

基于FCC-Deeplabv3+的城市地下管道缺陷语义分割方法

田淙文, 李波*, 蓝雯飞, 潘禹欣, 姚为

(中南民族大学 计算机科学学院, 武汉 430074)

摘要 城市地下管道图像缺陷具有种类多、背景复杂、噪声多、缺陷尺度变化大等特点, 导致目前城市地下管道缺陷分割算法精度不够高. 本研究提出了一种基于Deeplabv3+的改进分割模型FCC-Deeplabv3+, 并将该模型首次应用到城市地下管道缺陷分割. 结合十字交叉注意力机制, 使模型在预测时获取更丰富的上下文信息; 提出了改进的解码器上采样策略, 引入多尺度信息, 减少中间层信息的丢失; 使用基于增强的对比学习策略监督模型, 提升了模型分割能力. 此外, 针对目前城市地下管道缺陷分割领域没有公开数据集的情况, 基于Sewer-ML公开数据集, 进行数据标注工作, 构建了包含900张用于缺陷分割任务的数据集. 通过实验验证了提出的缺陷分割模型的有效性及时效性, 对比原始Deeplabv3+模型, mIoU提升了3.73%, mPA也提升了1.67%, 并且相比其他基于深度学习的语义分割算法, 也具有一定优势.

关键词 FCC-Deeplabv3+算法; 缺陷分割; 城市地下管道; 十字交叉注意力; 对比学习; 深度监督

中图分类号 TP391 **文献标志码** A **文章编号** 1672-4321(2025)01-0107-11

doi: 10.20056/j.cnki.ZNMDZK.20250113

Semantic segmentation method for urban underground pipeline defects based on FCC-Deeplabv3+

TIAN Congwen, LI Bo*, LAN Wenfei, PAN Yuxin, YAO Wei

(College of Computer Science, South-Central Minzu University, Wuhan 430074, China)

Abstract The image defects of urban underground pipelines have the characteristics of multiple types, complex background, high noise, and large scale changes, which lead to insufficient accuracy of current urban underground pipeline defect segmentation algorithms. This research proposes an improved segmentation model FCC-Deeplabv3+ based on Deeplabv3+ and applies this model for the first time to defect segmentation of urban underground pipelines. Combined with the criss-cross attention mechanism, the method can obtain richer context when making predictions. An improved decoder upsampling strategy is proposed to introduce multi-scale information to reduce the loss of intermediate layer information. The method is supervised based on the contrastive learning strategy, which improves the method segmentation capability. In addition, regarding the current situation that there is no publicly available dataset for defect segmentation in urban underground pipelines, based on the Sewer-ML dataset, we performed data annotation work and constructed a dataset containing 900 images for the defect segmentation. The effectiveness and real time of the proposed defect segmentation method was verified through experiments. Compared with the original Deeplabv3+ model, mIoU increased by 3.73%, and mPA also increased by 1.67%. It also has certain advantages compared with other semantic segmentation methods based on deep learning.

Keywords FCC-Deeplabv3+ algorithm; defect segmentation; urban underground pipelines; criss-cross attention; contrastive learning; deep supervision

收稿日期 2024-03-03

* 通信作者 李波, 男, (1975-), 教授, 研究方向: 模式识别和机器学习, E-mail: libo@mail.scuec.edu.cn

基金项目 国家自然科学基金资助项目(61976226)

城市地下管道系统是现代城市重要的基础设施之一,其运行情况关系到城市的公共卫生和基础设施建设,随着使用年限的增长,会出现各类缺陷^[1]. 其内部情况评估的相关技术有激光检测技术、声纳检测技术和管道闭路电视(Closed-Circuit Television, CCTV)技术等^[2]. 由于成本低、部署灵活等优势, CCTV 检测技术已成为地下管道检测任务中应用最广泛的技术^[3]. 该技术使用机器人进入到管道内部拍摄管道内部的图像,由专业人员对获取到的图像数据进行判读^[4],对管道内的缺陷进行定位、分类和等级评定. 人工判读费时且易出错,因此需要自动化的技术辅助人工检测^[5].

基于图像的地下管道检查算法主要包括 3 类: 缺陷分类、缺陷检测和缺陷分割^[6]. 其中缺陷分割不仅可以对多种缺陷同时进行检测、分类和定位,还可以提取各类缺陷的几何特征,为后期的缺陷评级工作提供支撑. 传统的分割方法主要基于数学的形态学, SU 等^[7]验证了数学形态学方法中的开顶帽方法能够很好地分割错口缺陷. 此外,他们还提出了一种基于边缘检测的形态学分割方法,并验证该方法比开顶帽方法有更好的分割性能^[8]. 其后基于深度学习的缺陷分割方法在该项任务上得到了应用^[9], WANG 等^[10]提出了 DilaSeg 模型,首次将深度学习模型应用于城市地下管道缺陷分割任务中. HAURUM 等^[11]为城市地下管道缺陷分类工作提供了 Sewer-ML 公开数据集. ZHOU 等^[12]将 Deeplabv3+ 模型应用于城市排水管道缺陷语义分割. PAN 等^[13]提出了 PipeUNet 模型,对 4 种特定缺陷类型的图像进行分割.

本文基于 Deeplabv3+^[14] 模型,提出 FCC-Deeplabv3+ 分割模型,并首次应用到城市地下管道缺陷分割. 主要贡献包括:(1)设计了融合多尺度特征的 Fusion 模块,减少高倍上采样导致的中间层信息丢失;(2)加入了带有交叉注意力^[15]模块的分割头部,平衡模型参数量和上下文信息的计算;(3)在模型训练过程中加入了对比深度监督^[16]的策略调整参数,优化模型的抗干扰能力. 此外,针对目前城市地下管道缺陷语义分割领域尚未有公开数据集的情况,本文基于 Sewer-ML^[11]城市地下管道缺陷图像数据集,构建了城市地下管道缺陷分割任务数据集,并在该数据集上进行实验.

1 Deeplabv3+模型原理

Deeplabv3+ 是一种编码器解码器架构的语义分割模型,编码器部分主要由骨干网络和空洞空间池化金字塔^[17](Atrous Spatial Pyramid Pooling, ASPP)模块组成,解码器部分由上采样和卷积模块组成. 基本结构如图 1 所示. 以 ResNet-50^[18]作为骨干特征提取网络,初步提取图像的特征,对提取到的特征使用 ASPP 模块进行不同感受野的特征提取. 将得到的特征通过一个 1×1 的卷积块,得到编码器的输出. 将编码器的输出 4 倍上采样后与骨干网络的第 3 阶段的浅层特征融合,而对于骨干网络的浅层特征会使用一个 1×1 的卷积降维到 48 通道,这样可以使网络融合不同尺度的特征,之后将融合的特征通过一系列的卷积块并再次 4 倍上采样,得到与原始图像大小相同的预测输出.

