

基于改进 DeepLabv3+的道路积水检测^①



贾军营, 吴兴宇, 杨海波

(沈阳工业大学 信息科学与工程学院, 沈阳 110870)

通信作者: 吴兴宇, E-mail: 1394316938@qq.com

摘要: 近年来, 随着城市化进程的加快, 城市排水系统在面对极端天气时常难以应对, 道路积水问题频繁发生. 为了解决道路积水的检测问题, 本文基于 DeepLabv3+模型提出改进算法. 首先, 在解码器端设计加权双向特征金字塔网络 (bidirectional feature pyramid network, BiFPN) 模块, 利用主干网络获取的不同尺度低层特征映射进行融合, 充分发挥从骨干网络获取的多尺度信息的潜力. 其次, 利用 Mamba 改进 Transformer 模块设计并行分支对高级特征映射进行处理, 构建全局依赖, 弥补 ASPP 中空洞卷积可能造成的局部信息丢失问题. 最后, 引入极化自注意力机制 (polarized self-attention, PSA) 模块, 减少双分支输出直接相加对于数据可能带来不同的影响. 实验结果表明, 在道路积水数据集上, 改进算法 $mIoU$ 为 87.54%, PA 为 96.61%, 与原算法相比, $mIoU$ 提高了 4.22%, PA 提高了 1.66%.

关键词: DeepLabv3+; 多尺度特征融合; 极化自注意力; Mamba; Transformer; 语义分割

引用格式: 贾军营, 吴兴宇, 杨海波. 基于改进 DeepLabv3+的道路积水检测. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/9799.html>

Road Water Accumulation Detection Based on Improved DeepLabv3+

JIA Jun-Ying, WU Xing-Yu, YANG Hai-Bo

(School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China)

Abstract: In recent years, with the acceleration of urbanization, urban drainage systems often struggle to cope with extreme weather, and road waterlogging occurs frequently. To solve the road waterlogging detection problem, this paper proposes an improved algorithm based on the DeepLabv3+ model. Firstly, a weighted bidirectional feature pyramid network (BiFPN) module is designed at the decoder side, which utilizes the different scales of low-level feature mapping obtained from the backbone network for fusion, giving full play to the potential of the multi-scale information obtained from the backbone network. Secondly, the Mamba-improved Transformer module is utilized to design parallel branches to process high-level feature mappings, construct global dependencies, and compensate for the possible local information loss caused by dilated convolution in ASPP. Finally, the polarized self-attention (PSA) module is introduced to mitigate the possible different effects of the direct addition of two-branch outputs on the data. The experimental results show that on the road waterlogging dataset, the improved algorithm has an $mIoU$ of 87.54% and a PA of 96.61%, which is an improvement of 4.22% in terms of $mIoU$ and 1.66% in terms of PA compared with the original algorithm.

Key words: DeepLabv3+; multi-scale feature fusion; polarized self-attention (PSA); Mamba; Transformer; semantic segmentation

随着城市化进程的加快, 城市排水系统在面对极端天气时常常难以应对, 从而导致道路积水问题. 道路

积水不仅影响交通通行效率, 还可能造成严重的交通事故, 威胁行人和车辆的安全^[1,2]. 此外, 长期的积水问

^① 基金项目: 2022 年度沈阳市科学技术计划“揭榜挂帅”产业共性技术项目 (22-316-1-07); 辽宁省应用基础研究项目 (2022JH2/101300243)

收稿时间: 2024-09-18; 修改时间: 2024-10-10; 采用时间: 2024-10-14; csa 在线出版时间: 2025-02-18

题还会对道路基础设施造成损害,增加维护成本^[3].因此,及时准确地检测和处理道路积水至关重要.目前,道路积水监测主要采用以下几种方法:人工观测、仪器监测以及结合移动互联网技术和传感器技术建立城市积水监测预警平台等.其中,人工观测方法虽然广泛使用,但需要大量的人力资源,效率较低.监测预警平台具有更高的准确性,但设备成本高,需要专人维护.

而随着深度学习技术的发展^[4-6],基于深度学习的检测算法处理道路状况成为新的研究热点.Sazara等^[7]使用预训练的深度学习神经网络(如卷积神经网络CNN)模型,通过迁移学习的方法将这些模型应用于道路积水检测.Witherow等^[8]采用了一种基于图像处理和机器学习的方法,通过分析来自用户上传的道路图像,自动识别积水区域.Huang等^[9]使用卷积神经网络(Mask R-CNN)检测积水中的轮胎,通过勾股定理计算积水深度.Jafari等^[10]使用安装在监测区域的实时摄像头捕捉水体的实时视频图像,利用计算机视觉算法处理摄像头图像,自动识别水体的边界,并通过与已知参考点的比较来计算实时水位高度.Shu等^[11]基于神经网络提出多地理单元卷积神经网络(MGCNN)模型,用于评估城区内涝风险.Liu等^[12]改进YOLOv5算法对道路交通内涝进行检测.白岗岗等^[13]提出了一种基于深度学习的智能监测方法,采用卷积神经网络(CNN)模型,结合视频监控数据,自动识别道路积水的深度和范围,分析水体的特征变化.胡昊等^[14]提出了一种基于视频影像的道路积水实时监测模型,采用YOLOv5s进行实时图像识别,计算积水区域面.廖宇鸿等^[15]提出一种利用深度学习技术和椭圆检测算法的城市道路积水深度监测方法,通过深度学习模型对不同类型的车辆的车轮进行检测和分割,利用椭圆检测算法对淹没车轮的几何特征参数进行提取,从而构建道路积水深度计算模型.张峥等^[16]提出一种基于改进YOLOv8的道路积水实时监测方法,在YOLOv8的颈部结构网络加入了注意力机制,提高识别道路积水的准确率,并利用透视变换和像素来计算积水面积.以上方法都取得了不错的结果,但仍然存在一些不足之处.为此,本文基于DeepLabv3+^[17]算法,提出了针对道路积水的改进算法,该算法主要有以下3方面的工作.

