

基于注意力机制及多尺度融合的红外船舶检测

张燊¹, 胡林^{1,2}, 孙祥娥^{1,2*}, 刘美华^{1,2}

¹长江大学电子信息学院, 湖北 荆州 434023;

²长江大学人工智能研究院, 湖北 荆州 434023

摘要 针对红外图像船舶目标检测方法的准确性和实时性还不能满足海防场景需求的问题, 提出一种基于改进 YOLOv7 算法的轻量级船舶检测算法。该算法首先在 Backbone 网络中引入 MobileNetv3 主干网络, 实现模型轻量化处理。然后在 Neck 网络引入注意力机制抑制噪声与干扰, 以提高网络的特征提取能力, 并采用双向加权特征金字塔, 以提升特征融合能力。最后引入 Wise IoU 优化损失函数, 提高模型的收敛速度与精度。在艾睿数据集上的实验结果表明, 相较于 YOLOv7, 所提改进算法的精确率、召回率、平均精度均值 (mAP)@0.5 与 mAP@0.5:0.95 分别提升 0.9 个百分点、2.5 个百分点、1.2 个百分点和 1.2 个百分点, 模型参数量降低了约 38.4%, 浮点运算数 (FLOPs) 降低了约 65.5%。所提改进算法在满足检测速度要求的同时得到了更优的检测精度, 有效地实现了高速、高精度的船舶检测。

关键词 YOLOv7; 注意力机制; 多尺度融合; 船舶检测

中图分类号 TP391.41; U675.79

文献标志码 A

DOI: 10.3788/LOP231462

Infrared Ship Detection Using Attention Mechanism and Multiscale Fusion

Zhang Shen¹, Hu Lin^{1,2}, Sun Xiang'e^{1,2*}, Liu Meihua^{1,2}

¹School of Electronic Information, Yangtze University, Jingzhou 434023, Hubei, China;

²Intelligence Research Institute, Yangtze University, Jingzhou 434023, Hubei, China

Abstract To address the issue of inadequate accuracy and real-time performance of infrared ship target detection methods on coastal defense scenarios, a novel lightweight ship detection algorithm based on improved YOLOv7 framework is proposed. This framework incorporates several enhancements to augment its capabilities. First, to achieve model lightweight processing, the algorithm integrates the MobileNetv3 network into the architecture of the Backbone network. This addition contributes to efficient computation and model size reduction. Second, an attention mechanism is introduced within the Neck network to mitigate noise and interference, thereby improving the network's feature extraction capability. In addition, we employ a bidirectional weighted feature pyramid to enhance feature fusion within the network, promoting more effective information integration. Finally, the algorithm incorporates Wise IoU to optimize the loss function, improving convergence speed and model accuracy. Experimental evaluations on the Arrow dataset demonstrate noteworthy improvements over the standard YOLOv7 approach. Specifically, the proposed enhanced algorithm exhibits a 0.9 percentage points increase in accuracy, 2.5 percentage points increase in recall, and 1.2 percentage points increase in mean average precision (mAP) at IoU thresholds of 0.5 and 0.5 : 0.95. In addition, it achieves approximately 38.4% reduction in model parameters and a 65.5% reduction in floating point operations per second (FLOPs). This enhanced algorithm delivers superior inspection accuracy while meeting the speed requirements for efficient ship inspection. Consequently, it effectively enables high-speed and high-precision ship detection.

Key words YOLOv7; attention mechanism; multiscale fusion; ship detection

1 引言

近年来,随着互联网、云计算等新技术的飞速发

展,智慧水上交通的建设得到了积极推进,例如智慧海事监管服务平台、智慧航道船舶流量检测系统等应用,这些应用都离不开对船舶实时准确的检测。船舶检测

收稿日期: 2023-06-05; 修回日期: 2023-06-28; 录用日期: 2023-07-24; 网络首发日期: 2023-08-18

基金项目: 国家自然科学基金(62273060)

通信作者: *xinges2000@yangtzeu.edu.cn

在民用领域具有重要价值,包括海上交通、航行安全、渔业管理和海洋监测;在军事领域同样具有重要作用,如目标搜索、航行安全和港口监测^[1-2]。因此,研究出精确、高效的船舶检测技术很有必要。

