

DOI: 10.13382/j.jemi.B2407348

# 基于改进 YOLOv8s 的光学遥感小型船舶检测算法 \*

朱圣博 魏利胜 高 港 郑泊文

(安徽工程大学电气工程学院 芜湖 241000)

**摘要:**针对海陆边界、近岸岩礁等复杂海洋场景下,光学遥感小型船舶检测成像特征不明显、目标占比小的问题,提出一种改进 YOLOv8s 的小型船舶检测方法。首先,在颈部层引入浅层特征图的基础上修改预测层,平衡浅层位置信息和深层语义信息的权重,增强模型对小目标的关注度;其次,采用融合 FasterNet Block 和高效多尺度注意力机制的 C2f-FE 模块,利用通道分组和跨通道信息交互,加强对小型船舶的特征提取,并降低模型参数;最后,采用动态检测头模块,在不同预测层级上提高模型对不同空间尺度、任务目标的检测能力。实验结果表明,与原始 YOLOv8s 模型相比,改进模型的参数量减少 42.3%,在 MASATI 数据集上,改进模型的检测精度  $mAP_{50}$  和  $mAP_{50,95}$  值分别提高 4.2% 和 2.2%,在 DOTA-Ship 和 DOTA-Small Vehicle 数据集上,改进模型的检测精度  $mAP_{50,95}$  值分别提高 1.7% 和 1.4%。由此可知,改进模型不仅有效地实现轻量化、高精度的小型船舶检测,而且满足在遥感场景下泛化小目标的高精度检测。

**关键词:**遥感图像;目标检测;小型船舶;特征提取;动态检测头;YOLOv8s

中图分类号: TN957.52; U675.79 文献标识码: A 国家标准学科分类代码: 510.40

## Optical remote sensing small ship detection algorithm based on improved YOLOv8s

Zhu Shengbo Wei Lisheng Gao Gang Zheng Bowen

(School of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, China)

**Abstract:** Aiming at the problem that the imaging features are inconspicuous and the proportion of objects is small in the optical remote sensing small ship detection under the complex marine scenes, such as sea-lean boundary and near-shore rocky reefs, an improved small ship detection method based on YOLOv8s is proposed. Firstly, the prediction layers are modified based on the introduction of shallow feature maps in the neck layers, which balances the weights of shallow locational information and deep semantic information, and enhances the attention of the model to small objects. Secondly, the C2f-FE module is adopted to utilize the channel grouping and the cross-channel information interactions, enhance the feature extraction of small ships, and reduce the model parameters, which merges the FasterNet Block and the efficient multi-scale attention mechanism. Finally, the dynamic detection head module is employed to improve detection capability of the model on different spatial scales and object tasks at different prediction layers. The experimental results show that compared with the original YOLOv8s model, the improved model reduces the number of parameters by 42.3%, the detection accuracy  $mAP_{50}$  and  $mAP_{50,95}$  values are improved by 4.2% and 2.2% on the MASATI dataset, and  $mAP_{50,95}$  values are improved by 1.7% and 1.4% on the DOTA-Ship and DOTA-Small Vehicle datasets, respectively. It can be concluded that the improved model not only achieves lightweight and accurate detection of small ships, but also satisfies the high-accuracy detection for the generalized of small objects in remote sensing scenarios.

**Keywords:** remote sensing image; object detection; small ship; feature extraction; dynamic detection head; YOLOv8s

## 0 引言

随着世界各国经济的深度互补、进出口贸易的增长,船舶运输的需求与日俱增,规避船舶碰撞、领海主权维护的遥感船舶检测技术引起广泛关注。光学遥感图像凭借成像面积大、宏观性强的优势,为船舶检测提供丰富的物体和背景空间信息。然而,光学遥感成像易受天气、光照等因素影响,同时船舶在图像中目标占比小、特征不明显,使其难以检测。因此,光学遥感图像中船舶目标检测的研究具有重要的理论意义与实际应用价值。

光学遥感船舶目标检测的研究工作,以机器学习和深度学习为主要研究方向。传统机器学习检测算法,以低分辨率遥感图像为研究目标,通过将船舶视为点目标,应用恒定虚警率检验、广义似然比检验和模板匹配等方法进行检测<sup>[1-2]</sup>。然而,手工设计视觉特征易受研究人员时间、精力以及专业程度的限制,难以适应现代大规模、多样化的遥感图像数据。深度学习检测算法分为双阶段和单阶段两种模型。R-CNN、Faster R-CNN 等经典双阶段的算法在船舶检测领域广泛应用<sup>[3-4]</sup>,并衍生出诸多变体。Li 等<sup>[5]</sup>引入选择性滤波层,将不同尺度的特征映射到同一尺度空间,为多尺度船舶生成通用特征,但滤波层的加入消耗大量计算资源。为此, Wang 等<sup>[6]</sup>从深度可分离卷积中汲取灵感,设计单次聚合与深度可分离网络,在实现船舶特征提取的同时,最大程度轻量化模型。然而,该方法并不适用于定向舰船的检测场景。于是, Zhang 等<sup>[7]</sup>利用旋转检测头获取船舶角度信息,以应对舰船位置倾斜的检测任务,并设计可变形感兴趣区域模块,通过重组特征形状解决特征提取中长宽比失调问题。由于候选框生成、候选框分类等冗长操作,使双阶段检测实时性仍然不高,难以满足实际高性能检测需要。因此,以 YOLO、SSD 为主的单阶段检测算法<sup>[8]</sup>在光学遥感船舶检测任务中更受关注。李晖晖等<sup>[9]</sup>基于 SSD 设计 CReLU 激活函数,解决模型对船舶目标浅层特征利用率不高的问题。然而,在云雾干扰等场景中, CReLU 函数不能引导模型进行抗干扰训练。为此, Tian 等<sup>[10]</sup>基于 YOLOv5 提出加权融合连接以聚合网络颈部,增强模型在极端天气下通道特征表示。但过度特征融合常伴随推理能力的削弱。Wang 等<sup>[11]</sup>为解决上述问题,通过高斯热图回归后直接输出特征图,并经简单后处理获取船舶位置,大幅提高推理速率。此外,胡欣等<sup>[12]</sup>基于 YOLOv5 提出多分支注意力模块,实现跨维度的船舶位置和语义信息融合,并引入圆形平滑标签实现旋转检测,弥补单阶段模型定向船舶检测领域空缺。

