

城市雨水管理

DOI: 10.19853/j.zgjsps.1000-4602.2024.09.019

基于YOLOv7算法的易涝区积水快速识别方法研究

金晓飞¹, 郭 帅¹, 黄 琼², 袁 伟³

(1. 合肥工业大学 土木与水利工程学院, 安徽 合肥 230009; 2. 安徽省职业技术学院
智能制造学院, 安徽 合肥 230011; 3. 合肥市测绘设计研究院有限公司, 安徽 合肥
230061)

摘 要: 为了能够高效便捷地监测识别城市易涝区积水情况,利用深度学习技术,提出了一种基于YOLOv7算法的快速识别积水方法。采用传统数据增强和Mosaic数据增强方法对训练集图像进行扩充,构建了YOLOv7积水检测模型,并与其他主流目标检测模型Faster R-CNN和YOLOv5m进行了对比分析。结果表明,YOLOv7模型取得了最好的效果,其精度、召回率、平均精度、 F_1 分数分别达到了92.9%、83.4%、88.8%和87.9%,且单张图片推理时间仅约为0.025 s。该方法在城市内涝积水识别与预警方面具有良好的应用前景。

关键词: 城市内涝; 积水识别; 目标检测; YOLOv7算法; 深度学习

中图分类号: TU992 **文献标识码:** A **文章编号:** 1000-4602(2024)09-0123-06

A Method for Rapid Identification of Waterlogging in Flood-prone Area Based on YOLOv7

JIN Xiao-fei¹, GUO Shuai¹, HUANG Qiong², YUAN Wei³

(1. School of Civil Engineering, Hefei University of Technology, Hefei 230009, China; 2. College of Intelligent Manufacturing, Anhui Vocational and Technical College, Hefei 230011, China;
3. Hefei Surveying and Mapping Design Institute Co. Ltd., Hefei 230061, China)

Abstract: This paper proposed a method for rapid identification of waterlogging based on YOLOv7 algorithm by using deep learning technology, so as to monitor and identify waterlogging in flood-prone area efficiently and conveniently. The traditional and Mosaic data augmentation methods were used to expand the training set images, and the YOLOv7 waterlogging detection model was established. Then, the model was compared with other mainstream object detection models (Faster R-CNN and YOLOv5m). The YOLOv7 model achieved the best performance. Its precision, recall, average precision and F_1 score reached 92.9%, 83.4%, 88.8% and 87.9% respectively, and the inference time of a single image was only about 0.025 s. This method demonstrates a good application prospect in the identification and early warning of urban waterlogging.

Key words: urban waterlogging; waterlogging identification; object detection; YOLOv7

基金项目: 安徽省重点研发计划项目(202104i07020012)

通信作者: 黄琼 E-mail: ahzy8093@126.com

algorithm; deep learning

近年来,随着我国城市化进程的快速发展与全球气候变化导致的极端天气明显增多,城市内涝灾害呈现急剧上升的趋势^[1-2],因此,对城市易涝地区进行实时、高效的内涝积水信息监测与预警意义重大。当前,获取城市积水信息的方法多为利用在线仪器监测,但高额的成本限制了其大规模应用。最近几年,随着卷积神经网络(CNN)的兴起^[3],基于深度学习技术的目标检测算法发展迅速,在智慧交通、工业检测、医疗影像等方面均得到了广泛应用^[4-6],使得便捷、低成本的积水信息智能监测成为可能。在众多目标检测算法中,YOLO(You Only Look Once)算法凭借其优秀的检测速度与准确率脱颖而出。赵德安等^[7]提出了基于YOLOv3复杂背景下机器人采摘苹果定位的方法;谈世磊等^[8]使用YOLOv5算法实现了对口罩佩戴情况的实时检测;宋立博等^[9]精简了YOLOv4-tiny网络,从而获得两类YOLOv4-tiny简化网络,并比较了其裂缝检测性能。

