由 1 896 张图像构成，验证集由 348 张图像构成，而测试集由 1 124 张图像构成。训练过程中，将 BatchSize 设为 16，初始学习率设为 0.000 1，训练 200 个 epoch，每 50 轮学习率衰减 10 倍，挑选出测试集上最优效果进行分析，如表 7 所示。

本研究所提出的 RDAN-Large 模型获得了最高的

准确率 95.47%, 最高的精确率 84.73% 及最高的 Dice 相似系数 83.17%。相比于其他方法来说, 本研究提出的 RDAN 网络更加关注提升裂缝检测的精确度, 同样随着网络的加深, 提升效果越明显, 进而在 CRACK500 数据集上取得了最好的效果。

表 7 不同模型在 CRACK500 数据集上的识别效果指标

模型	准确率	精确率	召回率	Dice 系数
CrackForest	0.230 3	0.184 9	0.995 7	0.290 7
ConvNet	0.899 7	0.711 9	0.609 6	0.600 3
FPHBN	0.907 8	0.617 8	0.951 9	0.730 5
Pyramid Residual Network	0.942 9	0.807 7	0.801 4	0.772 1
本研究提出的 RDAN-Small	0.950 5	0.830 2	0.845 4	0.817 5
本研究提出的 RDAN-Medium	0.952 3	0.831 4	0.856 6	0.825 4
本研究提出的 RDAN-Large	0.954 7	0.847 3	0.851 1	0.831 7

表 8 不同数据集交叉训练、测试结果

Tab. 8 Cross-training and testing results of different dataset

训练集	测试集	准确率	精确率	召回率	Dice 系数
CiCS50000	CFD	0.932 9	0.945 9	0.700 7	0.779 3
CiCS50000	AigleRN	0.958 8	0.885 1	0.782 6	0.792 1
CiCS50000	CRACK500	0.859 7	0.990 8	0.061 5	0.084 5
CFD	CiCS50000	0.793 0	0.041 8	0.553 8	0.027 6
AigleRN	CiCS50000	0.916 0	0.198 0	0.519 4	0.064 6
CRACK500	CiCS50000	0.744 1	0.070 3	0.627 2	0.059 6
CiCS50000+CFD	CFD	0.968 7	0.930	0.913 3	0.916 8
CiCS50000+AigleRN	AigleRN	0.974 8	0.933 4	0.873 7	0.884 9
CiCS50000+CRACK500	CRACK500	0.956 2	0.851 5	0.850 7	0.829 1
CFD+CiCS50000	CiCS50000	0.972 3	0.834 9	0.849 5	0.789 5
AigleRN+CiCS50000	CiCS50000	0.972 2	0.829 8	0.841 1	0.779 6
CRACK500+CiCS50000	CiCS50000	0.973 0	0.847 8	0.837 1	0.788 4

单纯在 CiCS50000 裂缝精修数据集上训练的模型, 只有在 AigleRN 数据集上表现出不错的泛化效果, 而在 CFD 与 CRACK500 数据集上, 相比之前均表现出较低的召回率和 Dice 相似度; 反之, 对于在 CFD, AigleRN, CRACK500 上训练的模型, 在 CiCS50000 裂缝精修数据集上直接进行检测, 各项指标同样很低。这说明了本研究选取的各个数据集图像特征差异较大, 具有对比价值, 在一个数据集上训练, 直接在另一个数据集上进行测试, 难以取得很好的裂缝检测效果。

而在 CiCS50000 裂缝精修数据集上进行预训练后, 再在 3 个公开数据集上进行迭代微调训练, 均取得了非常优秀的测试结果, 特别是在 CFD 与 AigleRN 公开数据集上, 整体指标相比表 5 和表 6 中

