光电技术及应用

DOI：10.16818/j.issn1001－5868.2023103002

基于改进DeepLabv3＋网络的沥青道路裂缝检测方法

陈昌川[1]，郝晓严[1]，龙虹毓[2]，孙　霞[3*]

（重庆邮电大学 1. 通信与信息工程学院，重庆 400065；2. 自动化学院，重庆 400065；3. 重庆工程学院，重庆 400056）

摘　要： 针对传统的语义分割技术对于沥青道路裂缝的检测存在检测精度低、误差大的问题，提出了一个基于改进DeepLabv3＋网络的语义分割方法。该方法在编码器阶段，采用轻量级MobileNetv2取代DeepLabv3＋的主干网络Xception，从而减少参数量；在解码器阶段，引入双注意力机制以进一步提高网络的分割精度；使用Dice Loss函数与原始交叉熵损失函数混合，以缓解样本中前景和背景不平衡问题。最后以道路实时检测的数据为对象进行了大量的实验，结果表明，该方法与原始DeepLabv3＋相比，平均交并比（mIoU）、平均像素精度（mPA）分别提升了8.98％和17.39％。与其他主流语义分割模型相比，改进后的DeepLabv3＋在沥青道路裂缝的检测上也取得了较好的效果。

关键词： 图像分割；DeepLabv3＋；注意力机制；特征金字塔

中图分类号： TP391.4　**文章编号：** 1001－5868(2024)03－0493－08

Asphalt Road Crack Detection Method Based on Improved DeepLabv3＋ Network

CHEN Changchuan[1], HAO Xiaoyan[1], LONG Hongyu[2], SUN Xia[3]

(1. School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, CHN; 2. School of Automation, Chongqing 400065, CHN; 3. Chongqing Institute of Engineering, Chongqing 400056, CHN)

Abstract: A semantic segmentation method based on an improved DeepLabv3＋network is proposed to address the issues of low detection accuracy and large errors associated with traditional semantic segmentation techniques for detecting asphalt road cracks. In the encoder stage, this method replaces the backbone network Xception of DeepLabv3＋ with lightweight MobileNetv2, thereby reducing the number of parameters. In the decoder stage, a dual attention mechanism is incorporated to further improve the segmentation accuracy of the network. The Dice loss function is combined with the original cross-entropy loss function to alleviate the imbalance between foreground and background in the sample. Extensive experiments were conducted on real-time road detection data. The results indicate that, compared with the original DeepLabv3＋, the average intersection-to-union ratio (mIoU) and average pixel accuracy (mPA) achieved by the proposed method were higher by 8.98％ and 17.39％, respectively. As compared with other mainstream semantic segmentation models, the improved DeepLabv3＋ also exhibits good performance in detecting asphalt road cracks.

Key words: image segmentation; DeepLabv3＋; attention mechanism; feature pyramid

0　引言

道路巡检是日常维护工作的重要组成部分，对中国经济稳定发展具有重要作用。由于道路裂缝情况复杂多变，人工巡检效率低且易受主观因素影响，已不能满足日趋增长的需求。2021年末，全国

收稿日期：2023－10－30.
基金项目：重庆市教委重点项目（KJZD-K202301901）；重庆市科委面上项目（2023NSCQ-MSX4308）.
* **通信作者：**孙　霞

公路总里程比上年末增加 8.26 万千米[1]，随着道路被频繁使用，难免会出现病害问题，例如不同类型的裂缝给交通发展带来了极大的安全隐患，因此开发一种精度高的道路裂缝检测算法具有十分重要的意义。

近年来，深度学习[2]在图像识别、目标分类和语义分割等领域都取得了进展。与传统的图像处理技术相比，深度学习也能够在人工神经网络的支持下，利用卷积运算自动学习关键特征，从而产生良好的结果[3]。随着深度学习技术[4]的迅猛发展，基于卷积神经网络[5]（Convolutional Neural Networks，CNN）的图像分割技术取得了很大的进步，CNN 也被愈来愈多的研究者们应用于图像分割领域中。2015 年，Long 等[6]率先设计了一个全卷积网络（Fully Convolutional Neural Networks，FCN），它可以接受任何大小的输入，并通过反卷积和合并操作产生相应大小的输出。受 FCN 的启发，人们提出了各种深度学习语义分割模型，如 PSPNet[7]，U-Net[8]，LeNet[9]，SegNet[10] 和 DeepLab 系列[11-14]。PSPNet 网络通过使用不同的池化操作来控制感受野，在提高检测速度上起到了一定的作用，但是并不能充分利用场景，其检测精度受到影响。U-Net 引入了跳跃式连接结构，在解码过程中将相应的底层特征连接到反卷积层，因此模型可以充分利用每一层的语义特征，即使在训练样本较少的情况下也能获得较好的分割结果。SegNet 采用 VGG-16 作为编码器层，每个编码器层都对应一个解码器层，最终解码器的输出被送入 soft-max 分类器以独立的为每个像素产生类概率。DeepLabv3＋属于 DeepLab 系列，它把空间金字塔（ASPP）模块和编码器-解码器架构组合在一起来捕捉多尺度特征，并获得更准确的对象边界。综上所述，深度学习在裂缝的检测上具有出色的性能。

