

基于CT-ST半监督模型的城市地下管道缺陷语义分割研究

潘禹欣,李波*,田淙文,姚为

(中南民族大学 计算机学院,武汉 430074)

摘要 利用机器视觉技术进行城市地下管道缺陷分割是工业化智能化发展趋势.由于常规的监督方法进行缺陷分割任务时需要大量的标注,提出了一种基于ST半监督模型改进的CT-ST半监督语义分割模型,并首次应用到城市地下管道缺陷分割领域.该模型基于半监督语义分割领域自训练方法,结合Co-teaching算法思想,区分不同质量伪标签,利用一次伪标签筛选策略代替传统设置阈值迭代方法,降低因低质量标签带来的错误特征训练影响;针对地下管道背景复杂、缺陷类别多、多尺度、多噪声等问题,在每个残差块引入NAM注意力机制,给每个重要特征增加权重,弱化不重要特征的占比.实验验证了CT-ST半监督分割模型的有效性,在不同比例有标签样本集上mIoU均有提升,其中1/2比例有标签数据集mIoU为67.36%,对比原模型增加了2.33%.与多种主流的伪标签、一致性正则化方法相对比,所提出的模型在精度上均有较好的表现.

关键词 半监督学习;ST模型;Co-teaching算法;注意力机制;伪标签;地下管道缺陷;缺陷分割

中图分类号 TP391 文献标志码 A 文章编号 1672-4321(2026)02-0221-10

doi:10.20056/j.cnki.ZNMDZK.20250848

Semantic segmentation of urban underground pipeline defects based on CT-ST semi-supervised model

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

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

Abstract Defect segmentation of urban underground pipes using machine vision technology is an industrialized intelligent development trend. Since conventional supervised methods require a large number of annotations for defect segmentation task, an improved CT-ST semi-supervised semantic segmentation model based on the ST semi-supervised model is proposed, which is firstly applied to the field of defect segmentation of urban underground pipelines. The model is based on the self-training method of semi-supervised semantic segmentation domain, combined with the idea of Co-teaching algorithm, distinguishes different quality pseudo-labels, and utilizes a one-time pseudo-label screening strategy instead of the traditional set-threshold iterative method, to reduce the impact of erroneous feature training due to low-quality labels; for the problems of complex background of underground pipelines, multiple defect categories, multiple scales, and multiple noises, we introduce a NAM attention mechanism into each residual block, to give each important defects a more accurate and more accurate labeling. NAM attention mechanism is introduced in each residual block to increase the weight of each important feature and weaken the proportion of unimportant features. The experiments verify the effectiveness of CT-ST semi-supervised segmentation model, and the mIoU is improved on different proportions of labeled sample sets, in which the mIoU of 1/2 proportion of labeled dataset is 67.36%, which is increased by 2.33% compared with the original model. Compared with many mainstream pseudo-labeling and consistency regularization methods, CT-ST has better performance in terms of accuracy.

收稿日期 2024-03-05

*通信作者 李波(1975-),男,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)

Keywords semi-supervised learning; ST model; Co-teaching algorithm; attention mechanism; pseudo labels; defects in urban underground pipelines; defect segmentation

地下管道系统是城市现代化建设的重要基础设施.随着使用年限的增长,管段呈现出一些缺陷(变形、堵塞、渗漏、坍塌等)^[1],而这些缺陷会导致地面塌陷、内涝积水、污水冒溢等事故,所以定期对管道内部进行检测是维护管道系统正常运行的必要手段.

随着科技进步和自动化技术不断发展,管道缺陷识别经历了三个阶段,分别为传统人工检测阶段,机器辅助检测阶段和深度神经网络检测阶段^[2].

传统人工检测方法主要有四方面缺点:人工检测花费时间长;管道缺陷种类多,受检测人员的主观影响,容易出现缺陷误判的情况;管道内部环境复杂,人工识别能力有限,容易漏检缺陷;长时间进行管道内部作业会影响检测人员身体健康.针对传统人工检测的缺点,管道缺陷检测进入机器辅助阶段,从业人员使用超声波、探地雷达、激光、闭路电视(Closed Circuit Television, CCTV)、管道潜望镜(Quick Video, QV)^[3-4]等方法进行缺陷检测.这些方法虽然优化了管道内部样本的采集方式,但仍需要技术人员对管道内部情况进行分析,自动化程度较低,不利于管道维护.

近年来,基于深度学习的检测技术因其良好的准确度和鲁棒性在管道缺陷检测任务上得到了应用,减轻人工作业的负担,而只对地下管道缺陷识别并不能满足工业需求,还需要将缺陷按照区域进行分割,提升缺陷分割的准确度也成为深度学习重要研究目标.与此同时,传统的深度学习模型需要充足的标注信息,目前业内并无针对地下管道缺陷分割任务的公开数据集,管道缺陷语义分割领域缺少半监督分割模型相关应用.针对上述问题,本文以半监督ST模型^[5]为基线进行优化改进,提出一种基于伪标签筛选方法的CT-ST半监督分割模型,主要贡献包括以下三个方面:

(1)在自训练模型的基础上增加伪标签筛选环节,引入Co-teaching^[6]思想,将生成的伪标签通过两个网络相互监督的方式,抉择出质量更好的伪标签.

(2)考虑到不同特征的重要程度不同,引入了NAM注意力机制^[7],给每个重要特征增加权重,弱化不重要特征的占比,从而提升模型分割精度.

(3)针对目前缺少像素级别语义分割标注的地

下管道缺陷公开数据集的现状,首次将半监督分割方法应用到城市地下管道缺陷分割领域,实现了一个具有实用价值的CT-ST半监督管道缺陷语义分割模型,并通过实验验证了模型的有效性.

