

· 创新与探索 ·

基于改进 MobileNetV3-SSD 的河道排污口目标检测研究

徐伟¹, 王建华^{1*}, 郑翔¹, 王昱博¹, 冯居¹, 姜洪岩¹, 田雨², 钱建华², 张欣尧²

(1. 上海海事大学, 航运技术与控制工程交通行业重点实验室, 上海 201306;

2. 上海遨拓深水装备技术开发有限公司, 上海 201306)

摘要:为实现对水系入河排污口有效、准确的自动检测,提出一种基于改进 MobileNetV3-SSD 的深度学习模型。在 MobileNetV3-SSD 模型的基础上,使用 K -means 聚类算法和遗传算法,对先验框的宽高比进行调整,使得预测框更好地匹配真实框。引入多尺度特征融合模块,提高模型对小排污口的检测能力。引入改进的 CBAM 注意力模块,减少模型在排污口检测时计算的参数数量。使用可变形卷积替代普通卷积,自适应地捕获不同排污口的形态与尺度信息,提升模型的特征提取能力。实验结果表明,改进后 MobileNetV3-SSD 模型的平均精度为 89.36%, F1 分数为 91.88%, 较改进前分别提升 4.83% 和 5.03%。

关键词: MobileNetV3-SSD 模型; 图像处理; 深度学习; 河道排污口

中图分类号: X52; TP391.41

文献标志码: B

文章编号: 1006-2009(2023)05-0063-05

Target Detection at Sewage Outlets Based on Improved MobileNetV3-SSD

XU Wei¹, WANG Jian-hua^{1*}, ZHENG Xiang¹, WANG Yu-bo¹, FENG Ju¹, JIANG Hong-yan¹,

TIAN Yu², QIAN Jian-hua², ZHANG Xin-yao²

(1. Key Laboratory of Marine Technology and Control Engineering, Shanghai Maritime University,

Shanghai 201306, China; 2. Shanghai Auto Subsea Vehicles Inc., Shanghai, 201306, China)

Abstract: In order to achieve effective and accurate automatic detection at sewage outlets of water system, a deep learning model based on improved MobileNetV3-SSD was proposed. On the basis of MobileNetV3-SSD model, K -means clustering algorithm and genetic algorithm were used to adjust the aspect ratio of the prior bounding box and make the prediction box match the ground truth box better. The introduction of a multi-scale feature fusion module improved the detection ability of the model at small sewage outlets. The introduction of an improved CBAM attention module reduced the number of parameters calculated by the model during the detection at sewage outlet. Deformable convolution was used instead of ordinary convolution to adaptively capture the shape and size of different sewage outlets, and improve the feature extraction ability of the model. The experimental results showed that the average accuracy of the improved MobileNetV3-SSD model was 89.36%, F1 score was 91.88%, which was 4.83% and 5.03% higher than before improvement, respectively.

Key words: MobileNetV3-SSD; Image processing; Deep learning; Sewage outlet

入河排污口是水系中主要的污染源,对污染源构建完善的检测体系,是水资源环境保护的核心工作。然而,我国排污口面广量大、形状多样、大小不一,这些因素加大了入河排污口排查整治的难度^[1],检测技术力量的增强对于污染防治工作至关重要^[2]。

目前,国内对入河排污口的排查主要有以下几

种方式:人工实地调查、卫星遥感检测和无人机排查。人工实地调查是我国早期排污口的主要排查方式,费时费力,排查覆盖面小^[3]。随着卫星遥感

收稿日期:2022-09-27;修订日期:2023-07-17

基金项目:国家自然科学基金资助项目(62176150)

