



基于双分支点流语义先验的路面病害分割模型

庞荣, 杨燕, 冷雄进, 张朋, 刘言

引用本文:

庞荣,杨燕,冷雄进,张朋,刘言. 基于双分支点流语义先验的路面病害分割模型[J]. *智能系统学报*, 2024, 19(1): 153–164.

PANG Rong, YANG Yan, LENG Xiongjin, et al. Segmentation model of pavement diseases based on semantic priori of double-branched point flow[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(1): 153–164.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202306037>

您可能感兴趣的其他文章

基于反馈注意力机制和上下文融合的非模式实例分割

Feedback attention mechanism and context fusion based amodal instance segmentation

智能系统学报. 2021, 16(4): 801–810 <https://dx.doi.org/10.11992/tis.202007042>

基于双注意力模型和迁移学习的Apex帧微表情识别

Apex frame microexpression recognition based on dual attention model and transfer learning

智能系统学报. 2021, 16(6): 1015–1020 <https://dx.doi.org/10.11992/tis.202010031>

双向特征融合与注意力机制结合的目标检测

Target detection based on bidirectional feature fusion and an attention mechanism

智能系统学报. 2021, 16(6): 1098–1105 <https://dx.doi.org/10.11992/tis.202012029>

基于注意力融合的图像描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

层次化双注意力神经网络模型的情感分析研究

Hierarchical double-attention neural networks for sentiment classification

智能系统学报. 2020, 15(3): 460–467 <https://dx.doi.org/10.11992/tis.201812017>

注意力机制和Faster RCNN相结合的绝缘子识别

Insulator recognition based on attention mechanism and Faster RCNN

智能系统学报. 2020, 15(1): 92–98 <https://dx.doi.org/10.11992/tis.201907023>

DOI: 10.11992/tis.202306037

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20240103.0911.002>

基于双分支点流语义先验的路面病害分割模型

庞荣^{1,2,3,4}, 杨燕^{1,2}, 冷雄进^{1,2}, 张朋^{3,4}, 刘言^{1,2}

(1. 西南交通大学计算机与人工智能学院, 四川 成都 611756; 2. 可持续城市交通智能化教育部工程研究中心, 四川 成都 611756; 3. 招商局重庆公路工程检测中心有限公司, 重庆 400067; 4. 国家山区公路工程技术研发中心, 重庆 400067)

摘要: 针对基于深度学习的真实路面病害图像识别算法主要面临的复杂道路背景与病害前景比例不同、病害尺度小等导致的类别严重不平衡、路面病害与道路的几何结构特征对比不明显导致其不易识别等问题, 本文提出一种基于双分支语义先验网络, 用于指导自注意力骨干特征网络挖掘背景与病害前景的复杂关系, 运用高效自注意力机制和互协方差自注意力机制分别对二维空间和特征通道进行语义特征提取, 并引入语义局部增强模块提高局部特征聚合能力。本文提出了一种新的稀疏主体点流模块, 并与传统特征金字塔网络相结合, 进一步缓解路面病害的类别不平衡问题; 构建了一个真实场景的道路病害分割数据集, 并在该数据集和公开数据集上与多个基线模型进行对比实验, 实验结果验证了本模型的有效性。

关键词: 语义先验信息; 高效注意力机制; 互协方差注意力机制; 稀疏主体点流; 类别不平衡; 语义分割; 路面病害; 深度学习

中图分类号: TP391 文献标志码: A 文章编号: 1673-4785(2024)01-0153-12

中文引用格式: 庞荣, 杨燕, 冷雄进, 等. 基于双分支点流语义先验的路面病害分割模型[J]. 智能系统学报, 2024, 19(1): 153-164.

英文引用格式: PANG Rong, YANG Yan, LENG Xiongjin, et al. Segmentation model of pavement diseases based on semantic priori of double-branched point flow[J]. CAAI transactions on intelligent systems, 2024, 19(1): 153-164.

Segmentation model of pavement diseases based on semantic priori of double-branched point flow

PANG Rong^{1,2,3,4}, YANG Yan^{1,2}, LENG Xiongjin^{1,2}, ZHANG Peng^{3,4}, LIU Yan^{1,2}

(1. School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu 611756, China; 2. Engineering Research Center of Sustainable Urban Intelligent Transportation, Ministry of Education, Chengdu 611756, China; 3. China Merchants Chongqing Road Engineering Inspection Center Co., Ltd, Chongqing 400067, China; 4. Mountain Highway Engineering Technology Research Center, Chongqing 400067, China)

Abstract: At present, the main problems faced by real road disease image recognition algorithms based on deep learning include serious imbalance in categories caused by different proportions of complex road background and foreground of diseases, and small disease scales. What's more, the inconspicuous contrast between pavement diseases and the geometric structure characteristics of roads leads to their difficulty in recognition. To address the above issues, we propose a semantic prior two-branch network to guide Transformer's backbone feature network in mining the complex relationship between background and foreground of pavement disease. It uses high-efficiency self-attention mechanism and cross-covariance image transformers(XCiT) to extract semantic features from two-dimensional space and feature channels, respectively, and a semantic locally-enhanced feed-forward (SLeff) module to improve the ability of local feature aggregation. We also propose a new sparse subject sampling point stream module, which is combined with the traditional FPN structure to further alleviate the category imbalance problem of pavement diseases. Finally, we constructed the road disease segmentation dataset based on real scene and compared it with multiple baseline models on this dataset and public dataset. The experimental results demonstrated effectiveness of this model.

Keywords: semantic priori information; efficient attention mechanism; cross-covariance image transformers attention mechanism; sparse subject sampling point flow; category imbalance; semantic segmentation; pavement diseases; deep learning

收稿日期: 2023-06-15. 网络出版日期: 2024-01-03.

基金项目: 国家自然科学基金项目(61976247); 国家重大研发计划项目(2019YFB-1310400); 重庆市技术创新与应用发展专项重点项目(CSTB2022TIAD-KPX0100); 重庆市交通科技自筹项目(CQJT20-22ZC05).

通信作者: 杨燕. E-mail: yyang@swjtu.edu.cn.

我国公路投资固定资产累计 10 万亿元, 公路总里程接近 520 万公里^[1]。完善的公路路网和交通体系为推动国家经济快速发展和加快城市化进程提供坚实的保障^[2]。但是目前车多路多的严峻

形势,对公路养护提出了更高的要求,其中裂隙和坑槽等道路病害严重威胁到行车安全和人民生命健康,及时发现道路病害能避免非人为因素,保障车辆驾驶安全^[3]。传统人工路面病害识别检测方法效率低下并且极具主观性,急需引入路面病害智能化高效检测,来加快智慧交通养护的发展,提高识别准确率的同时实现降本增效^[4]。

基于数字图像处理的路面病害识别算法需要先使用直方图均衡、灰度变换等技术进行预处理,再使用投影法或者边缘检测对图像进行分割,使用人工设计的卷积核进行特征提取,最后通过反向传播训练神经网络实现路面病害识别。基于机器学习的路面病害识别算法的特征提取需要依赖人工经验,通过机器学习图像特征提取算法自行挖掘特征表示,常见的算法有尺度不变特征变换算法和方向梯度直方图算法。但是传统机器学习算法是基于大量假设条件通过详细数学推导而设计出来的,不具有通用性。而基于深度学习的路面病害识别通过车载摄像头采集 RGB 路面图像,利用车载边缘端设备的数据传输服务将图片传输到云计算平台并利用深度学习图像算法进行识别,具有较强的鲁棒性。

虽然深度学习方法相对更加鲁棒,但是目前此类算法仍然存在许多难点:1)采集图像的质量和清晰度较低,路面病害与道路的几何结构特征对比不明显,影响算法识别率。2)路面病害尺度不一致且相对复杂的真实路面背景普遍较小,类别不平衡问题极其严重。3)目前大部分巡查设备的车载摄像头安装在车底并垂直地面进行病害采集,利用此方法采集的图像识别效果好,具有背景简单、病害相对背景占比大、清晰度高等特点,但是存在道路病害不完整,重复识别等问题。

