

文章编号 1004-924X(2025)08-1327-12

多光谱融合的红外舰船目标轻量化检测

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:针对红外多光谱融合舰船目标检测模型体积大、效率低、难以在嵌入式设备上部署等问题,本文基于 YOLOv8n-MF 提出了一种轻量级舰船目标红外多光谱融合检测模型 YOLOv8n-MFLW。首先,引入轻量级网络 HGNetv2 替换主干网络,基于 GSConv 卷积重构 HGBlock 和 C2f 模块,在降低模型参数数量的同时保留模型特征提取和融合能力;提出一种自适应稀疏结构化剪枝算法 La-Depgraph 修剪模型,进一步大幅降低模型参数数量;最后,采用中间特征知识蒸馏学习策略恢复剪枝带来的精度损失,提高模型的检测性能。实验结果表明,相较于原模型,改进后的轻量级舰船目标红外多光谱融合检测模型的检测精度可以达到 96.4%,提高了 1.2%,模型参数数量、计算量及内存占用量仅为 0.9 MB, 3.5 GFlops 和 2.3 MB,分别减少 88.1%, 81.2% 和 82.8%。因此,本文提出的模型体积小、精度高,具有更好的检测性能,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;轻量化;模型剪枝;知识蒸馏

中图分类号:TN291;TP391 文献标识码:A

doi:10.37188/OPE.20253308.1327

CSTR:32169.14.OPE.20253308.1327

A lightweight detection of multi-spectral infrared ship target

ZHAO Jiale¹, LOU Shuli^{1*}, 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)

* Corresponding author, E-mail: shulilou@sina.com

Abstract: In order to solve the problems of large size, low efficiency, and high deployment requirements for embedded devices of infrared multi-spectral ship target detection models, a lightweight ship target infrared multi-spectral detection model YOLOv8n-MFLW was proposed. Firstly, the model replaced the backbone network with a lightweight network, HGNetv2. Based on GSConv convolution, the modules of HGBlock and C2f were reconstructed to reduce the model parameter count while retaining the model's feature extraction and fusion capabilities. A self-adaptive sparse structured pruning algorithm, La-Depgraph, was proposed to prune the model, leading to a significant reduction in the model's parameters. Finally, an intermediate feature knowledge distillation learning strategy was employed to recover the accuracy loss caused by pruning and enhance the model's detection performance. Experimental results show that compared to the original model, the improved lightweight ship target infrared multi-spectral fusion detection

收稿日期:2025-01-13;修订日期:2025-02-25.

基金项目:国家自然科学基金资助项目(No. 62005318)

model achieves a detection accuracy of 96.4%, an increase of 1.2%. The model's parameter count, computational complexity, and memory usage are only 0.9 MB, 3.5 GFlops, and 2.3 MB, respectively, reduced by 88.1%, 81.2%, and 82.8%. Therefore, the proposed model is small in size and high in accuracy, it has a better detection performance and is capable for ship target detection tasks in complex environments.

Key words: infrared ships; multispectral object detection; lightweight; model Pruning; knowledge distillation

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、强大的穿透力、远距离作用、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。红外多光谱融合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。

传统的红外多光谱融合舰船目标检测算法采用手动方式提取多谱段特征并融合训练分类器,用于区分舰船和背景,存在特征提取能力弱、融合方法复杂和检测效率低等问题。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 (Support Vector Machine) 分类器进行舰船目标的识别。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 (Region-based Convolutional Neural Network)^[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 (Single Shot MultiBox Detector)^[7] 和 YOLO (You Only Look Once)^[8-12] 系列算法。Fei 等^[13]采用跨模态差异模块来表示不同模态之间的差异,去除多模态间共享信息的冗余,设计了一种基于 YOLOv5 的多模态检测算法;Hou 等^[14]引入注意力机制对融合过程进行动态处理,提高模型获取关键特征的能力,改善融