近年来,由于深度学习技术的发展,出现了许多目标检测框架,如两阶段算法(R-CNN^[3]、fast R-CNN^[4]、faster R-CNN^[5]),此类网络具有更高的准确性;单阶段算法(YOLO^[6-8]、SSD^[9]),此类网络由于省略了区域提议步骤,具有计算复杂度低、速度快等特点。针对船舶检测任务面临的挑战,许多研究者在上述网络的基础上进行了有效的改进。Miao等^[10]提出了一种基于多尺度特征提取和轻量级卷积神经网络(CNN)的船舶检测方法,采用多特征融合实现船舶分类。王慧赢等^[11]提出了一种以 ShuffleNetv2 为主干网络的轻量级光学遥感图像舰船检测网络。徐志京等^[12]针对遥感舰船小目标检测,设计了生成对抗网络来增强纹理特征。Liu等^[13]为解决 SAR 图像目标尺度的多样性问题,引入改进的感受野块(RFB)结构,增强网络的特征提取能力。

目前的船舶检测研究主要基于可见光图像和 SAR 图像,提升对小目标、多尺度目标的检测能力,重在关注目标识别性能而忽略了模型的大小和计算效率^[14]。为了解决上述问题,本文提出了一种改进模型,

改进如下:主干网络的改进,用更高效的轻量级网络 MobileNetv3 替换原网络,减少网络参数量,提高检测速度;特征融合网络的改进,采用双向特征金字塔网络(BiFPN)结构进行多尺度特征融合,集成更丰富的语义特征和保持空间信息,嵌入注意力模块,增强图像关键信息,抑制图像背景信息,提高船舶目标的置信度;引入一种新的回归损失函数(Wise IoU 损失)来增加检测框定位精度和模型收敛速度。

2 改进的 YOLOv7 网络结构

YOLOv7 算法^[15]由输入端、骨干网络(Backbone)、颈部网络(Neck)、头部网络(Head)4 部分组成。其中,输入端主要对图像进行一系列预处理操作;Backbone 的作用是提取图像三次下采样的特征;Neck 的作用是将图像的特征信息融合,最终得到 $80 \times 80 \times 36$ 、 $40 \times 40 \times 36$ 、 $20 \times 20 \times 36$ 三种不同尺寸的特征图;Head 的作用是对预测框进行处理,得到预测类别概率、置信度及预测定位信息。

本文以 YOLOv7 为基准模型,设计了融合注意力机制和加权双向金字塔的轻量级船舶检测模型。改进后的模型如图 1 所示,其中 k 为卷积核的大小, s 为卷积步长。

