

MEC-YOLOv11n:水面小目标漂浮物的检测算法<sup>\*</sup>安梦军<sup>1,2</sup> 吴朝明<sup>1,2</sup> 邓承志<sup>1,2</sup>

(1.南昌工程学院信息工程学院 南昌 330000;2.智慧水利江西省重点实验室 南昌 330000)

**摘 要:**为应对水面波动、光照变化以及漂浮物部分遮挡等复杂场景中水面小目标漂流物检测面临的准确性和鲁棒性问题,提出了 MEC-YOLOv11n 算法。MEC-YOLOv11n 算法包括 Backbone、Neck 和 Head 共 3 个部分。为了增大目标感受野的识别区域,本研究设计了 MSWTC 结构,并在 Neck 部分改进了 C3k2 结构,这一优化显著提高了水面小目标漂流物的提取能力,从而增强了模型在复杂背景下对细节的捕捉能力;其次,提出了一种 EUCB 上采样方法,替代了 v11 中的传统上采样模块,该方法在上采样过程中增强了图像边缘的清晰度,使得高分辨率特征图中的目标轮廓更加精确,尤其在处理复杂背景和小目标检测任务时,显著提升了模型对细节的捕捉能力;最后,在 Head 前设计了一种专门用于识别边缘特征的注意力模块 CCA,进一步优化了模型在边缘信息提取方面的表现。实验结果表明,经过优化后的模型,其精确率 P 相较之前提高了 3.3%,召回率 R 提高了 2.4%,mAP50 提升了 2.5%,mAP50-95 提高了 1.5%。

**关键词:**YOLOv11;漂流物检测;小目标检测;目标感受野;边缘特征

**中图分类号:**TN911.73;TN919.8 **文献标识码:**A **国家标准学科分类代码:**510.4

MEC-YOLOv11n:Detection algorithm for floating objects of  
small targets on the water surfaceAn Mengjun<sup>1,2</sup> Wu Zhaoming<sup>1,2</sup> Deng Chengzhi<sup>1,2</sup>

(1.School of Information Engineering, Nanchang Institute of Technology,Nanchang 330000, China;

2.Smart Water Conservancy Jiangxi Provincial Key Laboratory,Nanchang 330000, China)

**Abstract:** To address the accuracy and robustness issues of small floating object detection on water surfaces under complex scenarios such as wave disturbances, changes in lighting, and partial occlusion by floating debris, the MEC-YOLOv11n algorithm is proposed. The MEC-YOLOv11n algorithm consists of three parts: Backbone, Neck and Head. To increase the recognition area of the target receptive field, we designed the MSWTC structure and improved the C3k2 structure in the Neck part, which significantly enhances the extraction ability of small floating objects on water surfaces, thus strengthens the model's ability to capture details in complex backgrounds; next, we proposed a EUCB up-sampling method, replacing the traditional up-sampling module in v11, which enhances the clarity of image edges during up-sampling, making the object contours more accurate in high-resolution feature maps, especially when dealing with complex backgrounds and small target detection tasks, which significantly improves the model's ability to capture details; finally, we designed an attention module CCA specifically for recognizing edge features before the Head, further optimizing the model's performance in edge information extraction. Experimental results show that after optimization, the precision P of the model has increased by 3.3%, the recall R has increased by 2.4%, the mAP50 has increased by 2.5%, and the mAP50-95 has increased by 1.5%.

**Keywords:** YOLOv11;drift detection;small target detection;target receptive field;edge features

## 0 引 言

水面漂浮物检测在环境监测、海洋污染防治和航运安

全等领域中具有重要的实际意义。漂浮物,尤其是塑料垃圾及其他污染物,不仅会显著影响水质和生态环境,还可能对航运安全构成严重威胁。因此,开发高效且精准的水面

漂浮物检测技术,已成为水域监控与生态保护领域中的关键研究方向。早期的水面漂浮物检测方法主要依赖传统的图像处理技术,通过提取图像中的颜色、形状、纹理等视觉特征进行目标识别。然而,这些传统方法在复杂环境中的应用效果较为有限,尤其是在水面波动、光照变化及目标重叠等情况下,往往难以提供稳定和可靠的检测结果。随着深度学习技术<sup>[1]</sup>,特别是卷积神经网络(CNN)的兴起,水面漂浮物检测技术迎来了新的发展机遇。基于深度学习的方法能够从数据中自动学习到有效的特征表示,突破了传统方法的局限性,显著提升了检测精度和系统的鲁棒性。

在深度学习技术的推动下,双阶段检测框架与 Transformer 架构等创新方法持续演进,其中 YOLO 系列模型<sup>[2]</sup>凭借其检测精度与场景适应性的平衡优势,已成为水面漂浮物小目标检测中应用最广泛的算法之一。YOLO 系列通过将目标检测转化为回归问题实现实时检测;而双阶段方法,如 Faster R-CNN 则通过区域生成网络(RPN)实现更精细的定位;对于 Transformer 架构,DETR 模型摒弃传统锚框设计,在背景复杂的近岸场景中展现出独特优势。针对小目标尺寸小、特征信息少的问题,张杨等<sup>[3]</sup>在改进的 Faster R-CNN 中提出了一种轻量化特征提取网络与特征金字塔相乘融合策略,并结合双线性插值优化感兴趣区域池化的量化误差,从而有效缓解小目标因浅层特征信息缺失及定位偏差导致的检测性能下降问题;Zhang 等<sup>[4]</sup>提出了 DETR++,该方法通过整合双向特征金字塔(BiFPN)、专用检测头和局部注意力机制,显著提升了小目标的检测精度;马梁等<sup>[5]</sup>提出了一种基于动态选择机制的轻量化特征提取模块,并提出了基于自适应特征加权融合的 FPN(feature pyramid networks)模块,该模块运用分组卷积技术对所有特征通道进行分组处理,各组间相互独立,有效提升了图像特征的表达精度,同时该方法能有效解决过多下采样可能引起的小目标信息丢失问题。Qiao 等<sup>[6]</sup>提出的 DetectoRS 算法,通过将 FPN 的输出重新输入骨干网络,进一步提升了小目标检测的精度,同时自适应调整感受野大小,增强了算法对不同尺度目标的适应能力。针对目标分布不均和尺度差异大的问题,赵子婧等<sup>[7]</sup>在改进的 Libra R-CNN 中通过引入 GA-RPN 生成自适应锚框,结合平衡采样策略和多尺度特征融合方法,减少背景影响和目标分布不均的问题,提高检测准确率;Dai 等<sup>[8]</sup>提出的 Dynamic DETR,结合动态编码器(自适应多尺度特征融合)与动态解码器(ROI 稀疏注意力聚焦),通过优化多尺度特征的整合和目标区域的精确聚焦,有效提升了模型在处理尺度差异和分布不均问题时的鲁棒性;陈旭等<sup>[9]</sup>基于 YOLOv5s 模型引入了残差空洞卷积模块,该模块能够增大感受野,提高空间特征的利用率,此外陈旭等<sup>[9]</sup>还提出了特征融合模块 SCAM,通过自监督方法利用局部特征,强化了细节信息的利用效率,从而显著提升了对大尺度差异小目标的检测精度;戚玲珑等<sup>[10]</sup>对 YOLOv7 进行改进,加入

了 Acmix 注意力机制等方法,提升了水上垃圾检测的精度,但其检测精度仍未达到实际应用的要求。针对背景干扰严重、目标被遮挡的问题,张娜等<sup>[11]</sup>基于改进的 Cascade R-CNN 算法,通过引入可切换空洞卷积模块、坐标注意力机制和 Repulsion Loss 损失函数,并构建了多模块协同的增强网络结构,在维持原算法高精度的同时,显著缓解了因目标遮挡导致的误检与漏检问题;Zhu 等<sup>[12]</sup>提出的 Deformable DETR,通过可变形注意力动态聚焦目标的关键区域,结合多尺度特征融合与迭代边界框优化,有效地提升了模型对小尺度目标的关注能力,增强了模型在复杂背景和部分遮挡条件下的鲁棒性;Xiao 等<sup>[13]</sup>基于 YOLOv7 设计了 C3GhostV2 特征提取模块,并引入无参数注意力机制(SimAm),增强了目标特征的提取能力,抑制了复杂背景对目标的干扰;张华卫等<sup>[14]</sup>提出的 GUS-YOLO 算法,充分利用全局信息,抑制了冗余的多尺度融合信息,从而提升了小目标检测的精度。针对同一问题,3 类方法呈现不同的解决路径:两阶段法侧重区域精细化,Transformer 依赖自注意力机制建立全局关联,YOLO 系列则追求效率与精度的平衡。

