液晶与显示
Chinese Journal of Liquid Crystals and Displays

文章编号:1007-2780(XXXX)XX-0001-12

基于改进YOLO11n的海上小目标检测算法

任喜伟*,方明玉,郭紫阳,高　帆
(陕西科技大学 电子信息与人工智能学院,陕西 西安 710021)

摘要:无人机视角下,针对海上小目标检测目标尺度小、背景复杂,易出现漏检率高、检测精度不足的问题,提出了一种改进YOLO11n的海上小目标检测算法——SEW-YOLO。首先,在骨干网络中设计分裂通道注意力瓶颈模块SCAB,通过通道分组策略与坐标注意力机制增强小目标特征提取能力;其次,设计增强快速空间金字塔池化模块SPPF_E,针对小目标场景以小核池化代替大核池化并用高效通道注意力自适应校准多尺度感受野;同时,引入WIoU损失函数替代原始CIoU,提升小目标边界框回归精度。最后,在SeaDronesSee公开数据集上进行了消融实验和对比实验。实验结果表明,SEW-YOLO算法的mAP50和mAP50-95分别达到68.1%和40.6%,较基线模型分别提升2.8和1.5个百分点,保持了模型轻量化特性,实现了检测精度的提升,证明了算法的有效性。

关　键　词:小目标检测;YOLO11n;轻量化;海上搜救;无人机影像
中图分类号:TP391.41　　**文献标识码**:A　　**doi**:10.3724/CJLCD.2026-0150　　**CSTR**:32172.14.CJLCD.2026-0150

A maritime small object detection algorithm based on improved YOLO11n

REN Xiwei*, FANG Mingyu, GUO Ziyang, GAO Fan

(*School of Electronic Information and Artificial Intelligence*, *Shaanxi University of Science & Technology*, *Xi'an* 710021, *China*)

Abstract: To address the high missed-detection rates and insufficient accuracy in UAV-based maritime small object detection caused by small scales and complex backgrounds, we propose SEW-YOLO, an improved YOLO11n-based algorithm. First, a Split Channel Attention Bottleneck(SCAB) module is designed in the backbone, enhancing small object feature extraction through channel grouping and coordinate attention mechanisms. Second, an Enhanced Spatial Pyramid Pooling Fast(SPPF_E) module replaces large-kernel pooling with small-kernel pooling for small object scenarios, using efficient channel attention to adaptively calibrate multi-scale receptive fields. Meanwhile, WIoU loss replaces the original CIoU to improve bounding box regression accuracy for small objects. Finally, ablation and comparative experiments on the publicly available SeaDronesSee dataset show that SEW-YOLO achieves mAP50 of 68.1% and mAP50-95 of 40.6%—improvements of 2.8 and 1.5 percentage points over the baseline,

收稿日期:2026-08-04;**修订日期**:2026-10-29.
基金项目:国家自然科学基金(No.61971272);陕西科技大学科研启动基金(No.2020BJ-01)
Supported by National Natural Science Foundation of China (No.61971272); Scientific Research Launch Fund of Shaanxi University of Science and Technology (No.2020BJ-01)
*通信联系人,E-mail:renxiwei@sust.edu.cn

respectively—while maintaining model lightweight nature, validating the algorithm's effectiveness.
Key words: small object detection; YOLO11n; lightweight; maritime search and rescue; UAV imagery

1 引　　言

基于无人机平台的海上目标检测在海洋监测、搜索救援等领域具有重要的应用价值[1-2]。无人机凭借机动性强、覆盖范围广、部署成本低等优势,已成为海上态势感知的重要工具。然而,在无人机高空视角下采集的影像中,待检测目标(如落水人员、小型船只等)往往仅占极少像素,呈现典型的小目标特征。这类面积占整图不足0.1%的小目标普遍存在尺度微小、纹理稀疏、边界模糊的特征,且易受海面反光、波浪杂波和天气变化的干扰,给准确稳定检测带来严峻挑战。如何在有限计算资源约束下实现海上小目标的高精度实时检测,是当前亟需解决的关键问题。

近年来,围绕小目标检测问题,学界已从感知成像和检测算法开展了系统性研究。在感知成像与计算层面,新型光电技术为弱小目标的探测、成像与处理提供了新手段,Wu等人[3]利用宽带量子压缩感知成像,在信号背景比低至1/332的强背景噪声下实现了10km外无人机目标的被动探测。Dun等人[4]将超表面光学与神经阵列成像模型结合,以大幅缩短的光学总长实现了媲美商用镜头的成像质量,并支持目标检测等视觉任务。Dong等人[5]基于衍射神经网络实现了每秒数百万帧的光学图像处理,为高吞吐、低功耗视觉计算提供了新途径。此外,纳米光子学、结构光与人工智能的结合展现了广阔的应用前景[6-7]。然而,上述成像与计算新技术多面向专用探测或实验系统,短期内难以直接嵌入机载通用视觉平台,在既有成像与算力条件下,算法层面的改进更为直接可行。在检测算法层面,早期以经典图像处理方法为主,许开宇等人[8]结合背景预测与改进Top-Hat变换实现红外弱小目标检测,陈鹏鹏等人[9]提出基于多特征与决策树的海上搜救目标检测算法,解决了单一特征易误检的问题。随着深度学习技术的快速演进,基于卷积神经网络的目标检测方法逐渐成为主流技术路线。研究者围绕多尺度特征融合、注意力机制与轻量化设计等方向提出多种改进方案。Wang等人[10]提出CF-YOLO,通过跨尺度特征金字塔网络与感受野注意力卷积缓解小目标深层特征衰减问题。Zheng等人[11]提出LAM-YOLO,引入光照-遮挡注意力机制与辅助检测头增强密集小目标感知能力。况子晶等人[12]针对海上搜救场景提出SR-Det,结合Transformer解码器与多尺度感受野增强模块,并构建海面聚焦模块抑制波浪与反光干扰。吕学涵等人[13]提出HMD-YOLO,设计高分辨率多尺度卷积注意力模块增强小目标特征表达。杨晓敏等人[14]针对SAR影像小目标检测提出超轻量化检测网络,通过多分支高效层聚合与细节增强共享检测头增强多尺度感知与小目标细节表达,并结合剪枝与知识蒸馏压缩模型,以不足1×10^6的参数量实现了较高检测精度。上述检测方法虽各有提升,但多数在引入新模块的同时增加了计算开销,在无人机有限算力约束下仍面临精度与效率的权衡。

