

• 四川大学建校 130 周年特刊约稿 •

DOI:10.12454/j.jsuese.202600021



本刊网刊

基于视频监控及视觉 AI 的城市内涝积水过程识别

秦友伟^{1,2}, 李心昊¹, 郑飞飞^{3*}, 朱 坚¹, 赵钟琳¹, 曹成龙³, 黄 源⁴

(1. 河海大学 水文水资源学院, 江苏 南京 210098; 2. 长江保护与绿色发展研究院, 江苏 南京 210098;
3. 浙江大学 建筑工程学院, 浙江 杭州 310009; 4. 河海大学 水利水电学院, 江苏 南京 210098)

摘 要:针对城市内涝监测中水尺/传感器布设不足、过程观测缺测导致难以获得连续水深演变数据的问题,提出一种基于视觉 AI 的积水等级识别与积水过程复原方法,实现积水深度的动态量化。构建以车辆为参照物的内涝积水图像数据集,涵盖 1 911 张公开平台图像及道路监控拍摄图像。积水等级识别模型在该数据集上表现出良好的收敛性与稳定性,平均检测精度(mAP)达到 0.673。过程复原采用“关键帧标定+线性插值”:选取具有明确物理意义的关键帧,依据车辆特征点与几何参照关系标定关键帧时刻水深,在相邻关键帧间线性插值得到连续水深-时间曲线;同时,以车辆参照的人工图像量测水深为参考真值,对插值曲线进行误差统计验证。利用上述方法复原了两个典型场景下的内涝积水过程:日间道路短历时积水过程呈典型“S”形曲线,积水初期扩张、快速加深以及高水位维持阶段的平均水位上涨速率分别为 3.05、8.69 与 1.05 cm/min;该场景下过程复原的均方根误差(RMSE)为 3.81 cm、平均绝对误差(MAE)为 3.30 cm、最大偏差为 6.91 cm。夜间停车场长历时积水过程呈“线性累积+加速增长”的非线性特征,线性累积与加速增长阶段的平均上涨速率分别为 0.56 与 1.74 cm/min;该场景下 RMSE 为 2.34 cm、MAE 为 1.99 cm、最大偏差为 4.61 cm。结果表明,该方法在不同光照条件与降雨模式下均能以厘米级误差复原积水过程,验证了积水识别与过程重建的精度与可靠性,可依托既有监控网络为缺乏直接内涝观测设备的地区提供非接触式、动态化的积水过程监测手段,为内涝风险研判、应急调度与雨洪模型率定提供低成本、可推广的数据支撑。

关键词:城市内涝;内涝积水;积水过程复原;视觉 AI;YOLOv8;防灾减灾

中图分类号:TV122

文献标志码:A

文章编号:2096-3246(2026)05-0349-11

城市内涝引发道路积水,不仅会降低道路的通行效率,还会对居民的生命财产安全构成威胁^[1-3]。不同淹没深度和淹没范围的内涝,对应的内涝风险也不尽相同^[4]。如何高效准确地获取城市内涝积水状态及发展态势,对于城市防灾减灾和应急管理至关重要。

2013 年,住房和城乡建设部发布了《城市排水防涝设施普查数据采集与管理技术导则(试行)》,各城市陆续开始布设内涝监测设备,基本实现了低洼积水点内涝的监测覆盖^[5-6]。然而,水位传感器的安装、率定及维护成本较高,限制了其在城市空间上的规模化布设^[7]。以 2020 年深圳市为例,据统计全市仅有 184 个内涝专用水位传感器,离内涝积水精细化监测的需求尚存在一

定差距。受热岛和雨岛效应的影响,城市暴雨常显示出明显的空间异质性^[8]。特别是短历时的强降雨,其空间异质性尤为显著^[9-10]。内涝监测点位与内涝实际发生位置在空间上往往存在不匹配的状况,导致内涝积水缺测,尤其是特大暴雨期间。

事实上,在社交媒体动态、视频监控等城市大数据中存储着大量城市内涝积水信息^[11-13]。内涝发生初期,通过社交媒体发布的内涝受灾信息与内涝事件几乎同步^[14-15]。此外,城市道路交通摄像头能够完整地记录城市道路的积水过程,且密度远大于内涝积水传感器,成为城市内涝积水的潜在智能监测终端^[16-17]。以北京市为例,2018 年交通摄像头数量约 115 万个,平均覆盖度约

收稿日期:2026-01-07 修回日期:2026-02-14 网络出版日期:2026-03-04

基金项目:国家重点研发计划项目(2022YFC3202602);中央高校基本科研业务费专项项目(226-2025-00245);国家自然科学基金项目(52261160379)

作者简介:秦友伟(1987—),男,副教授,博士。研究方向:城市内涝大数据挖掘。E-mail:youwei.qin@hhu.edu.cn

* 通信作者:郑飞飞,教授,E-mail:feifeizheng@zju.edu.cn

71个/km²,密度远大于内涝专用传感器^[18]。

随着视觉AI和深度学习的快速发展,利用视频和图像大数据开展内涝积水动态识别成为内涝研究的热点。当前研究主要围绕车辆、行人以及交通参照物来开展内涝积水的深度识别。以汽车为参照物的研究,常利用关键点检测算法识别汽车关键部位,如车轮、排气管、车把手等,并以此来估算积水等级^[11,19-20]。以行人为参照物的研究,常将目标检测与人体关键点检测相结合,通过识别涉水图像中人脸、手部、身体、脚踝等关键点位来推求内涝积水等级^[21-23]。以交通标志物为参照的研究,常引入具有一定标准尺寸的交通标志,如交通警示桶^[24]、停车让行标志^[25],来开展内涝积水等级识别。

现有研究多聚焦于利用互联网公开图像或社交媒体数据开展城市内涝积水识别,主要通过不同的深度学习模型或图像增强算法提升检测精度。然而,此类研究的数据来源分散、获取时序不连续,难以在固定空间位置上持续追踪积水变化过程,因而难以支撑内涝积水演变特征的动态反演与定量分析。本文提出了一种基于视觉AI的城市内涝积水特征动态识别方法,依托高清摄像头实时采集的视频监控数据,对同一监测点的内涝积水过程进行时序化识别与反演,重建积水深度随时间变化的动态演变过程,从而刻画内涝发生、发展与维持阶段的关键特征,显著提升了城市内涝监测的时空连续性与工程应用价值。

1 实验算法

1.1 算法流程

本文算法流程按“模型构建—训练策略—性能评估”展开。首先,基于YOLOv8的通用网络架构构建积水等级识别模型;其次,按既定参数迭代训练,采用主干网络冻结等策略提高收敛稳定性;最后,以准确率、平均精度等指标进行性能评估。