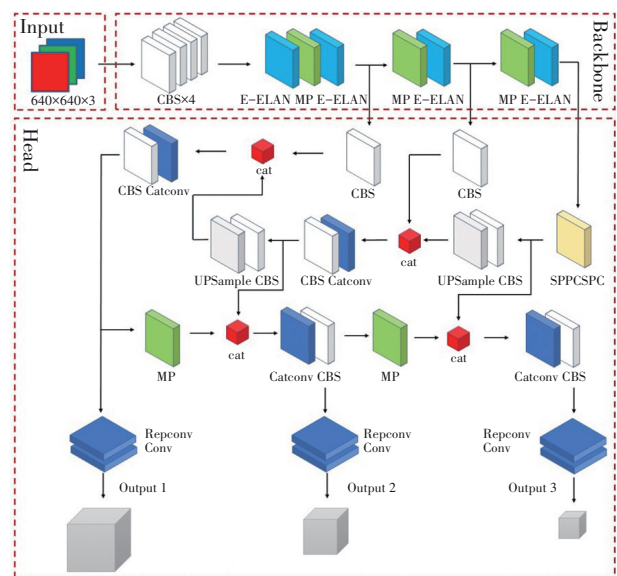
笔者基于YOLOv7算法^[10],提出了一种积水快速识别技术。利用收集到的积水图像,详细阐明了训练YOLOv7网络模型的过程,并与其他几种目标检测算法进行了比较,显示了YOLOv7模型在积水检测方面的优良性能。该方法可以实现积水的快速准确识别,可为城市内涝积水信息监测与预警提供方法和技术依据。

1 YOLOv7原理

目标检测的任务是找出图像中所有感兴趣的目标,并确定它们的类别和位置。目标检测算法分为两类——以R-CNN^[11]为代表的two-stage算法和以YOLO系列为代表的one-stage算法。YOLOv7是YOLO系列的最新作品,作为一个单阶段目标检测模型,YOLOv7以图像为输入,通过反向传播直接回归目标类和坐标。处理后的图片被映射为三通道图像作为输入,图像中存在相应目标时,模型输出 (x,y,w,h,c) 数组,分别表示目标矩形中心所在水平和垂直坐标、目标矩形长度和宽度、目标类别。模型由三部分组成:Input、Backbone和Head。Input层首先对输入图像进行预处理,对齐成640×640大小的RGB图片,处理后的图像被馈入Backbone网络以

提取特征;基于Backbone网络中三层的输出,在Head层融合它们,并执行分类和边界框坐标预测,作为最终结果。

YOLOv7网络在先前工作的基础上进一步提高了检测速度和准确性。具体而言,在总体架构方面,YOLOv7使用了E-ELAN(Extended-ELAN,扩展的高效长程注意力网络),对基数进行了扩展、乱序、合并,以实现在不破坏原始梯度路径的情况下持续增强网络的学习能力,E-ELAN还可以引导不同组的计算块学习更多样化的特征。YOLOv7还采用了一种复合模型缩放方法,以保持模型在初始设计时的特性,并保持最佳结构。在网络优化策略方面,YOLOv7就模型重参数化和动态标签分配存在的问题进行了改进。对于前者,由于RepConv^[12]具有身份连接,直接访问ResNet^[13]或DenseNet^[14]的级联将为不同的特征图提供更多的梯度多样性,从而破坏网络结构。在YOLOv7中删除了RepConv中的身份连接,并设计了计划的重参数化卷积,实现了重参数化的卷积与不同网络的有效组合。对于后者,YOLOv7使用深度监督^[15]的思想,并在网络的中间层添加一个额外的辅助头结构作为辅助损耗,以引导浅层网络的权重,为此结构设计了一种新的标签分配方法。YOLOv7的网络架构如图1(a)所示,各模块的具体结构如图1(b)所示。



a. 网络架构

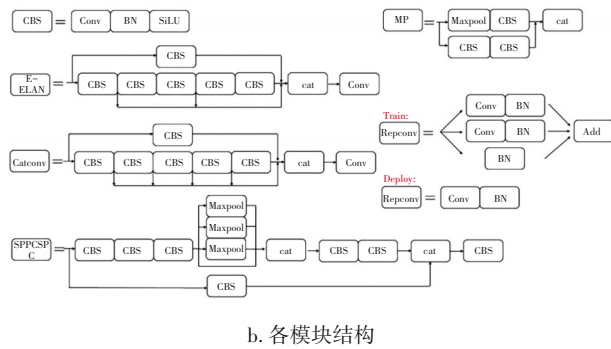


图1 YOLOv7网络架构及各模块结构

Fig.1 Architecture of YOLOv7 network and structure of each module

2 数据处理

2.1 图片采集与预处理

本研究所使用的积水图片主要来源于谷歌浏览器爬取和智能手机实际拍摄。经过筛选,保留了800张图片用以训练,并按8:2的比例随机划分为训练集和验证集。数据集相比于大多数实验室较为单一的背景场景,背景更加复杂多样、处理难度大,但这有助于提高训练模型的鲁棒性和泛化能力。

