

基于YOLOv5的水面垃圾旋转目标检测模型

赵 政, 李 蓉, 黄鹏辉

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年4月22日; 录用日期: 2024年5月22日; 发布日期: 2024年5月30日

摘 要

水面垃圾污染给生态环境带来巨大威胁, 无人船抓捞为水面垃圾清理提供了一种更为高效环保的方法。水面垃圾目标检测是无人船研究的关键技术之一。传统的水平目标检测在环境复杂, 目标角度不规则的情况下检测精度较低且抓捞失误率较高。针对此问题, 本文提出一种基于改进YOLOv5的水面垃圾旋转目标检测模型。该模型优化了YOLOv5模型的检测头, 可以对任意角度的水面垃圾生成定向检测框。在此基础上, 本文进一步从三个方面改进了YOLOv5在水面垃圾检测中的性能: 提出动态平滑角度损失函数以增强角度预测能力; 通过增强浅层特征并引入注意力机制来优化骨干网络; 改进特征聚合网络, 增强特征多尺度融合效果。实验结果表明, 该方法对水面垃圾的目标检测性能优于其它深度学习方法。

关键词

水面垃圾, 旋转目标检测, YOLOv5, 动态平滑

Rotating Target Detection Model of Water Surface Garbage Based on YOLOv5

Zheng Zhao, Rong Li, Penghui Huang

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 22nd, 2024; accepted: May. 22nd, 2024; published: May. 30th, 2024

Abstract

Water surface garbage pollution poses a significant threat to the ecological environment, and unmanned boat grabbing offers a more efficient and environmentally friendly method for water surface garbage cleanup. In the field of unmanned boat research, the precise detection of water surface garbage is one of the core technologies. Traditional horizontal target detection has lower detection accuracy and a higher rate of grabbing errors in complex environments with irregular

target angles. To address this issue, this paper proposes a water surface garbage rotating target detection model based on an improved YOLOv5. The model optimizes the detection head of the YOLOv5 model to generate oriented detection boxes for water surface garbage at any angle. On this basis, the paper further improves the performance of YOLOv5 in water surface garbage detection from three aspects: proposing a dynamic smooth angle loss function to enhance angle prediction capability; optimizing the backbone network by enhancing shallow features and introducing attention mechanisms; and improving the feature aggregation network to enhance the effect of multiscale feature fusion. Experimental results show that this method outperforms other deep learning methods in the target detection performance of water surface garbage.

Keywords

Water Surface Garbage, Rotating Target Detection, YOLOv5, Dynamic Smoothing

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着全球工业化进程的加速,水污染问题日益严重,其中水面垃圾问题尤为突出。水面垃圾不仅破坏了水域环境的美观,更重要的是,它对水生生物健康和人类的生活质量构成了严重威胁。因此,有效的水面垃圾监测和清理是保护环境和管理水资源的重要手段。

随着智能交通技术的快速发展,智能驾驶的水面无人船悄然兴起[1],尤其是专门从事水面清洁的无人船,它们使得漂浮垃圾的清理工作实现全自动化、全天候操作。这些无人船还能够进入人类无法难以到达的危险区域,从而大幅提升漂浮垃圾清理的效率和工作安全性[2][3]。

在水面垃圾清理的无人船运用中,垃圾检测系统的检测精度和速度成为了评估其清洁效率的关键指标。得益于计算机视觉和深度学习技术的广泛应用,无人船配备的图像采集系统能够充分利用所捕获的视觉图像信息。因此,基于视觉图像的水面垃圾检测成为了目标最经济、最有效的检测手段[4]。文献[5]通过在 YOLOv3 主干网络后添加 SE-PPM 模块的方式加强网络提取水面目标特征的能力。文献[6]将多尺度特征与注意力机制相结合,提出一种水面小目标检测算法,解决复杂海面场景下小目标特征信息匮乏,纹理信息不清晰等问题。这些方法都采用水平检测框的方式进行目标定位。而在本文研究的水面垃圾检测场景中,无人船检测的角度通常带有俯视倾角,这就意味着观测点和目标物体不处于相同的水平面。加之无人船的航行路径并非固定,因此无人船捕捉到的图像会随着观察角度的变化而出现缩放和旋转等多种情况。若仍然采用传统的水平检测框识别目标,将会在检测过程中引入大量背景噪声,不仅增加了错拣的可能,而且在目标较为密集的情况下导致检测框之间的严重重叠,从而出现漏检的情况。

针对上述问题,本文改进 YOLOv5 模型提高了其对水面垃圾检测应用场景的适用性。主要贡献如下:

- 1) 增加角度预测网络和损失函数,以及用于角度分类的动态角度平滑算法以改进角度预测能力。
- 2) 优化主干网络中的 BottleneckCSP 模块来增强模型提取定向水面垃圾特征的能力。
- 3) 优化特征聚合网络以提高多尺度特征融合效果。

