

基于 YOLOv8-n 的多光谱红外舰船检测算法

赵佳乐¹, 娄树理¹, 林超²

(1. 烟台大学 物理与电子信息学院, 山东 烟台 264005; 2. 海军航空大学 航空作战勤务学院, 山东 烟台 264000)

摘要:为解决在复杂战场环境下准确检测出舰船目标这一技术难题,结合舰船目标与背景、干扰物在不同谱段之间存在信息互补的特点,本文提出了一种基于 YOLOv8-n 的多光谱红外舰船检测算法。首先将骨干网络设计成多流网络,分别提取多光谱图像特征,为后续的特征融合奠定基础;其次在特征提取阶段构造交叉融合残差块 SaCF,通过 Transformer 的自注意力机制捕获不同谱段之间的潜在关联来增强单谱段特征图;最后设计了基于 CBAM 注意力机制的分级自适应融合模块 AFM,通过注意力机制生成融合权重指导多光谱特征的融合过程,提高网络对舰船目标的识别能力。实验结果表明,相较于单谱段舰船目标检测算法,本文提出的多光谱融合舰船目标检测算法的检测精度可以达到 94.8%,提高了 6.1%,该算法具有较好的检测性能,能够胜任复杂环境下的舰船目标检测任务。

关键词: 红外舰船; 多光谱目标检测; Transformer; 注意力机制; 特征融合

中图分类号: TN291; TP391 **文献标识码:** A **DOI:** 10.3969/j.issn.1001-5078.2025.04.021

Multispectral infrared ship detection algorithm based on YOLOv8-n

ZHAO Jia-le¹, LOU Shu-li¹, LIN Chao²

(1. School of Physics and Electronic Information, Yantai University, Yantai 264005, China;
2. Aviation Operations and Service Institute, Naval Aviation University, Yantai 264000, China)

Abstract: Aiming at the technical challenge of accurately detecting ship targets in complex battlefield environments, combined with the characteristics of information complementarity between ship targets and background and interferences in different spectral bands, a multispectral infrared ship detection algorithm based on YOLOv8-n is proposed in this paper. Firstly, the backbone network is designed as a multi-stream network to extract multispectral image features separately, setting the stage for subsequent feature fusion. Secondly, the cross-fertilization residual block (SaCF) is constructed in the feature extraction stage, and the single-spectral segment feature maps are enhanced by capturing the potential correlation between different spectral segments through the self-attention mechanism of the Transformer. Finally, an adaptive hierarchical fusion module (AFM) based on the attention mechanism is designed, which guides multispectral feature fusion by generating fusion weights through the attention mechanism to improve the network's ability to recognize ship targets. The experimental results show that compared with the single-spectrum ship target detection algorithm, the detection accuracy of the multispectral fusion ship target detection algorithm proposed in this paper can reach 94.8%, which is 6.1% higher. The improved algorithm exhibits superior detection performance and is capable of handling ship target detection tasks in complex environments.

Keywords: infrared ships; multispectral object detection; Transformer; attention mechanisms; feature fusion

基金项目: 国家自然科学基金项目 (No. 62005318) 资助。

作者简介: 赵佳乐 (2001 -), 男, 硕士研究生, 主要从事深度学习目标检测等方面的研究。E-mail: 1659443934@qq.com

通讯作者: 娄树理 (1976 -), 男, 副教授, 硕士生导师, 主要从事机器视觉算法研究与实现方面的研究。E-mail: shulilou@ sina.com

收稿日期: 2024-06-27

1 引言

在现代化海战中,红外成像制导技术因其具备全天候工作能力、反隐身伪装识别和抗干扰能力等多项优点而得到广泛应用。然而,随着精确制导武器作战环境的进一步复杂化和隐身技术的持续提升,单一谱段的红外成像制导易受到背景、干扰物、海杂波等诸多因素的影响,难以适应日益复杂的战场环境。多光谱红外舰船目标检测可以集成多个谱段图像的互补信息,增强目标、抑制背景,提高舰船目标检测的准确性和鲁棒性。