为此,本文提出一种基于改进YOLO11n的海上小目标检测算法,旨在增强小尺度目标特征提取与识别能力,在提升精度的同时兼顾部署适用性,为无人机海上搜救场景提供工程可行的解决方案。本文主要贡献如下:

(1)设计分裂通道注意力瓶颈模块(SCAB)。采用通道分组策略,仅对部分通道施加注意力增强的轻量卷积运算,其余通道恒等传递,在强化小目标空间特征响应的同时有效抑制了冗余计算。

(2)提出增强快速空间金字塔池化模块(SPPF_E)。以小核池化替代原始SPPF的大核池化,缩小最大感受野以降低背景噪声对小目标特征的淹没风险,并引入高效通道注意力机制,自适应校准各感受野分支的通道响应权重,增强多尺度感知鲁棒性。

(3)引入Wise-IoU(WIoU)损失函数替代Complete-IoU(CIoU)。利用离群度感知的梯度增益分配策略,动态抑制低质量锚框对回归训练的干扰,使优化过程聚焦于小目标的高质量边界框回归,提升定位精度。

2 本文算法

2.1 SEW-YOLO整体结构

海面波浪杂波、光照反射等复杂背景经网络提取后在特征通道中形成强响应，易引发虚警，弱小目标特征则被淹没其中。同时，海上远距离成像使目标仅占极少像素、对比度低、纹理稀疏，小目标特征在向深层传递时易被背景淹没。此外，目标边界模糊使标注不确定性高，训练样本中低质量锚框占比大，边界框回归易受干扰。针对上述影响，本文提出轻量化海上小目标检测算法SEW-YOLO，完整网络整体架构如图1所示。

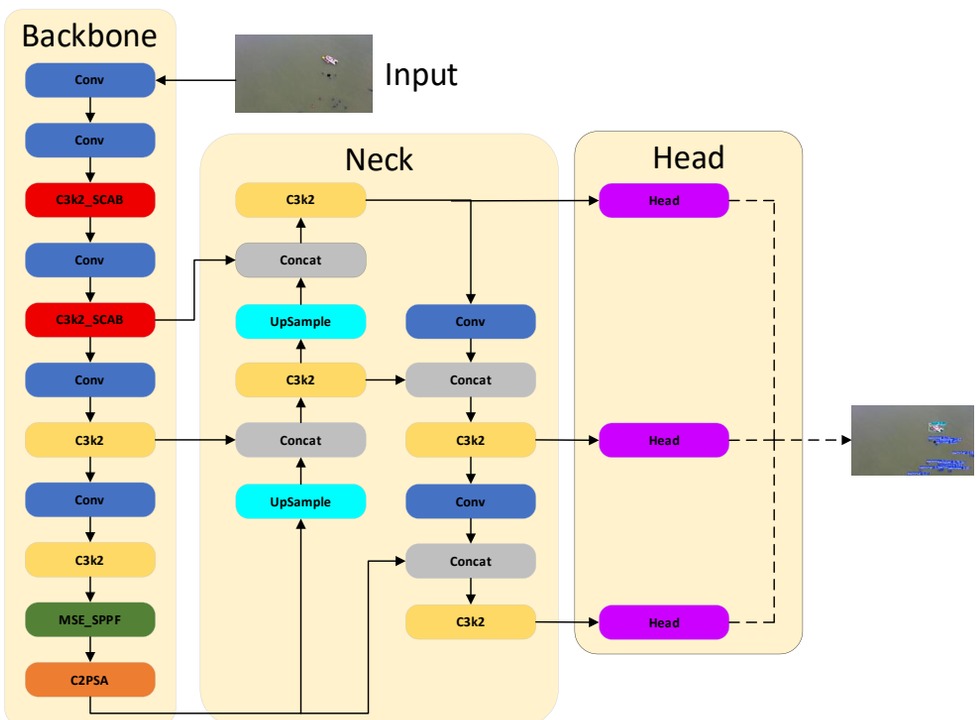


图1 SEW-YOLO结构

Fig. 1 Framework of SEW-YOLO

网络由骨干Backbone、颈部Neck、检测头Head三部分构成。骨干网络采用SCAB模块替换原生Bottleneck，其坐标注意力可增强模型区分小目标与海面杂波的判别能力。同时，骨干网络嵌入SPPF_E模块，以小核池化保留小目标空间细节，并通过高效通道注意力抑制编码海面波浪与水面反光的背景干扰通道。二者协同缓解背景强响应对弱小目标的淹没。训练阶段采用WIoU损失优化边界框回归，利用离群度感知的梯度增益抑制低质量锚框对回归训练的干扰。下文分模块详细介绍三项改进策略。

2.2 分裂通道注意力瓶颈(SCAB)

在YOLO11n的骨干网络中，C3K2模块内部的Bottleneck模块对所有输入通道施加完整的3×3空间卷积，其计算过程如公式(1)所示：

$$[Y = X + \mathrm{Conv3}\times 3\mathrm{Conv3}\times 3X], \quad (1)$$

其中$X \in \mathbb{R}^{B\times C\times H\times W}$为输入特征。对于以小目标为主的海上检测场景，输入特征中包含了大量稀疏的背景信息，对所有通道均匀执行空间卷积会导致不必要的计算冗余。受FasterNet[15]中部分卷积思想的启发——仅对输入的部分通道进行卷积即可保持相当的表示能力，基于此本文提出分裂通道注意力瓶颈模块(Split Channel Attention Bottleneck，SCAB)，图2分别展示了原始Bottleneck基准结构与本文所提SCAB模块的结构示意图。

SCAB的核心设计是将输入特征沿通道维度划分为主分支与保留分支。主分支承担核心特征映射功能：首先通过1×1卷积实现通道降维，随后采用深度卷积完成高效空间特征建模，

最终引入的坐标注意力模块(CA)[16]沿水平、垂直两个空间维度进行特征编码,输出位置感知型通道注意力权重。保留分支恒等传递,不施加任何空间变换。两分支输出拼接后经1×1卷积融合,并引入残差连接。

