

基于RT PSPNet的路面条状修补实时语义分割研究

陈文柯¹, 李乐成², 刘国玺², 黄苾¹, 代飞¹

(1. 西南林业大学大数据与智能工程学院, 云南 昆明 650224;

2. 西南林业大学林学院, 云南 昆明 650224)

摘要: 路面病害中的条状修补检测是路面智能养护的关键技术之一。现有基于深度学习的语义分割模型在此任务中普遍存在模型参数量大、推理速度慢、细节特征丢失等问题, 难以同时满足精度与实时性要求。针对上述问题, 提出了一种轻量化高精度的实时语义分割模型——RT PSPNet。该模型在PSPNet基础上进行了多维改进: 首先, 对ResNet50主干进行深度与通道数的结构化剪枝, 得到浅层轻量ResNet, 在显著减少参数量的同时保留更多细节特征; 其次, 设计轻量化小型金字塔池化模块(CPPM)替代原PPM, 减少池化分支并引入逐深度卷积, 以提升特征提取效率; 最后, 引入多尺度特征融合机制与多分割头复合损失函数, 增强模型对细长目标的完整分割能力与训练稳定性。实验结果表明, 在自建条状修补数据集上, RT PSPNet相比原PSPNet在分割精度(mIoU)上提升7.06个百分点, 参数量减少96.4%, 计算量降低76.1%, 推理速度达到147 fps, 结合混合精度加速策略后可达258 fps, 为基准模型的8.6倍。与当前多种主流及轻量化先进模型相比, RT PSPNet在分割精度、推理速度与模型轻量化3个方面表现出综合的领先优势。RT PSPNet的模型代码和权重已在GitHub上供免费下载使用。

关键词: 语义分割; 路面病害; 轻量化; 实时推理; PSPNet

中图分类号: TP391

文献标志码: A

doi: 10.11959/j.issn.2096-6652.ZNKX26016

Real-time semantic segmentation of strip mending based on RT PSPNet

Chen Wenke¹, Li Lecheng², Liu Guoxi², Huang Bi¹, Dai Fei¹

1. College of Big Data and Intelligent Engineering, Southwest Forestry University, Kunming 650224, China

2. Forestry College, Southwest Forestry University, Kunming 650224, China

Abstract: The detection of strip mending in pavement diseases provides crucial technical support for intelligent road maintenance. Existing deep learning-based semantic segmentation models often suffer from issues such as large model parameters, slow inference speed, and loss of detailed features in strip mending detection tasks, making it difficult to balance accuracy and real-time performance. To address these challenges, RT PSPNet, a lightweight and high-precision real-time semantic segmentation model, was proposed. Based on PSPNet, RT PSPNet incorporated multi-dimensional improvements. Firstly, structured pruning was applied to the ResNet50 backbone in terms of depth and channel number, resulting in a shallow lightweight ResNet that significantly reduced parameters while preserving more detailed features. Secondly, a lightweight compact pyramid pooling module (CPPM) was designed to replace the original PPM, which re-

收稿日期: 2026-01-30; 修回日期: 2026-05-15

通信作者: 代飞, daifei@swfu.edu.cn

基金项目: 云南省重大科技专项计划项目 (No.202502AD080004); 云南省高校服务重点产业科技项目 (No.FWCY-ZNT2024020); 云南省基础研究专项项目 (No.202301AU070223)

Foundation Items: Key Research and Development Program of Yunnan Province (No.202502AD080004), Key Industry Science and Technology Projects for University Services in Yunnan Province (No.FWCY-ZNT2024020), Foundational Research Special Projects of Yunnan Province (No.202301AU070223)

duced pooling branches and introduced depth-wise convolutions to further enhance feature extraction efficiency. Finally, a multi-scale feature fusion mechanism and a multi-head composite loss function were introduced to improve the model's ability to segment slender targets completely and enhance training stability. Experimental results on a self-built strip mending dataset show that RT PSPNet improves the segmentation accuracy (mIoU) by 7.06 percentage points compared to the original PSPNet, while reducing the parameter count by 96.4%, computational cost by 76.1%, and achieving an inference speed of 147 fps, and reaching 258 fps while using a mixed precision speeding strategy, which is 8.6 times that of the baseline model. Compared with many mainstream and lightweight state-of-the-art models, RT PSPNet demonstrates comprehensive advantages in accuracy, speed, and model light-weighting. RT PSPNet's code and weights are available for free accessing and downloading on GitHub.

Key words: semantic segmentation, pavement diseases, light-weighting, real-time inference, PSPNet

0 引言

路面病害是指水泥和沥青路面出现的各类破损,主要包括修补、裂缝、坑槽等^[1]。条状修补作为修补类病害的常见形式,具有形态细长、宽度不一、轮廓不规则等特点。与裂缝相比,条状修补宽度更大,轮廓更平滑;与坑槽相比,条状修补形态更细长,且表面与道路基本持平。条状修补不仅影响道路本身的结构安全和耐久性,还会对行车安全、养护成本及环境造成负面影响,因此需要高效、准确的检测方法以支持道路养护工作。

条状修补的检测方法主要分为人工检测和计算机自动检测两类。人工检测依赖专业人员实地勘查,耗费人力且效率低,检测结果受主观因素影响大,不适用于大规模、长周期的道路巡检。计算机自动检测可分为传统图像处理方法与基于深度学习的方法。传统方法通常依赖手工设计特征,对环境变化敏感、泛化能力有限;深度学习方法通过数据驱动自动学习特征,在检测精度、推理速度与泛化能力上具有明显优势,更适用于当前路面病害的智能识别任务^[2],有利于传统研究与现有先进技术的有机结合,并促进多学科交叉发展^[3]。

语义分割作为计算机视觉的基础任务之一^[4],旨在实现像素级别的图像前景与背景分离,精确刻画目标轮廓。基于深度学习的条状修补语义分割模型能够自动区分修补区域与正常路面,具备高精度与强泛化能力。然而,现有主流语义分割模型在面向条状修补这类特定场景时,常因参数冗余导致推理效率低下,或因结构设计不合理而丢失细节特征,难以在精度与效率之间取得良好平衡^[5]。

随着语义分割技术的发展,兼顾精度与速度的模型已成为该领域的重要研究方向。在路面病害的

自动检测方面,徐文远等^[6]提出了一种新模型 RT-FormerF10,其在保证推理速度的前提下利用双模态数据进行训练,强化了不同环境下分割路面裂缝的泛化能力。唐源等^[7]提出了一种改进于 SegFormer 的模型,重新设计了融合不同尺度特征的解码器等操作,获得了比原模型更好的路面裂缝分割精度。韩建峰等^[8]提出了一种改进于 U-Net 的航拍路面破损检测模型,通过改进编码器和添加注意力机制的方式,提升了路面破损的分割效果。杨萍等^[9]提出了一种改进 DeepLabv3+ 的道路裂缝检测方法,在保证实时性的同时还在分割精度上优于基准模型。上述研究在裂缝等路面病害自动检测和分割方面有所进展,但是缺乏针对条状修补路面病害的研究以及对语义分割模型本身实时性推理的研究。在轻量化模型结构的探索方面,Yu 等^[10]提出了一个双分支的轻量化实时语义分割模型 BiSeNet V2,同时具有良好的分割精度和推理速度。Mehta 等^[11]提出了一种通用的轻量化 ViT 特征提取主干 MobileViT,适合部署在边缘设备上。Xu 等^[12]提出了一个能够捕获长距离上下文信息的实时语义分割模型 SCTNet,这是近两年的轻量化语义分割 SOTA。上述模型虽然具有不同程度的轻量化优势,但是分割精度通常无法和主流语义分割模型相比,而且缺乏定制化的优化和改良工作,无法在特定的路面病害检测识别和分割任务上兼顾效率与精度,难以达到路面养护的标准要求。

