



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目： 利用社会公众拍摄图像构建城市内涝深度等级评估模型
作者： 薛丰昌, 陈笑娟, 吕鑫, 成沿年
DOI: 10.13203/j.whugis20240290
收稿日期: 2024-11-10
网络首发日期: 2024-11-29
引用格式: 薛丰昌, 陈笑娟, 吕鑫, 成沿年. 利用社会公众拍摄图像构建城市内涝深度等级评估模型[J/OL]. 武汉大学学报(信息科学版), <https://doi.org/10.13203/j.whugis20240290>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI:10.13203/j.whugis20240290

引用格式：

薛丰昌, 陈笑娟, 吕鑫, 等. 利用社会公众拍摄图像构建城市内涝深度等级评估模型[J]. 武汉大学学报(信息科学版), 2024, DOI:10.13203/J.whugis20240290 (XUE Fengchang, CHEN Xiaojuan, LÜ Xin, et al. Constructing an Urban Waterlogging Depth Assessment Model Using Public Video Information[J]. Geomatics and Information Science of Wuhan University, 2024, DOI:10.13203/J.whugis20240290)

利用社会公众拍摄图像构建城市内涝深度等级评估模型¹

薛丰昌^{1,2}, 陈笑娟³, 吕鑫², 成沿年²

1 南京信息工程大学气象灾害地理信息工程实验室, 江苏 南京 210044

2 中国气象局流域强降水重点开放实验室, 湖北 武汉 430205

3 河北省气象灾害防御和环境气象中心, 河北 石家庄 050021

摘要：衡量城市内涝严重程度的标准除了积水范围外, 积水深度是更为重要的一个标准。面向传统城市内涝积水状态监测方法存在成本高昂、监测范围有限以及仅能针对特定易涝区域进行监测的局限性, 提出了一种利用社会公众拍摄的图像信息构建的城市内涝深度等级评估模型。我们利用爬虫技术以“城市内涝”和“车辆受淹”等词语作为关键词, 从互联网上收集社会公众拍摄的包含涉水小型车辆的图像和视频构建样本集, 采用车辆淹没深度作为参照标准, 将内涝深度划分为安全(0-20cm)、不安全(20-60cm)和危险(>60cm)三个等级, 并为这些等级创建了相应的标签。应用 Yolov8 模型构建了 Yolov8l, Yolov8m, Yolov8n, Yolov8s 四组不同网络复杂度的模型, 训练过程中, 利用混淆矩阵和精确度(Precision)和召回率(Recall)、平均精度均值(mAP)、平均精度均值精度(AP)四种精度指标对模型的准确性进行了综合评估。实验结果证实, Yolov8 算法能够有效识别不同场景下车辆的淹没状态, 且对各危险等级的车辆识别平均精确度达到了70%。这一精度水平表明, 该模型能够作为一种有效的工具, 用于监测和评估城市内涝的积水深度等级。本研究不仅提高了内涝灾害评估的时效性和准确性, 还拓展了社会信息在灾害管理中的应用, 为未来的城市灾害管理提供了一种人工智能与社会公众信息相结合的解决方案。

关键词：城市内涝; 公众图像; 人工智能; 评估模型

Constructing an Urban Waterlogging Depth Assessment Model Using Public Video Information

收稿日期：2024-11-10

基金项目：中国气象局流域强降水重点开放实验室基金项目(2023BHR-Y29); 河北省省级科技计划项目(22375421D); 河北省资源环境灾变机理及风险监控重点实验室开放基金项目(FZ248101)

第一作者：薛丰昌, 教授, 博士生导师, 从事气象 GIS、气象灾害监测预警与风险管理、城市内涝数值模拟、灾害遥感等研究。001968@nuist.edu.cn

通讯作者：陈笑娟, 高级工程师, 从事气象防灾减灾研究。juan2044335@163.com

XUE Fengchang^{1,2}, CHEN Xiaojuan³, LÜ Xin², CHENG Yannian²

1 Laboratory of Geo-Information Engineering for Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing 210044

2 Key Open Laboratory of Heavy Precipitation over Basin of China Meteorological Administration, Wuhan 430205

3 Hebei Province Meteorological Disaster Prevention and Environmental Meteorological Center, Shijiazhuang 050021

Abstract: Objectives: In addition to the extent of water accumulation, the depth of water accumulation is a more important criterion for measuring the severity of urban waterlogging. In response to the limitations of traditional urban waterlogging monitoring methods, such as high costs, limited monitoring range, and the ability to only monitor specific flood-prone areas, a model for assessing the depth levels of urban waterlogging using image information captured by the public has been proposed. Methods: We used web crawling techniques with keywords like "urban waterlogging" and "vehicles submerged" to collect images and videos from the internet that include small vehicles in water, creating a sample set. Using the depth of vehicle submersion as a reference standard, we divided the waterlogging depth into three levels: safe (<20cm), unsafe (20-60cm), and dangerous (>60cm), and created corresponding labels for these levels. We applied the YOLOv8 model to construct four models of different network complexities: YOLOv8l, YOLOv8m, YOLOv8n, and YOLOv8s. During training, we used a confusion matrix and four accuracy metrics—Precision, Recall, mean Average Precision (mAP), and Average Precision (AP)—to comprehensively evaluate the model's accuracy. Results: (1) For the training set, all four YOLOv8 models converged after about 40 training epochs, with Precision, Recall, and mAP50 values stabilizing around 60%-70%, and mAP50:95 values around 35%. (2) In the test set, for the Precision metric, YOLOv8m scored the highest for the safe class at 80.7%, YOLOv8l for the unsafe class at 55.4%, and YOLOv8s for the dangerous class at 84.8%. For the Recall metric, YOLOv8m scored the highest for both the safe and unsafe classes at 66.3% and 68.8%, respectively, while YOLOv8n scored the highest for the dangerous class at 68.3%. For the mAP50 metric, YOLOv8s scored the highest for all classes, with 74.4% for safe, 61.7% for unsafe, and 76.2% for dangerous. For the mAP50:95 metric, YOLOv8m scored the highest for the safe class at 43.1%, while YOLOv8s scored the highest for both the unsafe and dangerous classes at 37.7% and 36.7%, respectively. (3) It should be noted that the depth and width of the network, as well as the ability to extract features, also have a significant impact on model performance. Although increasing the width and depth of the network significantly increases the number of model parameters and computational load, not all evaluation metrics increase with the width and depth of the network. Although the YOLOv8l model is the most complex, its Precision and mAP50 scores are not dominant, ranking fourth and third at 69.3% and 68.0%, respectively, while its Recall and mAP50:95 scores are the highest. Despite the YOLOv8n model having the fewest parameters and computational load, its mAP50 score also achieved a second-place ranking. Overall, YOLOv8m performed the best on the validation set, with the highest scores in Precision, mAP50, and mAP50:95, with Precision and mAP50 being 2.2% and 2.6% higher than the second place, respectively, and without much sacrifice in Recall, making it the best overall performer. (4) The detection accuracy of the four models for the unsafe class is lower than that of the other two levels. On the one hand, this is due to the imbalance in the number of samples (unsafe class samples are 68% of safe class samples and 59% of dangerous class samples). On the other hand, it is also due to certain limitations in the model's detection performance, such as vehicles in motion easily splashing water, leading to misjudgment of water levels; when the water level is just at the halfway point of the tire, the unsafe class is easily confused with the safe class, and when the water level is just at the exhaust outlet of the engine, it is easily confused with the dangerous class,

thus the probability of correctly detecting the unsafe class is the lowest. To address this issue, future research will focus on targeted data augmentation techniques for sample processing to improve the model's recognition ability for unsafe class samples, and introduce attention mechanism modules into the model network structure to enhance the model's learning ability for key features.