2. 水面垃圾旋转目标检测模型

如图 1 所示,水面垃圾检测方案框架包括五个部分:输入层、特征提取主干网络、特征聚合网络、

检测头和动态平滑模块。其中，主干网络主要由 Focus 模块、卷积模块和基于优化 BottleneckCSP 的 OP BottleneckCSP 模块组成。Focus 模块通过切片操作减少计算量并提升运算速度。BottleneckCSP 是一种卷积结构，在模型学习中表现出良好的性能。主干网络用于从水面垃圾图像中提取特征，并生成三种不同大小的特征图。特征聚合网络融合来自主干网络的多尺度特征，以提高水面垃圾旋转角度特征的学习能力。检测头基于多尺度特征图生成水面垃圾的类别、位置和旋转角度。最后，动态平滑模块部分加密角度标签的独热标签编码用于模型训练。

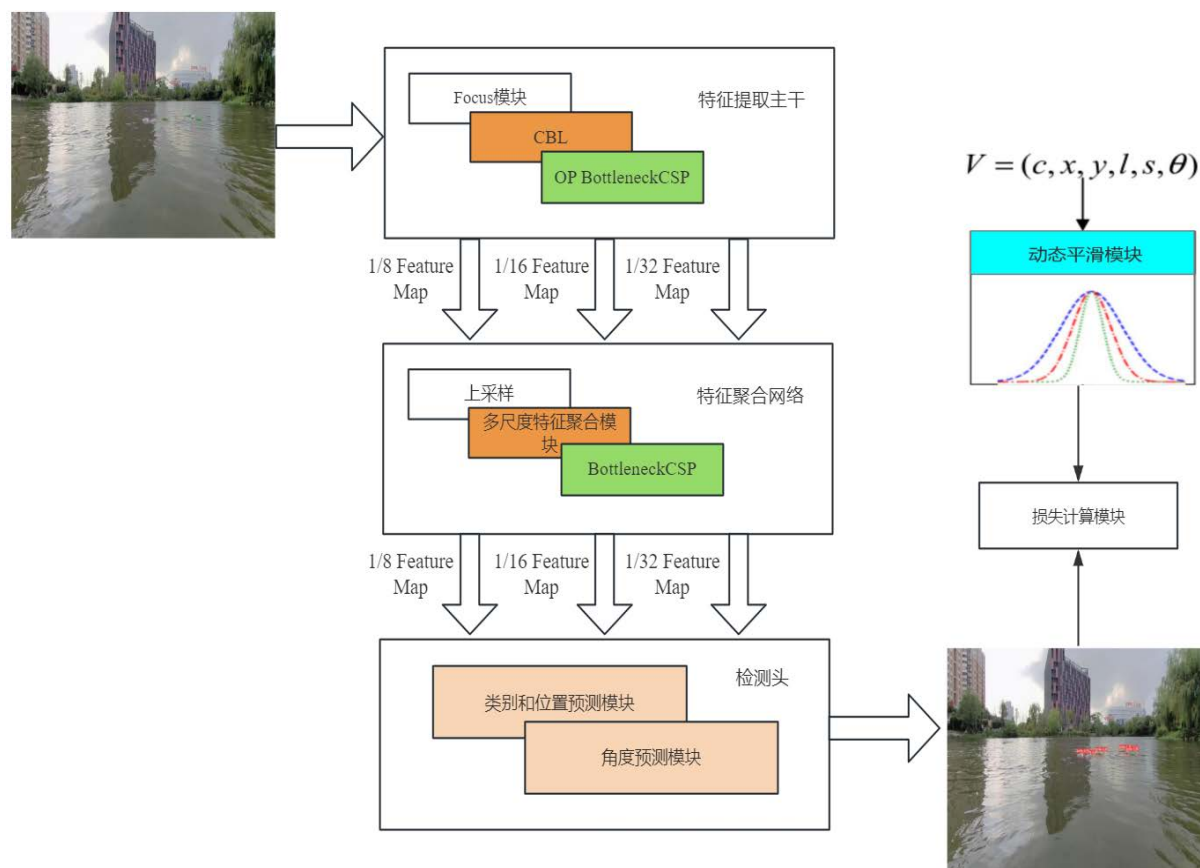


Figure 1. Overall structure of model

图 1. 模型总体框架图

2.1. 检测头网络的改进

原始的 YOLOv5 检测器缺乏用于角度预测的网络结构，并且不能提供水面垃圾物体的抓取角度信息。因此，水上清污机器人无法根据垃圾的实时角度设置最佳的抓取方案，这样在抓拣过程中容易导致物体掉落或者抓取失败。因此，我们优化了检测头的结构。角度预测可以通过回归或分类实现。回归中会对角度产生连续的预测值，但是这样会存在周期性边界问题，这会导致损失函数在周期性变化边界处的值突然增加，进而增加学习难度。例如，在 180° 长边定义方法中，定义的标签范围是 $(-90^\circ, 90^\circ)$ 。当垃圾的真实角度是 89° 而预测值为 -90° 时，模型学习到的误差为 179° ，但是实际误差应该是 1° ，这影响了模型的学习。