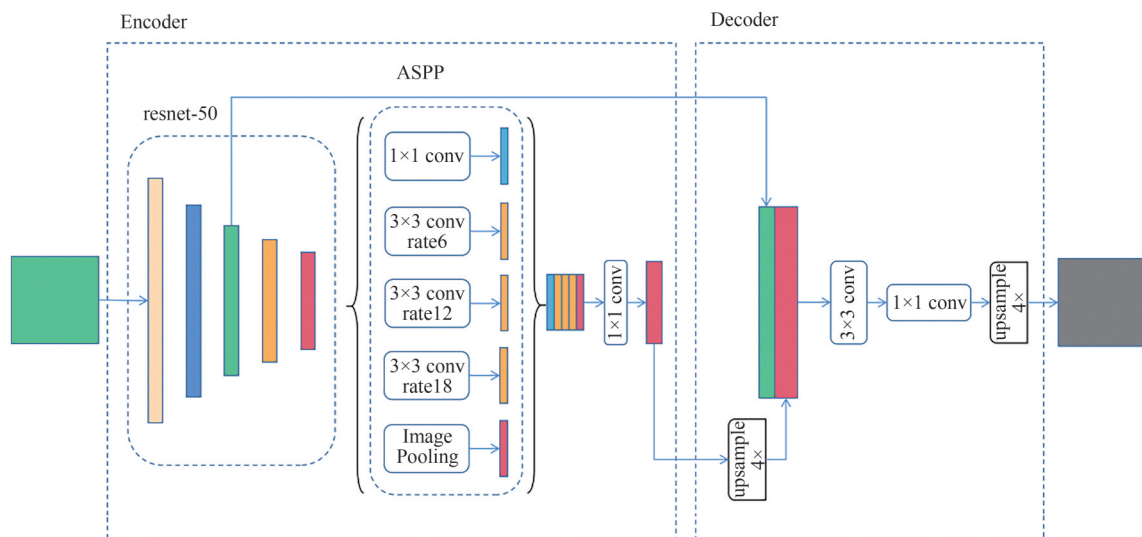


图 1 Deeplabv3+结构图

Fig. 1 Structure of Deeplabv3+

2 本文提出的 FCC-Deeplabv3+模型

本文基于 Deeplabv3+模型提出了一种用于城市地下管道缺陷分割的改进模型 FCC-Deeplabv3+,其结构如图 2 所示,红色加粗虚线区域为本文模型的改进工作.对比 Deeplabv3+模型,本文改进了模型的解码器上采样策略,融合了更丰富的多尺度特征,从骨干网络部分引入了 L1、L2、L3 三个尺度的浅层特征,将这些浅层特征与编码器部分的输出特征图通过 Fusion Module 融合上采样,并对编码器部分

的输出使用更高倍率的上采样后与 Fusion Module 输出的特征图拼接.在此基础上,拼接后的特征图在经过一个 3×3 的卷积之后会通过带有十字交叉注意力的 RCCA Module,获取更丰富的上下文信息.最后进行 2 倍上采样得到最终的预测.此外,引入了对比学习的策略.在模型训练的过程中,对同一张图像使用两种不同的增强方法变换,将变换后的图像同时输入到模型中,对模型编码器和解码器部分的特征图输出计算对比学习损失,优化模型的抗干扰能力.