尽管现有方法在水面漂浮物检测中取得了一定的进展,但由于水面波动、光照变化以及漂浮物的部分遮挡等因素,尤其是在小目标检测<sup>[15]</sup>中,这些方法的检测精度和鲁棒性仍然受到显著影响。针对水面波动、光照变化以及漂浮物的部分遮挡等因素影响导致的水上小目标检测面临的准确性及环境适应性问题,本文提出了 MEC-YOLOv11n 算法,提高模型对水上小目标的检测性能。

本文主要包括 3 个创新点:

1) 设计了多尺度注意力小波变换卷积(multiscale attention wavelet transform convolution,MSWTC)模块,对 C3k2 进行了优化,在不显著增加参数数量的情况下增大对小目标漂浮物的感受野,提升网络的特征处理能力来提高准确性。

2) 在 neck 网络中采用了高效的上采样(edge-aware upsampling with channel boosting,EUCB)模块,用于逐步上采样特征图,使其尺寸和分辨率与后续跳跃连接相匹配,这种对齐增强了特征图中的目标轮廓在不同层级和阶段间的信息融合,保留了更多的特征,提高了模型的鲁棒性。

3) 在检测头前设计了边缘卷积注意力(contour convolution attention,CCA)模块,进一步优化了模型对小目标的细节捕捉,显著提高了对小目标漂流物的检测性。

## 1 Yolov11 算法

YOLOv11<sup>[16]</sup>是由 Ultralytics 推出的最新版本,集成了目标检测、语义分割、图像分类等多项任务。与 YOLOv8<sup>[17]</sup>相比,YOLOv11 在多个方面进行了显著的改进。首先,YOLOv11(如图 1 所示)采用了改进的模型结构,以提升图像处理能力和预测精度,进一步优化了特征提

取与预测生成过程。其次, YOLOv11 针对 GPU 进行了优化, 显著提高了训练速度和准确性, 尤其在大规模数据集上的表现尤为突出。通过模型结构优化和 GPU 加速, YOLOv11 在推理和训练过程中相较于前版本的速度提高了多达 25%, 延迟显著降低, 从而使其在实时应用场景中的表现更加优异, 此外, YOLOv11 通过减少不必要的参数, 提升了模型的运行效率, 同时确保了准确性不受影响。这一优化使得 YOLOv11 能够在保持高性能的同时, 更好地适应资源受限的应用场景。

在主干网络方面, YOLOv11 引入了 C3k2 模块, 该模块继承并优化了 C2f 模块。C2f 模块是一种特征融合模块, 而 C3k2 模块在此基础上进行了改进, 通过设置 C3k 参数为 True 或 False, 模型可以灵活选择不同的特征提取方式。当 C3k 为 False 时, C3k2 模块作为传统的 C2f 模块, 包含常规的瓶颈结构; 而当 C3k 为 True 时, 瓶颈模块将被更

高效的 C3 模块所替代。C3 模块的设计旨在更有效地提取图像特征, 灵活可切换的设计提升了模型的适应性和特征提取效率, 使得 YOLOv11 能够在不同场景下做出最佳选择。此外, YOLOv11 还引入了 C2PSA 模块, 这是对 C2f 模块的扩展, C2PSA 模块结合了 PSA 块<sup>[18]</sup>, PSA 利用多头注意力机制和前馈神经网络, 能够选择性地关注输入特征中的重要部分, 同时抑制无关信息。这对于处理复杂背景下的目标检测任务<sup>[19]</sup>尤为重要。通过将 PSA 机制融入 C2f 模块, C2PSA 模块不仅能更有效地捕捉图像中的关键特征, 提升模型对目标区域的关注能力, 还能够通过引入残差结构<sup>[20]</sup>优化梯度传播, 进一步改善网络训练效果, 提高模型的鲁棒性和准确性。总的来说, YOLOv11 在多个方面进行了优化, 使其在速度和准确性上超越了前版本, 并通过引入更高效的特征提取与注意力机制, 进一步增强了在复杂环境下的表现。

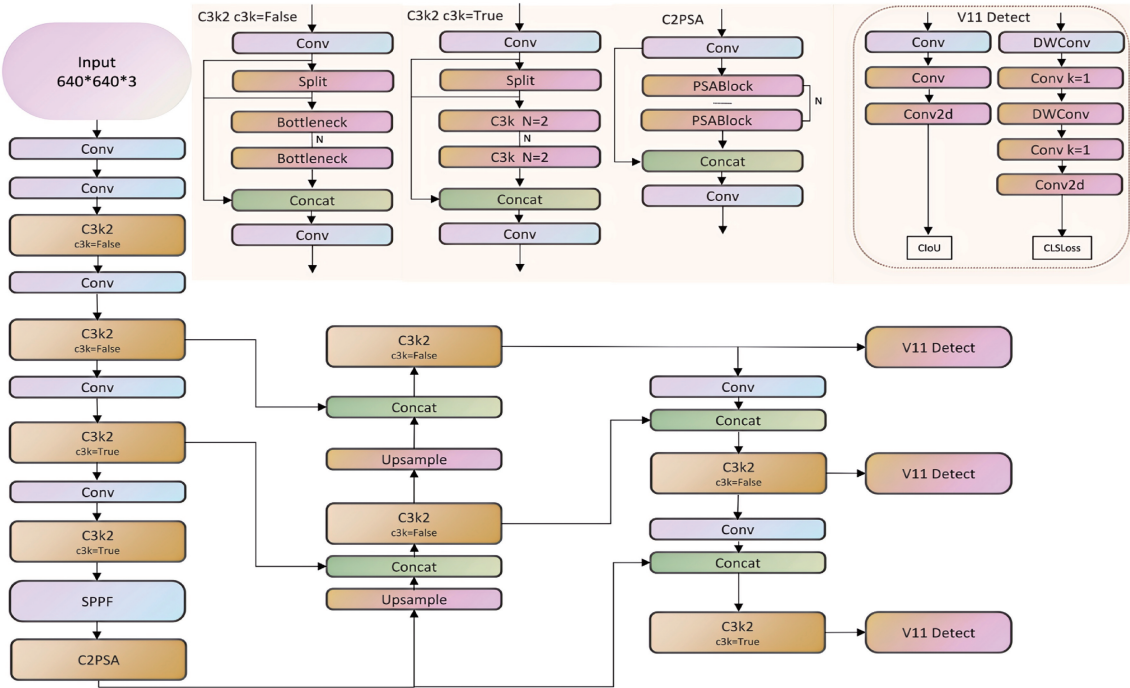


图 1 YOLOv11n 模型  
Fig. 1 YOLOv11n model

2 MEC-YOLOv11n 算法

2.1 MEC-YOLOv11n 算法结构

WEC-YOLOv11n 算法整体结构如图 2 所示。该算法在保留 YOLOv11n 整体架构的基础上, 设计了 MSWTC, 以优化 Neck 部分的 C3k2 结构; 此外采用了高效的 EUCB 上采样方法, 取代了传统的 Upsample 上采样策略; 最后在 Head 部分前, 本文设计了 CCA 模块, 使检测头能够更加专注于小目标漂浮物。通过上述改进, WEC-YOLOv11n 算法在计算复杂度、参数量和计算效率之间实现了性能与

效率的平衡, 从而确保了在水上小目标漂浮物边缘设备<sup>[21]</sup>部署时, 能够提供高性能的检测能力。

2.2 多尺度注意力小波变换卷积

MSWTC 结构(如图 3 所示)的思想来源 Finder<sup>[22]</sup>等提出的小波变换卷积(wavelet transform convolutions, WTC)结构。WTC 结构主要利用小波变换(WT)有效增加卷积神经网络(CNN)的感受野, 而不增加模型的参数量。传统意义上, 增加感受野通常意味着使用更大的卷积核, 这可能导致过多的参数和计算成本, 小波变换作为一种在时频分析中广泛应用的工具, 能够在保持空间分辨率