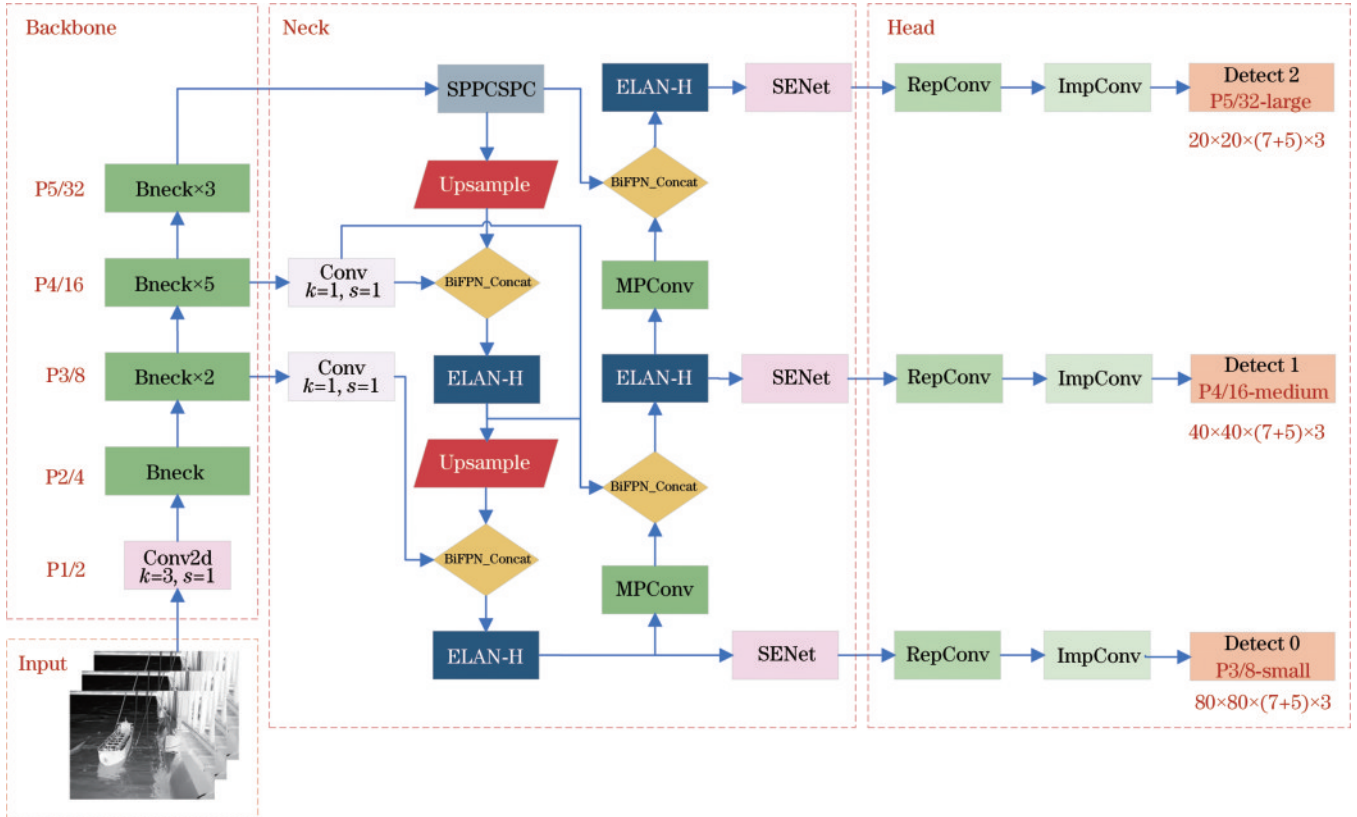


图 1 改进算法的网络结构

Fig. 1 Network structure of improved algorithm

2.1 主干网络改进

MobileNetv3 是谷歌 MobileNet 系列的最新成

果^[16],作为 MobileNet 系列的最新版本,具有参数量少、精度高、实时检测速度快等特点,广泛应用于嵌入式和

移动终端。与同样是轻量级网络的 ShuffleNetv2^[17]和 GhostNet^[18]相比, MobileNetv3 在计算能力、内存消耗和推理延时方面具有一定的优势, 表 1 列出了相关数据, 这些数据使用 torchstate 工具得到统计, 推理延时指标是使用 NCNN 推理框架在树莓派 4B 上进行推理测试得到的。为了减少模型的参数个数和浮点运算数 (FLOPs), 使用轻量级网络 MobileNetv3 替换原有的 YOLOv7 主干网络进行特征提取, 实现模型的轻量化。

表 1 轻量级网络性能对比

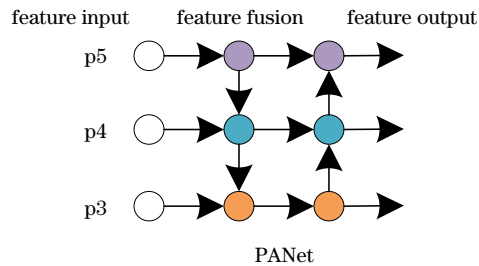
Table 1 Performance comparison of light weight network

Network	FLOPs / 10 ⁶	Memory / MB	Latency / ms
MobileNetv3	68.62	18.99	27
ShuffleNetv2	149.58	20.84	41
GhostNetv1	148.17	40.04	90