针对上述问题,本文提出了一种基于双分支语义先验的裂隙、坑槽路面病害分割模型,可以对前置摄像头采集的低质量路面图像进行裂隙、坑槽病害分割。该网络模型由双重自注意力的语义先验模块、语义局部增强模块和稀疏主体采样点流模块组成,并采用双分支结构同步进行训练。其中,双重自注意力的语义先验模块运用高效自注意力机制和互协方差自注意力机制分别对图像的二维空间和特征通道进行语义特征的提取。语义局部增强模块利用深度可分离卷积对语义先验特征进行局部聚合,并利用全连接层进行升维和降维操作,切换特征空间维度进行特征聚合。稀疏主体采样点流模块可以融合跨层和本层语义特征并利用稀疏采样的方式对病害前景进行特征提取和传播,以缓解类别不平衡问题。

本文的主要贡献如下:1)提出一种基于双分支语义先验网络,并运用高效自注意力机制和 XCI_T 自注意力机制分别对二维空间和特征通道进行语义特征提取,以提高低质量复杂背景的病害识别率。2)引入语义局部增强 SLeff 模块提高局部特征聚合能力。3)提出了一种新的稀疏主体点流模块,并与传统特征金字塔网络相结合,进一步缓解路面病害的类别不平衡问题。4)构建了一个真实场景的道路病害分割数据集,并在该数据集和公开数据集上与多个基线模型进行对比实验,验证了模型的有效性。

1 路面病害识别相关工作

传统路面病害分割检测主要依赖数字图像处理技术,其过程一般包括以下步骤:1)定义每个图像像素的各种梯度特征;2)使用二值分类器判断像素点所属类别。路面病害分割检测算法总体可以分成基于阈值化,显著性和图像边缘 3 大类^[5-7]。Sheng 等^[8]提出一种基于梯度提升决策树的路面裂缝检测方法,该方法通过结合多层次特征来描述裂缝,并充分利用了裂缝的结构化信息,使用梯度提升决策树算法训练模型。Sun 等^[9]使用领域加权对病害附近像素进行加权计算,然后使用局部阈值化对加权结果进行精细化分割,最后为了解决病害连接问题将膨胀法运用在病害特征上,取得良好结果。曹建农等^[10]提出将均值漂移与直方图模式判别方法对已分块图像进行平滑处理,根据图像直方图模式特征判断有无病害,结合多方向搜索进行分割。然而上述方法都是基于人工设计特定特征提取模块,无法自适应不同病害环境,不具有通用性。

随着神经网络在计算机视觉领域的不断发展,高效智能化路面病害识别成为可能。基于深度学习的路面病害检测主要分为深度病害目标检测^[11-13]和深度病害语义分割^[14-16]两大类。这两类检测方法根据特征提取网络的不同又可以分为基于卷积神经网络和基于视觉 Transformer 的病害识别算法。实际项目中对路面病害识别采用目标检测的方法占主流,而基于语义分割的病害识别方法相对较少,因此具有较大的研究空间。

Zhang 等^[17]提出一种基于深度卷积神经网络的路面裂缝检测方法,将切割独立的小块图像输入神经网络进行分类,并将分类结果组合成一张概率图,其中高亮部分作为病害区域。韩静园等^[18]将全卷积神经网络用于路面裂缝分割,分析通道特征的关系,在特征金字塔网络的池化层加

入挤压和激励模块, 利用权重对原始特征在通道上进行重标定, 自适应地为裂缝边缘、图案和形状等特征分配权重。Yang 等^[19]提出一种用于路面裂缝检测的特征金字塔和层次增强深度卷积网络, 该网络以特征金字塔的方式将上下文信息集成到低层特征中进行裂纹检测。

由于 Transformer 所具备的长距离上下文建模能力和并行计算能力使其在自然语言处理领域取得巨大成功, 因此相关学者开始尝试将其应用在计算机视觉领域^[20]。Xie 等^[21]提出一种简单高效且鲁棒性强的语义自注意力分割模型 (segmentation transformer, SegFormer), 该模型由层次化自注意力编码器层和仅由几个全连接层构成的解码器 2 部分组成。Strudel 等^[22]提出的 Segmenters 是基于最新的视觉自注意力模型的研究成果, 将图像分割成块并将它们映射为一个线性嵌入序列, 用编码器进行编码, 再由 Mask Transformer 将编码器和类嵌入的输出进行解码, 解码器可以通过用对象嵌入代替类嵌入来直接进行全

景分割。Carion 等^[23]提出的基于自注意力的目标检测算法通过在解码器上附加一个掩码头并扩展到全景分割任务, 并获得有竞争力的结果。

2 基于双分支点流语义先验的复杂路面病害模型

模型整体框架如图 1 所示, 主要由自注意力层、语义先验层与稀疏主体采样点流模块组成。模型整体流程如下: 输入图像通过图像块划分层变成统一大小的图像块, 利用通过线性投射层和分层窗口自注意力模块对前置摄像头采集的真实路面病害图进行特征提取; 然后通过基于语义先验模块指导后续模块的特征更新方向; 通过稀疏主体点流模块对病害前景特征点进行采样, 并利用采样点进行训练。最后利用双分支训练结构同步完成自注意力层和语义先验层的更新优化。另外, 本文利用语义局部增强模块对语义先验层输出的语义特征张量在更高维度的特征空间进行局部语义聚合。

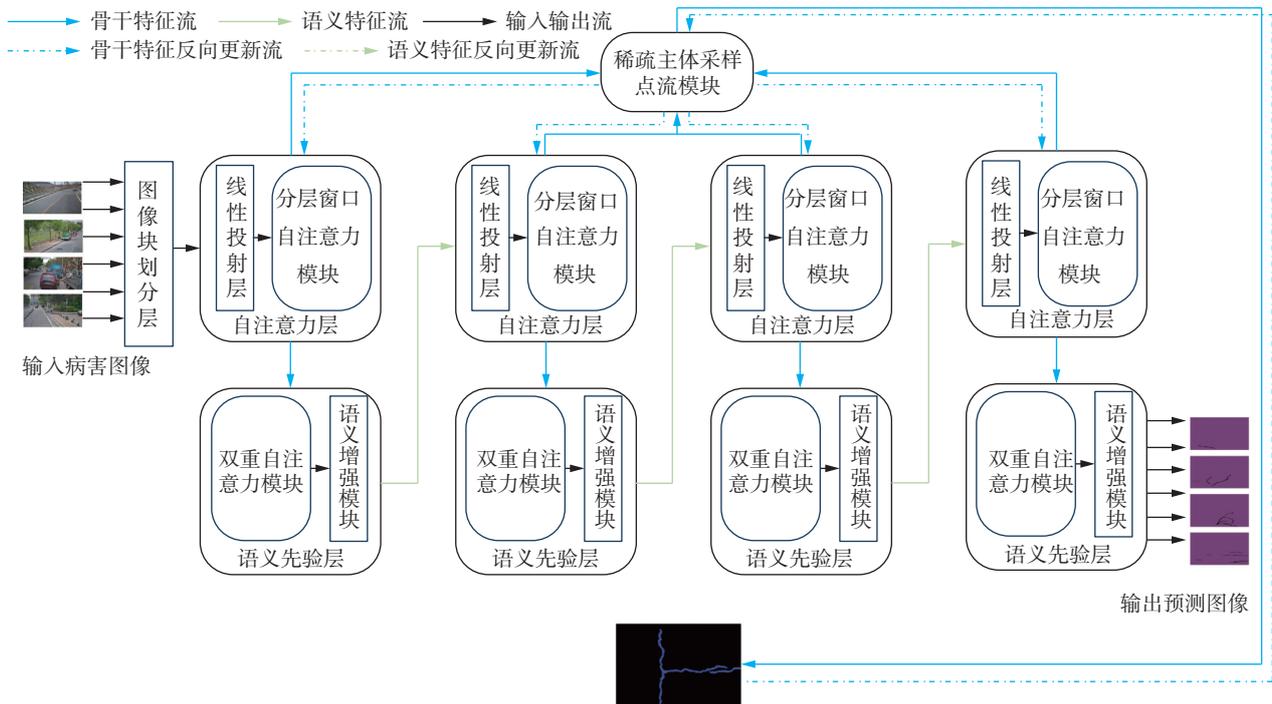


图 1 基于双分支点流语义先验的复杂路面病害(裂隙、坑槽)分割模型框架

Fig. 1 Segmentation model diagram for complex road surface diseases (cracks, potholes) based on semantic priori of double branched point flow