尽管上述 SSD 和 YOLO 系列算法,在诸多具有挑战性的遥感船舶数据集中表现出色。但是,面向遥感图像

前景和背景难以区分的复杂海洋场景,在检测分辨率较小的船舶时,现阶段的研究易出现误报和漏报,同时难以实现轻量化和准确率两者平衡。相较现有的目标检测模型, YOLOv8s 能很好地兼顾模型轻量化和准确率。因此,通过选取 YOLOv8s 作为基底模型,提出一种适用于光学遥感小型船舶检测的改进算法,进一步优化检测小型船舶的性能。首先,改进模型结构,引入浅层特征图,实现与浅层船舶位置信息的融合,并修改预测层,改善预测时对目标船舶尺度的误判;其次,引入特征提取 C2f-FE 模块,通过融合 FasterNet Block 模块<sup>[13]</sup> 和高效多尺度注意力 (efficient multi-scale attention, EMA) 模块<sup>[14]</sup>, 分流冗余通道信息,聚焦跨通道信息的特征交互;最后,引入动态检测头 (dynamic head, Dyhead) 模块<sup>[15]</sup>, 通过在不同预测层上部署动态感知模块,集成更具有特征表示能力的小目标检测头。

## 1 基于 YOLOv8s 算法的改进

YOLO 系列的不断优化改进,使 YOLO 算法来到 v8 版本。YOLOv8s 算法由主干网络、颈部网络和检测头组成。主干用于提图像的目标特征信息,由 CBS 模块、C2f 模块、SPPF 模块组成,其中,CBS 模块包含卷积 Conv 层、批归一化 BN 层和激活函数 SiLU 层;颈部对不同层级的特征图进行融合;检测头根据目标的不同尺寸,在对应预测层的特征图上进行目标分类和位置回归。

考虑到遥感图像的船舶目标占比小、密度比例大, YOLOv8s 算法不能满足更高的小型船舶检测需求。针对这一问题,对模型结构、通道特征提取和检测头 3 个方面对 YOLOv8s 算法进行优化。首先,通过颈部网络添加小目标采样层,引入浅层特征图,构建与主干网络的通道连接,使其能融合浅层目标位置信息,在此基础上修改检测头的预测层级,通过剔除大目标预测层与添加小目标预测层,优化预测过程中对小目标尺度的判定;其次,引入特征提取 C2f-FE 模块,以替换主干网络的 C2f 模块,增强主干网络的特征提取能力,在提升检测精度的同时最大程度地降低参数量;最后,采用 Dyhead 模块,以替换原始的解耦头模块,增强了检测头对小目标中不同特征层级的适应能力。

### 1.1 模型结构的改进

特征图在检测过程中具有重要作用,如颈部层融合不同尺度特征图、预测层构建分类与回归特征图等。根据特征图与输入层的距离,细分为浅层与深层。如图 1 (a) 所示,浅层特征图距输入层较近,具有更高分辨率,包含更多利于小型船舶检测的位置信息。然而,由于遍历的卷积较少,其获取的语义信息较低。相反,如图 1 (b) 所示,深层特征图距输入层较远,随着图像信息的压

缩,深层网络感受野及重叠区域逐渐增加,使其包含更多语义信息,但对小目标位置感知较弱。

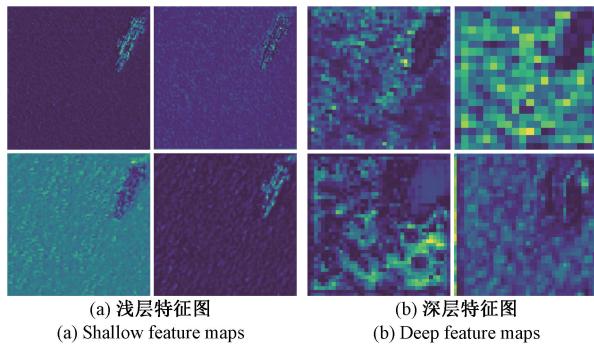


图 1 特征图可视化

Fig. 1 Visualization of feature maps

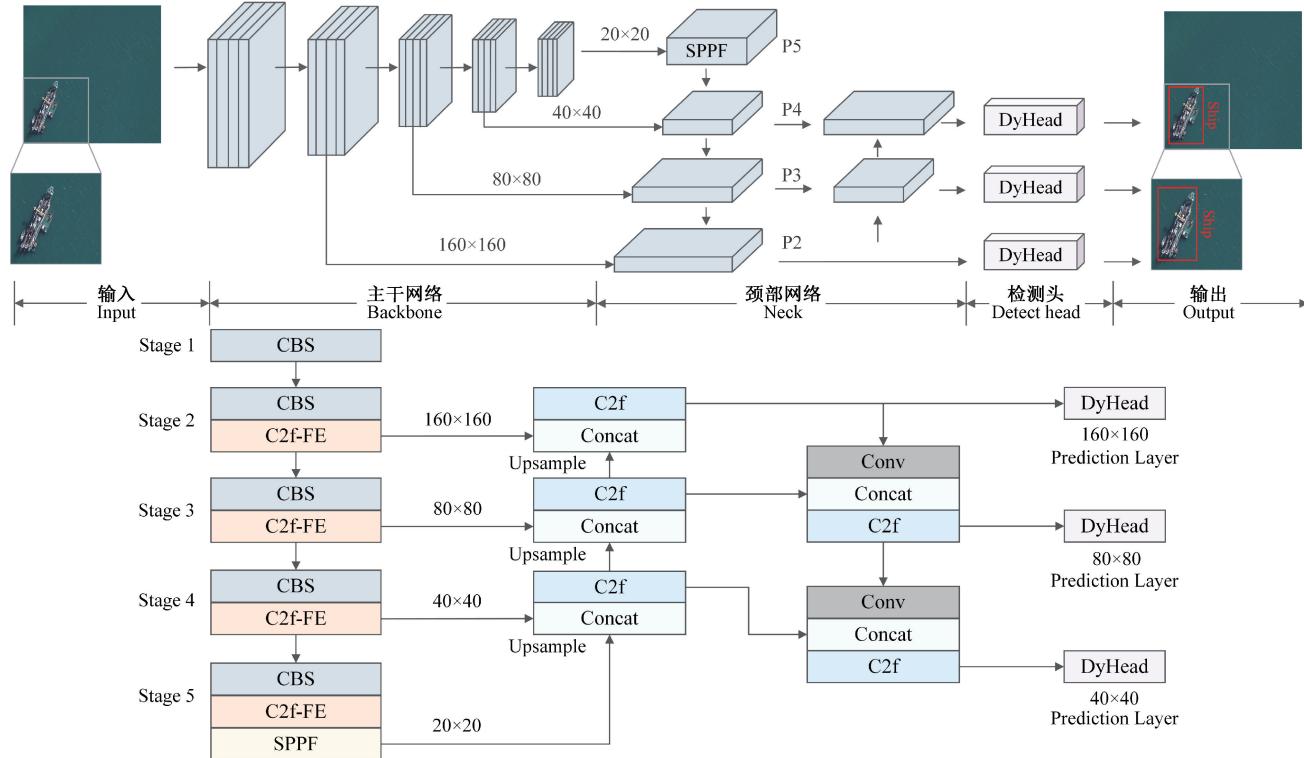


图 2 改进 YOLOv8s 模型整体结构

Fig. 2 The overall structure of the improved YOLOv8s model

对比不同模型结构的实验结果,最终得到结构 2 为检测小型船舶的最佳结构,且保持较小内存占用。结构 2 通过颈部层引入浅层特征图,避免特征融合时船舶位置信息的丢失。同时,颈部层保留深层特征图,使特征融合时获取深层语义信息,并降低引入浅层特征图的噪声累积。最终,小目标预测层的引入和大目标预测层的剔除,提高预测时对小型船舶的尺度感知能力,避免模型对目标尺度的误判。