Keywords: Urban waterlogging; Public image; Artificial intelligence; Evaluation model

衡量城市内涝严重程度的标准除了积水范围外，积水深度是更为重要的一个标准。当前，物联网技术与 GIS 建模相结合在城市内涝积水监测中取得了显著进展，GIS 技术结合电子水尺、雨量计等实现了对城市积水水位的实时监测和数据自动采集，这些数据通过无线传输至云平台，城市管理者能够及时获取积水情况并采取相应措施^[1-7]。然而这些技术在应用中仍存在局限性：首先，城市内涝的复杂性要求积水传感监测设施必须具有足够的代表性、准确性和覆盖度，而高精度积水传感监测设备购置成本、安置和维修成本都很高，很难实现在城市中大规模布设；其次独立积水传感监测设备只能针对易涝点进行监测，难以覆盖大范围地段。以上技术局限性导致基于积水传感监测设施的城市内涝监测技术方法尚不能完全满足城市内涝防灾减灾需求。

小型汽车作为城市重要交通工具，在城市内涝监测中也发挥着重要作用，通过在小型车辆上搭载监测设备，在积水区域进行巡回监测，可以对城市内涝情况进行实时监测和数据采集，而利用小型汽车自身淹没状态进行城市内涝状况判断目前尚属空白。社会公众拍摄的图像信息包含众多城市内涝信息，小型车辆在城市洪涝发生时有很高的出现几率，适合作为判断城市洪涝严重等级的动态参照物。

Yolov8 是 2023 年所提出的一个全新的 SOTA 模型，是目标检测领域中 YOLO 系列的最新模型，它在多个方面相较于其他深度学习模型展现出其优越性。YOLOv8 抛弃了传统的基于锚点（Anchor-Based）的检测方式，直接预测物体的中心点和尺寸，简化了检测流程并减少了超参数的调整，提高了检测的准确性和效率；YOLOv8 引入了新的网络主干和颈部架构，能够更有效地提取图像特征，从而提高目标检测的性能；YOLOv8 引入了新的损失函数，有助于在训练过程中更有效地优化模型参数，进一步提升检测精度^[8]。与 Faster R-CNN 相比，YOLOv8 在准确性和速度方面均优于 Faster R-CNN；与 EfficientDet 比较，在相同的评估中，EfficientDet 的 mAP@50 为 0.47，低于 YOLOv8 的 0.62^[9-10]。

本文基于 YOLOv8 提出了一个成本低、具有广泛适用性的城市内涝积水深度监测方法。具体思路是利用 Yolov8 目标检测模型对图像中的涉水车辆的状态进行识别，不同涉水状态对应一个积水深度等级，从而借助车辆这一参照物来表达积水深度。该方法具有普适性，可以广泛应用社会公众拍摄的视频图像，为城市内涝的深度监测提供有益方法，具有良好应用价值。

1 数据集与实验环境

本文的主要研究内容为对图像中在城市内涝环境下的小型涉水车辆的淹没

状态进行目标检测，根据车辆的淹没程度设定判定等级作为区分标准，用以表达该区域的积水深度情况，不同于正常条件情况下的车辆检测数据集，需要自行收集数据并制作标签。考虑到所需数据的需求条件，使用 Python 爬虫技术以“城市内涝”、“车辆淹没”为关键词在网络上爬取图片，最终爬取了 2314 张社会公众拍摄的图片，经过人工目视解译筛选出了 1141 张图片进行标注。

积水深度按照 27 厘米、40 厘米、60 厘米，风险等级由低到高划分为四级，积水深度小于 27 厘米属于较低风险区域；积水深度大于等于 27 厘米但小于 40 厘米属于中等风险区域；积水深度大于等于 40 厘米但小于 60 厘米属于较高风险区域；积水深度大于等于 60 厘米属于高风险区域^[11]。汽车自身高度固定，随着镜头视角和焦距的不同，汽车虽然在不同图像中大小不同，但其自身形状特征并没有变化。汽车在垂直地面方向上随着高度的变化，表现出不同的外形构造，作为参照物，在不同的淹没状态下能够表现出不同的外形特征，同时，随着汽车涉水深度的变化，汽车的驾驶安全等级也在变化。我们根据汽车的高度节点的主要特征以及汽车涉水安全等级变化节点将汽车的淹没状态分为三种。查阅相关资料，普通小轿车和 SUV 的轮胎半径大小在 33 厘米到 38 厘米之间，引擎进气口高度在 50-65 厘米之间。以此为参考，本研究将积水深度分为安全（0-20cm）、不安全（20-60cm）和危险（>60cm）三个等级，分别为淹没至轮胎半径及以下设为安全(safe)，轮胎半径到引擎进气口之间设为不安全(unsafe)，引擎进气口之上设为危险(danger)。分为三个等级的理由是：小型车辆有小轿车、SUV、小型皮卡等，不同车型轮胎大小亦有差异，导致内涝淹没到不同车型的同一位置时深度会存在波动，等级划分过多会导致各个等级特征不明显，模型预测时容易误分。数据集的分类标准如图 1 所示，三类样本数据如图 2 所示。

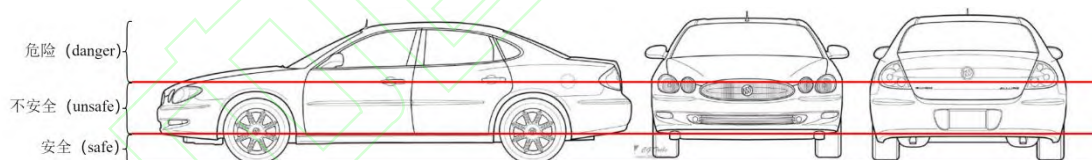


图 1 积水深度等级划分标准

Figure 1 Classification Standard for Floodwater Depth Levels





图 2 样本实例 (a、安全类样本; b、不安全类样本; c、危险类样本)

Figure 2 Sample instances

根据上文确定的数据集划分标准, 利用 Labelimg 手动标注边界框, 标注完成后会生成一个 xml 格式文件, 里面包含了一张图片中所标注的边界框四个顶点在该图片中的坐标位置, 边界框的中心坐标、宽度和高度, 以及该边界框选择物体的类别。图 3 为标签制作界面, xml 为 VOC 数据格式, 不同的目标检测模型需要不同的数据格式, 例如 Yolo 系列模型训练需要 txt 格式文件, 可以通过编程脚本完成数据格式之间的转化。



图 3 标签制作界面

Figure 3 Label Creation Interface

本研究将 1141 张图片按照 9:1 的比例划分训练集和测试集, 再从训练集中按照 9:1 的比例划分成训练集和验证集, 最终得到 923 张训练集, 103 张验证集和 115 张测试集。其中, 安全(safe)类别总共 1659 个边界框, 不安全(unsafe)类别总共 1128 个边界框, 危险(danger)类别总共 1884 个边界框。数据量的分布, 边界框高宽和中心点的分布如图 4 所示(图例表示标签中心点出现在图像某个区域的密度)。本文实验采用 Windows10 操作系统、Pytorch1.10 深度学习框架, 在 GPU NVIDIA RTX4080(16GB)完成网络的训练, batch_size 设置为 16, 学习率设置为 0.001, 使用 SGD 优化器, 迭代次数 Epoch 设为 100 次。

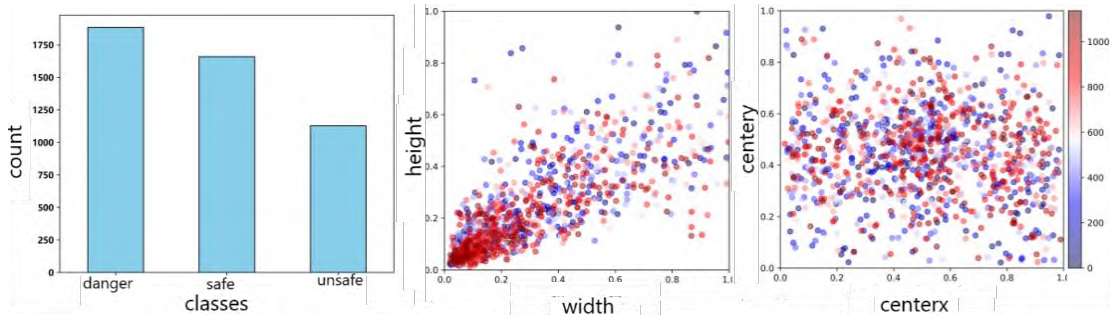


图 4 数据集样本分布情况

Figure 4 Distribution of Dataset Samples

2 研究方法

2.1 Yolov8 网络模型

Yolov8 主要由四个部分组成，分别是输入端(Input)、主干特征提取网络(Backbone)、加强特征提取网络(Neck)和检测头(Head)。Yolov8 的整体网络结构如图 5 所示。