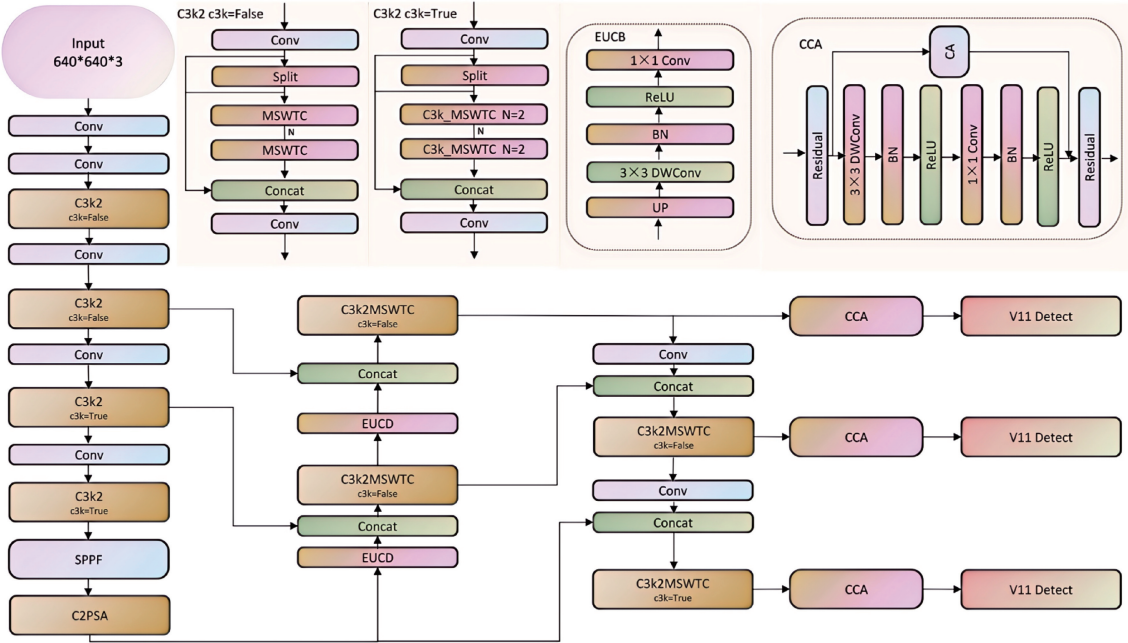


图 2 MEC-YOLOv11n 网络结构

Fig. 2 MEC-YOLOv11n network structure

的同时,扩展感受野。通过使用小波变换级联分解执行一系列小卷积核操作,能够同时提取不同层次的低频(整体轮廓)和低频(细节纹理)信号,从而聚焦于输入数据的不同频率带,并逐步增大感受野,这种方法的核心思想是通过将输入信号分解为多个频率带,每个卷积操作专注于不同的频率范围,网络能够在多个尺度<sup>[23]</sup>上捕获信息。通过这种方式,网络能够在不直接增加大量参数的情况下,扩大其感受野,从而能够捕捉更广泛的上下文信息。在此基础上,进一步引入了多尺度注意力<sup>[24]</sup>特征加权机制,该机制根据不同尺度特征的重要性动态调整其注意力权重,使得模型能够更加关注任务中最关键的特征。重要尺度上的特征可能包含更多有用信息,而其他尺度的特征可能包含冗余或噪声。通过动态加权 MSWTC 模块能够智能筛选信息<sup>[25]</sup>并优化特征表示,从而在扩展感受野的同时,更加聚焦于小目标漂浮物的细节,显著提高模型在水上小目标漂浮物检测中的准确性与鲁棒性。

在前述内容中,本文介绍了 MSWTC 模块的基本原理和优势,接下来本文将通过公式详细说明该模块的工作机制以及如何优化信号处理,提升小目标漂浮物检测的精度。对于给定的图像  $X$ ,一阶小波变换在一个空间维度(宽度或高度)上通过深度卷积操作如下:

$$WT(X) = \text{Conv}\left(\left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right], \left[\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right], X\right) \downarrow 2 \quad (1)$$

其中,  $\downarrow 2$  表示以 2 倍因子进行下采样操作。

对于二维图像,将一维变换应用于图像的两个空间维度(宽度和高度),得到 4 个不同频率带的输出:

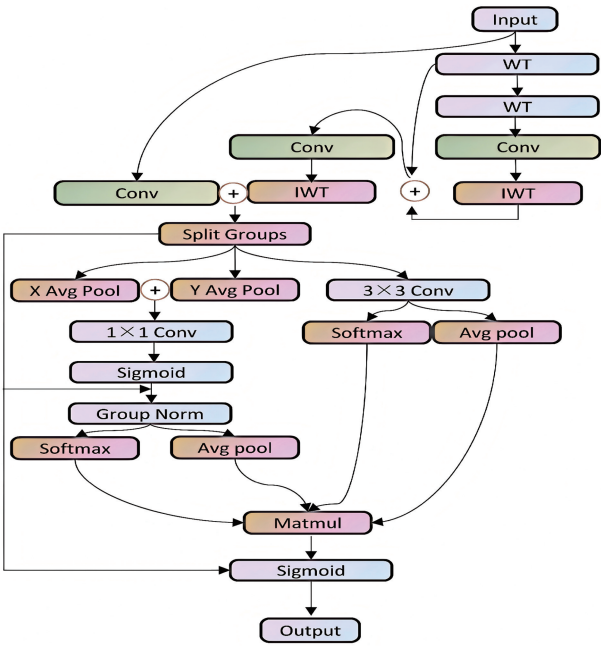


图 3 MSWTC 模块

Fig. 3 MSWTC module

$$f_{LL} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, f_{LH} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix},$$
$$f_{m} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, f_{m} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad (2)$$

其中,  $f_{LL}$  是低通滤波器,  $f_{LH}$ ,  $f_{m}$ ,  $f_{m}$  是高通滤波器。



通过这些滤波器对输入图像  $X$  进行卷积,得到4个输出通道:

$$\begin{bmatrix} X_{LL}, X_{LH}, \\ X_{HL}, X_{HH} \end{bmatrix} = \text{Conv} \left( \begin{bmatrix} f_{LL}, f_{LH}, f_{HL}, f_{HH} \end{bmatrix}, X \right) \quad (3)$$

其中,  $X_{LL}$  是低频部分,而其他三个是高频成分,分别对应水平、垂直和对角线的高频信息。

由于上述小波基是正交的,逆小波变换(IWT)是通过转置卷积操作来完成的,它通过恢复小波变换后的低频和高频成分来重建原始图像:

$$X = \text{Conv-transposed} \left( \begin{bmatrix} f_{LL}, f_{LH}, f_{HL}, f_{HH} \end{bmatrix}, \begin{bmatrix} X_{LL}, X_{LH}, X_{HL}, X_{HH} \end{bmatrix} \right) \quad (4)$$

为了进一步扩大感受野,级联小波分解通过递归地对低频成分进行小波变换来进行,每一层的分解结果包含低频和高频分量,逐步减小空间分辨率,增加频率分辨率:

$$X_{LL}^{(i)}, X_{LH}^{(i)}, X_{HL}^{(i)}, X_{HH}^{(i)} = \text{WT}(X_{LL}^{(i-1)}) \quad (5)$$

其中,  $X_{LL}^{(0)} = X$  是原始输入,  $i$  表示当前的小波分解层级。

而在小波域进行卷积时,首先对输入进行小波变换(WT)以分解低频和高频成分,然后在每个频率成分上使用小卷积核进行深度卷积,最后使用逆小波变换(IWT)将卷积结果恢复:

$$Y = \text{IWT}(\text{Conv}(W, \text{WT}(X))) \quad (6)$$

其中,  $W$  是  $k \times k$  的深度卷积核,具有四倍于输入通道数的参数,通过这种方法,小卷积核可以在较大的区域内操作,从而扩大感受野。

从而进一步将级联小波变换与小卷积核结合,以递归地对每一层的低频部分进行处理,每一层的卷积过程如下:

$$\begin{aligned} X_{LL}^{(i)}, X_H^{(i)} &= \text{WT}(X_{LL}^{(i-1)}), \\ Y_{LL}^{(i)}, Y_H^{(i)} &= \text{Conv}(W^{(i)}, (X_{LL}^{(i)}, X_H^{(i)})) \end{aligned} \quad (7)$$

其中,  $X_H^{(i)}$  表示第  $i$  层的所有高频分量。

最后,通过逆小波变换将不同频率层次的结果结合:

$$Z^{(i)} = \text{IWT}(Y_{LL}^{(i)} + Z^{(i+1)}, Y_H^{(i)}) \quad (8)$$

其中,  $Z^{(i)}$  是从第  $i$  层开始的所有卷积结果的聚合输出。

在小波变换卷积之后,引入注意力机制来对不同尺度的特征进行加权,注意力机制的目标是根据各尺度特征对当前任务的相关性,动态调整它们的权重。设定每个尺度  $\alpha$  的注意力权重为  $\alpha_{att}(\alpha)$ ,该权重可以通过以下注意力机制进行计算:

$$\alpha_{att}(\alpha) = \frac{\exp(f_\alpha)}{\sum_{\alpha'} \exp(f_{\alpha'})} \quad (9)$$

其中,  $f_\alpha$  是通过某种方式计算得到的尺度特征的相关性得分,这些权重  $\alpha_{att}(\alpha)$  反映了每个尺度特征在当前任务中的重要性。