2.1 基于双重注意力机制的语义先验模块

大部分语义分割任务都是对已经预训练好的骨干网络进行微调, 以适配当前任务。但是直接对预训练好的骨干网络进行微调缺乏语义指导, 忽略了整体网络在编码阶段提供的语义上下文信息。基于此, 本模型引入 Semask 模型^[24]的思想,

在特征提取骨干网络中添加图像语义信息, 每个先验语义模块都为后续特征提取模块的微调提供指导意义, 使得整体网络的语义优化朝着更加合理的方向。具体来说, 本模型设计了一个新的语义解码器对中间语义先验信息提供监督, 使用基于双重自注意力的语义先验模块, 用来捕获全局

整体图像特征在空间和通道 2 个维度的关系。对于图像空间级别的特征语义提取, 本文采用高效自注意力机制^[25]; 对于特征通道维度级别的特征提取; 本文采用 XCiT 自注意力机制^[26]。相较于传统的注意力计算方式, 同时对这 2 个维度分别建模可以获得更多的上下文信息和关系。基于双重自注意力的语义先验模块整体结构图 2 所示。

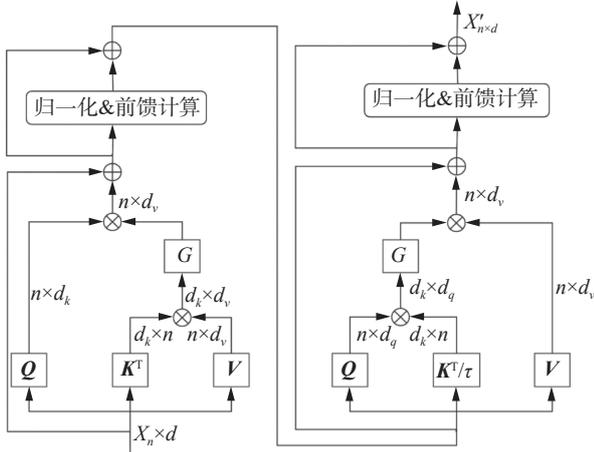


图 2 基于双重注意力的语义先验模块整体结构

Fig. 2 The overall structure diagram of semantic prior modules based on dual attention

双重自注意力模块中的高效注意力机制的运算方式为

$$D(Q, K, V) = \rho_q(Q) (\rho_k(K^T) V) \quad (1)$$

式中: 张量 $Q \in \mathbf{R}^{n \times d_q}$, 张量 $K \in \mathbf{R}^{n \times d_k}$, 张量 $V \in \mathbf{R}^{n \times d_v}$, ρ_q 、 ρ_k 分别表示张量 Q 和张量 K 沿矩阵每一行和每一列应用 Softmax 归一化函数。高效注意力改变传统注意力机制中间特征矩阵的乘法顺序, 将归一化的张量 K 与张量 V 相乘, 再将生成的全局上下文向量与归一化的张量 Q 相乘生成新的语义特征。其具有线性的时间和空间复杂度, 同时与传统的自注意力机制有相同的表达能力, 使用矩阵乘法的结合律来交换 2 个矩阵乘法的顺序, 大大减小注意力机制运算次数, 从 $n \times n$ 减小到 $d_k \times d_v$ 。

由于格拉姆矩阵 (Gram) 和协方差矩阵的特征谱非零部分等价, 因此 Gram 矩阵和协方差矩阵能够互相用对方的特征向量表示, 而原始自注意力计算过程可以看成类似 Gram 矩阵的计算过程。因此对于特征通道维度的语义信息提取, 本文抛弃了传统沿图像块维度进行自注意力矩阵乘法的计算方式, 改用基于互协方差自注意力机制来重新对特征维度进行建模, 进一步提高运算效率, 减少运算次数。互协方差自注意力机制利用特征通道之间的注意力操作替代图像块维度之间显示完全成对交互计算, 将图像块维度投影的键标准化张量 K 与图像块维度投影的查询标准化张

量 Q 做乘法操作, 得到关于特征通道的互协方差矩阵, 并将该矩阵与值张量 Q 相乘得到特征通道的注意力张量。最后使用分组自注意力机制, 对输入特征在通道维度进行分组, 只对位于对角线的通道分组应用互协方差注意机制, 避免对所有通道进行计算降低运算复杂度。值得注意的是, 为了使得互协方差矩阵权重分布更加合理, 互协方差自注意力机制引入了可学习的温度参数 τ 。基于通道特征的互协方差自注意力公式为

$$XA(Q, K, V) = VC_T(K, Q)C_T = \text{soft max} \left(\tilde{K}^T \cdot \frac{Q}{\tau} \right) \quad (2)$$

综上所述, 整个语义先验模块的完整流程为

$$\begin{aligned} E_{A_{res}} &= EA(Q_s, K_s, V_s) + I \\ S_{A_{res}} &= \text{MLP}_1(\text{LN}(E_{A_{res}})) + E_{A_{res}} \\ X_{A_{res}} &= XA(Q_c, K_c, V_c) + S_{A_{res}} \\ C_{A_{res}} &= \text{MLP}_2(\text{LN}(X_{A_{res}})) + X_{A_{res}} \end{aligned} \quad (3)$$

式中: EA 为二维图像高效注意力操作, XA 为通道 XCiT 注意力操作, $E_{A_{res}}$ 和 $X_{A_{res}}$ 分别为注意力计算中间结果, LN 为层标准化。

2.2 融合局部增强信息的语义先验模块

虽然 Transformer 网络架构能够克服卷积的归纳偏置所带来的局限性, 但是其训练需要大量数据集^[27], 其模型无法通过简单几层网络轻易捕捉本文所构建的小样本路面病害图像的几何低维拓扑结构信息。受到 CeiT 模型的启发^[28], 本文在语义先验模块中引入语义局部增强 SLeff 结构对语义信息进行局部聚合操作, 该模块的详细设计如图 3 所示。