1.2 算法介绍

YOLO(you only look once)算法^[26]作为一阶段检测算法的代表,自2015年提出以来经过不断优化迭代,在检测速度和精度上均取得了持续的突破。YOLOv8算法以其实时性强、推理效率高的特点,被广泛应用于各类目标检测场景中。YOLOv8网络由骨干网络(backbone)、特征融合层(neck)、检测头(head)3部分组成,YOLOv8网络结构如图1所示。图1中:Input表示输入图像;Conv表示卷积模块,用于提取与变换特征;C2f表示特征提取/融合模块,用于增强特征表达并改善梯度流;SPPF表示快速空间金字塔池化模块,用于引入多尺度上下文信息;Upsample表示上采样操作;Concat表示特征拼接;检测头采用解耦结构,其中Conv2d表示二维卷积层;BboxLoss表示边界框回归损失;ClsLoss表示分类损失,用于指导模型参数优化。

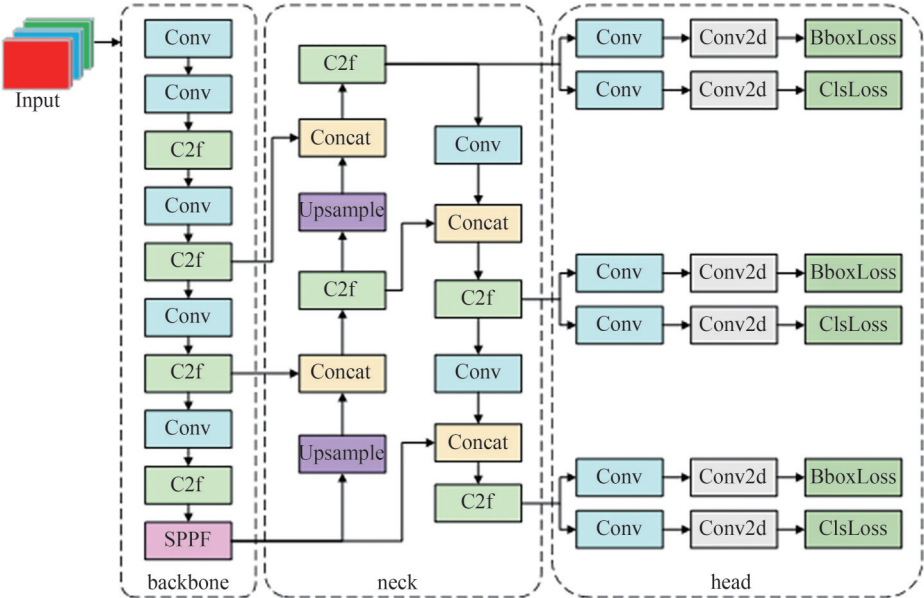


图1 YOLOv8网络结构

Fig. 1 Network structure of YOLOv8

1.3 算法训练

实验在 Windows 10 操作系统中开展 YOLOv8 算法训练,深度学习框架采用 PyTorch 1.10 版本,训练硬件采用 NVIDIA RTX 3050 GPU 显卡。

YOLOv8 模型的主要训练参数设置如下:

- 1) 批次大小(batch size)为 16;
- 2) 输入图像尺寸(input size)为 640 像素×640 像素;
- 3) 优化器(optimizer)为 AdamW;

4)初始学习率(initial learning rate)为0.001,采用余弦退火策略进行动态学习率调整;

5)最大允许迭代轮数(epochs)为300;

6)backbone冻结策略:前10个epoch阶段冻结主干网络参数,仅训练neck和head部分,以加快早期收敛速度并稳定特征提取;

7)模型检查点(checkpoint)保存频率:每5个epochs保存一次,同时在验证集上评估模型性能。

1.4 算法评估

为有效评估YOLOv8算法在城市内涝积水特征动态识别任务上的表现,本文采用5项关键评估指标,即准确率(Precision,记为 P)、召回率(Recall,记为 R)、F1分数(F1 Score,记为 F_1)、精确率-召回率(P - R)曲线及平均精度均值(mAP, mean Average Precision,记为 m_{AP})。

1)准确率,衡量模型对阳性类别预测的准确性,定义为真阳性预测与总预测阳性的比率。计算公式为:

$$P = \frac{T_p}{T_p + F_p} \quad (1)$$

式中, T_p 为真阳性数量, F_p 为假阳性数量。

2)召回率,评估模型检测数据集中所有相关样本的能力,定义为真阳性与真阳性和假阴性总和的比率。计算公式为:

$$R = \frac{T_p}{T_p + F_N} \quad (2)$$

式中, F_N 为假阴性数量。

3) F_1 分数,精确度和召回率的调和平均值,综合反应模型整体检测性能。计算公式为:

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (3)$$

4)精确率-召回率(P - R)曲线,通过绘制不同置信度阈值下的 P 和 R 变化,反应模型在不同检测敏感度下的性能表现。曲线下的面积越大,说明模型整体检测效果越好。

5)平均精度,反应模型在所有类别上的综合检测精度。对单类别而言,平均精度(AP,记为 A_p)的定义为精确率-召回率曲线下的面积:

$$A_p = \int_0^1 P(R) dR \quad (4)$$

式中, $P(R)$ 为准确率随召回率变化的关系。

多分类情况下,取各类别AP的平均值得到mAP:

$$m_{AP} = \frac{\sum_{i=1}^N A_{P_i}}{N} \quad (5)$$

式中, A_{P_i} 为第 i 个分类的平均精度, N 为分类总数。

2 实验数据

2.1 样本数据集

本文构建了一个以汽车为参照物的内涝积水样本数据集,共计1911张,来源包括两部分:互联网公开视频/图片平台收集的典型内涝场景图像1780张,是以城市内涝、车辆淹水为关键词搜索得到的典型城市内涝事件相关公开图片;城市道路监控摄像头视频的抽帧图像131张,主要来自11场强降雨过程的固定监控视角样本,包含了强降雨期间16场内涝积水过程。两类样本均按“画面清晰、车辆涉水特征可辨识、可标注等级”的标准筛选,剔除重复与模糊帧。总体上,白天/弱光场景样本约占90%,夜间路灯/车灯照明条件下样本约占10%。

将尺寸统一为640像素×640像素、色彩归一化,使训练样本数据集的所有原始图像均完成标准化预处理,以适配模型的数据输入要求。采用YOLOv8默认在线增强设置(Mosaic马赛克拼接、随机仿射、HSV扰动等)。

2.2 样本数据标注

使用LabelImg工具对训练样本数据集中的车辆参照物进行手动标注,通过绘制矩形边界框框定车辆参照物及其积水级别,内涝积水等级标签制作如图2所示。对1911张内涝积水图片进行标注后,共生成了5260个车辆样本实例。每个车辆样本实例包含了标注框左上角与右下角顶点像素坐标、中心点坐标、宽度、高度信息,以及对应的积水级别。



图2 内涝积水等级标签制作

Fig. 2 Annotation of urban inundation levels

依据积水程度由轻到重,结合车辆行车安全,将积水等级划分为5级,每个级别对应不同的积水等级和车辆通行风险:

第0级(Level 0),无明显积水,车辆可正常行驶;

第1级(Level 1),积水低于轮胎一半,车辆可正常

行驶,但存在轻微进水风险;