3.1.5 交叉训练测试试验

不同数据集之间的路面背景和裂缝特征均有所不同, 为了更加全面地验证本研究所使用的实际工程数据集 CiCS50000 的有效性, 及该数据集上训练出模型的泛化能力, 基于 RDAN-Large 网络设计了不同数据集之间的交叉训练和验证试验, 包括: (1) 在 CiCS50000 裂缝精修数据集上训练, 在 3 个公开数据集 CFD, AigleRN, CRACK500 上测试; (2) 在 3 个公开数据集 CFD, AigleRN, CRACK500 上训练, 在 CiCS50000 裂缝精修数据集上测试; (3) 在 CiCS50000 裂缝精修数据集上预训练模型, 分别在 CFD, AigleRN, CRACK500 数据集上进行迭代训练和测试; (4) 在 CFD, AigleRN, CRACK500 数据集上预训练模型, 在 CiCS50000 裂缝精修数据集上进行迭代训练和测试。试验结果如表 8 所示。

单独训练的结果, 均取得了进一步的突破。而由于 3 个公开数据集与 CiCS50000 裂缝精修数据集相比图像数量较少, 因此在公开数据集上预训练后, 再在本研究的 CiCS50000 裂缝精修数据集上迭代微调, 提升效果并不明显。

3.2 生成对抗网络预处理试验

(1) DCGAN 预处理后对于不同模型的训练效果提升对比试验

本研究在 CiCS50000 数据集上设计对比试验。首先使用训练好的生成对抗网络的生成模型, 对 CiCS50000 数据集整体进行预处理, 然后使用几种效果最好的 CNN 进行训练。如表 9 所示, 本研究在使用 DCGAN 预处理后的 CiCS50000 数据集上, 分别使用 ResNet34, ResNet50, ResNet101 及本研究所提出

表9 使用 DCGAN 预处理前后裂缝识别效果对比

Tab.9 Comparison of crack recognition effect before and after DCGAN pretreatment

模型	无 DCGAN 预处理				DCGAN 预处理			
	准确率	精确率	召回率	Dice 系数	准确率	精确率	召回率	Dice 系数
ResNet34	0.971 0	0.824 8	0.807 6	0.748 3	0.970 6	0.866 5	0.796 8	0.768 3 (+0.020)
ResNet50	0.969 9	0.806 8	0.815 8	0.743 6	0.969 8	0.850 7	0.799 4	0.760 9 (+0.017 3)
ResNet101	0.970 3	0.823 6	0.809 1	0.750 0	0.970 3	0.846 1	0.809 7	0.763 6 (+0.013 6)
本研究提出的 RDAN-Large	0.973 5	0.868 1	0.827 1	0.793 0	0.975 1	0.870 5	0.833 6	0.810 2 (+0.017 2)

的 RDAN-Large 网络进行训练和测试。通过对比发现,使用 DCGAN 对试验数据进行预处理后,相比不使用预处理方法,对上述各个模型的在测试集上的 Dice 相似系数分别提升了 2.00%, 1.73%, 1.36%, 1.72%, 进而验证了本研究提出的基于生成对抗网络的路面裂缝预处理算法,在实际工程数据上能够带来有效的数据质量提升。

(2) DCGAN 与传统预处理方法对比试验

本研究在 CiCS50000 数据集上设计对比实验。使用不同的传统预处理和 DCGAN 预处理方法分别对 CiCS50000 数据集整体进行预处理,使用 RDAN-Large 网络模型分别进行预处理后训练和测试,探究哪种预处理方法对于裂缝识别网络的训练效果提升最大。

如表 10 所示,相比不使用任何预处理方法,传统预处理方法均可以对模型训练效果带来一定提升,其中中值滤波方法对于识别召回率的提升最为明显,达到 85.39%。本研究采用 DCGAN 预处理,裂缝识别模型的训练效果相比之前取得了进一步提升,在准确率、精确率和 Dice 相似系数上分别达到了最高的 97.51%, 87.05%, 81.02%, 这也验证了本研究所提出预处理算法的先进性和实际可用性。