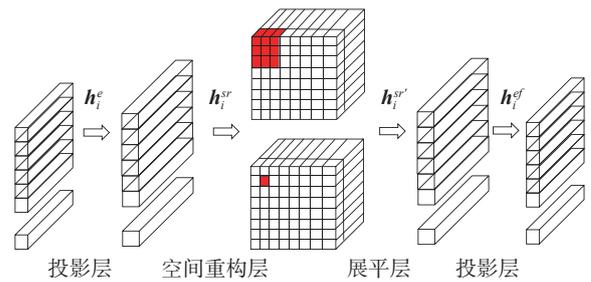


图 3 语义局部增强模块结构

Fig. 3 The structure diagram of semantic local enhancement module

假设输入图像通过第 i 层特征提取网络和双注意力语义先验模块得到语义特征向量 $h_i \in \mathbf{R}^{n \times c}$, 该向量首先经过线性投影层将维度扩展到 $n \times e$, 然后根据图像原始 Token 的排列顺序将二维语义向量 $h_i^e \in \mathbf{R}^{n \times (c \times e)}$ 重组为三维张量 $h_i^{sr} \in \mathbf{R}^{\sqrt{n} \times \sqrt{n} \times (c \times e)}$ 。为了聚合局部特征, SLeff 模块利用深度可分离卷积将新重组的特征矩阵进行卷积运算, 最后将特征向量展平, 并经过一个线性投影层恢复成与输入大小相同的语义特征向量 $h_i^{ef} \in \mathbf{R}^{n \times (c \times e)}$, 其中

n 表示中间特征语义向量的行数, c 表示特征向量通道的维度。语义局部增强模块 SLeff 流程为

$$\begin{cases} \mathbf{h}_i^e = \text{GELU}(\text{BN}(\text{Linear1}(\mathbf{h}_i))) \\ \mathbf{h}_i^{sr} = \text{SpatialRestore}(\mathbf{h}_i^e) \\ \mathbf{h}_i^{sr'} = \text{GELU}(\text{BN}(\text{DWConv}(\mathbf{h}_i^{sr}))) \\ \mathbf{h}_i^{ef} = \text{Flatten}(\mathbf{h}_i^{sr'}) \\ \mathbf{h}_i^f = \text{GELU}(\text{BN}(\text{Linear2}(\mathbf{h}_i^{ef}))) \end{cases} \quad (4)$$

式中: GELU 为 GELU 激活函数, BN 表示层归一化函数, Linear1 和 Linear2 表示线性投影层, SpatialReStore 表示像素重排函数, DWConv 表示深度可分离卷积函数, Flatten 表示特征矩阵展平函数。

2.3 稀疏主体点流采样模块

基于自注意力的语义分割模型需要图像中的每个像素点与全局像素点进行交互计算, 但对于复杂背景占比较大的路面病害图, 如果直接进行全局关系的建模就会造成大量计算的冗余, 无法充分挖掘小面积前景的特征表示。此外, 对于目前主流的特征提取网络, 随着网络层数的增加语义特征包含的信息逐渐抽象和高级, 但是语义特征分辨率却在逐渐减小, 该特性会导致小面积的前景特征随着网络层数的增加而逐渐消失, 不利于病害的识别。本文提出的稀疏主体点流模块通过在当前 FPN 层采样部分病害前景特征像素点作为重点关注对象, 只将该部分像素点与全局像素点进行建模。同时该模型将相邻层特征进行下采样融合, 利用下层高分辨率特征的具象语义信息补充本层低分辨率的抽象语义信息。针对传统点流方法没有考虑融合本层特征, 本模块又对采样点所在层进行采样点自注意力聚合, 并将其与跨层特征相结合, 真正意义上完成了不同阶段的特征融合。稀疏主体点流采样模块通过稀疏注意力机制将相邻特征层的信息相关联, 避免大量背景噪声点参与全局建模, 减少运算次数, 提高运算效率, 缓解类别严重不平衡问题。同时跨层特征融合将不同层的分辨率信息与语义信息进行互补, 兼顾高分辨率信息和高语义信息的相互统一, 提高对复杂背景下小面积病害的识别能力^[29], 具体原理和操作如图 4 所示。

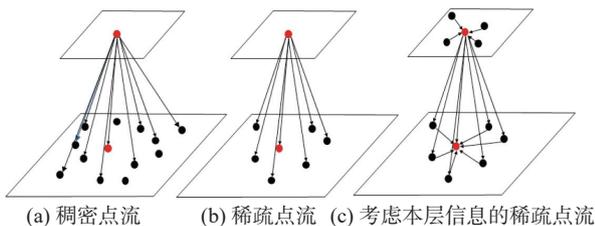


图 4 考虑本层节点聚合信息的稀疏主体点流优化

Fig. 4 Sparse subject point flow optimization graph considering aggregation information of local nodes

假设图片经过编码骨干网络得到 4 组不同的中间特征张量 t_i 。首先将本层语义特征张量 t_i 与相邻层特征张量 t_{i-1} 相叠加, 叠加后的特征通过卷积操作得到融合低层具象特征和高层抽象特征的主体特征选择信息流张量 c_i 。然后对原始特征图的主体特征选择信息流张量 c_i 的非 0 特征值点与对应二维坐标相映射, 并根据该映射完成对特征张量 t_i 的更新。为了达到稀疏特征点的目的, 选择更新后语义特征张量 t_i 中特征值大小前 top- n 的点作为兴趣点, 并根据兴趣点的坐标信息进行采样, 最终得到稀疏语义特征向量 t'_i 。根据 FPN 结构的相邻特征大小倍率关系, 可以计算出语义特征张量 t_{i-1} 中对应主体特征选择信息流张量 c_i 的非 0 特征值点集合的二维坐标。为了关联相邻 FPN 网络的特征, 进一步提高其表达能力, 还将本层稀疏采样点之间进行自注意力操作, 通过 2 次矩阵乘法操作计算不同层之间语义特征的注意力矩阵, 并将该矩阵与当前层稀疏语义特征向量 t'_i 相乘后求和, 得到最终更新完成的全局稀疏语义特征张量 t''_i 。最后根据语义特征向量 t_{i-1} 的采样坐标将对应位置的 t''_i 值插入到 t_{i-1} 中 (对应图 5 扩散操作), 进而完成语义特征向量 t_{i-1} 的更新操作, 最终得到 $i-1$ 层融合了 i 层语义信息的特征语义向量 t''_{i-1} 。稀疏主体点流结构如图 5 所示。

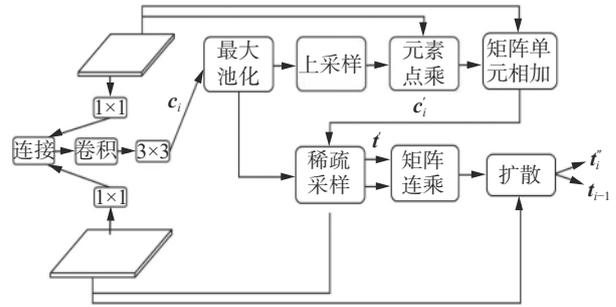


图 5 稀疏主体点流结构

Fig. 5 Sparse subject point flow structure diagram

稀疏主体点流模块具体操作流程为

$$\begin{cases} \mathbf{c}_i = \text{Conv}(\text{Concat}(\mathbf{t}_{i-1}, \mathbf{t}_i)) \\ \mathbf{c}'_i = \text{Interpolation}(\text{AdaptiveMaxPool2d}(\mathbf{c}_i)) \\ \mathbf{t}'_i = \text{Update}(\mathbf{t}_i, \text{Sample}(\mathbf{c}'_i)) \\ \mathbf{t}'_{i-1} = \text{Sample}(\mathbf{c}'_i) \\ \mathbf{self_atten} = \text{relu}(\text{Bmm}(\mathbf{t}'_i{}^T, \mathbf{t}'_i)) \\ \mathbf{cross_atten} = \text{soft max}(\text{Bmm}(\mathbf{t}'_i{}^T, \mathbf{t}'_{i-1})) \\ \mathbf{t}''_i = \mathbf{t}'_{i-1} + \text{Bmm}(\mathbf{self_atten} + \mathbf{cross_atten}, \mathbf{t}'_i) \\ \mathbf{t}''_{i-1} = \text{Scatter}(\mathbf{t}''_i, \mathbf{t}'_{i-1}) \end{cases} \quad (5)$$

式中: Interpolation 表示插值操作, AdaptiveMaxPool2d 表示自适应最大池化操作, Update 表示根据采样点进行更新操作, Bmm 表示矩阵乘法操