尽管上述研究在路面病害检测或模型轻量化方面取得了一定进展,但现有模型在条状修补分割任务中常难以兼顾精度、轻量化与推理效率。条状修补实时检测面临以下挑战。

(1) 现有语义分割模型在条状修补检测任务中参数量大、推理效率低,需进行针对性轻量化

设计。

(2) 条状修补形态细长多变, 深层网络易丢失细节与语义信息, 且主流模型如PSPNet^[13]缺乏多尺度特征融合, 导致分割结果出现明显的漏分割、误分割现象。

(3) 常规训练仅采用单一分割头与损失函数, 缺乏对网络多层次特征的协同监督, 限制模型性能的充分挖掘。

针对上述问题, 本文提出了一种基于PSPNet改进的实时语义分割模型——RT PSPNet, 用于路面条状修补的实时语义分割。本文主要改进包括: 对ResNet50^[14]主干进行深度与通道数的结构化剪枝, 构建浅层轻量ResNet; 设计轻量化小型金字塔池化模块(compact pyramid pooling module, CPPM)以替换原金字塔池化模块(pyramid pooling module, PPM), 提升特征提取效率; 引入多尺度特征融合机制, 增强对细长目标的特征保持能力; 采用多分割头复合损失函数进行多层次训练监督; 在推理阶段采用混合精度加速策略以提高效率。具体地, 本文主要贡献如下。

(1) 构建了面向条状修补的语义分割数据集, 包含1 000张涵盖不同光照、背景干扰与病害形态的高质量图像, 为相关研究提供数据支持。

(2) 提出了轻量化高精度实时语义分割模型RT PSPNet, 通过浅层轻量ResNet、CPPM、多尺度特征融合与多分割头损失函数, 在分割精度、推理速度与模型轻量化之间实现优异平衡。

(3) 在自建数据集上进行了全面实验验证, RT PSPNet相比原PSPNet在mIoU上提升7.06个百分点的同时, 参数量减少96%, 计算量降低76%~92%, 推理速度为147 fps, 在采用无精度损失的混合精度加速策略后, 推理速度最高可达258 fps。与FCN^[15]、DeepLabv3+^[16]、HRNet^[17]、SCTNet和SegFormer^[18]等主流模型和轻量化模型相比, 所提模型在多项指标上均表现突出。

1 RT PSPNet条状修补分割网络

1.1 网络结构

基准PSPNet采用ResNet50作为特征提取主干, 虽具备较强的特征提取能力, 但参数量与计算量都较大, 且其解码器中的PPM缺乏与浅层特征的融合, 不利于细长目标的完整分割。为此, 本文提出RT PSPNet, 其整体结构如图1所示。

RT PSPNet主要从编码器和解码器两方面进行改进: 首先, 在编码器中, 对ResNet50主干进行深度与通道数的结构化剪枝, 构建浅层轻量ResNet, 在减少参数量的同时保留更多的细节信息; 其次, 在解码器中, 设计轻量化CPPM替代原PPM, 减少池化分支并引入逐深度卷积(depthwise convolution)^[19], 同时融合浅层特征图, 提升特征表达能力和分割精度。

1.2 浅层轻量ResNet主干

为适应条状修补的细长形态, 避免深层特征图尺寸过小导致细节丢失, 同时达到轻量化的效果, 本文对ResNet50主干进行了结构化剪枝, 在保留有利于细长特征留存的高分辨率特征图的同时, 舍弃了不利于细长特征完整性保持的低分辨率特征图, 同步聚焦主干网络的轻量化改进和细长目标提取能力, 并得到浅层轻量ResNet主干, 具体改进如图2所示。

首先, 将原PSPNet主干的第一个7×7大小的卷积核替换为3个3×3的卷积核, 整合为deepbase模块, 该模块在感受野上保持不变, 参数量和计算量有一定程度的减少。

其次, 本文运用了结构化剪枝的方法, 以bottleneck层为剪枝的基本单位, 将ResNet50中的bottleneck数量从16缩减到了7, 主干的深度减少了一半以上, 实现了显著的轻量化。同时, 由于原本深层的低分辨率特征图被剪枝并舍弃, 较浅层的高分辨率特征图得以保留, 使改进后的主干深层特征图具有较大的64×64分辨率, 相比原主干最深层的16×16分辨率扩大了16倍, 更加有利于保留细长条形的条状修补所需的重要语义和细节信息。同时, bottleneck中的常规卷积核被统一替换为逐深度卷积核, 得到dw_bottleneck。dw_bottleneck不仅大幅减少了参数量和计算量, 还能保证基本的特征提取能力, 以达到非常显著、理想的轻量化效果。

最后, 将两阶段bottleneck的卷积通道数分别从原来的128和256缩减至96和128, 以实现进一步的轻量化, 得到浅层轻量ResNet。改进的浅层轻量ResNet主干和原ResNet50主干的参数量和计算量对比见表1。

从表1中可以看出, 本文提出的浅层轻量ResNet在参数量和计算量上相比原ResNet50具有显著优势, 其中, 浅层轻量ResNet的参数量仅约为原主干的4.7%, 减少了约95.3%的参数量; 另