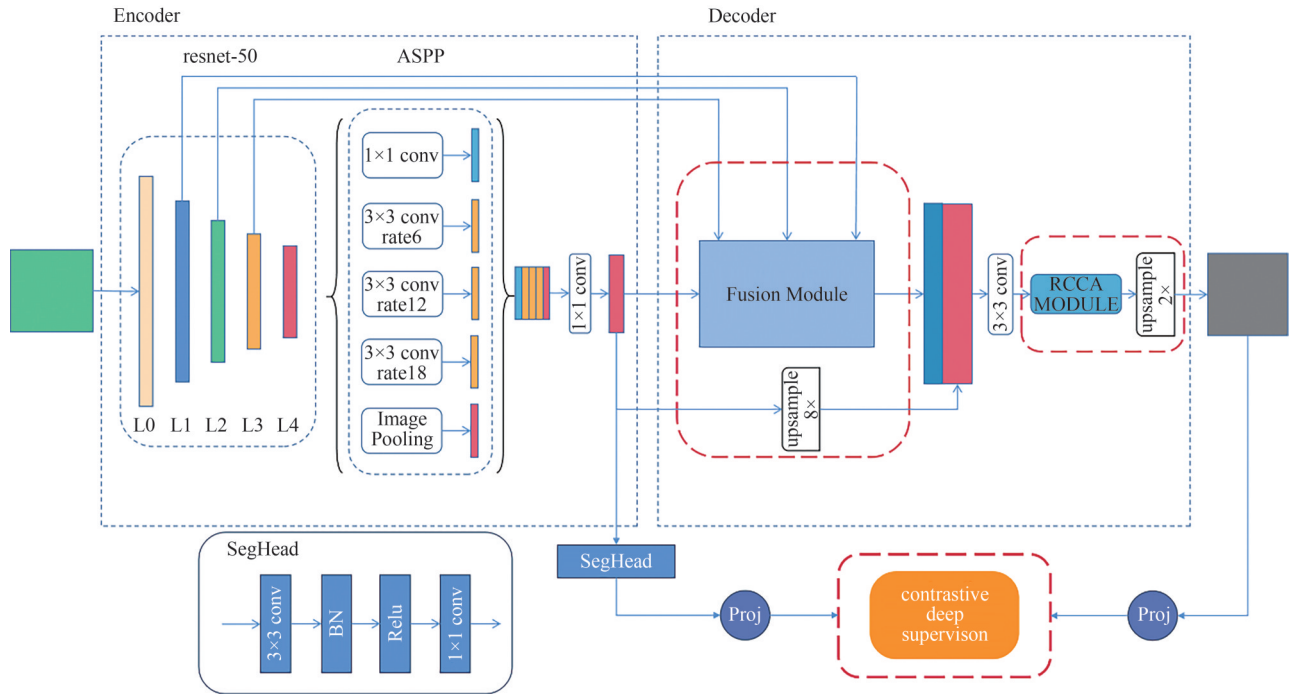


图2 FCC-Deeplabv3+结构图

Fig. 2 Structure of FCC-Deeplabv3+

2.1 解码器上采样模块的改进

考虑到城市地下管道缺陷类别形态特征丰富,本文尝试引入多尺度空间信息.FPN^[19]是目标检测任务中融合多尺度信息的模块,改进的上采样模块 Fusion Module 采用了与 FPN 类似的融合策略,共使用 3 次 2 倍上采样,每次上采样后都会与骨干网络的浅层特征分支进行融合, Fusion Module 如图 3 所示.输入图像大小为 $512 \times 512 \times 3$,模型编码器部分的输出特征图大小为 $32 \times 32 \times 256$,经过 2 倍上采样操作后宽高与骨干网络的 L3 特征层输出相同,对 L3 特征层通过 1×1 卷积,将通道数缩减到与主分支相同,之后再相加并进行 2 倍上采样,此时的特征层宽高与 ResNet 的 L2 特征层输出相同,即可与 L2 特征层进行融合.经过 3 次特征融合操作后,将得到的特征图通过 1×1 卷积降维到 48 通道,与模型编

码器输出 8 倍上采样的输出拼接,弥补高倍上采样导致的信息损失,并融合多尺度信息.

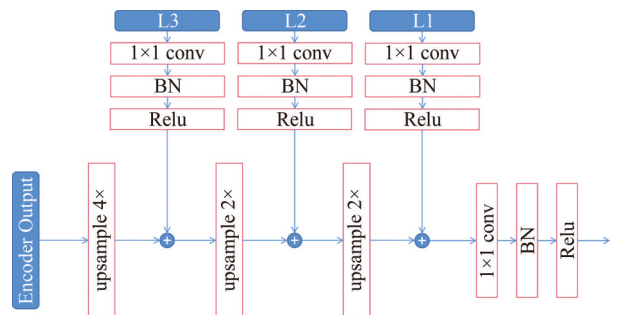


图3 Fusion Module 结构图

Fig. 3 Structure of Fusion Module

2.2 引入十字交叉注意力的分割头

注意力机制使深度神经网络关注更重要的信息^[20].不少研究表明,上下文信息在计算机视觉任务

中起到了至关重要的作用^[21-22]. 在语义分割任务中引入上下文信息的方式主要包括基于扩张卷积的方法、基于池化的方法^[23]和基于非局部的方法^[24]. 然而基于扩张卷积的方法不能生成密集的上下文信息;基于池化的方法对所有像素都使用相同的上下文信息提取

方式,不能满足不同像素对上下文信息的需求;基于非局部的方法虽然能生成密集上下文信息,但又会带来计算量的大量提升. 为了平衡计算量和上下文信息的获取,本文引入交叉交叉注意力(Criss-Cross Attention)模块,交叉交叉注意力模块如图 4 所示.

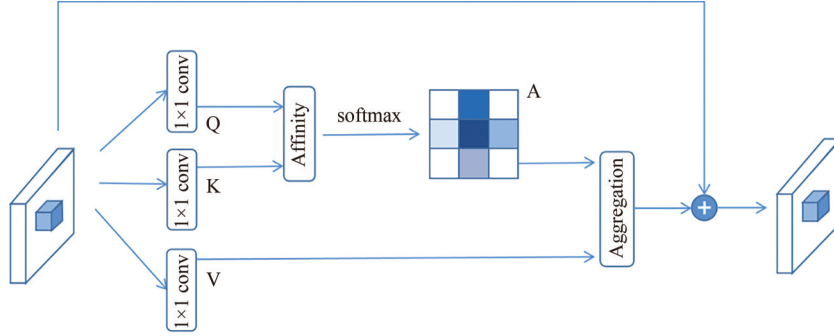


图 4 交叉交叉注意力模块

Fig. 4 Criss-cross attention module

对于一个输入特征图 $M \in \mathbb{R}^{C \times H \times W}$,先通过两个 1×1 的卷积层分别生成特征图 Q 和 K , $\{Q, K\} \in \mathbb{R}^{C' \times H \times W}$, 本文设置 C' 为 C 的 $\frac{1}{8}$. 然后进行 Affinity 操作,得到特征图 $A \in \mathbb{R}^{(H+W-1) \times H \times W}$. Affinity 操作可以表示为:

$$b_{i,u} = Q_u \Pi_{i,u}^T, \quad (1)$$

其中 $b_{i,u} \in B$, $Q_u \in \mathbb{R}^{C'}$ 表示特征图 Q 空间上位置 u 的向量, $\Pi_u \in \mathbb{R}^{(H+W-1) \times C'}$ 表示特征图 K 空间上同一位置 u 的集合,集合包含了特征图 K 位置 u 水平方向和垂直方向上的所有向量. $\Pi_{i,u} \in \mathbb{R}^{C'}$, $i = [1, \dots, H+W-1]$, 表示集合 Π_u 的第 i 个元素. 如图 5 所示,通过 Affinity 操作可以得到特征图 $B \in \mathbb{R}^{(H+W-1) \times H \times W}$, 然后对 B 计算 softmax,得到特征图 A (图 5).

再将特征图 M 通过 1×1 的卷积层得到特征图 $V \in \mathbb{R}^{C \times H \times W}$, 进行 Aggregation 操作,得到最终的输出

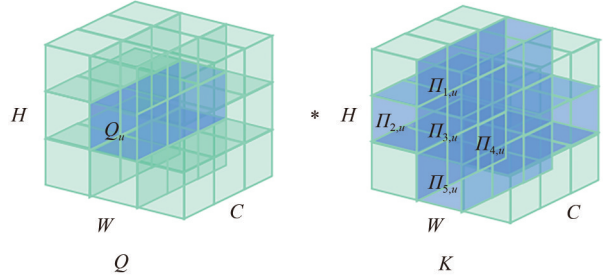


图 5 Affinity 操作

Fig. 5 Affinity operation

$M' \in \mathbb{R}^{C \times H \times W}$. Aggregation 操作如图 6 所示. 该操作可以表示为:

$$M'_u = \sum_{i=0}^C A_u \Omega_{u-i} + M_u, \quad (2)$$

其中 $A_u \in \mathbb{R}^{H+W-1}$ 表示特征图 A 位置 u 的向量, $\Omega_{u-i} \in \mathbb{R}^{H+W-1}$ 表示特征图 V 第 i 通道 ($i = [1, 2, \dots, C]$) 位置 u 与水平和垂直元素组成的向量, $M_u \in \mathbb{R}^C$ 表示特征图 M 位置 u 的特征向量(图 6).