遍历连续上采样的卷积层较少,融合的特征图尺寸远远小于上采样补偿,导致小型船舶的位置信息丢失。同时,预测层选择 3 个深层特征图进行预测,使模型预测过程对小目标的尺度感知较弱。为解决上述问题,通过不断调整浅层特征图与深层特征图所参与特征融合和目标预测的权重,改进模型结构以提高小目标检测敏感度。基础结构与权重调整的变体结构,如图 3 所示,结构 1 在基础结构上引入大小为  $160 \times 160$  的浅层特征图和大目标预测层;结构 2 在结构 1 基础上剔除  $20 \times 20$  的大目标预测层;结构 3 在结构 2 基础上剔除  $20 \times 20$  的深层特征图。通过改变特征图连通性,浅层特征图参与特征融合和预测权重比例从基础结构到结构 3 逐渐增加。

## 1.2 特征提取模块的改进

网络层数的加深导致特征图通道数的扩增,使特征图携带相似信息造成特征冗余,同时使模型在特征提取时对小目标的敏感度不高。为缓解上述问题,采用特征提取 C2f-FE 模块,通过引入轻量化主干 FasterNet 的 FasterNet Block 结构,并在 FasterNet Block 中融合 EMA 注意力机制,经通道重塑和维度分组最大化保留小型船舶的特征信息,并进一步轻量化模型。

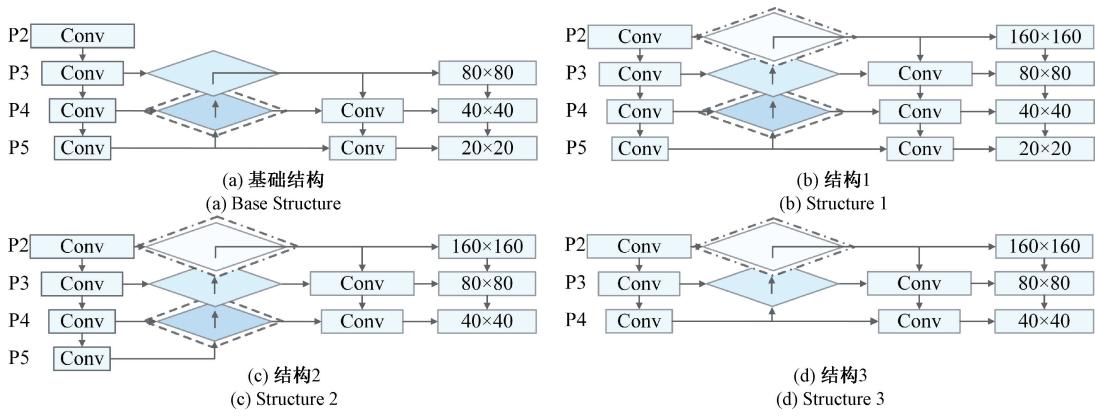


图 3 3 种变体模型结构(菱形框代表不同尺度特征图,虚线框代表  $20 \times 20$  浅层特征图,点划线框代表  $160 \times 160$  深层特征图)

Fig. 3 Three different model structures (The diamond boxes represent feature maps at different scales, the dashed box represents the  $20 \times 20$  deep feature map, and the dotted line box represents the  $160 \times 160$  shallow feature map)

### 1) EMA 注意力机制

注意力机制能有效提高模型对船舶目标敏感度,使模型提取更多有效特征。EMA 提出并行网络和跨空间学习思想,即不进行通道降维,在每个并行子网络中构建局部跨通道交互,在降低计算开销的同时不增加模型大小。如图 4 所示,通过跨空间学习聚合并行分支的输出特征,EMA 能捕获像素级成对关系。

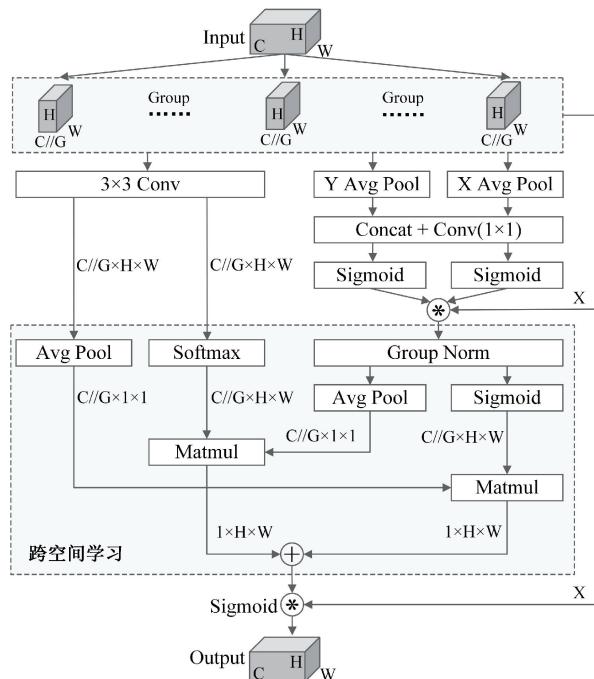


图 4 EMA 注意力机制结构

Fig. 4 Structure of the EMA Attention Mechanism

EMA 注意力首先采用 Group 特征分组,将输入特征  $X \in R^{C \times H \times W}$  按通道维度划分为  $G$  个子特征以学习不同语义;其次,采用并行网络,双路径分别沿  $1 \times 1$  分支依次经