以上模型都取得了较好的效果，但在准确率上还是存在一些不足，欠缺对复杂道路情况的全面考究。本文针对道路裂缝的准确率低的问题提出了一个改进的 DeepLabv3＋模型，该模型采取了更加轻量级的主干网络——MobileNetv2，该网络相比于原始的 Xception 网络具有更少的参数，使模型更加轻量化；为解决前景背景不平衡带来的消极影响，引入了 Dice Loss 函数，选择了交叉熵损失函数与 Dice Loss 函数加和的方式；最后在编码器阶段采取 DenseASPP 代替原始的 ASPP 模块，在解码器阶段使用 CARAFE 上采样和引入双注意力机制以进一步提高模型的分割精度。

1 研究方法

改进的 DeepLabv3＋的结构如图 1 所示，该结构由编码器和解码器组成。输入图片经过骨干网路后输出相应的深层特征和浅层特征，其中深层特征送入 DenseASPP 后再送入基于通道的 ECA 注意力模块中，最后进行 1×1 卷积运算，获得多尺度上下文信息；在解码阶段先对来自编码阶段的深层语义特征进行 CARAFE 上采样，之后再与送入 ECA 模块后的浅层特征融合，进行 3×3 卷积运算，4 倍上采样得到最终的分割结果。

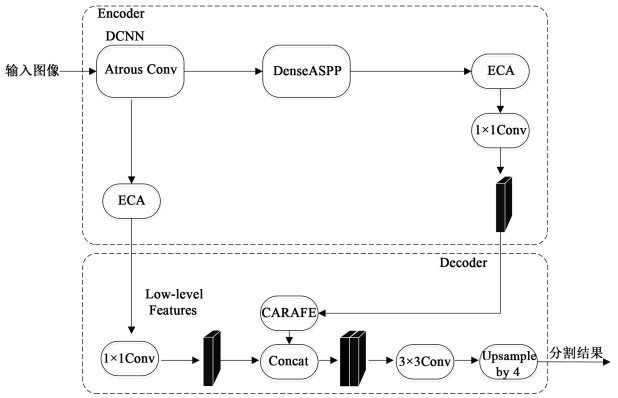


图 1 改进 DeepLabv3＋网络结构图

1.1 原始 DeepLabv3＋模型

原始 DeepLabv3＋模型的网络结构如图 2 所示。在 Encoder 结构中，首先通过骨干网络对输入图像进行特征提取，得到的特征图像分别传送到空洞空间金字塔池化金字塔和 Decoder 结构。ASPP 结构采用不同空洞率的空洞卷积层获得输入特征图

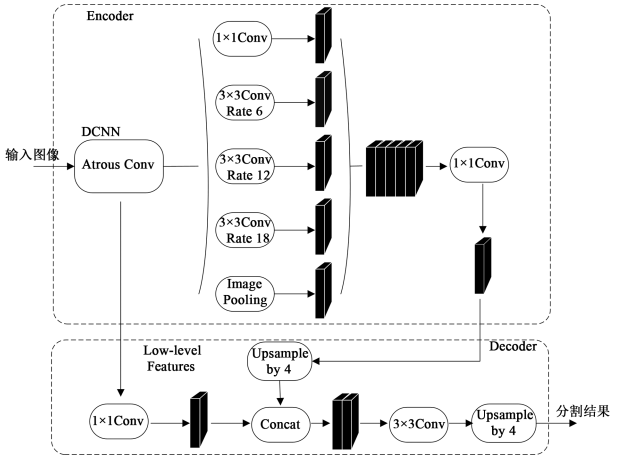


图 2 原始 DeepLabv3＋网络结构图[15]