第 2 级(Level 2),积水超过轮胎一半,但未完全覆盖保险杠,存在较高进水风险;

第 3 级(Level 3),积水超过保险杠位置,但未淹没发动机前盖,车辆通行困难;

第 4 级(Level 4),积水超过发动机前盖,但未完全淹没车辆,车辆无法通行。

2.3 样本标注评估

图 3 为积水等级标注数据分布分析。图 3(a)中,内涝积水等级标注样本在 5 个不同积水等级均超过 800 个,其中:Level 2 样本数量最多,占比约 35%;其余等级分别在 15%~20%之间,整体分布相对稳定。图 3(c)中, X 、 Y 为车辆目标标注框中心点在图像中的归一化坐标。由图 3(b)、(c)可以看出,目标中心主要集中于图像中心区域,同时涵盖了从小尺度到大尺度的多类目标,这与图像拍摄的行为习惯一致。图 3(d)中,大多数目标框的尺寸集中在 0.1~0.4 之间,代表以小中型目标为主,同时也包含部分宽度和高度较大的目标,体现出一定的尺寸延展性,这种尺度分布符合城市内涝视频监控中车辆的特点。

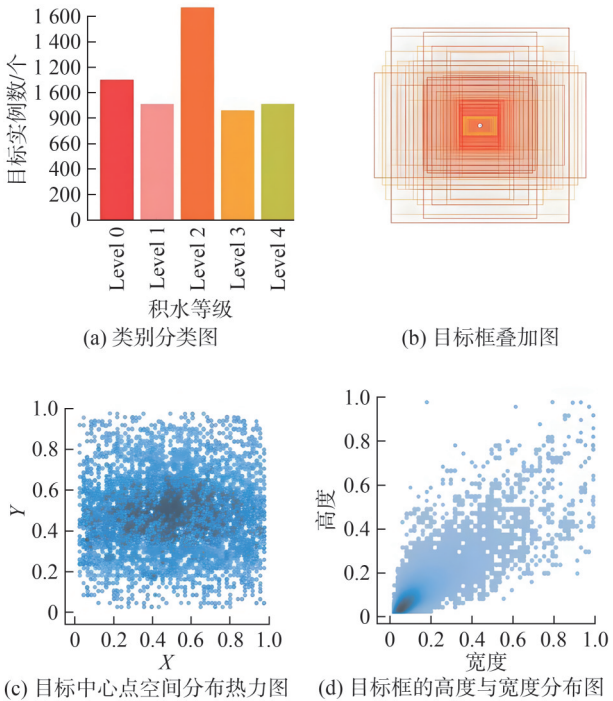


图 3 积水等级标注数据分布分析

Fig. 3 Distribution analysis of annotation water depth levels

2.4 视频监控数据

本文选取两段监控视频作为测试数据,分别覆盖日间道路暴雨内涝与夜间低照度长历时积水两类典型场景(视频链接为 <https://pan.baidu.com/s/1gFd2g371tPlWdv-TjLM0Gg?pwd=szzv>):

1) 日间道路监控视频(2024-04-27,广东省某市):由道路监控高清摄像头拍摄,时间跨度为 37 min(采用倍速存储,获取的视频文件实际时长为 1 min55 s),完整地记录了强降雨条件下积水从初现、加深至形成的全过程。视频中一辆灰色面包车因积水过深熄火,成为整个积水过程中的固定参照物,为后续基于车辆的积水等级识别提供了稳定目标。

2) 夜间停车场监控视频:由停车场监控摄像头拍摄,时间跨度为 66 min(采用倍速存储,获取的视频文件实际时长为 55 s),记录了长历时降雨条件下积水由浅到深的演变过程。该场景受路灯/车灯照明、频闪与水面反光影响,积水边缘特征更模糊。画面中多辆静止轿车构成稳定参照系。

3 实验结果

3.1 积水等级识别模型

模型训练设置最大迭代轮数为 300,模型在第 274 个训练轮次到达收敛,自动终止。积水等级识别模型性能评估如图 4 所示。图 4 中, $@0.5$ 表示预测框与真实框的交集面积与并集面积之比,即预测框与真实框的 IOU。由图 4(a)、(b)可见,模型在训练初期(第 1 个训练轮次)损失较高,分别为 3.69(训练)和 4.12(验证);随着迭代进行,损失持续下降,第 274 个训练轮次稳定在 0.80(训练)和 1.08(验证),表现出良好的收敛性和训练稳定性。图 4(c)表明,积水等级识别算法对 Level 4 及 Level 2 识别的表现较好,曲线整体位于图表上方。不同积水等级在 IOU 大于 0.5 时检测性能存在显著差异,在 Level 4 严重积水检测精度最高达 0.774,而 Level 1 轻微积水仅为 0.577。图 4(d)显示,不同积水等级在置信度 0.5 附近的检测性能存在差异,其中,Level 4(严重积水)的 F1 分数最高,而对 Level 1 的识别效果较差。这主要是因为 Level 1 样本的视觉特征与 Level 0 相似,易造成模型混淆,从而导致了较低的检测性能,说明模型对轻微积水的识别能力相对较弱,但在中低置信度下仍具有较好的稳定性和泛化性。整体而言,模型对道路积水等级的识别具备较好的检测效果。

图 5 为积水等级识别模型在道路积水等级识别中的部分代表性结果。由图 5(a)可知,积水等级识别模型能够较为准确地识别图像中央横向区域内的车辆,而对边缘位置车辆检测中仍存在漏检现象,在处理小尺度或远距离目标时的识别能力有限。从图 5(b)可以进一步确认,积水等级识别模型能准确识别图像中心区域的道路积水等级,而对边缘区域的车辆目标识别能力有限。