传统的多光谱红外舰船目标检测算法采用手动方式提取多谱段特征并融合来训练分类器,用于区分舰船和背景。仇荣超^[1]等基于高斯模型对五个中红外波段的特征向量进行融合,采用线性 SVM 分类器进行舰船目标的识别。传统的多光谱检测算法存在特征提取能力弱、融合方法复杂和检测效率低等问题。随着深度学习的蓬勃发展,基于深度学习的目标检测算法因其高精度、高效率、鲁棒性较好的特点被广泛应用于多光谱目标检测领域^[2-4]。基于深度学习的目标检测算法主要分为两大类:一类为双阶段目标检测算法,例如 R-CNN^[5]和 Faster R-CNN^[6]等算法,此类算法检测速度较慢。另一类为单阶段目标检测算法,如 SSD^[7]和 YOLO^[8-12]系列算法。方康等^[13]采用多尺度特征增强模块提取多光谱特征图信息,设计了一种基于 YOLOX 的多光谱行人检测网络;孙昆等^[14]引入条件卷积对融合后的特征信息进行动态处理;Bao 等^[15]基于 YOLOV7 算法,在多光谱特征融合过程中加入扩展卷积,增加特征融合的感受野。这类算法的融合策略多是采用通道堆叠和简单相加,忽略了多光谱特征之间的差异性。目前,基于深度学习的红外舰船检测算法都是针对单谱段红外舰船图像,李海军等^[16]基于坐标注意力机制提出了改进 YOLOV5s 的单谱段红外舰船检测算法,单谱段红外舰船目标算法受天气和背景干扰影响较大,容易出现误检和漏检的情况。

针对传统的多光谱红外舰船目标检测算法存在融合过程复杂、检测效率低等问题,本文结合多光谱特征融合算法和深度学习目标检测算法,选择一阶段无锚框检测算法 YOLOV8-n 作为基础检测器,提出了一种基于 YOLOV8-n 的多光谱红外舰船检测算法,首先采用多流特征提取网络 MF-CSPDarknet53

替换掉 CSPDarknet53 网络,实现多光谱图像的特征提取,为后续的特征融合奠定基础;其次在特征提取阶段构造基于自注意力机制的交叉融合残差块 SaCF,通过 Transformer 的自注意力机制捕获不同谱段之间的潜在关联来增强单谱段特征图;最后设计了一种基于 CBAM 注意力机制的分级自适应融合模块 AFM,通过注意力机制生成融合权重指导多光谱特征的融合过程,提高网络对舰船目标的识别能力,降低舰船目标的误检率和漏检率。

2 多光谱图像特性分析

本研究通过成像仪采集了多组红外舰船图像,分辨率为 320×256 ,包含 $4.4 \sim 4.8 \mu\text{m}$ 、 $3.7 \sim 3.9 \mu\text{m}$ 、 $3.7 \sim 4.8 \mu\text{m}$ 、 $3.7 \sim 4.1 \mu\text{m}$ 四个谱段(分别标记为 B1、B2、B3、B4)。典型复杂背景下的多光谱红外舰船目标图像如图 1 所示,从舰船目标和背景的红外特性上分析,目标和背景反射红外波谱与太阳光谱比较接近,主要集中在小于 $4 \mu\text{m}$ 的谱段,因此, B2、B3 和 B4 谱段图像更易受太阳光谱辐射的影响; B3 谱段是一个中红外宽谱段,能量信号最强;舰船目标与岸岛背景、海天背景和海杂波在不同谱段的光谱辐射差异表现各不相同^[17],多谱段之间存在信息互补的特点,使得多光谱图像的特征融合有益于复杂环境下的舰船目标检测。

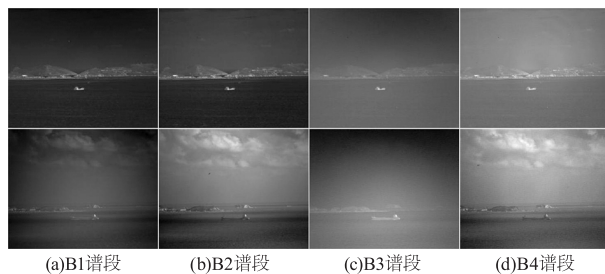


图 1 多光谱图像

Fig. 1 Multi-spectral images

3 YOLOV8-n 算法模型改进

YOLOV8 是 Ultralytics 公司推出的一阶段检测算法,包含主干网络、颈部网络和检测头网络。其主干网络采用 CSPDarknet53 作为特征提取网络,用于提取多尺度特征图以便检测不同大小的目标。YOLOV8 的颈部网络为 ELAnet^[18]结构,通过先自顶向下后自底向上的方式,将多尺度特征进行双向融合,建立多尺度特征图之间的联系。检测头网络将融合后的多尺度特征聚合到相同尺寸,然后分别进行类别和位置的预测。