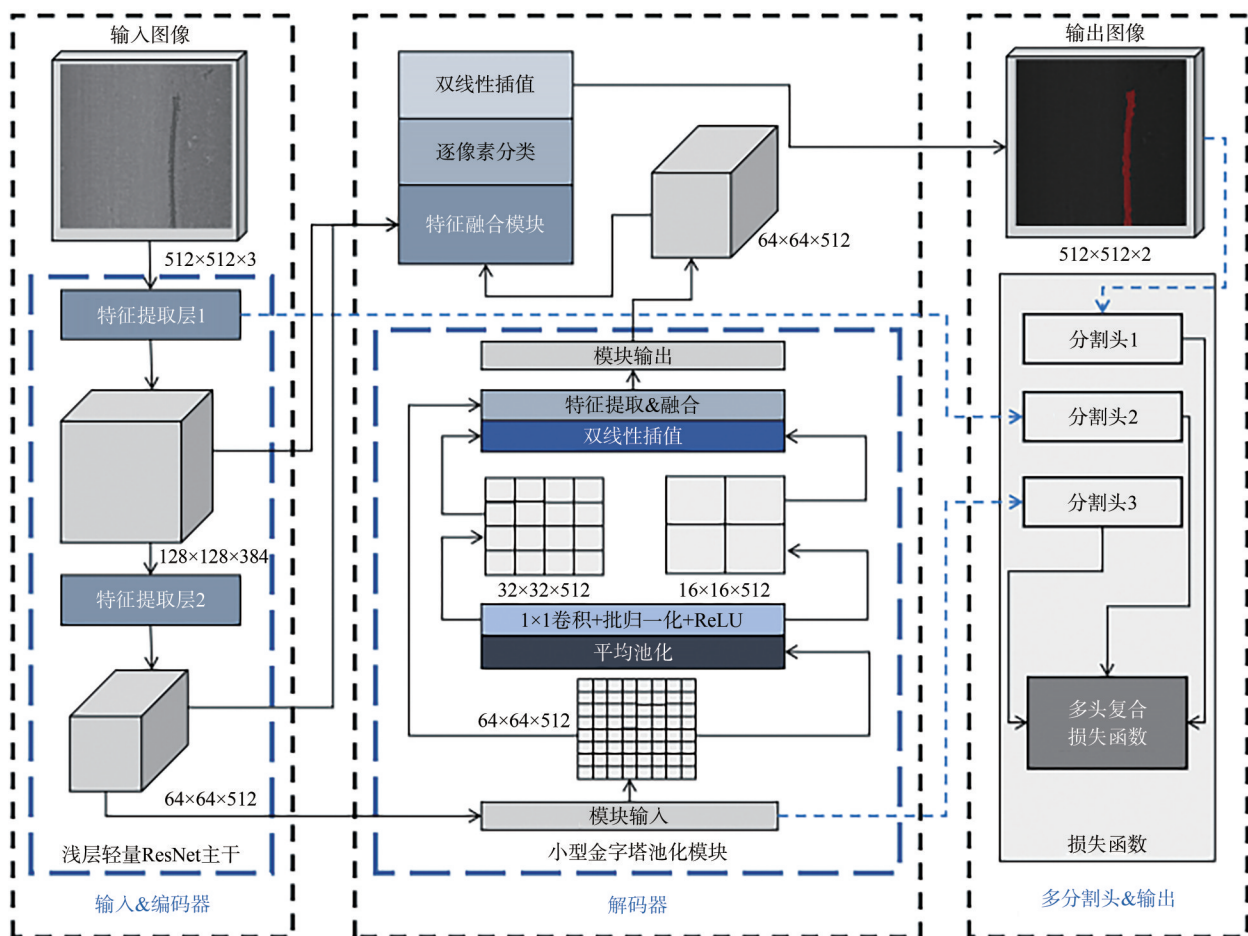


图1 RT PSPNet的网络结构

表1 浅层轻量ResNet主干和原ResNet50主干的参数数量和计算量对比

模型	参数量	计算量/GFLOPs
ResNet50	23.508×10^6	43.171
浅层轻量ResNet	1.104×10^6	23.326

外, 浅层轻量ResNet的计算量仅约为原主干的54%, 减少了约46%的计算量。综上, RT PSPNet采用的浅层轻量ResNet相比原主干拥有非常显著的轻量化优势, 更有助于提升模型的实时推理性能。

1.3 轻量化CPPM

由于条状修补具有形态细长、结构连续的特点, 原PSPNet解码器中的PPM在设计中存在与目标特性不匹配的问题。PPM通过多个并行的池化分支(通常为4个)捕获多尺度上下文信息, 但其较大的池化窗口与密集的池化操作容易对细长目标造成过度降维与细节平滑, 导致深层特征图中的条状修补语义信息丢失, 进而影响分割精度。此外,

PPM的多分支结构也引入了较多的参数量与计算量, 降低了模型的推理效率。在条状修补分割任务中, 部分池化分支可能存在冗余, 且过度的池化操作会进一步削弱对细长结构特征的保留能力, 从而制约分割性能的提升。为此, 本文改进PPM, 提出了CPPM。PPM和CPPM的网络结构对比如图3所示。

由图3可知, CPPM通过结构化剪枝方法, 将原PPM中的4个池化分支精简为2个, 并相应增大了池化特征的尺寸(分别采用 16×16 和 32×32 的平均池化), 分别是模块输入特征图的 $1/4$ 和 $1/16$ 。采用具有更大感受野的池化分支替代原有较小分支, 在有效减少参数量与计算量的同时, 更有利于保留条状修补具有的细长结构特征。此外, 模块中原有的常规卷积被替换为逐深度卷积, 可以显著提升模块的训练和推理速度, 并且将多分支输出的特征拼接操作改为逐像素相加(element-wise addition), 所得的新特征图相比采用通道对齐的特征拼接方式拥

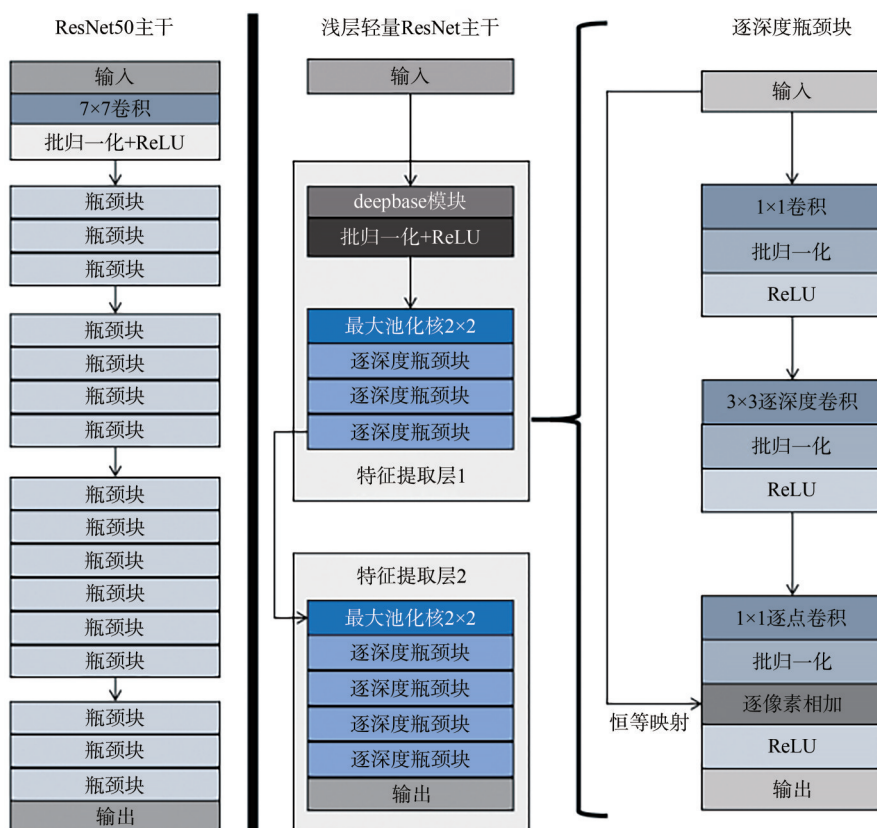


图2 ResNet50和浅层轻量ResNet结构的对比展示

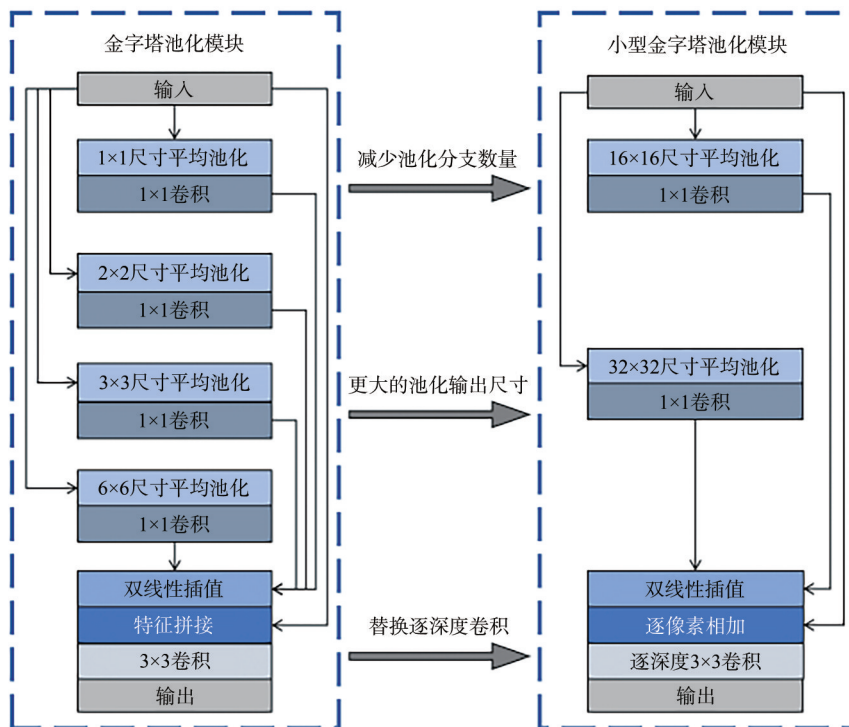


图3 PPM和CPPM的网络结构对比

有更少的通道数，同时有利于改善该模块的内部运行效率。这些改进不仅显著降低了解码器的计算开

销，也进一步提升了模型在细节保持与分割精度方面的性能。CPPM的计算过程可形式化表示为：

$$\text{output} = \text{dwconv}3 \times 3 \left(\text{conv}1 \times 1 \left(\text{Pool}_{16}(\text{input}) \right) + \text{conv}1 \times 1 \left(\text{Pool}_{32}(\text{input}) \right) \right) \quad (1)$$

