

基于改进的 YOLOv11 海上人员搜救的目标检测算法

史娜^{1,3}, 马俊杰^{2,3}, 陈凯源^{2,3}, 周雨聪^{2,3},
景森阳^{2,3}, 杨光^{2,3}, 李凯^{2,3}

(1. 中北大学 数学学院, 太原 030051; 2. 中北大学 信息与通信工程学院, 太原 030051;
3. 极限环境光动态测试技术与仪器全国重点实验室, 太原 030051)

摘要: 针对无人机海上人员搜救任务中复杂环境下目标检测精度与实时性的需求, 对改进 YOLOv11 算法进行了研究; 通过结合风车状卷积优化网络主干, 设计特征增强模块 (FEM) 与自适应权重的双向特征金字塔网络 (BiFPN), 并引入动态注意力机制, 实现了对海上人员微小目标及遮挡目标的特征增强与噪声抑制; 采用 SeaDronesSee 数据集进行实验分析, 测试结果表明, 改进后模型的检测精度 ($mAP@0.5$) 达到 78.47%, 推理速度 (FPS) 为 511.79 Hz, 优于传统的 YOLO 系列算法; 经实际应用验证, 该算法能够满足海上搜救任务的高精度与实时性要求, 为智能化应急救援提供了有效技术支持。

关键词: 海上搜救; 目标检测; 风车状卷积; YOLOv11; 注意力机制

Object Detection Algorithm for Maritime Search and Rescue Based on Improved YOLOv11

SHI Na^{1,3}, MA Junjie^{2,3}, CHEN Kaiyuan^{2,3}, ZHOU Yucong^{2,3},
JING Senyang^{2,3}, YANG Guang^{2,3}, LI Kai^{2,3}

(1. College of Mathematics, North University of China, Taiyuan 030051, China;
2. College of Information and Communication Engineering, North University of China, Taiyuan 030051, China;
3. State Key Laboratory of Extreme Environment Optoelectronics Dynamic Measurement Technology and Instrument, Taiyuan 030051, China)

Abstract: In response to the detection accuracy and real-time performance of targets in complex environments for drone-based maritime search and rescue missions, research has been conducted on improved YOLOv11 algorithm. By integrating the pinwheel-shaped convolution to optimize the network backbone, designing a feature enhancement module (FEM) and a bidirectional feature pyramid network (BiFPN) with adaptive weights, and introducing a dynamic attention mechanism, the feature enhancement and noise suppression for small targets of maritime personnel and occluded targets are achieved. Experiments are conducted on the SeaDronesSee dataset, and the results indicate that the improved model achieves a detection accuracy ($mAP@0.5$) of 78.47% and an inference speed of 511.79 FPS, outperforming traditional YOLO series algorithms. Practical applications show that this algorithm can meet the high-precision and real-time requirements of maritime search and rescue missions and provide an effective technical support for intelligent emergency rescue.

Keywords: maritime search and rescue; object detection; pinwheel-shaped convolution; YOLOv11; attention mechanism

收稿日期:2025-06-05; 修回日期:2025-06-27。

基金项目: 中央引导地方科技发展资金项目 (YDZJSX2025D034, YDZJSX2024D032); 山西省重点研发计划项目 (202202010101007), 山西省科技成果转化引导专项 (202204021301044 和 202304021301028)。

作者简介: 史娜 (1978-), 女, 博士, 副教授。

通讯作者: 杨光 (1978-), 男, 硕士, 讲师。

引用格式: 史娜, 马俊杰, 陈凯源, 等. 基于改进的 YOLOv11 海上人员搜救的目标检测算法[J]. 计算机测量与控制, 2025, 33(8): 54-59, 71.