MobileNetv3 网络参数分析和计算如下: 大小为 $D_F \times D_F \times M$ 的输入图像经过 N 个大小为 $D_K \times D_K$ 的卷积核, 输出深度为 N 的特征图, 其经过普通卷积的计算量为 $D_K \times D_K \times M \times N \times D_F \times D_F$, 其经过深度可分离卷积的计算量为深度卷积的计算量 $D_K \times D_K \times M \times D_F \times D_F$ 加上点卷积的计算量 $M \times N \times D_F \times D_F$ 的结果。综上可知, 理论上普通卷积的计算量是深度可分离卷积的 8~9 倍。改进的具体步骤为: 采用倒残差结构, 先利用 1×1 卷积进行升维, 再进行 3×3 深度可分离卷积, 并通过注意力机制调整每个通道的权重, 最后通过 1×1 卷积进行降维。所提算法的具体实现方式为: 通过 torchvision 调用 MobileNetv3-small 网络的 feature 模块, 通过切片的形式得到对应于 YOLOv7 三次下采样的特征, 以达到替换主干网络的目的。

2.2 特征融合网络改进

SENet 是 Hu 等^[19]于 2018 年在 CVPR 会议上提出的一种通道注意力机制。其核心思想是: 对通道之间的相互依赖进行建模, 并为每个通道生成相应的权重,



以聚焦重要特征, 并压制不重要的特征。SENet 的过程分为两个步骤, 包括 Squeeze 和 Excitation。Squeeze 通过对提取的特征进行全局平均池化 (GAP) 操作, 得到当前特征图的全局压缩特征向量。Excitation 采用包含两个全连接层的 bottleneck 结构来进行非线性变换, 主要为了达到降低模型复杂度以及提升泛化能力的目的。其中第一个 FC 层起到降维的作用, 然后采用 ReLU 函数进行激活。后一个 FC 层恢复原始的维度, 采用 h-swish 函数进行激活。改进算法在 Neck 末端增加了 SENet, 由于 SENet 可以显著加强对目标的关注, 增强检测的细化能力, 因此在多尺度预测前使用它来提高性能。SENet 注意力机制结构如图 2 所示。

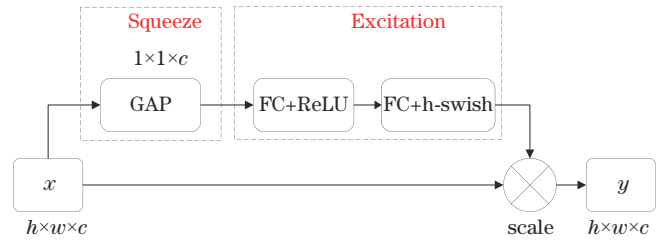


图 2 SENet 注意力机制

Fig. 2 SENet attention mechanism

不同尺度特征层的融合是提高检测网络识别性能的重要途径, YOLOv7 采用 PANet 结构。为了更好地融合特征, 采用 BiFPN 结构。BiFPN^[20]的主要思想是高效的双向跨尺度连接和加权特征融合。此前有研究者将该方法引入 YOLOv5 算法, 以此提升对遥感图像舰船的分类检测效果。周旗开等^[21]将该算法引入 YOLOv5 改进算法, 采用一次堆叠策略及自下而上分支优化高层特征, 提升了检测效果。所提方法在 40×40 特征层增加一条额外的边, 用加权特征融合代替不同特征层之间的连接; 并在融合过程中引入自适应权重来学习不同层次特征的重要性。PANet 和 BiFPN 的结构如图 3 所示。

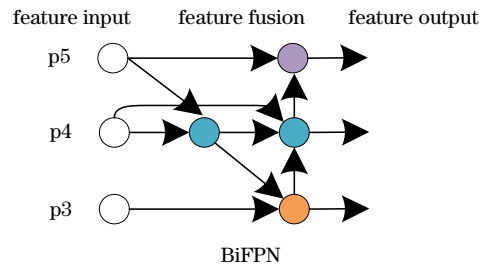


图 3 PANet 和 BiFPN 的结构

Fig. 3 Structure of PANet and BiFPN

2.3 损失函数改进

所提方法采用一种基于动态非单调聚焦机制的边界框损失, 以实现海上船舶检测中预测框与真实框更准确的边框定位损失计算。

在目标检测中, 边界框回归是决定目标定位性能的关键步骤, 可以很好地提升模型的检测精度和收敛

速度。在 YOLOv7 模型中, 损失被用作边界框回归的损失函数。在 YOLOv7 的基础上增加了对纵横比一致性的考虑^[22], 进一步增加了预测框和实际框之间长度与宽度之比的惩罚项, 公式为

$$\mathcal{L}_{\text{CloU}} = 1 - R_{\text{IoU}} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \left[\frac{v}{(1 - R_{\text{IoU}}) + v} \right] v, \quad (1)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2, \quad (2)$$