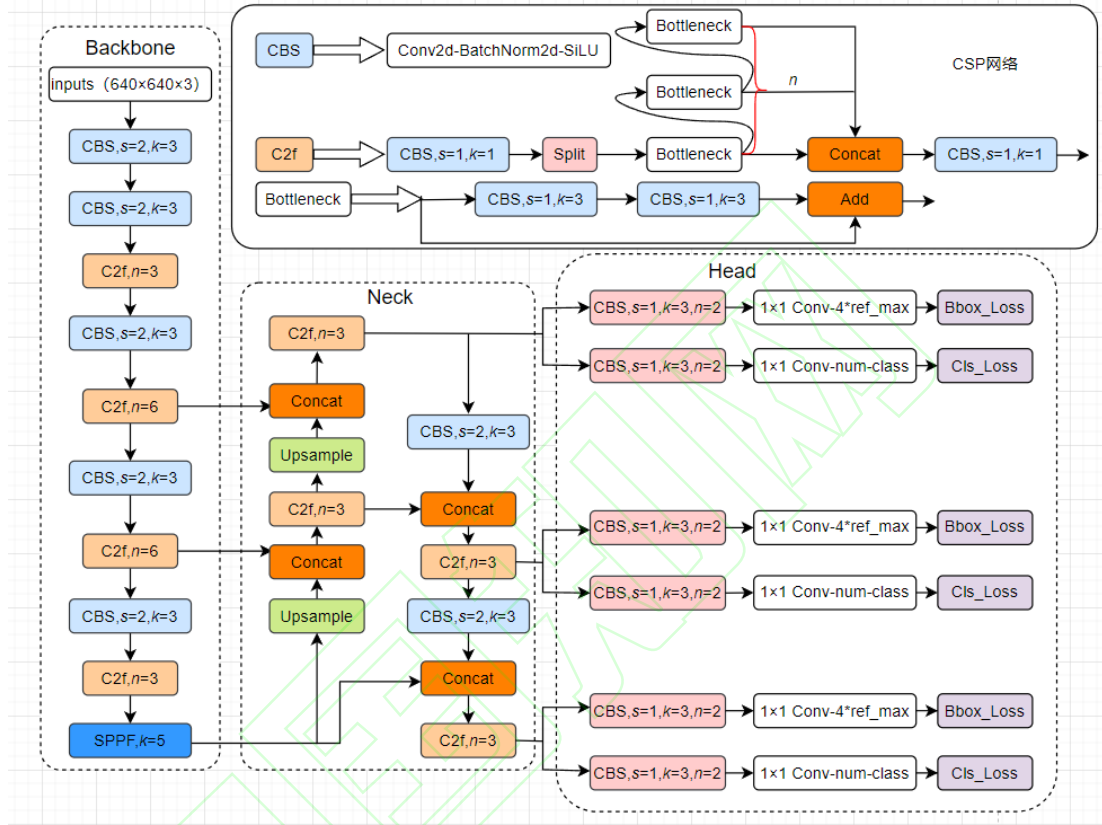


图 5 Yolov8 整体网络结构

Figure 5 Overall Network Structure of Yolov8

2.1.1 输入端结构

Yolov8 网络模型将图像输入主干特征提取网络前还有一个输入端操作，其目的在于对数据集进行预处理，起到数据增强的作用，主要包括图片自适应缩放、马赛克数据增强(Mosaic)、自适应锚框(Anchor)计算和 Mixup 数据增强^[12]。其中，马赛克数据增强的方法流程如图 6 所示，它包括马赛克增强、随机缩放和色彩空间(HSV)的调整。本文研究中，一张图片中不同区域的积水深度会不一样，而远处小目标车辆的受淹状态往往会难以识别，马赛克数据增强方法将一个 batch_size 中取出四张图片进行随机缩放、裁剪、打乱分布方式，然后重新拼接，可以有效解决小目标车辆状态难以识别的问题。自适应锚框计算会为识别目标设定最适合其尺寸大小的锚框。Mixup 数据增强会随机选取两个样本及其对应的标签进行特征线性插值，作为给模型引入的先验知识，以提升模型的泛化性。

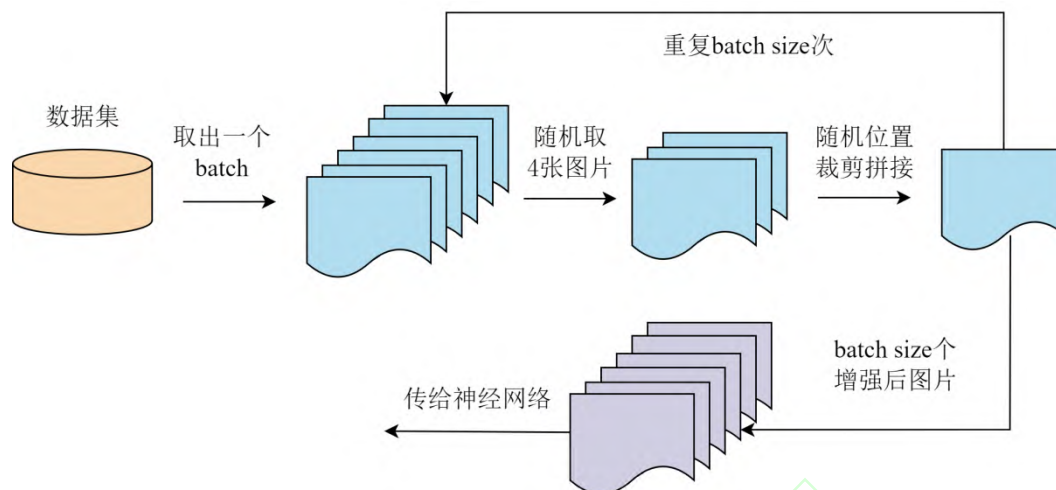


图 6 马赛克数据增强方法

Figure 6 Mosaic Data Augmentation Method

2.1.2 主干网络与加强特征提取网络

Yolov8 主干网络和 Neck 结构分别如图 5 左侧 Backbone 部分和中间 Neck 部分所示, 主要由跨阶段局部网络(Cross Stage Partial Network,CSP)、注意力模块和空间金字塔池化结构组成。CSP 借鉴了 Yolov7 的多堆叠结构^[13], 并将 Yolov5 的 C3 网络替换成了 C2f 网络, 其中 CBS 是卷积层, 而 Bottleneck 部分使用了 1×1 的卷积, 针对不同尺度的模型调整了通道数, 能使模型保持轻量化的同时获得更加丰富的梯度流信息, 从而使网络学习到更多的特征。Neck 模块采用了路径聚合网络(Path Aggregation Network, PAN)和特征金字塔网络(Feature Pyramid Network, FPN), 对主干网络提取的特征进行上采样和下采样跨层连接使特征充分融合^[8], 增强多个尺度上的语义表达和定位能力。Head 结构如图 5 右侧 Head 部分所示, 该结构采用了解耦检测头(Decoupled-Head)^[14], 将检测与分类分离, 解决了分类和定位关注侧重点不同的问题。同时采取了无锚框目标检测方法(Anchor-Free), 提升了检测速度。

2.1.3 任务对齐学习

在深度学习目标检测任务中, 正负样本不平衡是一个常见的问题^[15], 例如本研究的任务是识别图像中是否存在被积水淹没的车辆, 并且定位其位置。在这个过程中, 正样本是包含受淹车辆的图像区域, 而负样本是不包含受淹车辆的图像区域, 该问题会体现在以下几个方面: 一是样本分布不均衡, 存在受淹车辆的区域往往远远少于没有受淹车辆的区域, 这会导致样本分布不均匀。二是训练不稳定, 模型可能更偏向于学习负样本, 因为在数量上占主导地位, 而对于正样本的学习相对困难。三是影响性能评估, 当模型测试表现良好但未能准确识别少数类别目标时, 可能会掩盖模型对于整体性能的真实表现。

针对上述问题 Yolov8 采取了任务对齐学习(Task Alignment Learning, TAL)^[16]的动态分配策略, 根据训练的进展和样本的特点动态调整权重。TAL 根据分类和回归的分数加权选择正样本, 样本的分配能力可以随着训练不断地增强, 让模型更加专注受淹车辆正样本的学习。其核心计算公式如下:

$$t = s^\alpha \times u^\beta \quad (1)$$

利用上述公式可以衡量预测框和真实框的对其程度，其中 s 为分类得分， u 为 IoU 的值， α 和 β 为权重超参数。

2.2 评价指标

2.2.1 精度评价指标

本文采用精确度(Precision)和召回率(Recall)、平均精度均值(mAP)，平均精度均值精度(average precision, AP)进行精度评价，其计算公式如下：

$$AP = \int_0^1 P(R) dR \quad (2)$$

$$mAP = \frac{\sum_{N=1}^n AP}{N} \quad (3)$$

式中，AP 的值为精确度和召回率曲线(Precision-Recall, PR)所围成区域的面积， N 表示任务中目标检测的类别数量。在模型性能评估时，IoU 表示预测框与真实标注框之间的重叠程度，通常使用不同的 IoU 阈值来确定预测结果的正确与否。当 IoU 大于等于设定阈值时，认为预测正确，当 IoU 小于设定阈值时，认为预测错误。mAP50 表示 IoU 取值为 0.5 时的总类别平均精度，mAP50:95 表示不同 IoU 取值 (从 0.5 到 0.95，步长 0.05) 的平均精度。