为了对多光谱舰船目标图像进行特征融合检测,本文将YOLOV8的骨干网络设计成多流特征提取网络MF-CSPDarknet53,算法的网络结构如图2所示,在C3、C4、C5层后分别对提取的多光谱特征进行融合。为了捕获多光谱特征之间的潜在关联,

本文在特征提取阶段嵌入自注意力交叉融合模块(Self-attention Cross Fusion, SaCF)。同时,针对多光谱特征融合不充分的问题,本文设计了一种基于CBAM注意力机制的分级自适应融合模块(Adaptive Fusion Model, AFM)。

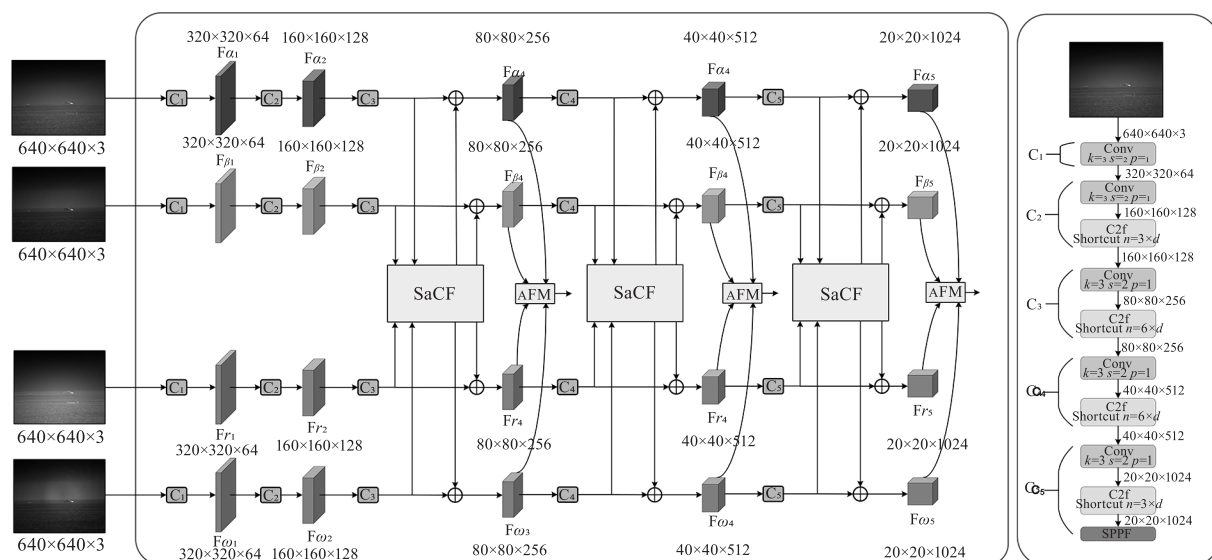


图2 多光谱检测网络

Fig. 2 Multispectral detection network

3.1 SaCF 自注意力交叉融合残差块

由于配准后的多光谱舰船图像之间存在潜在的关联性,4个谱段图像的全局语义信息基本一致,如果仅在简单的特征提取之后就进行特征融合,并不能完全利用多光谱图像的互补性。基于自注意力机制的Transformer可以建模长距离依赖关系,实现全谱段信息的关联^[19],因此,本文提出了一种基于Transformer的交叉融合模块SaCF(Self-attention Cross Fusion),如图3所示,同时,为了加强多谱段之间的关联性,本文将SaCF模块输出的特征图以残差的方式重新映射到原特征图上,以此丰富单谱段特征图的信息量。具体步骤如下:

①将主干网络提取出的多谱段特征图 $F_\alpha, F_\beta, F_\gamma, F_\omega$ ($C \times H \times W$)逐个展开并链接,得到输入序列 I ($4HW \times C$),并嵌入位置编码($4HW \times C$),位置编码在模型训练时可以标记序列的相对位置。将输入序列 I 分别映射到三个权重矩阵 W^Q, W^K, W^V 上并相乘得到查询 Q 、键 K 和值 V ^[19]:

$$\begin{cases} Q = IW^Q \\ K = IW^K \\ V = IW^V \end{cases} \quad (1)$$

②自注意力层通过计算矩阵 Q 和 K 之间的内积

并进行SoftMax归一化得到注意力权重矩阵,其次将注意力权重矩阵与值 V 相乘得到输出矩阵 Z :