作者简介:徐伟(1996—),男,江苏南通人,在读研究生,研究方向为图像处理。

* 通信作者:王建华 E-mail: jian-hua.wang@163.com

技术和无人机航拍技术的发展,可以得到高分辨率的排污口卫星影像信息和无人机航摄图像信息^[4]。由于视角原因,卫星检测和无人机检测都难以发现存在于草下、桥下等隐蔽性强的排污口,且对于一些小型 PVC 塑料排污口也可能存在漏检的情况^[5]。今以汾湖水域内的排污口为研究对象,建立较为全面的排污口图像数据集,提出改进的 MobileNetV3-SSD 模型,引入多尺度特征融合模块和改进的 CBAM 注意力模块,提升模型对于小型、形状不规则、被部分遮挡排污口的检测能力。训练好的模型可布置于无人船上,用于河道排污口的检测,为排口整治、水体治理、水环境提升提供帮助。

1 网络结构

1.1 基础网络

为实现无人船搭载相机对排污口的实时检测,需要建立轻量化的检测模型。MobileNet 系列模型计算量少、检测速度快,适宜在移动端部署。SSD 模型在检测精度上表现优异。因此,用 MobileNetV3 的主干网络替换 SSD 模型的主干网络,构建了 MobileNetV3-SSD 目标检测模型,并在该模型的基础上加以改进,以提升其对河道排污口的检测精度。

1.2 改进网络

MobileNetV3-SSD 模型的先验框宽高比是依靠经验手动设定,虽适用于多物体的情况,但主要针对排污口单一目标进行检测时,默认先验框的宽高比适用性不强。此外,在特征提取方面,MobileNetV3-SSD 模型对不同尺度的特征图直接进行目标识别和定位,不同层次之间的信息未能进行融合,从而造成提取特征信息的缺失,故提出对应的

改进办法。先利用改进的 K-means 聚类算法,依据排污口的特性进行自适应学习,对先验框的宽高比进行重新设定。再对 MobileNetV3-SSD 模型的结构进行修改,在主干网络中,先通过特征融合模块将网络不同层次的信息进行融合,再将经过特征融合后的特征图输入改进的 CBAM 注意力模块中,增强网络的特征提取能力。改进后的网络结构见图 1。

1.2.1 基于改进 K-means 聚类的先验框设定

先验框的作用是对预测框的回归位置进行优化,使预测框更好地匹配真实框。先验框的具体大小由尺度和宽高比决定,其宽高比虽可根据经验选取 $a_r = \left\{ 1, 2, 3, \frac{1}{2}, \frac{1}{3} \right\}$ 这 5 个比例,但该设置方法对上述排污口检测任务缺乏特异性。为得到适用于排污口检测的宽高比设置,先用 K-means 聚类算法对样本宽高比进行聚类,再用遗传算法,在 K-means 聚类的结果上进行变异,提高聚类结果的泛化能力。K-means 聚类得到的 6 个先验框的大小分别为 $[13, 10][16, 17][22, 24][33, 31][57, 43][92, 82]$, 适应度为 78.2%。使用遗传算法,在聚类结果上进行变异后的先验框大小分别为 $[14, 12][18, 22][28, 28][37, 35][58, 48][82, 79]$, 适应度为 79.3%,相较于遗传变异之前,适应度提升了 1.1%,能让预测框更好地匹配真实框。

1.2.2 多尺度特征融合模块

由于初始的 MobileNetV3-SSD 模型直接检测多尺度特征信息,缺少不同层次的信息融合,导致特征之间的相关性丢失,不利于待检测目标细节特征的提取。分析排污口图像数据发现,部分排污口