作, Scatter 表示扩散操作。

3 基于双分支点流语义先验的路面病害分割模型实验分析

3.1 数据集和指标

实验使用公开数据集 Crack500 和采集自自研车载自动化巡查项目的路面病害语义分割两种数据集。其中, Crack500 数据集包含训练集 1 896 张、验证集 348 张、测试集 1 124 张图像。而自研巡查车路面病害数据集通过前置摄像头录制路面视频并传输到边缘服务器上, 并隔固定时间间隔进行帧采样。该数据集收集了从 2022 年 3 月-9 月重庆部分路面病害数据, 每张图片的分辨率规整为 1080×720 的 PNG 格式图像, 一共包含 1000 张路面病害图, 训练集验证集的比例为 3:1, 包含 3 个语义类别, 分别为横向裂缝、纵向裂缝和坑槽。本实验采用的语义分割评价指标包括平均交并比 (mIoU)、平均准确率 (mPrecision)、召回率 (recall)、 F_1 得分和参数量 (param)。各项指标表示如下:

$$m_{IoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}}$$

$$m_{Precision} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \quad (6)$$

$$r_{recall} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ji}}$$

$$F_{1-score} = \frac{1}{k+1} \frac{2 \times m_{Precision} \times r_{recall}}{m_{Precision} + r_{recall}}$$

式中: k 表示种类数 (不包含背景), p_{ii} 、 p_{jj} 、 p_{ij} 和 p_{ji} 分别表示预测正确属于该病害、预测正确不属于该病害、预测该类病害不属于该类的假类和预测不属于该类的病害属于类 (如果第 i 类为正类)。本实验如果在 mIoU 相差不大的情况下, 即

使出现 mIoU 偏低的情况, 也着重观察平均准确率、召回率及 F_1 得分, 原因如图 6 所示, 由于人工对低质量路面病害图仅仅靠肉眼观察后打标不可能做到像素级别, 而所提供数据集的 Ground Truth 相较于预测结果而言反而有漏打少打的情况, 因此仅仅依靠计算 mIoU 会出现效果好的模型反而 mIoU 较低的情况, 无法说明模型的准确性, 因此着重考虑后面 3 个指标。

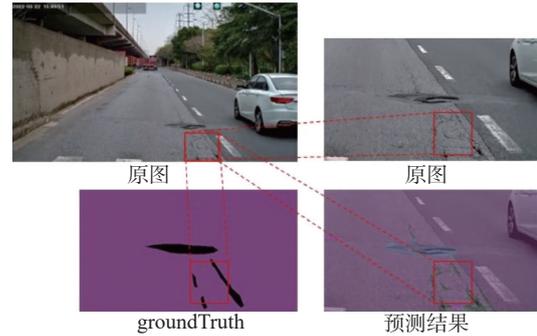


图 6 人工路面病害打标 Ground Truth 与模型预测结果对比可视化

Fig. 6 Visualization of comparison between Ground Truth and model prediction results for artificial pavement disease marking

所有实验均在 Pytorch-GPU-1.8.1 和 Python-3.7 上实现, 并在单张 NVIDIA GeForce RTX 3090 (显存 24 G) 上运行。为了实验的公平, 所有实验模型均采用基准学习率为 $\gamma_0=0.001$ 的 AdamW 优化器, 并采用 poly 学习率衰减, 设置训练迭代次数为 80 000, 批处理大小为 4。

3.2 实验结果对比

本文在上述 2 个数据集上将所提出模型与其他语义分割模型进行比较, 如表 1 和表 2 所示。比较方法分为 2 类, 基于 CNN 的语义分割算法和基于视觉 Transformer 的语义分割算法, 基于卷积网络包括以 ResNet 为骨干网络的 Ccnet^[30]、Deeplabv3^[31]、Deeplabv3+^[32] 及 Hnet 系列^[33] 和 Cgnet^[34]; 基于视觉 Transformer 的分割模型包括以 ViT 为骨干网络的 Setr^[35]、Segmenter^[22] 和 Segformer^[21]。

表 1 不同语义分割模型在自研路面病害数据集上的结果对比

Table 1 Comparison of results of different semantic segmentation models on self-developed road surface disease datasets

方法	参数量/M	浮点运算次数	平均交并比/%	平均准确率/%	召回率/%	F_1 得分/%
Ccnet-ResNet50	49.83	451.57	46.47	58.16	56.67	57.34
Ccnet-ResNet100	71.42	638.69	37.42	55.74	58.39	58.71
Deeplabv3-ResNet50	68.10	607.36	15.82	49.78	16.97	38.78
Deeplabv3-ResNet101	87.09	782.63	5.89	49.38	6.36	19.07
Deeplabv3+-ResNet50	43.59	405.67	47.25	60.84	56.34	58.29

续表 1

方法	参数量/M	浮点运算次数	平均交并比/%	平均准确率/%	召回率/%	F_1 得分/%
Deeplabv3+-ResNet101	62.58	584.62	46.23	63.16	54.14	57.21
Hrnet18	9.64	37.31	1.04	25.06	19.76	1.04
Hrnet48	65.95	95.43	6.62	13.41	12.18	14.36
Cgnet	58.74	527.86	42.32	54.31	53.66	51.79
Setr-Tiny	10.27	18.58	17.87	20.69	12.85	20.36
Setr-Base	92.35	678.01	27.67	30.63	30.37	30.23
Segmenter-tiny	26.03	144.93	19.79	16.36	33.33	24.57
Segmenter-small	25.33	454.78	23.17	14.33	28.74	4.02
Segformer	13.72	12.27	7.39	13.88	13.62	15.22
本文方法	60.68	209.61	40.12	63.96	59.91	69.13

表 2 不同语义分割模型在 Crack500 公开数据集上的结果对比

Table 2 Comparison of results of different semantic segmentation models on Crack500 public dataset

方法	参数量/M	浮点运算次数	平均交并比/%	平均准确率/%	召回率/%	F_1 得分/%
Ccnet-ResNet50	39.84	404.92	51.77	56.37	57.62	52.97
Ccnet-ResNet100	62.32	558.79	51.82	54.29	55.81	55.38
Deeplabv3-ResNet50	57.79	578.47	55.79	58.72	51.94	60.79
Deeplabv3-ResNet101	77.67	739.58	55.37	50.13	49.10	48.30
Deeplabv3+-ResNet50	34.88	369.19	58.49	56.28	57.00	59.94
Deeplabv3+-ResNet101	52.18	529.77	59.31	57.27	58.92	60.17
Hrnet18	8.78	31.65	16.64	22.30	18.66	21.99
Hrnet48	54.73	95.43	16.88	25.79	22.07	28.36
Cgnet	49.60	492.55	59.71	59.38	48.67	51.79
Setr-Tiny	8.62	16.78	28.24	30.45	27.63	32.36
Setr-Base	77.84	622.41	31.37	29.38	35.53	30.23
Segmenter-tiny	18.97	111.28	24.74	21.89	28.46	23.52
Segmenter-small	17.33	391.93	29.14	29.31	27.83	25.77
Segformer	9.72	10.59	25.96	25.01	22.37	27.39
本文方法	49.86	151.61	65.72	76.77	69.34	72.78

在 2 个数据集上, 基于 Transformer 的语义分割模型的效果要远差于 CNN, 是因为自注意力网络训练需要大规模数据集才能充分挖掘全局关系, 而 2 个数据集中训练集数量远远不足。但是, 本模型使用 Transformer 骨干特征网络却在 2 个少量样本的数据集上比所有卷积方法在 Flops、mPrecision、Recall 和 F_1 -score 4 个重点指标上都高, 说明了该算法在训练少样本的路面病害样本上具有巨大潜力。同时可以观察到在自研数据集上, 本文方法的 mPrecision 指标与 Deeplabv3+ 接近; 但是 Flops 远低于该方法, 说明本文方法在相同精准度的情况下具有更快的运算速度; 同理, 本文方法与基于 ResNet101 的 Ccnet 模型在 Recall 指标接近, 但是 Flops 减少了接近 2 倍, 进一