坐标注意力机制将通道注意力中的空间特征聚合进行维度解耦,分解为沿水平与垂直方向的两个一维特征编码,在捕获跨通道关系的同时保留精确的位置信息。CA模块的实现分为坐标信息嵌入与坐标注意力生成两个步骤。在坐标信息嵌入阶段,给定输入特征图$X\in\mathbb{R}^{B\times C\times H\times W}$,使用尺寸为$(H,1)$和$(1,W)$的池化核分别沿水平和垂直方向对每个通道进行编码。第c个通道在高度h处的输出如公式(2)所示:

$$z_c^h(h)=\frac{1}{W}\sum_{0\leqslant i<W}x_c(h,i)\,. \tag{2}$$

同理,第c个通道在宽度w处的输出如公式(3)所示:

$$z_c^w(w)=\frac{1}{H}\sum_{0\leqslant j<h}x_c(j,w)\,. \tag{3}$$

上述变换沿两个空间方向分别聚合特征,生成一对方向感知的特征图,使注意力模块能够沿一个空间方向捕获长程依赖,同时沿另一方向保存精确的位置信息。在坐标注意力生成阶段,先将z^h和z^w沿空间维度拼接,经共享的1×1卷积F_1和非线性激活δ处理后得到中间特征图,如公式(4)所示:

$$f=\delta(F_1([z^h,z^w]))\,, \tag{4}$$

其中$[\cdot,\cdot]$表示沿空间维度的拼接操作,$f\in R^{(C/r\times(H+W))}$为编码了水平和垂直方向空间信息的中间特征图,r为通道缩减比。随后将f沿空间维度拆分为$f^h\in R^{(C/r\times H)}$和$f^w\in R^{(C/r\times W)}$,分别经两个1×1卷积F_h和F_w及Sigmoid函数生成注意力权重,如公式(5)和公式(6)所示:

$$g^h=\sigma\left(F_h\left(f^h\right)\right)\,, \tag{5}$$

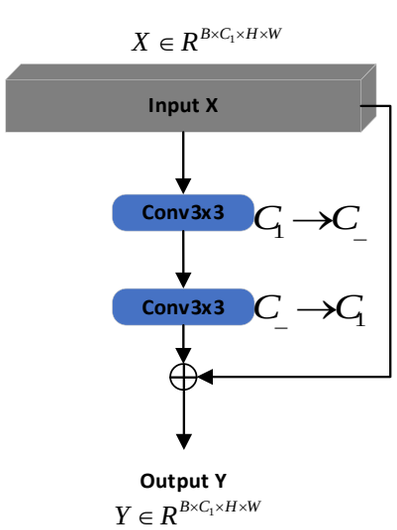
$$g^w=\sigma(F_w(f^w))\,. \tag{6}$$

最终,CA模块的输出如公式(7)所示:

$$y_c(i,j)=x_c(i,j)\times g_c^h(i)\times g_c^w(j)\,. \tag{7}$$

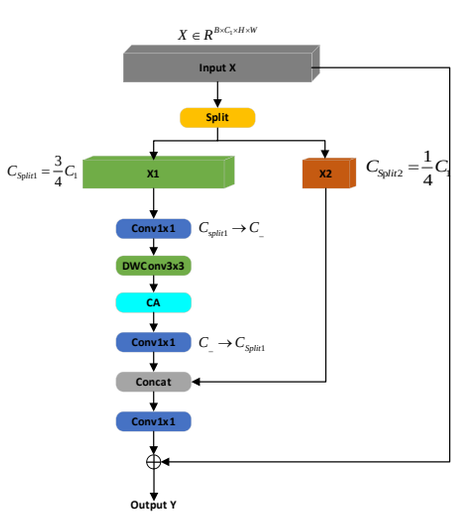
CA通过将二维全局池化分解为两个并行一维编码操作,以较小计算开销实现通道关系与空间位置信息的联合建模,增强了模型对感兴趣区域的空间判别能力。

SCAB通过保留分支传递未失真上下文信息,坐标注意力赋予模型方向感知的特征增强能力,有利于稀疏分布小目标的精确定位。本文将SCAB嵌入骨干网络前两个阶段的C3K2模块中,以最大化浅层特征图上的效率收益。



(a) Bottleneck模块的结构图

(a) The structure of Bottleneck module



(b) SCAB模块的结构图

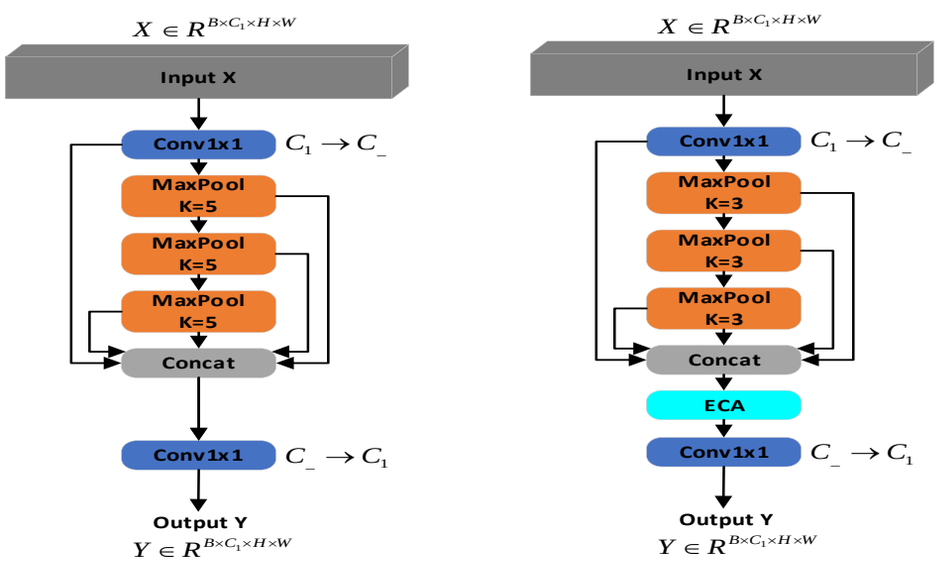
(b) The structure of SCAB module

图2 Bottleneck模块与SCAB模块的结构图

Fig. 2 The structure of Bottleneck and SCAB modules

2.3 增强快速空间金字塔池化（SPPF_E）

SPPF位于YOLO11骨干网络末端，负责聚合多尺度感受野信息，其结构如图3(a)所示，在骨干网络末端，特征图分辨率仅为20×20，标准SPPF的最大池化核经三次串行堆叠后感受野扩展至13×13，覆盖约42%的画面区域，细粒度空间细节丧失严重，小目标特征极易被大范围背景噪声淹没，整体感受野偏向大尺度，与小目标尺寸失配。此外，原始结构在拼接多分支特征后直接通过卷积模块进行通道混合，未对各感受野分支进行差异化建模。在无人机海上搜救场景中，海面杂波、光照反射等复杂背景经池化后在特征通道中形成强响应，与弱小目标特征共存于拼接特征中，均等化融合使背景噪声通道与前景目标通道获得相同权重，不利于小目标特征的凸显。类似地，何梦云等人[17]在水下声呐小目标检测中也指出传统空间金字塔池化易忽视通道信息。