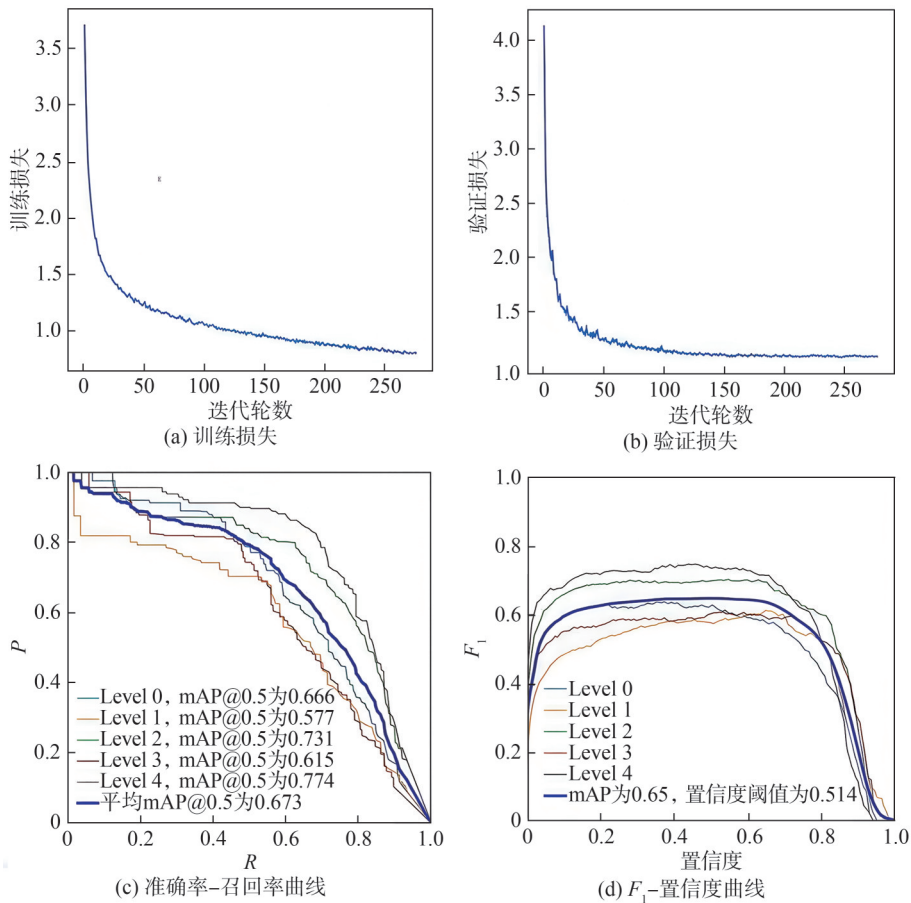


图 4 积水等级识别模型性能评估

Fig. 4 Performance evaluation of water level recognition model



(a) 模型漏检示意图



(b) 模型错检示意图

图 5 积水等级识别模型检测代表性结果示意图

Fig. 5 Schematic diagram of detection results for water level recognition model

积水等级识别模型总体 mAP 为 0.673, 整体性能处于中等水平。其性能差异与图 3 的样本分布特征密切相关。一方面, 图 3(a) 的 Level 2 样本数量最多, 为模型提供了充分的训练数据, 因此训练表现最为稳定; Level 4 样本虽数量不及 Level 2, 但因其视觉特征最为显著, 模型仍能较好识别。而 Level 1 样本虽占一定比例, 但其视觉特征与 Level 0 相似, 易造成混淆, 导致检测性能较低。另一方面, 图 3(b)~(d) 表明训练数据中车辆样本主要分布在图片中央位置, 边缘区域样本数量相对不足, 是造成边缘区域小目标的错检和漏检的一个主要原因。考虑到交通监控摄像头多采用中近景拍摄, 且车辆主要分布于监控画面的中心区域, 因此该模型在基于监控视频的实际应用中受边缘样本偏差影响较小, 具有良好的适用性。

为评估本文提出的积水等级识别模型性能水平, 进一步与近期基于 YOLOv8 的相关研究进行对比。Panja 等^[27]在雨雾等恶劣天气条件下的车辆检测任务中报告 YOLOv8 的基准 mAP 约为 0.685; 本文在暴雨监控场景下叠加积水等级判别取得的 mAP 为 0.673, 整体精度与该基准水平接近。Qiu 等^[28]针对公交车积水深度分级任务获得最佳 mAP 为 0.662; 相比之下, 本

文在多车型混合场景下 mAP 达到 0.673,处于同类研究的可比水平。

3.2 实景应用验证

本文的积水过程复原采用“关键帧+线性插值”的技术路线:首先,在监控视频中选取具有明确物理意义的关键帧,基于车辆特征点与参照物几何关系对关键时刻的积水深度进行标定;随后,在相邻关键帧之间采用线性插值,得到连续的积水深度-时间演变曲线。由于研究区域缺乏同位置水尺/传感器水位观测数据,以车辆参照的图像量测水深为“参考真值”,并在关键帧间隔内选取若干时刻进行人工量测复核;以插值曲线在对应时刻的取值为“模拟值”,两者差异用于表征过程重建误差与趋势非一致性。

本文选取两类具有代表性的真实监控场景开展验证:其一为 2024-04-27 日间道路监控场景(短历时强降雨触发的快速积水过程);其二为夜间停车场监控场景(长历时降雨背景下的缓慢累积过程)。

3.2.1 日间道路监控场景

为复原 2024-04-27 日间内涝积水时序演变过程,本研究选取了 4 个关键帧,日间场景内涝积水深度识别结果(积水深度-时间曲线)如图 6 所示,关键帧①~

④分别对应第一次达到 Level 1~Level 4 积水等级的时刻,以及两个具有明显物理意义的人工参考点(图 7),涵盖从“首次漫过路缘石(图 7(a)关键帧⑤)”到“积水达到最深(图 7(b)关键帧⑥)”的典型阶段。参照《城市道路工程设计规范》(CJJ 37—2012),此次路缘石外露高度选取 10.0 cm。面包车不同积水等级和典型特征点高度由实际测量获得,Level 2~Level 4 以及积水最高点尾灯上沿对应深度分别对应 28.5、58.0、98.0、113.0 cm。

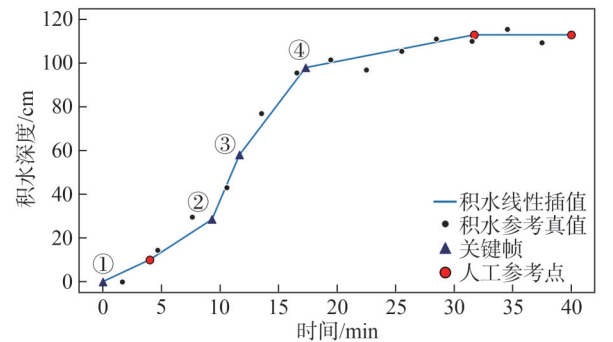


图 6 日间场景内涝积水深度识别结果(积水深度-时间曲线)

Fig. 6 Results of waterlogging depth for day scenarios (water depth-time curves)

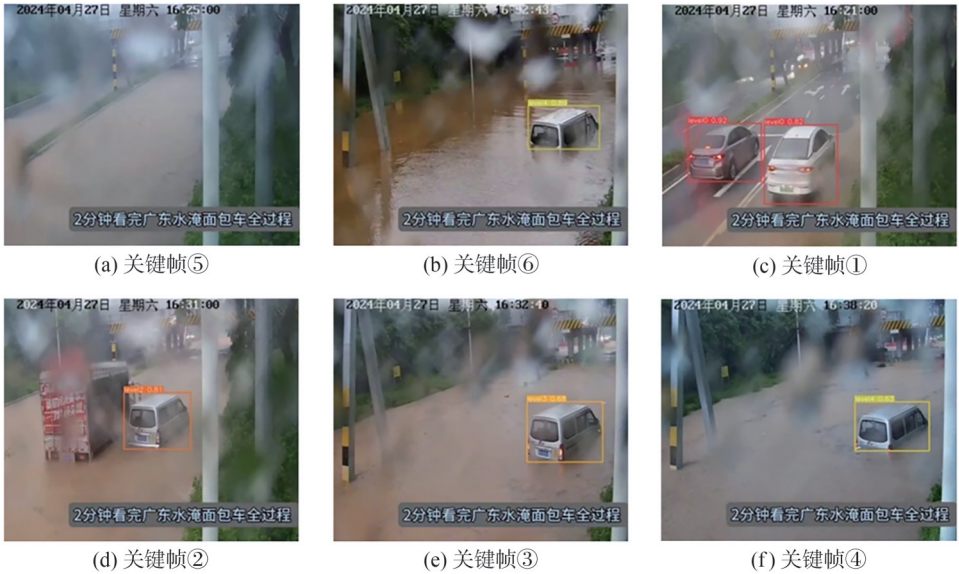


图 7 日间场景积水深度识别关键帧及人工参考点

Fig. 7 Key frames and reference frames of water depth recognition for day scenarios