0 引言

随着“智能航运发展指导意见”的深入推进,海上船舶航行、渔业作业及海上资源开发等活动规模不断扩大,安全事故风险也随之增加,如何快速精准地搜救海上落水人员成为保障生命安全的关键任务。目前,无人机海上救援具有定位精准、长时间滞空、低风险等优势,其搭载的高分辨率光学相机和红外热成像仪能够在复杂环境中精准定位被困人员,并且可以长时间的空中监控和作业,能够在海上救援中发挥重要作用。

针对海上目标的检测,传统的目标检测方法主要包括基于船舶雷达^[1]、合成孔径雷达^[2]、红外热成像^[3]、可见光成像^[4]、声呐^[5]的检测方法。由于海上波浪的干扰,基于雷达的目标检测方法容易漏检小目标;而红外热成像在温差较小时会导致目标检测失效;可见光成像检测的算法主要依赖光照,虚警率较高;声呐检测的部署复杂,只适用于距离较短的目标。而海上目标检测主要以船舶为研究对象,以海上落水人员为对象的相关研究较少,究其原因主要是落水人员目标一般在 10×10 像素以下,且易受光照、波浪、天气(如雾、雨、夜暗)等条件影响,落水人员的目标特征难以准确提取与识别。近年来,基于计算机视觉的智能目标识别算法因其高效的实时检测能力备受关注。

2016年,YOLO模型^[6-8]通过FPN/PANet结构和浅层高分辨率网格,有效捕捉小目标的细节特征,显著提升小目标检出率。通过在YOLOv7模型的主干中加入ResNet-ACmix^[9],以及在Head部分增加ACmix,有效提取了复杂场景下的落水人员小目标的特征信息。文献[10]提出一种改进的YOLOv3目标检测算法,采用k-means聚类算法对自有落水人员数据集进行聚类,得到更适合落水人员的锚框,从而提高检测速度与精度。文献[11]对YOLOv5模型进行改进,通过新增小目标检测层以增强型路径聚合网络结构,并引入坐标注意力机制提升小目标特征融合能力。文献[12]有效解决了印刷电路板缺陷检测中的背景干扰问题。文献[13]利用改进的YOLOv5模型和车辆安装图像对路面裂缝进行了检测,加权注意力机制增强了裂缝的聚焦,同时优化了网络效率。文献[14]应用YOLOv5-IMPROVEMENT模型对焊接射线图的智能缺陷进行识别,显著提高检测效率并降低误判率。文献[15]将YOLO模型集成到无人机(UAV)中,以实现铺砖人行道的实时、远程、可靠和快速的裂缝检测。

针对YOLO系列方法的改进模型广泛应用在智能交通与安防防控方面,包括遥感图像和红外图像的船舶检测^[16-18],多尺度和遮挡感知的面部检测^[19],以及智能城市交通管理的行人和车辆检测^[20]。

基于以上分析,针对海上落水人员检测精度较低和漏检的问题,提出改进的YOLOv11混合模型,通过在网络Backbone部分引入风车状卷积(PConv, pinwheel-shaped convolution)^[21]模块,增强海上落水人员目标的特征信息,在Neck部分采用加权双向特征金字塔(BiFPN, bidirectional feature pyramid network)替换PANet模块,提升多尺度特征融合能力以及落水人员识别的准确率。

1 海上小目标检测方法

1.1 YOLOv11 模型

YOLO(You Only Look Once)是一种基于深度学习的目标检测算法,其主要特点是将目标检测任务视为一个回归问题,直接从图像像素到边界框坐标和类概率的映射,该方法使YOLO能够平衡实时性和准确性的要求。YOLOv11作为该系列的最新成员,特别适合于复杂场景下的海难船舶和落水人员的实时检测任务。

YOLOv11的网络架构包含输入(Input)、主干网络(Backbone)、颈部(Neck)和检测头(Head)4个主要部分。相对于YOLOv8模型,在主干部分用3层卷积核结构的C3K2代替C2f模块,采用跨层特征复用机制使C3K2模块允许特征图在3个不同尺度上进行交互融合,这样既优化了计算效率又增强了特征提取能力;在SPPF模块后面加入一个C2PSA注意力机制模块,通过采用双分支并行结构,同时处理通道和空间两个维度的注意力,通过可变形卷积自适应聚焦关键区域;而在检测头内部增加了两个深度可分离卷积(DW-Conv)模块,这样可以大幅减少参数量和计算量。