(a)SPPF模块的结构图 (b)SPPF_E模块的结构图

(a) The structure of SPPF module (b) The structure of SPPF_E module

图3 SPPF模块与SPPF_E模块的结构图

Fig. 3 The structure of SPPF and SPPF_E modules

针对上述缺陷，本文提出增强快速空间金字塔池化模块(Enhanced SPPF，SPPF_E)，其结构如图3(b)所示。该模块将串行池化核由$k=5$缩减为$k=3$，最大感受野由13×13降至7×7，在20×20特征图上覆盖面积由42%压缩至约12%，感受野范围与小目标尺寸更为匹配。为进一步增强多尺度特征的判别力，SPPF_E在拼接后引入高效通道注意力(ECA)[18]进行自适应校准，其计算过程为如公式(8)和(9)所示：

$$\omega=\sigma(\mathrm{Conv1d}_k(\mathrm{GAP}(Y\mathrm{pool}))), \quad (8)$$

$$Y_{\mathrm{attn}}=Y_{\mathrm{pool}}\times\omega, \quad (9)$$

其中GAP为全局平均池化，σ为Sigmoid函数，Conv1d为一维卷积，卷积核大小k根据通道维度自适应确定。ECA通过融合四个不同感受野的特征后，学习各通道的重要性权重，抑制编码海面波浪、水面反光等背景干扰的通道，放大携带小目标细节的通道，使网络能根据输入场景动态调整感受野侧重。

2.4 损失函数WIoU

YOLO11默认采用CIoU[19]作为边界框回归损失函数，综合考虑了重叠面积、中心点距离和宽高比三项几何因素，其计算过程如公式(10)所示：

$$L_{\mathrm{CIoU}}=1-\mathrm{IoU}+\frac{\rho^2(p,p^{gt})}{c^2}+\alpha v, \quad (10)$$

其中$\rho(p,p^{gt})$为预测框与真实框中心点的欧氏距离，c为覆盖两框的最小包围框对角线长度，$v=\frac{4}{\pi^2}(\arctan\frac{w^{gt}}{h^{gt}}-\arctan\frac{w}{h})^2$用于衡量宽高比的

一致性，$\alpha=\frac{v}{(1-IoU)+v}$为权衡参数。CIoU综合考虑了重叠面积、中心点距离和宽高比三项几何因素，在常规目标检测任务中取得了良好的回归效果。然而，在无人机视角的小目标检测场景中，小目标像素占比低、特征信息稀疏，锚框质量普遍偏低，CIoU中对中心距离和宽高比的惩罚项会过度放大低质量锚框的几何偏差，导致模型泛化能力下降，边界框回归难以充分收敛。

针对上述问题，本文引入Wise-IoU[20](WIoUv3)损失函数替代原始CIoU。WIoU v3通过构建动态非单调聚焦机制，以离群度评估锚框质量并设计智能梯度增益分配策略，有效降低低质量锚框的有害梯度，使优化过程聚焦于高质量边界框回归。

首先定义锚框的离群度β为单个样本的IoU损失L_{IoU}^{*}与全局L_{IoU}损失的指数滑动平均值$\overline{L_{IoU}}$之比，计算过程如公式(11)所示：

$$\beta=\frac{L_{\mathrm{IoU}}^{*}}{\overline{L_{\mathrm{IoU}}}},\beta\in[0,+\infty). \tag{11}$$

较小的β表示锚框质量较高，较大的β则表示锚框质量过低。WIoU v3据此构建非单调聚焦系数r，完整损失函数计算过程如公式(12)所示：

$$L_{\mathrm{WIoUv3}}=r\cdot R_{\mathrm{WIoU}}\cdot L_{\mathrm{IoU}},r=\frac{\beta}{\delta\alpha^{\beta-\delta}}, \tag{12}$$

其中R_{WIoU}为WIoU的距离注意力因子，α控制聚焦曲线的陡峭程度并决定梯度增益峰值对应的离群度位置，δ控制聚焦系数$r=1$时的离群度值。实验中采用WIoU论文最优配置$\alpha=1.9$、$\delta=3$。由于$\overline{L_{IoU}}$随训练动态更新，质量评判标准能够实时适配当前阶段的锚框质量分布——这一动态特性是WIoU区别于此前静态聚焦损失的核心优势。

在海上小目标检测场景中，受目标像素占比极低、标注边界模糊的影响，大量极低质量锚框均为标注噪声样本。WIoU v3通过非单调聚焦机制可自动抑制此类样本的有害梯度，避免模型对标注误差产生过拟合。在训练过程中，锚框质量分布随模型性能提升动态变化，基于动态滑动均值$\overline{L_{IoU}}$计算的离群度可精准反映当前训练阶段的锚框相对质量，从而摆脱初始锚框分布的约束。

3 实验与分析

3.1 数据集

本文实验采用SeaDronesSee[21]公开数据集，该数据集面向海上智能救援与海事监管场景，样本包含远距离、小尺度、部分遮挡的目标，各类目标普遍像素占比低。鉴于MS COCO[22]将目标面积小于32×32像素的目标定义为小目标，也有研究采用相对面积口径，将占整图面积不足0.1%的目标视为相对小目标[23]。考虑到目标相对整幅图像的比例更能反映其检测难度，本文采用相对面积口径。数据集图像包含海面波浪、阳光反光等复杂海面背景，可用于验证算法在海上复杂环境下的鲁棒性。所有图像均由低空无人机于真实海域采集。