因此，我们在检测头中增加了卷积网络分支，并用旋转目标框长边绕中心顺时针旋转得到的 180 个类别来定义角度标签。角度卷积网络使用特征聚合网络获得的多尺度特征中提取的信息生成角度预测。

在检测头中, 角度卷积网络和原始网络共享特征聚合网络的输出作为输入特征图。角度预测网络和原始网络的输出合并如式(1)所示:

$$V = (\hat{c}, \hat{x}, \hat{y}, \hat{l}, \hat{s}, \hat{\theta}) \quad (1)$$

其中, $\hat{c}$ 是水面垃圾的预测类别, $\hat{x}$ 和 $\hat{y}$ 是目标框预测的中心坐标, $\hat{l}$ 和 $\hat{s}$ 为预测框长边和短边的预测长度, $\hat{\theta}$ 为定向水面垃圾的预测角度。

2.2. 角度平滑与损失函数

将角度预测作为一个分类任务可以避免回归引起的周期性边界问题, 但仍然存在一些局限性。传统的分类任务的损失函数计算大多为交叉熵损失, 标签的形式为独热标签编码, 如式(2)和(3)所示:

$$y_{ic} = \begin{cases} 1, & c = \theta \\ 0, & c \neq \theta \text{ \& } c \in \{0, 1, 2, \dots, 179\} \end{cases} \quad (2)$$

$$L = -\frac{1}{N} \sum_i \sum_{c=0}^{179} y_{ic} \log(p_{ic}) \quad (3)$$

其中, y_{ic} 是样本 i 角度的独热标签编码, θ 是定向垃圾的角度, p_{ic} 是检测模型的预测值。

如式(2)和式(3)所示, 对于不同角度的错误预测, 会获得相同的损失值, 并且无法量化错误的角度差值, 这使得模型训练难以确定定向水面垃圾的角度。为了解决这一问题, 本文基于 CSL 算法提出了一种动态平滑标签算法, 来优化角度的“独热标签编码”标签。

平滑标记算法如式(4)所示, 其中 θ 为旋转角度值, r 为平滑范围, $g(x)$ 为平滑函数。角度标签向量为密集分布因为 $g(x)$ 在平滑范围内。平滑函数值如式(5), 其中当 $\varepsilon = 0$ 时, 函数最大值为 1; 当 $\varepsilon = r$ 时, 函数最大值为 0。

$$\text{CSL}(x) = \begin{cases} g(x), & \theta - r < x < \theta + r \text{ \& } x \in \{0, 1, 2, \dots, 179\} \\ 0, & \text{others} \end{cases} \quad (4)$$

$$0 < g(\theta - \varepsilon) = g(\theta + \varepsilon) \leq 1, |\varepsilon| \leq r \quad (5)$$

CSL 算法部分增加了独热标签编码。当模型的角度预测在平滑范围内时, 可以获得不同预测角度的不同损失值, 所以它可以量化角度类别预测中的错误。然而, CSL 算法对平滑范围很敏感, 如果平滑范围太小, 平滑标签将退化为独热编码并失去效果, 很难学习角度信息。如果范围过大, 角度预测的偏差将会很大, 很容易导致抓取失败, 特别是对于那些长宽比较大的水面垃圾。因此, 本文提出动态平滑函数, 用于调整角度标签的平滑幅度和范围。

动态平滑函数使用动态高斯函数来平滑角度标签。从图 2 可以看出, 高斯函数的平滑幅度和范围受到均方根(RMS)值的控制。RMS 的值越大, 曲线越平坦; RMS 的值越小, 曲线越陡峭, 平滑范围越小。因此, 本文通过逐步缩小高斯函数的 RMS 值的方法实现动态平滑, 如式(6)所示。本文采用线性退火和余弦退火的方式来调整 RMS, 如式(7)和式(8)所示:

$$\text{DSM}(x) = \exp\left(-\frac{d^2(x, \theta)}{2 \times b^2}\right), x \in \{0, 1, \dots, 179\} \quad (6)$$

$$b = c + e \times \cos\left(\frac{0.5 \times \pi \times \text{epoch}}{\text{epochs}}\right) \quad (7)$$

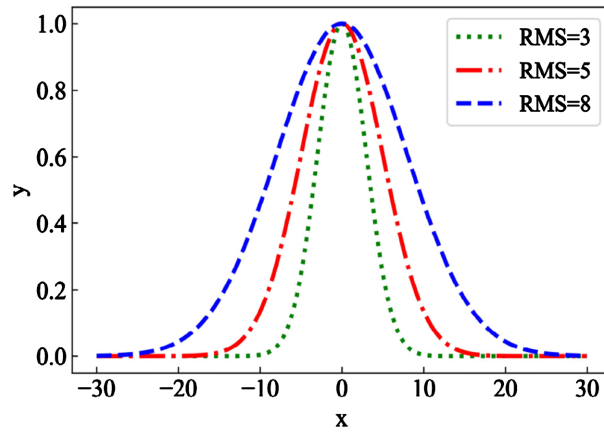


Figure 2. Gaussian function curves with different root-mean-square values
图 2. 不同均方根值的高斯函数曲线