表 10 各种预处理方法预处理后裂缝识别效果对比

Tab.10 Comparison of crack recognition effect after pretreatment with various pretreatment methods

预处理方法	准确率	精确率	召回率	Dice 系数
无预处理	0.973 5	0.868 1	0.827 1	0.793 0
均值滤波	0.974 9	0.858 8	0.848 5	0.805 8
中值滤波	0.974 3	0.854 3	0.853 9	0.809 0
高斯滤波	0.974 7	0.852 5	0.847 0	0.800 1
双边滤波	0.974 6	0.856 0	0.845 4	0.801 9
直方图均衡	0.973 6	0.853 3	0.839 0	0.794 2
自适应直方图均衡化	0.974 3	0.857 8	0.843 6	0.801 9
DCGAN	0.975 1	0.870 5	0.833 6	0.810 2

4 结论

本研究在实际工程化应用过程中,针对以浅色裂缝和模糊裂缝为代表的特殊裂缝识别准确率低的

问题,提出一种基于 DCGAN 预处理和 RDAN 网络的路面裂缝识别算法。

(1) 在图像预处理过程中,提出了一种基于生成对抗网络的路面裂缝图像预处理算法。通过设计 DCGAN 作为基本的生成模型和判别模型网络结构,实现特殊裂缝的增强和重建,识别效果优于传统图像预处理方法。

(2) 通过结合传统 ResNet 和 DenseNet 的基本结构并加以改进,引入 CBAM 注意力机制,分别提出了基于层级的基本连接单元 RDAL,基于模块级的基本连接单元 RDAB 和完整的裂缝识别网络 RDAN。通过模块化的基本结构设计,可以根据不同的算力水平和数据量大小,对 RDAN 的网络深度进行动态设计调整。

(3) 本研究提出的 RDAN 网络在公开数据集和实际工程数据集上,相较于其他网络,均能取得最优的路面裂缝识别效果。并且通过不同数据集之间的交叉训练、验证试验,证明了在本研究提出的 CiCS50000 数据集上预训练的模型,在其他数据集上迭代训练后表现出良好的泛化能力。

参考文献:

References:

- [1] 潘玉利. 路况自动检测原理 [M]. 北京: 人民交通出版社, 2021.
PAN Yu-li. Automated Pavement Condition Surveys [M]. Beijing: China Communications Press, 2021.
- [2] AGGARWAL S, LAHOTI M, SINGH A P, et al. Pothole Recognition Using DeepCNN with Layer Permutation Scheme [J]. Journal of Electronic Imaging, 2022, 31 (4): 043011. <https://doi.org/10.1117/1.JEI.31.4.043011>.
- [3] 李雯雯, 李喜媛, 周健, 等. 基于全卷积神经网络和朴素贝叶斯数据融合的桥梁裂缝识别算法 [J]. 公路交通科技, 2023, 40 (2): 44-52.
LI Wen-wen, LI Xi-yuan, ZHOU Jian, et al. An Algorithm for Recognizing Bridge Cracks Based on Full Convolution Neural Network and Naïve Bayesian Data