的多尺度信息并拼接融合，得到图像的高层语义特征，然后使用 1×1 的卷积调整输出通道数。在 Decoder 结构中，对 Encoder 结构输出的高层语义特征进行 4 倍上采样，将分辨率恢复到原图的 1/4，并对 Backbone 网络输出的低层特征经过 1×1 卷积调整通道数，与高层语义特征图进行特征融合。之后对融合特征图进行 3×3 卷积运算和 4 倍上采样得到预测结果[15]。

1.2　轻量化网络 MobileNetv2

MobileNetv2 与 DeepLabv3＋原始的主干网络 Xception 相比，具有更精简的网络结构，是一个轻量级的网络模型，以参数量少、预测精度高而被人们知晓且使用，两者的共同点在于关键步骤都使用了深度可分离卷积，不同点在于 Xception 在使用深度可分离卷积的同时加大了参数量，而 MobileNet 使用深度可分离卷积是为了压缩模型和提高检测速度，使模型更加轻量化[16]。

MobileNetv2 的基本结构由三部分组成，深度可分离卷积、倒残差和线性瓶颈结构。其中最具有亮点的是反向残差结构，且在反向残差结构中，最后一层的卷积层，采用了线性的激活函数，而不是 ReLU 激活函数。其网络结构如图 3 所示。

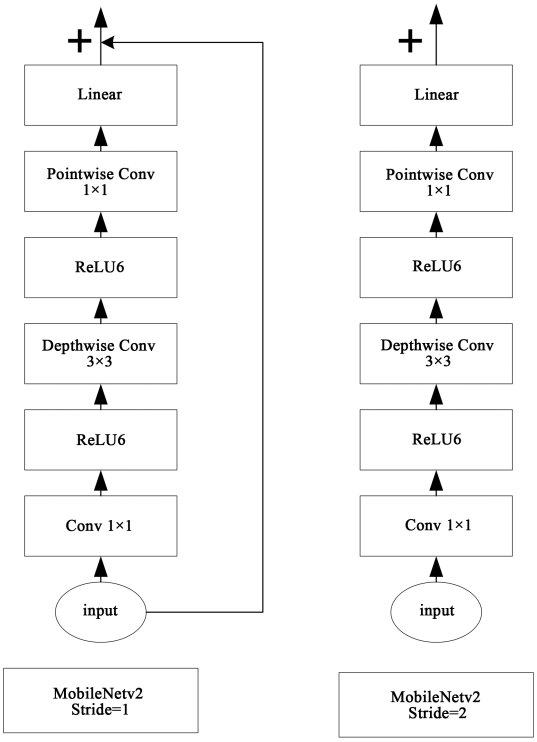


图 3　MobileNetv2 网络结构图[16]

1.3　ECA 注意力机制模块

ECA(Efficient Channel Attention Module)是 2020 年由 Wang 等[17]提出的一种高效通道注意力模块，该模块只设计少数几个参数，但是具有明显的效果增益，适当的跨通道交互可以在保持性能的同时显著降低模型的复杂性，不降维的局部跨通道交互策略有效避免了降维对于通道注意力学习效果的影响。

其整体结构如图 4 所示，针对特征图进行操作，通过通道注意力模块提取到更加精炼的特征，从而提升了模型的表现力。

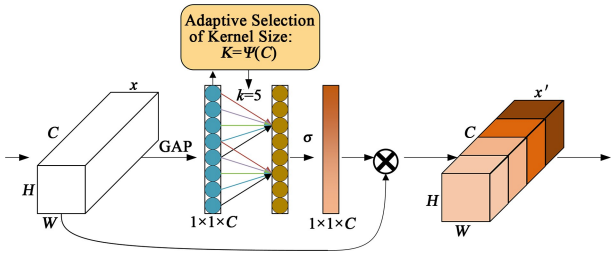


图 4　ECA 网络结构图[17]

1.4　CARAFE 上采样

最广泛使用的特征上采样算子是最近邻和双线性插值，它们采用像素之间的空间距离来指导上采样过程。然而，最近邻插值和双线性插值只考虑亚像素邻域，无法捕获密集预测任务所需的丰富语义信息。本文使用的 CARAFE 上采样是一种通用的、轻量级的、高效的算子，它能够在较大的接受域内聚合上下文信息，动态适应特定于实例的内容且能够保持计算效率[18]，网络结构如图 5 所示。CARAFE 主要由上采样核预测模块和特征重组模块组成。在每个位置，CARAFE 可以利用底层内容信息来预测重组核并且在预定义的附近区域内重组特征。