$$b = c - e \times \frac{\text{epoch}}{\text{epochs}} \quad (8)$$

其中, θ 是垃圾旋转角度的值, 与函数的峰值位置相对应; x 是垃圾角度的编码范围; b 是 RMS 的值; $d(x, \theta)$ 是编码位置与角度值之间的圆周距离。例如, θ 为 179° 时, 当 x 为 0 时, $d(x, \theta)$ 等于 1; epoch 和 epochs 分别表示模型当前训练轮数和最大轮数, c 和 e 是超参数。从式(6)可以看出, DSM 根据编码位置和角度值之间的距离动态密集化角度标签。在模型训练前期, 由于 epoch 较小, b 的值较大, 此时平滑范围较大, 模型对角度的学习在窗口域内。平滑范围越大, 模型学习就越能接近最优点的邻域内。所以, DSM 降低了角度学习的难度, 并提高了水面垃圾检测的召回率。随着 epoch 值的增加, 角度平滑范围减小。模型学习目标从最优区域变为了最优点, 这使得角度预测的偏差更小。角度预测的准确性越高, 水面垃圾的召回率就越高, 特别是在对长宽比较大的垃圾角度预测上。

本文采用基于动态平滑算法的交叉熵损失函数计算水面垃圾的角度损失, 公式如式(9)所示:

$$\text{loss}(a) = - \sum_{i=0}^{s^2} I_{ij}^{obj} \sum_{t=0}^{179} \left\{ \hat{p}_i(t) \log[p_i(t)] + [1 - \hat{p}_i(t)] \log[1 - p_i(t)] \right\} \quad (9)$$

式中, $\hat{p}_i(t) = \text{DSM}(t)$, $p_i(t)$ 代表角度的预测值, s^2 是图片子域量化值, 该模型能够预测所有子域。 I_{ij}^{obj} 的值为 0 或者 1, 表示图像中是否存在目标。当预测值越接近真实值, 交叉熵的值就越小。

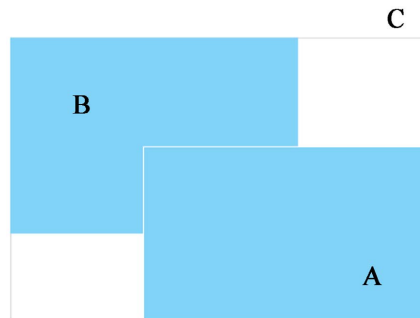


Figure 3. GIoU loss function
图 3. GIoU 损失函数

此外, 本文采用 GIoU 损失函数来计算检测边界框的回归损失。图 3 中 A 和 B 分别为检测目标的真

实框和预测框。C 是 A 和 B 外接的最小矩形，白色区域为 $|C| - |A \cup B|$ 。计算公式如式(10)-(12)所示。GIoU 不仅能够表示真实框与预测框重叠部分，还能够表示非重叠部分，这样能够解决由于 A 与 B 不重叠而导致梯度不计算的问题。

$$\text{IoU}(A, B) = |A \cap B| / |A \cup B| \quad (10)$$

$$\text{GIoU}(A, B) = \text{IoU}(A, B) - (|C| - |A \cup B|) / |C| \quad (11)$$

$$\text{loss}(r) = 1 - \text{GIoU}(A, B) \quad (12)$$

上式中，A 和 B 分别为检测目标的真实框和预测框。C 是 A 和 B 外接最小矩形，置信度损失函数和类别损失函数如式(13)和(14)所示。其中 I_{ij}^{obj} 和 I_{ij}^{noobj} 表示网格 i 的预测框 j 是否为目标框， l_{noobj} 表示权重系数。

$$\text{loss}(o) = -\sum_{i=0}^{s^2} \sum_{j=0}^B I_{ij}^{obj} [\hat{c}_i \log(c_i) + (1 - \hat{c}_i) \log(1 - c_i)] - l_{noobj} \sum_{i=0}^{s^2} \sum_{j=0}^B I_{ij}^{noobj} [\hat{c}_i \log(c_i) + (1 - \hat{c}_i) \log(1 - c_i)] \quad (13)$$

$$\text{loss}(c) = -\sum_{i=0}^{s^2} I_{ij}^{obj} \sum_{c \in \text{class}} \{ \hat{p}_i(c) \log[p_i(c)] + [1 - \hat{p}_i(c)] \log[1 - p_i(c)] \} \quad (14)$$

改进后模型的损失函数是上述损失函数的加权和，如式(15)所示：

$$\text{Loss} = \text{loss}(r) + \text{loss}(o) + \text{loss}(c) + \text{loss}(a) \quad (15)$$

2.3. 特征提取网络的改进

特征提取骨干网络可以提取图像中垃圾目标的特征信息。因为在定向水面垃圾检测中加入了角度预测，所以对于进行有效的特征提取提出了更高的要求，特别是在垃圾长宽比较大、垃圾重叠的情况下。