(1) 设计BiFPN模块,促进低层特征映射间信息交互,使得网络能够在不同层级上进行特征的共享和利用.充分利用多尺度信息,发挥骨干网络获取的多尺度

信息的潜力,缓解边界不连续问题.

(2) 基于Mamba模块改进Transformer编码器结构设计并行分支对高级特征映射进行处理,构建全局依赖关系,弥补ASPP中空洞卷积可能造成的局部信息丢失问题.

(3) 为防止双分支输出直接相加对于数据可能带来不同的影响,引入极化自注意力机制(polarized self-attention, PSA)模块,避免简单相加带来的信息丢失或无效信息累积.

1 本文方法

1.1 本文算法结构

本文算法以DeepLabv3+模型为主体进行改进.使用ResNet50替换MobileNetV2作为特征提取主干网络.并设计双向加权特征金字塔网络、并行双分支结构,引入极化自注意力机制模块.改进后的模型如图1所示.

在编码器阶段,利用ResNet50网络提取4个不同尺度的特征映射,分别为layer 1、layer 2、layer 3、layer 4,前3个不同尺度的特征映射传入加权双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)模块,进行有效的特征融合,以充分利用骨干网络提取的不同特征.

同时,将高层特征映射layer 4分别通过空间卷积池化金字塔模块(atrous spatial pyramid pooling, ASPP)和改进的Transformer模块,通过ASPP中空洞率卷积进一步提取特征,并利用Transformer编码器模块对layer 4构建全局依赖关系,将两者输出相加,弥补ASPP中空洞卷积可能造成的信息遗漏问题.为防止双分支输出直接相加对于数据可能带来不同的影响,通过极化自注意力机制(PSA)模块,减少直接相加带来的负面影响.并将获取的高级特征映射直接4倍上采样以匹配BiFPN模块输出的特征映射的大小,并且将它们通道维度上堆叠.然后应用两个深度可分卷积来获得最终有效特征图.最后,将特征映射上采样到输入图像的大小.

1.2 双向加权特征金字塔

在提取图像特征信息时,DeepLabv3+解码器部分采用简单的方法将低级特征映射与高级特征进行简单结合,以实现多尺度特征融合.然而,这种方式可能无法充分利用多尺度信息,导致无法充分发挥从骨干网

络获取的多尺度信息的潜力.为了解决这个问题,受到 EfficientDet^[18]模型启发,本文根据参考 BiFPN 模块设计 了多尺度融合模块作为处理浅层特征的组件, BiFPN 是一种加权的 双向特征金字塔网络,由自上而下传播 高层语义信息的路径和自下而上传递低层位置信息的路径组成.它考虑了不同分辨率特征图的贡献,通过引入可学习权重 $\omega_1、\omega_2、\dots、\omega_n$ 来确定不同输入特征的重要性.为了避免在特征融合过程中某些特征图的

权重过大或者过小, BiFPN 对权重 $\omega_1、\omega_2、\dots、\omega_n$ 进行归一化处理,通常类似于 Softmax 方法,将权重被 缩放到[0, 1]的范围内,如式 (1) 所示:

$$\omega'_i = \frac{\text{ReLU}(\omega_i)}{\varepsilon + \text{ReLU}(\omega_1) + \dots + \text{ReLU}(\omega_n)} \quad (1)$$

其中, ω_i 是一个可学习的权重,在每个 ω_i 之后应用 ReLU 激活函数,以确保权重为正值,用 $\varepsilon=0.0001$ 的小值 来防止数值不稳定.