针对落水人员这类小目标检测任务,YOLOv11通过增强的特征提取和融合能力,能够更好地应对复杂水域环境中的各种干扰因素,如波浪反射、光线变化等,为实际应用提供更可靠的解决方案。后续研究可以针对复杂应用场景,对YOLOv11的主干和颈部模块进一步改进,以获得更好的网络检测性能。

1.2 PConv 风车状卷积模块

风车状卷积是一种新颖的卷积(PConv),将替代主干网络低层的标准卷积,它能够更好地与微弱小目标的像素高斯空间分布对齐,虽然参数量有少量增加,但是通过增加感受野可以增强小目标的特征提取能力。

PConv采用非对称填充模式,为图像的不同区域创建水平和垂直的卷积核,并且卷积核可以向外扩散。为增强训练稳定性和速度,PConv在第一层的左、右、上、下4个方向分别进行3个像素或1个像素的填充,再进行并行卷积运算,在每次卷积后应用批归一化(BN)和线性单元(SiLU)。对第一层交错卷积的输出特征进行拼接但不填充图像区域,通过 2×2 的卷积核

进行归一化。在网络主干低层将 Conv 替换成 PConv, 可以作为一种通道注意力机制分析 4 个不同卷积方向的贡献。当所检测微小目标的特征较集中时, 类似于高斯分布, PConv 能以最小的参数增加来实现高效的感受野扩展, 使视觉感受野的有效性和卷积次数会向外逐渐减小。PConv 增强了微小目标与复杂背景之间的对比度, 同时抑制了杂波样信号。

1.3 BiFPN 模块

双向特征金字塔网络 (BiFPN)^[22] 是一种高效的多尺度特征融合模块, 通过双向跨尺度连接和加权特征融合机制, 显著提升了目标检测模型对多尺度目标的识别能力。通过结合自上而下与自下而上的双向信息流, 并引入可学习权重动态平衡不同层级特征的贡献, 既保留了高层语义信息, 又增强了底层细节特征。相比传统 FPN 或 PANet, BiFPN 通过轻量化设计能够保持高效的计算, 尤其对小目标检测效果显著。

2 改进的 YOLOv11 模型

2.1 改进的 PConv 特征提取模块

海上落水人员的搜救任务受到复杂海洋环境和多变天气条件的影响, 导致人员和小型船只的目标识别非常困难。虽然 YOLOv11 模型在目标检测的实时性和准确性方面已经具有较优的成绩, 但海上人员的目标通常体积十分微小, 且受海面光斑的影响, 增加了漏检的可能性。而风车状卷积在检测微小目标时, 可以突出微小目标的特征, 因此将 YOLOv11 主干网络中的 Conv 替换为 PConv, 如图 1 所示。

PConv 模块的具体算法如下:

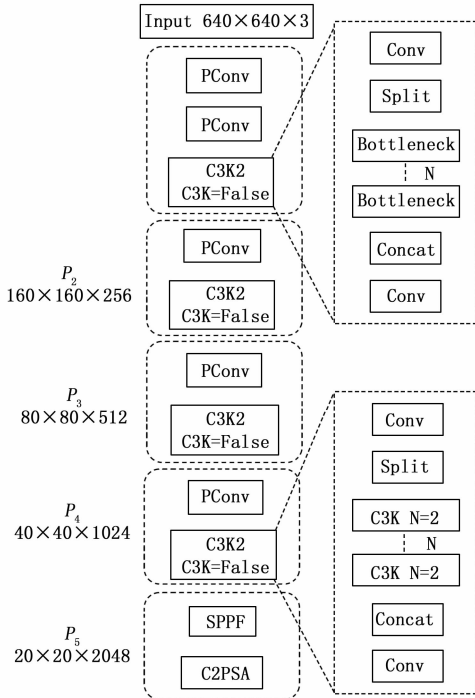


图 1 海上人员检测模型的主干框架

假设输入图像为 $X^{(h_1, w_1, c_1)}$, h_1 、 w_1 、 c_1 分别表示输入图像的高度、宽度和通道数。在水平和垂直的 4 个方向的卷积计算为:

$$X_1^{(h', w', c')} = \text{SiLU}\{\text{BN}[X_{P(1,0,0,3)}^{(h_1, w_1, c_1)} \otimes W_1^{(1,3,c')}] \};$$

$$X_2^{(h', w', c')} = \text{SiLU}\{\text{BN}[X_{P(0,3,0,1)}^{(h_1, w_1, c_1)} \otimes W_2^{(3,1,c')}] \};$$

$$X_3^{(h', w', c')} = \text{SiLU}\{\text{BN}[X_{P(0,1,3,0)}^{(h_1, w_1, c_1)} \otimes W_3^{(1,3,c')}] \};$$

$$X_4^{(h', w', c')} = \text{SiLU}\{\text{BN}[X_{P(3,0,1,0)}^{(h_1, w_1, c_1)} \otimes W_4^{(3,1,c')}] \}$$

其中: $\otimes$ 是卷积运算, BN 表示批量归一化, SiLU 为深度学习网络的激活函数, $W_1^{(1,3,c')}$ 是一个 1×3 的卷积核, 输出通道为 c' , 填充参数 P (1, 0, 0, 3) 分别表示在左、右、上、下方向的填充像素数, 其他填充参数类似。

将第一层交错卷积的结果进行拼接 $\text{Cat}(\cdot, \cdot)$, 输出计算结果为:

$$X'^{(h', w', c')} = \text{Cat}[X_1^{(h', w', c')}, \dots, X_4^{(h', w', c')}]$$

通过卷积核 $W^{(2,2,c_2)}$ 对 $X'^{(h', w', c')}$ 进行批量归一化, 但不进行水平与垂直填充。输出特征图的高度为 h_2 , 宽度为 w_2 , 最后的输出 $Y^{(h_2, w_2, c_2)}$ 为:

$$Y^{(h_2, w_2, c_2)} = \text{SiLU}\{\text{BN}[X'^{(h', w', c')} \otimes W^{(2,2,c_2)}]\}$$

结合风车状卷积对 YOLOv11 模型的改进算法如下:

$$\text{Backbone}(X) = \{P_2, P_3, P_4, P_5\}$$

其中: $X \in R^{H \times W \times 3}$ 为输入图像, $P_k \in R^{H/2^k \times W/2^k \times C_k}$ 为 k 级特征图:

$$P_2 = \text{PConv}(X);$$

$$P_3 = \text{C3K2}[\text{Downsample}(P_2)];$$

$$P_4 = \text{C3K2}[\text{Downsample}(P_3)];$$

$$P_5 = \text{SPPF}\{\text{C2PSA}[\text{Downsample}(P_4)]\}$$

由于海上人员搜救为微小目标检测, 在多尺度特征强化阶段需要保留高分辨率 P_2 特征层, 同时各个阶段的局部网络中都采用风车状卷积增强微小目标的特征信息。本文的海上人员搜救算法采用改进的骨干架构, 能够增强微小目标的特征提取能力, 提高人员和船只检测的精确度和复杂任务的表现。相较于 YOLOv8 模型, 改进的 YOLOv11 模型将 CF2 模块改成 C3K2, 同时在 SPPF 模块后面添加了一个 C2PSA 模块, 该轻量化网络使用深度可分离的方法, 减少冗余计算, 从而提高海上人员微小目标的检测效率。

2.2 改进的 BiFPN 特征融合模块

针对海上人员微小目标在深层特征中易丢失的问题, 本文在 Neck 部分设计了一种基于多尺度特征增强的双向特征金字塔网络架构。在特征融合阶段, 结合 FEM (Feature Enhancement Module) 特征增强模块, 通过并行跨阶段特征增强来捕获微小人员目标的语义信息, 显著提升特征图的判别能力。首先, FEM 模块通过多分支结构和扩展卷积来增强海上人员小目标的局部