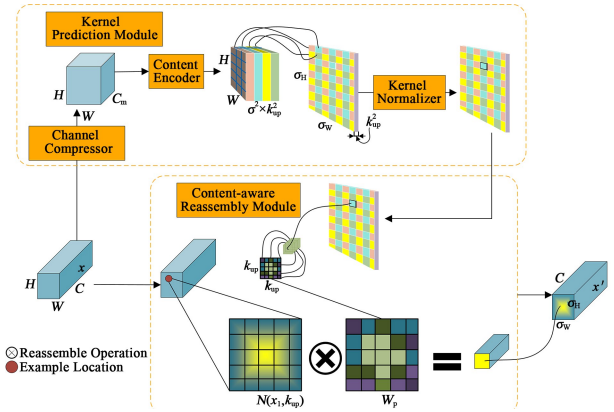


图 5　CARAFE 上采样网络结构图[18]

1.5　DenseASPP 模块

DenseASPP 的结构如图 6 所示。atrous 卷积

层以级联方式组成，其中每层膨胀率逐层增加。膨胀率小的层放在下部、大的层放在上部。每个atrous层的输出与输入特征图和所有来自较低层的输出进行级联，并将级联的特征图馈送到下一层。DenseASPP的最终输出是由多速率、多尺度atrous卷积生成的特征图，所提出的结构可以同时组成一个更密集和更大的特征金字塔，只使用几个atrous卷积层。DenseASPP中的各atrous层可表示为：

$$y_1 = H_{K,d_l}([y_{l-1}, y_{l-2}, \cdots, y_0]) \tag{1}$$

式中，d_l 表示 l 层的膨胀率，[…]表示拼接操作，$[y_{l-1}, y_{l-2}, \cdots, y_0]$表示将之前所有层的输出串联起来形成的特征图。与原来的结构相比，DenseASPP将所有膨胀层堆叠在一起，并通过密集连接将它们连接起来。这种变化给我们带来了两个好处：更密集的特征金字塔和更大的接受域[19]。

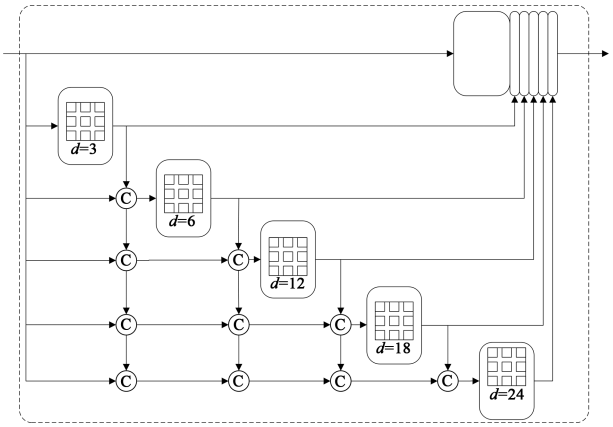


图6 DenseASPP网络结构图[19]

1.6 损失函数

传统的损失函数，如交叉熵、均方等，在数据类别不平衡的情况下性能较差。比例较小的类别会被比例较大的类别所淹没，导致训练时预测准确率较高，但应用效率较差。由于道路大部分都是完好无损的，只有很少的部分是有裂缝的，这些很少的部分大都是由于大型公共交通工具频繁挤压造成的，这就造成了一个问题，即前景背景不平衡，也就是图片中大部分区域都是不包含目标的，只有一小部分区域包含目标。然而Dice Loss函数可以有效缓解前景背景不平衡带来的消极影响。因此，我们在改进的DeepLabv3+模型中引入了Dice Loss函数。