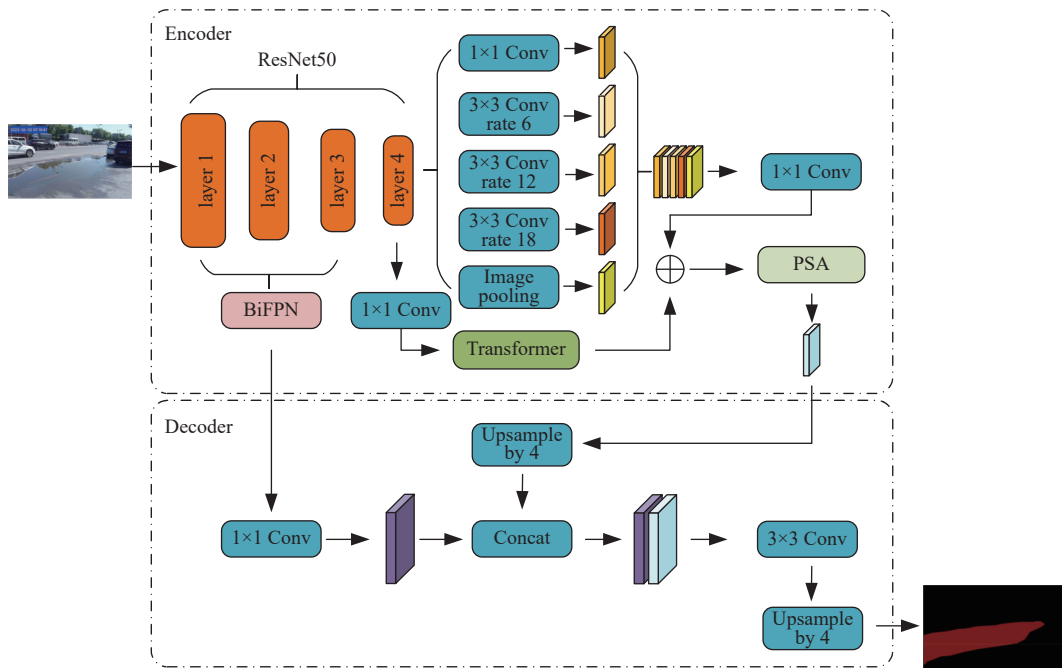


图1 本文算法结构图

本文设计的 BiFPN 结构如图 2 所示,首先通过 Res-Net50 骨干网络获得不同度的低层特征映射 $\text{layer } 1 \in R^{C \times H \times W}$ 、 $\text{layer } 2 \in R^{C/2 \times H/2 \times W/2}$ 、 $\text{layer } 3 \in R^{C/4 \times H/4 \times W/4}$ 、 $\text{layer } 4 \in R^{C/8 \times H/8 \times W/8}$,选取前 3 个低层特征映射,假设为 $P1-P3$.以 $P2$ 和 $P3$ 为例, $P2_mid$ 和 $P2_out$ 的表达式如式 (2) 和式 (3) 所示.由图 2 可知, $P2_mid$ 由 $P2$ 、 $P3$ 融合而成,首先,使用卷积大小为 3, padding 为 1 的卷积进一步提取 $P3$ 的特征信息和改变通道数,将其通道数改为和 $P2$ 通道数相一致,同时,采用双线性插值法将 $P3$ 特征映射上采样和 $P2$ 大小相同.最后,利用式 (1) 生成可学习权重 $\omega'_1、\omega'_2$,决定是更关注 $P2$ 还是 $P3$.而 $P2_out$ 则是由 $P2、P2_mid、P1_out$ 这 3 个融合而成,与 $P2_mid$ 过程大致相同,唯一区别就是 $P1_out$ 特征映射采用具有相同填充的二维最大池化操

作下采样到和 $P2_out$ 大小相同.在权重机制选择后,其他特征进行类似的叠加和权重选择操作,形成一个 BiFPN 模块.根据需求,可经过多次重复操作获得多级 BiFPN 模块.

$$P2_mid = \omega'_1 \cdot P2 + \omega'_2 \cdot \text{Resize}(P3) \quad (2)$$

$$P2_out = \omega'_3 \cdot P2 + \omega'_4 \cdot P2_mid + \omega'_5 \cdot \text{Resize}(P1_out) \quad (3)$$

其中, Resize 代表上采样或下采样操作, ω'_i 是生成的归一化后的可学习权重,下标不同是为了区分.

1.3 改进 Transformer 模块

对于高层特征映射, DeepLabv3+网络通过 ASPP 模块中空洞卷积进行处理.虽然可以扩大对特征整体的感受野,但是空洞卷积会造成图像中采样的点不连续,引入不准确的特征表示,并且无法有效地捕捉上

下文信息. 为此, 本文通过改进 Transformer 模块, 对网络高层特征映射进行并行处理, 在建立全局依赖关系

的同时, 弥补空洞卷积带来的问题. 改进模块如图 3 所示.