在计算出每个尺度的注意力权重后,对小波变换卷积得到的多尺度特征进行加权。每个尺度的加权特征  $\hat{x}_\alpha$  可以表示为:

$$\hat{x}_\alpha(t) = \alpha_{att}(\alpha) Z_\alpha^{(i)} \quad (10)$$

其中,  $Z_\alpha^{(i)}$  是小波变换卷积得到的某一尺度  $\alpha$  上的信号特征,  $\alpha_{att}(\alpha)$  是通过注意力机制计算的加权系数。

最后,将加权后的多尺度特征进行融合,得到最终的输出特征  $\hat{x}(t)$ ,可以表示为:

$$\hat{x}(t) = \sum_{\alpha} \hat{x}_\alpha(t) \quad (11)$$

结合以上部分,MSWTC模块的整体公式可以表达为:

$$\hat{x}(t) = \sum_{\alpha} \alpha_{att}(\alpha) \text{IWT}(Y_{LL}^{(i)} + Z^{(i+1)}, Y_H^{(i)}) \quad (12)$$

改进后的C3k2MSWTC的优势不仅在于能够同时捕捉信号的多层次和多频率特征,还通过注意力加权机制<sup>[26]</sup>有效避免了对不重要信号成分的过度关注,从而显著减少了噪声干扰和信息丢失。此外,C3k2MSWTC通过扩展感受野,使模型能够获取更大范围的上下文信息,显著提升了在复杂环境中对关键信号的识别能力,这一特性在水上小目标<sup>[27]</sup>漂流物检测中表现尤为突出,因为其能够在广阔且动态变化的水面背景中更准确地识别和定位小型漂浮物体。这样的结合使得该模块在处理复杂信号时,能够更高效地识别和提取关键信息,提升了对快速变化和突发事件的响应能力,并增强了对全局趋势的捕捉能力,表现出显著的精度和适应性。特别是在信号去噪和图像处理等任务中,C3k2MSWTC能够在多个尺度上对信号进行优化<sup>[28]</sup>,提供更加精细和准确的处理结果,确保在各种复杂环境下都能够有效地提取并利用信号中的关键信息,通过多尺度注意力机制动态调整不同尺度特征的权重,进一步提升了模块的自适应能力,使得其在不同应用场景下的性能得到了显著增强。

### 2.3 上采样

目前常用的上采样方法主要包括最邻近插值法、双线性插值法和反卷积法。前两种方法主要集中于亚像素空间,因其感受野较小,无法有效获取丰富的语义信息;而反卷积法尽管能够进行较为灵活的上采样,但在局部变化感知方面存在局限,且其引入的参数较多,计算量较大。特别是在处理不同图像内容、尺度变化或复杂场景时,传统上采样方法通常无法根据图像的具体特征自动调整上采样策略。因此,针对这一问题,本文使用了一种更高效的EUCB<sup>[29]</sup>方法,如图4所示。该方法能够逐步上采样当前阶段的特征图,并通过自适应调整其维度和分辨率,从而有效匹配来自下一个跳跃连接的特征图。

EUCB首先使用一个具有比例因子2的上采样 $Up$ 来放大特征图,然后它通过应用一个 $3 \times 3$ 深度卷积DWC后跟一个BN和一个ReLU激活来增强上采样的特征图。最后,使用一个 $1 \times 1$ 卷积来减少通道数以匹配下一阶段。

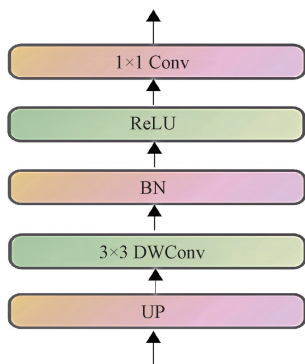


图 4 EUCB 模块

Fig. 4 EUCB module

EUCB 的公式为:

$$\begin{aligned} EUCB(x) &= C_{1 \times 1}(ReLU(BN(x))), \\ BN(x) &= BN(DWC(UP(x))) \end{aligned} \quad (13)$$

与传统上采样方法(如双线性插值法、转置卷积法等)相比,EUCB 上采样方法不仅具有更强的边缘感知能力,还在扩展感受野的过程中更加关注小目标的边缘细节。传统方法通常难以有效保留图像中的边缘信息,容易导致细节丢失或图像模糊<sup>[30]</sup>,尤其在复杂背景和小物体检测任务中,这一问题尤为突出。而 EUCB 上采样通过引入边缘感知模块,在上采样过程中增强图像边缘的清晰度,使得在高分辨率特征图中目标的轮廓更加精确。与此同时,随着感受野的扩展,模型能够捕捉更大范围的上下文信息,这一过程与边缘感知相辅相成,有助于在复杂场景中精准定位小目标。此外,EUCB 上采样还通过通道增强机制优化了各通道信息的融合<sup>[31]</sup>,提升了模型在多样性特征表达方面的能力。通过边缘感知和通道增强策略的结合,EUCB 上采样能够更好地恢复不同尺度<sup>[32]</sup>下的特征细节,还通过扩展感受野,增强了模型对全局和局部信息的敏感度。尤其对于小物体和细节丰富的场景,EUCB 上采样能够生成更高质量的特征图,从而有效提升检测精度,特别是在小目标检测任务中,展示出明显的优势。

## 2.4 边缘卷积注意力

为增强检测头对小目标的识别能力,在检测头前引入了一种更高效的注意力机制。受到 Hou<sup>[33]</sup>等提出的坐标注意力(coordinate attention, CA)启发,结合其核心思想,创新性地设计了一个新型注意力模块。CA 模块通过在水平和垂直方向上应用一维卷积来有效捕捉长程依赖关系及精确的空间位置信息,从而显著提升了网络在多种视觉任务中的表现,尤其在处理图像中具有显著空间结构依赖的任务时,该模块能够较好地平衡计算效率与性能之间的矛盾,进一步优化了目标检测的效果。

基于此,本文在坐标注意力机制的基础上,进行了关键性的创新与优化,以更好地满足水面漂浮物检测任务的特定需求。具体而言,对坐标注意力模块进行了以下几方

面的改进:首先,引入了一个专门用于边缘特征提取的模块,采用高效的卷积<sup>[34]</sup>操作来精准捕捉水面漂浮物的边缘信息。该模块通过逐通道卷积的方式,显著提升了漂浮物轮廓的识别精度,对于定位任务具有重要作用,能有效增强漂浮物的识别<sup>[35]</sup>与定位准确性。其次,为了更好地处理漂浮物在图像中可能存在的不同尺度和空间结构,对注意力图的生成策略进行了优化,进一步强化了模型对不同方向空间信息的建模能力,同时保持了对长程依赖的捕捉能力。通过这些优化,模型在处理复杂背景和小目标时,能够更加高效地集中注意力于目标区域。最后,在模块中引入了多层次的空间与通道注意力<sup>[36]</sup>机制,使得模型能够更加精准地关注漂浮物的关键区域,从而进一步提升了模型在水面漂浮物检测中的精度。通过这些创新和优化,本文的模型在水面漂浮物检测任务中表现出了显著的性能提升。CCA 模块(如图 5 所示)不仅能够在常规视觉任务中提升网络性能,更在水面漂浮物检测这一特定应用场景中展现了独特的优势,该模块显著增强了模型在复杂背景下对小目标边缘纹理特征的识别能力,同时提升了检测的准确性与鲁棒性,尤其在处理细粒度的边缘信息时,展现了更高的精度和稳定性。

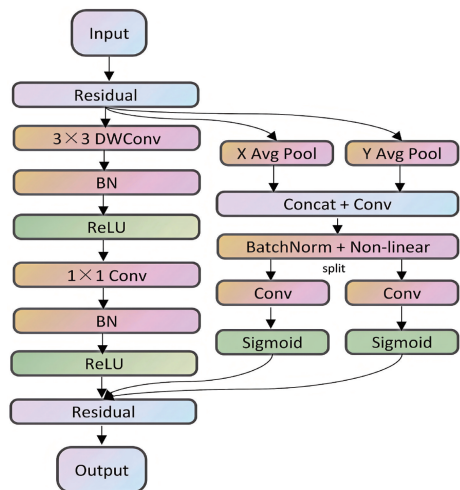


图 5 CCA 模块

Fig. 5 CCA module

## 3 实验结果分析

### 3.1 FloW-IMG 数据集

为了验证改进后算法的性能,本实验选择了 FloW-IMG 数据集<sup>[37]</sup>作为评估对象。该数据集由欧卡智船公司开发,是全球首个从无人船视角出发,专门针对真实内陆水域环境中漂浮物垃圾检测而构建的数据集。FloW-IMG 数据集(如图 6 所示)内容丰富,共包含 2 000 张高清图像,并精确标注了 5 271 个目标对象,为算法的性能评估提供了坚实的数据基础。该数据集涵盖了多种具有挑战性的场景,包括复杂光照条件、大范围视角、多物体交织的环境