其中, $\text{dwconv}3 \times 3$ 表示包含 3×3 逐深度卷积的模块, $\text{conv}1 \times 1$ 为 1×1 标准卷积, Pool_{16} 和 Pool_{32} 分别为 16×16 和 32×32 的平均池化操作, input 为输入特征图, output 为输出特征图。

1.4 多尺度特征融合

原 PSPNet 的解码器仅包含 PPM 和上采样操作, 缺乏对特征提取主干中多层次信息的有效利用。这导致模型难以同时捕获浅层的细节特征与深层的长距离上下文依赖, 无法通过多尺度特征融合充分提升表达能力。因此, 在处理具有“细长、不规则”形态的条状修补目标时, 其分割精度与完整性经常受限。

为缓解上述问题, 本文在 RT PSPNet 中引入多尺度特征融合机制。具体而言, 本文将主干网络中提取的浅层特征图与解码器中 CPPM 输出的深层特征图进行融合, 使融合后的特征同时具备丰富的细节信息与高层语义信息, 从而显著改善对条状修补结构的分割效果。RT PSPNet 的特征融合结构如图 4 所示。

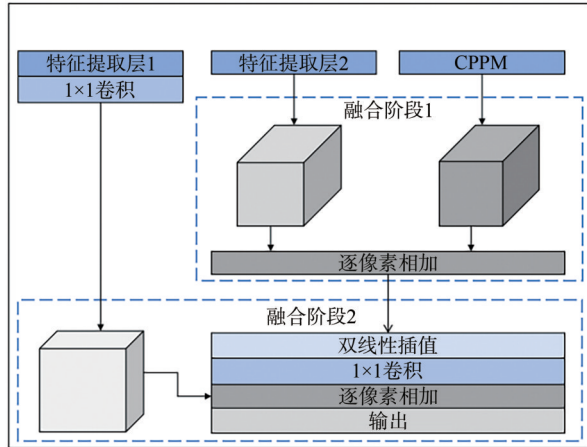


图4 RT PSPNet的特征融合结构

此外, 为减少信息在传递过程中的损失, 本文在 CPPM 内部添加了残差连接, 将模块输入与输出直接相加, 以保留重要的原始语义信息。在最终分类之前, 还引入一个 3×3 卷积层进行进一步特征细化与噪声抑制, 提升输出质量。该特征融合解码过程可形式化表示为:

$$\text{output} = \text{Classconv} \left(\text{conv}3 \times 3 \left(\text{conv}1 \times 1 \left(\text{Up} \left(\text{CPPM}(\text{d_feat}) \right) + \text{conv}1 \times 1(\text{s_feat}) \right) \right) \right) \quad (2)$$

其中, Classconv 表示 1×1 分类卷积, $\text{conv}3 \times 3$ 和 $\text{conv}1 \times 1$ 分别为 3×3 深度卷积和 1×1 标准卷积, Up 为 2 倍上采样操作, CPPM 为改进后的轻量化金字塔池化模块, d_feat 和 s_feat 分别为本文主干网络输出的深层特征图和浅层特征图。

1.5 多分割头复合损失函数

基准 PSPNet 在训练阶段采用交叉熵损失 (cross-entropy loss, CE Loss) 函数, 这也是语义分割任务中广泛使用的基础损失函数。为提升 RT PSPNet 的分割精度, 本文在 CE Loss 的基础上引入 Dice Loss^[20], 即 Dice 损失函数, 通过结合两者优势以加强对前景区域的拟合能力。其中, CE Loss 倾向于平等地对待并计算不同类别中每个像素点的损失值, 但容易受类别不平衡情况影响导致模型精度不高; Dice Loss 更侧重于衡量预测结果和真值标签的区域重叠程度, 能有效忽略类别不平衡干扰, 但训练过程可能不平稳, 且易出现振荡。本文将二者在训练阶段的值进行结合以提升损失函数对模型的监督水平, 初步构建的单分割头复合损失函数为:

$$\text{SH_hybridloss} = \text{CE Loss} + \text{Dice Loss} \quad (3)$$

其中, SH_hybridloss 为单分割头的复合损失函数, CE Loss 为交叉熵损失函数, Dice Loss 为 Dice 损失函数。

为了进一步提高 RT PSPNet 的精度, 本文尝试对其解码器的 3 个特征图输出位置添加分割头, 以进行多层次的训练监督, 即在模型的训练阶段使用多分割头的复合损失函数, 以此监督和提高不同深度位置下特征图的准确性, 在模型的推理阶段只采用最终的输出结果图, 并不使用训练过程中辅助的分割头。RT PSPNet 的 3 个分割头具体位置如图 5 所示。

为进一步提升模型在不同层次特征上的表征一致性, 本文在编码器中部、底部以及解码器输出端分别增设辅助分割头 (如图 5 所示), 构成多分割头监督机制。训练时, 各辅助头计算 CE Loss, 主输出头额外计算 Dice Loss, 并通过加权融合形成最终的多分割头复合损失函数:

$$\text{MH_hybridloss} = \text{Dice Loss} + \lambda \times (\text{CE Loss}_1 + \text{CE Loss}_2 + \text{CE Loss}_3) \quad (4)$$

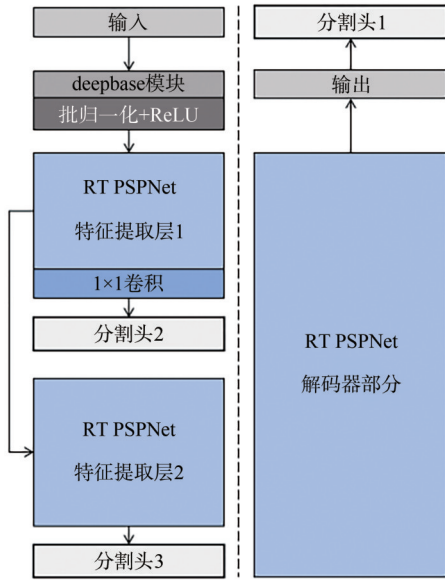


图5 RT PSPNet的3个分割头具体位置展示

其中, $MH_hybridloss$ 为多分割头的复合损失函数值, $CE Loss_1$ 、 $CE Loss_2$ 和 $CE Loss_3$ 分别为3个分割头计算的交叉熵损失值, $Dice Loss$ 为Dice损失函数, λ 为对损失值的加权放大系数。

1.6 推理阶段的混合精度加速策略

在模型训练阶段, 为保持数值稳定性并确保梯度收敛效果, RT PSPNet采用单精度(FP32)进行训练, 这与当前主流语义分割模型的常规做法一致。然而, FP32格式在推理时会对计算与内存带宽造成较大负担, 从而限制模型的实时性能。

为兼顾推理效率与精度保持, 本文在模型部署阶段引入半精度(FP16)推理机制。该策略可以在不影响分割精度的前提下, 显著提升模型的前向计算速度。具体而言, 将网络中的权重与激活值转换为FP16格式, 能够大幅减少显存占用与计算量, 从而充分发挥现代GPU张量核心的计算优势。

2 实验和结果分析

2.1 数据集介绍

本文构建了一个专门用于条状修补检测的语义分割数据集。原始图像通过路面采集车获取, 均为2 048像素×2 048像素的灰度JPEG文件, 内容为沥青路面及其上的条状修补病害^[21]。使用LabelMe软件对图像进行像素级的标注, 生成对应的PNG格式二值掩码标签, 最终形成一个包含1 000张图像-标签对的标准数据集。为扩充数据, 首先对500张初始数据集图像按8:1:1的比例进行划分, 再统一

逆时针旋转90°, 扩充得到训练集、验证集和测试集。值得注意的是, 为了避免数据泄露并保证数据增强的严谨性, 同一张图像及其增强版本只会统一被归纳至同一子集(如验证集), 而不会分布在其他不同的子集里, 以确保测试集是训练完的模型从未见过的图像, 有助于更科学地衡量各个模型对未知条状修补病害分割的鲁棒性和泛化能力。