模型训练前需要对积水数据集进行标注,使用LabelImg标注工具进行标签制作,根据积水周围最小的矩形对图像进行标记,以确保矩形中包含的背景面积尽可能小,并将其保存为YOLO文件格式。

2.2 数据增强

深度学习依赖于大量高质量图片数据集,由于数据的获取与标注耗费大量人力物力,数据增强可让有限的数据产生更多的数据,从而减少网络的过拟合现象,提高模型的泛化能力。为兼顾样本有效性与训练速度,本研究对训练集的640张图片采用多种数据增强方法,其中主要采用了Mosaic增强方法^[16],扩充了一倍数据集。数据增强方法包括旋转(随机旋转90°、180°或270°,100张)、镜像(左右或上下镜像,100张)、随机HSV增强(色相 H 、饱和度 S 、明度 V 变化倍率范围分为0.015、0.7和0.4,100张)、Mosaic数据增强(340张)。Mosaic数据增强方法简单来说就是把4张图片通过随机缩放、随机裁减、随机排布的方式进行拼接,图2反映了其中一个例子。这样做有诸多优点,一是丰富数据集:随机选取4张图片进行组合,组合得到的图像个数比原图个数要多,同时也增强了模型鲁棒性,混合4张具有不同语义信息的图片,可以让模型检测超出常规

语境的目标;二是有利于提升小目标检测性能:Mosaic数据增强图像由4张原始图像拼接而成,这样每张图像会有更大概率包含小目标;三是加强批归一化层(Batch Normalization, BN)的效果:BN原理为计算每一个特征层的均值和方差,4张图片拼接在一起变相地提高了批大小,在进行BN操作时可以更好地统计均值和方差。

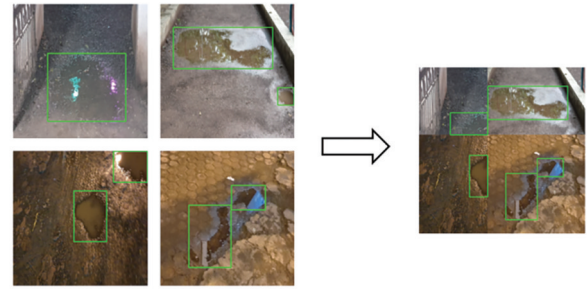


图2 Mosaic数据增强示例

Fig.2 Example of Mosaic data augmentation

3 模型训练与结果分析

3.1 模型训练

基于开源深度学习框架PyTorch,在Windows 10系统上使用Conda构建运行环境:Python 3.10+PyTorch Stable 1.12.1+ CUDA 11.6,硬件配置如下:CPU为Intel(R) Xeon(R) CPU E5-2630 V4 @ 2.20 GHz, GPU为NVIDIA Quadro P4000,显存为8 GB。

使用官方提供的预训练模型进行迁移学习训练YOLOv7积水检测模型。使用SGD with momentum(基于动量的随机梯度下降)优化算法,初始学习率设置为0.001,动量因子设置为0.937,权重衰减系数为0.0005,批大小为8,训练轮数为200轮。

3.2 评价指标

采用精确率(P)、召回率(R)、 F_1 分数(F_1)、平均精度(AP)以及反映模型推理速度的FPS(单位时间内检测的图片数量)等指标评价试验所训练模型的性能。精确率也叫查准率,表示在所有预测值为正例的样本中,正确预测为正例的比例;召回率也叫查全率,即正确预测为正例的样本占有所有真实值为正例的比例; F_1 分数和平均精度综合反映了精确率和召回率指标,前者为精确率和召回率的调和平均数,后者为PR曲线(以召回率作为横坐标轴、精确率作为纵坐标轴所绘制的曲线)与坐标轴围成的图形面积, F_1 分数和AP的取值范围为 $[0,1]$,越接近1说明模型性能越好。这些指标的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

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

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

式中:TP表示 $IoU > IoU_{threshold}$ [IoU, 交并比, 即“预测的边框”和“真实的边框(GT)”的交集和并集的比值]的检测框数量;FP表示 $IoU \leq IoU_{threshold}$ 的检测框或检测到同一个GT的多余检测框的数量;FN表示未检测到的GT的数量, 本研究中 $IoU_{threshold}$ 取0.5。