以及远距离目标等。这些多样化的场景不仅反映了漂浮物在不同环境条件下的分布特征,还充分考察了算法在处理不同尺度、不同距离以及复杂背景下的鲁棒性和精度。

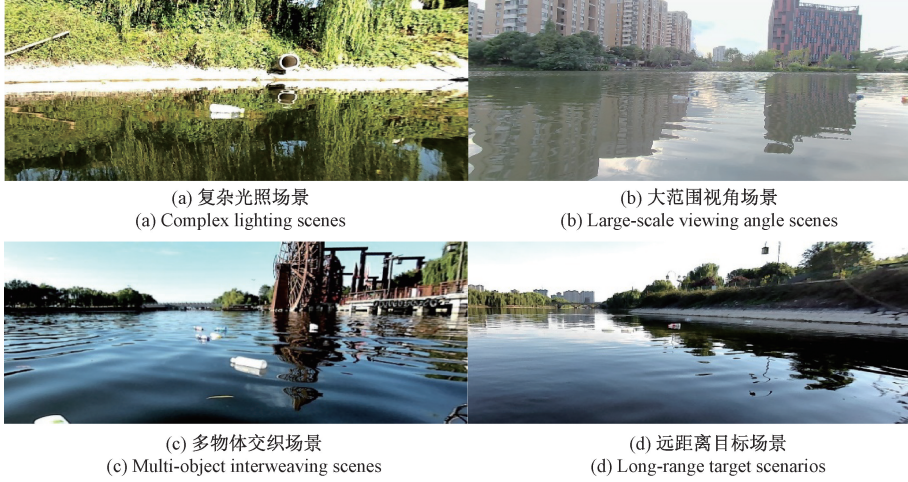


图 6 FloW-IMG 数据集不同场景

Fig. 6 FloW-IMG dataset for different scenarios

通过 Python 脚本,本文将 FloW-IMG 数据集中 2 000 张图像文件及其对应的标签文件进行随机打乱,并按 6:2:2 的比例划分为训练集、验证集和测试集。此比例的划分确保了数据集在不同子集中的分布均匀,有助于提升模型训练的稳定性与泛化能力。

为了进一步增强模型的鲁棒性并减小水波、倒影等环境干扰对漂浮物识别精度的影响,本文对训练集的图像进行了多种数据增强处理。具体而言,应用了旋转、混合、亮度调节、色彩调整等增强方法,以增强训练数据的多样性。这些增强手段不仅扩大了训练数据的有效性,还能在一定程度上模拟不同环境下的实际情况,从而帮助模型更好地应对变化的环境和复杂的背景噪声。此外,为了提高模型在不同尺度目标上的表现在训练过程中采用了多尺度训练(multiscale)策略,这一策略使得模型能够在多个尺度下进行目标检测,进一步提升了其对小物体及复杂场景的适应能力,从而为模型的训练和性能评估提供了更为坚实的基础。

### 3.2 实验环境及参数设置

本实验在 Windows 11 操作系统上进行,使用深度学习框架 Pytorch 2.4.1 和 CUDA 12.4,实验所用的硬件配置包括 Intel i5-12600KF 处理器、NVIDIA RTX 4060 显卡以及 32 GB 内存。为了确保实验结果的鲁棒性,设置了 500 个训练周期(epochs)以确保网络模型的充分收敛。为了防止过度训练,在训练过程中设置了早期停止机制,若模型在连续 50 轮训练中未出现性能提升(patience),则提前终止训练。在每个训练周期中,批次大小(batch size)设定为 8 张图像,以平衡训练效率和计算资源占用。同时使用了 4 个线程加载数据,以优化训练过程中的数据读取速

因此,FloW-IMG 数据集为评估水域漂浮物检测算法在实际应用中的适应性和有效性提供了全面的验证平台。

度,并减少训练时间的延迟。最后在测试集上对模型的各项性能指标进行了评估,以全面衡量模型在实际应用中的表现。

### 3.3 评价指标

本实验旨在优化水上垃圾目标检测算法,提升精度。为准确评估改进效果,设定了多个评估指标,具体包括精确率(precision, P)、召回率(recall, R)和平均精度均值(mean Average Precision, mAP)。这些指标的计算公式如下:

精确率(Precision, P):

$$P = \frac{TP}{TP + FP} \times 100\% \quad (14)$$

其中,TP 表示真正例(True Positives),FP 表示假正例(False Positives)。

召回率(Recall, R):

$$R = \frac{TP}{TP + FN} \times 100\% \quad (15)$$

其中,TP 表示真正例,FN 表示假负例(False Negatives)。

平均精度均值(mAP):

$$mAP = \frac{\sum_{i=1}^N \int_0^1 P dR}{N} \times 100\% \quad (16)$$

其中,N 是类别数,P 是精确率, $\int_0^1 P dR$  是平均精度(AP)。

在目标检测任务中,精确率 P 越高,表示模型在检测到的目标中误报较少,即能够准确区分目标与非目标;召回率 R 越高,意味着模型能检测出更多的真实目标,但这



可能会导致误报增加,即模型可能会将一些非目标误识别为目标。平均精度均值(mAP)作为一个综合性评估指标,能够平衡精度与召回率的表现,高 mAP 值意味着模型在多个类别上都能实现较高的精度和较强的召回能力,从而能够更加全面和准确地检测目标。

3.4 消融实验

为了全面评估所提出算法的检测性能并验证各改进模块对模型性能提升的作用,本研究首先进行了 C3k2MSWTC 模块的实验(如表 1 所示),探索了小波分解层数(wt\_levels)和注意力机制超参数(groups)的影响。在此基础上,冻结了小波滤波器层(wt\_filter 和 iwt\_filter),避免训练这些固定的数学变换层,以减少计算负担并集中调整其他可调超参数。具体而言,wt\_levels 控制小波分解的层数,影响模型在不同尺度上提取特征的能力,通过调整此超参数,测试不同分解层数对性能的影响;另一方面,groups 是通道分组数,影响注意力机制的效果,通过调整 groups,优化了通道间的注意力分配,进一步提升了模型的性能。

基于 C3k2MSWTC 模块的初步实验结果(其中选择了最佳配置,即在 wt\_levels:1,groups:8 下获得的性能),后

表 1 C3k2MSWTC 消融实验				
Table 1 C3k2MSWTC ablation experiment				
wt_levels	groups	mAP50	mAP50-95	Params(总)
1	4	0.816	0.396	2 526 139
2	4	0.822	0.394	2 531 259
3	4	0.822	0.393	2 536 379
1	8	0.824	0.393	2 522 235
2	8	0.821	0.401	2 527 355
3	8	0.816	0.387	2 532 475
1	16	0.800	0.385	2 521 243
2	16	0.819	0.393	2 526 363
3	16	0.798	0.386	2 531 483

续实验在相同实验配置下,逐一引入其他改进模块,进一步分析它们对模型性能的具体贡献。实验结果如表 2 所示,详细展示了每个改进模块对整体检测性能的具体贡献,通过一系列实验分析,结果表明每个模块的引入均有效提升了模型的性能,且各模块间的组合在多个方面展现了协同效应。

表 2 消融实验结果对比				
Table 2 Comparison of ablation experiment results				
Model	P	R	mAP50	mAP50-95
YOLOv11n	0.82	0.751	0.812	0.384
C3k2MSWTC	0.863	0.748	0.824	0.393
EUCB	0.845	0.757	0.819	0.394
CCA	0.836	0.755	0.822	0.387
C3k2MSWTC + EUCB	0.868	0.743	0.812	0.396
C3k2MSWTC + CCA	0.858	0.759	0.836	0.394
EUCB + CCA	0.841	0.767	0.823	0.387
C3k2MSWTC + EUCB + CCA(our)	0.853	0.775	0.837	0.399

具体而言,单独使用 MSWTC、EUCB 和 CCA 模块时,均能够显著增强模型的检测能力,验证了每个模块在特定任务中的有效性;而在实验 C3k2MSWTC+EUCB 中,虽然 mAP50 未能显著提高,但在 mAP50-95 的指标上却提高了 1.2%,这表明该模块在提高模型对不同 IoU(intersection over union)阈值下目标检测的鲁棒性方面起到了关键作用;将 C3k2MSWTC+CCA 模块结合使用时,mAP50 提升了 2.4%,mAP50-95 提升了 1.0%,该组合优化了模型在多尺度特征提取和长程依赖建模方面的能力,进一步提升了对复杂场景和小物体的检测精度;在实验 EUCB+CCA 模块结合的实验中,mAP50 提升了 1.1%,mAP50-95 提升了 0.3%,这一结果表明,实验 EUCB+CCA 能够在提升模型对小目标和边缘特征的检测精度方面具有显著效果,尤其是在高 IoU 阈值下,能够增强模型