BottleneckCSP 是 YOLOv5 骨干网络中的主要模块。该模块使用 bottleneck 架构进行堆叠，随着网络的深入，特征提取能力逐渐增强，但是浅层特征逐渐丢失。浅层特征具有较低的语义信息，由于较少的卷积操作所以它可以被细化。采用多层特征融合的方式在计算机视觉任务中能取得不错的效果。旁路路径被认为是简化网络训练的关键因素。将不同层学习到的特征图串联起来，可以增强特征信息。此外，加入注意力机制即根据不同特征的重要性分配不同权重的方法在应对图像识别任务时效果良好。CA (Coordinate Attention Mechanism) 注意力机制就是其中一种表现良好的注意力机制。所以将 BottleneckCSP 中的中间特征进行融合并加入 CA 注意力机制模块，增强特征提取能力。

由于 bottleneck 中的残差块连接，由 bottleneck 间过多的特征融合导致的特征冗余不适合模型训练，参数量的增加意味着要消耗更多资源。因此，本文采用层间合并的方法连接特征层，如图 4 所示，优化后的模块命名为 OPBottleneckCSP。

OPBottleneckCSP 中的 CA 模块结构如图 5 所示。输入特征图沿水平坐标和垂直坐标进行编码，得到全局字段并对位置信息进行编码操作，有助于网络更准确地检测目标位置。

如式(16)所示，CA 模块将输入的特征图生成水平和垂直方向的特征图，然后进行 1×1 的卷积变换，生成的 $A \in R^{C/r \times (H+W)}$ 为水平和垂直方向空间信息的中间层特征图， r 为下采样尺度， F_1 为卷积运算。

$$A = \delta(F_1(F^h, F^w)) \quad (16)$$

其中 A 在空间维度上分为 $A^h \in R^{C/r \times H}$ 和 $A^w \in R^{C/r \times W}$ 。如式(17)和式(18)所示，通过卷积操作将其转化为与输入特征图相同的通道数， g^h 和 g^w 作为注意力权值参与特征图运算，CA 模块的输出结果如式(19)所示。