数据集包含5类海上典型检测目标：船只(boat)、浮标(buoy)、摩托艇(jetski)、救生设备(life saving appliances)与游泳者(swimmer)，覆盖了海上搜救、水域巡逻等核心应用场景。

数据集采用官方标准划分，训练集8930张、验证集1547张。经统计，该数据集共含67374个标注目标实例，其中78.3%为小目标，即目标面积占整图不足0.1%。本文遵循官方划分开展实验，确保结果可复现与公平对比。

3.2 实验环境

所有实验在统一软硬件平台下完成以保证结果可复现。硬件采用Intel Xeon Platinum 8474C中央处理器与NVIDIA GeForce RTX 4090D(24GB显存)图形处理器，CUDA版本为12.4；软件基于Ubuntu22.04系统，Python版本为3.11.15，PyTorch框架为2.5.0。训练超参数为：优化器选用SGD，初始学习率设为0.01，动量系数取0.937，权重衰减系数设定为0.0005，输入图像统一设置为640×640尺寸，训练批次大小设置为16，总计训练200个轮次。

3.3 评价指标

为客观、全面地评估所提算法的检测性能，本文从检测精度与模型复杂度两个维度构建多维度评价体系。精度维度选取精确率(Precision，P)、召回率(Recall，R)、mAP50与mAP50-95作为核心评价指标。效率维度选取模型参数量(Parameters)、浮点运算次数(GFLOPs)以及

每秒帧数(FPS)综合衡量算法的检测精度与效率。

(1)精确率(P),用于衡量模型的查准能力,计算过程如公式(13)所示:

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}\ . \tag{13}$$

(2)召回率(R),用于衡量模型的查全能力,计算过程如公式(14)所示:

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}\ , \tag{14}$$

式中,TP表示正确识别的正类样本数量,FP表示被错误识别为正类的负类样本数量,FN表示被错误识别为负类的正类样本数量。

(3)平均精度均值(mAP),综合衡量模型在不同类别上的检测精度,计算过程如公式(15)所示:

$$\mathrm{mAP}=\frac{1}{\mathrm{n}}\sum_{i=1}^{n}AP(i)\ , \tag{15}$$

式中,n为数据集的目标类别总数,$AP(i)$为第i类目标的平均精度。本文采用mAP50与mAP50-95两项评价指标,mAP50对应IoU阈值为0.5的平均检测精度,mAP50-95则是对IoU从0.5到0.95以0.05为步长计算每个阈值下的mAP值并取平均。

(4)参数量与浮点运算次数,用于评估模型轻量化程度与计算效率。卷积层是计算开销的主要来源,其参数量和计算量如公式(16)和(17)所示:

$$\mathrm{Params}_{\mathrm{Conv}}=C_{\mathrm{out}}\times C_{in}\times K_h\times K_w\ , \tag{16}$$

$$\mathrm{GFLOPS}_{\mathrm{Conv}}=\frac{2\times \mathrm{Params}_{\mathrm{Conv}}\times H_{\mathrm{out}}\times W_{\mathrm{out}}}{10^9}\ , \tag{17}$$

其中C_{in}和C_{out}为输入与输出通道数,K_h和K_w为卷积核尺寸,H_{out}和W_{out}为输出特征图尺寸。其余部分的参数量和运算量按各自结构特性统一纳入总量。

(5)每秒帧数(FPS),用于衡量模型的实际推理速度。计算过程如公式(18)所示:

$$FPS=\frac{1000}{T_{\mathrm{inf}}}\ , \tag{18}$$

其中,T_{inf}为单帧推理时间(ms),涵盖图像预处理、网络前向推理与非极大值抑制后处理全过程。

3.4 消融实验

为验证各改进模块的有效性及其协同作用,本文设计系统性的消融实验。以YOLO11n为基线,依次添加SCAB、SPPF_E和WIoU,共进行6组实验,结果如表1所示。

表1 消融实验结果

Tab.1 Ablation experiment results

Base	C3K2_SCAB	SPPF_E	WIoU	P/%	R/%	mAP50%	mAP50−95/%	Params/10^6	GFLOPs
√				80.4	61.1	65.3	39.1	2.6	6.3
√	√			**84.5**	60.7	66.4	39.6	2.6	6.2
√		√		81.5	63.0	66.5	40.2	2.6	6.3
√			√	82.2	63.1	66.8	40.1	2.6	6.3
√	√	√		83.8	62.2	67.5	40.4	2.6	6.2
√	√	√	√	82.5	**63.4**	**68.1**	**40.6**	**2.6**	**6.1**

单独引入SCAB后,mAP50从65.3%提升至66.4%,精确率从80.4%提升至84.5%,可知SCAB的坐标注意力增强了模型区分小目标与海面杂波的判别能力,保留分支跳过空间变换则有效保持了轻量化特性。单独引入SPPF_E后,mAP50提升至66.5%,mAP50-95达40.2%,召回率从61.1%提升至63.0%,这表明,较小池化核保留了更精细的空间细节以匹配小目标尺度,而ECA注意力机制则动态调整通道权重,两者协同作用,共同增强了对多尺度小目标的特征捕获能力。引入WIoU v3后,mAP50达66.8%,召回率升至63.1%,说明离群度感知的梯度分配策略有效抑制了海上小目标标注噪声对边界框回归的干扰,使优化聚焦于有效样本。

SCAB与SPPF_E结合时,mAP50达67.5%(+2.2%),增益呈现正向协同效应。叠加WIoU

后，完整模型mAP50达68.1%（+2.8%），mAP50-95提升至40.6%（+1.5%），三项核心指标达到最优，验证了三个模块在特征提取、多尺度感受野融合与回归优化三个维度的互补性。完整改进模型参数量保持2.6×10^6不变，GFLOPs从6.3降至6.1，表明本文方法在提升精度的同时保持了良好的轻量化特性。