1 相关工作

1.1 管道缺陷分割方法

随着神经网络的发展,一系列高性能的分割模型被相继提出,成为语义分割任务的主流解决方案^[8],FCN^[9],U-Net^[10],PSPNet^[11],SegNet^[12]和DeepLab^[13-15]等模型被充分应用到管道缺陷分割当中,FCN实现了端到端、像素对像素的分割,但是对图像细节不敏感,分割精度不够高;U-Net是一种先下采样再上采样的U形结构分割网络,通过跳跃连接方法融合不同层的特征;PSPNet采用PPM更好的获得全局先验信息;SegNet为对称结构,通过编码到解码的方式实现分割,这种轻量级的网络结构训练过程简单、训练速度快、可扩展性强,可用于多种分割任务;Deeplab系列提出空洞空间池化金字塔模块,以捕捉不同感受野下的空间细节信息,更好的应对多尺度目标分割问题.WANG等人^[16]提出DilaSeg语义分割网络来分割下水道缺陷,并采用条件随机场(Conditional Random Field, CRF)对原始网络进行了改进,提出DilaSeg-CRF^[17].PAN等人^[18]在U-Net的基础上提出PipeUNet来解决高级特征和低级特征之间的语义差距.LI等人^[19]通过引入Res2Net-Mish-BN-101高效骨干结构,并改进BiFPN设计出增强的EBiFPN,从而提出Pipe-SOLO管道缺陷实例分割模型,该模型能够有效定位和分割缺陷.

1.2 半监督分割方法

半监督方法介于完全监督学习和无监督学习之间,仅处理部分标注的数据集,半监督深度学习分割模型主要分为五种方法:生成对抗方法、一致性正则化方法、对比学习方法、伪标签方法和混合方法^[20].

经典的生成对抗网络由生成器和鉴别器两个网络组成,GCT^[21]、S4GAN^[22]均利用对抗思想实现分割任务,其分割网络承担生成器的角色,鉴别器将分割网络生成的分割图与真实分割图区分开.一致性正则化方法,基于平滑假设和聚类假设,利用相

似的数据点具有相似输出的特性进行图像分割. CHEN 等人^[23]提出的 CPS 模型使用两个相同结构、但是不同初始化方式的网络,添加约束使得两个网络对同一样本的输出相似. LIU 等人^[24]对经典的平均教师模型进行改进,引入一个新的辅助教师分支,用更严格的置信度加权交叉熵损失代替均方误差,提出 PS-MT 模型. ALONSO 等人^[25]基于对比学习提出 MemoryBank 模型,将标记和未标记数据中的类和像素特征与存储库中已标记的所有特征进行对齐,该模型在半监督语义分割和半监督领域自适应方面取得良好效果. 伪标签方法主要有两种模式,一种是集成多个不同网络来提高整个框架的性能,另一种是自训练,ST 模型简化了自训练方法,取消了迭代过程,在有标签的样本上训练一个初始的教师模型,再为没有标签的样本生成伪标签,最后将有标签样本和无标签样本混合训练,并在半监督训练阶段,利用 colorjitter、blur、grayscale 和 Cutout 对无标签样本进行强数据增强. 混合方法是将多种半监督方法进行综合应用提升分割精度.

半监督分割可以有效地在少量标注数据集基础上进行实验,同时管道缺陷分割领域缺少像素级标注数据集,符合半监督学习的应用条件,利用半监督的方式训练模型可以在一定程度上改善管道缺陷分割任务标注少的问题. 在标注数据量少但原始样本数据量充足的情况下,伪标签方法可以有针对性地应对管道缺陷分割任务的现状.

2 本文提出的 CT-ST 半监督管道缺陷分割模型

2.1 模型整体结构

本文模型以 ST 半监督分割模型为基础,整体流程如图 1 所示,其中红色虚线框区域为改进部分. 为提高管道缺陷分割精度,对分割网络部分增加 NAM 模块,图中表现为 NAM-初始教师模型、NAM-可靠教师模型、NAM-学生模型. 增加伪标签筛选环节,利用 Co-teaching 思想作为筛选高质量标签的依据,并沿用了 ST 模型的强增强方法. 改进的 ST 模型先将数据集划分为有标签数据集和无标签数据集,第一阶段在有标签数据集上进行完全监督训练,生成初始教师模型,使用该初始教师模型为无标签数据集生成伪标签. 第二阶段将有标签数据和伪标签数据混合输入 Co-teaching 网络中,在 Co-teaching 的双重网络监督下被遗忘次数越高的伪标签样本越不稳定,被遗忘次数低的样本质量较高,选取质量高的样本组成高质量伪标签数据集. 第三阶段为半监督训练阶段,将高质量伪标签数据集通过强增强后与有标签数据集混合,作为半监督训练的输入数据集,训练一个可靠教师模型,利用该模型为第二阶段淘汰的样本生成伪标签,即为无标签数据集全部生成伪标签. 第四阶段将有标签样本、高质量伪标签数据集和可靠伪标签数据集全部混合,作为最终训练的数据集,训练出的学生模型作为最终测试模型.