2.2.2 损失函数

Yolov8 的损失函数由分类损失和回归损失构成，分类损失使用的是二元交叉熵损失，回归损失中的 CIoU 损失和计算 IoU 的思想类似，是一种描述预测框和真实框之间重合度的方式，通过比值来衡量框的回归程度。而当模型引入了 Anchor-Free 的基于中心点预测后，从输出锚框大小偏移量变成了预测框左、上、右、下边框与中心点的距离，因此 Yolov8 引入了分布焦点损失(Distribution Focal Loss, DFL)^[17-20]，它以交叉熵的形式，优化与标签最接近的左右的两个位置的概率，从而让网络更快聚焦到目标位置及临近区域的分布。假设 y 为真实标签，第 i 次和第 $i+1$ 次的预测为靠近 y 的预测，则有：

$$S_i = \frac{y_{i+1} - y}{y_{i+1} - y_i}, S_{i+1} = \frac{y - y_i}{y_{i+1} - y_i} \quad (4)$$

DFL 损失函数的公式表达为：

$$DFL(S_i, S_{i+1}) = -((y_{i+1} - y) \log(S_i) + (y - y_i) \log(S_{i+1})) \quad (5)$$

3 实验结果与分析

对 Yolov8 的 Yolov8l, Yolov8m, Yolov8n, Yolov8s 四种不同大小的网络算法训练过程中的变化指标和对测试集进行预测的结果指标进行分析，以评估模型在车辆受淹状态识别中的应用表现。其中，模型的训练过程选择 Precision、Recall、mAP50 和 mAP50:95 四个指标。如图 7 所示，可以观察到 Yolov8 四种不同大小网络算法随着训练次数的增加，最终训练轮次达到 40 次左右时都能趋于收敛状

态，并且随着迭代次数的上升四项指标都是呈现出振荡上升的状态，到训练结束时四个模型的 Precision、Recall、和 mAP50 值都维持在 60%-70%左右，mAP50:95 的值都维持在 35%左右。模型刚开始训练时，Yolov8s 的各指标的初始值最高，Yolov8l 的各指标初始值最低，但 Yolov8l 指标训练前期涨幅最大，到训练的中后期各指标达到相对平稳状态时，Yolov8l 的四项指标都略高于 Yolov8s。在整个训练周期中 Yolov8m 的四项指标变化与 Yolov8l 最为接近，Yolov8n 的四项指标都表现得最低。

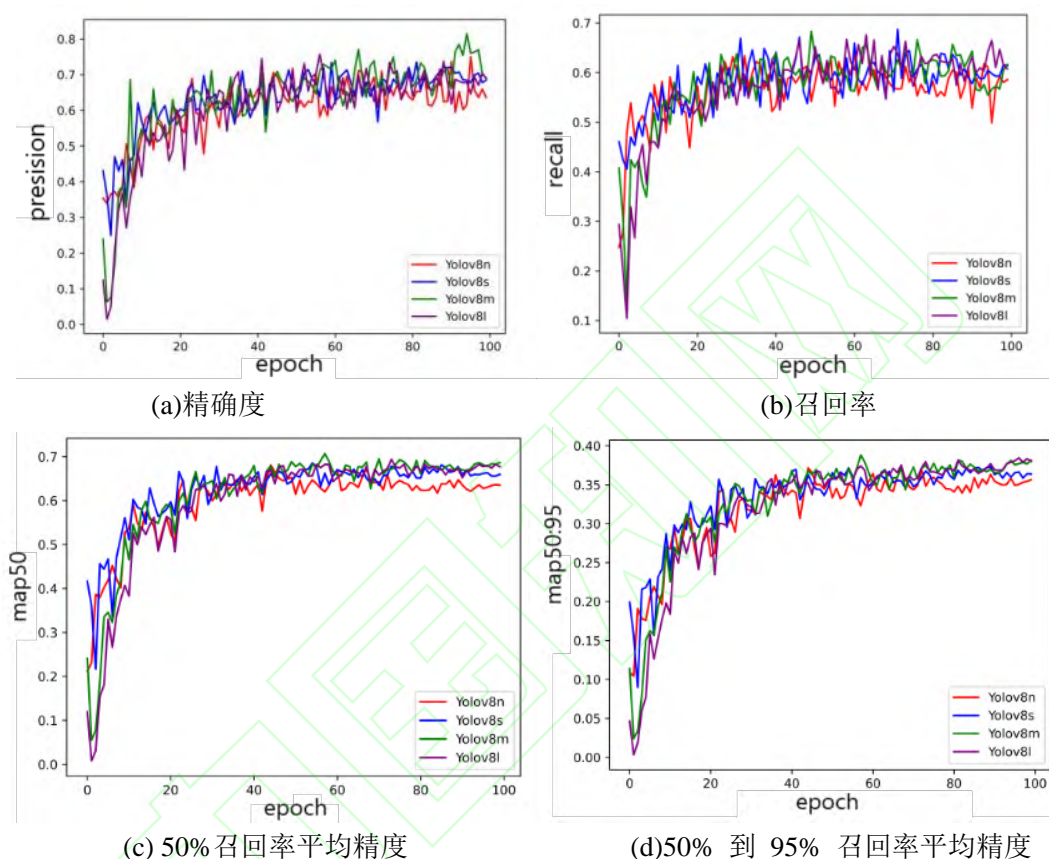


图 7 各模型训练参数变化

Figure 7 Variation of Training Parameters for Each Model

将四种 Yolov8 模型整个训练过程中对验证集进行测试的各类别混淆矩阵进行可视化展示，如图 8 所示（图例表示识别精度在整体样本识别结果中出现的概率）。在验证集中，四个模型均对 danger 类检测正确的概率最高，safe 类次之，unsafe 类最低，并且对 danger 类的漏检概率最高，由于验证集中存在许多俯视拍摄的大范围积水场景，这种场景积水很深，小型车辆被淹没到只剩顶棚的少部分露出在水外，人眼目视解译能够理解到是受淹程度严重的汽车，但训练集中这种极端情况的数据集较少，导致这些目标未能检测出来导致漏检率较高。当近距离拍摄时，danger 类的淹没汽车特征较为明显，因此正确识别的概率较高。反之，由于在被拍摄过程中多数车辆处于行动状态，容易将地面的积水见其水花，这种情况下人眼也难以判别积水等级，并且当积水正好淹没至轮胎一半附近时 unsafe 类容易和 safe 类混淆，当积水正好淹没至引擎排气口附近时容易与 danger 类混淆，所以正确检测 unsafe 类的概率最低。