对复杂背景中多尺度目标的适应能力;而本文提出的 MEC-YOLOv11n 结构中,mAP50 提高了 2.5%,mAP50-95 提高了 1.5%,尽管增加了 0.12 M 参数量和 0.6 GFLOPs 的计算复杂度,但通过高效的模块优化,模型在实时检测性能上保持了优异表现,并在复杂环境下进一步提升了鲁棒性和精度。

3.5 可视化分析

为了全面评估本文提出的算法在实际应用场景中的检测性能,将所提模型与 YOLOv11n 模型在 FloW-Img 数据集的挑战性场景中进行了对比测试。这些场景涵盖了目标距离较远、光照变化显著、图像模糊以及帧丢失等复杂条件。本文的实验结果包括在训练集和验证集上,模型的 box loss、cls loss 和 dfl loss 曲线,以及其他关键性能指标,如 P(精确率)、R(召回率)、mAP50 和 mAP50-95

的收敛曲线(如图 7 所示),横坐标表示迭代次数,实验在迭代 350 轮时触发了早停机制。通过观察曲线可以发

现,在训练 200 轮后,各项损失曲线趋近于零,模型开始收敛。

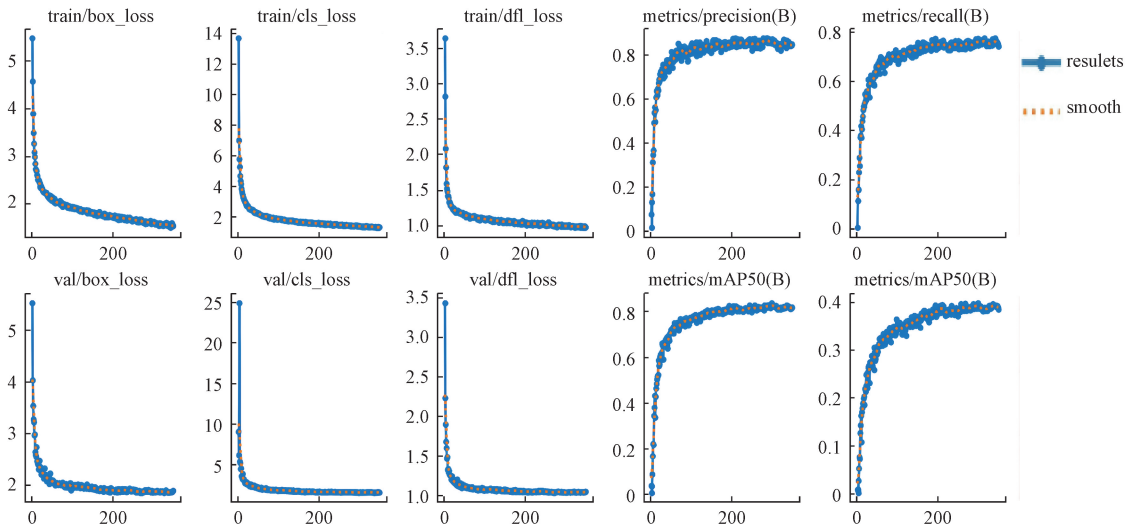


图 7 MEC-YOLOv11n 在 FloW-Img 训练与验证集上损失曲线和性能指标曲线图

Fig. 7 MEC-YOLOv11n loss curve and performance indicator curve on FloW-Img training and validation sets

在验证过程中,MEC-YOLOv11n 的精确率 P、召回率 R、mAP50 和 mAP50-95 收敛曲线如图所示,横坐标表示迭代次数。从图 8 中可以看出,MEC-YOLOv11n 在

FloW-Img 数据集上的表现优于 YOLOv11n。具体来说,经过 200 轮迭代后,召回率(R)逐渐稳定在 77%左右, mAP50 稳定在约 83%,而 mAP50-95 接近 40%。

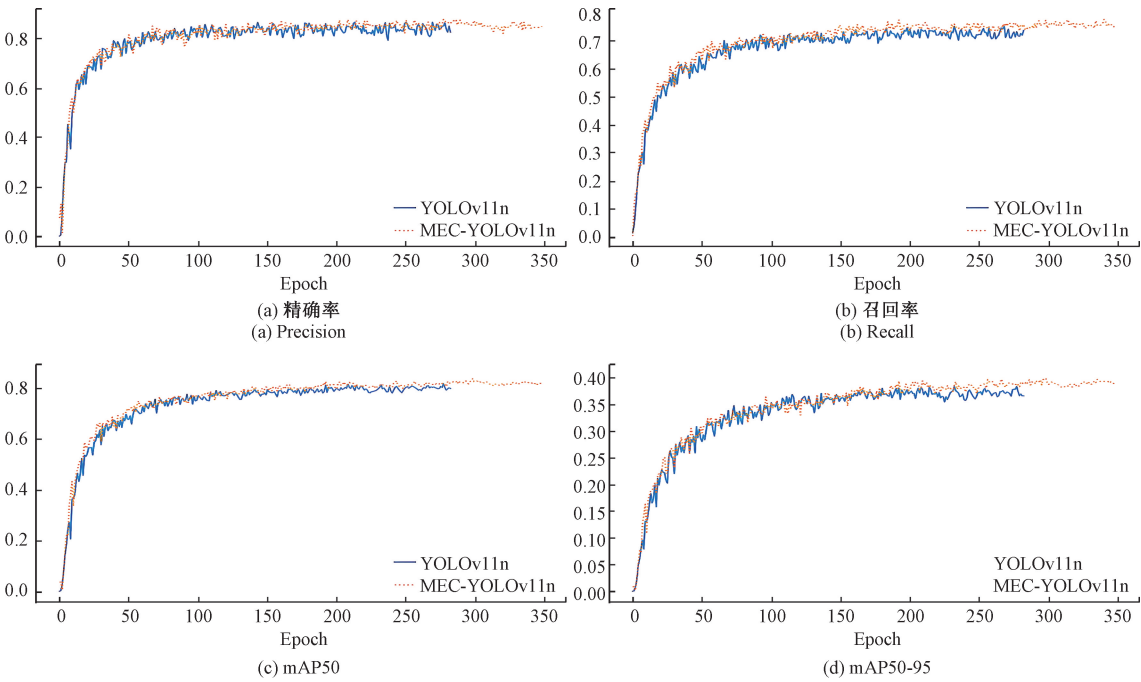


图 8 精确率 P、召回率 R、mAP50 和 mAP50-95 曲线对比图

Fig. 8 Precision P, recall R, mAP50 and mAP50-95 curve comparison chart

此外,为了更直观地展示本文提出的算法在检测过程中对目标关注度的优化效果,采用热力图的形式对检测结果进行可视化。通过热力图(如图 9 所示),可以清晰地观

察到算法在处理过程中对不同区域的关注程度,特别是在关键目标的识别与定位方面的表现。热力图通过颜色变化直观地反映了模型在不同区域的关注强度,其中蓝色和

绿色表示模型的关注度较低或激活值较弱,而黄色和红色则表示模型的关注度较高或激活值较强,这种颜色的层次

对比能够突出模型在目标区域的聚焦情况,从而有效评估算法在复杂背景下的检测性能和准确性。

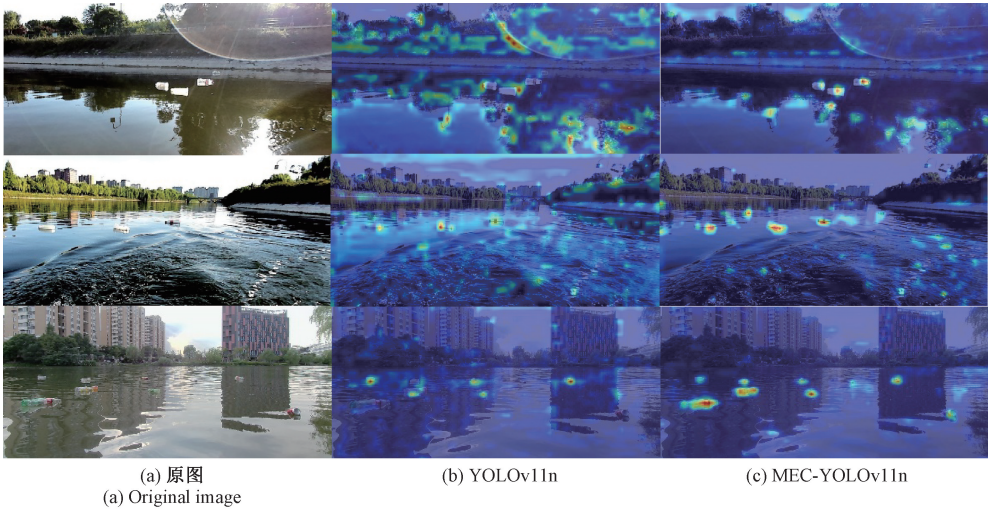


图 9 不同算法热力图与原图对比结果

Fig. 9 Comparison results of heat maps of different algorithms and original images