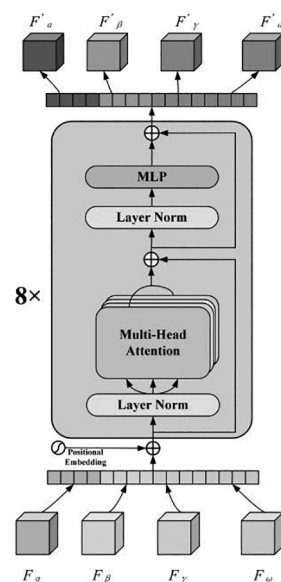


图3 自注意力交叉融合网络结构

Fig. 3 Self-attention cross fusion network structure

$$Z = \text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{D_K}}\right)V \quad (2)$$

式中, $\sqrt{D_K}$ 是一个缩放因子,用于防止内积过大。

③采用多头注意力机制捕捉不同位置和不同子

空间中的复杂关系,得到增强矩阵 Z' :

$$\begin{cases} Z' = \text{MultiHead}(Q, K, V) = \text{Concat}(Z_1, \dots, Z_h) W^O, \\ Z_i = \text{Attention}(Q W_i^Q, K W_i^K, V W_i^V) \end{cases} \quad (3)$$

式中,下标 h 表示注意力头的数量; W^O 表示 $\text{Concat}(Z_1, \dots, Z_h)$ 的投影矩阵。

④使用含 GELU 激活函数的两层全连接前馈神经网络来计算输出序列 O ($4HW \times C$):

$$\begin{aligned} O &= \text{MLP}(Z' + I) + Z' + I, \\ &= \text{FC}_2(\text{GELU}(\text{FC}_1(Z' + I))) + Z' + I \end{aligned} \quad (4)$$

⑤将输出序列 O 切分并逐个折叠为原始尺寸 ($C \times H \times W$),得到关联后的特征图 $F'_\alpha, F'_\beta, F'_\gamma, F'_\nu$ 并以残差的方式映射到原特征图上,得到增强后的特征图 $F''_\alpha, F''_\beta, F''_\gamma, F''_\nu$:

$$\begin{cases} F''_\alpha = F_\alpha + F'_\alpha \\ F''_\beta = F_\beta + F'_\beta \\ F''_\gamma = F_\gamma + F'_\gamma \\ F''_\nu = F_\nu + F'_\nu \end{cases} \quad (5)$$

由于 Transformer 的计算复杂度较高,考虑到内

存访问的限制^[20],为了减少参数量和计算复杂度,本文将输入 SaCF 模块的特征图下采样到低分辨率 ($H = W = 8$),SaCF 模块输出的特征图使用双线性插值恢复到原始分辨率。

3.2 基于 CBAM 注意力机制的自适应分级融合

舰船目标与背景、干扰物在不同谱段的反射特性、光谱辐射的差异性表现不同。背景、干扰物的反射特性、光谱辐射在相邻子谱段之间差异较为明显,特征信息相关性较低;而舰船目标区域在相邻子谱段之间差异较小,特征信息具有高度相关性;如果只是等比例融合多光谱特征图,往往忽略了这种差异性。为了突出目标特征,抑制杂波和背景特征,结合注意力机制可以调整特征图关注信息的特点,本文引入 CBAM(Convolutional Block Attention Module)注意力机制^[21],设计了一个注意力指导的自适应融合模块(Adaptive Fusion Model, AFM),采用分级融合策略,将 4 个谱段的特征图按层级进行自适应融合,如图 4 所示。