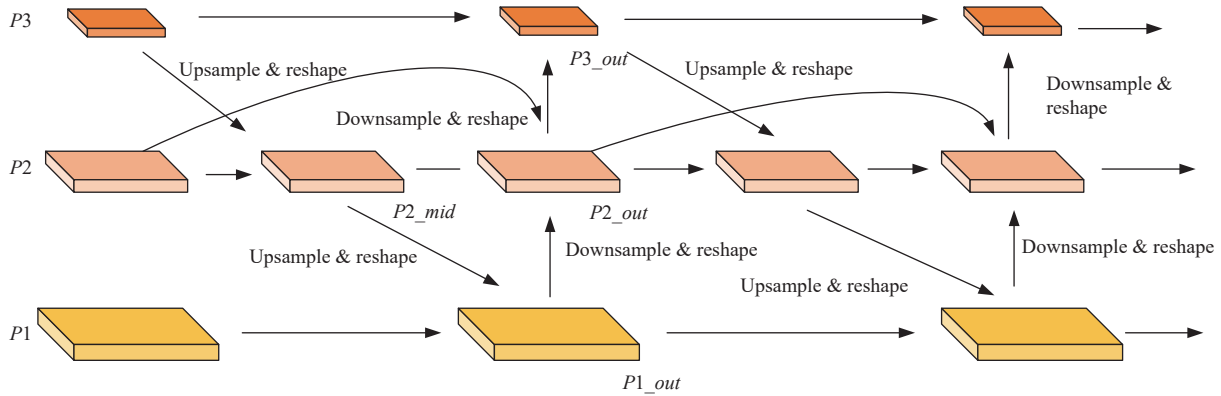


图 2 本文设计的 BiFPN 模块

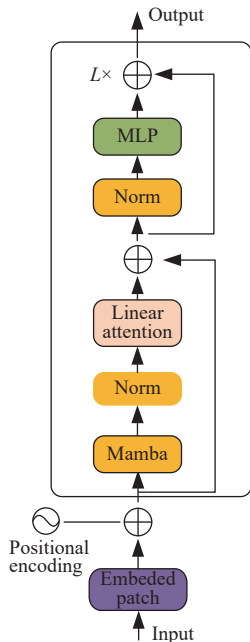


图 3 改进 Transformer 模块图

由于 Transformer 可以很好地建立全局依赖, 但无法跨长序列更高效和有效地捕获相关信息, 而 Mamba^[19-22] 可以跨长序列捕捉信息. 因此, 本文在 Transformer 注意力模块前插入 Mamba 块捕获长序列间的关系, 其中 Mamba 模块如图 4 所示.

Mamba 模块会将输入通过投影层扩展, 分别经过 1D 卷积和 SiLU 激活函数激活, 再通过 SSM 模块实现序列线性的扩展, 最终通过残差连接. SSM 模块如图 5 所示.

传统状态空间模型的时序结构导致了其输出状态完全依赖有序的输入数据. 一旦输入数据增减, 或者顺

序有所变化, 那么状态空间模型就无法进行处理. Mamba 通过设计类似门控结构, 打破传统 SSM 的限制约束, 利用离散可学习变量 Δt , 直接作用于 B 和 C 中, 将其变为变化状态, 通过数据的流动间接干扰 A 的状态. SSM 主要由状态方程和观测方程组成, 如式 (4)、式 (5) 所示:

$$h_t = Ah_{t-1} + Bx_t \quad (4)$$

$$y_t = Ch_t \quad (5)$$

其中, h_{t-1} 表示系统 $t-1$ 时刻状态; A 表示状态转移矩阵, 描述系统状态随时间的演变; x_t 代表输入; B 代表描述外部输入 x_t 对状态变化的影响; C 表示观测矩阵, y_t 代表观测结果.

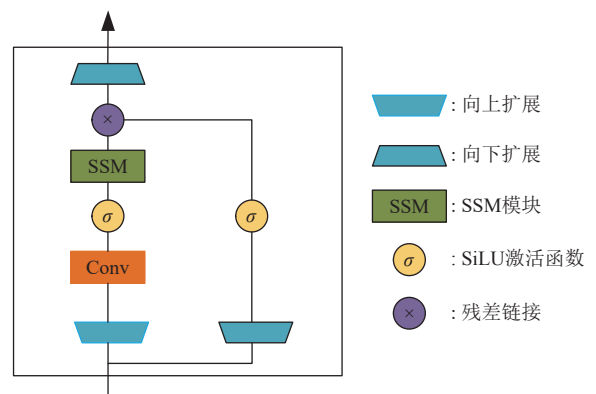


图 4 Mamba 模块图

经过 Mamba 处理的数据可以跨长序列有效地捕获相关信息, 再通过注意力机制计算相关性 (注意力权重), 然后使用这些权重对序列进行加权求和, 得到更新的表示. 并经过多层感知机进一步处理和变换特征,

最后, 将经输出的特征图恢复到原空间顺序, 并重塑到特征映射大小, 将其与 ASPP 模块输出进行相加.

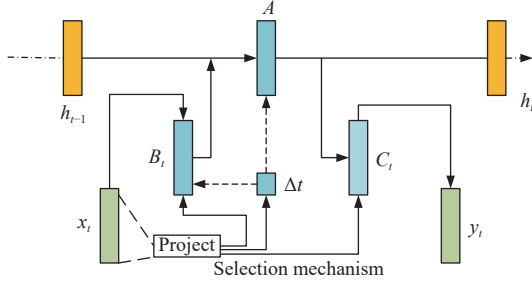


图5 SSM 模块图

1.4 极化自注意力

尽管高层特征映射通过并行处理, 两个分支输出进行相加可以很好地弥补空洞卷积的弊端, 但考虑到两者直接相加对于数据可能带来不同的影响, 仍需要一些额外的处理措施以确保数据的稳定. 为此, 本文引入极化自注意力机制 (PSA), 减少直接相加带来的负面影响. PSA 模块, 如图 6 所示.

它通过分别处理特征图的通道维度和空间维度, 增强模型对不同特征的捕捉能力, 并通过联合处理进一步优化特征表达.