在二元分类中，交叉熵常用于评价标签与预测之间的差异。计算公式如式(2)所示，式中 y 表示标签，y'表示预测，N 表示像素数。

$$L_{ce} = \frac{1}{N}\sum_i -[y\log_2 y' + (1-y)\log(1-y')] \tag{2}$$

Dice Loss是由Dice系数而得名的，Dice系数是一种用于评估两个样本相似性的度量函数，其值越大意味着这两个样本越相似，Dice系数的计算公式如式(3)所示。

$$Dice = \frac{2|X \cap Y|}{|X|+|Y|} \tag{3}$$

式中，X 代表裂缝的真实情况，Y 代表预测的情况。$|X \cap Y|$表示 X 和 Y 之间的前景相交像素数，$|X|$和$|Y|$分别表示标签和预测中前景(裂缝)在特定图像中的像素数。Dice Loss表达式如式(4)所示。

$$I_{\text{Dice}} = 1 - Dice = 1 - \frac{2|X \cap Y|}{|X|+|Y|} \tag{4}$$

如式(5)所示，最终损失函数为交叉熵函数与Dice损失函数的加和。

$$L = L_{ce} + L_{\text{Dice}} \tag{5}$$

2 实验与分析

2.1 参数设置和评价指标

本文的实验环境和网络参数设置如表1所示。

表1 实验环境和训练参数

名称	参数
操作系统	Windows 10
处理器	12th Gen Intel(R) Core(TM) i5-12400
显卡	NVIDIA GeForce RTX 3060 Ti
RAM/GB	8
网络框架	pytorch
Batchsize	8
迭代次数	54 950

实验采用cos的学习衰减策略，并使用Adam优化器，优化器内部使用的momentum参数为0.9，并将扩充后的数据集以9∶1的比例划分训练集和验证集。实验设置的初始学习率为0.000 5，最小学习率为 5×10^{-6}，使用Dice Loss函数进行实验，每5轮进行一次模型评估并保存最优权重，以平均交并比(mIoU)和平均像素准确率(mPA)作为评价指标。PA表示预测类别正确的像素总量除以像素总数，而mPA则是对PA的改进，它是先对每个类计算PA，然后再对所有类的PA求平均，其计算公式如式(6)所示。

$$mPA = \frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{j=0}^{k}p_{ij}} \tag{6}$$

mIoU是分别对每个类计算(真实标签和预测结果的交并比)IoU，然后再对所有类别的IoU求均值，其计算公式如式(7)所示。

$$mIoU=\frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{j=0}^{k}p_{ij}+\sum_{j=0}^{k}p_{ji}-p_{ii}} \tag{7}$$

2.2　数据集

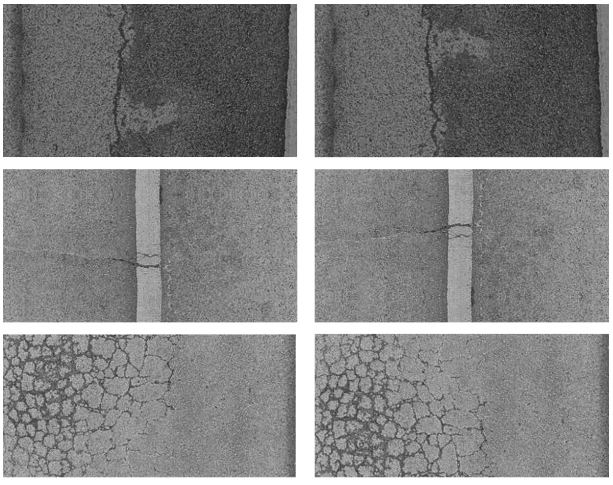
本文所使用的数据集都是通过公共道路实拍的图片，拍摄方式是两个摄像头并行排列并且固定在拍摄使用的交通工具尾部上方，交通工具匀速前进并每隔 10 米拍摄一次，整体拍摄装置如图 7(a)、(b)所示。



(a) 拍摄设备头部　　(b) 拍摄设备尾部　　(c) 软件界面

图 7　整体设备

由于摄像头拍摄的原始图片亮度很暗，且是由两个摄像头进行拍摄的，所以在此之后需要进行图片的拼接和增强操作，如图 7(c)所示，拼接后的图片大小为 3 846×1 000。由于数据集原始图片只有 1 257 张图片，所以送入模型训练之前需要进行数据集扩充，扩充的方式主要是上下翻转、添加公共数据集和增加负样本。考虑到输入图片的大小对模型的识别精度会有影响，本文所使用的数据集都是经过剪裁后再送入模型训练的，剪裁后的图片大小为 1 923×1 000。扩充和剪裁后的数据集的总数是 8 789 张图片，正负样本的比例为 1∶2。送入模型的输入图片大小为 512×512，对图像进行分割时分为裂缝和背景两类，每张图片都有对应的人工标注图。扩充后的图片如图 8 所示。