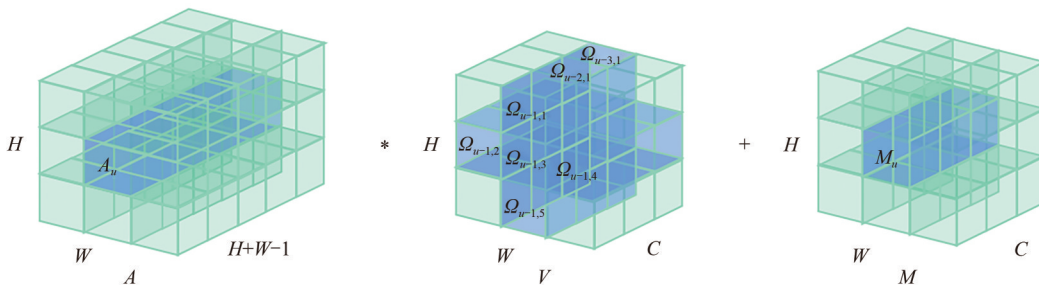


图 6 Aggregation 操作

Fig. 6 Aggregation operation

单次的交叉交叉注意力的计算只能获取水平和垂直方向上的上下文信息,而重复计算交叉交叉注意力能够将水平和垂直方向上的上下文信息

射到周围. 为了获取完整的上下文信息,使用到了如图 7 所示的 RCCA 模块(Recurrent Criss-Cross Attention Module). 对于输入特征图,先通过一个 3×3

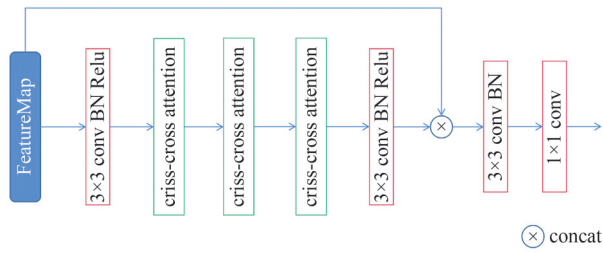


图 7 RCCA 模块

Fig. 7 Recurrent criss-cross attention module

的卷积进行通道上的降维,然后进行 R (本文取 $R=3$)次的交叉交叉注意力计算,然后通过一个 3×3 的卷积并与原始特征图进行拼接,再通过一个 3×3 的卷积,最后通过一个 1×1 卷积,将特征图降维到与类别数量相同。

2.3 对比深度监督优化网络参数

传统的深度学习模型训练只在网络的最后一层计算任务相关的损失,并逐层反向传播,这导致

模型的中间层难以得到优化. LEE 等提出了深度监督网络^[25],在模型不同阶段加入分类器,对这些分类器的输出与原始模型的分器一同计算任务损失. 图像领域中的深度神经网络浅层更倾向于学习形状之类的低级特征,深层更倾向于学习高级语义特征. 深度监督的方式迫使网络的浅层学习与任务相关的高级特征,这可能会导致模型的效果下降. 针对这样的问题, ZHANG 等提出了对比深度监督(Contrastive Deep Supervision)的方法^[16],引入对比学习的方法来监督模型的中间层,将模型各层的输出使用投影头投影到同一纬度并计算对比学习损失. 实验表明,对比深度监督的方法在一些与图像相关的任务中取得了高于深度监督方法的效果. 本文采用了对比深度监督的策略,如图 8 所示,对于模型编码器和解码器两个部分的特征图输出分别使用对比学习方式计算损失,优化模型的参数.

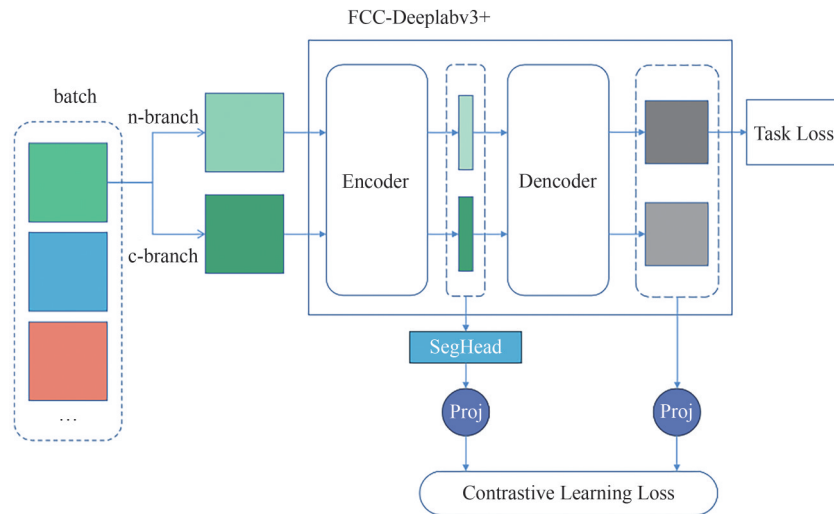


图 8 对比深度监督

Fig. 8 Contrastive deep supervision

在模型训练的过程中,对于同一张输入图像使用两个不同分支输入,对 n 分支图像使用原始模型的图像增强方法,对 c 分支图像使用不同于 n 分支的图像增强方法,将两个分支增强后的图像同时输入到 FCC-Deeplabv3+ 网络中,对编码器的输出特征图通过分割头后投影到相同维度计算对比学习损失. 对于解码器输出,先对两个分支的输出投影到相同维度计算对比学习损失,再取 n 分支的输出计算任务损失,最后整体损失相加,如下:

$$L = L_t + \alpha L_c, \quad (3)$$

其中 L_t 为任务损失, L_c 为对比学习损失,本文设置 α 为 0.01.

由于该方法仅在训练的过程中加入了对比学

习损失的计算,不会带来参数量和推理速度的提升. 对一个批量大小为 N 的输入图像,进行两种不同的数据增强,得到 $2N$ 张图像输入,一张图像 v 和该图像原图的另一种增强 v^+ 视作正对,其他 $2(N-1)$ 张图像记为 v^- , v 与 v^- 视作负对,本文采用的对比损失函数如下:

$$L_c = -\log \frac{\exp(\text{sim}(v, v^+)/\tau)}{\exp(\text{sim}(v, v^+)/\tau) + \sum_{v^- \in K_t} \exp(\text{sim}(v, v^-)/\tau)}, \quad (4)$$

其中 $\text{sim}(m, n) = m^T n / (\|m\| \|n\|)$,表示余弦相似度,即 m 和 n 经过 l_2 正则化后的点积. K_t 表示输入该批量中除去正对图像外其他 $(2N-1)$ 张图像的集合. τ 为温度超参数,本文设置 τ 为 0.07.