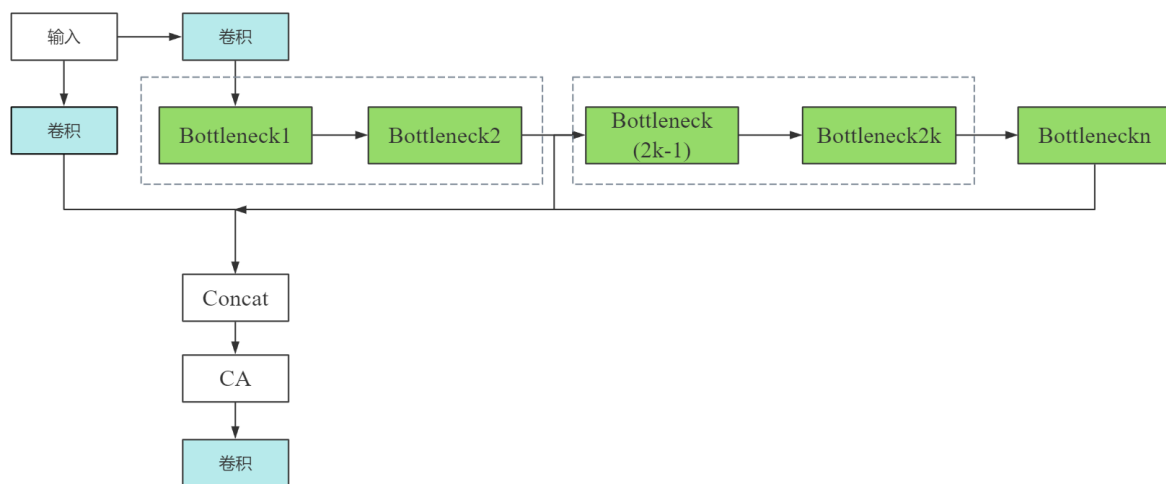


Figure 4. The optimized BottleneckCSP module

图 4. 优化后的 BottleneckCSP 模块

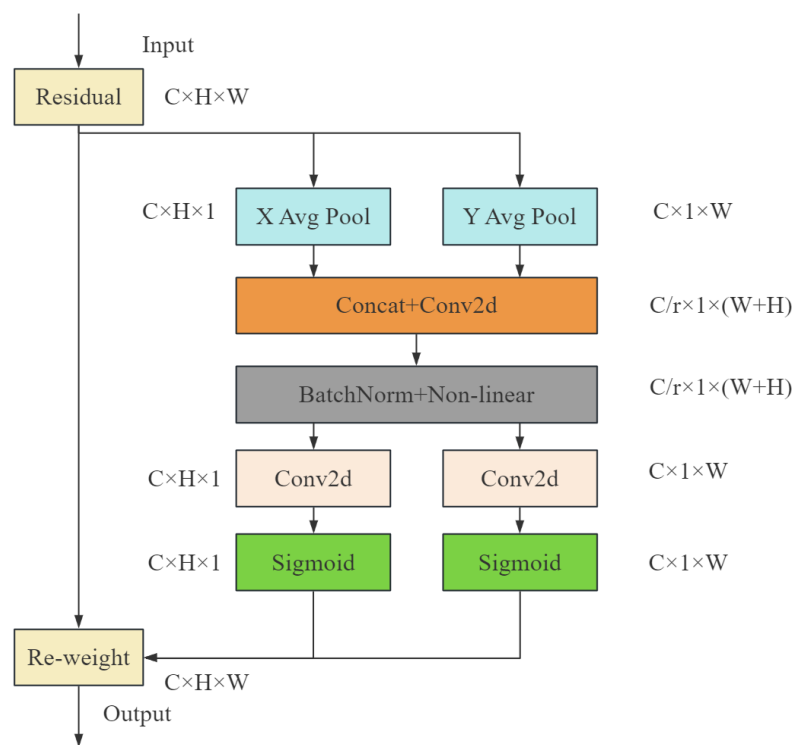


Figure 5. CA attention module

图 5. CA 注意力模块

$$g^h = \delta \left(F_h \left((A^h) \right) \right) \quad (17)$$

$$g^w = \delta \left(F_w \left((A^w) \right) \right) \quad (18)$$

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (19)$$

优化后的特征提取骨干网络结构如图 6 所示。它通过卷积模块和 OPBottleneckCSP 模块提取特征，并通过 1/8,1/16 和 1/32 三次下采样操作生成三种尺寸的特征图。

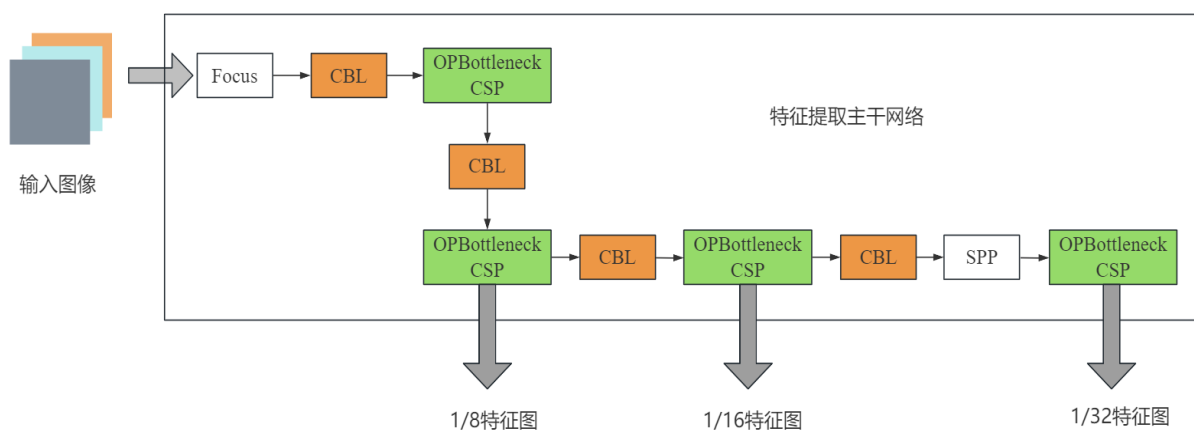


Figure 6. Improved feature extraction network

图 6. 改进后的特征提取网络

2.4. 特征聚合网络的改进

YOLOv5 特征聚合网络由特征金字塔网络(FPN) [7]和路径聚合网络(PANet) [8]组成。PANet 的结构如图 7(a)所示。PANet 沿着自顶向下和自底向上两条路径进行特征聚合。然而聚合的特征是高语义的深度特征，而高分辨率的浅层特征没有融合。为了更有效地利用输入特征，本文对 PANet 进行改进，如图 7(b)所示。

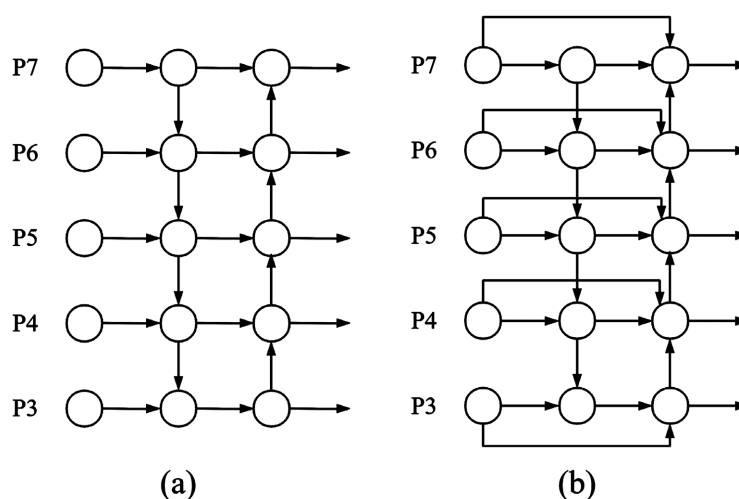


Figure 7. Network structure before and after PANet improvement

图 7. PANet 改进前后的网络结构

改进后的 PANet 增加了输入特征映射与输出特征映射之间的端到端连接，建立了低级特征到高级特征点对点连接路径，可以在不增加太多成本的情况下实现图像中高分辨率和复杂语义特征的融合。通过对高分辨率和低分辨率特征图的语义信息的提取，进一步增强了旋转水面垃圾目标的角度特征信息，提高了模型的检测能力。