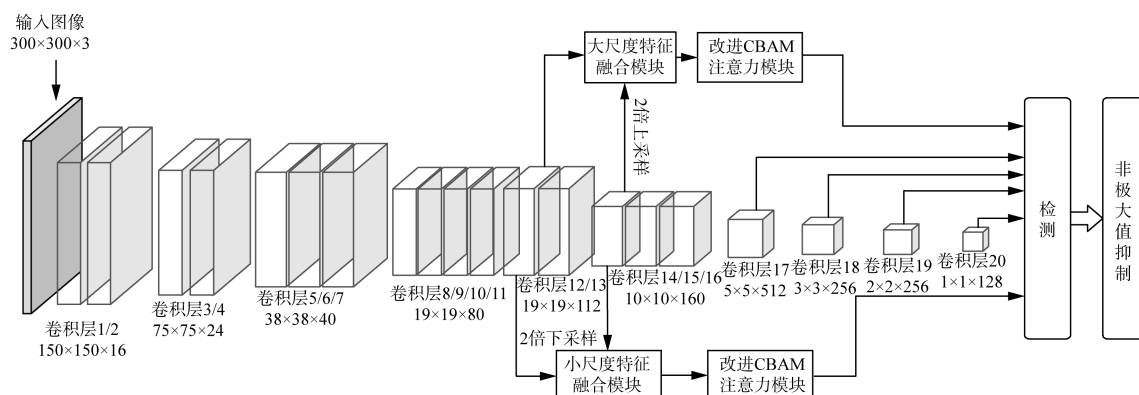


图 1 改进 MobileNetV3-SSD 网络结构

Fig. 1 Improved MobileNetV3-SSD network structure

为小型排污口,其细节信息需要在浅层、分辨率大的特征图上得到。虽然浅层特征包含较多细节信息,但所提取的语义信息可能仍然不够充分^[6]。

故提出两种特征融合模块使卷积层 12 与卷积层 14 的特征层建立联系,即大尺度特征融合模块和小尺度特征融合模块,分别见图 2(a)(b)。

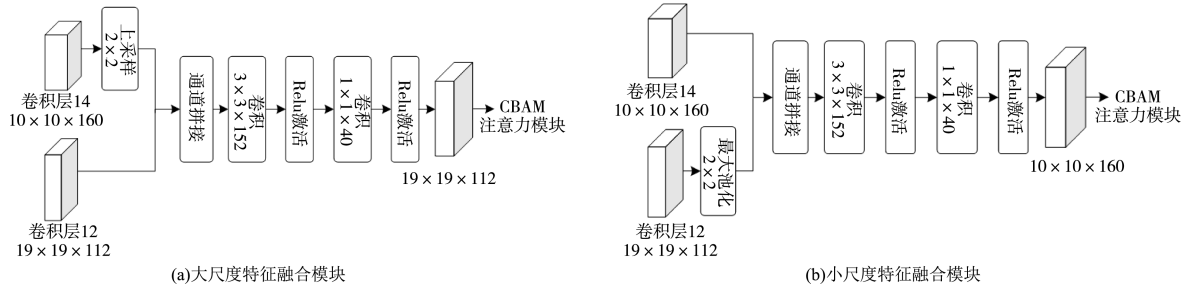


图 2 多尺度特征融合模块
Fig. 2 Multi-scale feature fusion module

由图 2(a)可见,对卷积层 14 先进行 2 倍的上采样操作后,使其特征图的大小与卷积层 12 保持相同,再与卷积层 12 进行通道拼接,进行特征融合,并利用 3×3 卷积来消除融合后特征的混叠效应,接着用 1×1 卷积进行通道降维操作,得到 $19 \times 19 \times 112$ 的特征层。由图 2(b)可见,在特征融合的过程中,将卷积层 12 进行步长为 2、大小为 2×2 的最大池化操作,降低特征图的维度,得到与卷积层 14 相同尺寸的 10×10 大小特征图,与卷积层 14 进行通道拼接,通道拼接后的操作与图 2(a)类似,最终得到 $10 \times 10 \times 160$ 的特征层。

1.2.3 改进 CBAM 注意力模块

为提升改进 MobileNetV3-SSD 模型对排污口细节特征的提取能力,使其能更好地检测不规则、被遮挡的排污口,在模型中引入改进的 CBAM 注意力模块。

虽然 CBAM 模块在深度学习模型中被广泛使用,但其在通道注意力模块中使用全连接层对通道特征进行映射会产生冗余计算;在空间注意力模块中,通过使用一个 7×7 的卷积增大感受野,使得输出特征层的每个单元对应输入层的区域面积变大,获得更多的信息。然而,普通卷积其几何形变建模能力有限,对于一些不规则目标的表征能力不强^[7]。而试验所需检测的排污口形态多样,且有被遮挡的情况。针对上述问题,提出改进 CBAM 模块,见图 3。