(a) 原图　　(b) 上下翻转

图 8　部分数据集扩充结果

2.3　结果与分析

2.3.1　不同损失函数的对比

损失函数反映了真实值与预测值的差值，本文分别利用不同的损失函数进行对比实验，实验结果如表 2 所示。

表 2　不同损失函数对分割精度的影响

损失函数	*mIoU*/%	*mPA*/%	准确率/%
L_{ce}	66.43	71.65	94.32
$L_{ce}+L_{Dice}$	68.32	82.43	96.55
$L_{ce}+L_{Dice}+L_{Focal}$	68.15	82.87	96.45

由表 3 可知，只采用原始交叉熵损失函数模型的均交并比只能达到 66.43%，而采用交叉熵、Dice Loss 和 Focal Loss 三者加和的方式则能达到 68.15%。经过实验证明，采用交叉熵和 Dice Loss 两者加和的方式对模型的识别效果起着最大的提升作用，平均交并比达到了 68.32%，故采取了交叉熵和 Dice Loss 函数混合的方式。

2.3.2　不同改进策略对模型的影响

改进的 DeepLabv3＋网络训练结果如图 9 所示。

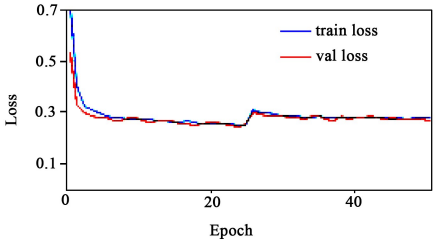


图 9　模型训练结果

由图 9 可知，训练开始后，损失函数值出现急剧下降的情况，特别是在前 10 个 Epoch 时效果最为明显，在第 10～20 个 Epoch 之间时损失函数值下降的比较缓慢，但是验证集在第 24 个 Epoch 时下降到了最低点为 0.23，之后又急剧上升到了 0.32，大约在第 25 个 Epoch 后训练集和测试集损失值逐渐缓慢下降，最后在第 50 个 Epoch 时趋于稳定，达到了 0.27。因此，随着训练次数的增加，验证集和

训练集的损失值都下降至了 0.27 左右，说明该网络取得了比较不错的损失函数收敛性。

为了验证本文的改进策略对道路裂缝的检测有一定的优化效果，设计了如下的消融实验。在原始的 DeepLabv3＋网络进行主干网络的替换，在编码器使用 DenseASPP 替换原始的 ASPP 网络，在解码器阶段使用 CARAFE 上采样，在编码和解码阶段分别引入基于通道的注意力机制，将以上几种改进方式分别与原始 DeepLabv3＋网络进行对比实验，并分析平均交并比、平均像素准确率。

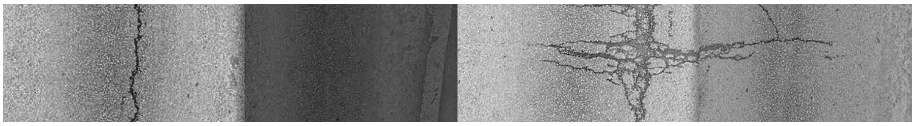
表 3 特征提取网络和注意力机制的消融实验结果

DeepLabv3＋	特征提取网络	优化器	DenseASPP	CARAFE	ECA	损失函数	*mIoU*/%	*mPA*/%
√	Xception	Sgd	——	——	——	L_{ce}	62.73	67.93
√	MobileNetv2	Sgd	——	——	——	L_{ce}	65.31	71.93
√	MobileNetv2	Sgd	——	——	——	$L_{ce}+L_{Dice}$	68.32	82.43
√	MobileNetv2	Adam	——	——	——	$L_{ce}+L_{Dice}$	70.46	83.71
√	MobileNetv2	Adam	√	——	——	$L_{ce}+L_{Dice}$	70.64	84.92
√	MobileNetv2	Adam	√	√	——	$L_{ce}+L_{Dice}$	70.55	86.52
√	MobileNetv2	Adam	√	√	√	$L_{ce}+L_{Dice}$	71.71	85.32