过 Avg Pool 池化、拼接、激活、相乘实现跨通道特征交互,单路径沿  $3 \times 3$  分支经  $3 \times 3$  卷积捕获多尺度特征表示;随后,采用跨空间学习,一方面,经分组归一化 Group Norm 后利用 Avg Pool 对  $1 \times 1$  分支编码全局空间信息,与  $3 \times 3$  分支输出聚合后统一维度为  $R_1^{1 \times 1 \times C//G} \times R_3^{C//G \times H \times W}$ 。另一方面,利用 Avg Pool 对  $3 \times 3$  分支编码全局空间信息,与  $1 \times 1$  分支输出聚合后统一维度为  $R_3^{1 \times 1 \times C//G} \times R_1^{C//G \times H \times W}$ 。其中, Avg Pool 如下:

$$Z_c = \frac{1}{H \times W} \sum_j^H \sum_i^W x_c(i, j) \quad (1)$$

最后,采用 Matmul 矩阵相乘对每组输出特征映射进行计算,生成两个空间注意力权重值集合,经 Sigmoid 函数后得到输出特征图。EMA 采用并行网络重新校准并行分支的权重,突出强调小型船舶特征,通过跨空间学习进一步聚合并行分支特征,利于改善遥感图像有限感受野导致局部信息交互不良的问题。

### 2) C2f-FE 模块

如图 5 所示,在原始的特征提取 C2f 模块中,假设首个 CBS 模块输入通道数为  $2x$ ,经 Split 层将其分为  $2$  份  $x$  个通道的特征图,后续串联的  $n$  个 Bottleneck 模块<sup>[16]</sup> 的输入均为  $x$  个通道的特征图,且每个 Bottleneck 的输出均作为下一个 Bottleneck 的输入。最后  $n$  个 Bottleneck 模块按通道拼接得到  $(2+n)x$  维的特征图,经第 2 个 CBS 模块将其通道压缩为  $2x$  输出。

受轻量级主干网络 FasterNet 和 EMA 注意力机制的启发,特征提取 C2f-FE 模块通过引入 Faster-EMA Block 替换 Bottleneck 结构,在降低参数量和计算量的同时增强小型船舶目标的特征提取,缓解因通道信息分流导致精度下降的问题,如图 6 所示。

首先, Faster-EMA Block 用部分卷积 (partial convolution, PConv) 选取  $1/4$  通道进行卷积处理,余下

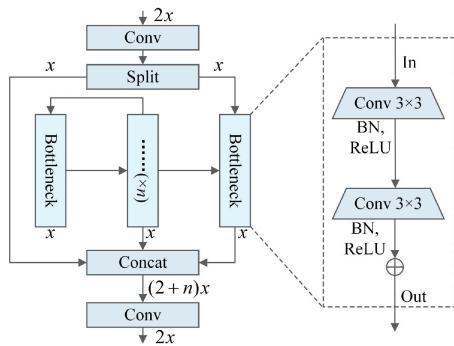


图 5 C2f 模块及 Bottleneck 结构

Fig. 5 Structure of the C2f and Bottleneck block

3/4 通道保持维度不变,并与卷积处理的 1/4 通道进行拼接输出。与 Bottleneck 结构相比,PCConv 在减小冗余计算的同时最大程度保留原有通道信息;其次,由卷积、归一化和激活函数组成的 CBR 模块采用  $1 \times 1$  逐点卷积将高层输出特征图通道数扩张两倍,避免通道信息浪费的同时缓解因 PCConv 造成特征图信息丢失的问题;最后,利用  $1 \times 1$  卷积将高层输出通道数还原,保证捷径分支和主干处理后的特征图具备相同维度。此外,在  $1 \times 1$  卷积后加入 EMA 的操作,实现其余通道信息的充分利用和跨通道船舶特征信息的畅通交互。

尽管串联 Bottleneck 经不同尺度特征融合后利用多层次的信息,提高了特征图的表达能力。然而,每增加

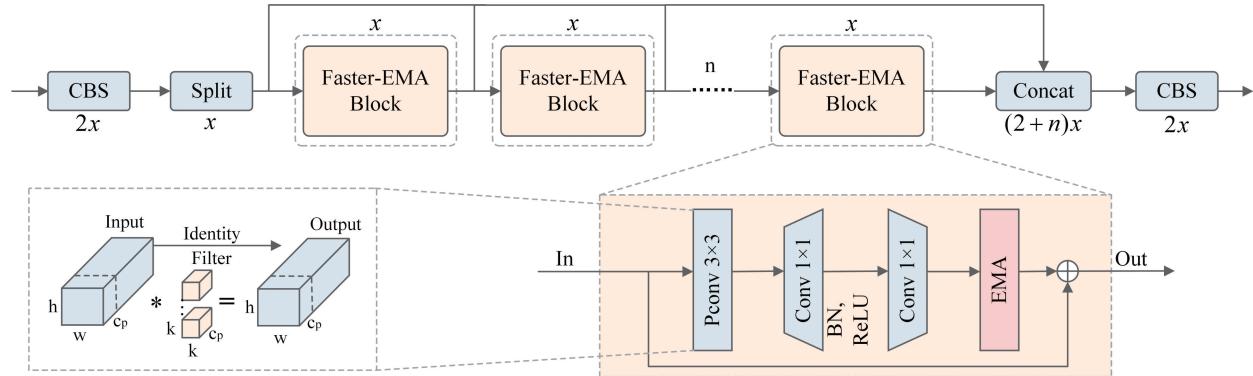


图 6 C2f-FE 模块及 Faster-EMA Block 结构

Fig. 6 Structure of the C2f-FE module and Faster-EMA Block

一个 Bottleneck 会伴随额外参数量的增加。为此,C2f-FE 模块使用 Faster-EMA Block 替换原始 C2f 模块的 Bottleneck。每个 Faster-EMA Block 仅需对 1/4 通道进行卷积操作,而  $1 \times 1$  卷积较原始  $3 \times 3$  卷积参数量更小,同时 EMA 注意力的融入最大程度限制参数量的增加。最终, Faster-EMA Block 的参数量近似为 Bottleneck 的  $1/16$ 。

### 1.3 检测头的改进

在遥感图像中小型船舶与海洋背景具有特征相似的特点,导致模型易将目标与背景相互混淆,同时空间位置和尺度差异也为预测的定位与分类带来挑战。由 Dyhead 通过将多头自注意力分别部署在尺度感知的特征层级、空间感知的空间层级、任务感知的输出通道层级上,并结合在同一框架内。因此,Dyhead 有助于在同等算力下有效提升检测头区分目标和背景的性能。

如图 7 所示,Dyhead 将特征金字塔输出的四维张量  $L \times W \times H \times C$  重塑为三维张量  $L \times S \times C$ ,即特征层级、空间层级、输出通道层级。随后,在特征层级部署尺度感知注意力  $A_L$ ;在空间层级部署空间感知注意力  $A_S$ ;在输出通道层级部署任务感知注意力  $A_C$ ;并依次经过  $A_L, A_S, A_C$

。其中,  $A_L$  根据语义的重要性,动态融合不同尺度船舶的特征;  $A_S$  通过关注船舶的空间位置与特征层级间的共通分辨率区域,更好地定位图像中船舶的位置;  $A_C$  通过联合学习并概括船舶的不同特征,动态打开或关闭特征通道,以适应不同场景检测任务。通过分离式注意力, Dyhead 对每个维度进行特征感知,使检测头对小型船舶的检测更为友好。