式中: b 和 b^{gt} 表示预测框和实框的中心点; ρ 表示预测框与实框中心点之间的欧几里得距离; c 表示既可以包含预测框又可以包含真实框的最小闭合区域的对角线距离; w^{gt} 和 h^{gt} 是真实框的宽度和高度; w 和 h 表示预测框的宽度和高度。

完全交并比(CIoU)考虑到了预测帧和真实帧的尺寸信息,如重叠面积、质心距离、纵横比等,但由于训练数据不可避免地包含低质量的样例,距离和纵横比等几何因素会加重对低质量样例的惩罚,从而降低模型的泛化性能。因此选择使用 Wise IoU,其构建了距离注意力,很好地解决了此问题。Wise IoU 包含 3 个版本,是 Tong 等^[23]于 2023 年在 CVPR 会议上提出的一种损失函数。Wise IoU v1 的公式定义为

$$\mathcal{L}_{\text{Wise IoU v1}} = R_{\text{Wise IoU}} \mathcal{L}_{\text{IoU}}, \quad (3)$$

$$R_{\text{Wise IoU}} = \exp \frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}. \quad (4)$$

为了防止 $R_{\text{Wise IoU}}$ 产生阻碍收敛的梯度,将 W_g 和 H_g 从计算图(上标*表示此操作)中分离,有效地消除了阻碍收敛的因素。Wise IoU 因为没有对纵横比进行计算反而有更快的速度,Wise IoU 的计算耗时为 CIoU 的 87.2%。Wise IoU v2 类似 Focal EIoU v1^[24],构造出了基于动态梯度增益的单调聚焦机制的边界框损失函数,公式为

$$\mathcal{L}_{\text{Wise IoU v2}} = \left(\frac{\mathcal{L}_{\text{IoU}}^*}{\mathcal{L}_{\text{IoU}}} \right)^{\gamma} \mathcal{L}_{\text{Wise IoU v1}}, \quad (5)$$

式中: $\overline{\mathcal{L}_{\text{IoU}}}$ 为 \mathcal{L}_{IoU} 的均值,该均值作为归一化因子; $\left(\frac{\mathcal{L}_{\text{IoU}}^*}{\mathcal{L}_{\text{IoU}}} \right)^{\gamma}$ 为梯度增益,动态更新归一化因子使该梯度增益保持在较高水平,从而解决训练后期速度慢的问题。

Wise IoU v3 则重新定义离群度以描述锚框的质量,利用该离群度构造了非单调聚焦系数,让梯度增益随着损失值的增加呈非单调变化。Wise IoU v3 公式为

$$\beta = \frac{\mathcal{L}_{\text{IoU}}^*}{\mathcal{L}_{\text{IoU}}} \in (0, +\infty), \quad (6)$$

$$\mathcal{L}_{\text{Wise IoU v3}} = r \mathcal{L}_{\text{Wise IoU v1}}, \quad r = \frac{\beta}{\delta \alpha^{\beta - \delta}}, \quad (7)$$

式中: β 为离群度,离群度小意味着锚框质量高,对离群度较大的锚框分配较小的梯度增益,将有效防止低质量示例产生较大的有害梯度; r 为非单调聚焦系数,由于 $\overline{\mathcal{L}_{\text{IoU}}}$ 是动态的,锚框的质量划分标准也是动态的,这使得 Wise IoU v3 在每一时刻都能做出最符合当前情况的梯度增益分配策略。

3 实验结果与分析

3.1 实验环境搭建

实验操作平台为 Windows11 系统,采用

PyTorch1.10.1 深度学习框架、CUDA 11.3 和 CUDNN 8.2,CPU 的型号为 Inter 13 代 13400f,显卡(GPU)的型号为 NVIDIA GeForce RTX 3060,显卡内存 12 GB,编程语言为 Python。在训练阶段使用 SGD 优化器,初始学习率为 1×10^{-5} ,权重衰减为 5×10^{-3} ,此外还有 3 个 0.8 动量的预热期和余弦退火方法来衰减学习率,每个实验有 200 个 epoch,批量大小为 16。模型的训练过程耗时约 10 h。