3.6 对比试验

为了验证 MEC-YOLOv11n 在复杂水平面小目标漂浮物检测中的性能,在实验环境和数据集保持不变的情况下,选取了最新的目标检测算法 YOLOv11n 以及主流的目标检测算法,如 YOLOv3-tiny、YOLOv5n、YOLOv6n、YOLOv8n、YOLOv9-tiny、YOLOv10n 和 RT-DETR-R18 进行对比实验,实验结果如表 3 所示。

表 3 对比实验结果对比  
Table 3 Comparison of test results

算法	P	R	mAP50	mAP50-95
YOLOv3-tiny	0.807	0.579	0.689	0.33
YOLOv5n	0.846	0.748	0.812	0.39
YOLOv6n	0.838	0.747	0.805	0.389
YOLOv8n	0.855	0.752	0.821	0.383
YOLOv9-tiny	0.831	0.76	0.818	0.393
YOLOv10n	0.852	0.752	0.815	0.375
YOLOv11n	0.82	0.751	0.812	0.384
RT-DETR-R18	0.818	0.772	0.828	0.377
本文	0.853	0.775	0.837	0.399

除精确率 P 以外,本文提出的模型在召回率 R、mAP50 和 mAP50-95 等指标上均表现出明显的优势。从召回率 R 的对比来看,MEC-YOLOv11n 在 9 种模型中表现最佳,较 YOLOv8n、YOLOv9-tiny、YOLOv10n、YOLOv11n 和 RT-DETR-R18 分别高出 2.3%、1.5%、2.3%、2.4%和 0.3%,表明我们的模型在减少漏检方面具有优势,能够识别出更多的正类样本;从 mAP50 和 mAP50-95 的对比来看,MEC-YOLOv11n 同样在 9 种模型

中排名最高,较 YOLOv8n、YOLOv9-tiny、YOLOv10n、YOLOv11n 和 RT-DETR-R18 在 mAP 指标上分别高出 1.6%、1.9%、2.2%、2.5%和 0.9%,在 mAP50-95 指标上分别高出 1.6%、0.6%、2.4%、1.5%和 2.2%,这些结果表明,MEC-YOLOv11n 在检测精准度方面具有最优表现。

3.7 扩展实验验证

针对公开的 FloW-IMG 数据集中目标场景单一的问题,通过网络搜索和实地拍摄,自行构建了水面 Trash 数据集。该数据集包含 3 090 张图像,涵盖 11 种类别,共标注了 8 101 个实例,如图 10、11 所示。

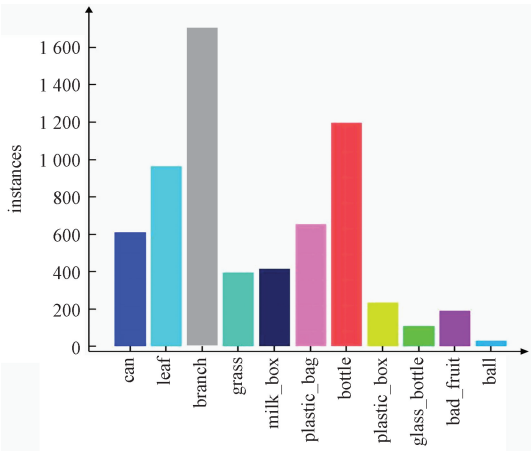


图 10 数据标签分布  
Fig. 10 Data label distribution

为了更好地验证 MEC-YOLOv11n 算法在 Trash 数据集上的有效性,在保持实验参数基本一致的前提下,我们将 Trash 数据集按 8:1:1 的比例划分为训练集、验证集和测试集,并对 YOLOv11n 和 MEC-YOLOv11n 算法进行了



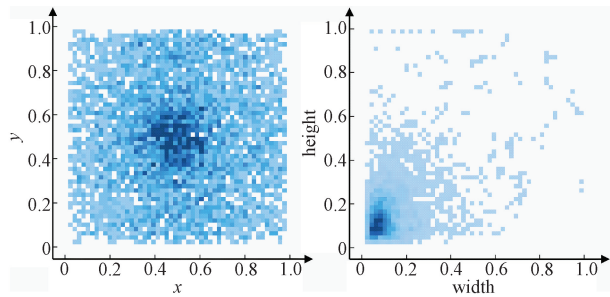
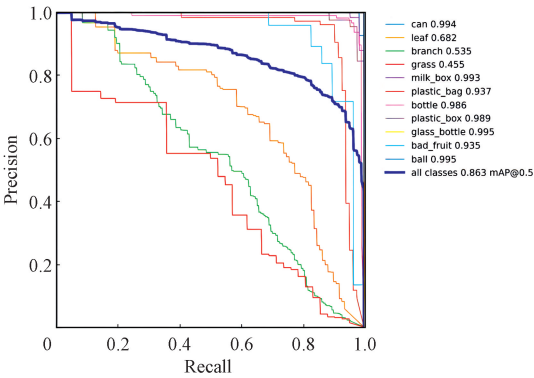


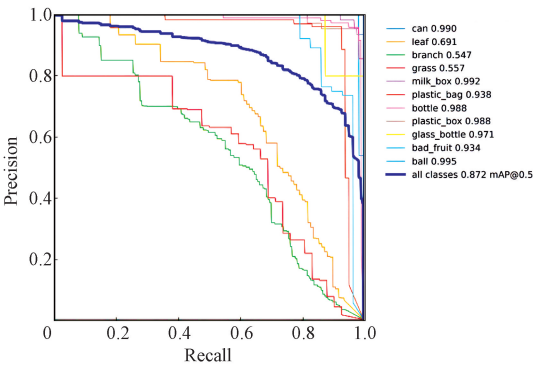
图 11 数据集标签坐标及宽高分布

Fig. 11 Dataset label coordinates and width and height distribution

训练和测试。图 12(图例数字表示 mAP)为 YOLOv11n 和 MEC-YOLOv11n 的 P-R 曲线图。



(a) YOLOv11n P-R 曲线  
(a) YOLOv11n P-R curve



(b) MEC-YOLOv11n P-R 曲线  
(b) MEC-YOLOv11n P-R curve

图 12 P-R 曲线对比图

Fig. 12 Curve comparison chart

从 P-R 曲线图中可以观察到,规则几何体(如罐体 can、奶盒 milk\_box)的曲线几乎紧贴右上角,表明其检测性能优越,主要得益于其形态规则且纹理特征稳定;而植被类目标(如叶片 leaf、枝干 branch、草 grass)的曲线在初始阶段明显偏左,并且随着召回率的提升迅速下降,反映出植被类目标易受边缘模糊和背景干扰的影响。尽管如此,MEC-YOLOv11n 算法在植被类目标的 mAP 值始终优于 YOLOv11n 算法,且 MEC-YOLOv11n 的整体平均曲线

(蓝色粗线)向右上方的位移更加显著,进一步验证了其在综合性能上的优势。

## 4 结 论

针对水上小目标漂流物受环境因素影响和检测精度较低等问题,本文提出了一种改进的 MEC-YOLOv11n 算法,以提升检测性能。首先设计了 C3k2MSWTC 模块,通过扩大模型在识别小目标特征时的感受野,增强了对小目标的检测能力;其次采用了高效的上采样模块 EUCB,改善了不同尺度特征图中的目标轮廓,使得细节更清晰;最后引入了 CCA 模块,进一步提升了对小目标细节纹理的捕捉能力。通过这些优化,本文提出的 MEC-YOLOv11n 算法在水上小目标漂流物检测中展现了显著优势。

FlOW-IMG 数据集的实验结果表明,与其他现有算法相比,MEC-YOLOv11n 算法在精度和鲁棒性方面显著提升,特别是在水上小目标漂流物场景下,检测精度和细节识别能力表现尤为突出。