因此,对于给定的三维特征张量  $H \in R^{L \times S \times C}$ ,Dyhead 将整体注意力转化为三个序列化的注意力,每个注意力只聚焦于一个角度,其定义如下:

$$W(H) = A_C(A_S(A_L(H) \cdot H) \cdot H) \quad (2)$$

由于遥感图像中感知场较大,而 YOLOv8s 的解耦头结构没有动态感知的能力。因此,引入 Dyhead 模块提升小型船舶检测的效果。此外,3 种应用于检测头的感知注意力能进行多次堆叠,考虑到计算开销等因素,改进模型仅使用一个 Dyhead 模块。

## 2 实验与分析

### 2.1 实验数据集

MASATI 数据集<sup>[17]</sup>由阿利坎特大学提出,由

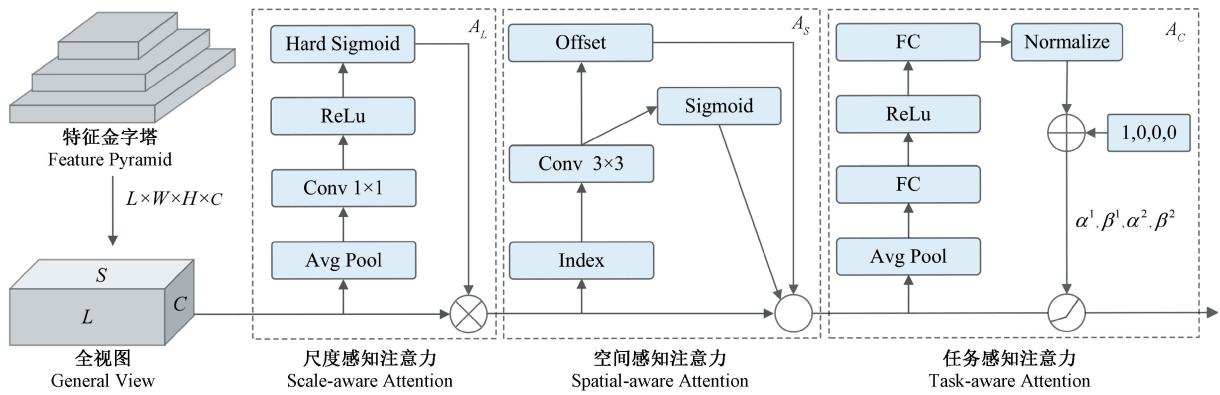


图 7 动态检测头结构

Fig. 7 Structure of the dynamic detection head

Microsoft 必应地图获取的 7 389 张光学卫星遥感图像组成, 分辨率均为  $512 \times 512$  像素。数据集不仅包含不同光照条件下各类小型船舶图像, 还包含各种海洋场景, 如纯海洋、礁石和港口等。通过 7 389 张图像中剔除无目标图像, 选取 2 368 张包含船舶的图像作为实验数据集, 按 7:2:1 随机划分训练集、验证集和测试集。

DOTA 数据集<sup>[18]</sup>由国土资源卫星与谷歌地球应用中心共同提供, 共计 2806 张航拍图, 标注船舶、港口、储槽、网球场、篮球场、大型车辆、小型车辆等共计 15 个类别与 188 282 个实例。数据集像素尺寸在 800~4 000 范围间。本文选取 DOTA 数据集中船舶和小型车辆两个类别, 分别构建 DOTA-Ship 和 DOTA-Small Vehicle 数据集。同时, 根据图像金字塔调整图像尺寸, 将图像裁剪为  $1 024 \times 1 024$  像素。

## 2.2 实验配置与测评标准

实验采用的硬件配置为 i5-12490F 处理器, NVIDIA GeForce RTX 3060 12 G 显卡; 操作系统为 Windows 11; 软件平台为 PyCharm; 深度学习框架为 Pytorch1.12.0; Python 版本为 3.8, 在 CUDA11.4 上进行加速训练。实验参数的设置: batch size 为 16; 训练迭代次数为 300; 采用随机梯度下降作为模型优化器; 采用余弦退火学习衰减方案; 初始学习率为 0.01, 权重衰减设为  $1 \times 10^{-2}$ 。

目标检测算法评价指标主要分为模型复杂度和检测精度。模型复杂度由计算量 GFLOPs、参数量 Params 体现, 数值越大代表模型复杂程度越高。模型检测精度由准确率 (precision, Pre)、召回率 (recall, Re) 和平均准确率 (mean average precision, mAP) 衡量。准确率是所有预测为真样本结果中预测正确的概率, 召回率是根据所有实际正样本中正确预测为正样本的比例计算, 准确率和召回率的定义分别如下:

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

式中:  $TP$  为真正例,  $FN$  为假反例,  $FP$  为假正例。

平均准确率是所有类别准确率的平均值, 在  $n$  个类别的前提下, 计算公式如下:

$$mAP = \frac{1}{n} \sum_{i=1}^n \int_0^1 Precision(Recall) d(Recall) \quad (5)$$

## 2.3 消融实验

为探索不同改进模型结构对遥感图像小型船舶检测任务的效果, 并获取最佳的模型结构。本节在 MASATI 数据集上对不同模型结构进行消融实验, 实验结果如表 1 所示。

表 1 不同模型结构消融结果

Table 1 Ablation results of different model structures %

| 模型结构 | Pre         | Re          | mAP <sub>50</sub> | mAP <sub>50:95</sub> |
|------|-------------|-------------|-------------------|----------------------|
| 基础结构 | 80.9        | 74.3        | 77.2              | 31.8                 |
| 结构 1 | 82.5        | 72.9        | 77.5              | 32.0                 |
| 结构 2 | <b>83.9</b> | <b>74.7</b> | <b>79.4</b>       | <b>33.2</b>          |
| 结构 3 | 81.9        | 73.4        | 78.2              | 32.2                 |

由表 1 可知, 结构 1 引入了浅层特征图和小目标预测层, 尽管提升了小型船舶的检测精度, 但是大目标预测层干扰了模型对目标尺度的判断, 使精度提升不明显; 结构 2 在结构 1 的基础上, 保留深层特征图并剔除大目标预测层, 使模型能融合深层语义特征, 又适当分配预测层权重并聚焦小目标尺寸, 与基础结构相比, 结构 2 在 mAP<sub>50</sub> 和 mAP<sub>50:95</sub> 分别提升 2.2% 和 1.4%; 结构 3 在结构 2 的基础上剔除深层特征图, 由于深层船舶语义信息的减少, 结构 3 相对基础结构在精度上的提升也不显著。