3 实验与结果分析

3.1 实验数据

3.1.1 实验数据来源

目前城市地下管道缺陷分割领域尚未有具有分割标注的公开数据集,本文的实验数据来源于2021年CVPR会议论文发布的Sewer-ML数据集^[11].该数据集主要应用于地下管道多标签缺陷分类任务,没有缺陷分割任务的标注,不能直接用于分割

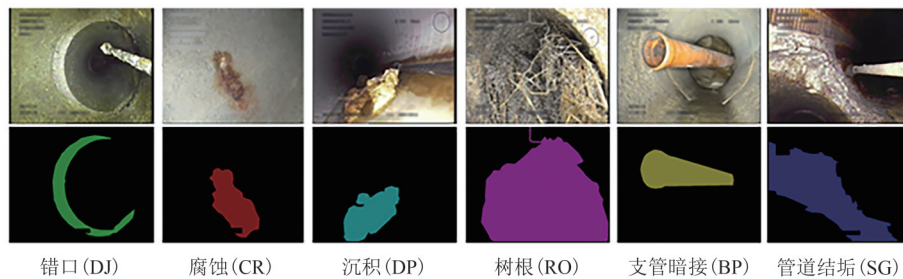


图9 地下管道缺陷样本

Fig. 9 Underground pipeline defects samples

3.1.2 数据增强

为了更好地验证模型的泛化能力,本文取原始数据的20%作为测试集,由于数据集图片数量有限,用于训练的数据较少,对剩余数据集采用随机色彩抖动、随机角度旋转、垂直翻转的方式进行数据增广,如图10所示,增广后的数据集达到4320张.对增强后的数据集按照9:1划分为训练集和验证集.

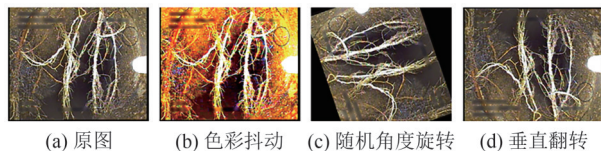


图10 数据增强

Fig. 10 Data augmentation

3.1.3 实验环境及参数设置

实验硬件环境为:操作系统为64位的Windows11, CPU为Intel i5-11400, GPU为NVIDIA GeForce RTX 3070, 显卡内存为8 GB;软件环境为:Pycharm平台下Python 3.8, 深度学习框架为Pytorch 1.10.1和CUDA 11.4. 设置最大学习率为 5×10^{-4} , 最小学习率为 3×10^{-6} , 使用Adam优化器, 损失函数使用Focal Loss和Dice Loss, 学习率优化策略采用带有热身机制的余弦下降策略. 迭代次数设置为120次, 前50次迭代冻结骨干网络权重, 冻结阶段批处理大小设置为8, 后70次迭代对模型整体进行训练, 批处理大小设置为4.

实验,本文选取了在实际占比中较多、较为典型的缺陷类型进行缺陷语义分割数据的标注,共标注了900张图像.缺陷类型包括管道错口(DJ, Displaced Joint)、腐蚀(CR, Corrosion)、沉积(DP, Deposit)、树根(RO, Root)、支管暗接(BP, Branch Pipe)、管道结垢(SG, Surface Gathering)6种,加上背景类别(BG, Background)共7种类别,为了更易于区分,对不同类别使用不同颜色展示,如图9所示,第1行为原始缺陷样本,第2行为缺陷标签.

3.2 评价指标

语义分割任务中,常用的评价指标有平均交并比(mIoU)、平均像素精确度(mPA)、频率权重交并比(FWIoU).本文主要使用mIoU和mPA作为模型性能的评价指标.

IoU用于衡量模型的预测区域和真实标签区域的重合程度,是两个区域的交集与并集的比值, mIoU则是数据集中每一个类别的IoU的平均值,计算过程如公式(5):

$$mIoU = \frac{1}{k} \sum_{i=1}^k \frac{TP}{FN + FP + TN}, \quad (5)$$

其中 k 为类别数量,TP表示预测为正样本且真实标签也为正样本的像素, FN是预测为负样本但真实标签是正样本的像素, FP是预测为正样本但真实标签是负样本的像素, TN是预测为负样本真实标签也为负样本的像素.

PA用于计算预测正确的像素数占总像素数的比例,可以被认为是召回率, mPA就是对每一个类别取平均,如公式(6):

$$mPA = \frac{1}{k} \sum_{i=1}^k \frac{TP + TN}{TP + FN + FP + TN}. \quad (6)$$

Precision用于计算预测为正例的样本中实际为正例的比例, mPrecision就是对每个类别取平均,如公式(7):

$$mPrecision = \frac{1}{k} \sum_{i=1}^k \frac{TP}{TP + FP}. \quad (7)$$

Accuracy 用于计算预测正确的样本占全部样本的比例,如公式(8):

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \cdot \quad (8)$$

3.3 实验结果分析

3.3.1 不同语义分割算法对比实验

为了验证本文的 FCC-Deeplabv3+ 算法对城市地下管道缺陷分割能力,将改进模型与目前常用的

语义分割模型进行了对比实验,结果如表 1 所示,表中的第 1、2 列分别表示语义分割模型、该模型选取的特征提取网络. 由表 1 可知,本文的算法与原始 Deeplabv3+ 算法相比,在参数上略微增加,但在精度上有不错的提升,并且与其他常用的语义分割算法相比,本文算法在城市地下管道缺陷分割任务上效果最好. 其中 BiSeNetV2 以最小的参数量,取得不错的分割效果.