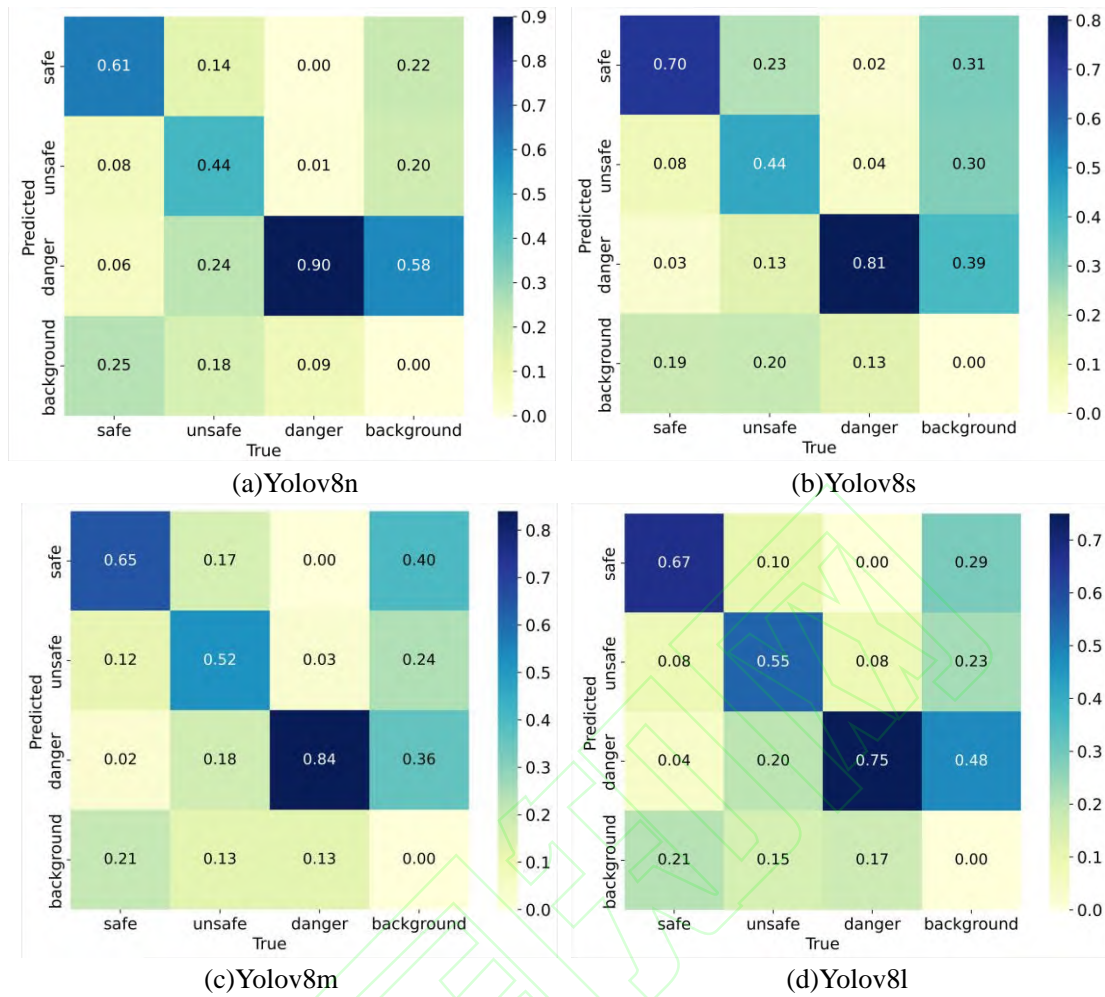


图 8 验证集混淆矩阵对比

Figure 8 Comparison of Confusion Matrices for the Validation Set

将四种 Yolov8 模型整个训练过程中验证集三个类别的 Precision、Recall、mAP50 和 mAP50:95 四个指标求平均，并且计算出各模型的数量和计算量结果如表 1 所示。可以看出，随着网络宽度和深度的增加，模型参数数量和计算量都在大幅度增加，但并不是所有的评价指标都随着网络宽度和深度的增加而增加。其中，虽然 Yolov8l 模型最为复杂，但其精确度和 mAP50 评分都不具优势，分别以 69.3%和 68.0%排第四和第三，相对的其召回率和 mAP50:95 评分最高。尽管 Yolov8n 模型参数数量和计算量最小，但其 mAP50 评分也取得了第二的排名。整体来看，Yolov8m 在验证集上表现最好，其精确度、mAP50 和 mAP50:95 三项指标评分第一，其中精确度和 mAP50 分别高第二名 2.2%和 2.6%，并且没有太多召回率的牺牲，综合评分最优。

表 1 验证集各精度指标对比(%)

Table 1: Comparison of Accuracy Metrics for the Validation Set (%)

Yolov8	Precision	Recall	mAP50	mAP50:95	Params(M)	Flops(G)
n	70.4	59.3	68.1	37.2	3.0	8.2
s	71.9	59.1	67.8	37.3	11.1	28.7
m	74.1	60.5	70.7	38.8	25.9	79.1
l	69.3	61.0	68.0	38.4	43.6	165.4

将训练完成的四组模型均选择最优权重对 115 张测试集进行预测，并绘制出

混淆矩阵如图 9 所示。分析可知，四个模型对与 danger 类的漏检率都比较高，但相对的对于 danger 类正确检测的概率也最高，均高于 65%。Yolov8n 和 Yolov8s 正确检测 unsafe 类的概率最低，分别为 46%和 60%，Yolov8m 和 Yolov8l 正确检测 safe 类的概率最低，均为 63%。整体来看，四个模型对正样本正确识别的概率较为良好，平均精度能够达到 60% 以上。

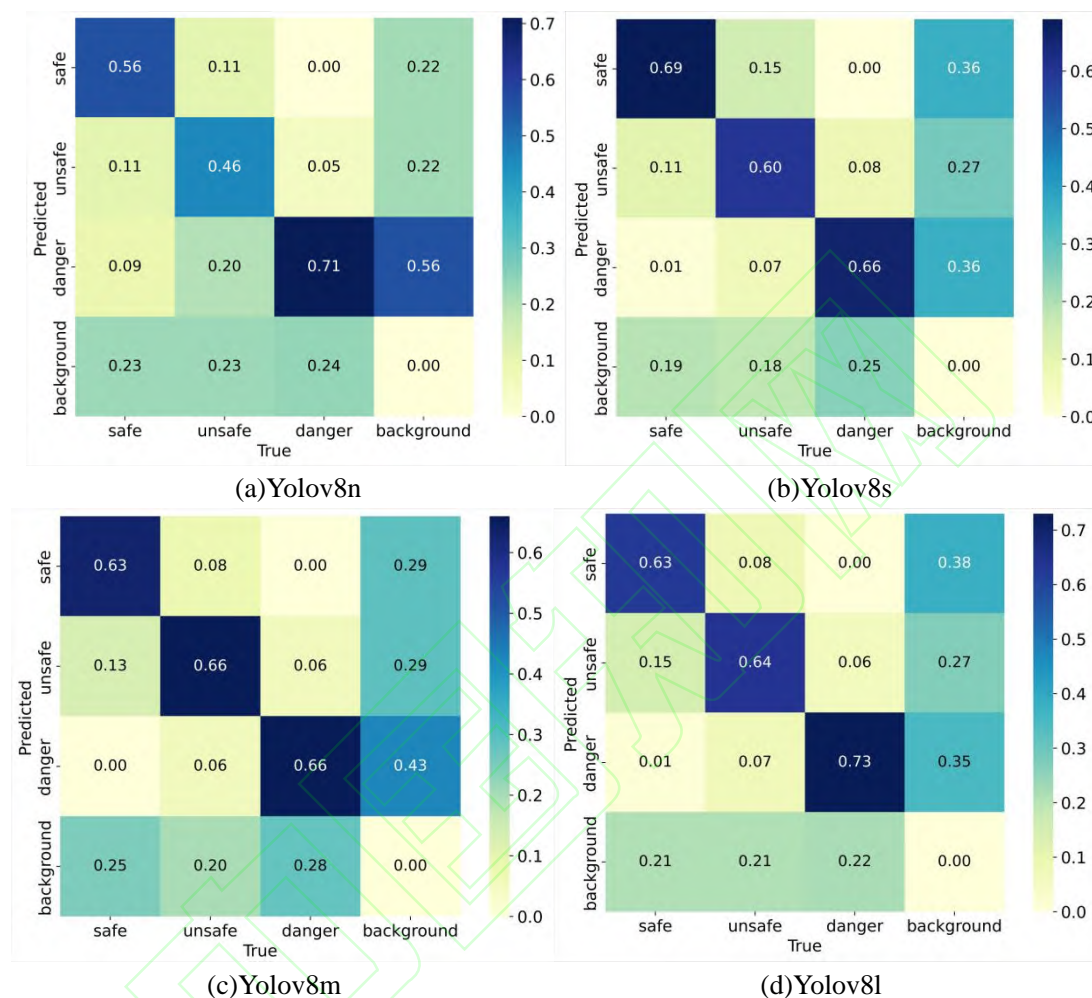


图 9 测试集混淆矩阵对比

Figure 9 Comparison of Confusion Matrices for the Test Set

将训练完成的四组模型均选择最优权重对 115 张测试集进行预测，并将三个类别的 Precision、Recall、mAP50 和 mAP50:95 四个评价指标进行计算，结果如表 2 所示。为了更加直观地表现四组模型在测试集上各指标的计算结果，对其进行热力图可视化如图 10 所示（图例表示四组模型 Precision、Recall、mAP50 和 mAP50:95 四个评价指标识别的不同类别样本准确率百分比）。分析可知，在测试集中各指标的评分并不是随着模型复杂度和计算量的增加而随之增加。对于 Precision 指标，Yolov8m 对 safe 类的评分最高，达到了 80.7%，Yolov8l 对 unsafe 类的评分最高，达到了 55.4%，Yolov8s 对 danger 类评分最高，达到了 84.8%。对于 Recall 指标，Yolov8m 对 safe 类和 unsafe 类的评分均最高，分别达到了 66.3%和 68.8%，Yolov8n 对 danger 类评分最高，达到了 68.3%。对于 mAP50 指标，Yolov8s 对 safe 类、unsafe 类和 danger 类的评分均最高，分别达到了 74.4%、61.7%