3.5 对比实验

为验证本文方法的综合性能，将改进模型与主流检测模型在SeaDronesSee数据集上进行对比实验，所有模型训练均采用统一超参数配置，结果如表2所示。

表2 对比实验结果

Tab. 2 Comparative experimental results of different models

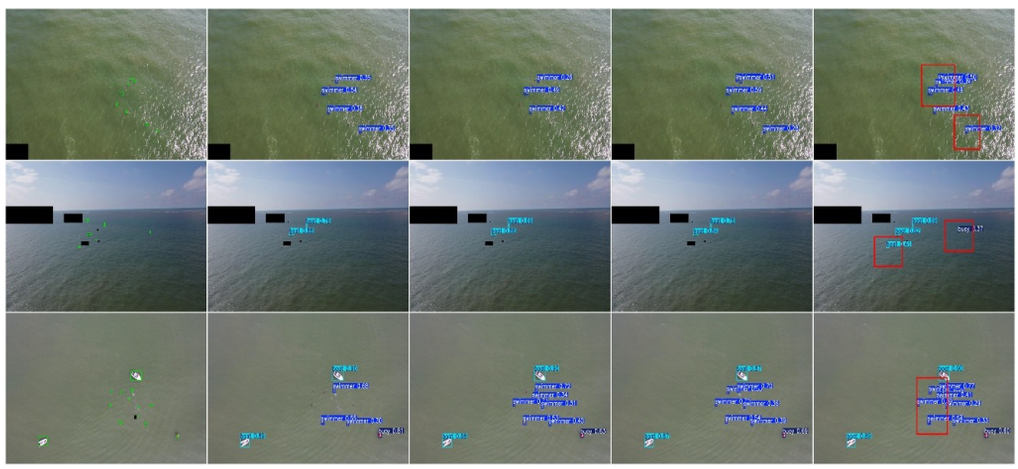
Model	P/%	R/%	mAP50/%	mAP50−95/%	Params/10^6	GFLOPs	FPS
YOLOv5n	82.2	60.9	66.0	40.0	2.5	7.1	230.4
YOLOv6n	81.1	54.8	60.1	36.4	4.2	11.7	**245.0**
YOLOv8n	81.6	62.7	67.8	40.4	3.0	8.1	234.5
YOLOv8nGhostp2	78.8	65.1	69.1	39.3	**1.6**	8.6	174.9
YOLOv9t	78.9	58.9	63.6	38.0	2.0	7.6	123.5
YOLOv10n	81.3	62.0	67.6	40.1	2.3	6.5	214.3
YOLO11n	80.4	61.1	65.3	39.1	2.6	6.3	193.2
YOLO12n	75.7	63.3	66.0	40.0	2.6	6.3	134.7
Drone−YOLO	81.3	**68.5**	70.7	39.7	3.0	12.4	221.4
RT−DETR	81.9	61.7	**71.3**	38.4	19.8	36.7	129.6
Ours	**82.5**	63.4	68.1	**40.6**	2.6	**6.1**	177.1

由表2可知，在mAP50-95指标上，本文模型以40.6%的检测精度位列所有对比方法之首，较基线YOLO11n[24]提升1.5个百分点，较RT-DETR[25]与Drone-YOLO[26]分别领先2.2和0.9个百分点。mAP50-95综合反映模型在不同定位精度要求下的稳定检测能力，该指标取得最优验证了SEW-YOLO在多尺度感知与边界框定位方面的综合优势。精确率方面，本文模型以82.5%位居所有对比模型首位，表明SCAB模块与SPPF_E中的注意力机制有效增强了模型对前景与背景的判别能力，在复杂海面干扰下显著抑制了误检。召回率方面，本文模型达到63.4%，较基线提升2.3个百分点，在精确率领先的前提下仍实现了查全能力的同步增长。

就mAP50而言，本文模型达68.1%，虽略低于RT-DETR（71.3%）和Drone-YOLO（70.7%），但二者的计算开销（GFLOPs分别为36.7和12.4）显著高于本文模型，对机载设备的算力与功耗提出了更高要求。在轻量级模型中，YOLOv8nGhostp2通过引入额外检测头增强小目标召回能力（召回率达65.1%），以1.6×10^6的参数量取得了69.1%的mAP50，较本文模型高出1.0个百分点。然而，在精确率与多阈值检测稳定性方面本文模型更具优势：精确率达82.5%，领先3.7个百分点，mAP50-95为40.6%，领先1.3个百分点，且GFLOPs仅为6.1，较其降低29.1%，推理速度亦略快。上述对比表明，YOLOv8nGhostp2以更高的计算开销换取了召回率与mAP50的提升，而SEW-YOLO在更低的浮点运算次数下实现了更优的查准率与跨阈值检测鲁棒性。在其余参数量$\leqslant3.0\times10^6$的模型中，SEW-YOLO的mAP50优于YOLOv5n（66.0%）、YOLOv8n（67.8%）、YOLOv10n（67.6%）及YOLO11n基线。SEW-YOLO的推理速度为177.1FPS，较基线YOLO11n（193.2 FPS）仅下降8.3%，但以此换来了四项检测指标的全面提升，SEW-YOLO在轻量级检测模型中实现了精度与效率的良好均衡。

3.6 检测可视化与误差分析

为直观展示改进效果，本文选取典型海上小目标复杂场景进行可视化对比分析，图4给出了各基线模型与本文改进模型的检测结果对照。可视化结果表明，各基线模型普遍存在微小目标



原图　YOLOv8n　YOLO11n　YOLO12n　SEW-YOLO

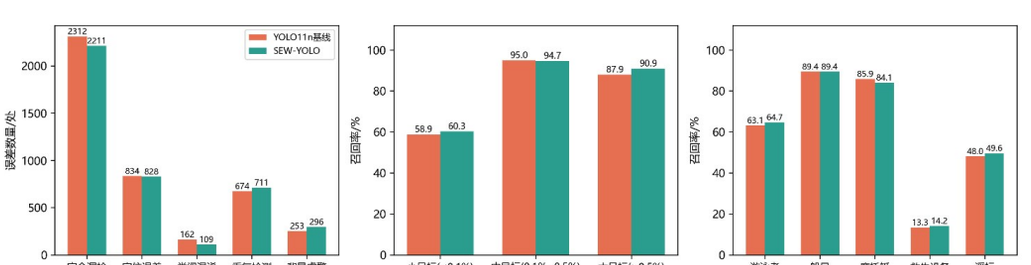
图4　检测结果可视化

Fig. 4　Detection result visualization

漏检、预测框置信度偏低等不足。而本文改进模型能够完整检出绝大多数尺度极小的海上目标，有效抑制漏检，且输出检测框置信度整体更高，对复杂环境下目标特征的失真与弱化表现出更强的抗干扰能力。