3.2 数据集

实验数据集来自 infiRay 公司发布的红外开源数据库中的红外海上船舶数据集。该数据库包含了多个不同场景、不同时段、不同分辨率的海上、港口和海面上的邮轮、散货船、军舰、帆船、皮划艇、集装箱船、渔船目标,共计 8000 张图片。数据库使用 liner、bulk carrier、warship、sailboat、canoe、container ship、fishing boat 分别作为邮轮、散货船、军舰、帆船、皮划艇、集装箱船、渔船的标签,数据集按照 7:2:1 的比例划分。

3.3 评价指标

使用损失函数曲线、平均精度均值(mAP)、每秒帧数、参数量(Parameter)和浮点排列(FLOPs)作为评价指标来描述每个网络的性能。

查准率(precision)和查全率(recall)往往是一对相互矛盾的性能指标。因此,引入 mAP 参数综合两个参数来检测网络的算法性能。mAP 的计算公式为

$$R_{\text{mAP}} = \frac{\sum_{k=1}^N P(k) \Delta R(k)}{C}, \quad (8)$$

式中: C 表示数据集中待检测目标的类别数目;本实验中, $N=7$; $P(k)$ 是同时识别 k 个样本的数量,查准率的大小; $\Delta R(k)$ 表示检测样本数从 $k-1$ 到 k 时查全率的变化。

3.4 目标检测网络对比实验结果

在选择海上船舶的目标检测模型时,将现有的流行目标检测模型应用于船舶数据集进行训练和测试,并比较了查准率、查全率、F1 得分、mAP@0.5 等评估指标,最后选择 YOLOv7 作为本研究中的目标检测模型,用于后续实验,实验结果如表 2 所示。如表 2 所示, YOLOv7 的整体性能优于其他测试检测算法。

3.5 消融实验

为了对 YOLOv7 算法与所提改进算法进行比较,保证实验的可靠性,所有训练数据和测试数据都在相同的训练环境下进行训练,并保证每次训练的 epoch 为 200。训练完成后,使用最优模型权值进行测试。

为了验证各种改进策略的有效性,在船舶数据集上进行消融实验,其中“√”表示使用该方法。以原始 YOLOv7 算法为基础模型,依次引入 MobileNetv3、BiFPN 模块、SENet 模块、Wise IoU 损失函数。测试集上,改进后的 YOLOv7 算法的评价指标数据如表 3 所示。

表 2 目标检测网络对比实验

Table 2 Comparison experiment of target detection networks

Method	Precision / %	Recall / %	F1 score	mAP@0.5 / %	mAP@0.5:0.95 / %	Speed / (frame·s ⁻¹)
CenterNet	88.16	91.21	0.90	91.71	58.80	33
SSD	82.03	80.50	0.81	86.86	42.80	39
EfficientDet	83.66	88.98	0.86	91.85	57.40	26
RetinaNet	85.00	86.17	0.85	90.88	53.30	13
YOLOv3	87.35	75.66	0.81	84.87	39.75	20
YOLOv4s	89.26	78.04	0.83	86.71	41.30	22
YOLOv5s	91.50	88.70	0.90	92.78	60.88	62
YOLOv7	91.80	89.64	0.91	93.11	61.45	60
YOLOv8s	92.20	87.33	0.90	92.21	64.72	105

表 3 消融实验结果

Table 3 Results of ablation experiments

Alias	Improvement strategy				Precision	Recall	mAP@0.5	mAP@0.5:0.95
	MobileNetv3	BiFPN	SENet	Wise IoU				
A	×	×	×	×	0.914	0.884	0.923	0.609
B	√	×	×	×	0.894	0.847	0.899	0.582
C	√	√	×	×	0.911	0.884	0.919	0.604
D	√	√	√	×	0.921	0.894	0.928	0.613
E	√	√	√	√	0.923	0.909	0.935	0.621