合信息质量。上述两种算法受卷积算子非全局感受野的影响,多光谱融合过程容易忽略全局信息的关联,特征融合不充分,且参数数量和计算量较大。Qing 等^[15]首次将 Transformer 应用于多光谱融合,利用 Transformer 的自注意力机制关注全局特征指导多模态特征融合。然而,由于 Transformer 的计算复杂性,该算法冗余信息更多,模型体积较大。

针对基于深度学习的多光谱融合目标检测网络参数数量和计算量较大,冗余信息较多,难以在嵌入式设备部署的问题,本文优化网络结构,压缩模型大小,提出了一种轻量化多光谱检测模型 YOLOv8n-MFLW。本文技术路线如图 1 所示。首先,引入 RT-DETR 模型的轻量化主干网络 HGNetv2 替换原主干网络,并基于 GSConv (Group Shuffle Convolution) 卷积重构主干网络 HGBlock 和颈部网络 C2f 模块,在降低模型参数量的同时保留模型特征提取和融合能力;然后,在分组稀疏化的基础上引入 Lamp 非结构化评分算法,提出一种自适应稀疏结构化剪枝算法 La-Depgraph,对模型进行剪枝,进一步降低模型参

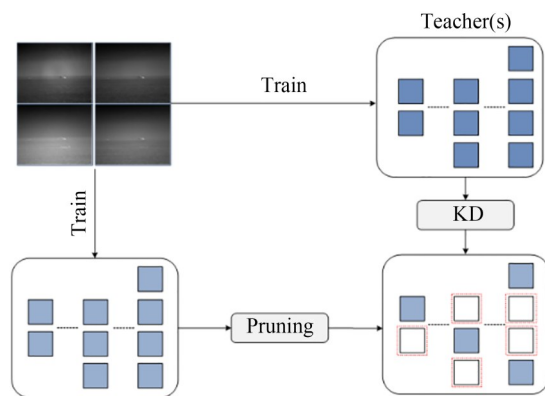


图 1 技术路线图

Fig. 1 Technical Roadmap

数量,使其更易在资源有限的硬件平台部署;最后,采用中间特征知识蒸馏恢复剪枝带来的精度损失,提高模型性能,从而获得精度高、体积小、泛化性强的轻量级红外多光谱融合舰船目标检测算法。

2 多光谱数据集与基础检测模型

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), 如图 2 所示。数据集包含太阳光谱干扰、渔船和岸岛等背景干扰和海杂波等噪声干扰, 适合复杂环境下的舰船目标检测研究。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^[16], 多谱段之间存在信息互补的特点。

本文采用红外多光谱融合目标检测网络 YOLOv8n-MF 作为基础检测模型, 网络结构如图 3 所示, 多分支主干网络可提取多光谱特征并融合, 颈部网络对融合的多尺度特征进行增强, 检测头网络负责预测多尺度目标的位置和类别。