图 6 复原了该场景下积水深度随时间的动态演变规律,展示了城市内涝积水过程从积水形成、积水扩张到积水达到巅峰的全过程。由图 6 可知,该场积水深度变化表现出显著的阶段性,整体近似呈“S”形曲线。整个内涝积水过程可划分为 3 个典型阶段。1)积水初期扩张阶段(积水开始至第 9 min20 s):随着降雨开始,地表径流迅速汇集并形成路面积水。此刻作为图 6 积水过程的 0 点时刻(视频时间为 16:21:00),如图 7

(c)关键帧①所示。随后,地表径流开始横向扩张,约 4 min 后(视频时间为 16:25:00)漫过 10.0 cm 的路缘石,迅速扩散至整个路面(图 7(a)关键帧⑤)。第 9 min20 s (视频时间为 16:30:20),积水首次达 Level 2 等级,对应深度为 28.5 cm,如图 7(d)关键帧②所示。此阶段内,积水深度在 9 min20 s 内上涨了 28.5 cm,平均上涨速度约 3.05 cm/min。2)积水快速加深阶段(第 9 min20 s—第 17 min20 s):积水进入快速上涨期,积水深度在随后的

2 min20 s内从Level 2上涨至Level 3,对应视频中16:32:40、积水深度为58.0 cm,如图7(e)关键帧③所示。随后在第17 min20 s第一次达到积水等级Level 4,此时对应视频时间为16:38:20、积水深度为98.0 cm,如图7(f)关键帧④所示。此阶段内,积水深度在8 min内从28.5 cm上涨至98.0 cm,平均上涨速度约8.69 cm/min。

3)高水位维持阶段(第17 min20 s以后):随后,积水深度继续缓慢攀升,最终在第31 min43 s(视频时间为16:52:43)达到积水深度最高点,即面包车尾灯上沿高度113.0 cm处(图7(b)关键帧⑥)。此阶段内,积水上涨速度锐减至1.05 cm/min。在随后几分钟内,水位几乎未见明显回落,形成持续性的高水位状态。这一特征反映出道路排水系统已完全饱和,雨后汇水无法及时排出,呈现“高位滞水—缓慢退水”的典型特征。

图6中,关键帧及人工参考点水深点线性插值得到的过程曲线(实线)与人工量测参考真值点(每隔3 min采点)整体趋势一致,能够刻画积水由缓慢增长到快速加深并进入高水位维持阶段的演变特征。以参考真值为基准,插值曲线在对应时刻的取值作为模拟值,计算误差:RMSE为3.81 cm、MAE为3.30 cm、最大偏差为6.91 cm;以峰值水深113.0 cm计,3个指标相对误差分别约为3.37%、2.92%与6.12%。

3.2.2 夜间停车场监控场景

为复原2025-07-21夜间长历时降雨内涝积水过程,本文选取3个关键帧作为积水深度演变节点,分别对应第一次达到积水等级Level 1~Level 3的时刻,夜间场景内涝积水深度识别结果(积水深度-时间曲线)如图8所示。小轿车在积水等级Level 2、Level 3对应的积水深度分别为28.5、39.0 cm。夜间场景积水深度识别关键帧如图9所示。

由图8可见,该场积水过程亦呈现分段演变特征,可划分为两个阶段。1)积水缓慢累积阶段(积水

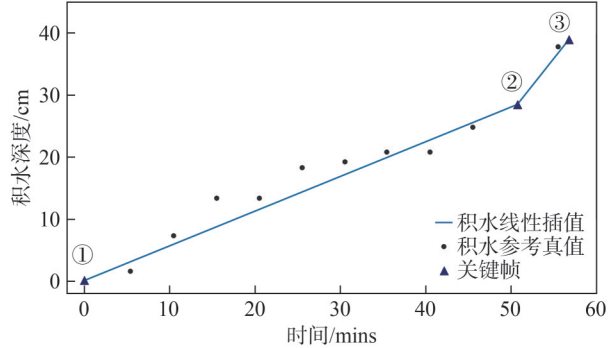


图8 夜间场景内涝积水深度识别结果(积水深度-时间曲线)

Fig. 8 Results of waterlogging depth for night scenarios (water depth-time curves)

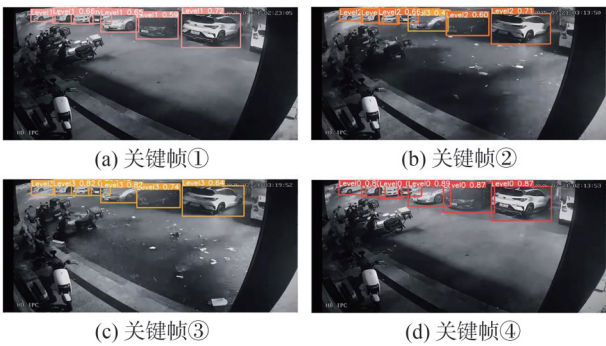


图9 夜间场景积水深度识别关键帧

Fig. 9 Key frames of water depth recognition for night scenarios

开始至第50 min45 s):将首次可辨识积水(Level 1)时刻作为图8积水过程的0点,视频时间为02:23:05,如图9(a)关键帧①所示。在第50 min45 s(视频时间为03:13:50)积水首次达到Level 2等级,对应深度28.5 cm,如图9(b)关键帧②所示。此阶段内,积水深度在50 min45 s内上涨了28.5 cm,平均上涨速度约为0.56 cm/min。

2)积水加速增长阶段(第50 min45 s—第56 min47 s):积水深度在随后的6 min2 s内由Level 2上涨至Level 3,对应的视频时间为03:19:52。此阶段水深从28.5 cm快速攀升至39.0 cm,平均上涨速率提高至1.74 cm/min。

在误差验证方面,该场景下视频中每间隔约5 min选取特征时刻开展人工图像量测复核(图8)。经统计计算得到RMSE为2.34 cm、MAE为1.99 cm、最大偏差为4.61 cm。以该场峰值水深39.0 cm来计算,3个指标相对误差分别为6.00%、5.10%与11.82%。上述结果表明,关键帧标定结合线性插值的方法在夜间低照度场景下仍能以较低误差复原积水深度的时序演变过程。