## 参考文献

[1] 郭庆梅,刘宁波,王中训,等.基于深度学习的目标检测算法综述[J]. 探测与控制学报,2023,45(6):10-20,26.  
GUO Q M, LIU N B, WANG ZH X, et al. Review of target detection algorithms based on deep learning[J]. Journal of Detection and Control, 2023, 45(6): 10-20,26.  
[2] 毛少华,王文东.基于深度学习的 YOLO 系列物体检测算法研究综述[J]. 延安大学学报(自然科学版),2024,43(2):88-95.  
MAO SH H, WANG W D. Review of research on YOLO series object detection algorithms based on deep learning [J]. Journal of Yan' an University (Natural Science Edition), 2024, 43(2): 88-95.  
[3] 张杨,程正东,朱斌.基于改进 Faster R-CNN 的无人机小目标检测算法[J]. 量子电子学报,2022,39(3):354-363.  
ZHANG Y, CHENG ZH D, ZHU B. UAV small target detection algorithm based on improved Faster R-CNN[J]. Journal of Quantum Electronics, 2022, 39(3): 354-363.  
[4] ZHANG C, LIU L, ZANG X, et al. Detr++: Taming your multi-scale detection transformer[J]. ArXiv preprint arXiv:2206.02977, 2022.  
[5] 马梁,苟于涛,雷涛,等.基于多尺度特征融合的遥感图像小目标检测[J]. 光电工程,2022,49(4):49-65.  
MA L, GOU Y T, LEI T, et al. Small object detection based on multi-scale feature fusion using remote sensing images[J]. Opto-Electronic Engineering, 2022, 49(4):

- 49-65.
- [6] QIAO S, CHEN L C, YUILLE A. Detectors: Detecting objects with recursive feature pyramid and switchable atrous convolution [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 10213-10224.
- [7] 赵子婧, 刘宏哲, 曹东璞. 基于 Libra R-CNN 改进的交通标志检测算法[J]. 机械工程学报, 2021, 57(22): 255-265.
- ZHAO Z J, LIU H ZH, CAO D P. Improved traffic sign detection algorithm based on Libra R-CNN[J]. Journal of Mechanical Engineering, 2021, 57(22): 255-265.
- [8] DAI X, CHEN Y, YANG J, et al. Dynamic detr: End-to-end object detection with dynamic attention [C]. IEEE/CVF International Conference on Computer Vision, 2021: 2988-2997.
- [9] 陈旭, 彭冬亮, 谷雨. 基于改进 YOLOv5s 的无人机图像实时目标检测[J]. 光电工程, 2022, 49(3): 69-81.
- CHEN X, PENG D L, GU Y. Real-time object detection for UAV images based on improved YOLOv5s [J]. Opto-Electronic Engineering, 2022, 49(3): 69-81.
- [10] 戚玲珑, 高建瓴. 基于改进 YOLOv7 的小目标检测[J]. 计算机工程, 2023, 49(1): 41-48.
- QI L L, GAO J L. Small object detection based on improved YOLOv7[J]. Computer Engineering, 2023, 49(1): 41-48.
- [11] 张娜, 包梓群, 罗源, 等. 改进的 Cascade R-CNN 算法在目标检测上的应用[J]. 电子学报, 2023, 51(4): 896-906.
- ZHANG N, BAO Z Q, LUO Y, et al. Application of improved Cascade R-CNN algorithm in target detection[J]. Journal of Electronics, 2023, 51(4): 896-906.
- [12] ZHU X, SU W, LU L, et al. Deformable detr: Deformable transformers for end-to-end object detection[J]. ArXiv preprint arXiv:2010.04159, 2020.
- [13] XIAO Z, WAN F, LEI G, et al. FL-YOLOv7: A lightweight small object detection algorithm in forest fire detection[J]. Forests, 2023, 14(9): 1812.
- [14] 张华卫, 张文飞, 蒋占军, 等. 引入上下文信息和 Attention Gate 的 GUS-YOLO 遥感目标检测算法[J]. 计算机科学与探索, 2024, 18(2): 453-464.
- ZHANG H W, ZHANG W F, JIANG ZH J, et al. GUS-YOLO remote sensing target detection algorithm introducing context information and Attention Gate[J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(2): 453-464.
- [15] 徐尽达, 陈慈发, 张上. 基于轻量级算法的水上垃圾小目标检测研究[J]. 电子测量技术, 2024, 47(18): 145-154.
- XU J D, CHEN C F, ZHANG SH. Research on detection of small targets of garbage on water based on lightweight algorithm [J]. Electronic Measurement Technology, 2024, 47(18): 145-154.
- [16] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements[J]. ArXiv preprint arXiv:2410.17725, 2024.
- [17] REIS D, KUPEC J, HONG J, et al. Real-time flying object detection with YOLOv8 [J]. ArXiv preprint arXiv:2305.09972, 2023.
- [18] WANG AO, CHEN H, LIU L, et al. Yolov10: Real-time end-to-end object detection[J]. ArXiv preprint arXiv:2405.14458, 2024.
- [19] 邝先验, 程福军, 吴翠琴, 等. 基于改进 YOLOv7-tiny 的高效轻量遥感图像目标检测方法[J]. 电子测量与仪器学报, 2024, 38(7): 22-33.
- KUANG X Y, CHENG F J, WU C Q, et al. Efficient and lightweight remote sensing image target detection method based on improved YOLOv7-tiny[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(7): 22-33.
- [20] ZHAO Y, LYU W Y, XU SH L, et al. Detrs beat yolos on real-time object detection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [21] 刘子洋, 徐慧英, 朱信忠, 等. Bi-YOLO: 一种基于 YOLOv8n 改进的轻量化目标检测算法[J]. 计算机工程与科学, 2024, 46(8): 1444-1454.
- LIU Z Y, XU H Y, ZHU X ZH, et al. Bi-YOLO: An improved lightweight target detection algorithm based on YOLOv8n [J]. Computer Engineering and Science, 2024, 46(8): 1444-1454.
- [22] FINDER S E, AMOYAL R, TREISTER E, et al. Wavelet convolutions for large receptive fields [C]. European Conference on Computer Vision. Cham: Springer, 2025: 363-380.
- [23] LI L, LI B, ZHOU H. Lightweight multi-scale network for small object detection [J]. PeerJ Computer Science, 2022, 8: e1145.
- [24] OUYANG D L, HE S, ZHANG G ZH, et al. Efficient multi-scale attention module with cross-spatial learning [C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and

- Signal Processing(ICASSP). IEEE, 2023: 1-5.
- [25] WANG C Y, YE H I, MARK LIAO H Y. Yolov9: Learning what you want to learn using programmable gradient information [C]. European conference on computer vision. Cham: Springer, 2025: 1-21.
- [26] JIANG Z, WU B, MA L, et al. APM-YOLOv7 for small-target water-floating garbage detection based on multi-scale feature adaptive weighted fusion [J]. Sensors, 2023, 24(1): 50.
- [27] LUO J, LIU Z, WANG Y, et al. Efficient small object detection you only look once: A small object detection algorithm for aerial images [J]. Sensors, 2024, 24(21): 7067.
- [28] BU Y, YE H, TIE Z, et al. OD-YOLO: Robust small object detection model in remote sensing image with a novel multi-scale feature fusion[J]. Sensors, 2024, 24(11): 3596.
- [29] RAHMAN M M, MUNIR M, MARCULESCU R. Emcad: Efficient multi-scale convolutional attention decoding for medical image segmentation[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 11769-11779.
- [30] CAI X, LAI Q, WANG Y, et al. Poly kernel inception network for remote sensing detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 27706-27716.
- [31] 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]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [32] LIN Z, LENG B. SSN: Scale selection network for multi-scale object detection in remote sensing images [J]. Remote Sensing, 2024, 16(19): 3697.
- [33] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [34] WANG H, QIAN H, FENG S, et al. L-SSD: lightweight SSD target detection based on depth-separable convolution[J]. Journal of Real-Time Image Processing, 2024, 21(2): 33.
- [35] LIN F, HOU T, JIN Q, et al. Improved YOLO based detection algorithm for floating debris in waterway[J]. Entropy, 2021, 23(9): 1111.
- [36] SHENG W, YU X, LIN J, et al. Faster rcnn target detection algorithm integrating cbam and fpn [J]. Applied Sciences, 2023, 13(12): 6913.
- [37] CHENG Y, ZHU J, JIANG M, et al. Flow: A dataset and benchmark for floating waste detection in inland waters [C]. IEEE/CVF International Conference on Computer Vision, 2021: 10953-10962.

## 作者简介

**安梦军**, 硕士研究生, 主要研究方向为人工智能、计算机视觉等。

E-mail: anmengjun1025@163.com

**吴朝明**(通信作者), 博士, 副教授, 硕士生导师, 主要研究方向为人工智能、机器视觉、视觉伺服等。

E-mail: zmwunit@foxmail.com

**邓承志**, 博士, 教授, 硕士生导师, 主要研究方向为图像处理与机器视觉、模式识别与机器学习、多媒体计算技术、高光谱遥感影像处理等。

E-mail: dengcz@nit.edu.cn