- Fusion [J]. *Journal of Highway and Transportation Research and Development*, 2023, 40 (2): 44-52.
- [4] 杭伯安. 基于改进 K-means 算法的沥青路面裂缝分割方法 [J]. *公路交通科技*, 2023, 40 (4): 1-8.
HANG Bo-an. A Method for Asphalt Pavement Crack Segmentation Based on Improved K-means Algorithm [J]. *Journal of Highway and Transportation Research and Development*, 2023, 40 (4): 1-8.
- [5] 高凡, 贺春光, 王鹏章, 等. 数字图像处理在混凝土路面裂缝中的应用 [J]. *中外公路*, 2023, 43 (1): 44-51.
GAO Fan, HE Chun-guang, WANG Peng-zhang, et al. Application of Digital Image Processing in Cracks of Concrete Pavement [J]. *Journal of China & Foreign Highway*, 2023, 43 (1): 44-51.
- [6] 刘建华, 董家修, 王念念, 等. 基于 Crack Mask R-CNN 模型的路面裂缝像素级分割及测量算法应用分析 [J]. *中外公路*, 2023, 43 (5): 47-52.
LIU Jian-hua, DONG Jia-xiu, WANG Nian-nian, et al. Analysis on Application of Pixel Level Segmentation and Measurement Algorithm Road Pavement Cracks Based on Crack Mask R-CNN Model [J]. *Journal of China & Foreign Highway*, 2023, 43 (5): 47-52.
- [7] 苏卫国, 王景霄. 基于 YOLO v3 深度学习算法的道路裂缝识别模型研究 [J]. *中外公路*, 2023, 43 (2): 58-63.
SU Wei-guo, WANG Jing-xiao. Research on Road Crack Recognition Model Based on YOLO v3 Deep Learning Algorithm [J]. *Journal of China & Foreign Highway*, 2023, 43 (2): 58-63.
- [8] PIZER S M, AMBURN E P, AUSTIN J D, et al. Adaptive Histogram Equalization and Its Variations [J]. *Computer Vision, Graphics, and Image Processing*, 1987, 39 (3): 355-368.
- [9] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 770-778.
- [10] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely Connected Convolutional Networks [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 4700-4708.
- [11] YANG F, ZHANG L, YU S J, et al. Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 21 (4): 1525-1535.
- [12] FENG H, XU G S, GUO Y H. Multi-scale Classification Network for Road Crack Detection [J]. *IET Intelligent Transport Systems*, 2019, 13 (2): 398-405.
- [13] NGUYEN H T, YU G H, NA S Y, et al. Pavement Crack Detection and Segmentation Based on Deep Neural Network [J]. *The Journal of Korean Institute of Information Technology*, 2019, 17 (9): 99-112.
- [14] TONG Z, YUAN D D, GAO J, et al. Pavement Defect Detection with Fully Convolutional Network and an Uncertainty Framework [J]. *Computer-aided Civil and Infrastructure Engineering*, 2020, 35 (8): 832-849.
- [15] HUYAN J, LI W, SUSAN T, et al. CrackU-net: A Novel Deep Convolutional Neural Network for Pixelwise Pavement Crack Detection [J]. *Structural Control & Health Monitoring*, 2020 (5): e2551. <https://doi.org/10.1002/stc.2551>.
- [16] 何宇超, 段中兴, 高静. 基于多尺度空洞卷积结构的路面裂缝分割方法 [J]. *公路交通科技*, 2024, 41 (1): 1-9, 17.
HE Yu-chao, DUAN Zhong-xing, GAO Jing. A Method for Pavement Crack Segmentation Based on Multi-scale Cavity Convolution Structure [J]. *Journal of Highway and Transportation Research and Development*, 2024, 41 (1): 1-9, 17.
- [17] MILLETARI F, NAVAB N, AHMADI S A. V-net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation [C] // *Proceedings of 2016 Fourth International Conference on 3D Vision (3DV)*. Stanford: IEEE, 2016: 565-571.
- [18] SHI Y, CUI L M, QI Z Q, et al. Automatic Road Crack Detection Using Random Structured Forests [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 17 (12): 3434-3445.
- [19] AMHAZ R, CHAMBON S, IDIER J, et al. Automatic Crack Detection on Two-dimensional Pavement Images: An Algorithm Based on Minimal Path Selection [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 17 (10): 2718-2729.
- [20] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional Networks for Biomedical Image Segmentation [C] // *Proceedings of International Conference on Medical Image Computing and Computer-assisted Intervention*. Cham: Springer, 2015: 234-241.
- [21] ZHANG L, YANG F, ZHANG Y D, et al. Road Crack Detection Using Deep Convolutional Neural Network [C] // *Proceedings of 2016 IEEE International Conference on Image Processing (ICIP)*. Phoenix: IEEE, 2016: 3708-3712.