与日间道路监控场景相比,夜间低照度场景在积水等级判别与过程复原上更具挑战。其一,在积水初始形成阶段,浅水在弱光条件下对比度低,且易受路灯/车灯反射干扰,导致水体边界与地面纹理的区分度显著下降,Level 0与Level 1的临界状态即便人工目视亦难以稳定判定,如图9(a)关键帧①与图9(d)关键帧④所示。其二,夜间案例属于长历时缓变过程,单位时间水深增量较小,图像量测误差与水面微波动在相对尺度上更易被放大,从而使相对误差指标较日间快速上涨过程更为敏感(本文夜间RMSE/MAE约为6.00%/5.10%,均高于日间的3.37%/2.92%)。总体而言,夜间场景在浅水分级与缓变过程的敏感性方面更具挑战,但“关键帧标定+线性插值”的路线仍能在低照度条件下为无资料地区提供可用的积水过程。

4 结论与展望

本文针对内涝过程观测缺测问题,提出了一种基于道路监控视频与视觉AI的积水等级识别与过程复原方法。构建以车辆为参照物的积水数据集(1 911张),基于YOLOv8训练得到识别模型,总体mAP为0.673,可实现多等级积水自动识别。实景应用中采用“关键帧标定+线性插值”复原了两场降雨积水过程:日间暴雨场景积水曲线呈典型“S”形特征,积水初期扩张、积水快速加深和高水位维持3个阶段平均上涨速率分别为3.05、8.69和1.05 cm/min;夜间低照度长历时场景呈“缓慢累积—加速增长”特征,关键阶段速率分别为0.56和1.74 cm/min。重建的积水过程误差处于可控范围,表明该方法可依托既有监控网络为缺乏水尺/传感器的区域提供低成本、非接触式的积水过程数据支撑,为内涝风险研判、应急调度与城市雨洪模型参数率定提供数据支撑。

从鲁棒性与适用性看,夜间弱光与反光背景会显著削弱浅水阶段可辨识性,Level 0、Level 1的视觉特征尤其接近,即便人工目视亦难稳定判别;同时,长历时缓变过程单位时间水深增量较小,使量测扰动在相对尺度上更易放大,导致夜间的相对误差较日间快速上涨,场景更敏感。

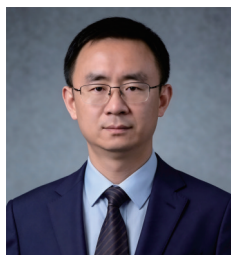
后续研究将从“数据扩充—反演增强—机理耦合”3方面深化:一是补充夜间与雨雾等恶劣天气样本,形成日间道路暴雨与夜间低照度两类典型应用场景;二是引入路缘石、护栏、交通标线/标志等多参照物联合标定与插值,提升积水过程复原的精度与可靠性;三是尝试与雨洪模型耦合用于率定与精细化模拟,逐步实现监测—模拟—预警的闭环应用。

参考文献:

- [1] Liu Jiahong, Mei Chao, Liu Hongwei, et al. Key scientific and technological issues of joint prevention and control of river flood and urban waterlogging disaster chain in megacities[J]. *Advances in Water Science*, 2023, 34(2): 172–181. [刘家宏, 梅超, 刘宏伟, 等. 特大城市外洪内涝灾害链联防联控关键科学技术问题[J]. *水科学进展*, 2023, 34(2): 172–181.]
- [2] Zhang Jianyun, Wang Yintang, He Ruimin, et al. Discussion on the urban flood and waterlogging and causes analysis in China[J]. *Advances in Water Science*, 2016, 27(4): 485–491. [张建云, 王银堂, 贺瑞敏, 等. 中国城市洪涝问题及成因分析[J]. *水科学进展*, 2016, 27(4): 485–491.]
- [3] Wang Hao. Thoughts on urban flood prevention and control in China under the new situation[J]. *China Flood & Drought Management*, 2019, 29(8): 1–2. [王浩. 新形势下我国城市洪涝防治的几点认识[J]. *中国防汛抗旱*, 2019, 29(8): 1–2.]
- [4] Gao Kai, Yang Zhiyong, Gao Xichao, et al. A review of the evaluation methods of urban flood loss[J]. *Water Resources and Hydropower Engineering*, 2021, 52(4): 57–68. [高凯, 杨志勇, 高希超, 等. 城市洪涝损失评估方法综述[J]. *水利水电技术*, 2021, 52(4): 57–68.]
- [5] Huang Jing, Kang Jinle, Wang Huimin, et al. A novel approach to measuring urban waterlogging depth from images based on mask region-based convolutional neural network[J]. *Sustainability*, 2020, 12(5): 2149.
- [6] Luan Qinghua, Qin Zhiyu, Wang Dong, et al. Review on monitoring technology of urban road waterlogging after rainstorm and its application[J]. *Water Resources Protection*, 2022, 38(1): 106–116. [栾清华, 秦志宇, 王东, 等. 城市暴雨道路积水监测技术及其应用进展[J]. *水资源保护*, 2022, 38(1): 106–116.]
- [7] Paul J D, Buytaert W, Sah N. A technical evaluation of lidar-based measurement of river water levels[J]. *Water Resources Research*, 2020, 56(4): e2019WR026810.
- [8] Muller C L, Chapman L, Johnston S, et al. Crowdsourcing for climate and atmospheric sciences: Current status and future potential[J]. *International Journal of Climatology*, 2015, 35(11): 3185–3203.
- [9] Wasko C, Sharma A. Steeper temporal distribution of rain intensity at higher temperatures within Australian storms[J]. *Nature Geoscience*, 2015, 8(7): 527–529.
- [10] Zheng Feifei, Westra S, Leonard M. Opposing local precipitation extremes[J]. *Nature Climate Change*, 2015, 5(5): 389–390.
- [11] Park S, Baek F, Sohn J, et al. Computer vision-based estimation of flood depth in flooded-vehicle images[J]. *Journal of Computing in Civil Engineering*, 2021, 35(2): 04020072.
- [12] Jiang Yunzhong, Ye Yuntao, Zhao Hongli, et al. Research status and prospects on water conservancy big data[J]. *Journal of Hydroelectric Engineering*, 2020, 39(10): 1–32. [蒋云钟, 冶运涛, 赵红莉, 等. 水利大数据研究现状与展望[J]. *水力发电学报*, 2020, 39(10): 1–32.]
- [13] Song Huimeng, Pan Zhian. Summary of urban waterlogging topic mining based on social media data[J]. *Technology Innovation and Application*, 2024, 14(32): 79–81. [宋会萌, 潘志安. 基于社交媒体数据的城市内涝主题挖掘综述[J]. *科技创新与应用*, 2024, 14(32): 79–81.]
- [14] Wang Huimin, Huang Jing, Liu Gaofeng, et al. Research paradigm of big data driven urban flood disaster risk perception, early warning and decision[J]. *Frontiers of Science and Technology of Engineering Management*, 2022, 41(1): 35–41. [王慧敏, 黄晶, 刘高峰, 等. 大数据驱动的城市洪涝灾害风险感知与预警决策研究范式[J]. *工程管理科技前沿*, 2022, 41(1): 35–41.]
- [15] Huang Huabing, Wang Xianwei, Liu Lin. A review on urban pluvial floods: Characteristics, mechanisms, data, and re-