为进一步验证更多的可能性,在引入 MobileNetv3 和 BiFPN 的基础上将其他的注意力机制加入同样的位置上,在该数据集上进行多组对比实验,实验结果如表 4 所示。分析表 4 实验结果可知:相比引入 MobileNetv3 和 BiFPN 的基础网络,引入 SENet、CA 和 CBAM 的网络的 mAP@0.5 提升了 0.9 个百分点、0.8 个百分点和 0.6 个百分点;SENet 的表现效果最佳,即 SENet 注意力模块能够有效提升所提算法对船舶检测的精度。由表 5 可见:相比原始 YOLOv7,改进后的模型参数量降低至 22.9×10^6 ,降低了约 38.4%;FLOPs 降至 36.3×10^9 ,降低了约 65.5%;在 GPU 端的推理速度提升了 1 倍多,在 CPU 端的推理速度提升了 2.5 倍多,大幅降低了模型计算量和参数量,同时大幅加快推

表 4 注意力机制对比实验

Table 4 Comparative experiments of attention mechanisms

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv7-C	0.911	0.884	0.919	0.604
YOLOv7-C+SENet	0.921	0.894	0.928	0.613
YOLOv7-C+CA	0.925	0.890	0.927	0.612
YOLOv7-C+CBAM	0.928	0.886	0.925	0.615

理速度,特别是 CPU 端。

图 4 为模型训练 200 个 epoch 后,改进前后算法在红外船舶数据集上训练的定位损失 (Box_Loss)、分类损失 (Cls_Loss) 和置信度损失 (Obj_Loss)。由图 4 可得,相比原始 YOLOv7 算法,改进后算法的三种损失

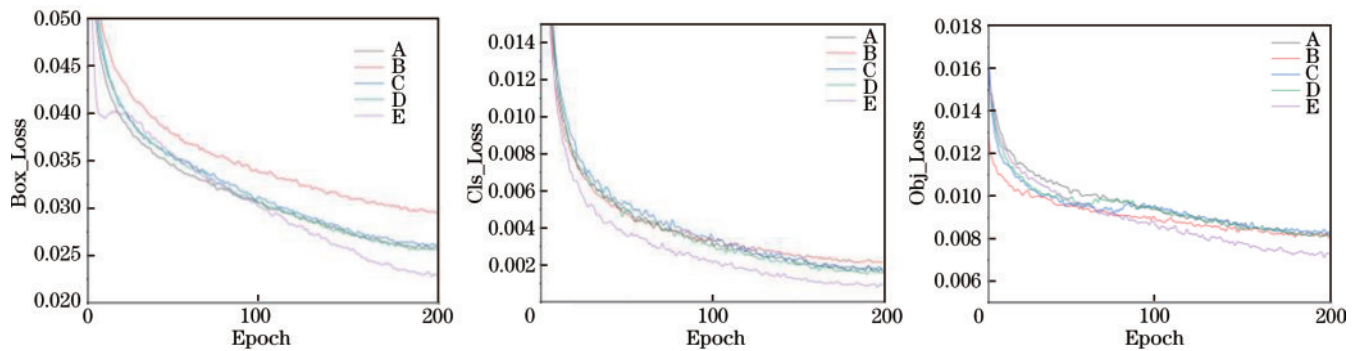


图 4 训练损失函数曲线图

Fig. 4 Training loss function graphs

表 5 轻量化结果

Table 5 Lightweighting results

Model	Inference time for GPU /ms	Inference time for CPU /ms	FLOPs / 10^9	Parameter / 10^6
YOLOv7	12.1	554.6	105.2	37.2
Proposed algorithm	5.1	150.5	36.3	22.9