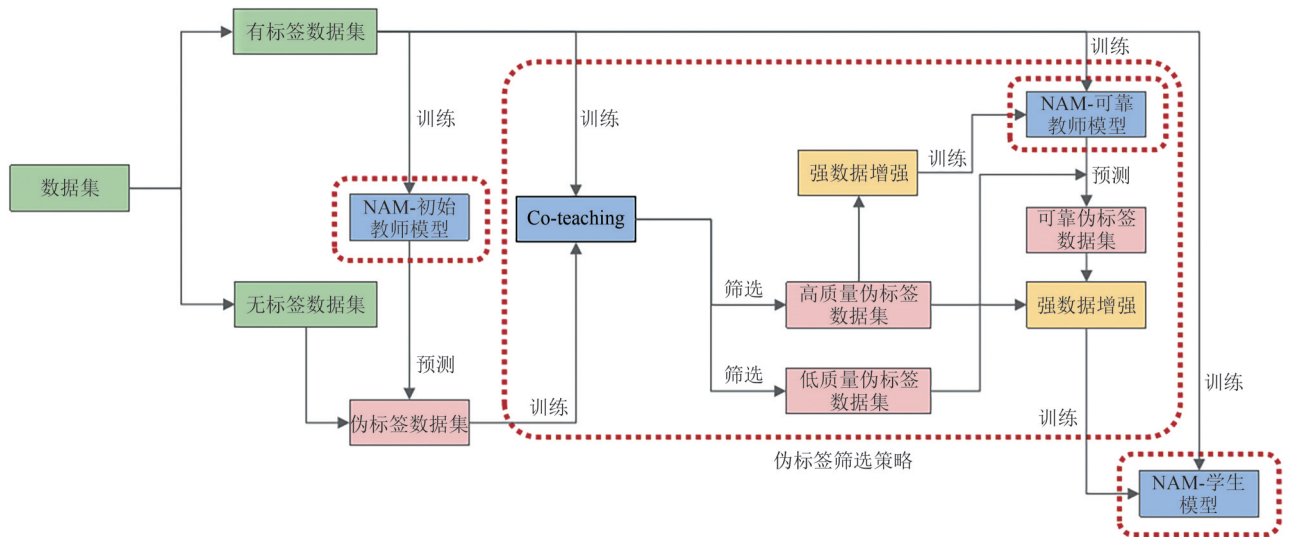


图 1 CT-ST 模型流程图

Fig. 1 Structure of CT-ST

2.2 伪标签筛选策略

ST 模型在训练过程中没有通过设置阈值的方

法进行迭代训练,而采用预测一次伪标签方法进行半监督训练,这种方法降低了时间和内存的消耗,

但生成的伪标签质量不均衡. 神经网络有先记忆干净数据, 后记忆噪声数据的特性, ST 模型利用全部生成的伪标签进行半监督训练, 训练过程中生成的质量较差的伪标签相当于噪声标签, 这些质量较差的伪标签将会对模型训练结果造成影响. 本文引入 Co-teaching 思想对 ST 模型进行改进, 减少低质量标签带来的错误特征学习. Co-teaching 同时训练两个神经网络 A 和 B, A 和 B 在每个 mini-batch 中的相互教学, 如图 2 所示, 首先每个网络前馈所有数据, 并选择高质量标签的数据, 之后两个网络相互交换较高质量数据信息, 记录当前 mini-batch 中哪些数据应该用于训练, 最后每个网络将其对等网络选择的数据进行反向传播用于更新自己的参数.

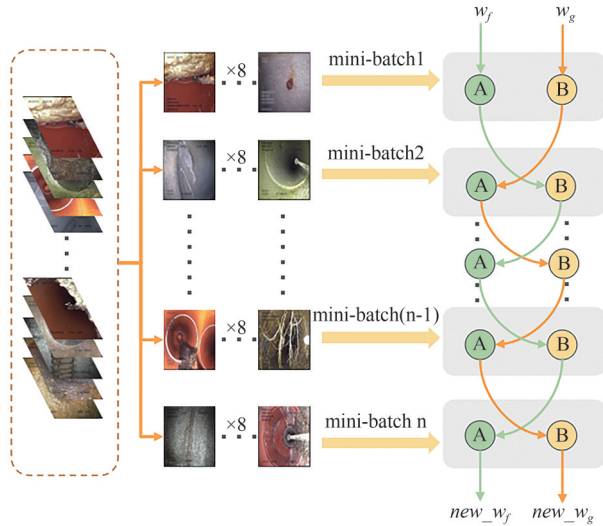


图 2 Co-teaching 流程图

Fig. 2 Structure of Co-teaching

本文改进的 ST 模型采用 Co-teaching 的思想作为筛选伪标签的评判标准, 算法流程如算法 1 所示^[6], 定义两个网络 f 和 g , 两个网络的参数分别为 w_f 和 w_g , 含有噪声标签的数据集为 D , 每轮训练创建 mini-batch 为 \bar{D} , 网络 f 和网络 g 分别选择 \bar{D} 中 loss 较小的实例, 被不同网络选择的实例集合分别被定义为 \bar{D}_f 和 \bar{D}_g , 而被选择的实例的多少由 $R(T)$ 选择率决定, $R(T)$ 可表示为公式 (1):

$$R(T) = 1 - \min \left\{ \frac{T}{T_k} \tau, \tau \right\}, \quad (1)$$

其中, τ 为遗忘率, 且 τ 属于 $[0, 1)$, T_k 为遗忘率稳定轮次, 即超过 T_k 轮次后 f 和 g 网络每轮选择的实例数量不再更改, 每一个轮次结束后更新 $R(T)$, 记录两个网络被丢弃的实例 \bar{D}_l , 然后将两个网络所选择的实例送到其对等网络中, 作为参数更新的有用知识. 统计每轮 \bar{D}_l 中遗忘样本的集合为 D_l , 计算 D_l 中每个

样本的出现频率. 被遗忘次数越高的样本质量越差, 被遗忘次数低的样本伪标签质量较高, 选择质量高的标签作为 ST 模型下一轮训练的输入样本. 统计所有轮次 \bar{D}_l 的样本, 选择多个轮次都被遗忘的样本作为低质量图像样本集合, 不参与高质量伪标签数据集训练, 通过可靠教师模型为低质量图像集合生成伪标签, 参与最终学生模型训练.

算法 1 Co-teaching 算法

Algorithm. 1 Co-teaching algorithm

Co-teaching Algorithm.