$$F_f = F_3(F_2(F_1(F_\alpha, F_\beta), F_\gamma), F_\nu) \quad (6)$$

式中, $F_{\alpha, \beta, \gamma, \nu}$ 代表不同谱段的特征图; F_f 代表融合后的特征图; F_i 代表自适应融合函数。

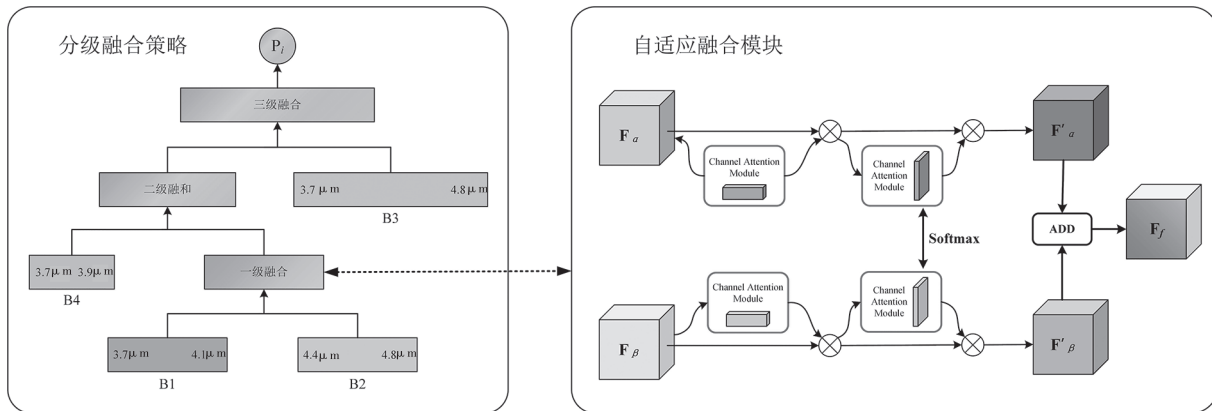


图 4 分级自适应融合模块

Fig.4 Hierarchical adaptive fusion module

在每个层级的融合过程中,两个谱段的特征图分别通过注意力机制生成各自的注意力权重矩阵。

$$A_\alpha = \Phi_{A_\alpha}(F_\alpha) \quad (7)$$

$$A_\beta = \Phi_{A_\beta}(F_\beta) \quad (8)$$

式中, A_α 和 A_β 分别表示 B1 谱段特征图和 B2 谱段特征图的注意力向量, $\Phi_{A_\alpha}(\cdot)$ 和 $\Phi_{A_\beta}(\cdot)$ 表示跨谱段注意力权重矩阵的生成过程,然后对两个跨谱段注意力权重矩阵进行 Softmax 归一化实现信息交互。

$$A_{c\alpha}, A_{c\beta} = \Phi_{CA}(A_\alpha, A_\beta) \quad (9)$$

式中, $\Phi_{CA}(\cdot)$ 表示跨谱段注意力权重矩阵归一化过程; $A_{c\alpha}$ 和 $A_{c\beta}$ 分别表示归一化后的注意力权重矩阵,归一化后的注意力权重矩阵对单谱段特征图进行重新标定即赋予单谱段特征图融合权值,最后将多谱段特征图进行加权求和。

$$F_f = \Phi_{\text{scale}}(F_\alpha, A_{c\alpha}) + \Phi_{\text{scale}}(F_\beta, A_{c\beta}) \quad (10)$$

式中, F_f 表示加权融合后的特征图; $\Phi_{\text{scale}}(\cdot)$ 表示赋予单谱段融合权值的过程。

3.3 损失函数

损失函数用于衡量预测值与真实值之间的差距,通过最小化损失函数可以调整模型参数、优化模型性能,总损失函数(L_{total})包括边界框回归损失函数(L_{IOU})、类别损失函数(L_{cls})和置信度损失函数(L_{conf})。

$$L_{\text{total}} = L_{\text{IOU}} + L_{\text{cls}} + L_{\text{conf}} \quad (11)$$

YOLOv8n 原网络使用 L_{CIOU} 作为边界框回归损失,但 CIOU 不能有效的测量目标盒与锚框之间的差异,这导致边界框回归损失优化的收敛速度慢,定位不准确。针对上述问题,本文引入 EIOU loss^[22],EIOU 通过拆分预测框和真实框的纵横比影响因子,分别计算预测框和真实框的长和宽, L_{EIOU} 的定义如下^[22]:

$$L_{\text{EIOU}} = L_{\text{IOU}} + L_{\text{dis}} + L_{\text{asp}} \\ = 1 - \text{IOU} + \frac{\rho^2(b, b^{\text{gt}})}{(w^e)^2 + (h^e)^2} + \frac{\rho^2(w, w^{\text{gt}})}{(w^e)^2} + \frac{\rho^2(h, h^{\text{gt}})}{(h^e)^2} \quad (12)$$

式中, w_e 和 h_e 是最小封闭框的宽度和高度;EIOU 损失函数包括 IOU 损失(L_{IOU})、中心点距离损失(L_{dis})和方向损失(L_{asp}),通过拆分预测框与真实框的纵横比影响因子,使目标盒与锚框宽度和高度的差值最小,加速模型的收敛,定位更加准确。