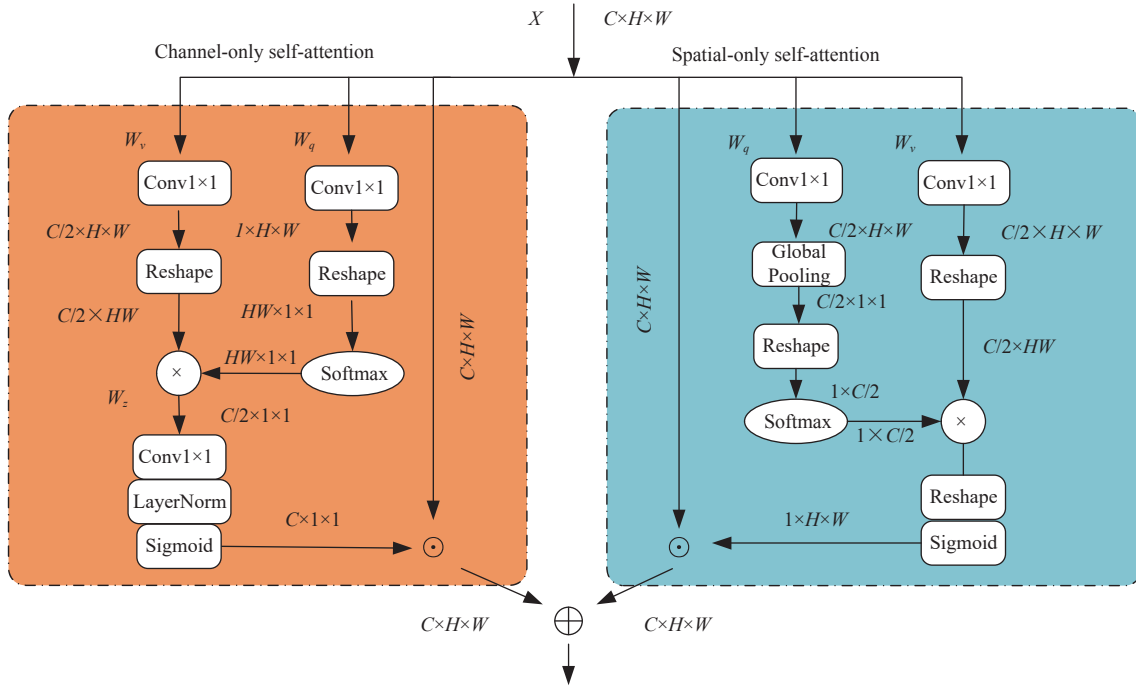


图6 PSA 模块图

图 6 中, 通道极化的操作是通过调整输入特征图的各个通道的权重来增强或抑制特定的通道特征. 首先, 通过 1×1 卷积将输入特征的通道维度压缩, 生成特征 W_q , 并通过 W_v 保留较高通道信息. 使用 Softmax 对压缩后的 W_q 进行信息增强, 确保重要特征不丢失. 通过变换将 W_q 与 W_v 进行矩阵乘法, 类似自注意力机制中的相关性计算, 捕捉通道间的依赖关系. 然后, 通过 1×1 卷积和层归一化恢复通道维度, 同时保持特征的稳定性. 最后, 通过 Sigmoid 激活, 将权重控制在 $[0, 1]$ 区间, 为后续特征加权提供参数. 通道分支计算式 (6) 为:

$$A^{\text{ch}}(X) = F_{\text{SG}} \left[W_{z|\theta_1} (W_v(X)) \times F_{\text{SM}} (\sigma_2 (W_q(X))) \right] \quad (6)$$

其中, $A^{\text{ch}}(X) \in \mathbb{R}^{C \times 1 \times 1}$, F_{SM} 表示 Softmax 函数, F_{SG} 表示

Sigmoid 函数, W 代表各种卷积, σ 表示降维, \times 表示点积运算. 通过这种方式, 通道极化注意力机制.

空间分支的权重计算与通道分支有一些相似之处, 但其主要处理的是空间维度的特征. 首先, 利用 1×1 卷积将输入特征转换为 W_q 和 W_v , W_q 的空间维度被压缩为 1×1 , W_v 的空间维度保持较大. 对 W_q 的空间信息进行全局池化, 并用 Softmax 进行增强, 确保压缩后的 W_q 能够保留重要的全局信息. 将增强后的 W_q 与 W_v 进行矩阵乘法, 捕捉全局空间特征之间的依赖关系. 将结果重塑为与输入特征相同的空间维度, 然后通过 Sigmoid 将权重压缩到 $[0, 1]$ 范围, 生成最终的空间权重图, 使得模型能够自动聚焦于图像中最有信息的空间区域, 从

而提升对局部特征的捕捉能力. 空间分支的权重计算式 (7) 为:

$$A^{sp}(X) = F_{SG} \left[\sigma_3 \left(F_{SM} \left(\sigma_1 \left(F_{GP} \left(W_q(X) \right) \right) \times \sigma_2 \left(W_v(X) \right) \right) \right) \right] \quad (7)$$

其中, $A^{sp}(X) \in R^{1 \times H \times W}$, W 代表各种卷积, σ 改变通道数, F_{GP} 代表全局池化.

2 实验结果与分析

2.1 实验数据集

(1) 道路积水数据集

由于道路积水数据集较少, 并没有公开的大型数据集图片. 因此本文采用某市城市管理真实图片作为数据集, 总共抽取道路积水图片 1543 张, 采用 labelme 工具对图片进行标注. 为方便验证模型效果, 本文将数据集分成训练集 1080 张、验证集 309 张、测试集 154 张.

2.2 评价指标

实验主要通过平均交并比 ($mIoU$)、像素精确度 (PA) 来考量算法性能. 平均交并比是用以展示预测成果和原本图像标签的匹配程度. 累加每一个种类的 IoU 并求平均, 最后得到的数值就是最终的 $mIoU$, 计算公式如下所示:

$$mIoU = \frac{1}{N} \sum_{i=0}^N \frac{P_{ij}}{\sum_{j=0}^N P_{ij} + \sum_{j=0}^N P_{ji} - P_{ii}} \quad (8)$$

关于 $mIoU$ 的数值, 它在 0-1 之间变动, 数值大小可以直接反映图像分割的精确度, 数值越接近 1, 其精确度越高.