在通道注意力模块中,使用 k 个 1×1 卷积取代原用于特征映射的两个全连接层,进行通道特征聚合。由于卷积操作的参数共享性质,引入 1×1 卷

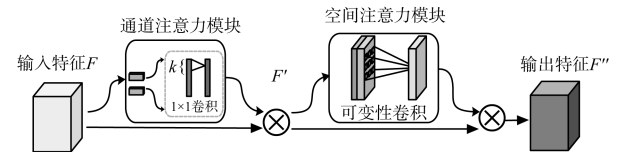


图 3 改进 CBAM 模块结构
Fig. 3 Structural of improved CBAM module

积使得所需计算的参数数量降低到常数级。具体过程可表示为:

$$F' = (\delta(f_{1D}^k(P_{avg}(F))) + (\delta(f_{1D}^k(P_{max}(F)))))) \otimes F \quad (1)$$

式中: F 表示输入特征图; F' 表示经过通道注意力调整后的特征图; δ 为 Sigmoid 激活函数, $\delta(x) = \frac{1}{1 + e^{-x}}$, 激活函数对通道特征进行非线性组合,增强模型的非线性表达能力; f_{1D}^k 为 k 个 1×1 卷积; P_{avg} 为全局平均化,在每个通道的特征图上计算特征图的平均值作为输出; P_{max} 为全局最大池化,用于在每个通道的特征图上找到最大值作为输出^[8]。

对于输入特征图 F ,先分别进行全局平均池化 P_{avg} 和全局最大池化 P_{max} ,然后分别经过两个 1×1 卷积操作并通过 Sigmoid 激活函数生成两个不同的通道注意力特征图,将两个特征图相加形成注意力权重,最后将其与输入特征图 F 逐像素相乘 (\otimes) 得到 F' 。

在空间注意力模块中,提出使用具有可变形卷积来替代 7×7 的普通卷积。首先,对输入的排污口图像,利用传统的卷积核提取特征图;其次,在传

统卷积核中每一个卷积核单元增加自适应学习的水平方向和垂直方向的位置偏移量,通过对偏移量的学习,调整卷积核的方向;最后,使得卷积核能够根据排污口的形状适应性采样,从而提升对排污口形态特征的提取能力^[9]。

2 结果与讨论

2.1 数据来源和特点

所建数据集的图像数据由实验室的无人船“海翔”号搭载工业相机和人工使用智能手机,在江苏省苏州市汾湖水域不同地点、不同时段进行采集。水域内河网密布,湖荡众多,河道规模较小,水系错综复杂。排污口特征丰富多样,外观上以圆形和方形为主,少部分排污口外端为不规则形状;布置形式上既有露出管道的明口,也有不露管道的暗口;材质方面包括微小 PVC 塑料排口和大型水泥排口。

2.2 评价指标

针对排污口的检测问题,采用的评价指标为精确率(Precision)、召回率(Recall)、平均精确率(AP)、F1 分数及检测速度(FPS)。精确率为被正确分类的排污口样本占全部预测样本的比例;召回率为被检测到的排污口样本占总样本的比例;F1 分数为精确率和召回率的调和平均数,使用 F1 分数来兼顾这两个评价指标;平均精确率为 P-R (Precision-Recall) 曲线下方面积的积分值^[10]。

2.3 模型训练

将训练集在初始模型和改进模型中分别迭代 100 次,并记录每次迭代的训练损失。随着训练次数的增加,改进后模型的训练损失由最高 15% 逐渐减少并稳定在 2% 左右。改进后模型的损失函数曲线在模型收敛后的波动幅度与改进前相比更加平稳,表明对 MobileNetV3-SSD 的改进有效,改进后的模型具有更好的稳定性。