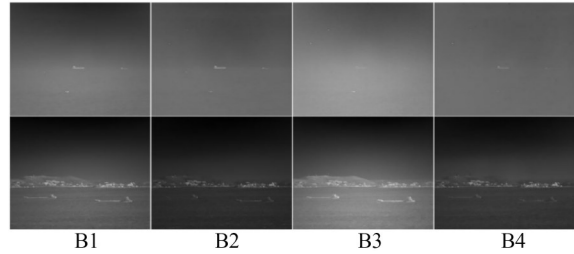


图 2 多光谱图像

Fig. 2 Multispectral images

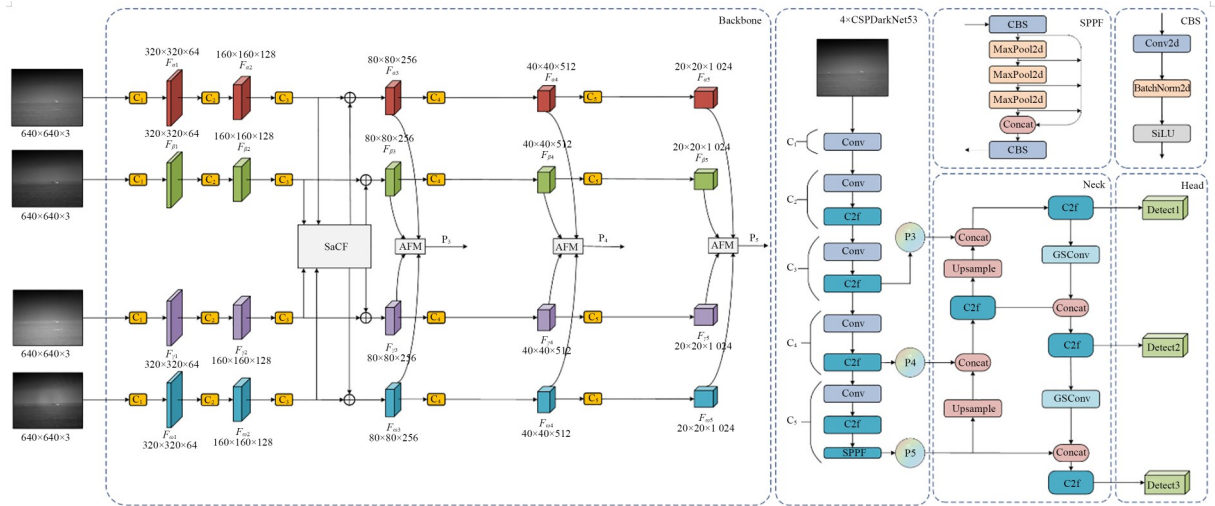


图 3 红外多光谱检测模型网络结构

Fig. 3 Network architecture of infrared multispectral detection model

3 轻量化 YOLOV8-MFLW 模型

3.1 轻量级网络结构

在实时目标识别任务中, 模型需满足嵌入式设备部署的低计算需求, 同时需兼顾复杂场景下多光谱目标的特征提取能力。YOLOv8n-MF 主干网

络 CSP-Darknet53 由多个下采样模块和固定尺度的 C2f 模块组成, 未针对多光谱目标设计差异化提取策略, 难以捕捉小渔船与舰船目标的特征差异, 且参数数量和计算复杂度较高, 不利于实时检测任务。本文引入 RT-DETR 的轻量级主干网络 HGNetv2^[17] 替换原有网络, 消除模型冗余参数, 降

低计算量,改进后的多光谱轻量级主干网络如图4所示。首先,采用HGStem多光谱降维模块作为网络的预处理层,通过降低初始特征映射的维数来提高计算效率。由于传统卷积计算过程复杂,HGNetv2网络采用深度可分离卷积DSConv(stride=2)代替传统下采样模块,通过对特征图的每个通道独立采样,进一步降低模型参数量。最

后,HGNetv2主干网络采用多尺度的HGBlock的作为网络的特征提取模块,相比于C2f模块,HGBlock通过堆叠多个卷积模块来提取多尺度特征,以较低的计算成本提取更深层次的多光谱目标特征,同时避免因参数冗余导致特征混淆,并采用一个大小为 1×1 的Conv对不同深度的特征进行混合,为后续多光谱特征融合检测奠定基础。