上下文信息,对各个阶段提取到的特征图 $E_{out} \in R^{H \times W \times C}$ 进行多尺度卷积运算,不同尺度的卷积核为 $K_1, K_2, \dots, K_n$, 增强后的特征图为 $Y_1, Y_2, \dots, Y_n$, 即 $Y_i = E_{out} \odot K_i + b_i$ ($i=1, 2, \dots, n$)。而 BiFPN 模型采用自上而下 (Top-Down) 和自下而上 (Bottom-Up) 的方式进行特征融合。这意味着每个特征图不仅可以从上层特征图中获取目标信息,也可以从下层特征图中获取细节信息,从而增强海上人员小目标特征的表达能力。同时,通过引入动态特征权重自适应机制,使网络自适应地调整不同特征图的权重系数,从而在特征金字塔的每一层级实现智能化的特征选择与融合。一方面,高层特征提供的丰富上下文信息通过权重分配得到有效传递;另一方面,底层特征包含的精细空间细节通过动态加权得以保留。这种智能化的特征融合机制,能显著提升深度学习网络对复杂背景下的海上人员微小目标的表征能力和检测性能,如图2所示。

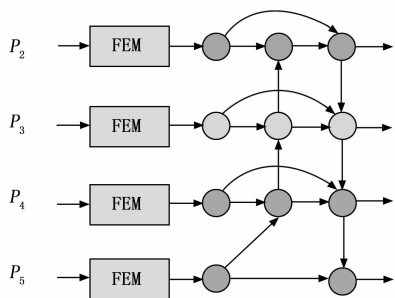


图2 海上人员检测模型的颈部框架

2.3 分类损失函数联合优化

海上人员小目标具有低对比度、光斑噪声干扰、复杂背景和小目标特征多样性等因素,目前常用的两种方法是重采样和重加权。重采样主要是通过生成新的样本集来增强模型评估和处理不平衡数据的能力。而重加权则是通过调整样本的权重来影响模型训练的结果,常用于处理不平衡数据和样本重要性调整,比较适合微小目标的检测。

Focal Loss 函数就是一种重加权方法,通过引入调节因子来解决训练过程中不同类别样本不均衡的问题。分类损失函数的计算如下式所示:

$$Loss = -\alpha(1 - p_i)^\gamma \log(p_i)$$

其中: p_i 表示正确类别的预测概率, α 为平衡因子, γ 表示调解因子。

EMA SlideLoss 是一种结合了指数移动平均和平滑机制的分类损失函数,旨在减轻模型对训练数据的过拟合。通过动态调整损失,适用于红外小目标类别不平衡和多类别分类任务,它提高模型的鲁棒性和泛化能力,具体的函数表示如下:

$$EMA_i = \alpha x_i + (1 - \alpha) EMA_{i-1}$$

其中: x_i 为当前时刻的损失值, α 为平滑因子, EMA_i 为当前时刻的 EMA 值。

结合 Focal Loss 和 EMASlideLoss 函数,设计的综合分类损失函数为:

$$ourLoss = \beta Loss + (1 - \beta) EMA_i$$

其中: $\beta \in (0, 1)$ 表示学习权重。

3 实验及结果分析

3.1 数据集

SeaDronesSee v2 数据集是一个专注于海洋环境搜索与救援场景的水上无人机大规模数据集,其涵盖了不同光照条件、海况和飞行高度下的航拍图像,包含近海、远海、港口等多种环境场景,并标注了游泳者、船只、摩托艇、救生设备、浮标五类关键目标的位置和类别,尤其关注小目标检测。该数据集包含了 8 930 张训练图像、1 547 张验证图像和 3 750 张测试图像,小目标 (小于 20×20 像素) 占比达 70%, 每张图像都经过人工标注,旨在为海洋监测、搜索与救援、以及海洋生态研究等应用提供丰富的数据支持。