步说明本文模型的有效性(注: 虽然 Hrnet18 网络的参数量仅 9.64 M, 但是本文不将其作为最佳参数量结果的原因是该网络表现其他指标极差, 没有参考意义。同理认为所有模型中浮点运算次数最低的为本文的模型, 而不是 Segformer)。而在公开数据集上, 由于数据集质量大幅度提高, 所有方法的指标相较于自研数据集都有大幅度提升, 但是本文所提模型表现大幅优于基于所有对比方法, 进一步验证了本模型在高质量语义分割数据集上更具鲁棒性和有效性。

表 3 给出不同大小的 Swin-Transformer 骨干网络细节参数。在表 4 中本文对 Swin Transformer 3 种不同编码器变体进行实验, 其中 Swin-T-FPN 表示利用 Swin-Tiny 作为编码器, FPN 作为解码器

展示使用不同骨干相同网络效果对比。变体具体参数如表 3 所示, 基于 Swin-Transformer 的 tiny 版本的骨干网络在所有指标表现最好; 相反, 使用更大版本的 Swin-Transformer 特征骨干网络反而会影

响模型的精度, 在各项指标中都有大幅度的下滑。原因是更大的模型虽然拟合效果更好, 但是由于实验使用的数据集质量较低, 数量较少, 针对小数据集, 大模型更容易出现欠拟合的现象。也就是说, 相同训练参数配置的下, 大模型可能会一直陷入一个局部次优解, 无法对模型全局参数有个更好的更新。

表 3 不同大小的 Swin-Transformer 骨干网络细节参数表
Table 3 Detailed parameter table of Swin-Transformer backbone network of different sizes

主干网络	窗口大小	特征维度	块数量	头数量	参数量/M
Swin-T	7	[96,192,384,768]	[2,2,6,2]	[3,6,12,24]	28
Swin-S	7	[96,192,384,768]	[2,2,18,2]	[3,6,12,24]	50
Swin-B	12	[128,256,512,1024]	[2,2,18,2]	[4,8,16,32]	88

表 4 Dual-SePointFlow 模型使用不同大小 Swin-Transformer 骨干网络效果对比

Table 4 Comparison table of effects of Dual-SePointFlow model using Swin-Transformer backbone networks of different sizes

方法	参数量/M	浮点运算数	平均交并比/%	平均准确率/%	召回率/%	F_1 得分/%
Dual-SePointFlow-T	66.12	209.61	69.91	42.63	59.91	69.13
Dual-SePointFlow-S	82.05	259.51	64.61	40.39	53.77	41.64
Dual-SePointFlow-B	115.37	343.98	36.40	37.18	43.52	43.78

3.3 消融实验

在本节实验中, 将对双分支训练结构、语义先验模块和稀疏点流主体模块进行消融实验, 分析本文设计的模块对于真实复杂路面病害识别的有效性。表 5 展示了 3 种结构的消融结果, 最简单的 Swin-Transformer+FPN 模型出人意料的召回率最高, 表明该方法查全率高, 但是查准率非常低, 综合两者的 F_1 得分非常低; 原始的 semask 模型相较于第 1 个模型添加了最简单的语义先

验模块, 其模型牺牲较多的查全率去提高查准率, 但是效果甚微, F_1 得分非常低, 效果不理想。在原始 semask 基础上添加稀疏点流模块, 查准率小步上升, 但是召回率进一步下降。而采用新设计的双重语义先验模块牺牲较少查准率的同时大幅度提高查全率, F_1 得分也得到显著提高, 证明了双重语义先验模块的有效性。最后结合局部增强模块得到最好的结果, 验证该模块的有效性。

表 5 Dual-SePointFlow 不同模块消融验证实验效果对比
Table 5 Comparison of ablation validation experiments on different modules of Dual-SePointFlow

方法	平均交并比/%	平均准确率/%	召回率/%	F_1 得分/%	参数量/M
Swin-T-FPN(No Semantic Attention)	43.68	38.14	63.23	43.68	32.61
Semask-T-FPN(Original Semantic Attention)	46.91	40.47	58.73	44.94	62.43
Semask-SePointFlow-T-FPN(Original Semantic Attention)	42.68	45.66	54.97	47.36	64.57
Dual-SePointFlow-T(Dual Attention)	44.32	60.97	58.41	66.74	60.68
Dual-SePointFlow-T(Dual Attention+LeFF)	40.12	63.96	59.91	69.13	66.12

3.4 可视化结果分析

本节主要展示在自研车载自动化巡查项目的路面病害语义分割数据集上的可视化结果分析。图 7 给出了不同经典语义分割模型的实验结果对照。基于卷积的分割模型中 Cgnet 预测的细小细节最多, Ccnet 和 Hrnet 基本上无法预测细小裂缝细节。而对于较大的裂缝, 所有卷积模型

表现良好。基于 Transformer 的分割模型对于细小裂缝识别总体表现很差, 但是对于较大裂缝的识别效果比卷积更好, 预测结果更加完整。总体而言, 在小数据集上基于卷积的分割模型效果更胜一筹。而本文方法效果最好, 基本上准确预测了大坑槽和细小的裂缝细节, 证明了本文模型的有效性。