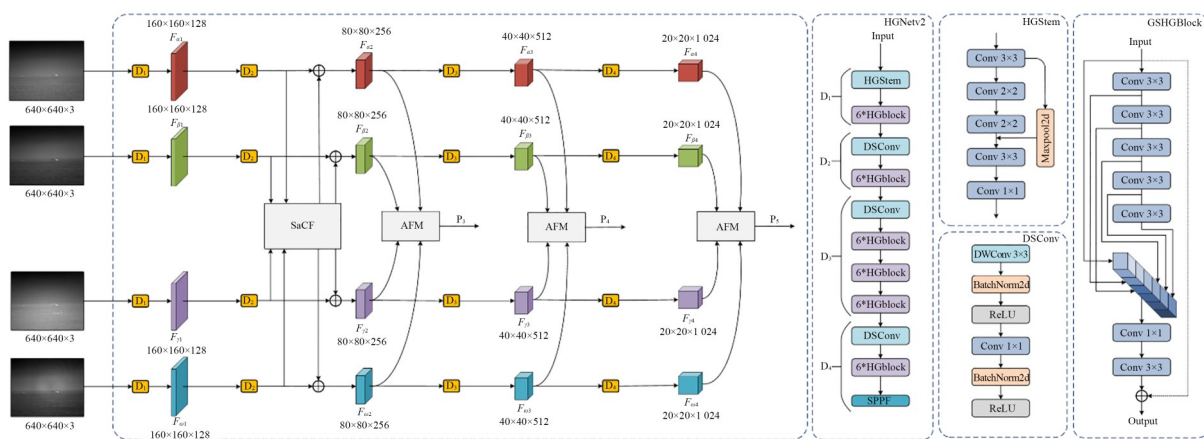


图4 HGNetv2主干网络

Fig. 4 HGNetv2 Backbone Network

标准卷积Conv可保留输入特征图各通道之间的隐藏链接,具有较强的特征提取能力,但其参数量和计算复杂度较高,深度可分离卷积DSConv在降低计算复杂度的同时会弱化主干网络对多光谱特征的提取和融合能力。因此,为了平衡模型的精度和计算复杂度,本文引入GSConv卷积,结构如图5所示^[18],首先,特征图经过轻量级标准卷积($1/2 C_2$ channel)捕捉多光谱目标局部特征,再通过DWConv卷积提取跨尺度特征,最后,通过Shuffle模块对拼接后的特征图进行混洗,建立光谱-空间特征的异构耦合关系,

在保留网络多光谱特征提取能力的同时降低模型的参数量。

GSConv与Conv之间的参数量和计算量比值如公式(1)和公式(2)所示^[18]:

$$ratio_p = \frac{D_K \times D_K \times \frac{N}{2} (M+1)}{D_K \times D_K \times M \times N} = \frac{M+1}{2M}, \quad (1)$$

$$ratio_c =$$

$$\frac{D_K \times D_K \times \frac{N}{2} (M+1) \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F} = \frac{M+1}{2M}, \quad (2)$$

其中: D_K 为卷积核大小; M 为输入通道数; N 为输出通道数; $D_F \times D_F$ 为输出特征图的大小。

如果将检测网络中的Conv全部替换为GSConv,网络层会变得更深,增加模型推理时间。在多光谱检测网络中,主干网络的特征提取模块HGBlock和颈部网络的特征增强模块C2f的参数量和计算量占比较高,因此,本文基于GSConv重构HGBlock和C2f模块,改进后的网络能够兼顾检测精度和速度,有效减少模型参数量。GS-HGBlock和GS-C2f结构如图6所示。