表 1 不同语义分割算法对比实验

Tab. 1 Comparative experiments of different semantic segmentation algorithms

Method	Backbone	mIoU/%	mPA/%	FWIoU/%	mPrecision/%	Accuracy/%	Params
UNet ^[26]	VGG-16	64.65	73.19	84.45	83.32	91.25	24.9M
PSPNet ^[23]	ResNet-50	67.59	75.97	85.54	85.30	92.02	46.7M
DANet ^[27]	ResNet-50	65.04	77.03	84.22	79.50	91.03	47.4M
BiSeNet ^[28]	ResNet-101	67.76	79.62	85.37	80.82	91.63	90.8M
BiSeNetV2 ^[29]	-	68.52	78.25	85.69	83.92	92.03	13.5M
DDRNet ^[30]	-	67.01	79.26	85.18	80.97	91.69	32.65M
Deeplabv3+ ^[14]	ResNet-50	69.18	79.95	86.01	83.25	92.16	40.4M
Ours	ResNet-50	72.91	81.62	87.23	86.38	92.95	41.7M

图 11 展示了不同语义分割模型对 7 种类别缺陷的分割 IoU 和 PA. 由图 11(a) 和图 11(b) 可以发

现,本文提出的模型在大部分的缺陷类别上的分割 IoU 和 PA 均优于其他模型,尤其在支管暗接(BP)类

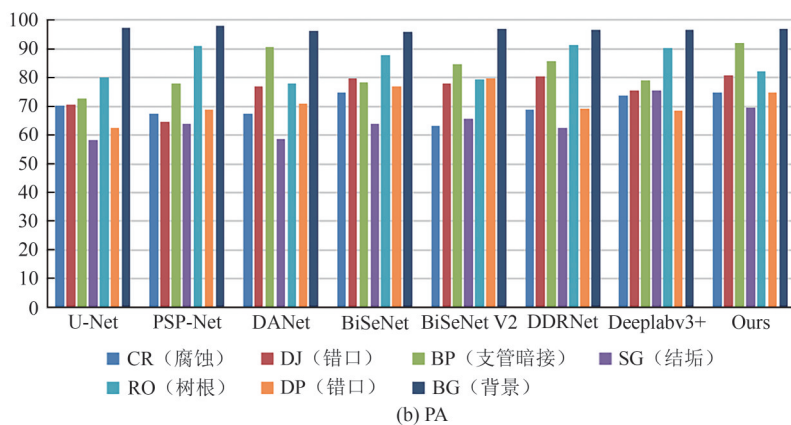
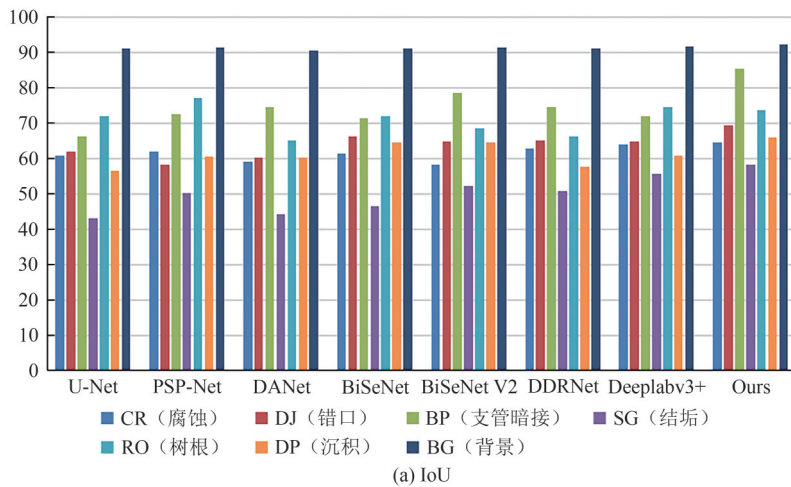


图 11 不同模型对各类别缺陷的分割 IoU 和 PA 对比

Fig. 11 Comparison of IoU and PA of different models on different defect categories

别上, IoU 达到了 85.53%。结合两张图可以推断出, 在树根缺陷类别上, PSP-Net 拥有最好的分割效果。DDRNet 在树根缺陷类别上拥有不错的召回率, BiSeNetV2 对各类别的分割精度较均衡。

图 12 展示了不同语义分割算法在城市地下管道缺陷图像上的分割效果, 其中第 1 行样本为管道错口缺陷, 第 2 行样本为错口和沉积缺陷, 第 3 行样本为腐蚀和沉积缺陷, 第 4 行样本为支管暗接和管道结垢缺陷, 第 5 行样本为管道错口和树根缺陷。由图 12 可知, 本文改进的算法对比几种常用的语义分割算法, 在分割精度 (mIoU) 和召回率 (mPA)

上都拥有最好的表现。对比第 1 行可以发现 UNet 和 PSPNet 对错口缺陷检测效果不佳, DANet 和 BiSeNetV2 都存在误检的情况; 第 2 行样本中 UNet 和 PSPNet 没有识别出错口缺陷; 对比第 3 行样本, PSP-Net 没有识别出腐蚀区域, DANet 效果不佳, BiSeNetV2 没有识别出沉积区域; 在第 4 行的样本上, DANet 对结垢类别分割效果较优, 对支管暗接类本文模型拥有最好的效果; 第 5 行的管道错口和树根两个缺陷类别上只有 DDRNet 接近本文方法的分割效果, 但仍在部分区域误测出了错口缺陷。

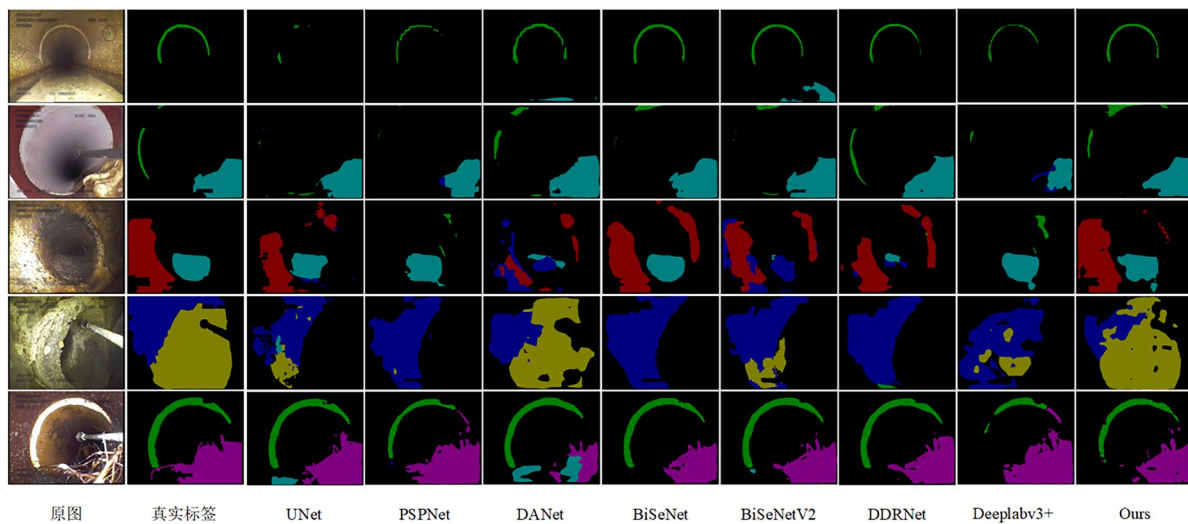


图 12 对比实验分割结果可视化

Fig. 12 Visualization of segmentation results in contrast experiments

3.3.2 消融实验

为了验证本文添加的各个模块的有效性, 在原 Deeplabv3+ 算法中依次添加改进的几种模块, 以基本的 Deeplabv3+ 算法为基线, 对改进的 Fusion Module 和 RCCA Module 以及对比深度监督策略进行了一