该数据集任务定义清晰, 为前景(条状修补)与背景(沥青路面及其他物体)的二分类逐像素语义分割。然而, 任务面临显著挑战: 前景目标形态细长且不规则; 背景环境复杂, 常受光照变化、阴影、行道线、水渍及其他杂物干扰, 极易导致模型误分割。条状修补数据集部分示例如图6所示, 直观展示了其复杂性与分割难度。

2.2 实现环境

实验基于Windows 11操作系统, 使用Python 3.12.3作为编程语言, 深度学习框架为PyTorch 2.5.1, CUDA并行计算框架为12.6。实验的硬件测试环境采用12代Intel Core i7-12800HX, 频率为2.00 GHz, GPU为NVIDIA的RTX 4070, 显存为8 GB。本文实验使用adam优化器, 初始的学习率设置为 1×10^{-4} , 最小学习率设置为 $1 \times 10^{-4} \times 0.01$, 共训练40个epoch, 批次大小为2或4。

2.3 评价指标

为了全面评估语义分割模型的性能, 本文采用平均交并比(mean intersection over union, mIoU)、模型参数量、计算量和推理速度作为评价指标, 其中, mIoU作为模型的分割性能评价, 数值越高说明模型的分割效果越好; 模型参数量和计算量作为模型的轻量化程度评价, 数值越小说明模型轻量化的效果越好; 推理速度作为实时性评价, 数值越大说明实时性越好。mIoU计算式为:

$$mIoU = \frac{1}{k} \sum_{i=1}^k \frac{P_{ii}}{\sum_{j=1}^k P_{ij} + \sum_{j=1}^k (P_{ji} - P_{ii})} \times 100\% \quad (5)$$

其中, k 为包括前景和背景的总类别数量, 在本文中 k 为2, 包括条状修补和路面背景两类; P_{ii} 为正确的预测类别的像素数量, P_{ij} 和 P_{ji} 均为错误的预测类别的像素数量。

2.4 与基模型PSPNet的对比分析

基准PSPNet采用的是ResNet50特征提取主干, 同时它还设计了两种不同的深层特征提取方式, 即在对应层次同时改变卷积核的步长和膨胀系

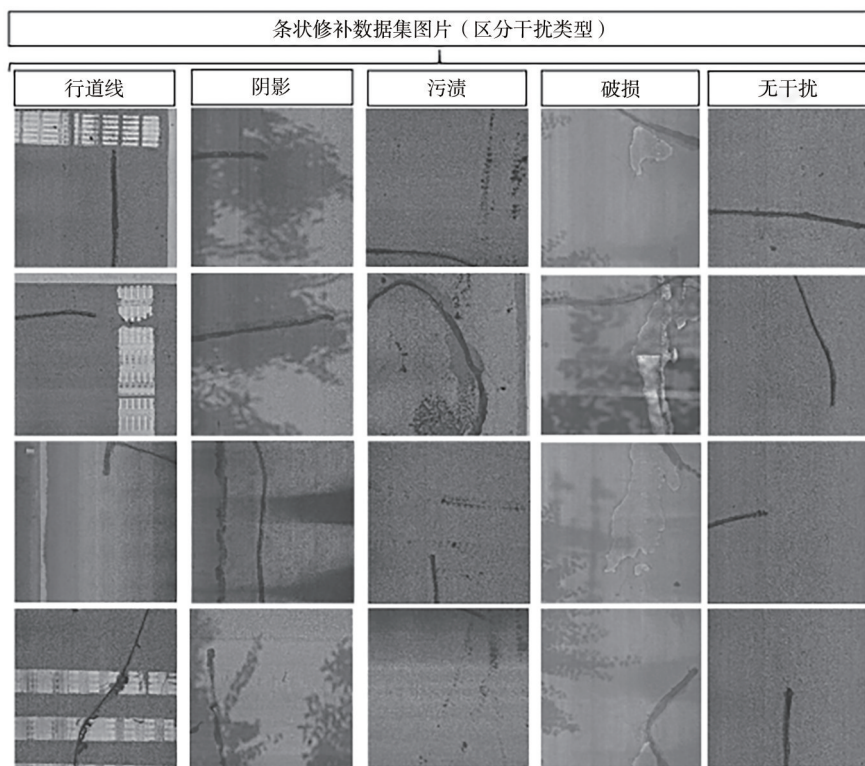


图6 条状修补数据集示例

数，以在不缩短特征图的长和宽的前提下，还能提取到全局特征。当下采样因子为8时，ResNet50的第8到第16个bottleneck的卷积核步长都为1，其中第8到第13个bottleneck卷积核的膨胀系数由1变为2，第14到第16个bottleneck卷积核的膨胀系数由1变为4；当下采样因子为16时，第8到第16个bottleneck的步长统一为1，然后将第8到第13个bottleneck步长变为2，且第14到第16个bottleneck卷积核的膨胀系数由1变为2。通常来说，当PSPNet的下采样系数为8时，其分割精度会更高，但是推理速度比下采样系数为16时更慢。由于本文的研究方向是路面病害条状修补的实时、高精度推理，因此在选择基准模型进行后续优化时，不仅需要选择推理速度较高的模型，而且要避免难以进行轻量化优化的结构（如结合Transformer架构的SegFormer模型等）。综合来看，PSPNet以较高的

分割精度和优秀的推理效率，成为本文进行轻量化改进和实时推理优化的一个理想对象。

同时，为了检验原PSPNet在使用不同主干情况下的轻量化可能性，本文还尝试将ResNet50主干替换为MobileNet v2主干，该主干在具有不错的特征提取能力的同时，还兼备了轻量化的优势。RT PSPNet和其他版本PSPNet在条状修补数据集上的各项参数对比见表2。

通过表2可以看出，PSPNet在使用默认主干ResNet50时，其参数量和计算量明显比使用MobileNet v2主干时更高，不过分割精度mIoU指标也相对更优，这表明轻量化主干虽然轻巧，但是在特征提取性能上还是略有欠缺。对比使用两种不同主干的原PSPNet，本文提出的RT PSPNet不仅拥有更高的分割精度、更小的参数量和次优的计算量，还拥有大幅领先的推理速度，这使其几乎在各个方

表2 RT PSPNet和其他版本PSPNet在条状修补数据集上的各项参数对比

模型	MIoU	参数量	推理速度/fps	计算量/GFLOPs
PSPNet-dsfactor=8 (MobileNet v2)	80.95%	2.376×10^6	110	6.031
PSPNet-dsfactor=16 (ResNet50)	84.70%	46.707×10^6	63	118.427
PSPNet-dsfactor=8 (ResNet50)	89.07%	46.707×10^6	30	369.468
RT PSPNet	91.76%	1.64×10^6	147	27.342

面都全面领先于采用默认主干的PSPNet模型，其分割精度更是远超使用MobileNet v2主干的PSPNet 10.81个百分点。值得注意的是，虽然MobileNet v2主干的计算量比本文提出的浅层轻量ResNet主干更低，但该网络层数更深卷积通道数更多，以及逐点卷积造成内存访问频繁等因素，导致在现有GPU的硬件环境下，浅层轻量ResNet即使计算量更大，也能够推理速度上显著领先于MobileNet v2，有利于RT PSPNet高效的实时性推理性能。对比以ResNet50为主干的PSPNet，RT PSPNet的mIoU提升了2.69~7.06个百分点，参数量减少了约96%，计算量减少了77%~93%，推理速度大幅提升了133%~390%，在实现理想轻量化和大幅提升推理速度的同时，分割精度也有了显著的提升。