Input w_f and w_g , learning rate η , forgetting rate τ , epoch T_k and T_{\max} , iteration N_{\max} ;

Output w_f 和 w_g , D_l

for $T = 1, 2, \dots, T_{\max}$ do

Shuffle training set D ;

for $N = 1, 2, \dots, N_{\max}$ do

Fetch mini-batch \bar{D} from D ;

Obtain $\bar{D}_f = \operatorname{argmin} D': |D'| \geq R(T) |\bar{D}| l(f, D')$;

Obtain $\bar{D}_g = \operatorname{argmin} D': |D'| \geq R(T) |\bar{D}| l(g, D')$;

Obtain $\bar{D}_l = (\bar{D} - \bar{D}_f) + (\bar{D} - \bar{D}_g)$;

Update $D_l += \bar{D}_l$;

Update $w_f = w_f - \eta \nabla l(f, \bar{D}_g)$;

Update $w_g = w_g - \eta \nabla l(g, \bar{D}_f)$;

end

Update $R(T) = 1 - \min \left\{ \frac{T}{T_k} \tau, \tau \right\}$

end

2.3 NAM 注意力机制模块

许多注意力机制的研究都着力于利用不同维度特征信息来寻找显著特征, 忽略了不显著特征对模型的负面影响. NAM 注意力机制认为要考虑不同维度特征信息贡献因素的权重, 进一步抑制不重要的特征. NAM 是一种轻量级和高效的注意力机制, 与 SE^[26]、CA^[27] 和 CBAM^[28] 注意力机制不同的是, NAM 没有使用完全连接层和卷积层, 一定程度上减轻了模型计算负担. 本模型骨干网络采用的是 ResNet-50, 在每个残差块中添加 NAM 注意力机制, 有助于提高模型的训练精度, 插入 NAM 模块的残差结构如图 3 所示.

NAM 采用 CBAM 注意力机制的模块集成方式, 重新设计了通道注意力和空间注意力, 使用批量归一化 (BN) 的缩放因子, 该缩放因子使用标准偏差来表示权重的重要性, 如公式 (2) 所示:

$$B_{\text{out}} = \text{BN}(B_{\text{in}}) = \gamma \frac{B_{\text{in}} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} + \beta, \quad (2)$$

其中, μ_B 和 σ_B 分别为小批量 B 的平均值和标准偏

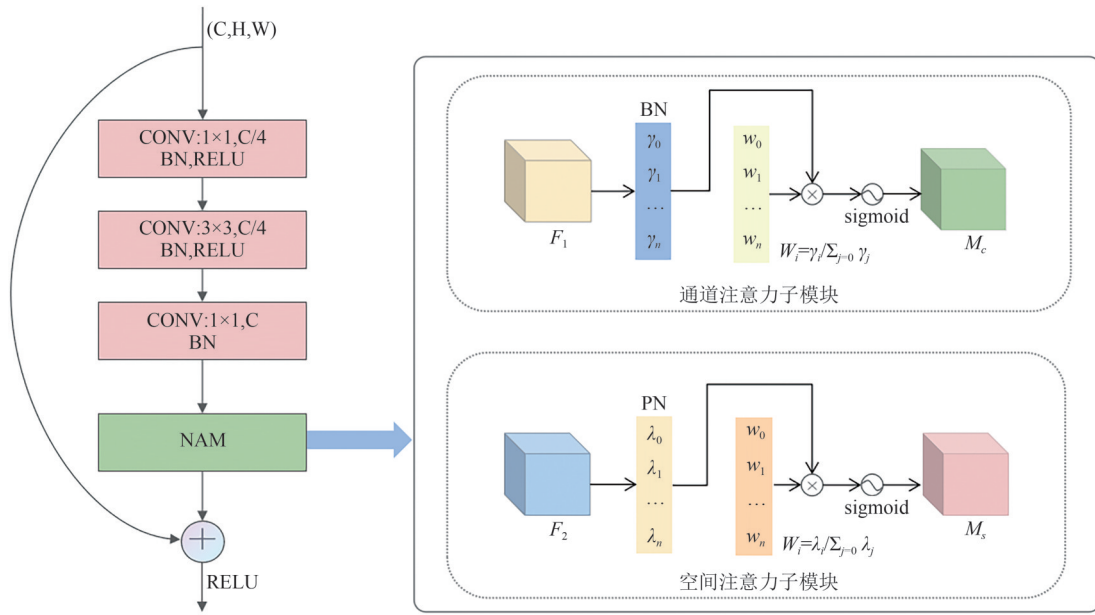


图 3 ResNet-50 插入 NAM 模块后的残差结构

Fig. 3 Residual structure of ResNet-50 after insertion into the NAM module

差 γ 和 β 是可训练的仿射变换参数, 分别为尺度和偏移. ϵ 为极小常数, 防止分母为 0. 通道注意力子模块可表示为公式(3):

$$M_c = \text{sigmoid}(W_\gamma(\text{BN}(F_1))), \quad (3)$$

其中, F_1 为输入特征, M_c 表示输出特征. γ 是每个通道的比例因子, 其权重为 $W_\gamma = \gamma_i / \sum_{j=0}^n \gamma_j$, 结构如图 3 中通道注意力子模块所示.

NAM 还将 BN 的比例因子应用于空间维度, 以测量像素的重要性, 命名为像素归一化(PN). 相应的空间注意力子模块如图 3 和公式(4)所示:

$$M_s = \text{sigmoid}(W_\lambda(\text{PN}(F_2))), \quad (4)$$

其中, 输出特征为 F_2 , 输出为 M_s . λ 是比例因子, 权重为 $W_\lambda = \lambda_i / \sum_{j=0}^n \lambda_j$.

为了抑制不太显著的权重, NAM 在损失函数中添加了一个正则化项, 如式(5)所示:

$$\text{Loss} = \sum_{(x,y)} l(f(x, W), y) + p \sum g(\gamma) + p \sum g(\lambda), \quad (5)$$

其中 x 表示输入, y 是输出, W 表示网络权重, $l(\cdot)$ 是损失函数; $g(\cdot)$ 是 l_1 范数, p 用来平衡 $g(\gamma)$ 和 $g(\lambda)$.