尽管如此，改进模型在部分极端场景下仍存在不足：漏检多为尺度极小的浮标、游泳者等目标，定位偏差与误检则集中于波浪反光形成的高亮纹理区域。究其原因，极小目标可用特征信息不足与复杂背景干扰仍是主要制约因素。

为量化上述误差的分布规律，在验证集上按置信度阈值0.25、IoU≥0.5的匹配口径，对全部9630个标注目标实例进行检测误差统计分型。如图5(a)所示，基线YOLO11n共漏检（纯漏检和定位偏差）3146处、误检（剩余三项）1089处，SEW-YOLO漏检降至3039处。其中纯漏检由2312处降至2211处，类间混淆由162处降至109处，基线中游泳者误判为浮标的现象在改进模型中完全消除。背景虚警与重复框略有增加，与模型对弱小目标响应增强后输出框增多有关，改进模型的误差结构整体向好。按目标尺寸分档统计如图5(b)，小目标召回率由58.9%提升至60.3%，漏检减少101处，占全部漏检减量的94%以上，中、大目标召回基本不变，表明改进收益集中于小目标，与本文改进目标一致。分类别看如图5(c)，游泳者类改善最为明显，召回率达64.7%、较基线提升1.6个百分点，改进模型新增命中的目标主要来自该类。救生设备与浮标类召回率仍偏低，分别为14.2%与49.6%，二者成像尺寸小、样本占比低，是后续工作的重点方向。此外，mAP50-95涵盖IoU阈值0.5～0.95，高阈值段对边界框回归误差尤为敏感，本文模型该指



(a)检测误差分型对比　(b)不同尺寸目标召回率对比　(c)不同类别目标召回率对比

(a) Detection error types　(b) Recall by object sizes　(c) Recall by category

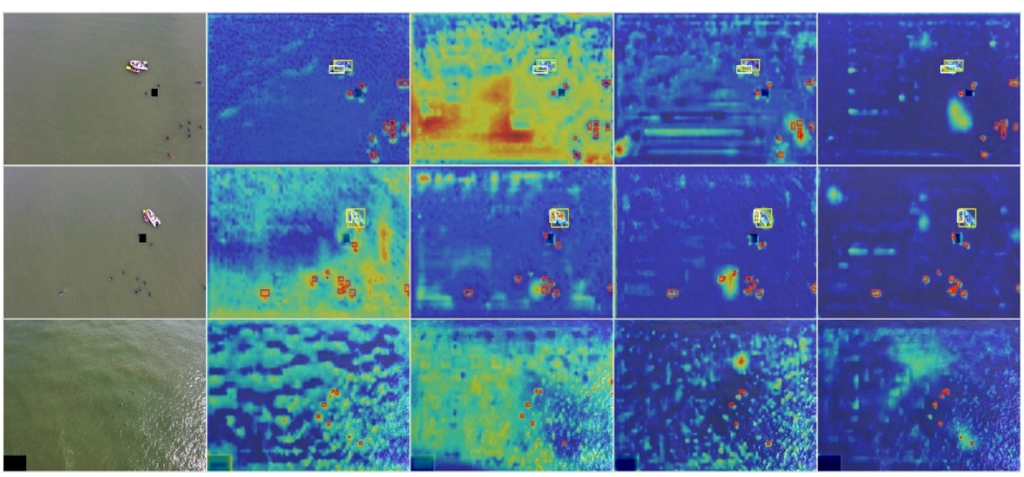
图5　检测误差分析

Fig. 5　Detection error analysis

标提升1.5个百分点，间接反映了WIoU损失对边界框定位精度的改善。

3.7 感受野可视化对比

为直观比较基线模型与本文改进模型在特征关注区域上的差异，本文采用Grad-CAM++[27]方法对YOLOv8n、YOLO11n、YOLO12n及SEW-YOLO进行感受野可视化对比，结果如图6所示。三种YOLO系列模型在复杂海面环境下均存在不同程度的问题，YOLOv8n虽能在目标位置形成有效响应，但在大面积海面波浪纹理上呈现大范围的青绿色弥散激活，背景特征被明显抬高，存在虚警风险。YOLO11n的热力响应在空间分布上与YOLOv8n较为相似，对海面背景的抑制能力同样不足，在密集小目标场景下目标与背景的响应对比度偏低。YOLO12n的背景响应较前两者有所收敛，但其热力图整体呈现块状弥散形态，空间定位精度不够，对小目标的响应零散且不完整。相比之下，SEW-YOLO的热力响应明显向目标局部区域收敛，海面背景的干扰响应被显著抑制。感受野可视化结果表明，本文提出的改进策略能够有效引导骨干网络的特征激活向目标本体聚焦，显著增强复杂海面背景下小目标的特征表达能力。



原图 YOLOv8n YOLO11n YOLO12n SEW-YOLO

图6 感受野可视化对比

Fig. 6 Receptive field visualization comparison

4 结 论

针对无人机海上搜救场景中海面杂波干扰强烈、小目标检测精度偏低的问题，本文基于YOLO11网络架构提出一种改进检测算法SEW-YOLO。针对骨干网络浅层特征提取阶段存在的计算冗余问题，设计SCAB模块，将输入特征沿通道维度分裂为主分支与保留分支，仅对主分支施加坐标注意力以增强方向感知的通道响应，保留分支恒等传递原始特征，在强化小目标空间判别能力的同时抑制了浅层特征提取中的冗余计算。针对基准模型在多尺度感受野特征融合过程中小尺度目标特征易被淹没的问题，提出SPPF_E结构，以小核池化替代原始SPPF的大核池化，抑制背景噪声对小目标的干扰，并嵌入高效通道注意力对多尺度感受野进行自适应校准。在损失函数层面，引入WIoU v3替代原始CIoU，利用离群度感知的梯度增益策略动态抑制低质量锚框对回归训练的干扰，改善了小目标边界框回归的稳定性与收敛质量。在SeaDronesSee数据集上，改进模型以2.6M参数量和6.1GFLOPs的计算开销，mAP50达68.1%，mAP50-95达40.6%，较基线分别提升2.8和1.5个百分点。消融实验验证了SCAB、SPPF_E与WIoU三项改进各自有效且叠加后增益稳定。所提方法在计算开销与基线相当的前提下实现了精度提升，为资源受限的无人机平台的海上小目标检测提供了可行方案。

参 考 文 献:

[1] 郑旭霞. 无人机视角下海洋垃圾监测方法与空间分析[J]. 地理空间信息,2026,24(3):99-102.
ZHENG X X. Monitoring method and spatial analysis of marine debris from UAV perspective [J]. *Geospatial Information*, 2026, 24(3): 99-102. (in Chinese)

[2] 欧丽婷,王云恪,陈硕. 面向海上环境下无人机搜救技术[J]. 中国军转民,2025,4:32-33.
OU L T, WANG Y K, CHEN S. UAV search and rescue technology for maritime environment [J]. *China Military to Civilian Conversion*, 2025, 4: 32-33. (in Chinese)

[3] WU S X, HU J Y, GE J Q, *et al*. 10-km passive drone detection using broadband quantum compressed sensing imaging [J]. *Light: Science & Applications*, 2025, 14(9): 2583-2594.