经过优化后的水面垃圾旋转目标检测方法命名为 YOLOv5m-OW。当图片输入模型中时,模型使用主干网络提取特征,并生成三种不同尺寸的下采样特征图,用于检测水面垃圾。特征聚合网络进行特征聚合和融合,增强模型学习特征的能力。检测头根据多尺度特征生成水面垃圾目标的预测信息,在模型的训练阶段,使用动态平滑模块对训练集数据进行平滑处理,并通过损失函数计算预测中的损失,包括类别、角度和位置进行迭代学习。

3. 实验设计与结果分析

3.1. 实验数据集

本文采用水面垃圾数据集 FloW-Img [9]。该数据集标注框采用的水平框的方式进行标注,因此本文使用 roLabelImg 工具进行旋转目标框标注,最终数据集中包括 8068 张水面垃圾图像,垃圾类别主要为塑料瓶(Bottle),模型训练时,对数据集进行划分,其中训练集 6194 张图片,验证集 1005 张图片,测试集 869 张图片。

3.2. 评价指标

为了评估 YOLOv5m-OW 的性能,本文采用召回率(Recall)和平均准确率(mAP),公式如式(20)和式(21)所示。mAP50 表示模型在 IOU 大于等于 0.5 的条件下检测目标的平均准确性。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (20)$$

$$\text{mAP} = \frac{1}{m} \sum_{i=1}^m \text{AP}_i \quad (21)$$

3.3. 实验设计与结果分析

本实验将 YOLOv5m-OW 与其它主流的水平框检测方法如 SSD [10], YOLOv3 [11], YOLOv5s, YOLOv5m 进行对比实验结果如表 1 所示。SSD, YOLOv3, YOLOv5s 以及 YOLOv5m 都是单阶段目标检测模型,通过单个神经网络预测目标位置和类别,具有较高实时性。但是它们本身仅支持水平目标检测。

Table 1. Comparison between YOLOv5m-DSM and horizontal frame detection methods

表 1. YOLOv5m-OW 与水平框检测方法比较

方法	Recall/%	mAP50/%
SSD	82.3	74.8
YOLOv3	82.8	75.7
YOLOv5s	84.5	77.4
YOLOv5m	87.1	82.4
YOLOv5m-OW	94.6	93.2

从表 1 中可以得出,与 SSD, YOLOv3 和 YOLOv5s 相比, YOLOv5m 的召回率和平均准确率都更高。主要原因在于水平框在面对垃圾重叠,垃圾长宽比较大的情况下常出现误检的情况。与原网络 YOLOv5m 相比, YOLOv5m-OW 在召回率和平均准确率上分别提升了 7.5%和 10.8%,这说明改进后的 YOLOv5 算法对于水面垃圾检测效果更好,也证明了动态平滑标记的有效性。

表 2 为 YOLOv5m-OW 模型与其它主流旋转目标检测方法在该数据集上的结果对比图。在表中可以

看到, 与 Gliding-Vertex [12]相比, 召回率和平均准确率分别提高了 2.3%和 3.7%。与 R3Det [13]相比, 分别提高了 3.3%和 3.2%。同时, 召回率和平均准确率也优于 S2A-Net [14], 而且参数更少。本文模型的 GFLOPS 也低于其它比较算法, 使得该模型更适合嵌入在无人船设备中使用。

Table 2. Comparison between YOLOv5m-DSM and horizontal frame detection methods

表 2. YOLOv5m-OW 与主流旋转目标检测算法的比较

方法	Recall/%	mAP50/%	Params/M	GFLOPs
Gliding-Vertex	92.3	89.5	41.4	225.4
R3Det	91.3	90.0	47.9	250.3
S2A-Net	94.1	93.0	38.8	154.2
YOLOv5m-OW	94.6	93.2	23.7	76.5

通过对网络模型进行消融实验, 对各个改进方案的优化效果进行结果分析, 实验对比结果如表 3 所示。其中 opt1a 表示线性退火, opt1b 表示余弦退火。opt2 表示基于 OPBottleneckCSP 模块对特征提取骨干网络进行的改进, opt3 表示基于 YOLOv5 对 PANet 结构的特征聚合网络的改进, 实验结果如表 3 所示。

Table 3. The ablation experiment of each improved scheme

表 3. 各个改进方案的消融实验

方法	角度	opt1a	opt1b	opt2	opt3	Recall/%	mAP50/%
YOLOv5m	×	×	×	×	×	-----	-----
YOLOv5m-OW	√	×	×	×	×	87.1	82.1
YOLOv5m-OW	√	√	×	×	×	92.4	90.3
YOLOv5m-OW	√	√	×	√	×	93.4	91.8
YOLOv5m-OW	√	√	×	×	√	93.3	91.6
YOLOv5m-OW	√	√	×	√	√	94.6	93.2