- search methods[J]. *Progress in Geography*, 2021, 40(6): 1048–1059. [黄华兵, 王先伟, 柳林. 城市暴雨内涝综述: 特征、机理、数据与方法[J]. *地理科学进展*, 2021, 40(6): 1048–1059.]
- [16] Hao Xin, Lyu Heng, Wang Ze, et al. Estimating the spatial-temporal distribution of urban street ponding levels from surveillance videos based on computer vision[J]. *Water Resources Management*, 2022, 36(6): 1799–1812.
- [17] Xue Fengchang, Song Xiaoyi, Tang Buxing, et al. Urban waterlogging monitoring based on urban video image[J]. *Science of Surveying and Mapping*, 2018, 43(8): 50–55. [薛丰昌, 宋肖依, 唐步兴, 等. 视频监控的城市内涝监测预警[J]. *测绘科学*, 2018, 43(8): 50–55.]
- [18] Jiang Jingchao, Liu Junzhi, Qin Chengzhi, et al. Extraction of urban waterlogging depth from video images using transfer learning[J]. *Water*, 2018, 10(10): 1485.
- [19] Wang Youyi. Detection and risk assessment of submerged vehicles in urban waterlogging scenarios based on deep[D]. Ganzhou: Jiangxi University of Science and Technology, 2022. [王有毅. 基于深度学习的城市内涝场景受淹车辆检测及危险度评估[D]. 赣州: 江西理工大学, 2022.]
- [20] Zhong Pengcheng, Liu Yueyi, Zheng Hang, et al. Detection of urban flood inundation from traffic images using deep learning methods[J]. *Water Resources Management*, 2024, 38(1): 287–301.
- [21] Chaudhary P, D'Aronco S, Leitão J P, et al. Water level prediction from social media images with a multi-task ranking approach[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 167: 252–262.
- [22] Li Jingru, Cai Ruying, Tan Yi, et al. Automatic detection of actual water depth of urban floods from social media images[J]. *Measurement*, 2023, 216: 112891.
- [23] Quan K C, Nguyen V T, Nguyen T C, et al. Flood level prediction via human pose estimation from social media images[C]// *Proceedings of the 2020 International Conference on Multimedia Retrieval*. Dublin Ireland: ACM, 2020: 479–485.
- [24] Jiang Shijie, Babovic V, Zheng Yi, et al. Advancing opportunistic sensing in hydrology: A novel approach to measuring rainfall with ordinary surveillance cameras[J]. *Water Resources Research*, 2019, 55(4): 3004–3027.
- [25] Alizadeh Kharazi B, Behzadan A H. Flood depth mapping in street photos with image processing and deep neural networks[J]. *Computers, Environment and Urban Systems*, 2021, 88: 101628.
- [26] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]// *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE, 2016: 779–788.
- [27] Panja E, Hendry H, Dewi C. YOLOv8 analysis for vehicle classification under various image conditions[J]. *Scientific Journal of Informatics*, 2023, 11(1): 127–138.
- [28] Qiu Yanbin, Zhou Xudong, Wan Jiaquan, et al. Automated urban flood level detection based on flooded bus dataset using YOLOv8[J]. *Natural Hazards and Earth System Sciences*, 2025, 25(9): 3525–3544.

【校友简介】



郑飞飞, 浙江大学求是特聘教授、长聘教授、博士生导师。郑飞飞为教育部高层次人才、国家基金委优秀青年科学基金获得者、国家海外高层次人才青年项目获得者、浙江省杰出青年科学基金获得者。长期从事城市水文和城市洪涝领域的基础理论和关键技术研究。主持国家自然科学基金6项, 包括青年、面上(2)、国际(地区)合作和原创探索项目等, 主持国家重点研发课题1项; 共发表学术论文80余篇, 其中第一/通讯作者SCI论文70余篇, 包括 *Water Resources Research* 19篇和 *Water Research* 6篇; 出版排第一的学术专著2部, 授权排第一国家发明专利22项(其中美国和日本发明专利4项)且转让2项; 主编国家标准《城镇排水管网数字化通用技术要求》, 参编国家标准1部(排第2), 主编省行业标准3部; 担任SCI期刊 *Journal of Hydroinformatics* 领域主编和两个中文核心期刊的青年编委, 获澳大利亚“GN Alexander Medal”奖(排第1)、2025年大禹水利科学技术进步奖一等奖(排第1)和2022年大禹水利科学技术进步奖二等奖(排第1)。

Recognition of Urban Flooding Process Based on Surveillance Video and Visual AI

QIN Youwei^{1,2}, LI Xinhao¹, ZHENG Feifei^{3*}, ZHU Jian¹, ZHAO Zhonglin¹, CAO Chenglong³, HUANG Yuan⁴

(1. College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China;

2. Yangtze Institute for Conservation and Development, Nanjing 210098, China;

3. College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310009, China;

4. College of Water Conservancy & Hydropower Engineering, Hohai University, Nanjing 210098, China)

Abstract:

Objective Continuous, time-resolved water-depth hydrographs are often unavailable during urban pluvial flooding because currently installed

water-depth gauges/sensors are sparse and event observations are incomplete. Therefore, this study uses an AI-based approach to (i) recognize multi-level waterlogging severity from widely deployed surveillance imagery and (ii) reconstruct the continuous temporal evolution of water depth at fixed locations, enabling the derivation of flood-process information for rapid situational assessment and the calibration/validation of urban rainfall-runoff models in data-scarce areas.