由表 3 第一二行分析可知，用 MobileNet 替换原始主干网络 Xception 后，平均交并比提高了 2.58%；由表 3 第二三行可知，使用 Dice Loss 函数与交叉熵损失函数进行加和的方式，平均交并比提高了 3.01%；通过使用不同的优化器进行对比实验，选择性能最优者，结果表明，Adam 优化器效果优于 Sgd 优化器，*mIoU* 提升了 1.73 个百分点；由表第四五行可知，使用 DenseASPP 结构后，模型的平均交并比分别提升了 0.18%，0.21%，在第五行的基础上采用 CARAFE 上采用代替原始双线性插值上采样后*mIoU*虽然降低了 0.09%，但是*mPA*却提升了 1.6%；由表最后一行可知，在编码器和解码器分别引入 ECA 注意力机制后，*mIoU* 提升了 1.16%，将以上几种方式结合后改进效果最佳，*mIoU* 和 *mPA* 分别提升了 8.98%，17.39%，进一步证明了本文方法的有效性与可行性。

2.3.3 与其他模型的分割效果对比

为了进一步验证本文改进算法与其他语义分割算法的有效性，分别使用了 PSPNet，U-Net，ResNet101，DeepLabv3＋等语义分割网络来进行对比实验，各种网络的预测结果如图 10 所示。



(a)原图



(b) 标记图像



(c) PSPNet



(d) U-Net



(e) ResNet101



(f) DeepLabv3＋



(g) 本文模型

图 10　改进 DeepLabv3＋与其他模型预测效果对比

由图 10 可知，PSPNet 几乎能够大致识别出裂缝的基本轮廓，但是在细节上就不能很精确地分割裂缝；而图 10(d)的 U-Net 网络不仅能够把裂缝的轮廓准确地分割出来，就连里面的很小的裂缝也可以很好地识别出来，效果相比于 PSPNet 要好很多；对于 Resnet101 和原始 DeepLabv3＋而言，裂缝的识别效果都不是很好，而本文模型不仅能够很精确地识别出不同类型的裂缝，而且还能在此基础上进一步提升，把较小的不易识别出的裂缝也精确地识别到了，这一点相比于其他几个模型都是最好的，并且平均交并比也是最高的。

改进 DeepLabv3＋模型与 PSPNet，U-Net，ResNet，DeepLabv3＋预测速度以及参数量进行对比，详细结果如表 4 所示。

表 4　改进 DeepLabv3＋与其他网络实验结果对比

方法	*mIoU*/%	*mPA*/%	预测速度/(f·s^{-1})	参数量/MB
PSPNet	61.71	73.53	59.85	177.54
ResNet	57.47	81.8	43.65	45.3
U-Net	68.43	81.75	36.87	94.54
DeepLabv3＋	62.73	67.93	33.53	208.64
DeepLabv3＋(MobileNetv2)	65.31	71.93	52.76	26.54
改进 DeepLabv3＋	71.71	85.32	68.04	24.87

由表 4 可知，DeepLabv3＋网络对沥青道路的分割精度优于前 3 组对比模型，但是模型的参数量很大，超过了 200MB，预测速度还需提升。用 MobileNetv2 替换原始 DeepLabv3＋骨干网络后，参数量明显减少且预测速度得到很大提升，达到了 52.76 f/s，参数量仅为 26.54 MB。本文提出的改进 DeepLabv3＋相较于其他模型在平均交并比和平均像素精度上均提升了 8.58%，9.93%，这证明本文算法更有利于道路裂缝的检测。

3　结束语

该文针对道路裂缝检测精度不高的问题，提出了基于改进 DeepLabv3＋网络的道路裂缝检测方法，该方法在原始 DeepLabv3＋网络的基础上，替换主干网络 Xception，对优化器和损失函数进行改进，在编码器阶段引入 DenseASPP，在解码器采用 CARAFE 上采样代替原始双线性插值上采样，并在编码器和解码器分别引入 ECA 注意力机制，以进一步提高网络的分割精度，有效改善了针对道路裂缝识别精度较低的问题，最后通过消融实验验证了所提方法的合理性和有效性。实验结果表明该方法相比于原始算法，平均交并比和平均像素精度分别提升了 8.98%，17.39%，在道路裂缝检测方面有较好的鲁棒性和泛化能力。

参考文献：

[1] 商讯. 交通运输部《公报》：2021 年末我国公路总里程 528.07 万千米，公路营运汽车 1 231.96 万辆，牵引车 346.68 万辆，挂车 359.25 万辆[J]. 商用汽车，2022(6)：5-6.
Business News. Ministry of Transport Bulletin: As of the end of 2021, China's total road mileage was 5. 280 7 million kilometers, with 12.319 6 million vehicles in operation, 3.466 8 million tractors, and 3.592 5 million trailers[J]. Commercial Vehicles, 2022(6): 5-6.