像素精确度是评估语义分割算法性能的重要指标之一. 它用于衡量模型在像素级别上的分类准确性, 即每个像素被正确分类到相应的语义类别的比例. 来衡量模型的分割精度, 其公式如下:

$$PA = \frac{\sum_{i=0}^k P_{ii}}{\sum_{i=0}^k \sum_{j=0}^k P_{ij}} \quad (9)$$

其中, N 为所有像素种类的个数, P_{ij} 表示实际类别为 i 类但预测为 j 类的像素总数, P_{ii} 表示实际类别为 i 类预测也为第 i 类的像素总数.

2.3 实验参数

本文算法使用 Python 3.8、PyTorch 1.13.0 编程实现. Ubuntu 18.04 操作系统, CPU 为 Intel(R) Xeon(R) CPU E5-2609 v4@1.70 GHz, GPU 为 Tesla P40, CUDA 11.7 加速运算.

道路积水数据集 crop-size 设置为 512, batch-size 设置为 16, 优化器为随机梯度下降, 初始学习率为 0.01, 学习衰减为 0.05, epoch 为 100.

损失函数采用交叉熵损失 (cross entropy loss) 函数, 其公式如下:

$$L(y, y') = - \sum_{i=1}^C y_i \log(y'_i) \quad (10)$$

其中, C 代表类别总数, y_i 表示样本属于第 i 类的真实标签, y'_i 表示模型对第 i 类的预测概率.

训练过程的损失收敛曲线如图 7 所示, 训练轮次为 100, 在 epoch=80 左右, 曲线趋于平稳, 参数训练至最佳, 模型得到有效收敛.

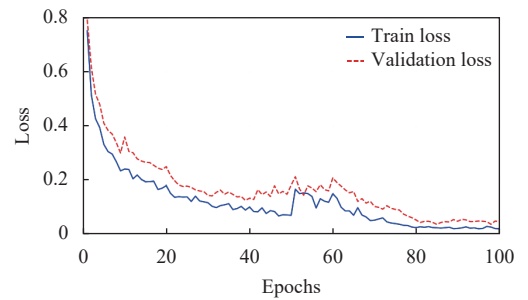


图 7 训练损失曲线图

2.4 消融实验

2.4.1 Transformer 参数选择

由于本文基于 Transformer 对高级特征映射并行处理, 考虑到 Transformer 结构中参数会对实验结果产生的影响. 本文通过控制变量设置了 6 组实验. 实验结果如表 1 所示.

表 1 不同参数设置的消融实验 (%)

Method	patch数量	depth数量	$mIoU$	PA
1	4	6	84.22	95.29
2	4	12	84.67	95.32
3	8	6	85.72	95.62
4	8	12	85.28	95.43
5	16	6	85.04	95.35
6	16	12	83.72	95.14

通过实验发现, 对于较小特征图采用较大 patch 会造成实验指标的下降. 当 patch 大小为 8、堆叠 6 层结

构时,实验取得最好效果。

2.4.2 Mamba 模块验证

为了探究改进后的 Transformer 模块,能否取得更好效果,本文追加了两组实验,验证 Mamba 模块的影响,patch 和 depth 设置为 8、6,实验结果如表 2 所示。

表 2 增加 Mamba 模块的消融实验 (%)

Method	Mamba	<i>mIoU</i>	<i>PA</i>
1	—	85.72	95.62
2	√	86.31	95.68

通过实验可以发现,当在 Transformer 注意力模块前插入 Mamba 模块捕获长序列间的关系,可以有效地提升实验效果。

考虑插入新模块会增加网络计算的复杂度,为了平衡性能和复杂度,本文对改进的 Transformer 的 depth 进行探索.实验结果如表 3 所示。

表 3 depth 探究的消融实验

Method	depth	Params (M)	<i>mIoU</i> (%)	<i>PA</i> (%)
1	4	67.3 ↓	86.01	95.57
2	6	73.4	86.31	95.68

通过实验发现 depth 为 4 层时,相比 depth 为 6 层 *mIoU* 仅下降了 0.3%,但是参数量降低了 6.1M。

2.4.3 极化注意力结构选择

关于极化自注意力机制,由于其由通道分支和空间分支组成,两者可以串联或并行.考虑到不同组合可能产生不同的影响,本文设置两组实验进行验证.实验结果如表 4 所示,根据实验结果表明,使用串联比并联效果更好。

2.4.4 不同模块性能

为了进一步的验证算法改进各模块的有效性,在道路积水数据集上进行了各模块的对比实验.实验结果如表 5,通过实验结果可发现,当仅使用 PSA 时,本

文算法获得了 84.03% 的 *mIoU*,在仅使用 BiFPN 模块时,本文算法获得了 85.10% 的 *mIoU*,使用改进的 Transformer 模块时,本文算法获得了 86.01% 的 *mIoU*.而同时使用 3 个模块时,本文算法获得了 87.54% 的 *mIoU*.可以得出,使 BiFPN 模块能够融合重要特征并充分利用低层特征映射.相较于单层特征映射的方法,采用 BiFPN 模使得模型能够更好地捕获不同尺度下的信息,并且有效地传递高层语义信息到低层特征图.同时,对高级特征映射并行处理,使得改进的 Transformer 建立全局依赖,弥补了 ASPP 中空洞卷积带来的负面影响.同时使用 3 个模块可以有效地提升算法分割性能。