3 实验及结果

3.1 数据集及实验环境

3.1.1 实验数据来源

本文实验采用的是 CVPR 会议论文提出的 Sewer-ML 公开数据集^[29], 该数据集主要用于多分类

任务, 而缺陷分类和本文的缺陷分割是两个完全不同的研究任务, 对于本研究而言, 原始数据集只提供了地下管道的图片素材集, 并未提供缺陷分割结果的标注信息, 因此不能直接用于本文的缺陷分割实验. 本文采用由课题组进行像素级标注的 900 张样本为原始数据进行实验. 缺陷代码与原始数据集提供的专业地下管道缺陷标注文档保持一致, 样本分为腐蚀(CR)、错口(DJ)、支管暗接(BP)、管道结垢(SG)、沉积(DP)和树根(RO)六种缺陷类和背景(BG)类, 共 7 种主要常见类别. 6 类缺陷样本是典型的管道缺陷样本, 其中腐蚀为材料性缺陷, 错口和支管暗接为结构性缺陷, 沉积导致阻塞为功能性缺陷, 比较充分的包括了多种情况导致的管道缺陷. 缺陷样本与标记如图 4 所示.

3.1.2 数据预处理

将原始数据集按 7:1:2 划分为训练集、验证集和测试集, 由于数据过少并结合 ST 模型训练流程中的强弱数据增强特性, 将训练集、验证集和测试集独立进行数据增强, 避免训练集、验证集和测试集样本重复导致模型训练精度虚高的问题.

本文采用水平翻转、垂直翻转、随机角度旋转、剪裁以及相互融合的方式进行数据增广, 如图 5 所示, 经过划分在增强后的训练样本数为 3780 张, 验证样本数为 540 张, 测试集样本数为 1080 张. 通过多种数据增强方法增广后的数据集, 数据量线性增长为原始数据的 6 倍, 有助于提高模型的泛化能力.

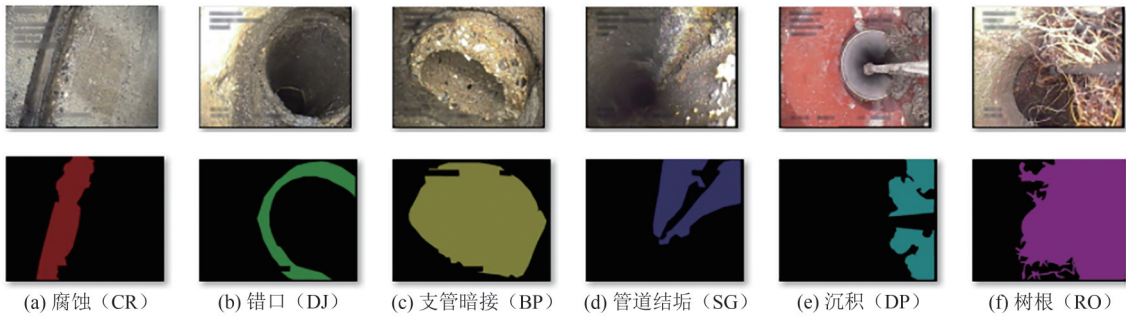


图4 地下管道缺陷样本类型

Fig. 4 Underground pipeline defects samples

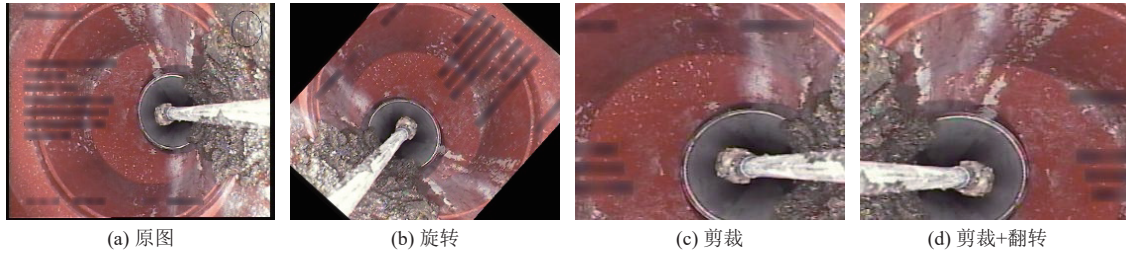


图5 图像增强

Fig. 5 Image enhancement

3.1.3 实验环境与参数设置

实验环境主要包括:硬件环境为Linux操作系统, CPU为Intel Xeon E5-2630 v4, GPU为NVIDIA Corporation Tesla P40, 显存为24 GB, 内存为32 GB; 软件环境为Python3.9、Pytorch1.12和CUDA11.7的深度学习框架. 完全监督和半监督训练阶段的初始学习率设为0.001, 优化器采用SGDM, 动量设置为0.9, 权重衰减设置为 1×10^{-4} , 批处理大小设置为16, 迭代80次. 在伪标签筛选阶段, Co-teaching网络的初始学习率 η 设置为 5×10^{-6} , 采用adam优化器, 批处理大小设置为8, 总轮数 T_{\max} 设为80次, 遗忘率 τ 设为0.5, 遗忘率稳定轮次 T_k 设为20, 即20轮后每轮遗忘批处理大小一半的样本量.

3.2 评价指标

为了评价缺陷分割算法的性能, 常用的指标有交并比(IoU)、平均交并比(mIoU)、频率权重交并比(FWIoU)以及平均像素精度均值(mPA).

语义分割属于密集预测任务, 即对每个像素点进行分类, 在分类过程中会出现四种情况TP、FP、FN、TN, 其中TP表示将正样本预测为正类的数量, 即真阳性; FP表示将负样本预测为正类的数量, 即假阳性; FN表示将正样本预测为负类的数量, 即假阴性; TN表示将负样本预测为负类的数量, 即真阴性.