[2] 邬　硕，汪海涛，姜　瑛，等. 基于混合神经网络的脑部 MRI 图像语义分割算法[J]. 重庆邮电大学学报(自然科学版)，2022，34(3)：423-432.
Wu Shuo, Wang Haitao, Jiang Ying, et al. A semantic segmentation algorithm for brain MRI images based on hybrid neural networks[J]. J. of Chongqing University of Posts and Telecommunications: Natural Science Edition, 2022, 34(3): 423-432.

[3] Al-Bayati A, Ali M, Nnaji C. Managing work zone safety during road maintenance and construction activities: Challenges and opportunities [J]. Practice Periodical on Structural Design and Construction, 2023, 28(1): 04022068.

[4] 肖　佳，杨　微，张志威. 基于图像识别技术的机器人移动中障碍物定位研究[J]. 自动化与仪表，2022，37(12)：34-38.
Xiao Jia, Yang Wei, Zhang Zhiwei. Research on obstacle

localization in robot movement based on image recognition technology[J]. Automation and Instrumentation, 2022, 37(12): 34-38.

[5] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[6] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]// Proc. of 2015 IEEE Conf. on Computer Vision and Pattern Recognition, 2015: 3431-3440.

[7] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network [C]// Proc. of 2017 IEEE Conf. on Computer Vision and Pattern Recognition, 2017: 6230-6239.

[8] Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation [C]// 2015 Inter. Conf. on Medical Image Computing and Computer-Assisted Intervention(MICCAI), 2015, 9351: 234-241.

[9] LeCun Y, Bottou L, Bengio B Y, et al. Gradient-based learning applied to document recognition[J]. Proc. of the IEEE, 1998, 86(11): 2278-2324.

[10] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.

[11] 秦 汉, 熊 凌, 肖林伟, 等. 基于改进 DeepLabv3+网络的氩花图像语义分割[J]. 武汉科技大学学报, 2023, 46(1): 25-32.
Qin Han, Xiong Ling, Xiao Linwei, et al. Argon flower image semantic segmentation based on improved DeepLabv3 + network[J]. J. of Wuhan University of Science and Technology, 2023, 46(1): 25-32.

[12] 徐 聪, 王 丽. 基于改进 DeepLabv3+网络的图像语义分割方法[J]. 激光与光电子学进展, 2021, 58(16): 225-232.
Xu Cong, Wang Li. Image semantic segmentation method based on improved DeepLabv3+ network[J]. Progress in Laser and Optoelectronics, 2021, 58(16): 225-232.

[13] 董亚盼, 高陈强, 谌 放, 等. 基于注意力机制的红外小目标检测方法[J]. 重庆邮电大学学报(自然科学版), 2023, 35(2): 219-226.
Dong Yapan, Gao Chenqiang, Chen Fang, et al. An infrared small target detection method based on attention mechanism [J]. J. of Chongqing University of Posts and Telecommunications: Natural Science Edition, 2023, 35(2): 219-226.

[14] 左安全, 秦伦明, 王 悉, 等. 基于改进 DeepLabv3+模型的电力线语义分割方法[J]. 无线电工程, 2023, 53(1): 96-104.
Zuo Anquan, Qin Lunming, Wang Xi, et al. A semantic segmentation method for power lines based on an improved DeepLabv3+ model[J]. Radio Engineering, 2023, 53(1): 96-104.

[15] Chen L, Zhu C, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]// The 15th European Conf. on Computer Vision (ECCV), 2018: 833-851.

[16] Sandler M, Howard A, Zhu M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]// Proc. of 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition, 2018: 4510-4520.

[17] Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]// Proc. of 2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition, 2020: 11531-11539.

[18] Wang J, Chen K, Xu R, et al. CARAFE: Content-aware reassembly of features[C]// Proc. of 2019 IEEE/CVF Inter. Conf. on Computer Vision, 2019: 3007-3016.

[19] Yang M, Yu K, Zhang C, et al. DenseASPP for semantic segmentation in street scenes[C]// Proc. of 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition, 2018: 3684-3692.

作者简介:

陈昌川(1978—),男,四川省广安市人,副教授,研究领域为智能信息处理、图像处理与人工智能、特高频局放检测;

孙 霞(1976—),女,河南省周口市人,副教授,研究领域为无线移动通信技术、算法研究和数字信号处理。