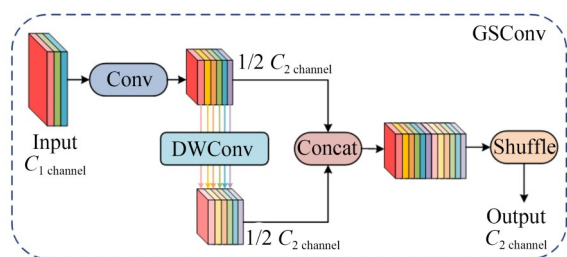


图5 GSConv结构图

Fig. 5 GSConv Structure Diagram

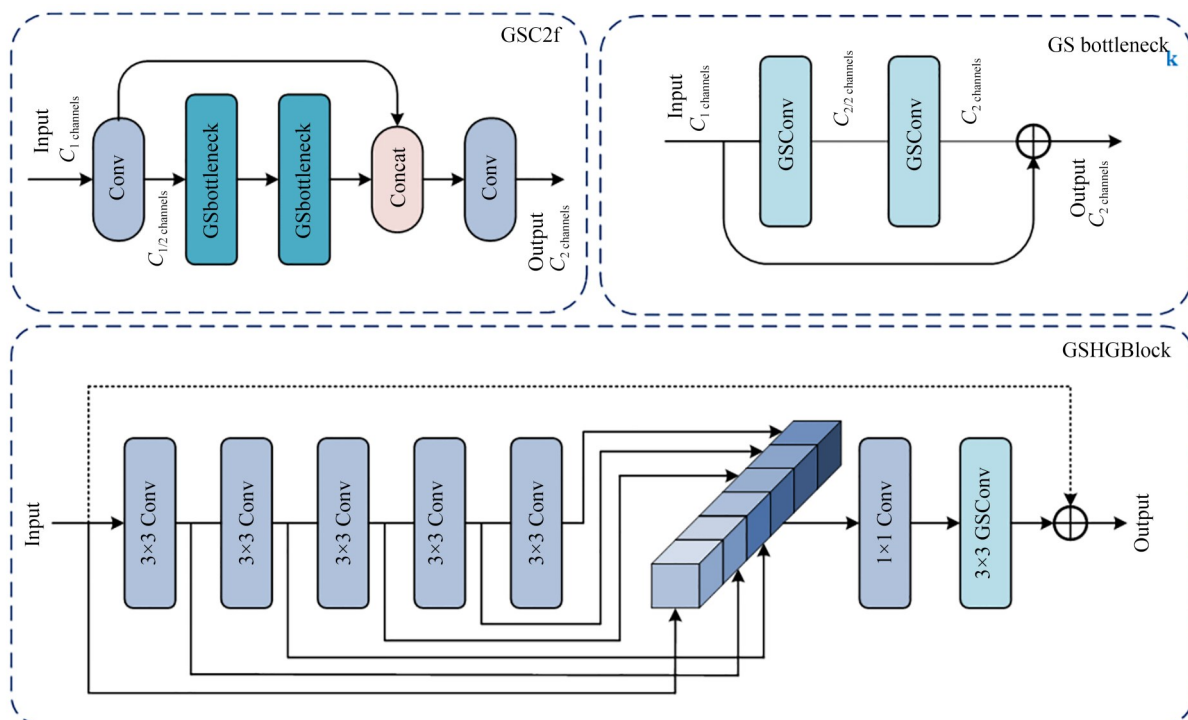


图 6 GS-HGBlock 和 GS-C2f 模块结构图

Fig. 6 GS-HGBlock and GS-C2f module structure diagram

3.2 La-Depgraph 通用结构化剪枝

在实际舰船目标识别任务中,多光谱检测模型含有大量冗余参数,难以在算力有限的嵌入式设备上部署,通过模型剪枝可去除网络中的冗余参数,提高多光谱检测模型的识别性能。

模型剪枝分为非结构化剪枝和结构化剪枝,非结构化剪枝算法以神经元为单位进行剪枝,破坏了模型的结构性,导致模型结构稀疏,现有的硬件架构也无法对其进行加速,实用性较低。结构化剪枝算法以通道为单位,根据稀疏化后的通道权重对层结构进行剪枝,能够去除冗余通道,保留完整的模型结构,可通过 GPU 或其他硬件进行加速。多光谱检测模型需融合多谱段特征以增强复杂环境下舰船特征的表达能,但常规的结构化剪枝算法对权重参数进行稀疏训练时,容易忽略不同光谱通道上目标层与其他层间的依赖关系,稀疏化过程中可能误删光谱敏感通道,降低模型的多光谱特征提取能力。

DepGraph 结构化剪枝算法^[19]可以将存在依赖关系的连接层分为一个剪枝组。根据依赖关系对组内通道权重进行一致的稀疏训练,保证多光谱特征表达的完整性。由于模型稀疏训练需

要花费大量的时间,本文在分组稀疏化的基础上引入非结构化剪枝算法 LAMP^[20]中的评分机制,提出了 La-Depgraph 通用结构化剪枝算法,该算法通过 LAMP 评分机制对组内通道权重进行一致稀疏化,如图 7 所示,跳过复杂耗时的模型稀疏训练过程,对模型的适配性更好。整体设计思路如下:

(1) 根据依赖关系对神经网络进行分组,采用一个权重正则项 R 表示组内每个通道权重的重要性,正则项 R 定义如式(3)所示^[19]:

$$R(k) = \sum_{k=1}^K \gamma_k \cdot I_k = \sum_{k=1}^K \sum_{w \in k} \gamma_k \|w[k]\|_F^2, \quad (3)$$

其中: $w[k]$ 表示组内第 k 层单个神经元的权重参数; $\|\cdot\|_F$ 为 Frobenius 范数; I_k 表示组内第 k 层的权重参数; γ_k 表示应用于第 k 层权重参数的缩放系数, γ_k 定义如下:

$$\gamma_k = 2^{\alpha(I_{\max} - I_k)/(I_{\max} - I_{\min})}, \quad (4)$$

其中: $\gamma_k \in [2^0, 2^\alpha]$, α 为一个超参数。

(2) 对于每个权重正则项 R , 根据给定通道索引 μ, ν 排序, $R[\mu], R[\nu]$ 表示通道索引 μ, ν 的正则项, 假设当 $\mu \leq \nu$ 时, $R[\mu] \leq R[\nu]$ 成立, LAMP 评分计算公式为:

$$sorce(\mu; R) = \frac{(R[\mu])^2}{\sum_{v \geq \mu} (R[v])^2}. \quad (5)$$

LAMP 评分可以反映一个剪枝组内每个通道的相对重要性,在设置好压缩比率后,La-Dep-graph 算法根据依赖关系对神经网络进行自动分组剪枝,组内 LAMP 评分较低通道权重会被剪除,直到满足所需的全局稀疏性要求。根据公式可知,每个组内都存在评分为 1 的通道,意味着每一组内至少会保留一个存活通道,有效避免神经网络崩溃。

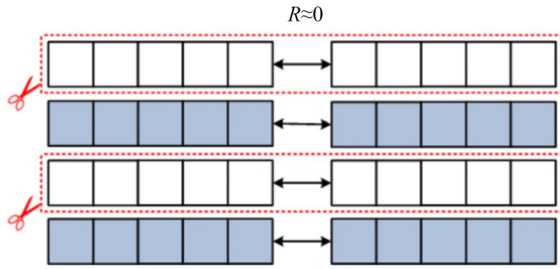


图 7 一致稀疏性

Fig. 7 Consistent sparsity in pruning groups

3.3 知识蒸馏

模型剪枝之后,特征通道的移除破坏了多光谱特征间的语义关联,导致模型识别精度降低,知识蒸馏可以从保留完整多光谱通道的教师模型中提取知识,将其传递给剪枝后的学生模型,提高学生模型的多光谱目标识别性能。本文引入特征知识蒸馏算法 CWD(Channel-Wise knowledge Dissillation)^[21],将剪枝前 s 版本和剪枝后 n 版本的模型分别设置为教师和学生模型。首先,选择蒸馏的特征层并在通道维度执行 softmax 标准化,将每个通道的特征映射转化为能反映相对重要性的概率映射,同时消除大模型和小模型之间的尺度差异。最后,将两个模型对应通道概率映射之间的 Kullback-Leibler(KL)散度作为损失函数,通过最小化损失函数使学生网络模仿教师网络。

通道蒸馏损失一般表示形式如下:

$$\varphi(\phi(y^T), \phi(y^S)) = \varphi(\phi(y_c^T), \phi(y_c^S)), \quad (6)$$

$$\phi(y_c) = \frac{\exp\left(\frac{y_{c,i}}{t}\right)}{\sum_{i=1}^{w \cdot H} \exp\left(\frac{y_{c,i}}{t}\right)}, \quad (7)$$