2.4 消融实验

通过消融实验,验证各改进部分对模型检测性能的影响。将 MobileNetV3-SSD、MobileNetV3-SSD+多尺度特征融合模块、MobileNetV3-SSD+CBAM 注意力模块和 MobileNetV3-SSD+改进 CBAM 注意力模块等模型做对比,分析各个改进部分对模型性能的作用,结果见表 1。由表 1 可知,引入多尺度特征融合模块和改进的 CBAM 注意力模块后,改进的 MobileNetV3-SSD 模型在所建立的排污口数据集上的精确率、召回率、AP 值、F1 分数这 4 个指

表 1 在排污口数据集上的消融实验结果

Table 1 Results of ablation experiments on sewage outlet dataset

模型	精确率/%	召回率/%	F1 分数/%	AP 值/%	检测速度/(帧·s ⁻¹)
MobileNetV3-SSD	81.13	93.45	86.85	84.53	35.56
多尺度特征融合	83.28	95.02	88.76	86.45	33.64
CBAM 注意力模块	84.65	94.73	89.41	86.89	30.69
改进注意力模块	86.87	94.57	90.55	88.26	32.24
改进 MobileNetV3-SSD	88.94	95.04	91.88	89.36	29.53

标,对比改进前,分别提高了 7.81%、1.59%、4.83%、5.03%。改进模型中的各个模块给各个评价指标所反映的检测效果都带来了一定提升。其中,相较于多尺度特征融合模块,改进后 CBAM 注意力模块检测效果更好,精确率提升了 3.52%,AP 值提升了 2.36%。由于使用可变形卷积,改进后的 CBAM 注意力模块能够更好地捕捉排污口的形态特征,从而更加准确地检测不同形状的排污口。然而,引入多尺度特征融合模块和注意力模块,使得模型的检测速度下降了 6 帧/s。考虑到无人船在进行排污口检测时一般为低速航行状态,改进后的 MobileNetV3-SSD 模型检测速度为 30 帧/s,能够满足实际需求。

对改进前后的 MobileNetV3-SSD 模型进行测试,发现改进后的模型能够较好地分辨真实排污口和排污口阴影的区别,这是由于引入了多层特征融合模块和改进后的 CBAM 注意力模块,对于网络浅层的语义信息提取更加丰富。此外,改进后模型能较好地识别小型排污口,一定程度解决了面积较小、不易察觉的小排污口的检测问题。同时,当排污口被水草、杂物等遮挡时,改进后的模型仍能对排污口进行较好的检测。综上所述,改进后的模型对小型、形状各异、部分遮挡的排污口均能实现较为精准的检测。

2.5 对比实验

为验证改进后 MobileNetV3-SSD 模型与其他目标检测模型在排污口检测效果上的区别,进行对比实验,结果见表 2。由表 2 可知,虽然改进 MobileNetV3-SSD 模型与二阶段检测模型 Faster RCNN 在 F1 分数和 AP 值这两个指标上相近,但在检测速度上有明显提升,更适用于排污口的实时检测需求。与其他几种一阶段目标检测模型相比,虽然 YOLOV3 的检测速度最快,但其 F1 分数和 AP 值

表 2 在排污口数据集上的对比实验

Table 2 Comparison experiment results on sewage outlet

模型	dataset				
	精确率/%	召回率/%	F1 分数/%	AP 值/%	检测速度/(帧·s ⁻¹)
VGG-SSD	71.82	80.69	75.99	78.96	27.39
MobileNetV2-SSD	80.50	87.64	83.91	82.73	34.73
YOLOV3	67.71	88.82	76.84	80.17	41.83
Faster RCNN	85.29	96.82	90.68	92.25	8.78
改进 MobileNetV3-SSD	88.94	95.04	91.88	89.36	29.35

较改进后 MobileNetV3-SSD 模型分别下降了 15.04% 和 9.19%。与传统 SSD 模型相比,改进后 MobileNetV3-SSD 模型由于采用轻量级主干网络,且引入多尺度特征融合模块和改进注意力模型,检测速度更快,精确率和召回率更高。综上所述,改进后 MobileNetV3-SSD 模型针对排污口的综合检测性能较其他模型具备一定优势。