系列消融实验, 结果如表 2 所示。最左侧的列为添加的模块, 其中 FS 代表 Fusion Module, RCCA 代表 RCCA Module, CDS 代表对比深度监督策略。中间 7 列为腐蚀、管道错口、支管暗接、管道结垢、树根、沉积、背景的分类 IoU, 最右侧 4 列为模型评价指标。

表 2 各模块消融实验

Tab. 2 Ablation experiments between modules

Method	IoU/%							mIoU/%	mPA/%	Params	FPS
	CR	DJ	BP	SG	RO	DP	BG				
Baseline	64.06	64.99	72.22	55.69	74.61	60.85	91.85	69.18	79.95	40.4M	48.50
+FS	62.66	64.73	78.24	55.23	74.47	66.52	91.94	70.54	79.65	40.8M	42.46
+RCCA	66.27	64.15	74.04	58.83	74.21	67.84	92.08	71.06	81.16	41.2M	37.80
+FS&RCCA	65.49	64.64	80.91	59.59	72.75	68.79	92.33	72.07	82.71	41.7M	26.13
+FS&RCCA&CDS	64.65	69.48	85.53	58.45	73.73	66.17	92.35	72.91	81.62	41.7M	26.13

由表 2 可知, 使用原始 Deeplabv3+ 模型的 mIoU 为 69.18%, 在 Deeplabv3+ 中添加改进的模块, 模型的分割 mIoU 都会有所提升, 只加入改进的上采样 Fusion Module 后 mIoU 提升了 1.36%, 只加入带有十字交叉注意力的 RCCA 模块后 mIoU 提升了 1.88%,

同时使用 Fusion Module 和 RCCA Module 后模型的 mIoU 提升了 2.89%; 同时使用 Fusion Module 和 RCCA 模块, 并加入对比深度监督策略后, mIoU 提升了 3.73%。模型引入的 RCCA 模块在提升分割精度的同时, 还会增加模型的参数, 因此加大了运行

的时间开销,目前基于深度学习的城市地下管道缺陷分割相关工作普遍存在精度不够高的问题,所有应更加重视分割的精确率的提升.本文提出的改进后的模型能以每秒26张的速度处理图片,证明能够支持实时分割.

图 13 展示了依次改进后的算法和原始 Deeplabv3+模型在城市地下管道缺陷图像上的分割效果,其中第 1 行样本为管道结垢缺陷,第 2 行样本为错口和沉积缺陷,第 3 行样本为错口和腐蚀缺陷,第 4 行样本为支管暗接缺陷,第 5 行样本为错口和树根缺陷.由图可知,对比原始方法,加入 3 种改进的方法在各类别的分割精度和召回率上都有优于原始算法的表现.对比第 1 行图像可以发现,原始模型和只加入一种改进的模型都无法很好地识别出图像右上角模糊缺陷区域,而同时加入 FS 和 RCCA

两种改进的模型虽然能检测出模糊区域的缺陷,但是对于缺陷的部分区域的识别效果不如加入 3 种改进方法的模型.对比第 2 行图像可以发现,原始模型和只加入一种改进的模型都存在误识别的情况,将错口识别成了支管暗接,加入两种改进的模型虽然正确识别出了两种缺陷,但对错口区域的缺陷的精确度明显不如加入 3 种改进方法后的模型.对比第 3 行图像可以发现加入 RCCA 改进的模型对腐蚀区域的马赛克边缘能够分割出来,其效果超过了同时加入两种改进(FS、RCCA)的模型.对比第 4 行可以发现,只有加入 3 种改进的模型对支管暗接缺陷的马赛克区域能够有较好的边缘分割效果.对比第 5 行图像可以发现,原始模型、加入一种改进的模型、同时加入两种改进的模型,都对错口区域产生了误判.