4 实验结果与分析

4.1 实验环境和数据集

本文的仿真实验在 Windows11 操作系统下完成,使用的 GPU 版本为 NVIDIA GeForce GTX 1650;CPU 版本为 i7-9750H;编程语言为 Python 3.8;深度学习框架为 Pytorch 1.10。并且利用 CUDA 11.2 进行加速,实验中的批量大小(Batch size)设置为 8,训练轮数(Epoch)设置为 300,优化器选用 Adam。所使用的数据集包含实际采集的 4 组不同谱段红外舰船图像,每个谱段 1143 张舰船图像,共 4572 张多光谱舰船图像,图像尺寸为 320×256 ,其中每个谱段训练集 914 张图像,测试集 229 张图像。

4.2 评价指标

为了综合评价所提算法在舰船目标检测任务中的准确性、实时性,本文采用平均精度 AP_{50} 、 $\text{AP}_{50 \sim 95}$ 和检测速度(frame per second, FPS)作为评价指标,其中 AP(average precision)是一种综合考虑精准率(precision)和召回率(recall)的指标,常用来评估模型在目标检测任务中的准确性,FPS 为模型每秒能处理的图

像帧数,常用来评估模型的实时性。本文设定多个 IOU 阈值(0.50 ~ 0.95, 0.05 为步长),每一个 IOU 阈值下都有一个 AP 值,求不同 IOU 阈值下的平均,即 $\text{AP}_{50 \sim 95}$,并且记录 IOU 为 0.50 时的 AP 值,即 AP_{50} 。由于本文算法为单目标检测,因此 $\text{mAP} = \text{AP}$ 。

4.3 算法性能比较与分析

4.3.1 对比实验

为了验证改进的多光谱目标检测算法在舰船目标检测任务中的性能,本文进行了对比实验,对比对象包括各个谱段图像在原网络的检测结果以及第三谱段图像在主流的目标检测网络上的检测结果。鉴于数据集的制作误差,实验在 $\text{IOU} = 0.75$ 的条件下进行了精度和速度对比。对比结果详见表 1。

表 1 不同算法效果对比

Tab. 1 Comparison of different algorithms

Band	Model	$\text{mAP}_{50}/(\%)$	$\text{mAP}_{50 \sim 95}/(\%)$	FPS
B3	Faster R-CNN	88.1	58.2	12.6
B3	YOLOv5-n	87.5	57.1	68.4
B3	YOLOv8-s	90.2	60.7	45.8
B1	YOLOv8-n	87.6	56.8	72.5
B2	YOLOv8-n	86.5	53.6	72.5
B3	YOLOv8-n	88.7	58.7	72.5
B4	YOLOv8-n	85.8	54.5	72.5
Fusion	Improved YOLOv8-n	94.8	66.5	47.3

通过对比实验可以得知,改进后的 YOLOv8-n 算法的 mAP_{50} 和 $\text{mAP}_{50 \sim 95}$ 分别达到了 94.8 % 和 66.5 %,相较于改进前算法检测效果最好的 B3 图像, mAP_{50} 和 $\text{mAP}_{50 \sim 95}$ 分别提高了 6.1 % 和 7.8 %,较好地满足了复杂海面战场的需求。改进的模型速度降低的原因,是由于多光谱特征提取网络需要对 4 个谱段的图像进行特征提取,这无疑增加了较多的参数量,且 Transformer 也会增加算法的计算复杂度。虽然改进模型牺牲了一部分速度,但是达到了较理想的检测精度。

4.3.2 消融实验

为测试每个改进模块对算法的影响,本文在相同的设备和数据集上进行了消融实验,实验结果如表 2 所示,表中“√”表示添加了该模块,“-”表示未添加该模块。

由表 4 可知,对比单谱段检测算法,本文提出的多光谱融合检测网络在检测精度上有所提升,且为

后续添加增强融合模块奠定了基础。后续单独添加交叉融合模块,虽然在检测速度上有所降低,但是精度的提升较为明显。提出的基于注意力的自适应融合模块进一步提高了算法的检测精度,最后引入 EIOU 加速模型收敛。多个模块组合对比单个模块的添加,算法的检测精度有较大提升。