2.5 与主流及轻量化模型的综合对比

为进一步评估RT PSPNet的性能，本节将其与主流轻量化语义分割模型进行对比。

2.5.1 与主流模型对比

从表3可见，在和主流的语义分割模型（FCN、DeepLabv3+、SegNet^[22]、HRNet和UNet++^[23]等）进行各项参数对比时，无论是在分割精度，还是在参数量和推理速度等方面，RT PSPNet显著优于绝大多数主流语义分割模型，在推理速度上甚至都是数倍或更高。

表3 RT PSPNet与主流语义分割模型的各项参数对比

模型	mIoU	参数量	推理速度/fps	计算量/ GFLOPs
SegNet (Vgg16)	78.42%	29.444×10^6	27	321.351
FCN (ResNet50)	83.13%	35.308×10^6	33	297.066
DeepLabv3+ (ResNet50)	86.07%	40.349×10^6	29	366.149
HRNet (w32)	87.71%	29.538×10^6	26	90.925
PSPNet (ResNet50)	89.07%	46.707×10^6	30	118.427
UNet++	92.30%	9.163×10^6	20	279.244
RT PSPNet	91.76%	1.64×10^6	147	27.342

在分割精度上，RT PSPNet高于绝大多数语义分割模型，如SegNet、DeepLabv3+和HRNet，仅次于UNet++，仅有0.54个百分点的差距；在参数量上，RT PSPNet比主流模型减少了82%~96%，不仅显著减小了模型的规模，还实现了非常理想的轻量化；在计算量上，RT PSPNet比主流模型减少了

67%~92%的计算量，显著缩减了计算开销，有利于模型运行速度的提升；在推理速度上，RT PSPNet以绝对优势（147 fps的极高帧数）领先于主流模型，比其他模型快了114~127 fps，甚至是对比模型中精度最高的UNet++的7.35倍；在参数量、计算量和推理速度全面领先于主流的语义分割模型的前提下，对条状修补的分割精度也达到了第一梯队，实现了理想的轻量化效果。

综上，RT PSPNet作为实时语义分割模型，在路面病害条状修补的语义分割任务上，各方面参数都明显优于大部分主流语义分割模型，尤其是推理速度和参数量，分割精度仅次于UNet++，明显高于对比的其他所有主流语义分割模型，具有显著的分割性能优势。

2.5.2 与轻量化模型对比

从表4可见，RT PSPNet在与多种轻量化分割模型（ENet^[24]、ICNet^[25]、BiSeNet^[26]、SCTNet等）进行各项参数对比时表现卓越。具体而言，RT PSPNet在精度（mIoU）和推理速度上均排名第一，其91.76%的mIoU显著优于其他模型，比历年轻量化SOTA模型SCTNet、BiSeNet和SegFormer分别高出6.83、18.71和0.84个百分点。尽管参数量（ 1.64×10^6 ）略高于ENet（ 0.349×10^6 ），但RT PSPNet在精度方面大幅领先了8.32个百分点，推理速度更是ENet的2.49倍。尽管模型计算复杂度较ENet、UNetFormer^[27]和SegFormer稍高，但是其依旧显著低于其他所有对比的轻量化模型，比计算量最高的BiSeNet降低了约77.6%的计算量。综上所述，RT PSPNet在路面条状修补的语义分割检测任务上具有优秀的分割精度和显著的轻量化效果，能够在“精度-速度”间取得较好平衡。

2.6 消融实验

为验证各改进模块的有效性，本节设计了系统消融实验。

2.6.1 主干网络轻量化

从表5可见，将ResNet50替换为浅层轻量ResNet后，在其他配置不变的情况下，尽管mIoU小幅下降了1.32个百分点，但参数量减少了94.6%，计算量降低了71.8%。这表明结构化剪枝能够在保留细节特征和实现轻量化之间达到较为优良的平衡。

2.6.2 解码器改进（特征融合与CPPM）

从表6可见，当RT PSPNet的解码器不加入特征融合以及改进的CPPM时，分割精度mIoU只有

表4 RT PSPNet与轻量化语义分割模型的各项参数对比

模型	MIoU	参数量	推理速度/fps	计算量/ GFLOPs
BiSeNet	73.05%	26.501×10^6	66	122.078
DeepLabv3+ (MobileVit)	75.57%	11.376×10^6	55	59.448
DeepLabv3+ (EdgeNext)	78.15%	4.662×10^6	70	46.157
ICNet	83.05%	26.243×10^6	56	73.931
ENet	83.44%	0.349×10^6	59	4.357
SCTNet	84.93%	16.287×10^6	105	28.105
UNetFormer	85.74%	11.682×10^6	89	23.476
HRNet (w18)	86.69%	9.637×10^6	26	37.319
SegFormer (b0)	90.92%	3.715×10^6	84	13.537
RT PSPNet	91.76%	1.64×10^6	147	27.342

表5 RT PSPNet的特征提取主干部分消融实验展示

浅层轻量 ResNet	MIoU	参数量	计算量/GFLOPs
×	84.70%	46.707×10^6	118.427
√	83.38%	2.502×10^6	33.387

83.38%，处于较低的水平；在将PPM替换为CPPM后，模型的参数量和计算量有较明显的下降，且mIoU也有小幅上升；当模型的解码器单独加入特征融合结构而保留PPM时，其mIoU为85.74%，精度改善明显，但参数量和计算量有小幅增加；当模型同时使用特征融合结构以及CPPM时，mIoU进一步提升到86.53%，且参数量和计算量也有一定的下降，这主要得益于CPPM更少的池化分支以及逐深度卷积的应用，使模型的轻量化程度和分割精度都有了提升。综上，RT PSPNet的解码器通过加入特征融合结构，以及改进原PPM为CPPM，使模型同时实现了轻量化和分割精度的改善。

表6 RT PSPNet的解码器部分消融实验展示

特征融合	CPPM	MIoU	参数量	计算量/GFLOPs
×	×	83.38%	2.502×10^6	33.387
×	√	85.82%	1.616×10^6	24.638
√	×	85.74%	2.593×10^6	35.886
√	√	86.53%	1.64×10^6	27.342

2.6.3 损失函数改进

从表7可见，使用单分割头&复合损失函数比单一损失函数精度更高，达到了88.51%；进一步采用多分割头&复合损失函数，精度达到最高的91.76%。这验证了多层次监督策略对提升模型判别能力的有效性。

表7 RT PSPNet的损失函数改进消融实验展示

损失函数	MIoU
单分割头&单一损失函数	86.53%
单分割头&复合损失函数	88.51%
多分割头&复合损失函数	91.76%

在模型的训练阶段，单独使用CE Loss函数分割精度是较低的，在引入Dice Loss函数后，损失值为CE Loss和Dice Loss二者之和，得到了单分割头&复合损失函数，RT PSPNet的模型分割精度有了明显的提升，增加了1.98个百分点；最后，当本文模型使用多分割头&复合损失函数进行训练时，模型的精度进一步提升，增加了3.25个百分点。综上，通过改进损失函数，RT PSPNet的分割精度最多可以提升5.23个百分点，有效改善了分割的效果。同时，本文也对MH_hybrid loss中CE Loss之和的值乘以不同的加权系数进行了消融实验，对比结果见表8。

表8 MH_hybrid loss中CE Loss之和的值乘以不同的加权系数精度对比

λ	MIoU
$\lambda=5$	91.64%
$\lambda=6$	91.76%
$\lambda=7$	91.55%

从表8可见，当MH_hybrid loss中的加权系数 $\lambda=6$ 时，RT PSPNet在条状修补数据集上的分割精度比加权系数 $\lambda=5$ 和 $\lambda=7$ 时分别提高了0.12个百分点和0.21个百分点，因此本文实验设置加权系数为6，这能更好地提升模型的分割效果。实验结果表明，使用本文提出的MH_hybrid loss函数进行RT PSPNet模型的训练，可以提升收敛效果，改善对条状修补的漏检和误检问题，并在一定程度上保证细长目标的分割边界完整性和条状修补的主体连贯性。