和 76.2%。对于 mAP50:95 指标，Yolov8m 对 safe 类评分最高，达到了 43.1%，Yolov8s 对于 unsafe 类和 danger 类评分均最高，分别达到了 37.7%和 36.7%。整体来看各模型对 unsafe 类的检测效果一般，整体的 mAP50 均值超过了 65%，说明模型在较为宽松的匹配标准下表现良好；整体的 mAP50:95 均值在 36%左右，说明各模型在精确定位这些类别目标方面还存在不足。

表 2 测试集各指标精度对比(%)

Table. 2 Comparison of Accuracy Metrics for the Test Set (%)

Yolov8	类别	Precision	Recall	mAP50	mAP50:95	mAP50 均值	mAP50:9 均值
n	safe	75.1	56.9	66.4	35.4	61.5	31.2
	unsafe	49.4	50.0	50.4	26.7		
	danger	69.7	68.3	67.5	31.3		
s	safe	74.4	65.7	74.4	41.9	70.8	38.7
	unsafe	53.9	65.6	61.7	37.7		
	danger	84.8	64.7	76.2	36.7		
m	safe	80.7	66.3	73.0	43.1	68.3	36.6
	unsafe	51.8	68.8	55.9	30.6		
	danger	83.2	66.0	75.9	36.2		
l	safe	77.1	62.3	72.1	41.1	68.9	36.7
	unsafe	55.4	62.5	58.3	32.5		
	danger	84.4	67.0+	76.2	36.4		

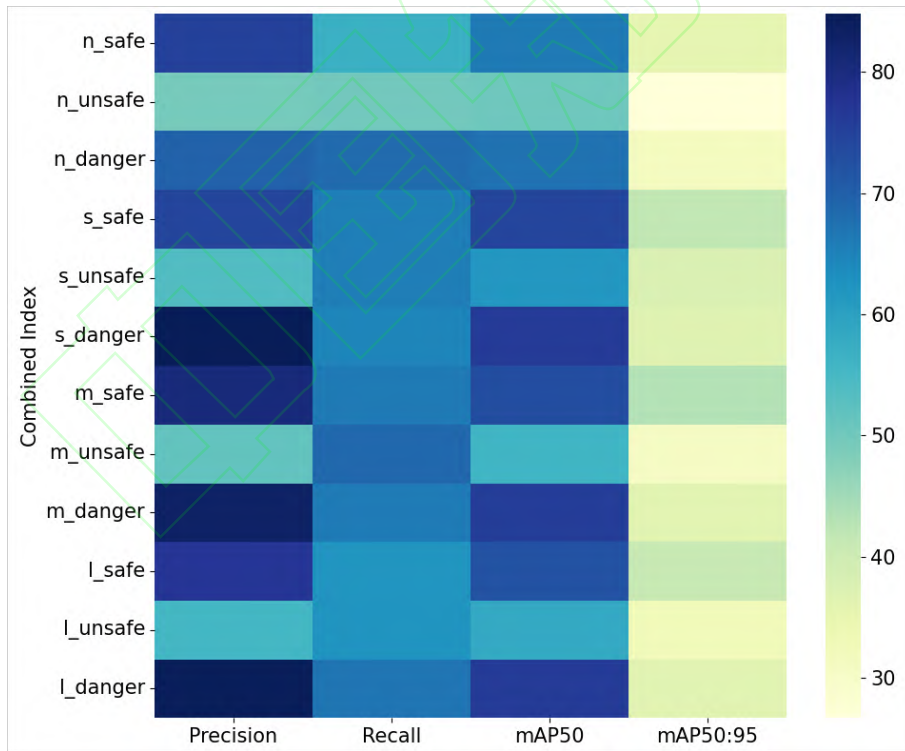


图 10 测试集各精度指标热力图

Figure 10 Heatmap of Accuracy Metrics for the Test Set

为了更加直观地展现模型的检测效果，选取几组具有代表性的不同场景的汽车淹没图片进行预测。如图 11，该图是一张多尺度混合场景下的汽车淹没状况图。其中，多尺度体现在淹没状况多尺度，目视解译可以明显看出近距离的车辆受淹程度较低，远距离的车辆受淹状况较为严重，说明距离拍摄地近处积水深度

低，远处积水深度高；车辆形态多尺度，近距离车辆拍摄不完整，远距离车辆拍摄完整；车辆大小多尺度，车辆由于拍摄距离呈现出不同的大小。Yolov8n 和 Yolov8l 对图片左下角区域不完整的车辆正确识别成了安全类，但整体置信度 Yolov8l 更高，而且 Yolov8n 将左上角远处的车辆误分为了 danger 类，实际为 unsafe 类。Yolov8s 和 Yolov8m 没有检测出左下角区域不完整的车辆，对于其余状态的车辆检测效果类似，并且 Yolov8s 检测出了更远出的小目标车辆。综合分析，在淹没状态、车辆大小以及车辆完整度等多尺度场景下，Yolov8 的四组模型能够比较准确地检测出不同车辆的淹没状态进而体现出积水深度。



图 11 多尺度混合场景下检测效果对比

Figure 11 Comparison of Detection Results in Multi-scale Mixed Scenarios

如图 12，该图是一张在俯视视角场景下拍摄的图像，并且拍摄相机高度与城市中的监控视频高度类似。该图片中发生了很严重的城市内涝，所有汽车都被积水淹没了引擎排气孔之上，目视解译图片中的积水等级都为 danger 类。根据检测效果可见 Yolov8 四组模型都准确地识别出了积水等级，其中 Yolov8n、Yolov8m 和 Yolov8l 对图片左上角区域有误检情况，但置信度都低于 50%，Yolov8s 对图片右侧小部分物体有误检验情况。整体来看四组模型对于俯视视角下的车辆