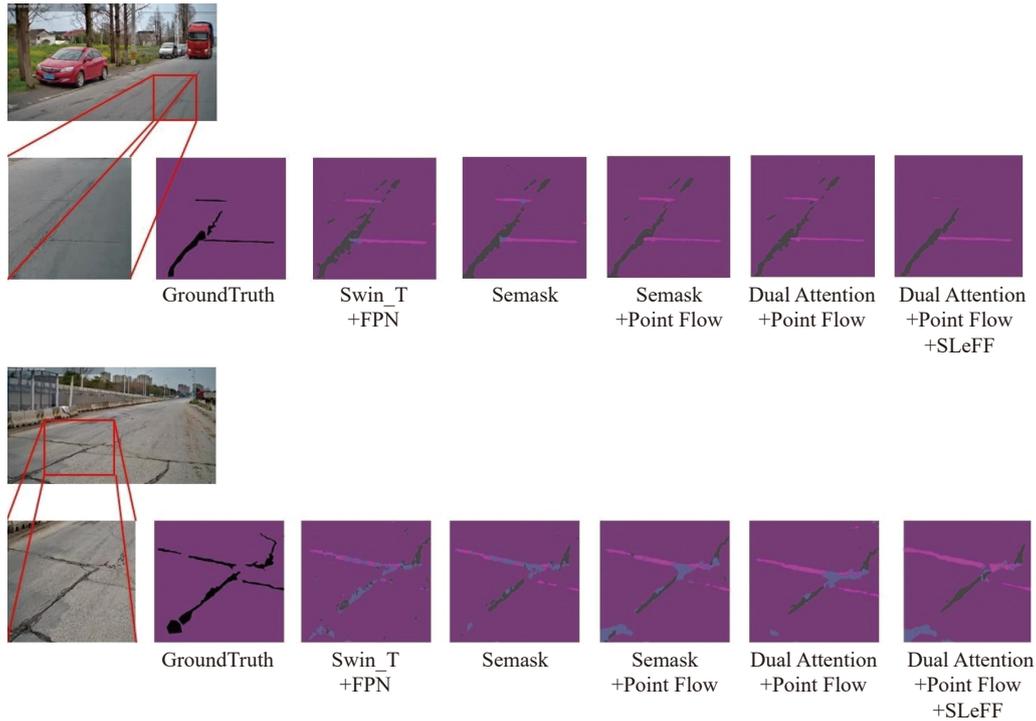
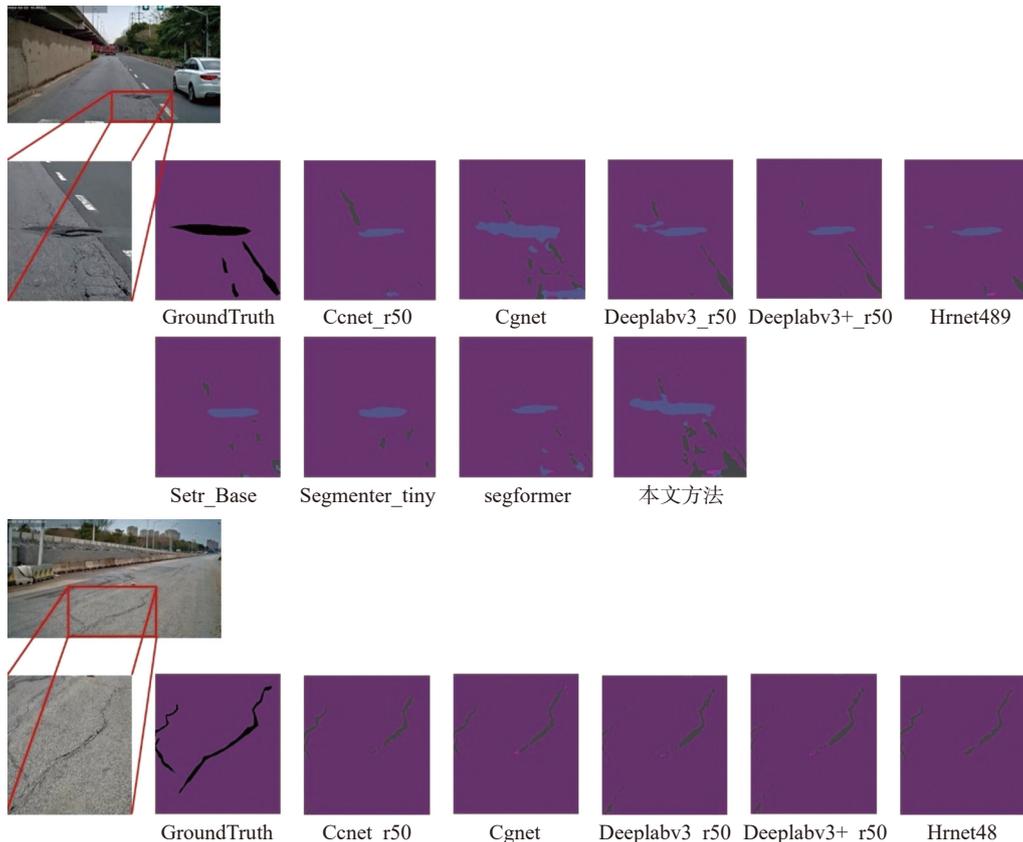


图 7 不同语义分割模型实验结果对照可视化

Fig. 7 Comparison visualization of experimental results of different semantic segmentation models

图 8 给出了本文模型在自研路面病害数据集上的消融实验结果对照。纯 Swin 骨干加 FPN 网络结构预测结果最差, 无法提取病害整体关系, 因此出现一条裂缝被识别成好几种病害。而添加最原始的语义先验模块, 缓解数据集不够导致无

法提取病害全局特征。进一步添加稀疏主体点流模型, 只关注病害特征点, 预测病害的完整度更加好, 最后使用本文的语义先验模块, 预测结果光滑平整, 基本上不会出现病害不连续的情况, 验证了本文所设计模块的有效性。



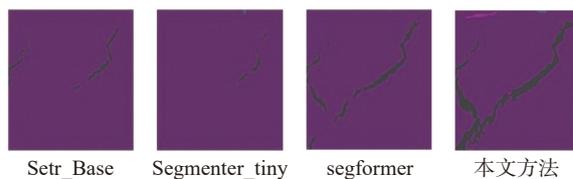


图 8 Dual-SePointFlow 模型消融实验结果对照可视化

Fig. 8 Comparison visualization of ablation experimental results of Dual-SePointFlow model

4 结束语

本文提出了一种基于双分支点流语义先验的真实路面病害识别模型。针对路面病害数据集小、图像质量差等问题,提出了双分支语义先验模块来指导骨干特征提取网络更好地优化;针对前置摄像头采集图像病害占比小的问题,提出了稀疏主体点流模块,只对部分病害前景采样点进行计算,同时融合不同层与自身层的特征信息进行更新,进一步提高识别精度的同时缓解了病害类别严重不平衡问题。为了训练和验证模型,本文构建了一个真实路面病害分割数据集,通过大量实验验证本文方法不仅能在真实环境下有效地识别路面病害,而且针对小样本和类比不平衡的情况也具有好的效果。由于类似的全景路面病害数据集很少,本文只在本数据集上表现良好,并验证本模型的有效性。但是想要模型做到强泛化能力和强鲁棒性,具有一定的挑战性和研究价值。未来如果有更多全景路面数据集公开,还可以进一步探究如何更好挖掘复杂背景与前景的关系,提高其检测和分割的识别率。

参考文献:

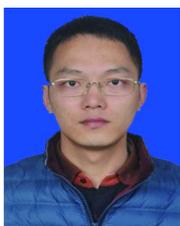
- [1] “十四五”现代综合交通运输体系发展规划 [J]. 铁道技术监督, 2022, 50(2): 9–23, 27.
Development plan of modern comprehensive transportation system in the 14th five-year plan[J]. Railway quality control, 2022, 50(2): 9–23, 27.
- [2] The Central People’s Government of the People’s Republic of China, v1.0[EB/OL]. (2023–01–11)[2023–06–15]. <http://www.gov.cn/>.
- [3] 侯越, 张慧婷, 高智伟, 等. 基于数据深度增强的路面病害智能检测方法研究及比较 [J]. 北京工业大学学报, 2022, 48(6): 622–634.
HOU Yue, ZHANG Huiting, GAO Zhiwei, et al. Research and comparison of intelligent detection methods of pavement distress based on deep data augmentation[J]. Journal of Beijing University of Technology, 2022, 48(6): 622–634.
- [4] 章世祥, 张汉成, 李西芝, 等. 基于机器视觉的路面裂缝病害多目标识别研究 [J]. 公路交通科技, 2021, 38(3): 30–39.
ZHANG Shixiang, ZHANG Hancheng, LI Xizhi, et al. Study on multi-objective identification of pavement cracks based on machine vision[J]. Journal of highway and transportation research and development, 2021, 38(3): 30–39.
- [5] 高建贞, 任明武, 唐振民, 等. 路面裂缝的自动检测与识别 [J]. 计算机工程, 2003, 29(2): 149–150.
GAO Jianzhen, REN Mingwu, TANG Zhenmin, et al. Automatic road crack detection and identification[J]. Computer engineering, 2003, 29(2): 149–150.
- [6] 徐威, 唐振民, 吕建勇. 基于图像显著性的路面裂缝检测 [J]. 中国图象图形学报, 2013, 18(1): 69–77.
XU Wei, TANG Zhenmin, LYU Jianyong. Pavement crack detection based on image saliency[J]. Journal of image and graphics, 2013, 18(1): 69–77.
- [7] 伯绍波, 闫茂德, 孙国军, 等. 沥青路面裂缝检测图像处理算法研究 [J]. 微计算机信息, 2007, 23(15): 280–282.
BO Shaobo, YAN Maode, SUN Guojun, et al. Research on crack detection image processing algorithm for asphalt pavement surface[J]. Control & automation, 2007, 23(15): 280–282.
- [8] SHENG Peng, CHEN Li, TIAN Jing. Learning-based road crack detection using gradient boost decision tree[C]// 2018 13th IEEE Conference on Industrial Electronics and Applications. Piscataway: IEEE, 2018: 1228–1232.
- [9] SUN Lu, KAMALIARDAKANI M, ZHANG Yongming. Weighted neighborhood pixels segmentation method for automated detection of cracks on pavement surface images[J]. Journal of computing in civil engineering, 2016, 30(2): 04015.
- [10] 曹建农, 张昆, 元晨, 等. 用 Mean Shift 实现路面裂缝损伤自动识别与特征测量 [J]. 计算机辅助设计与图形学学报, 2014, 26(9): 1450–1459.
CAO Jiannong, ZHANG Kun, YUAN Chen, et al. Automatic road cracks detection and characterization based on mean shift[J]. Journal of computer-aided design & computer graphics, 2014, 26(9): 1450–1459.
- [11] WANG Yanyan, SONG Kechen, LIU Jie, et al. RENet: rectangular convolution pyramid and edge enhancement