为直观地展示改进方法较原始模型更为优越的性能, 通过添加不同的改进点至原始模型, 在 MASATI 数据集上进行消融实验, 实验结果如表 2 所示。

表 2 消融实验结果

Table 2 The results of the ablation test

| 方法                      | 组 1  | 组 2  | 组 3  | 组 4  | 组 5        | 组 6         |
|-------------------------|------|------|------|------|------------|-------------|
| 原始模型                    | ✓    | ✓    | ✓    | ✓    | ✓          | ✓           |
| C2f-FE                  |      | ✓    |      |      | ✓          | ✓           |
| 结构 2                    |      |      | ✓    |      | ✓          | ✓           |
| Dyhead                  |      |      |      | ✓    |            | ✓           |
| mAP <sub>50</sub> /%    | 77.2 | 77.0 | 79.4 | 78.1 | 80.3       | <b>81.4</b> |
| mAP <sub>50,95</sub> /% | 31.8 | 31.9 | 33.2 | 32.7 | 32.9       | <b>34.0</b> |
| Params/M                | 11.1 | 9.7  | 7.4  | 10.9 | <b>6.0</b> | 6.4         |

对比实验组 1 与实验组 2 可知, C2f-FE 模块采用跨维度交互增强不同通道间的信息融合, 促使输入通道信息变窄, 在保持精度指标稳定的同时降低模型参数量; 对比实验组 1 与实验组 3 可知, 结构 2 平衡了浅层特征图与深层特征图在特征融合与目标预测的权重, 有效改善船舶位置特征丢失和尺度判定不准确的问题, mAP<sub>50</sub> 和 mAP<sub>50,95</sub> 分别提升 2.2% 和 1.4%, 同时参数量降低 30%;

对比实验组 1 与实验组 4 可知, Dyhead 模块动态增强不同尺度的小型船舶特征的预测, mAP<sub>50</sub> 和 mAP<sub>50,95</sub> 均提升 0.9%; 对比实验组 3 与实验组 5 可知, C2f-FE 模块在结构 2 的应用, 进一步提升了模型检测精度, 压缩了模型复杂度; 对比实验组 1 与实验组 6 可知, 相较原始模型, 改进模型参数量降低 42.3%, mAP<sub>50</sub> 和 mAP<sub>50,95</sub> 分别提升 4.2% 和 2.2%。由此可知, 改进模型不仅准确率得以提高, 而且更易于实际部署。

为更清晰地展示改进模型对小型船舶目标的特征增强效果, 如图 8(a<sub>1</sub>)~(d<sub>1</sub>) 所示, 输入 4 张背景各异、单多目标的遥感图像以充分验证改进模型对小型船舶特征提取的有效性。如图 8(a<sub>2</sub>)~(d<sub>2</sub>) 和图 8(a<sub>3</sub>)~(d<sub>3</sub>) 所示, 对比原模型和改进模型的热力图可知, 原模型对小目标的关注度较少, 尚未提取到丰富的船舶特征, 将海岸线树丛误检为船舶的现象较为严重。而改进模型对船舶特征提取效果得到增强, 尤其是小尺度船舶。

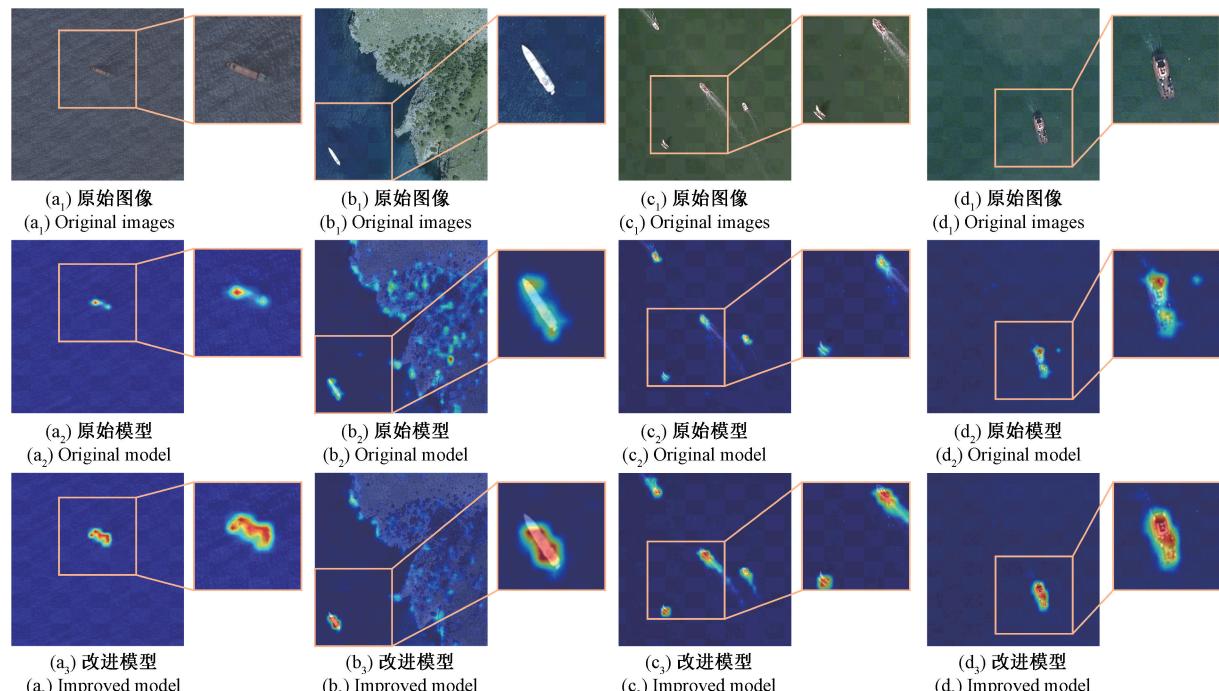


图 8 特征提取效果对比  
Fig. 8 Comparison of feature extraction effects

## 2.4 对比实验

本节将改进模型与 YOLOv5s、YOLOv7、YOLOv8s、RT-DETR<sup>[19]</sup>、文献[20]、文献[21]、文献[22]在 MASATI 数据集上进行实验对比, 结果如表 3 所示。

由表可知, 由于 YOLOv7 和 RT-DETR 的主干网络层数较深, 获取的深层语义信息较多, 忽略了浅层位置信息, 在检测小型船舶时效果不理想; 文献[20]采用 HR-FPN 进行特征提取, 但其对小型船舶的特征捕捉不完整,

既增加了计算量, 检测结果同样不理想; 文献[21]采用相似性掩码结构, 过滤部分无效区域, 但易造成小型船舶丢失, 整体性能不突出; 文献[22]在轻量化后 Re 值达到最佳 77.0%, 但是实际模型仍较大。而改进模型在指标 Pre、mAP<sub>50</sub>、mAP<sub>50,95</sub>、Params 上均达到最优, 分别为 85.3%、81.4%、34.0%、6.4 M, 相比原始模型 YOLOv8s, 改进模型分别优化了 4.4%、4.2%、2.2%、3.7 M, 实现了 YOLOv8s 检测准确率和模型轻量化的双向提升。