表 4 注意力组合方式的消融实验 (%)

Method	方式	<i>mIoU</i>	<i>PA</i>
1	并联	83.76	95.11
2	串联	84.03	95.28

表 5 不同模块的消融实验结果 (%)

Method	PSA	BiFPN	Transformer	<i>mIoU</i>	<i>PA</i>
1	√	—	—	84.03	95.28
2	—	√	—	85.10	95.67
3	—	—	√	86.01	95.57
4	√	√	—	85.31	95.72
5	√	—	√	86.26	95.63
6	—	√	√	86.87	95.91
7	√	√	√	87.54	96.61

2.5 道路积水数据集实验结果

为了验证改进算法的性能,本文基于道路积水数据集,对改进 DeepLabv3+模型进行实验验证,并选择了 PSPNet、U-Net^[19]、UNet++^[20]等模型作为对比模型.为确保实验数据的有效性,所有算法均在相同设置参数下完成实验.不同算法的实验对比结果如表 6 所示。

表 6 不同算法在道路积水数据集上的实验结果

Method	Backbone	Pretrained	Params ($\times 10^6$)	Speed (f/s)	GFLOPs	<i>mIoU</i> (%)	<i>PA</i> (%)
PSPNet	ResNet50	Y	46.70	28.17	59.21	69.45	89.72
SegNet	VGG16	Y	29.5	13.00	286.00	72.81	89.92
FCN	VGG16	Y	134	22.80	78.10	70.15	89.87
U-Net	Xception39	N	24.89	34.98	58.93	76.03	92.07
UNet++	ResNet50	Y	31.32	27.81	68.20	82.41	94.18
DeepLabv3+	ResNet50	Y	48.21	26.25	69.24	83.32	94.95
Ours	ResNet50	Y	76.43	24.74	83.16	87.54	96.61

实验结果显示,本文算法在道路积水数据集上取得了 87.54% 的 *mIoU*,*PA* 取得了 96.61%。DeepLabv3+ 算法取得了 83.32% 的 *mIoU*,*PA* 取得了 94.95%。与未

经改进的 DeepLabv3+算法相比,本文算法参数量增加 1/3,帧数仅下降 2 帧左右,*mIoU* 提高了 4.22%,*PA* 提高了 1.66%。与其他算法相比,本文算法在两个评估指

标上均有显著提升. 通过数据来看, 在高性能硬件和经过优化的条件下, 改进算法可能接近实时检测, 但在计算资源有限的边缘设备上, 可能需要进一步优化以达到理想的实时帧率.

道路积水数据集的可视化结果如图 8 所示. 可视化结果展示了本文算法与其他算法的分割结果对比. 第 1 行中, 本文算法对积水检测分割较为完整; 第 2 行

中, 本文算法对于整体信息保留较多, 其他算法都存在一些信息丢失问题; 第 3 行, 本文算法对积水边缘分割明显; 第 4 行, DeepLabv3+ 算法对上面积水并未分割完整, 本文算法较为充分的分割上面积水的细节特征. 综上所述可以得知, 改进后的算法有效地利用了低层特征, 在分割连续性上表现良好, 例如对物体边界的分割较为连续和光滑, 在分割特征上像素点保留较为还原.

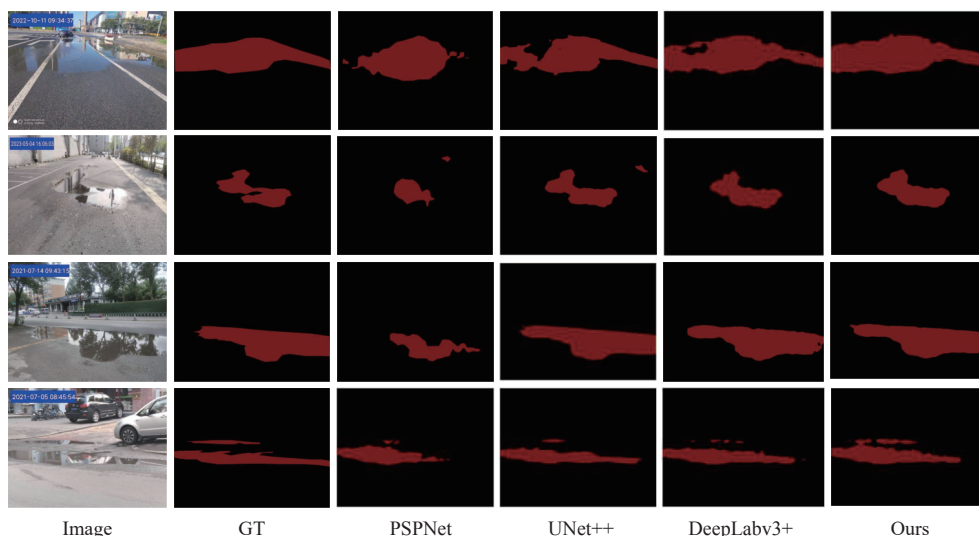


图 8 不同算法的分割结果图

3 结束语

为了解决道路积水的检测问题, 本文基于 DeepLabv3+ 网络模型提出改进算法, 通过在编码器端设计 BiFPN 模块, 促进在不同尺度和层级上进行特征融合; 并改进 Transformer 结构设计并行分支对高级特征映射进行处理, 构建全局依赖关系. 同时引入 PSA 模块, 确保双分支数据交互的稳定. 在道路积水数据集上进行实验, 结果表明改进网络可以有效地提升道路积水检测得性能, 在一定程度上缓解了边界不连续问题.