3 结语

针对城市河道内的排污口具有形状不同、大小各异、部分被遮挡等特点,利用车载摄像机结合人工拍摄制作了较为全面的排污口数据集,并提出改进 MobileNetV3-SSD 模型。通过改进 K-means 聚类算法对先验框的宽高比进行调整,使得检测框的大小更适用于排污口。引进多尺度特征融合模块,提升对小型、被遮挡的排污口的检测能力。使用可变卷积取代 CBAM 注意力模块中的普通卷积,使检测模型能够提取排污口的几何形变特征,可更好

地检测形状不同的排污口。实验结果表明,改进 MobileNetV3-SSD 模型相较改进前,在精确率、召回率、F1 分数等指标上均有所提升,一定程度上降低河道排污口的错检、漏检率。同时,对比其他检测模型,改进 MobileNetV3-SSD 模型更好地实现了精度与效率的平衡,能较为准确地对城市河道排污口进行实时检测。

[参考文献]

- [1] 嵇晓燕,王姗姗,杨凯,等. 长江水环境质量监测网络运行体系初步构建[J]. 环境监测管理与技术,2022,34(5):1-5.
- [2] 张明睿,郑俊,徐力刚,等. 城市小流域面源污染输出特征及污染负荷分类核算研究[J]. 环境监测管理与技术,2021,33(4):25-29.
- [3] 孙力. 无人机航拍图像排污口目标检测研究[D]. 西安:西安电子科技大学,2020.
- [4] 陈黄蓉. 入海河口及上游河流污染水体遥感监测研究[D]. 南京:南京信息工程大学,2021.
- [5] 祁舟. 基于深度学习的无人机航拍图像排污口检测方法研究[D]. 济南:山东大学,2021.
- [6] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]// ECCV. Proceedings of the European conference on computer vision. Munich: ECCV, 2018: 3-19.
- [7] WU X W, SAHOO D, HOI S C H. Recent advances in deep learning for object detection[J]. Neurocomputing, 2020, 396: 39-64.
- [8] ZHAO Z Q, ZHENG P, XU S, et al. Object detection with deep learning: A review[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(11): 3212-3232.
- [9] 陈晓雷,卢禹冰,曹宝宁,等. 轻量化高精度双通道注意力机制模块[J]. 计算机科学与探索, 2023(4): 857-867.
- [10] 赵彦杰,崔海斌,陈振,等. 基于可变形卷积的侦察视频增强方法[J]. 探测与控制学报, 2022, 44(5): 46-52.

· 征订启事 ·

欢迎订阅 2024 年《环境监控与预警》杂志

《环境监控与预警》创刊于 2009 年,是由江苏省生态环境厅主管、江苏省环境监测中心主办、南京大学环境学院协办的期刊。被中国科技核心期刊(中国科技论文统计源期刊)、JST 日本科学技术振兴机构数据库收录。期刊面向全国公开发行人,国内统一刊号 CN32-1805/X,国际标准刊号 ISSN1674-6732。

本刊致力于传播和推广先进的环保科技成果,聚焦环境前沿科技,介绍国内外环境监测、环境预警、环境信息等领域的新技术、新成果、新发展,跟踪国家及地方的环境政策、环境标准的变化。读者对象主要是从事环境管理、环境监测、环境监察、环境信息、环境治理、环境科学研究及其他领域的环境工作者。常设栏目有:前沿评述、环境预警、监测技术、解析评价、监管新论等。

本刊为双月刊,大 16 开国际标准版,108 页,每逢单月 30 日出版。国内定价(含邮费)35 元/期,全年 210 元。

订阅方法:1. 邮局订阅:邮发代号:28-414。

2. 编辑部订阅:汇款后将回执单 E-mail 至联系人:谭艳;电话:025-69586548;邮箱:hjkyjy@163.com。(电子版回执单下载地址:http://www.hjkyjy.com)

汇款信息:

单位名称:江苏省环境监测中心 开户行:中国银行南京凤凰花园城支行 账号:530058192469