表 3 对比实验结果(1)

Table 3 The results of the comparison test (1)

| 模型      | Pre/ %      | Re/ %       | mAP <sub>50</sub> / % | mAP <sub>50,95</sub> / % | Params/ M  |
|---------|-------------|-------------|-----------------------|--------------------------|------------|
| YOLOv5s | 82.3        | 73.3        | 77.1                  | 30.2                     | 7.1        |
| YOLOv7  | 75.1        | 57.7        | 60.8                  | 20.3                     | 36.5       |
| YOLOv8s | 80.9        | 74.7        | 77.2                  | 31.8                     | 11.1       |
| RT-DETR | 80.8        | 76.1        | 79.1                  | 33.0                     | 19.9       |
| 文献[20]  | 84.6        | 72.0        | 77.3                  | -                        | 16.6       |
| 文献[21]  | 82.5        | 76.5        | 78.7                  | -                        | 7.5        |
| 文献[22]  | 80.0        | <b>77.0</b> | 76.0                  | -                        | 36.1       |
| 改进模型    | <b>85.3</b> | 74.9        | <b>81.4</b>           | <b>34.0</b>              | <b>6.4</b> |

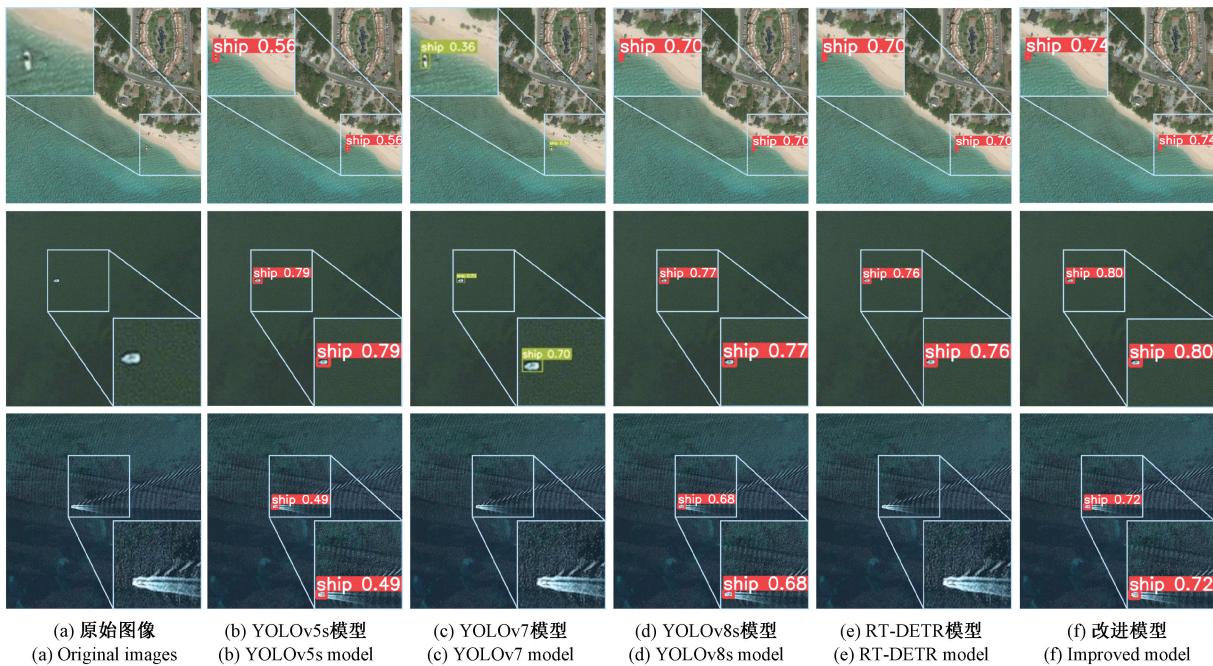


图 9 部分算法检测效果对比

Fig. 9 Comparison of detection effect with part algorithms

表 4 对比实验结果(2)

Table 4 The results of the comparison test (2)

| 模型                 | Pre         | Re          | mAP <sub>50</sub> | mAP <sub>50,95</sub> | % |
|--------------------|-------------|-------------|-------------------|----------------------|---|
| DOTA-Ship          |             |             |                   |                      |   |
| YOLOv5s            | 93.0        | 86.1        | 90.2              | 65.5                 |   |
| YOLOv7             | 92.9        | 85.5        | 89.3              | 63.1                 |   |
| RT-DETR            | 92.2        | 85.1        | 88.7              | 64.6                 |   |
| YOLOv8s            | 92.6        | 87.0        | 91.2              | 67.9                 |   |
| 文献[23]             | <b>93.9</b> | 81.9        | 90.9              | 64.5                 |   |
| 文献[24]             | -           | -           | 88.5              | 66.4                 |   |
| 改进模型               | 93.5        | <b>87.4</b> | <b>91.4</b>       | <b>69.6</b>          |   |
| DOTA-Small Vehicle |             |             |                   |                      |   |
| YOLOv5s            | 76.3        | 74.8        | 76.3              | 46.4                 |   |
| YOLOv7             | 77.1        | 78.2        | 79.3              | 45.0                 |   |
| RT-DETR            | 75.1        | 70.6        | 75.4              | 46.2                 |   |
| YOLOv8s            | 77.3        | 78.7        | 80.2              | 50.1                 |   |
| 文献[23]             | 73.6        | 76.0        | 75.3              | 48.3                 |   |
| 文献[24]             | -           | -           | 80.8              | 45.6                 |   |
| 改进模型               | <b>78.2</b> | <b>80.4</b> | <b>81.5</b>       | <b>51.5</b>          |   |

面的优越性;在 DOTA-Small Vehicle 数据集上,改进模型在 mAP<sub>50</sub> 和 mAP<sub>50,95</sub> 上分别达到 81.5% 和 51.5%,相比原模型 YOLOv8s,改进模型 mAP<sub>50</sub> 和 mAP<sub>50,95</sub> 分别提高 1.3% 和 1.4%,表明了改进模型在遥感图像中检测泛化小目标方面具有出色的泛化性。

### 3 结论

为解决遥感图像检测任务中小型船舶难以检测的问题,研究了一种改进 YOLOv8s 算法。首先,改进模型结构以平衡浅层特征图和深层特征图的权重,增强对小型船舶的定位能力和尺度判别能力;其次,采用 C2f-FE 模块,通过融合 Faster Block 模块和 EMA 注意力机制,分流特征通道,聚焦学习跨通道特征,实现跨空间信息聚合,增强特征提取能力;随后,引入 Dyhead 模块,在预测层的

不同层级部署注意力机制,动态增强特征融合后的小型船舶特征,加强模型的分类与定位能力。最后,在3个公开数据集上进行消融实验和对比实验。实验结果表明,与其他主流的检测算法相比,改进模型在检测效果和模型大小方面均表现优异,更易于实际部署应用。后续工作将进一步压缩模型大小,并保持检测准确率。

## 参考文献

[ 1 ] 陈艳浩. 基于全局背景减法滤波器与多形状特征的红外舰船检测算法 [J]. 电子测量与仪器学报, 2018, 32(10):138-146.