函数的收敛速度均更快,并且在模型损失曲线收敛时损失值更小,这表明改进后的算法提升了模型收敛的速度,更容易训练。

3.6 船舶检测效果

改进后的目标检测算法能有效识别船舶,检测效果如图 5 所示。



图 5 船舶检测效果图

Fig. 5 Ship inspection renderings

4 结 论

利用轻量级网络 MobileNetv3,重新设计 YOLOv7 的基础网络和特征提取网络,提出一种可用于海上船舶的实时目标检测算法。所提算法在海上船舶检测任务上取得了较好的检测效果。为了提高检测精度,提出了 SE-YOLOv7 算法,该模型在船舶数据集上的平均精度均值为 93.5%,参数量仅为 22.9×10^6 ,FLOPs 为 36.3×10^9 ,在 GPU 上测得的处理速度为 196 frame/s,在 CPU 上测得的处理速度为 6.65 frame/s。该模型能够满足对船舶的实时、快速、准确的智能化检测要求,大幅降低检测人员的工作量。未来的研究将继续优化所提算法的网络结构,并将网络部署到自动化监控使用的硬件环境中。

参 考 文 献

- [1] Hu J M, Zhi X Y, Zhang W, et al. Salient ship detection via background prior and foreground constraint in remote sensing images[J]. *Remote Sensing*, 2020, 12(20): 3370.
- [2] Hu J M, Zhi X Y, Shi T J, et al. PAG-YOLO: a portable attention-guided YOLO network for small ship detection[J]. *Remote Sensing*, 2021, 13(16): 3059.
- [3] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//*Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, June 23-28, 2014, Columbus, OH, USA. New York: ACM Press, 2014: 580-587.
- [4] Girshick R. Fast R-CNN[C]//*2015 IEEE International Conference on Computer Vision (ICCV)*, December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2016: 1440-1448.
- [5] Ren S Q, He K M, 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.
- [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)*, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [7] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2023-02-03]. <https://arxiv.org/abs/1804.02767>.
- [8] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2023-03-02]. <https://arxiv.org/abs/2004.10934>.
- [9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9905: 21-37.
- [10] Miao R, Jiang H X, Tian F Z. Robust ship detection in infrared images through multiscale feature extraction and lightweight CNN[J]. *Sensors*, 2022, 22(3): 1226.
- [11] 王慧赢,王春平,付强,等.面向嵌入式平台的轻量级光学遥感图像舰船检测[J]. *光学学报*, 2023, 43(12): 1212001. Wang H Y, Wang C P, Fu Q, et al. Lightweight ship detection based on optical remote sensing images for embedded platform[J]. *Acta Optica Sinica*, 2023, 43(12): 1212001.
- [12] 徐志京,柏雪.基于双重特征增强的遥感舰船小目标检测[J]. *光学学报*, 2022, 42(18): 1828002. Xu Z J, Bai X. Small ship target detection method for remote sensing images based on dual feature enhancement [J]. *Acta Optica Sinica*, 2022, 42(18): 1828002.
- [13] Liu S W, Kong W M, Chen X F, et al. Multi-scale ship detection algorithm based on a lightweight neural network for spaceborne SAR images[J]. *Remote Sensing*, 2022, 14 (5): 1149.
- [14] Ye J, Yuan Z Y, Qian C, et al. CAA-YOLO: combined-attention-augmented YOLO for infrared ocean ships

- detection[J]. *Sensors*, 2022, 22(10): 3782.
- [15] 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[EB/OL]. (2022-07-06) [2023-02-03]. <https://arxiv.org/abs/2207.02696>.
- [16] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Republic of Korea. New York: IEEE Press, 2020: 1314-1324.
- [17] Ma N N, Zhang X Y, Zheng H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design [M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer Vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11218: 122-138.
- [18] Han K, Wang Y H, Tian Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1577-1586.
- [19] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [20] Tan M X, Pang R M, Le Q V. EfficientDet: scalable and efficient object detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 10778-10787.
- [21] 周旗开, 张伟, 李东锦, 等. 基于改进 YOLOv5s 的光学遥感图像舰船分类检测方法[J]. *激光与光电子学进展*, 2022, 59(16): 1628008.
Zhou Q K, Zhang W, Li D J, et al. Ship classification and detection method for optical remote sensing images based on improved YOLOv5s[J]. *Laser & Optoelectronics Progress*, 2022, 59(16): 1628008.
- [22] 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.
- [23] Tong Z J, Chen Y H, Xu Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. (2023-01-24) [2023-02-03]. <https://arxiv.org/abs/2301.10051>.
- [24] Zhang Y F, Ren W Q, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. *Neurocomputing*, 2022, 506: 146-157.