2.6.4 推理精度优化

从表9可见，训练使用FP32，推理时切换至FP16半精度，可在精度无损失的前提下，将推理速度提升75.5%（从147 fps到258 fps），显著增强了模型的部署实用性。因此，本文采用“训练时FP32、推理时FP16”的混合精度方案，在确保模型分割精度的同时，大幅提升了其在实际部署中的实时处理能力。

表9 RT PSPNet的模型推理精度改进的消融实验展示

方案	推理速度/fps	MIoU
FP32	147	91.76%
FP16	258	91.76%

2.7 可视化对比分析

为直观对比模型性能,本文通过特征可视化与预测结果对比进行分析。如图7所示,基于类激活映射图(class activation map, CAM)的特征可视化表明,原PSPNet因ResNet50主干网络层次过深、高层特征图尺寸过小,从而对细长条状修补的激活响应弱且不连续,预测结果出现显著漏分割。相比之下,RT PSPNet采用的浅层轻量ResNet主干保留了更高分辨率的特征图,其CAM热力响应与目标轮廓高度吻合,再经CPPM与多尺度特征融合优化后,最终分割结果更完整和准确。

图8进一步展示了RT PSPNet与多种主流及轻量化模型在复杂场景下的分割结果对比。在光照不均、阴影遮挡、路面污染等多种干扰条件下,RT PSPNet均表现出更优的鲁棒性:在第1行存在明显局部污染干扰的图像中,UNetFormer出现大面积误分割,RT PSPNet表现出更强的抗干扰能力;在第2行类似污染干扰条件下,仅RT PSPNet与HRNet保持了较好的分割完整性;在第3行目标颜色较浅、形态不规则的案例中,DeepLabv3+、

UNetFormer与BiSeNet均出现显著漏分割,RT PSPNet则较好地维持了目标连贯性;在第4~6行受路面阴影严重干扰的案例中,UNetFormer、SCTNet及BiSeNet均出现大量漏检和误检,RT PSPNet对阴影具有更好的判别能力,分割结果更完整。

3 结束语

针对当前条状修补路面病害语义分割研究中存在的公开数据集匮乏、主流模型参数量与计算量大、推理速度慢,以及对细长目标分割不完整导致精度受限等问题,本文构建了专用的条状修补语义分割数据集,并提出了一种轻量化高精度的实时语义分割模型——RT PSPNet。该模型融合浅层轻量ResNet主干、轻量化CPPM、多尺度特征融合机制以及多分割头复合损失函数,在自建数据集上取得了91.76%的mIoU,当参数量仅为 1.64×10^6 时,在FP32和FP16精度下模型的推理速度分别可以达到147 fps和258 fps,在保持高精度的同时显著提升了运算效率与部署适应性,为在边缘及低算力设备上实现路面病害实时精准检测提供了可行方案。

然而,本文工作仍存在若干局限:RT PSPNet在分割精度上虽优于大多数对比模型,但仍略低于UNet++;损失函数设计及计算效率仍有优化空间。综上,虽然RT PSPNet的语义分割性能综合平衡较优,但无法实现最理想的轻量化程度并同时达到最

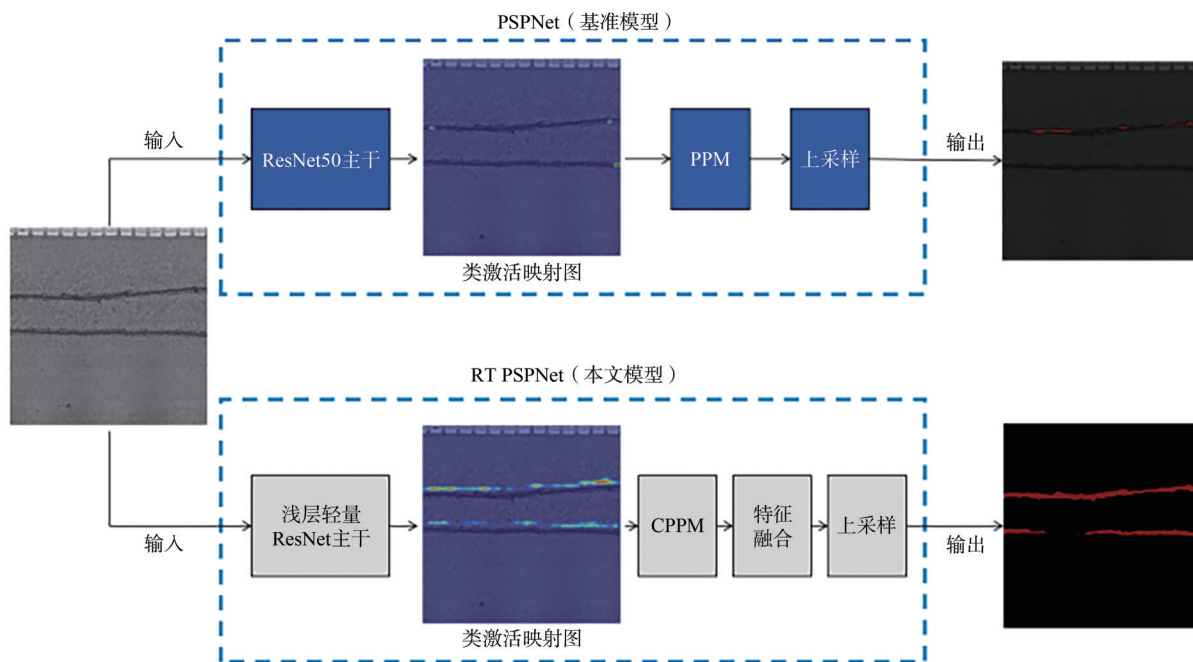


图7 PSPNet和RT PSPNet的特征可视化过程和分割结果对比

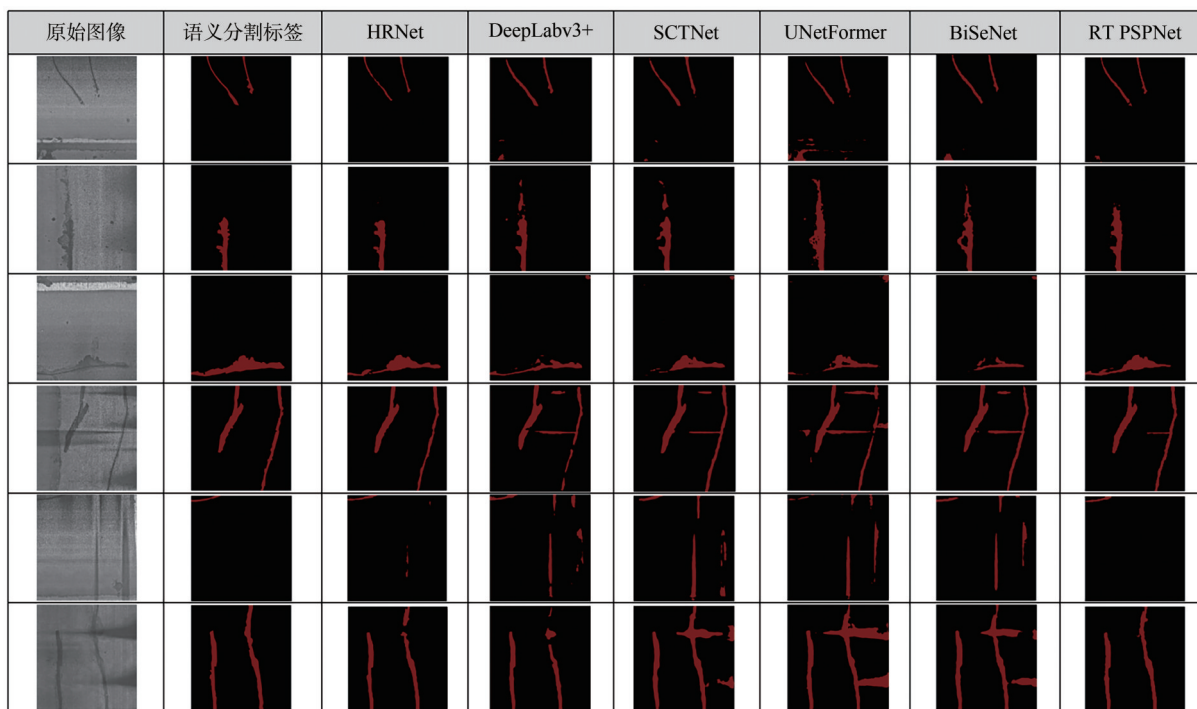


图8 RT PSPNet和其他主流语义分割模型的条状修补语义分割结果对比

高的分割精度。未来工作将在保持轻量化的基础上,进一步优化模型对细长目标的特征保持能力,探索更高效的损失函数与结构设计,并在更多的公共数据集与复杂场景下验证模型泛化能力,推动其在移动端与嵌入式设备中的实际部署。