淹没等级均能做出较好地识别并且置信度高,说明该模型对于用于城市监控视频来检测城市内涝深度情况有一定的应用前景。

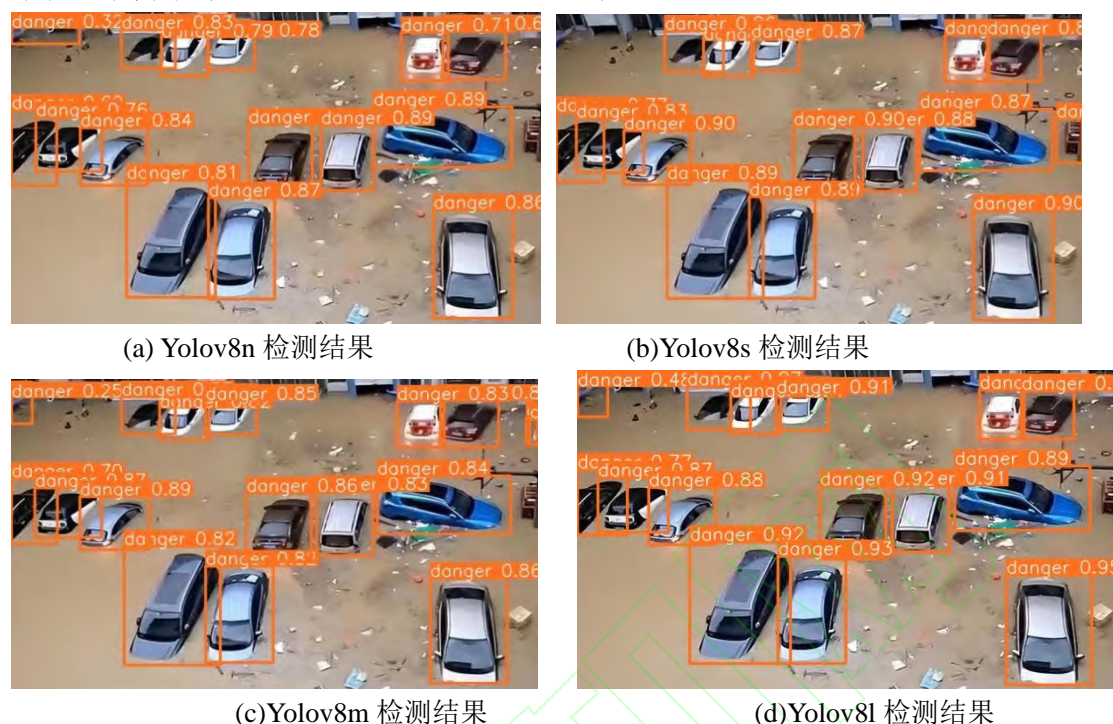
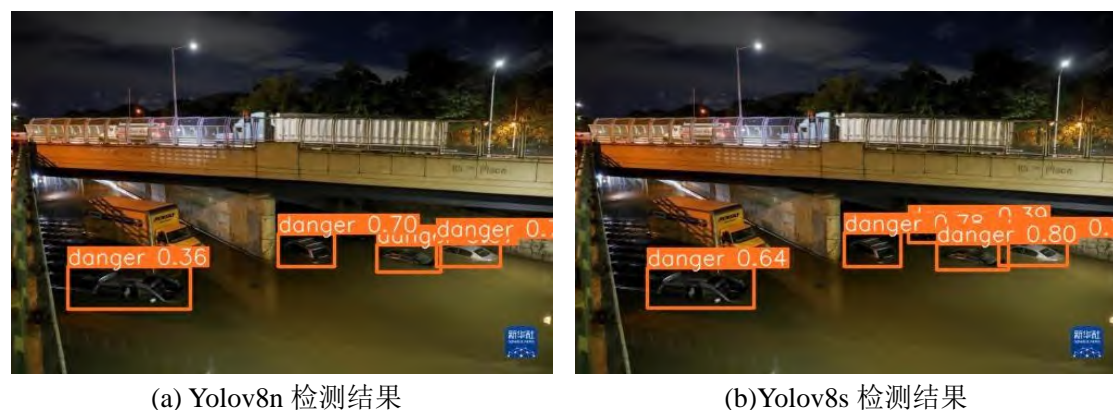


图 12 俯视视角场景下检测效果对比

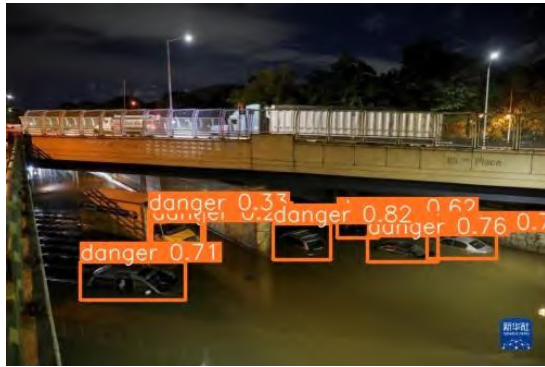
Figure 12 Comparison of Detection Results in Overhead View Scenarios

如图 13, 该图是一张在夜间场景下拍摄的图像。由于环境条件、相机质量和图像处理技术等各种因素, 评估对象检测模型在不同场景下的鲁棒性是很重要的, 因为它会影响它们在实际应用中的性能。城市内涝发生在晚上或伴随暴雨天气的概率较高, 致灾风险升高, 图像质量会因为环境条件影响而下降。图中为夜间场景的内涝场景, 图像本身较为模糊加上夜间灯光、桥梁遮挡阴影的影响, 部分淹没车辆的轮廓目视较难识别。图像左侧卡车有大半部分被积水淹没, 导致其前半部分的特征和小型车辆类似, Yolov8m 和 Yolov8l 将其识别成了小型车辆, 但置信度都偏低。图像右侧桥梁下有四辆车, 其中一辆被阴影遮挡人眼几乎识别不出车辆轮廓, Yolov8n 和 Yolov8l 未检测出该车辆, Yolov8s 和 Yolov8m 正确检测出了该车辆的淹没等级。

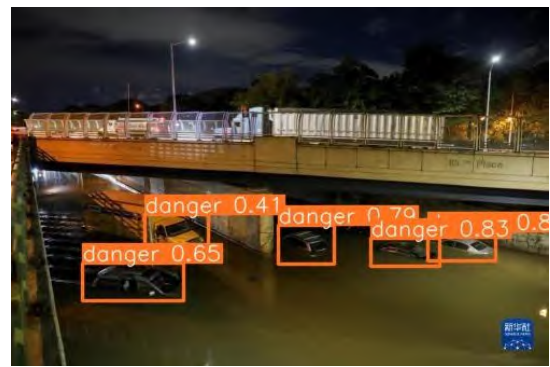


(a) Yolov8n 检测结果

(b) Yolov8s 检测结果



(c)Yolov8m 检测结果



(d)Yolov8l 检测结果

图 13 夜间场景下检测效果对比

Figure 13 Comparison of Detection Results in Nighttime Scenarios

4 结论与讨论

我们利用社会公众拍摄的包含涉水小型车辆的图像构建了训练样本,采用车辆淹没深度作为参照标准,将内涝深度划分为安全(0-20cm)、不安全(20-60cm)和危险(>60cm)三个等级,采用精确度(Precision)和召回率(Recall)、平均精度均值(mAP),平均精度均值精度(AP)作为识别精度评价指标,对 Yolov8 的 Yolov8l, Yolov8m, Yolov8n, Yolov8s 四种模型进行积水深度识别性能做了比较分析,结论与讨论如下:

(1) 对于训练集, Yolov8 四种不同大小网络算法随着训练次数的增加,最终训练轮次达到 40 次左右时都能趋于收敛状态,并且随着迭代次数的上升四项指标都是呈现出振荡上升的状态,到训练结束时四个模型的 Precision、Recall、和 mAP50 值都维持在 60%-70% 左右, mAP50:95 的值都维持在 35% 左右。

(2) 在测试集中,对于 Precision 指标, Yolov8m 对 safe 类的评分最高,达到了 80.7%, Yolov8l 对 unsafe 类的评分最高,达到了 55.4%, Yolov8s 对 danger 类评分最高,达到了 84.8%。对于 Recall 指标, Yolov8m 对 safe 类和 unsafe 类的评分均最高,分别达到了 66.3%和 68.8%, Yolov8n 对 danger 类评分最高,达到了 68.3%。对于 mAP50 指标, Yolov8s 对 safe 类、unsafe 类和 danger 类的评分均最高,分别达到了 74.4%、61.7%和 76.2%。对于 mAP50:95 指标, Yolov8m 对 safe 类评分最高,达到了 43.1%, Yolov8s 对于 unsafe 类和 danger 类评分均最高,分别达到了 37.7%和 36.7%

(3) 需要说明的是,模型的网络深度、网络宽度、特征提取能力也对模型性能有重要影响,随着网络宽度和深度的增加,模型参数量和计算量都在大幅度增加,但并不是所有的评价指标都随着网络宽度和深度的增加而增加。其中,虽然 Yolov8l 模型最为复杂,但其精确度和 mAP50 评分都不具优势,分别以 69.3%和 68.0%排第四和第三,相对的其召回率和 mAP50:95 评分最高。尽管 Yolov8n 模型参数量和计算量最小,但其 mAP50 评分也取得了第二的排名。整体来看, Yolov8m 在验证集上表现最好,其精确度、mAP50 和 mAP50:95 三项指标评分第一,其中精确度和 mAP50 分别高第二名 2.2%和 2.6%,并且没有太多召回率的牺牲,综合评分最优。

(4) 四种模型对于 unsafe 类的检测精度均低于其他两个等级的检测精度,一方面是由于样本数量不均衡(unsafe 类样本分别是 safe 类样本的 68%,是 danger 类样本的 59%),另一方面也是由于模型检测性能还存在一定局限性,如处于行动状态的车辆容易将地面的积水溅起水花,导致模型误判积水等级;如积水正好淹没至轮胎一半附近时 unsafe 类容易和 safe 类混淆,当积水正好淹没至引擎排气口附近时容易与 danger 类混淆,因此正确检测 unsafe 类的概率最低。对于这一问题,在未来的研究中,一方面对于样本的处理研究具有针对性的数据增强技术,提高模型对 unsafe 类样本的识别能力,另一方面在模型网络结构中