IoU用来衡量两个区域的重合程度, 是预测分割和标签之间的重叠区域除以预测分割和标签之

间的联合区域, 如公式(6)所示:

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}, \quad (6)$$

mIoU计算真实值和预测值两个集合的交集和并集之比, 一般都是基于类进行计算的, 将每一类的IoU计算之后累加, 再进行平均, 得到的就是基于全局的评价, 如公式(7)所示:

$$\text{mIoU} = \frac{1}{k} \sum_{i=0}^k \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}, \quad (7)$$

FWIoU为频率权重交并比, 是mIoU的改进版本, 它会根据每个分类出现频率, 对每个分类给予不同权重, 如公式(8)所示:

$$\text{FWIoU} = \sum_{i=0}^k \frac{\text{TP} + \text{FN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \times \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}, \quad (8)$$

mPA为平均像素精确度, 累计每个类别分类正确的像素数占有所有预测为该类别像素数的比例再平均, 如公式(9)所示:

$$\text{mPA} = \frac{1}{k} \sum_{i=0}^k \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}. \quad (9)$$

3.3 实验结果分析

3.3.1 不同半监督算法对比实验

为评估CT-ST的模型性能, 表1展示了包含不同比例的有标签样本在CT-ST模型中的mIoU、FWIoU、mPA, 并列出了不同缺陷种类的IoU精度, 其中mIoU为分割任务中最能体现分割准确性的评价指标.

表 1 本文 CT-ST 模型各类评价指标表现
Tab. 1 Performance of various evaluation indicators of CT-ST

Partition	mIoU/%	mPA/%	FWIoU/%	IoU/%						
				BG (背景)	CR (腐蚀)	DJ (错口)	BP (支管暗接)	SG (管道结垢)	RO (树根)	DP (沉积)
1/2	67.36	76.59	86.27	92.43	68.57	48.51	70.74	54.51	79.45	57.30
1/4	66.28	75.96	85.78	92.07	67.91	46.42	69.23	53.16	78.33	56.82
1/8	64.97	73.83	85.07	91.52	66.89	44.72	67.42	52.28	76.56	55.37
1/16	62.08	70.88	83.08	89.85	63.02	40.80	64.84	49.66	73.71	52.71

为了验证 CT-ST 模型对城市地下管道缺陷语义分割能力的准确性,将该模型和主流半监督语义分割模型进行对比实验,包括伪标签方法 ST 模型、ST++模型和 PCT^[30]模型,一致性正则化 CPS 模型和 PS-MT 模型,并为保证实验公平性,对比试验中的所有模型分割网络结构都采用 PSPNet,骨干网络均采用 ResNet-50,结果如表 2 所示,每一列分别表示半监督方法、分割网络和特征提取网络、不同比例有

标签样本的 mIoU.由表可知,CT-ST 模型在城市地下管道缺陷分割方向精度有不错的提升,在不同比例有标签样本中均有较好表现.

为直观感受不同半监督分割模型对城市地下管道缺陷的分割性能,在测试集上评估了不同半监督方法,图 6 为以 1/2 比例训练模型进行测试集可视化示例.第 1 列为原图、第 2 列为真实标签,3 至 8 列为对比模型分割效果,从第一行可以看出,其他方法在两类缺陷检测中均有混淆情况,而 CT-ST 对比真实标签表现较好;第二行中 CPS 模型仅能识别第一个错口,ST 和 ST++模型对第二个错口识别精度较差;第三行 CT-ST 对右下角沉积缺陷识别精度最高,但是对结垢(深蓝色)缺陷识别出现混淆;第四行只有 PCT 模型和本文模型能够对树根(紫色)和错口(绿色)两种缺陷分割较好,能对边缘信息进行细分;第五行 PS-MT 存在误检缺陷的情况,本文模型对暗接(黄色)分割比较完整,预测更加准确.

表 2 不同半监督模型对比实验

Tab. 2 Comparative experiments of different semi-segmentation model

Method	Network	mIoU/%			
		1/2	1/4	1/8	1/16
ST ^[5]	PSPNet+ResNet-50	65.03	64.57	63.43	60.23
ST++ ^[5]	PSPNet+ResNet-50	66.12	65.34	64.22	61.16
PCT ^[30]	PSPNet+ResNet-50	66.06	65.48	64.52	60.35
CPS ^[23]	PSPNet+ResNet-50	65.15	64.20	62.67	59.98
PS-MT ^[24]	PSPNet+ResNet-50	66.93	65.56	64.78	61.75
ours	PSPNet+ResNet-50	67.36	66.28	64.97	62.08