CHEN Y H. Infrared ship detection algorithm based on global background subtraction filter and multi-shape feature [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(10): 138-146.

[ 2 ] 赵泉华,王肖,李玉,等. 基于散射梯度矢量的PolSAR数据舰船检测方法 [J]. 仪器仪表学报, 2020, 41(3): 39-48.

ZHAO Q H, WANG X, LI Y, et al. PolSAR ship detection method based on the scattering gradient vector [J]. Chinese Journal of Scientific Instrument, 2020, 41(3): 39-48.

[ 3 ] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[ 4 ] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[ 5 ] LI Q P, MOU L CH, LIU Q T, et al. HSF-Net: Multiscale deep feature embedding for ship detection in optical remote sensing imagery [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(12): 7147-7161.

[ 6 ] WANG P, LIU J ZH, ZHANG Y B, et al. A novel cargo ship detection and directional discrimination method for remote sensing image based on lightweight network [J]. Journal of Marine Science and Engineering, 2021, 9(9): 932-959.

[ 7 ] ZHANG J, HUANG R F, LI Y, et al. Oriented ship detection based on intersecting circle and deformable RoI in remote sensing images [J]. Remote Sensing, 2022, 14(19): 4749-4760.

[ 8 ] 李科岑,王晓强,林浩,等. 深度学习中的单阶段小目标检测方法综述 [J]. 计算机科学与探索, 2022, 16(1):41-58.

LI K Q, WANG X Q, LIN H, et al. Survey of one-stage small object detection methods in deep learning [J]. Journal of Frontiers of Computer Science and Technology, 2022, 16(1): 41-58.

[ 9 ] 李晖晖,周康鹏,韩太初. 基于 CReLU 和 FPN 改进的 SSD 舰船目标检测 [J]. 仪器仪表学报, 2020, 41(4): 183-190.

LI H H, ZHOU K P, HAN T CH. Ship object detection based on SSD improved with CReLU and FPN [J]. Chinese Journal of Scientific Instrument, 2020, 41(4): 183-190.

[ 10 ] TIAN Y, WANG X, ZHU SH J, et al. LMSD-Net: A lightweight and high-performance ship detection network for optical remote sensing images [J]. Remote Sensing, 2023, 15(17): 4358-4374.

[ 11 ] WANG ZH Q, ZHOU Y, WANG F T, et al. SDGH-Net: Ship detection in optical remote sensing images based on gaussian heatmap regression [J]. Remote Sensing, 2021, 13(3): 499-515.

[ 12 ] 胡欣,马丽军. 基于YOLOv5的多分支注意力SAR图像舰船检测 [J]. 电子测量与仪器学报, 2022, 36(8): 141-149.

HU X, MA L J. Multi-branch attention SAR image ship detection based on YOLOv5 [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(8): 141-149.

[ 13 ] CHEN J, KAO S H, 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.

[ 14 ] OUYANG D, HE S, ZHAN G ZH, et al. Efficient multi-scale attention module with cross-spatial learning [C]. Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing, 2023.

[ 15 ] DAI X, CHEN Y, XIAO B, et al. Dynamic head: Unifying object detection heads with attentions [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021.

[ 16 ] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2016.

[ 17 ] ALASHHAB S, GALLEGOS A J, PERTUSA A, et al. Precise ship location with CNN filter selection from optical aerial images [J]. IEEE Access, 2019, 7: 96567-96582.

[ 18 ] XIA G S, BAI X, DING J, et al. DOTA: A large-scale dataset for object detection in aerial images [C].

Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2017.

[19] ZHAO YAN, LYU W Y, XU SH L, et al. DETRs beat YOLOs on real-time object detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024.

[20] ZHANG X W, LIU Q Y, CHANG H L, et al. High-resolution network with transformer embedding parallel detection for small object detection in optical remote sensing images [J]. *Remote Sensing*, 2023, 15(18): 4497-4507.

[21] ZHANG X, YUAN S, LUAN F, et al. Similarity mask mixed attention for YOLOv5 small ship detection of optical remote sensing images [C]. Proceedings of the WRC Symposium on Advanced Robotics and Automation, 2022.

[22] MISHRA S, SHALU P, SOMAN D, et al. Advanced ship detection system using YOLOv7 [C]. Proceedings of the IEEE Conference on Information and Communication Technology, 2023.

[23] KIM M, JEONG J, KIM S. ECAP-YOLO: Efficient channel attention pyramid YOLO for small object detection in aerial image [J]. *Remote Sensing*, 2021, 13(23): 4851-4871.

[24] 刘树东,任慧娟,张众维.基于特征融合和注意力机制的遥感目标检测[J].*遥感信息*,2023,38(5):1-7.  
LIU SH D, REN H J, ZHANG ZH W. Remote sensing image object detection based on feature fusion and

attention mechanism [J]. *Remote Sensing Information*, 2023, 38(5): 1-7.

### 作者简介



朱圣博,2022年于常州工学院获得学士学位,现为安徽工程大学硕士研究生,主要研究方向为遥感舰船检测、深度学习。  
E-mail: 2220320180@stu.ahpu.edu.cn

**Zhu Shengbo** received his B. Sc. degree from Changzhou Institute of Technology in 2022. Now he is a M. Sc. candidate at Anhui Polytechnic University. His main research interests include remote sensing ship detection and deep learning.



魏利胜(通信作者),2001年于安徽工程大学获得学士学位,2004年于中国航天科工集团061基地获得硕士学位,2009年于上海大学获得博士学位,现为安徽工程大学教授、硕士生导师,主要研究方向为图像识别与应用、智能化网络控制系统和仿真。

E-mail: lshwei\_11@163.com

**Wei Lisheng** (Corresponding author) received his B. Sc. degree in 2001 from Anhui Polytechnic University and received his M. Sc. degree in 2004 from China Aerospace Science and Industry Corporation 061 Base, and received his Ph. D. degree in 2009 from Shanghai University. Now he is a professor and master's supervisor at Anhui Polytechnic University. His main research interests include image recognition and application, intelligent network control system and simulation.