4.3.3 性能分析

为验证本文算法的有效性,将多光谱融合检测

效果与不同谱段的检测效果进行对比,如图 5 所示。B3 谱段为中波红外的宽谱段,其包含的特征信息更加丰富,因而基于 B3 谱段的舰船目标识别准确率往往高于其余窄谱段;而基于多谱段图像的舰船目标识别算法融合了不同谱段图像的特征信息,充分利用多光谱图像之间的互补特性,实现超高精度的舰船目标检测,且该算法抗干扰性更强,能够充分胜任复杂的海战检测任务。

表 2 消融实验

Tab.2 Ablation experiments

Fusion	AFM	SACF	EIOU	mAP ₅₀ /(%)	mAP _{50~95} /(%)	FPS
-	-	-	-	88.7	58.7	72.5
✓	-	-	-	90.6	61.8	61.2
✓	-	✓	-	92.8	64.1	50.1
✓	✓	-	-	92.2	63.7	57.6
✓	✓	✓	-	94.6	66.4	47.3
✓	✓	✓	✓	94.8	66.5	47.3

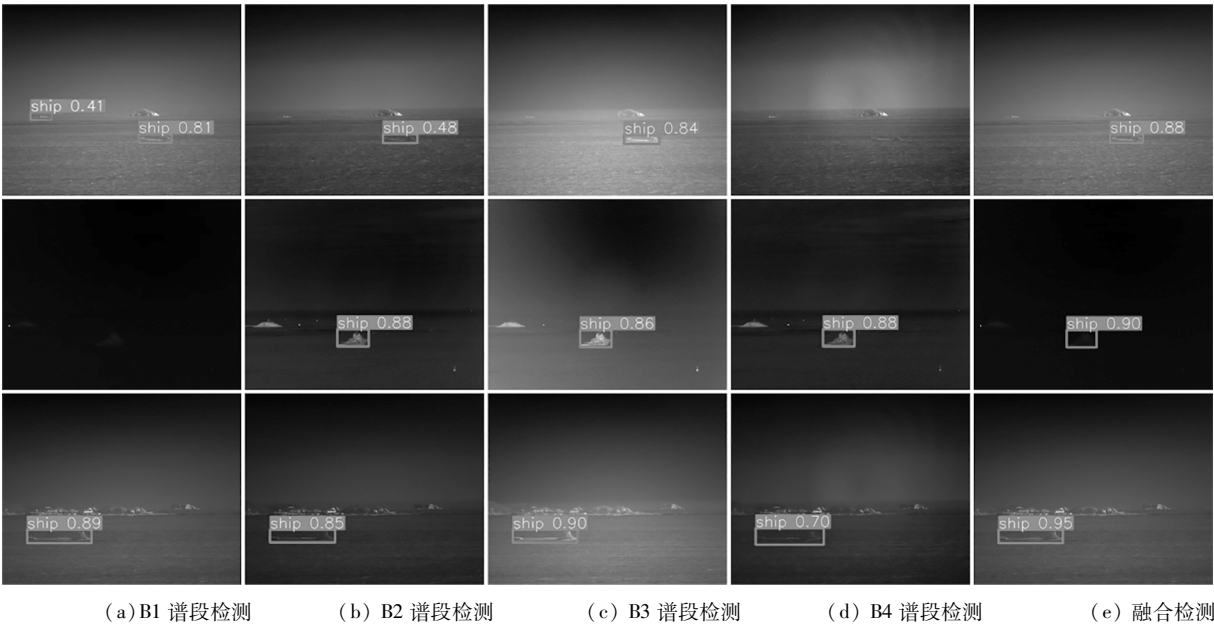


图 5 算法效果对比图

Fig.5 Comparison chart of algorithm effects

5 结 论

针对复杂环境下的舰船目标检测任务,本文在 YOLOv8-n 的基础上进行改进,提出了一种多光谱融合红外舰船检测算法,该算法使用多流特征提取网络对多光谱图像进行特征提取,同时利用 Transformer 的自注意力机制融合多光谱特征以增强单谱段特征,提高检测精度,最后使用自适应分级融合模块融合多光谱特征,进一步提高了算法的检测效果

和抗干扰能力。经过实验验证,相较于单谱段目标检测算法,本文提出的多光谱融合红外舰船目标检测算法的检测精度可以达到 94.8%,提高了 6.1%,证明该算法具有较好的检测性能,能够胜任复杂环境下的舰船目标检测任务。

参考文献:

[1] Qiu Rongchao, Lou Shuli, Li Tingjun, et al. Research on detection of ship target at sea based on multi-spectral in-