参考文献

- 胡昊, 孙爽, 马鑫, 等. 基于图注意力网络的城市内涝积水预测与研究. 人民黄河, 2024, 46(4): 43–48. [doi: 10.3969/j.issn.1000-1379.2024.04.007]
- 康得军, 温儒杰, 邱福杰, 等. 基于 SWMM 和 GIS 的城市内涝 4D 可视化研究. 中国给水排水, 2023, 39(13): 133–138.
- 肖桐, 余敦先, 夏军, 等. 不同降雨情况对武汉市主城区典型区域积水过程影响分析. 武汉大学学报(工学版), 2023, (8): 969–977. [doi: 10.14188/j.1671-8844.2023-08-008]
- 单志勇, 宫加辉. 基于改进的 Faster R-CNN 的车辆目标检

测. 现代计算机, 2021(20): 74–79. [doi: 10.3969/j.issn.1007-1423.2021.20.015]

- 周倩倩, 司徒祖祥, 腾帅, 等. 基于卷积神经网络的排水管道缺陷智能检测与分类. 中国给水排水, 2021, 37(21): 114–118.
- Gaur L, Bhatia U, Jhanjhi NZ, *et al.* Medical image-based detection of COVID-19 using deep convolution neural networks. *Multimedia Systems*, 2023, 29(3): 1729–1738. [doi: 10.1007/s00530-021-00794-6]
- Sazara C, Cetin M, Iftexharuddin KM. Detecting floodwater on roadways from image data with handcrafted features and deep transfer learning. *Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference*. Auckland: IEEE, 2019. 804–809.
- Witherow MA, Sazara C, Winter-Arboleda IM, *et al.* Floodwater detection on roadways from crowdsourced images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2019, 7(5–6): 529–540.
- Huang J, Kang JL, Wang HM, *et al.* A novel approach to measuring urban waterlogging depth from images based on

- mask region-based convolutional neural network. *Sustainability*, 2020, 12(5): 2149. [doi: [10.3390/su12052149](https://doi.org/10.3390/su12052149)]
- 10 Jafari NH, Li X, Chen Q, *et al.* Real-time water level monitoring using live cameras and computer vision techniques. *Computers & Geosciences*, 2021, 147: 104642.
- 11 Shu YQ, Zheng GB, Yan XW. Application of multiple geographical units convolutional neural network based on neighborhood effects in urban waterlogging risk assessment in the city of Guangzhou, China. *Physics and Chemistry of the Earth, Parts A/B/C*, 2022, 126: 103054. [doi: [10.1016/j.pce.2021.103054](https://doi.org/10.1016/j.pce.2021.103054)]
- 12 Liu JQ, Shang YJ, Li XY, *et al.* Road traffic waterlogging detection based on YOLOv5. *Proceedings of the 1st International Artificial Intelligence Conference*. Nanjing: Springer, 2023. 45–58.
- 13 白岗岗, 侯精明, 韩浩, 等. 基于深度学习的道路积水智能监测方法. *水资源保护*, 2021, 37(5): 75–80. [doi: [10.3880/j.issn.1004-6933.2021.05.012](https://doi.org/10.3880/j.issn.1004-6933.2021.05.012)]
- 14 胡昊, 李擎, 马鑫, 等. 基于影像的道路积水监测研究. *华北水利水电大学学报(自然科学版)*, 2023, 44(1): 62–70.
- 15 廖宇鸿, 黄国如. 基于深度学习和椭圆检测的城市道路积水深度监测方法研究. *人民珠江*, 2023, 44(6): 1–8, 17. [doi: [10.3969/j.issn.1001-9235.2023.06.001](https://doi.org/10.3969/j.issn.1001-9235.2023.06.001)]
- 16 张峥, 左向阳, 龙岩, 等. 基于改进 YOLOv8 的道路积水监测方法. *人民珠江*, 2024, 45(10): 44–50.
- 17 Chen LC, Zhu YK, Papandreou G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 801–818.
- 18 Tan MX, Pang RM, Le QV. EfficientDet: Scalable and efficient object detection. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 10781–10790.
- 19 Gu A, Dao T. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv:2312.00752*, 2023.
- 20 Zhu LH, Liao BC, Zhang Q, *et al.* Vision mamba: Efficient visual representation learning with bidirectional state space model. *Proceedings of the 41st International Conference on Machine Learning*. OpenReview.net, 2024.
- 21 Li SF, Singh H, Grover A. Mamba-ND: Selective state space modeling for multi-dimensional data. *Proceedings of the 18th European Conference on Computer Vision*. Milan: Springer, 2025. 75–92.
- 22 Lieber O, Lenz B, Bata H, *et al.* Jamba: A hybrid Transformer-Mamba language model. *arXiv:2403.19887*, 2024.

(校对责编: 张重毅)