3.2 实验环境

在 Python 语言上对本文提出的 PConv-YOLOv11 海上落水人员的目标检测方法进行实验,使用 NVIDIA GeForce GTX 4070TISURPER (16 GB 内存) 作为神经网络的训练平台,操作系统 Windows11。在网络训练过程中,根据图像的比例和裁剪,将每幅图像的大小调整到 640×640 , 对于 3 个通道的输入大小, PConv-YOLOv11 海上人员搜救的目标检测模型的每秒帧数 (FPS) 和参数计算成本分别为 511.19 Hz 和 8.7 MB。Adam 优化器用于批处理大小为 16, 训练轮次 epoch 为 300 的网络优化,初始学习率为 0.01。在网络训练之前,实验设定训练集为 8 930, 验证集为 1 547。同时考虑到样本图像中存在大量微小目标,以及在落水人员检测过程中平衡实时性能和准确性,将样本归一化为 640×640 。

3.3 评价指标

为验证模型的性能,本文采用准确率 (P)、召回率 (R)、平均精度均值 (mAP) 和模型参数量 Pare (M) 来评估模型检测和分类的准确性。通过计算受试者工作特征 (ROC) 曲线表明,该指标的值越高,意味着海上落水人员小目标检测任务的性能越好,而平均精度均值 (mAP) 越高,则目标定位的准确性越高,其中 P, R, AP 和 mAP 表达式为:

$$P = \frac{TP}{TP + FP}; R = \frac{TP}{TP + FN};$$

$$AP = \frac{TP + TN}{TP + TN + FP + FN};$$

$$mAP = \frac{\sum_{n=1}^{\text{Num(classes)}} AP(n)}{TP + TN + FP + FN}$$

其中： TP 是预测正样本中 IoU 大于 0.45 阈值的真阳性样本量， TN 为识别正确的负样本数量， FP 为预测正样本中的假阳性样本量， FN 为预测负样本中的假阴性样本量，实验使用的 SeaDronesSee v2 数据集上分类数为 5。

3.4 消融实验

本研究通过系统的消融实验验证了风车状卷积 (PConv)、特征增强模块 (FEM) 和双向特征金字塔网络 (BiFPN) 对深度学习网络性能的改进效果。实验在 SeaDronesSee v2 数据集上进行，以 YOLOv11 作为基准模型，分别测试了单独引入各模块和组合引入 3 种模块对海上人员目标检测性能的影响。为了全面评估各模块的贡献，设计了 5 组对比实验。首先在基准模型上单独测试每个改进模块的效果，随后将 3 个模块组合测试其协同作用。实验采用“√”标记法直观展示各模型的改进配置，如表 1 所示。

实验结果表明，仅引入 PConv 的模型 1 相比基准模型在 $mAP@0.5$ 指标上提升了 1.13%，这主要得益于 PConv 有效减少了特征提取过程中的信息冗余。值得注意的是，虽然模型 1 的参数量增加了 0.5 M，召回率略有下降，但由于 P2 层更专注于微小目标特征的提取，其计算复杂度 (GFLOPs) 反而有所降低。

仅引入 FEM 模块的模型 2 在检测准确率上有较快提升。相较于基准 YOLOv11 模型，该改进模型在 $mAP@0.5$ 和 $mAP@0.5 : 0.95$ 指标上分别提高了 5.24% 和 8.02%，同时召回率也上升了 3.94%。这种性能提升主要归因于 FEM 模块在多尺度特征提取过程中的优化作用，该模块通过增强特征表示能力，有效降低了不同尺度特征间的语义差异，从而显著改善了模型对海上微小目标的检测效果。

仅采用 BiFPN 特征融合网络的模型 3 准确率提高到 87.38%。该模型在 $mAP0.5$ 和 $mAP0.5 : 0.95$ 两个关键指标上分别实现了 12.84% 和 15.96% 的提升幅度。这种性能飞跃主要得益于 BiFPN 网络独特的并行跨阶段双向特征融合机制，该机制能够有效捕捉海上微小目标的深层语义特征，同时显著增强了多尺度特征图的表征能力。尽管模型参数规模增至 7.3 M，但其计算复杂度维持在 7.5 GFLOPs 的轻量级水平，完全满足海上落水人员实时检测的工程应用需求。