- frared images [J]. Spectroscopy and Spectral Analysis, 2019, 39(3): 698 - 704. (in Chinese)
- 仇荣超, 姜树理, 李廷军, 等. 多波段红外图像的海面舰船目标检测 [J]. 光谱学与光谱分析, 2019, 39(3): 698 - 704.
- [2] Zhang L, Liu Z, Zhang S, et al. Cross-modality interactive attention network for multispectral pedestrian detection [J]. Information Fusion, 2019, 50: 20 - 29.
- [3] Zhang H, Fromont E, Lefèvre S, et al. Guided attentive feature fusion for multispectral pedestrian detection [C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 72 - 80.
- [4] Qingyun F, Dapeng H, Zhaokui W. Cross-modality fusion transformer for multispectral object detection [DB/OL]. (2021 - 10 - 30). <http://arxiv.org/abs/2111.00278>.
- [5] A Dundar, J Jin, B Martini, et al. Embedded streaming deep neural networks accelerator with applications [J]. Neural Netw. & Learning Syst, 2017, 28(7): 1572 - 1583.
- [6] 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.
- [7] LIU W, Anguelov D, Erhan D, et al. SSD: single shot multi-box detector [C]// European Conference on Computer Vision, Springer International Publishing, 2016: 21 - 37.
- [8] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real - time object detection [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779 - 788.
- [9] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 6517 - 6525.
- [10] Ge, Zheng, Songtao Liu, Feng Wang, et al. YOLOX: exceeding YOLO series in 2021 [J]. arxiv preprint arxiv: 2107.08430, 2021.
- [11] Wang Chien-Yao, Alexey Bochkovskiy, Hong-Yuan Mark Liao. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 7464 - 7475.
- [12] Reis D, Kupec J, Hong J, et al. Real-time flying object detection with YOLOv8 [J]. arxiv preprint arxiv: 2305.09972, 2023.
- [13] Fang Kang, Huagn Qin, Wang Keqi. Detection dlgorithm of multispectral pedestrian based on improved YOLOX [J]. Journal of Chinese Computer Systems, 2024, 45(1): 185 - 191. (in Chinese)
- 方康, 黄琴, 王克琪, 等. 基于改进 YOLOX 的多光谱行人检测算法 [J]. 小型微型计算机系统, 2024, 45(1): 185 - 191.
- [14] Sun Kun, Wu Yi, Niu Yarui. Detection of multispectral pedestrian based on multi-stage cross-information fusion [J]. Electronic Measurement Technology, 2023, 46(15): 118 - 125. (in Chinese)
- 孙昆, 武一, 牛雅睿, 等. 基于多阶段交叉信息融合的多光谱行人检测 [J]. 电子测量技术, 2023, 46(15): 118 - 125.
- [15] Bao C, Cao J, Hao Q, et al. Dual-YOLO architecture from infrared and visible images for object detection [J]. Sensors, 2023, 23(6): 2934.
- [16] Li Haijun, Kong Fancheng, Lin Yun. Detection algorithm of infrared ship based on improved YOLOv5s [J]. Systems Engineering and Electronics, 2023, 45(8): 2415 - 2422. (in Chinese)
- 李海军, 孔繁程, 林云. 基于改进 YOLOv5s 的红外舰船检测算法 [J]. 系统工程与电子技术, 2023, 45(8): 2415 - 2422.
- [17] Zhan Wei, Zhang Jian, Guo Ming. A multilevel fusion method for multispectral infrared ship images [J]. Laser & Infrared, 2019, 49(2): 240 - 245. (in Chinese)
- 詹维, 张健, 郭明, 等. 一种多光谱红外舰船图像多级融合方法 [J]. 激光与红外, 2019, 49(2): 240 - 245.
- [18] Yi, Q, Dai, G, Shi, M, et al. ELANet: effective lightweight attention-guided network for real-time semantic segmentation [J]. Neural Process Lett 2023, 55: 6425 - 6442.
- [19] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C]// Advances in Neural Information Processing Systems, 2017, 30.
- [20] Sun Y, Xue B, Zhang M, et al. Completely automated CNN architecture design based on blocks [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 31(4): 1242 - 1254.
- [21] Woo S, Park J, Lee J Y, et al. Cham: Convolutional block attention module [C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3 - 19.
- [22] Yang Z, Wang X, Li J. EIoU: an improved vehicle detection algorithm based on vehiclenet neural network [J]. Journal of Physics: Conference Series, 2021, 1924(1): 012001.