3.3 训练结果

训练集和验证集的检测框损失及置信度损失拟合曲线分别如图3和图4所示。

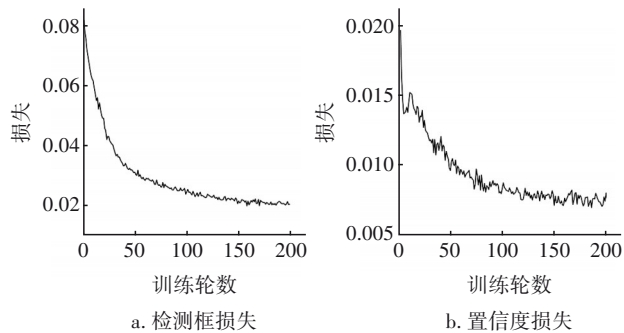


图3 训练集损失

Fig.3 Loss of training set

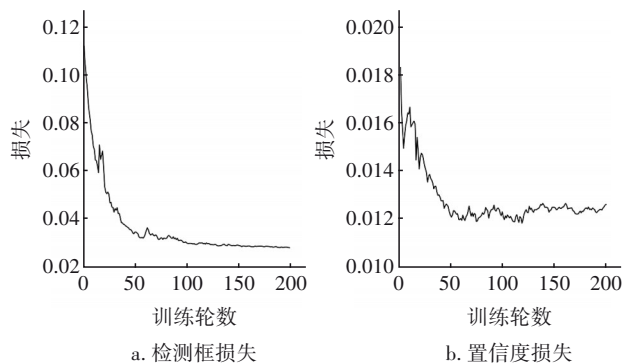


图4 验证集损失

Fig.4 Loss of verification set

由图3和图4可知,损失值在前50轮训练中迅速下降,在随后的100轮训练中缓慢下降,并且逐渐稳定,可以判断训练过程中没有发生欠拟合或过拟合。经过200轮训练之后,训练集的检测框损失稳定在0.02左右,置信度损失稳定在0.07左右;验证集的两类损失分别稳定在0.03和0.012左右。图5

反映了训练集的 P 、 R 和 AP 变化,可以看出, P 、 R 及 AP 在200轮训练之后都达到了较高的值。

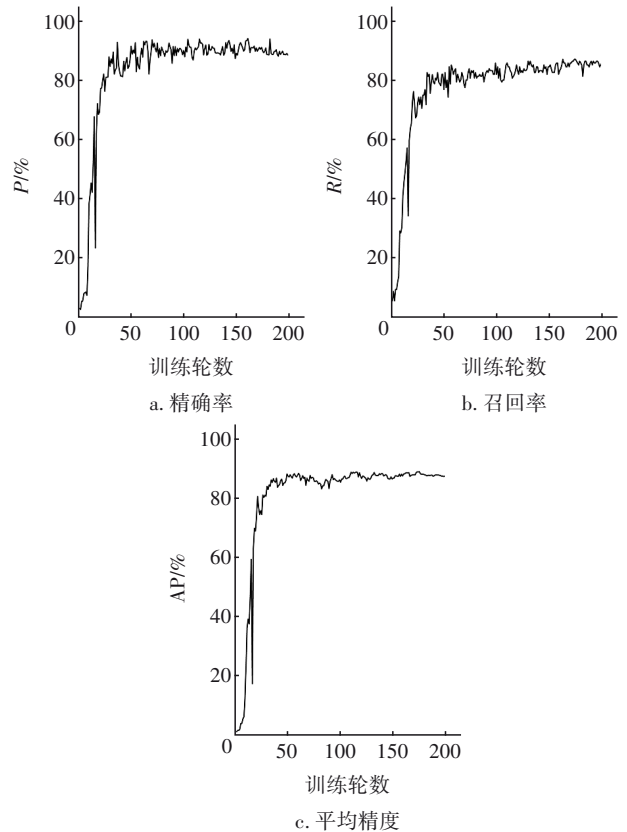
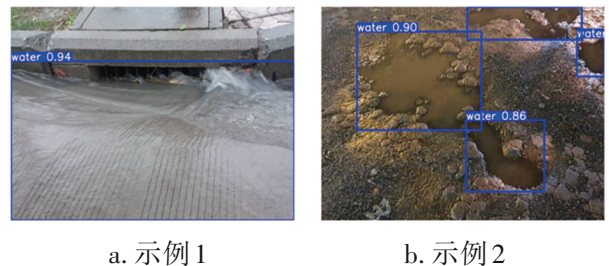