- network for salient object detection of pavement cracks[J]. *Measurement*, 2021, 170: 108698.
- [12] YANG Jing, FU Qin, NIE Mingxin. Road crack detection using deep neural network with receptive field block[J]. *IOP conference series: materials science and engineering*, 2020, 782(4): 042033.
- [13] YAO Hui, LIU Yanhao, LI Xin, et al. A detection method for pavement cracks combining object detection and attention mechanism[J]. *IEEE transactions on intelligent transportation systems*, 2022, 23(11): 22179–22189.
- [14] ZHANG Yujia, LI Qianzhong, ZHAO Xiaoguang, et al. TB-net: a three-stream boundary-aware network for fine-grained pavement disease segmentation[C]//2021 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2021: 3654–3663.
- [15] WU Junxian, ZHANG Yujia, ZHAO Xiaoguang. Multi-task learning for pavement disease segmentation using wavelet transform[C]//2022 International Joint Conference on Neural Networks. Piscataway: IEEE, 2022: 1–8.
- [16] YU Gui, DONG Juming, WANG Yihang, et al. RUC-net: a residual-unet-based convolutional neural network for pixel-level pavement crack segmentation[J]. *Sensors*, 2022, 23(1): 53.
- [17] ZHANG Lei, YANG Fan, ZHANG Y D, et al. Road crack detection using deep convolutional neural network[C]//2016 IEEE International Conference on Image Processing. Piscataway: IEEE, 2016: 3708–3712.
- [18] 韩静园, 王育坚, 谭卫雄, 等. 基于 FCN 的路面裂缝分割算法[J]. *传感器与微系统*, 2022, 41(6): 146–149.
HAN Jingyuan, WANG Yujian, TAN Weixiong, et al. Pavement crack segmentation algorithm based on FCN[J]. *Transducer and microsystem technologies*, 2022, 41(6): 146–149.
- [19] YANG Fan, ZHANG Lei, YU Sijia, et al. Feature pyramid and hierarchical boosting network for pavement crack detection[J]. *IEEE transactions on intelligent transportation systems*, 2020, 21(4): 1525–1535.
- [20] 刘文婷, 卢新明. 基于计算机视觉的 Transformer 研究进展[J]. *计算机工程与应用*, 2022, 58(6): 1–16.
LIU Wenting, LU Xinming. Research progress of transformer based on computer vision[J]. *Computer engineering and applications*, 2022, 58(6): 1–16.
- [21] XIE Enze, WANG Wenhai, YU Zhiding, et al. SegFormer: simple and efficient design for semantic segmentation with transformers[EB/OL]. (2021–10–28) [2023–06–15]. <https://arxiv.org/abs/2105.15203.pdf>.
- [22] STRUDEL R, GARCIA R, LAPTEV I, et al. Segformer: transformer for semantic segmentation[C]//2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2022: 7242–7252.
- [23] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]// European Conference on Computer Vision. Cham: Springer, 2020: 213–229.
- [24] JAIN J, SINGH A, ORLOV N, et al. SeMask: semantically masked transformers for semantic segmentation[C]//2023 IEEE/CVF International Conference on Computer Vision Workshops. Piscataway: IEEE, 2023: 752–761.
- [25] SHEN Zhuoran, ZHANG Mingyuan, ZHAO Haiyu, et al. Efficient attention: attention with linear complexities[C]//2021 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2021: 3530–3538.
- [26] EL-NOUBY A, TOUVRON H, CARON M, et al. XcIT: cross-covariance image transformers[EB/OL]. (2021–06–18) [2023–05–15]. <http://arxiv.org/abs/2106.09681.pdf>
- [27] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale[EB/OL]. (2021–07–03) [2023–06–15]. <http://arxiv.org/abs/2010.11929.pdf>.
- [28] YUAN Kun, GUO Shaopeng, LIU Ziwei, et al. Incorporating convolution designs into visual transformers[C]//2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2022: 559–568.
- [29] LI Xiangtai, HE Hao, LI Xia, et al. PointFlow: flowing semantics through points for aerial image segmentation[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 4215–4224.
- [30] HUANG Zilong, WANG Xinggang, HUANG Lichao, et al. CCNet: criss-cross attention for semantic segmentation[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2020: 603–612.
- [31] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017–12–05) [2023–06–15]. <http://arxiv.org/abs/1706.05587.pdf>
- [32] CHEN L C, ZHU Yukun, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//European Conference on Computer Vision. Cham: Springer, 2018: 833–851.
- [33] SUN Ke, XIAO Bin, LIU Dong, et al. Deep high-resolution representation learning for human pose estimation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2020: 5686–5696.
- [34] WU Tianyi, TANG Sheng, ZHANG Rui, et al. CGNet: a light-weight context guided network for semantic segmentation[J]. *IEEE transactions on image processing*,

2021, 30: 1169–1179.

- [35] ZHENG Sixiao, LU Jiachen, ZHAO Hengshuang, et al. Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 6877–6886.

作者简介:



庞荣, 博士研究生, 主要研究方法为人工智能、深度学习、大数据分析、挖掘和高速公路智能检测。E-mail: 519231410@qq.com。



杨燕, 教授, 博士生导师, 西南交通大学计算机与人工智能学院副院长、中国计算机学会杰出会员, 主要研究方向为人工智能、大数据分析、挖掘、多视图学习、云计算和云服务。主持国家自然科学基金等项目 10 余项。发表学术论文 230 余篇。E-mail: yyang@swjtu.edu.cn。



冷雄进, 硕士研究生, 主要研究方向为人工智能和计算机视觉。E-mail: 2932985761@qq.com。

第 20 届中国智能系统会议

中国智能系统会议是由中国人工智能学会智能空天系统专业委员会发起的系列学术会议, 其宗旨是为本领域的专家学者、工程技术人员以及研究生提供一个学术交流平台, 以推动我国智能系统相关理论、技术与应用的发展。第 20 届中国智能系统会议 (CISC2024) 将于 2024 年 10 月 26—27 日在广西桂林市召开。

本次会议由中国人工智能学会主办, 中国人工智能学会智能空天系统专业委员会与中国仿真学会人工智能仿真技术专业委员会协办, 北京邮电大学、桂林电子科技大学、航天伺服驱动与传动技术实验室与北京航空航天大学联合承办。会议论文集将由 Springer 出版社在 Lecture Notes in Electrical Engineering 系列正式出版, EI 收录。热忱欢迎海内外广大同仁踊跃投稿并出席本届会议, 交流学术成果。

投稿要求

- 1) 论文未曾在国内外期刊或会议上发表。
- 2) 稿件写作必须使用英文, 并严格按照 LaTeX 模板要求进行排版。
- 3) 论文采用网上投稿, 投稿请浏览网址: https://easychair.org/conferences/?conf=cisc2024_guilin。

重要日期

论文投稿截止日期: 2024 年 4 月 30 日

论文录用通知日期: 2024 年 6 月 20 日

会议注册/终稿提交截止日期: 2024 年 7 月 10 日