由表 3 可以看出, 加入动态平滑算法和相应的损失函数后, 模型召回率和平均精确度分别提高了 5.3%和 8.2%。加入动态平滑算法并加入 OPBottleneckCSP 模块后, 模型分别提高了 6.3%和 9.7%。加入动态平滑算法和改进 PANet 模块后, 模型分别提高了 6.2%和 9.5%, 最终的 YOLOv5m-OW 模型召回率和平均精度分别提高了 7.5%和 11.1%。

为了验证本文提出改进方法的水面垃圾检测性能, 利用该方法对一些较为复杂的水面垃圾场景进行了实际测试, 实验结果如图 8 所示。

4. 结论

本文主要研究基于 YOLOv5 的水面垃圾旋转目标检测, 为了实现无人船准确判断水面垃圾的角度以便机械臂能够设置合适的抓取角度, 本文在检测头中增加了角度预测网络, 提供水面垃圾角度信息, 并提出了计算角度损失的动态平滑算法, 增强了模型的角度预测能力。此外, 通过优化模型的主干网络和特征聚合网络, 提高了方法的特征提取和聚合能力。实验结果表明, 提出的改进旋转目标检测算法 YOLOv5m-OW 的召回率和平均准确率分别为 94.6%和 93.2%, 在水面垃圾检测上优于传统的水平目标检



Figure 8. Results of rotating target detection of water garbage

图 8. 水面垃圾旋转目标检测结果

测算法以及主流的旋转目标检测算法，具备将模型移植进无人船设备的可行性，能够准确高效地检测复杂场景的水面垃圾。

参考文献

- [1] 范涛. 海上垃圾收集船舶路径优化研究[D]: [硕士学位论文]. 兰州: 兰州交通大学, 2021.
- [2] 刘婷, 周柏均, 赵永生, 等. 基于卷积神经网络的无人船海面垃圾分类算法[J]. 船舶工程, 2021, 43(S1): 399-404.
- [3] 丁泽彦. 基于无人机一多无人船协同的水域自动清理策略研究[D]: [硕士学位论文]. 武汉: 华中科技大学, 2021.
- [4] Li, Q., Wang, Z., Li, G., et al. (2023) An Accurate and Adaptable Deep Learning-Based Solution to Floating Litter Cleaning up and Its Effectiveness on Environmental Recovery. *Journal of Cleaner Production*, **388**, Article ID: 135816. <https://doi.org/10.1016/j.jclepro.2022.135816>
- [5] 王一早, 马纪颖, 罗星, 等. 基于 SPMYOLOv3 的水面垃圾目标检测[J]. 计算机系统应用, 2023, 32(3): 163-170.
- [6] 童小钟, 魏俊宇, 苏绍璟, 等. 融合注意力和多尺度特征的典型水面小目标检测[J]. 仪器仪表学报, 2023, 44(1): 212-222.
- [7] Lin, T.-Y., Dollar, P., Girshick, R., He, K., et al. (2017) Feature Pyramid Networks for Object Detection. *Proceedings of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 936-944. <https://doi.org/10.1109/CVPR.2017.106>
- [8] Liu, S., Qi, L., Qin, H., Shi, J. and Jia, J. (2018) Path Aggregation Network for Instance Segmentation. *Proceedings of the 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, 18-23 June 2018, 8759-8768. <https://doi.org/10.1109/CVPR.2018.00913>
- [9] Cheng, Y., Zhu, J., Jiang, M., et al. (2021) Flow: A Dataset and Benchmark for Floating Waste Detection in Inland Waters. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 10953-10962. <https://doi.org/10.1109/ICCV48922.2021.01077>
- [10] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y. and Berg, A.C. (2016) SSD: Single Shot Multibox Detector. *Proceedings of the European Conference on Computer Vision*, Amsterdam, 11-14 October 2016, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [11] Redmon, J. and Farhadi, A. (2018) YOLOv3: An Incremental Improvement. arXiv 2018, arXiv:1804.02767.
- [12] Xu, Y., Fu, M., Wang, Q., Wang, Y., Chen, K., Xia, G.S. and Bai, X. (2020) Gliding Vertex on the Horizontal Bounding Box for Multi-Oriented Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **43**, 1452-1459. <https://doi.org/10.1109/TPAMI.2020.2974745>
- [13] Yang, X., Yan, J.C., Feng, Z.M. and He, T. (2021) R(3)Det: Refined Single-Stage Detector with Feature Refinement for Rotating Object. *Proceedings of the AAAI Conference on Artificial Intelligence*, Virtual, 2-9 February 2021, 3163-3171. <https://doi.org/10.1609/aaai.v35i4.16426>
- [14] Han, J., Ding, J., Li, J. and Xia, G.S. (2021) Align Deep Features for Oriented Object Detection. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, 1-11. <https://doi.org/10.1109/TGRS.2021.3062048>