图5 训练集的 P 、 R 、 AP 变化

Fig.5 Change of P , R and AP of training set

经过200轮训练之后,可以得到表现最好的模型权重文件,在160张验证图片集上运行测试,结果为: $P=92.9\%$ 、 $R=83.4\%$ 、 $AP=88.8\%$ 、 $F_1=87.9\%$ 、 $FPS=39.7$ 。这表明训练的模型在积水检测方面具有较高的准确率,并且检测速度非常快,证明了其在实时检测方面的潜力。相较于精确率,模型的召回率偏低,这可能是训练图片较少的缘故,导致一些背景下的积水不能被识别出来。训练好的模型可以对图片及视频进行积水检测推理,见图6。



a. 示例1

b. 示例2

图6 模型推理结果示例

Fig.6 Example of model inference results

3.4 模型比较

将YOLOv7模型与YOLOv5m和Faster R-CNN(ResNet50 FPN)模型进行比较,以验证其准确性和有效性。各模型都是基于与YOLOv7模型相同的MS COCO数据集建立的预训练权重训练得到的,并且训练数据集和训练参数也保持相同。3种模型的 P 、 R 、 AP 、 F_1 和FPS如表1所示。总体而言,3种模型在积水检测方面都达到了较好的效果。YOLOv7模型的精确率高于YOLOv5m和Faster R-CNN模型,分别提升了3.3%和7.5%;YOLOv7模型平均精度比YOLOv5m和Faster R-CNN模型高了0.7%和3.9%, F_1 分数高了2.2%和3.2%,反映检测速度的FPS高了4.6和26.7。YOLOv7模型在各个方面均优于过往版本YOLOv5m。Faster R-CNN模型在试验中获得了更高的召回率,但在其他方面的指标不及YOLOv7模型,并且作为一个两阶段目标检测算法,在检测速度上远低于YOLO算法。

表1 不同模型的性能比较

Tab.1 Performance comparison of different models

目标检测模型	$P/\%$	$R/\%$	$AP/\%$	$F_1/\%$	FPS
YOLOv7	92.9	83.4	88.8	87.9	39.7
YOLOv5m	89.6	82.2	88.1	85.7	35.1
Faster R-CNN	85.4	84.0	84.9	84.7	13.0

3.5 模型不足

本研究所训练的模型在积水识别方面取得了良好的结果,但对积水面积、积水深度等定量信息的获取仍需进一步研究。理论上,利用语义分割算法可将积水区域与背景分割开来,利用图像与实际成一定比例的相似关系可以提取其积水面积;积水深度的判读也可借助参照物如水尺来实现。上述方法在实际应用中的效果是今后研究的主要内容。

4 结论

基于深度学习的目标检测算法具有广泛的应用场景,本研究训练了YOLOv7模型以验证其在城市积水监测识别方面的可行性与有效性。利用谷歌浏览器爬取和手机拍摄收集了800张积水图像,其中640张作为训练集、160张作为验证集。经传统方法(旋转、镜像和随机HSV增强)和Mosaic方法进行数据增强,将训练集扩充至1280张。之后,使用合适参数对YOLOv7网络模型进行了迁移学习,并与YOLOv5m和Faster R-CNN模型进行了比较。YOLOv7模型在验证集上的 P 、 R 、 AP 、 F_1 和FPS分别

为92.9%、83.4%、88.8%、87.9%和39.7,优于YOLOv5m和Faster R-CNN模型。训练的YOLOv7模型取得了良好的结果,可用于城市积水的实时监测识别,为相关部门提供理论依据与技术支持。

参考文献:

- [1] 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.
ZHANG Jianyun, WANG Yintang, HE Ruimin, *et al.* Discussion on the urban flood and waterlogging and causes analysis in China [J]. Advance in Water Science, 2016, 27(4): 485-491 (in Chinese).
- [2] 鲁芳,刘艳丽,朱士江,等.2020年主汛期我国暴雨洪涝灾害概况分析与初步思考[J].水电能源科学,2022,40(5):68-71,150.
LU Fang, LIU Yanli, ZHU Shijiang, *et al.* Analysis and preliminary thinking of rainstorm and flood in China in the flood period of 2020 [J]. Water Resources and Power, 2022, 40(5): 68-71, 150 (in Chinese).
- [3] ALZUBAIDI L, ZHANG J L, HUMAIDI A J, *et al.* Review of deep learning: concepts, CNN architectures, challenges, applications, future directions [J]. Journal of big Data, 2021, 8(1): 53.
- [4] 曹诗雨,刘跃虎,李辛昭.基于Fast R-CNN的车辆目标检测[J].中国图象图形学报,2017,22(5):671-677.
CAO Shiyu, LIU Yuehu, LI Xinzhaoh. Vehicle detection method based on Fast R-CNN [J]. Journal of Image and Graphics, 2017, 22(5): 671-677 (in Chinese).
- [5] 周倩倩,司徒祖祥,腾帅,等.基于卷积神经网络的排水管道缺陷智能检测与分类[J].中国给水排水,2021,37(21):114-118.
ZHOU Qianqian, SITU Zuxiang, TENG Shuai, *et al.* Intelligent detection and classification of drainage pipe defects based on convolutional neural networks [J]. China Water & Wastewater, 2021, 37(21): 114-118 (in Chinese).
- [6] GAUR L, BHATIA U, JHANJHI N Z, *et al.* Medical image-based detection of COVID-19 using deep convolution neural networks [J]. Multimedia Systems, 2021, 29: 1729-1738.
- [7] 赵德安,吴任迪,刘晓洋,等.基于YOLO深度卷积神经网络的复杂背景下机器人采摘苹果定位[J].农业工程学报,2019,35(3):164-173.
ZHAO De'an, WU Rendi, LIU Xiaoyang, *et al.* Apple

- positioning based on YOLO deep convolutional neural network for picking robot in complex background [J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(3): 164–173 (in Chinese).
- [8] 谈世磊, 别雄波, 卢功林, 等. 基于YOLOv5网络模型的人员口罩佩戴实时检测[J]. 激光杂志, 2021, 42(2): 147–150.
- TAN Shilei, BIE Xiongbo, LU Gonglin, *et al.* Real-time detection for mask-wearing of personnel based on YOLOv5 network model [J]. Laser Journal, 2021, 42(2): 147–150 (in Chinese).
- [9] 宋立博, 费燕琼. 两类YOLOv4-tiny简化网络及其裂缝检测性能比较[J]. 同济大学学报(自然科学版), 2022, 50(1): 129–137.
- SONG Libo, FEI Yanqiong. Comparison of two types YOLOv4-tiny simplified networks and their crack detection performance [J]. Journal of Tongji University (Natural Science), 2022, 50(1): 129–137 (in Chinese).
- [10] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]// Proceedings of IEEE. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023: 7464–7475.
- [11] 王万国, 田兵, 刘越, 等. 基于RCNN的无人机巡检图像电力小部件识别研究[J]. 地球信息科学学报, 2017, 19(2): 256–263.
- WANG Wanguo, TIAN Bing, LIU Yue, *et al.* Study on the electrical devices detection in UAV images based on region based convolutional neural networks [J]. Journal of Geo-information Science, 2017, 19(2): 256–263 (in Chinese).
- [12] DING X H, ZHANG X Y, MA N N, *et al.* RepVGG: making VGG-style ConvNets great again [C]// IEEE. Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 13728–13737.
- [13] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]// IEEE. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 770–778.
- [14] HUANG G, LIU Z, VAN DER MAATEN L, *et al.* Densely connected convolutional networks [C]// IEEE. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 4700–4708.
- [15] LEE C Y, XIE S, GALLAGHER P, *et al.* Deeply-supervised nets [J]. Journal of Machine Learning Research: W&CP, 2015, 38: 562–570.
- [16] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [J/OL]. (2020-04-23) [2022-12-10]. <https://doi.org/10.48550/arXiv.2004.10934>.

作者简介: 金晓飞(1998–), 男, 安徽阜阳人, 硕士研究生, 主要研究方向为城市内涝防治。

E-mail: 1531086571@qq.com

收稿日期: 2022-12-26

修回日期: 2023-03-11

(编辑: 刘贵春)

大力推进水利薄弱环节建设,
提高防灾减灾能力