由改进的 YOLOv11 模型的实验可知，PConv 模块可以动态调整感受野以适应不同尺度的目标特征，FEM 模块增强海上目标的细节特征，在结合 BiFPN 特征融合网络共同作用时，深度学习网络的整体性能最优。虽然准确率为 83.33% 略低于模型 3，但是 $mAP@0.5$ 值和 $mAP@0.5 : 0.95$ 值相较于基准网络 YOLOv11 分别提升 15.18% 和 17.39%，召回率提升 15.51%，浮点运算总量 GFLOPs 为 8.1，能够满足海上搜救的轻量化与实时性需求。

3.5 对比实验

为验证改进的 YOLOv11 模型在无人机海上救援任务中的网络检测性能，选取用于多尺度目标和复杂场景检测的 YOLOv5、YOLOv8 和 YOLOv10 算法及改进的 YOLOv11 无人机海上救援目标检测算法在 SeaDronesSee v2 数据集上进行实验比较。

实验结果如表 2 所示，加粗数据为该列的最优值。由表 2 可见，在 $mAP@0.5$ 方面达到 78.47%，改进的 YOLOv11 比 YOLOv10 高出 1.9%，比 YOLOv5 高出 14.48%，而 $mAP@0.5 : 0.95$ 值仍达到 54.32%，虽然实时检测速度为 511.79 帧/秒，略逊色于 YOLOv5 的 522.98 帧/秒，实验说明改进的 YOLOv11 在检测精度和实时性方面有非常不错的表现。

表 1 改进模块的消融实验结果对比

模型	PConv	FEM	BiFPN	$P/\%$	$R/\%$	$mAP@0.5$	$mAP@0.5 : 0.95$	$Pare/M$	$GFLOPs$
YOLOv11	×	×	×	77.33	59.00	63.29	36.93	5.5	5.6
模型 1	√	×	×	81.15	58.73	64.42	38.38	6.0	5.5
模型 2	×	√	×	81.54	62.94	68.53	44.95	6.7	6.3
模型 3	×	×	√	87.38	67.53	76.13	52.89	7.3	7.5
改进的 YOLOv11	√	√	√	83.33	74.51	78.47	54.32	8.7	8.1

表 2 改进的 YOLOv11 模型与其他先进方法在 SeaDronesSeev2 数据集上检测性能对比

Methods	swimmer	Boat	jetski	life_saving_appliances	buoy	$mAP@0.5$	$mAP@0.5 : 0.95$	FPS
YOLOv5	66.46	91.70	91.89	23.50	56.40	63.99	37.92	522.98
YOLOv8	69.17	90.67	81.84	21.73	59.09	65.50	38.45	498.20
YOLOv10	75.15	93.84	89.11	44.35	81.42	76.57	52.74	495.44
改进的 YOLOv11	75.81	96.16	89.22	49.51	89.14	78.47	54.32	511.79

由于救生圈受海水背景干扰较大, 并且像素尺寸较小, 检测难度明显高于游泳者、船只、摩托艇和浮标, 检测准确率仅为 49.51%, 但明显高于 YOLOv5 的 23.5%。而海上落水人员目标尺寸非常小且数据量大, 而检测难度也显著高于船只、摩托艇和浮标, 检测精度仍达到 75.81%, 优于其他 YOLO 模型的检测方法。海上船只的尺度相对较大, 目标检测相对容易, YOLO 模型整体检测精度能达到 90% 以上, 效果令人满意。

从图 3 第五行的可视化海上小目标检测结果可见, 其他算法出现了多个微小目标的漏检, 这些微小目标与海面光斑高度相似, 而改进 YOLOv11 模型对微小目标的检测效果是十分理想的。

研究表明, 所提出的改进 YOLOv11 模型在无人机海上救援场景中展现出优越的综合性能, 针对最具挑战性的小目标检测的平均精度均值达到 78.47%, 显著优于 YOLOv5、YOLOv8 和 YOLOv10 等传统算法 (可视化对比见图 3)。相较于改进的 YOLOv11 模型在保持轻量化架构优势的同时, 显著提升了多尺度小目标的检测精度。特别值得注意的是, 所提模型对海面耀斑干扰具有更强的鲁棒性, 有效降低了复杂海况下的漏检率。改进后的模型在多目标识别任务中表现出更优的特征提取能力, 对落水人员与船只只有更好的识别能力, 其检测性能的全面提升为海上救援任务提供了可靠的技术支持。