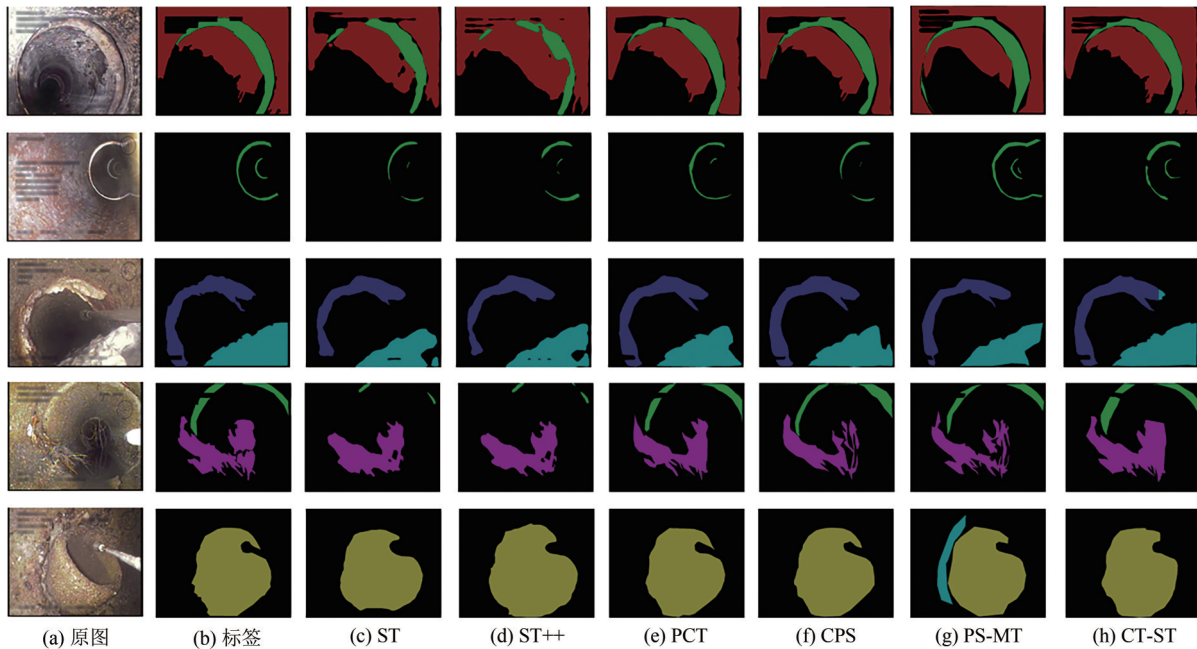


图 6 对比实验可视化结果

Fig. 6 Visualization of segmentation results in contrast experiments

3.3.2 消融实验

为了验证添加的各模块在 CT-ST 中的有效性, 本文在原 ST 模型中依次添加改进的模块, 对各模块之间做消融实验, 测试结果如表 3 所示. 表 3 中第一列和第二列为添加的模块, 分别是 Co-teaching 模块和 NAM 模块. 第三列为不同比例有标签样本的平均交并比. 由表 3 可知, 在 ST 模型中添加改进的模块, 精度都会有所提升, Co-teaching 算法在筛选标签方面起了正向作用, 高质量伪标签能帮助网络训练提

高训练精度. 添加 NAM 模块, 用赋予权重的方法, 导致网络对不重要特征进行忽略, 也对伪标签错误特征的学习起了阻碍作用, 抑制了错误特征的过度学习. 同时增加 Co-teaching 模块和 NAM 模块, 效果相对最好, 1/2 的有标签样本的 mIoU 由原来的 65.03% 提升到 67.36%, 1/4 的有标签样本的 mIoU 由原来的 64.57% 提升到 66.28%, 1/8 的有标签样本的 mIoU 由原来的 63.43% 提升到 64.97%, 1/16 的有标签样本的 mIoU 由原来的 60.23% 提升到 62.08%.

表 3 各模块之间的消融实验

Tab. 3 Ablation experiments between modules

Co-Teaching	NAM	mIoU/%			
		1/2	1/4	1/8	1/16
		65.03	64.57	63.43	60.23
√		66.70(+1.67)	65.85(+1.28)	64.46(+1.03)	61.59(+1.36)
	√	66.12(+1.09)	65.08(+0.51)	64.11(+0.68)	61.00(+0.77)
√	√	67.36(+2.33)	66.28(+1.71)	64.97(+1.54)	62.08(+1.85)

图 7 展示了改进后的 CT-ST 模型和 Baseline 在地下城市管道缺陷分割数据集上的分割结果, 其中蓝色线为 Baseline 分割结果、黄色线为添加 Co-Teaching 算法分割结果、绿色线为添加 NAM 注意力机制模块

分割结果、红色线为 NAM 与 Co-Teaching 共同作用分割结果, 由图 7 可知, 相比原算法, 两种模块在识别精度和分类准确率方面都有一定提升, 并且共同作用时模型效果达到最好.