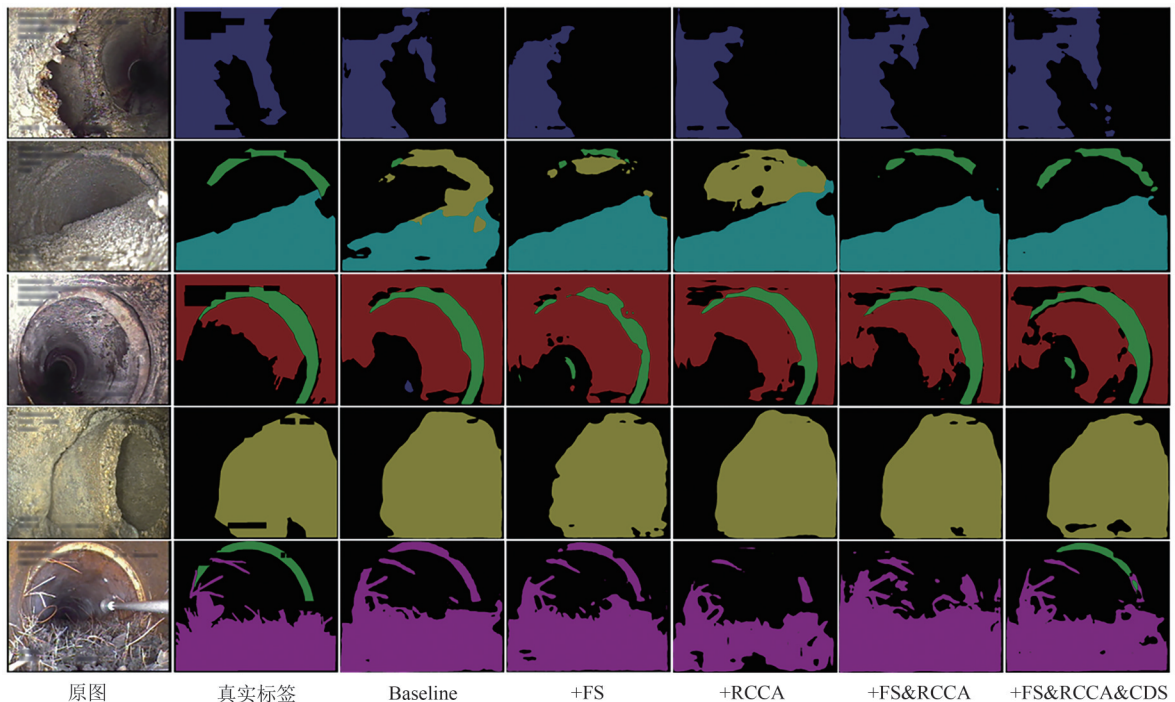


图 13 消融实验分割结果可视化

Fig. 13 Visualization of segmentation results in ablation experiments

4 总结

本文结合了多尺度融合上采样策略、交叉交叉注意力机制、对比深度监督的方法和 Deeplabv3+模型提出了基于 FCC-Deeplabv3+的城市地下管道缺陷语义分割方法,在模型的解码器部分改进了包含逐级上采样的特征融合模块 Fusion Module,增加了不同尺度的特征融合,在模型解码器部分融入多尺度信息,并分别对编码器解码器加入了对比深度

监督策略.对于目前没有城市地下管道缺陷语义分割数据集的情况,基于 Sewer-ML 数据集标注了用于该任务的数据集.在该数据集上的实验结果表明:基于交叉交叉注意力和对比深度监督的 FCC-Deeplabv3+模型具有良好的语义分割能力.在后续相关工作中,将从以下方面进行改进:(1)城市地下管道缺陷分割图像存在大量噪声和对比度问题,可以有针对性地引入降噪和对比度调整算法,优化数据集图像的质量,提高模型的分割效果;(2)模型在

分割精度提升的同时,参数量也变得更大,对图像的处理速度也变慢,后续将研究如何提高模型的推理速度.

参 考 文 献

- [1] WANG J, LIU G H, WANG J, et al. Current status, existent problems, and coping strategy of urban drainage pipeline network in China[J]. *Environmental Science and Pollution Research International*, 2021, 28(32): 43035-43049.
- [2] SADEGHKHAH A, AHMED E, KREBS P. Towards a decentralized solution for sewer leakage detection - a review[J]. *Water Science and Technology*, 2022, 86(5): 1034-1054.
- [3] HAURUM J B, MOESLUND T B. A survey on image-based automation of CCTV and SSET sewer inspections [J]. *Automation in Construction*, 2020, 111: 103061.
- [4] CHANDURE P, WAGHMARE C, ANSARI K. Review on guidelines of sewer inspection[J]. *CHEMIK*, 2020, 2: 160-167.
- [5] MORADI S, ZAYED T, GOLKHOOF. Review on computer aided sewer pipeline defect detection and condition assessment[J]. *Infrastructures*, 2019, 4(1): 10.
- [6] LI Y, WANG H, DANG L M, et al. Vision-based defect inspection and condition assessment for sewer pipes: A comprehensive survey[J]. *Sensors*, 2022, 22(7): 2722.
- [7] SU T C, YANG M D. Application of morphological segmentation to leaking defect detection in sewer pipelines[J]. *Sensors*, 2014, 14(5): 8686-8704.
- [8] SU T C, YANG M D, WU T C, et al. Morphological segmentation based on edge detection for sewer pipe defects on CCTV images[J]. *Expert Systems with Applications*, 2011, 38(10): 13094-13114.
- [9] SUN L, ZHU J, TAN J, et al. Deep learning-assisted automated sewage pipe defect detection for urban water environment management [J]. *Science of the Total Environment*, 2023, 882: 163562.
- [10] WANG M, CHENG J C P. Semantic segmentation of sewer pipe defects using deep dilated convolutional neural network [C]//*Proceedings of the International Symposium on Automation and Robotics in Construction (IAARC) Banff: IAARC*, 2019.
- [11] HAURUM J B, MOESLUND T B. Sewer-ML: A multi-label sewer defect classification dataset and benchmark [C]//*2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Nashville. IEEE*, 2021: 13456-13467.
- [12] ZHOU Q, SITU Z, TENG S, et al. Automatic sewer defect detection and severity quantification based on pixel-level semantic segmentation [J]. *Tunnelling and Underground Space Technology*, 2022, 123: 104403.
- [13] PAN G, ZHENG Y, GUO S, et al. Automatic sewer pipe defect semantic segmentation based on improved U-Net[J]. *Automation in Construction*, 2020, 119: 103383.
- [14] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [M]//*Computer Vision - ECCV 2018. Cham: Springer International Publishing*, 2018: 833-851.
- [15] HUANG Z, WANG X, HUANG L, et al. CCNet: Criss-cross attention for semantic segmentation [C]//*2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul. IEEE*, 2019: 603-612.
- [16] ZHANG L, CHEN X, ZHANG J, et al. Contrastive deep supervision [M]//*Lecture Notes in Computer Science. Cham: Springer Nature Switzerland*, 2022: 1-19.
- [17] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [18] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE*, 2016: 770-778.
- [19] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE*, 2017: 2117-2125.
- [20] NIU Z, ZHONG G, YU H. A review on the attention mechanism of deep learning[J]. *Neurocomputing*, 2021, 452: 48-62.
- [21] ZHANG H, DANA K, SHI J, et al. Context encoding for semantic segmentation [C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE*, 2018: 7151-7160.
- [22] DING H, JIANG X, SHUAI B, et al. Context contrasted feature and gated multi-scale aggregation for scene segmentation [C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE*, 2018: 2393-2402.
- [23] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE*, 2017: 2881-2930.

- [24] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City. IEEE, 2018: 7794-7803.
- [25] LEE C Y, XIE S, GALLAGHER P, et al. Deeply-supervised nets[C]// Artificial Intelligence and Statistics. San Diego: JMLR, 2015: 562-570.
- [26] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234-241.
- [27] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 3146-3154.
- [28] YU C, WANG J, PENG C, et al. BiSeNet: bilateral segmentation network for real-time semantic segmentation[M]//Computer Vision - ECCV 2018. Cham: Springer International Publishing, 2018:325-341.
- [29] YU C, GAO C, WANG J, 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.
- [30] PAN H, HONG Y, SUN W, et al. Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes [J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(3): 3448-3460.

(责编 曹东,校对 姚春娜)