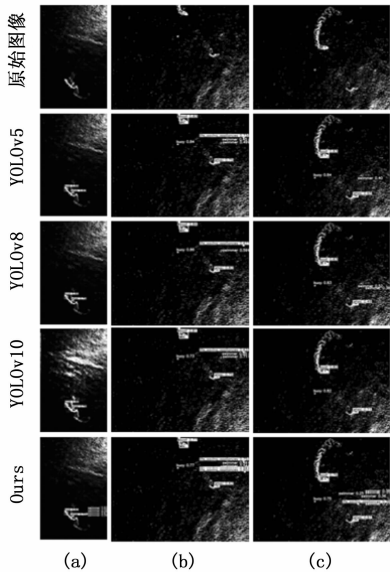


图 3 可视化对比图

4 结束语

针对海上落水人员小目标和多尺度海上目标, 本文提出一种轻量级改进的 YOLOv11 无人机海上救援目标检测算法。本研究提出的改进 YOLOv11 模型通过多项创新设计显著提升了海上救援目标的检测性能。首先,

引入的 PConv 模块采用自适应感受野机制, 有效增强了模型对多尺度特征的表达能力; 其次, 结合轻量级 FEM 模块和 BiFPN 网络的双向特征融合策略, 在减少特征冗余的同时, 显著提升了小目标的细节表征能力; 此外, 精心设计的四层检测结构 (P_2 、 P_3 、 P_4 、 P_5) 进一步强化了多尺度特征提取效果。在 SeaDronesSee v2 数据集上的实验结果表明, 该模型仅以 8.7 M 的参数量就实现了 78.47% 的 $mAP@0.5$ 检测精度, 相比 YOLOv5 提升了 14.48%, 在保持 511.79 FPS 实时处理速度的同时, 展现出卓越的检测性能优势。

在下一步的研究工作中, 可考虑将红外成像等多模态数据融合到现有框架中, 以提升复杂海况下的微小目标检测的鲁棒性; 而针对极端天气条件下的模型泛化能力, 需要构建更具挑战性的海事数据集进行验证, 这些改进将进一步提升算法在实际海上救援任务中的实用价值。

参考文献:

[1] 李家强, 姜庆刚, 陈金立, 等. 船舶导航雷达的海杂波自适应抑制算法 [J]. 现代雷达, 2016, 38 (2): 23 - 26.

[2] 熊伟, 徐永力, 崔亚奇, 等. 高分辨率合成孔径雷达图像舰船目标几何特征提取方法 [J]. 光子学报, 2018, 47 (1): 55 - 64.

[3] 明英, 蒋晶珏. 基于 Cauchy 分布的红外视频运动目标检测 [J]. 红外与毫米波学报, 2008 (1): 65 - 71.

[4] 王志虎, 沈小青, 桂伟龙. 光学成像小目标检测技术综述 [J]. 现代防御技术, 2020, 48 (5): 67 - 73.

[5] 郭戈, 王兴凯, 徐慧朴. 基于声呐图像的水下目标检测、识别与跟踪研究综述 [J]. 控制与决策, 2018, 33 (5): 906 - 922.

[6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016: 779 - 788.

[7] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [J]. Proceedings-30th IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517 - 6525.

[8] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. Arxiv, 2018, Arxiv.org/abs/1804.02767.

[9] 黄旭, 施闰虎, 曾孟佳. 一种基于改进 YOLOv7 的落水人员检测方法 [J]. 现代计算机, 2024, 30 (15): 25 - 31.

[10] 许晓峰, 陈姚节, 刘恒. 基于 YOLOv3 的落水人员检测 [J]. 计算机技术与发展, 2022, 32 (8): 49 - 54.

(下转第 71 页)