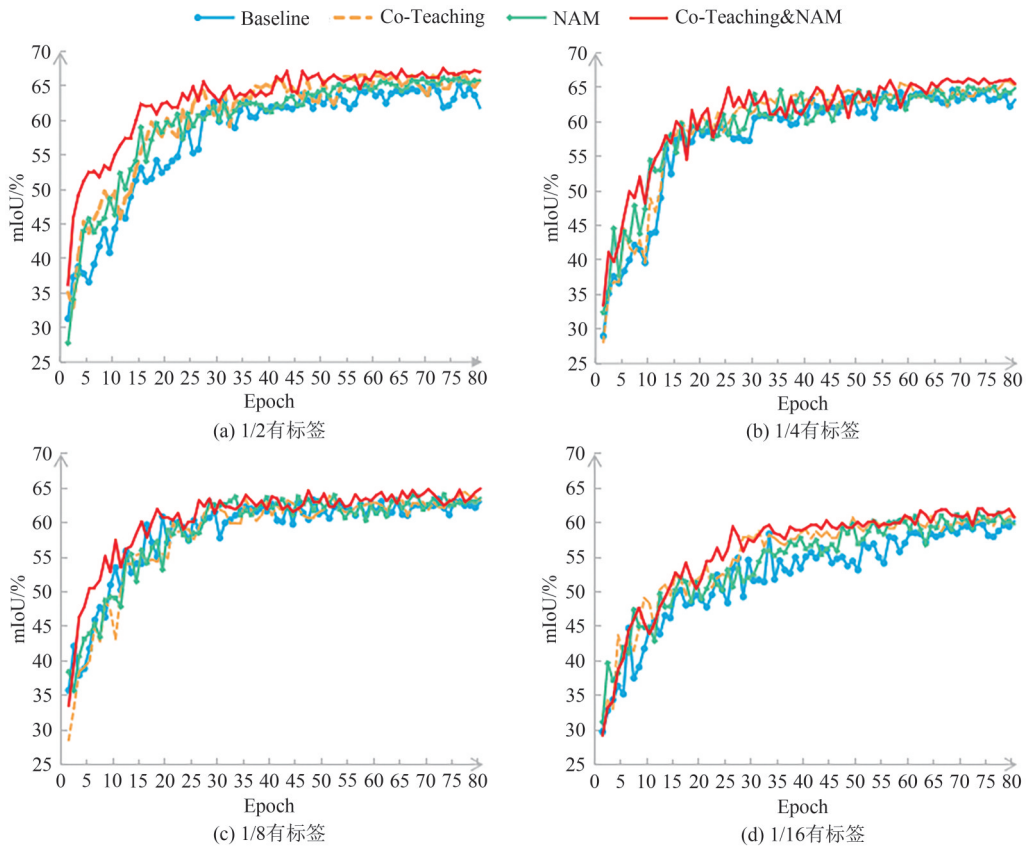


图 7 消融实验可视化结果

Fig. 7 Visualization of segmentation results in ablation experiments

4 结语

本文提出了一种新的用于语义分割的半监督学习方法,称为CT-ST半监督模型,该方法在传统自训练方法的基础上,利用Co-teaching思想筛选高质量的标签,充分利用了大量的未标记数据,通过NAM注意力机制降低不重要特征的学习比例,降低对伪标签错误特征的学习,从而提升整体训练精度.Co-teaching筛选标签部分基于两个不同初始化的并行网络,两个网络相互通信,传递本批次中使用哪些样本进行训练的信息,并通过对方网络中认为干净的数据来更新自己的网络参数,实验结果证明Co-teaching是保守可靠的,对不同比例有标签数据集的筛选都起了一定作用.NAM则对空间和通道两个模块进行设计,在通道注意力模块,通过缩放因子反映出各个通道的变化,方差越大变化越厉害,语义也就越丰富,那么该通道的重要性也越高,空间注意力模块也同样使用像素归一化的方法确定像素权重.后续的相关工作可以从以下两方面进行改进:(1)增加伪标签筛选环节虽然对精度提升有一定帮助,但是Co-teaching双重网络训练参数量过大,训练时长过长,后续可使用轻量级的筛选方法;(2)本文对错口和结垢两种缺陷识别准确率较低,原因在于样本不均衡,后续可采用对抗生成的半监督方式进行样本的生成.

参 考 文 献

- [1] HUANG R, TAO Z, LIN Y, et al. Current situation of drainage pipe network in China and its detection technology: A brief review [J]. Polish Journal of Environmental Studies, 2023, 33(1): 19-29.
- [2] LU H, XU Z D, ISELEY T, et al. Pipeline Inspection and Health Monitoring Technology: The Key to Integrity Management [M]. Singapore: Springer Nature Singapore, 2023.
- [3] ISLAM M R, AZAM S, SHANMUGAM B, et al. A review on current technologies and future direction of water leakage detection in water distribution network [J]. IEEE Access, 2022, 10: 107177-107201.
- [4] SADEGHKHAH A, AHMED E, KREBS P. Towards a decentralized solution for sewer leakage detection - a review [J]. Water Science and Technology: a Journal of the International Association on Water Pollution Research, 2022, 86(5): 1034-1054.
- [5] YANG L, ZHUO W, QI L, et al. ST: Make self-training Work better for Semi-supervised semantic segmentation [C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans: IEEE, 2022: 4258-4267.
- [6] HAN B, YAO Q, YU X, et al. Co-teaching: Robust training of deep neural networks with extremely noisy labels [C]// Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal: ACM, 2018: 8536-8546.
- [7] LIU Y C, SHAO Z R, TENG Y Y, et al. NAM: normalization-based attention module [C]//NeurIPS 2021 Workshop on ImageNet: Past, Present and Future. California: NIPS, 2021:1-12.
- [8] MO Y, WU Y, YANG X, et al. Review the state-of-the-art technologies of semantic segmentation based on deep learning [J]. Neurocomputing, 2022, 493: 626-646.
- [9] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4):640-651.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 234-241.
- [11] 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: 6230-6239.
- [12] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (12) : 2481-2495.
- [13] CHEN L C, PAPANDEOU 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.
- [14] CHEN L C, PAPANDEOU 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.
- [15] CHEN L C, ZHU Y, PAPANDEOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//European Conference on Computer Vision. Cham: Springer, 2018: 833-851.

- [16] 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; International Association for Automation and Robotics in Construction (IAARC), 2019: 586-594.
- [17] WANG M, CHENG J C P. A unified convolutional neural network integrated with conditional random field for pipe defect segmentation [J]. Computer-Aided Civil and Infrastructure Engineering, 2020, 35(2): 162-177.
- [18] 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.
- [19] LI Y, WANG H, DANG L M, et al. A robust instance segmentation framework for underground sewer defect detection [J]. Measurement, 2022, 190: 110727.
- [20] VAN ENGELEN J E, HOOS H H. A survey on semi-supervised learning [J]. Machine Learning, 2020, 109(2): 373-440.
- [21] KE Z, QIU D, LI K, et al. Guided collaborative training for pixel-wise semi-supervised learning [C]// Computer Vision - ECCV 2020: 16th European Conference, Glasgow: ACM, 2020: 429-445.
- [22] MITTAL S, TATARCHENKO M, BROX T. Semi-supervised semantic segmentation with high- and low-level consistency [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43 (4) : 1369-1379.
- [23] CHEN X, YUAN Y, ZENG G, et al. Semi-supervised semantic segmentation with cross pseudo supervision [C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 2613-2622.
- [24] LIU Y, TIAN Y, CHEN Y, et al. Perturbed and strict mean teachers for semi-supervised semantic segmentation [C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans: IEEE, 2022: 4248-4257.
- [25] ALONSO I, SABATER A, FERSTL D, et al. Semi-supervised semantic segmentation with pixel-level contrastive learning from a class-wise memory bank [C]// 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 8199-8208.
- [26] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132-7141.
- [27] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 13708-13717.
- [28] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision. Cham: Springer, 2018: 3-19.
- [29] 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: 13451-13462.
- [30] WANG Y, ZHANG J, KAN M, et al. Learning pseudo labels for semi-and-weakly supervised semantic segmentation [J]. Pattern Recognition, 2022, 132: 108925.

(责编&校对 雷建云)