引入注意力机制模块，增强模型对关键特征的学习能力。

参考文献

- [1] 于朝阳.基于物联网的城市内涝在线监测与精细化治理系统建设思路探索[J].中国建设信息化,2023,(24):63-67.
Yu Chaoyang. Exploration of the Construction Ideas for Urban Flood Online Monitoring and Fine-Grained Governance System Based on the Internet of Things [J]. China Construction Informatization, 2023, (24): 63-67.
- [2] Huang H, Wang X, Liu L. A review on urban pluvial floods: Characteristics, mechanisms, data, and research methods [J]. Science of the Total Environment, 2020, 732: 1048-1063.
- [3] 罗诗琦, 罗斌, 苏鑫, 张婧, 刘军. 利用模拟数据进行遥感图像目标检测模型训练[J]. 武汉大学学报 (信息科学版). DOI: 10.13203/j.whugis20230149
LUO Shi-qi, LUO Bin, SU Xin, ZHANG Jing, LIU Jun. Leveraging Synthetic Data for Object Detection in Remote Sensing Images[J]. Geomatics and Information Science of Wuhan University. DOI: 10.13203/j.whugis20230149
- [4] 徐美,刘舒,孙杨,等. 利用洪涝模型进行城市内涝风险快速识别与预警 [J]. 武汉大学学报(信息科学版), 2020, 45 (8): 1185-1194.
Xu Mei, Liu Shu, Sun Yang, et al. Rapid Identification and Early Warning of Urban Flood Risk Using Flood Models [J]. Journal of Wuhan University (Information Science Edition), 2020, 45(8): 1185-1194.
- [5] 朱晓寒, 李向阳, 刘昭阁. 大数据环境下城市内涝灾害的情景维度模型构建[J]. 武汉大学学报 (信息科学版), 2020, 45(11): 1818-1828.
ZHU Xiaohan, LI Xiangyang, LIU Zhaoge. Constructing Scenario Dimension Model of City Waterlogging Under Big Data Environment[J]. Geomatics and Information Science of Wuhan University, 2020, 45(11): 1818-1828.
- [6] 薛丰昌,吕鑫,陈笑娟. 基于改进 DeeplabV3+的城市内涝监测方法 [J]. 测绘科学, 2023, 48 (10): 216-224.
Xue Fengchang, Lü Xin, Chen Xiaojuan. An Improved DeeplabV3+ Method for Urban Flood Monitoring [J]. Surveying and Mapping Science, 2023, 48(10): 216-224.
- [7] 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+61.
Xue Fengchang, Song Xiaoyi, Tang Baxing, et al. Urban Flood Monitoring and Early Warning Based on Video Surveillance [J]. Surveying and Mapping Science, 2018, 43(8): 50-55+61.
- [8] Swathi, Y., Challa, M. YOLOv8: Advancements and Innovations in Object Detection. In: Senjyu, T., So-In, C., Joshi, A. (eds) Smart Trends in Computing and Communications. SmartCom 2024. Lecture Notes in Networks and Systems, vol 946. Springer, Singapore.
- [9] Sohan, M., Sai Ram, T., Rami Reddy, C.V. A Review on YOLOv8 and Its Advancements. In: Jacob, I.J., Piramuthu, S., Falkowski-Gilski, P. (eds) Data Intelligence and Cognitive Informatics. ICDICI 2023. Algorithms for Intelligent Systems. Springer, Singapore.
- [10] Terven, J. R., & Cordova-Esparaza, D. M. (2024). YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness. IEEE Conference Publication.
- [11] 北京发布城市积水内涝风险地图四个风险等级辅助市民汛期出行[R/OL]. <https://www.163.com/dy/article/HBMMGDHU0514R9NP.html>, 2022-7-7.
Beijing Releases Urban Flooding and Inundation Risk Map with Four Risk Levels to Assist Citizens During the Flood Season [R/OL]. <https://www.163.com/dy/article/HBMMGDHU0514R9NP.html>, 2022-7-7.
- [12] Ultralytic: Ultralytics YOLOv8 [EB/OL]. [2023-326]. <https://github.com/ultralytics/ultralytics>.

- [13] 刘瑞锦,何章鸣.基于 YOLOv8 的卫星遥感图像快速目标检测方法[J].空间控制技术与应用,2023,49(5):89-97.
- [14] 张潇艺,杨胜龙.基于 YOLOv8 的民用船舶影像分类方法研究[J].工业控制计算机,2024,37(4):72-73+76.
Zhang Xiaoyi, Yang Shengliang. Research on Civil Ship Image Classification Method Based on YOLOv8 [J]. Industrial Control Computer, 2024, 37(4): 72-73+76.
- [15] 康睿,徐海祥,冯辉. 基于改进 YOLOv8 的航海雷达图像目标检测算法 [J]. 大连海事大学学报, 2024, 50 (3): 23-30.
Kang Rui, Xu Haixiang, Feng Hui. Target Detection Algorithm for Marine Radar Images Based on Improved YOLOv8[J]. *Journal of Dalian Maritime University*, 2024, 50(3): 23-30.
- [16] Song G, Liu Y, Wang X. Revisiting the sibling head in object detector[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 11563-11572.
- [17] 胡惠娟,秦一锋,徐鹤,等. 面向无人机航拍图像的 YOLOv8 目标检测改进算法 [J/OL]. 计算机科学, 1-18[2024-10-02].
Hu Huijuan, Qin Yifeng, Xu He, et al. An Improved YOLOv8 Object Detection Algorithm for UAV Aerial Imagery [J/OL]. *Computer Science*, 1-18 [2024-10-02].
- [18] 程换新,矫立浩,骆晓玲,等.改进 YOLOv8 的遥感图像检测算法[J/OL]. 无线电工程,2024,05(7):1-9.
Cheng Huanxin, Jiao Lihao, Luo Xiaoling, et al. An Improved YOLOv8 Detection Algorithm for Remote Sensing Images [J/OL]. *Radio Engineering*, 2024, 05(7): 1-9.
- [19] 潘银,邵振峰,程涛,等. 利用深度学习模型进行城市内涝影响分析 [J]. 武汉大学学报(信息科学版), 2019, 44 (1): 132-138.
Pan Yin, Shao Zhenfeng, Cheng Tao, et al. Urban Flood Impact Analysis Using Deep Learning Models [J]. *Journal of Wuhan University (Information Science Edition)*, 2019, 44(1): 132-138.
- [20] 韦柳梅,罗雪梅,康健. 改进 YOLOv8 的航拍图像小目标检测算法 [J/OL]. 计算机工程与科学, 1-13[2024-10-02].
Wei Liumei, Luo Xuemei, Kang Jian. An Improved YOLOv8 Algorithm for Small Target Detection in Aerial Images [J/OL]. *Computer Engineering and Science*, 1-13 [2024-10-02].

网络首发:

标题: 利用社会公众拍摄图像构建城市内涝深度等级评估模型

作者: 薛丰昌, 陈笑娟, 吕鑫, 成沿年

收稿日期: 2024-11-10

DOI:10.13203/j.whugis20240290

引用格式:

薛丰昌, 陈笑娟, 吕鑫, 等. 利用社会公众拍摄图像构建城市内涝深度等级评估模型[J]. 武汉大学学报(信息科学版), 2024, DOI:10.13203/J.whugis20240290 (XUE Fengchang, CHEN Xiaojuan, LÜ Xin, et al. Constructing an Urban Waterlogging Depth Assessment Model Using Public Video Information[J]. *Geomatics and Information Science of Wuhan University*, 2024, DOI:10.13203/J.whugis20240290)

网络首发文章内容和格式与正式出版会有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

面向道路积水动态可视化的多源数据集成方法

史潇, 徐家鹏, 杜毅贤, 沈婕, 臧培岳

武汉大学学报(信息科学版), 2022, 47(5): 693-699.

<http://ch.whu.edu.cn/article/doi/10.13203/j.whugis20190350>

利用洪涝模型进行城市内涝风险快速识别与预警

徐美, 刘舒, 孙杨, 姚永慧

武汉大学学报(信息科学版), 2020, 45(8): 1185-1194.

<http://ch.whu.edu.cn/article/doi/10.13203/j.whugis20200122>

大数据环境下城市内涝灾害的情景维度模型构建

朱晓寒, 李向阳, 刘昭阁

武汉大学学报(信息科学版), 2020, 45(11): 1818-1828.

<http://ch.whu.edu.cn/article/doi/10.13203/j.whugis20190225>