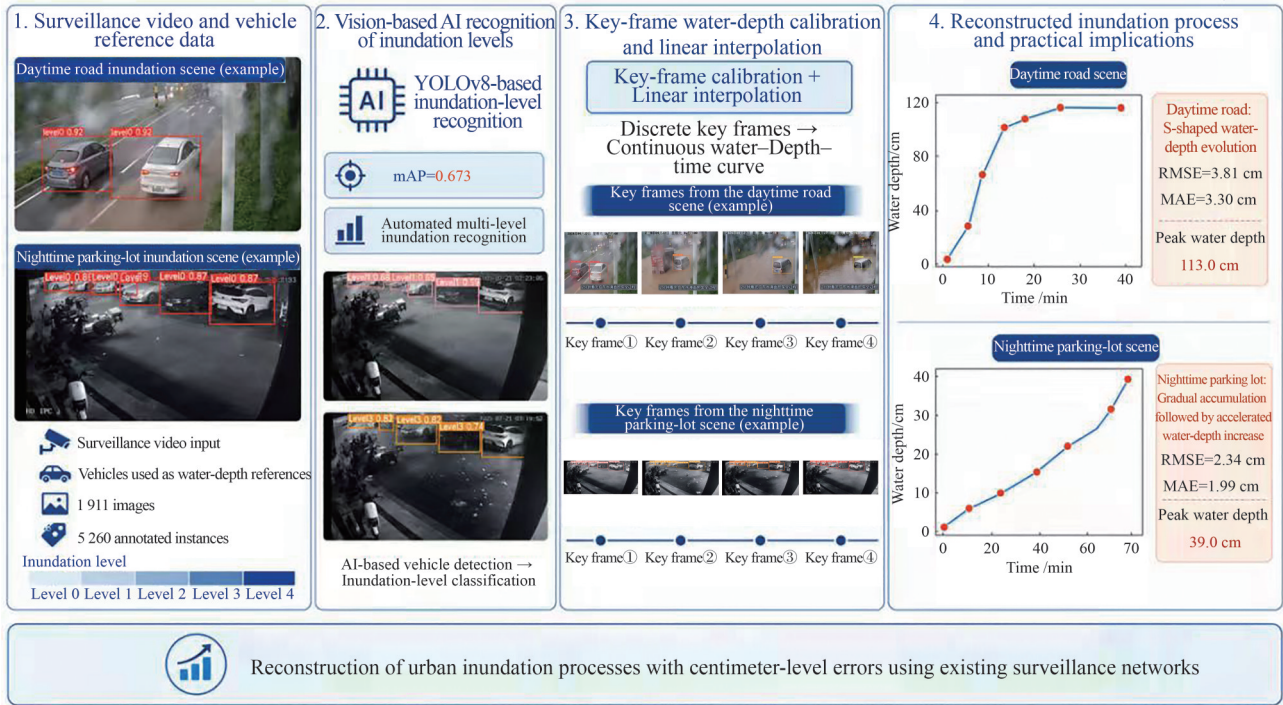
Methods A two-stage workflow was established, consisting of flood-level recognition and flood-process reconstruction. First, a vehicle-referenced waterlogging dataset was compiled using 1 911 images collected from public platforms and fixed surveillance cameras. Vehicles were annotated with bounding boxes, resulting in 5 260 labeled instances. Waterlogging severity was discretized into five classes (Level 0 to Level 4) based on visually interpretable vehicle-water interactions. Then, a YOLOv8 one-stage detector (backbone-neck-decoupled head) was trained using PyTorch. Online data augmentation techniques, including Mosaic, random affine transformation, and HSV perturbation, together with a short backbone-freezing schedule, were applied to stabilize the early training process. Model performance was evaluated using Precision, Recall, F1-score, P - R curves, and mean Average Precision (mAP). Class-wise performance was interpreted with respect to the sample distribution across severity levels and the spatial distribution of labeled instances within the image plane. Water-depth hydrographs were reconstructed using a “key-frame calibration + linear interpolation” approach. Key frames with clear physical significance were selected from surveillance videos. Water depth at each key frame was calibrated using vehicle feature points and reference geometry, and a continuous depth-time curve was generated through linear interpolation between adjacent key frames. Because installed water-depth gauge/sensor data were unavailable, manual image-based depth measurements obtained using the same vehicle references were treated as the reference truth. Additional timestamps were sampled within the key-frame intervals for manual measurements, and the discrepancies between the manual measurements and interpolated values were quantified using RMSE, MAE, and maximum absolute deviation. Two representative scenarios were analyzed, including a daytime road scene with short-duration intense rainfall and a nighttime parking-lot scene with long-duration rainfall.

Results and Discussions The model converged stably and achieved an overall mAP of 0.673 on the validation set. Performance varied across the severity levels. Severe inundation was detected more reliably because stronger visual cues were present, whereas the distinction between Level 0 and Level 1 remained challenging because shallow-water features were weak and visually ambiguous. Spatially, labeled vehicles were primarily concentrated near the image center, which explained the occasional missed detections and false alarms in edge regions. However, this spatial bias exerted less influence on typical mid-range traffic surveillance views, where vehicles predominantly appeared near the center of the surveillance images. Two representative real-world surveillance cases were selected for process-level validation, including a daytime road scene caused by short-duration intense rainfall and a nighttime parking-lot scene subjected to long-duration rainfall to evaluate whether the trained model and the subsequent hydrograph reconstruction strategy remained reliable under different illumination conditions and rainfall regimes. The reconstructed depth-time hydrograph for the daytime scenario exhibited a pronounced S-shaped pattern and was divided into three stages, with mean rising rates of 3.05 cm/min (initial expansion), 8.69 cm/min (rapid deepening), and 1.05 cm/min (high-water maintenance). Error verification against the manual measurements yielded RMSE (3.81 cm), MAE (3.30 cm), and a maximum deviation of 6.91 cm. Relative to the peak depth of 113.0 cm, the MAE and RMSE corresponded to 2.92% and 3.37%, respectively, indicating that the key-frame interpolation accurately captured both the magnitude and temporal trend at a level suitable for stage characterization. The hydrograph for the nighttime scenario followed a “slow accumulation-accelerated growth” pattern, with mean rising rates of 0.56 cm/min and 1.74 cm/min during the two stages. The corresponding errors were RMSE (2.34 cm), MAE (1.99 cm), and a maximum deviation of 4.61 cm. The relative errors were higher because the peak depth was smaller (39.0 cm), reaching 5.10% (MAE/peak), 6.00% (RMSE/peak), and 11.82% (max/peak). This increased sensitivity was consistent with the effects of low illumination and reflective backgrounds, which reduced the contrast of shallow-water boundaries, as well as the slow-varying flood process, in which small disturbances constituted a larger proportion of the overall signal.

Conclusions A surveillance video and vision AI framework was developed to recognize urban waterlogging levels and reconstruct water-depth hydrographs without co-located gauges. Using a vehicle-referenced dataset comprising 1 911 images, a YOLOv8-based model achieved an mAP of 0.673 for multi-level waterlogging recognition. The proposed “key-frame calibration + linear interpolation” strategy generated time-continuous depth-time curves with centimeter-level reconstruction errors for both a daytime rapid-evolution storm (RMSE 3.81 cm; MAE 3.30 cm) and a nighttime long-duration event (RMSE 2.34 cm, MAE 1.99 cm). The proposed framework enables the low-cost, non-contact extraction of flood-process information from existing surveillance infrastructure while providing actionable process data for event interpretation, emergency response support, and hydrologic model calibration in observation-limited urban environments.

Key words: urban waterlogging; flood inundation; water-depth hydrograph reconstruction; computer vision-based AI; YOLOv8; disaster prevention and mitigation

Graphical abstract:



(编辑 周璇)

引用格式: Qin Youwei,Li Xinhao,Zheng Feifei,et al.Recognition of urban flooding process based on surveillance video and visual AI[J].Advanced Engineering Sciences,2026,58(5):349–359.[秦友伟,李心昊,郑飞飞,等.基于视频监控及视觉AI的城市内涝积水过程识别[J].工程科学与技术,2026,58(5):349–359.]