参考文献:

- [1] 王双喜. 沥青混凝土路面病害分析及应对措施[J]. 黑龙江交通科技, 2011(11): 29-36.
Wang S X. Analysis of asphalt concrete pavement diseases and countermeasures[J]. Communications Science and Technology Heilongjiang, 2011(11): 29-36.
- [2] 王飞跃. 智能的本质与意义[J]. 智能科学与技术学报, 2026, 8(1): 1-11.
Wang F Y. On intelligence: the origin and the goal[J]. Chinese Journal of Intelligent Science and Technology, 2026, 8(1): 1-11.
- [3] 王飞跃, 缪青海, 张军平, 等. 探讨 AI for Science 的影响与意义: 现状与展望[J]. 智能科学与技术学报, 2023, 5(1): 1-6.
Wang F Y, Miao Q H, Zhang J P, et al. The DAOs to AI for science by DeSci: the state of the art and perspective[J]. Chinese Journal of Intelligent Science and Technology, 2023, 5(1): 1-6.
- [4] 严毅, 邓超, 李琳, 等. 深度学习背景下的图像语义分割方法综述[J]. 中国图象图形学报, 2023, 28(11): 3342-3362.
Yan Y, Deng C, Li L, et al. Survey of image semantic segmentation methods in the deep learning era[J]. Journal of Image and Graphics, 2023, 28(11): 3342-3362.
- [5] 田萱, 王亮, 丁琪. 基于深度学习的图像语义分割方法综述[J]. 软件学报, 2019, 30(2): 440-468.
Tian X, Wang L, Ding Q. Review of image semantic segmentation based on deep learning[J]. Journal of Software, 2019, 30(2): 440-468.
- [6] 徐文远, 李帅, 纪泳丞, 等. RTFormerF10: 基于RTFormer-slim模型融合红外与可见光图像的裂缝分割研究[J]. 中国公路学报, 2025, 38(4): 69-84.
Xu W Y, Li S, Ji Y C, et al. RTFormerF10: research on crack segmentation based on RTFormer-slim model fusion of infrared and visible light images[J]. China Journal of Highway and Transport, 2025, 38(4): 69-84.
- [7] 唐源, 董绍江, 刘超, 等. 基于改进SegFormer网络的路面裂缝分割算法[J]. 陕西科技大学学报, 2024, 42(3): 166-173.
Tang Y, Dong S J, Liu C, et al. An improved pavement crack segmentation algorithm based on SegFormer network[J]. Journal of Shaanxi University of Science & Technology, 2024, 42(3): 166-173.
- [8] 韩建峰, 张静, 宋丽丽, 等. 改进U-Net的无人机航拍路面破损检测方法[J]. 电子测量技术, 2024, 47(19): 137-145.
Han J F, Zhang J, Song L L, et al. Improved pavement damage detection method of UAV based on U-Net[J]. Electronic Measurement Technology, 2024, 47(19): 137-145.
- [9] 杨萍, 张汐. 改进DeepLabv3+的道路表面裂缝检测方法[J]. 计算机工程, 2025, 51(4): 262-268.
Yang P, Zhang X. Improved DeepLabv3+ road surface crack detection method[J]. Computer Engineering, 2025, 51(4): 262-268.
- [10] Yu C Q, Gao C X, Wang J B, et al. BiSeNet V2: bilateral network with guided aggregation for real-time semantic segmentation[J]. International Journal of Computer Vision, 2021, 129(11): 3051-3068.
- [11] Mehta S, Rastegari M. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer[PP]. V2. (2022-03-04) [2026-01-30]. arXiv: arXiv.2110.02178.
- [12] Xu Z Z, Wu D Y, Yu C Q, et al. SCTNet: single-branch CNN with transformer semantic information for real-time segmentation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2024, 38(6): 6378-6386.
- [13] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network[C]// Proceedings of the 2017 IEEE Conference on Computer Vision and

- Pattern Recognition (CVPR). Piscataway: IEEE Press, 2017: 6230-6239.
- [14] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 770-778.
- [15] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence. Piscataway: IEEE Press, 2016: 640-651.
- [16] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Computer Vision - ECCV 2018. Cham: Springer, 2018: 833-851.
- [17] Sun K, Zhao Y, Jiang B R, et al. High-resolution representations for labeling pixels and regions[PP]. V1. (2019-04-09)[2026-01-30]. arXiv: arXiv.1904.04514.
- [18] Xie E, Wang W, Yu Z, et al. SegFormer: simple and efficient design for semantic segmentation with transformers[C]//Advances in Neural Information Processing Systems. Massachusetts: MIT Press, 2021: 12077-12090.
- [19] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[PP]. V1. (2017-04-17)[2026-01-30]. arXiv: arXiv.1704.04861.
- [20] 王国忠, 陈明星, 姚辉, 等. 图像灰度处理对路面裂缝病害检测影响分析[J]. 市政技术, 2024, 42(4): 270-277.
- Wang G Z, Chen M X, Yao H, et al. Influence of image gray process on the detection of pavement cracks[J]. Journal of Municipal Technology, 2024, 42(4): 270-277.
- [21] Li X Y, Sun X F, Meng Y X, et al. Dice loss for data-imbalanced NLP tasks[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Kerrville: Association for Computational Linguistics, 2020: 465-476.
- [22] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[PP]. V3. (2016-10-10)[2026-01-30]. arXiv: arXiv.1511.00561.
- [23] Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N, et al. UNet++: a nested U-Net architecture for medical image segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer International Publishing, 2018: 3-11.
- [24] Paszke A, Chaurasia A, Kim S, et al. ENet: a deep neural network architecture for real-time semantic segmentation[PP]. V1. (2016-06-07)[2026-01-30]. arXiv: arXiv.1606.02147.
- [25] Zhao H S, Qi X J, Shen X Y, et al. ICNet for real-time semantic segmentation on high-resolution images[C]//Computer Vision - ECCV 2018. Cham: Springer, 2018: 418-434.
- [26] Yu C Q, Wang J B, Peng C, et al. BiSeNet: bilateral segmentation network for real-time semantic segmentation[C]//Computer Vision - ECCV 2018. Cham: Springer, 2018: 334-349.
- [27] Wang L B, Li R, Zhang C, et al. UNetFormer: a UNet-like transformer

for efficient semantic segmentation of remote sensing urban scene imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 190: 196-214.

[作者简介]



陈文柯 (1999-), 男, 西南林业大学大数据与智能工程学院硕士生, 主要研究方向为深度学习和语义分割。



李乐成 (1996-), 男, 西南林业大学林学院博士生, 主要研究方向为深度学习、交通预测和计算机视觉。



刘国玺 (1996-), 男, 西南林业大学林学院博士生, 主要研究方向为路面病害检测、森林资源监测和计算机视觉。



黄蕊 (1982-), 女, 西南林业大学大数据与智能工程学院副教授, 主要研究方向为业务过程、大数据挖掘与应用。



代飞 (1982-), 男, 西南林业大学大数据与智能工程学院教授、博士生导师, 主要研究方向为软件工程、服务计算、人工智能。