[4] DUN X, ZHANG J, CHEN F S, *et al*. Neural array meta-imaging [J]. *eLight*, 2025, 5(33).

[5] DONG Y B, BAI Y C, ZHANG Q M, *et al*. High-throughput optical neuromorphic graphic processing at millions of images per second [J]. *eLight*, 2025, 5(29).

[6] ZHANG Z L, KONG L Y, ZHANG L H Y, *et al*. Structured light meets machine intelligence [J]. *eLight*, 2025, 5(26).

[7] U N F, SUN Y X, HU J T, *et al*. Intelligent nanophotonics: when machine learning sheds light [J]. *eLight*, 2025, 5(5).

[8] 许开宇,冯德瀛. 海空背景下红外弱小目标检测方法[J]. 哈尔滨工程大学学报,2009,30(9):996-1000.
XU K Y, FENG D Y. Infrared dim small target detection method under sea-sky background [J]. *Journal of Harbin Engineering University*, 2009, 30(9): 996-1000. (in Chinese)

[9] 陈鹏鹏,冉鑫. 基于决策树的海上搜救目标检测算法[J]. 上海海事大学学报,2010,31(3):1-4.
CHEN P P, RAN X. Maritime search and rescue target detection algorithm based on decision tree [J]. *Journal of Shanghai Maritime University*, 2010, 31(3): 1-4. (in Chinese)

[10] WANG C, HAN Y, YANG C, *et al*. CF-YOLO for small target detection in drone imagery based on YOLOv11 algorithm [J]. *Scientific Reports*, 2025, 15: 16741.

[11] ZHENG Y C, JING Y X, ZHAO J F, *et al*. LAM-YOLO: drones-based small object detection on lighting-occlusion attention mechanism YOLO [J]. *Computer Vision and Image Understanding*, 2024, 261: 104489.

[12] 况子晶,饶瑶,张港平,等. SR-Det:融合轻量化卷积与Transformer的无人机海上搜救检测网络[J/OL]. 计算机工程与应用,1-12[2026-06-25]. https://link.cnki.net/urlid/11.2127.tp.20260328.1259.004.
KUANG Z J, RAO Y, ZHANG G P, *et al*. SR-Det: UAV maritime search and rescue detection network integrating lightweight convolution and Transformer [J/OL]. *Computer Engineering and Applications*, 1-12[2026-06-25]. https://link.cnki.net/urlid/11.2127.tp.20260328.1259.004. (in Chinese)

[13] 吕学涵,李富,祁铭瑞,等. 基于改进YOLOv11s的无人机小目标检测算法[J]. 液晶与显示,2025,40(11):1744-1756.
LV X H, LI F, QI M R, *et al*. Target detection algorithm based on improved YOLOv11s UAV aerial image [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(11): 1744-1756. (in Chinese)

[14] 杨晓敏,杨军. 超轻量化SAR影像小目标检测网络[J]. 光学精密工程,2025,33(10):1672-1690.
YANG X M, YANG J. Ultra lightweight SAR image small object detection network [J]. *Optics and Precision Engineering*, 2025, 33(10): 1672-1690. (in Chinese)

[15] CHEN J, KAO S, HE H, *et al*. Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.

[16] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C]// /Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13708-13717.

[17] 何梦云,何自芬,张印辉,等. 用于水下声呐目标检测的弱特征共焦通道调控方法[J]. 中国光学(中英文),2024,17(6):1281-1296.
HE M Y, HE Z F, ZHANG Y H, *et al*. Weak feature confocal channel regulation for underwater sonar target detection [J]. *Chinese Optics*, 2024, 17(6): 1281-1296. (in Chinese)

[18] WANG Q L, WU B G, ZHU P F, *et al*. ECA-Net: efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11531-11539.

[19] ZHENG Z H, WANG P, REN D W, *et al*. Enhancing geometric factors in model learning and inference for object detection and instance segmentation [J]. *IEEE Transactions on Cybernetics*, 2022, 52(8): 8574-8586.

[20] TONG Z J, CHEN Y H, XU Z W, *et al*. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [J/OL]. *arXiv*, 2023: 2301.10051.

[21] VARGA L A, KIEFER B, MESSMER M, *et al*. SeaDronesSee: a maritime benchmark for detecting humans in open water [C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2022: 3686-3696.

[22] LIN T Y, MAIRE M, BELONGIE S J, *et al*. Microsoft COCO: Common Objects in Context [C]//Proceedings of the European Conference on Computer Vision, 2014: 740-755.

[23] ALDUBAIKHI A, PATEL S. Advancements in Small-Object Detection (2023 - 2025): Approaches, Datasets, Benchmarks, Applications, and Practical Guidance.[J]. *Applied Sciences*, 2025,15(22): 11882.

[24] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements [J/OL]. *arXiv*, 2024: 2410.17725.

[25] LV W, XU S, ZHAO Y, *et al*. DETRs beat YOLOs on real-time object detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.

[26] ZHANG Z. Drone-YOLO: an efficient neural network method for target detection in drone images [J]. *Drones*, 2023, 7(8): 526.

[27] CHATTOPADHYAY A, SARKAR A, HOWLADER P, *et al*. Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks [C]//Proceedings of the IEEE Winter Conference on Applications of Computer Vision, 2018: 839-847.

作者简介：



任喜伟，男，博士，教授级高级工程师，2019年于陕西科技大学获得博士学位，主要从事机器视觉、过程系统工程等方面的研究。E-mail：renxiwei@sust.edu.cn