其中: c 为通道索引; i 为通道的空间位置; t 为蒸馏温度参数; y^T 和 y^S 分别为教师和学生网络激活图; $\phi(\bullet)$ 表示将激活值转化为概率分布。 t 值越大,概率分布更加平滑,每个通道的感受野越大。由于剪枝后小模型的通道数与大模型不匹配,本文使用 1×1 的卷积来对齐通道数。

KL 散度表达式如式(8):

$$\varphi(y^T, y^S) = \frac{t^2}{C} \sum_{c=1}^C \sum_{i=1}^{w \cdot H} \phi(y_c^T, i) \cdot \log \left[\frac{\phi(y_c^T, i)}{\phi(y_c^S, i)} \right]. \quad (8)$$

本文选择颈部网络的三个输出特征层作为蒸馏节点,通过最小化通道概率分布的 KL 散度,强制对齐多光谱细节特征,减小学生模型和教师模型之间的差异,使得学生模型的中间层特征表达与教师模型对应层特征表达相拟合,从而提高剪枝后模型的目标识别性能。

4 实验结果与分析

4.1 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} , $AP_{50 \sim 95}$, 参数量和计算量作为评价指标,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,参数量和计算量常用来评估模型的大小和计算复杂度。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 $AP_{50 \sim 95}$, 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 $mAP = AP$ 。

4.2 数据集与训练参数设置

本文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 1 143 张舰船图像,共 4 572 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,验证集 115 张图像,测试集 114 张图像。本文的仿真实验在 Windows 11 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

表 2 不同剪枝算法对比

Tab. 2 Comparison of different pruning algorithms

Model	Pruning	mAP ₅₀ /%	mAP _{50~95} /%	Param/M	Flops/G	MemFootprint /M
YOLOv5	L1-norm	80.1	50.9	1.47	1.3	3.3
	L2-norm	81.8	51.4	1.39	1.3	3.1
	Group-Slim	84.2	53.2	1.26	1.3	2.8
	DepGraph	86.1	55.3	0.39	1.2	0.9
	La-Depgraph	86.3	55.7	0.37	1.2	0.9
YOLOv8	L1-norm	80.3	51.2	1.25	1.7	3.1
	L2-norm	82.7	52.1	1.16	1.7	2.9
	Group-Slim	85.4	54.3	1.08	1.7	2.6
	DepGraph	88.2	58.3	0.52	1.6	1.2
	La-Depgraph	88.7	59.1	0.48	1.6	1.2
Ours	L1-norm	89.7	57.8	2.48	3.5	6.6
	L2-norm	91.1	59.2	2.35	3.5	6.1
	Group-Slim	92.2	61.3	2.23	3.6	5.6
	DepGraph	94.2	66.4	0.96	3.5	2.4
	La-Depgraph	94.9	67.4	0.90	3.5	2.3

多光谱检测模型,DepGraph剪枝算法能够有效降低模型参数量,但识别精度下降了1.1%。本文改进的La-Depgraph剪枝算法以0.4%的精度损失,消除了多光谱检测模型80.9%的冗余参数,降低了75.5%的计算量。

4.3.3 知识蒸馏实验分析

本文对比了知识蒸馏时不同比例的训练集和验证集对学生模型在测试集上检测性能的影响,结果如表3所示。当训练集和验证集比例为

9:1时,蒸馏后的学生模型在测试集上的泛化能力最优,模型检测精度为96.4%,相比于蒸馏前提高了1.5%。

4.3.4 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4所示,表中“√”表示添加了该模块,“-”表示未添加该模块。由表可以看出,轻量型网络结构可以在减少模型参数的同时提高了模型0.3%的识别精度。为了进一步压缩模型体积,利用结构化剪枝算法对模型冗余参数进行修剪,剪枝后的模型参数量仅为剪枝前的19.1%,但模型精度

表 3 不同训练集比例模型性能对比

Tab. 3 Comparison of different training set ratios

Ratio	Model	mAP ₅₀ /%	mAP _{50~95} /%
7:3	Teacher	95.6	70.1
	Student	93.8	66.2
	Distill-Student	95.1	69.6
8:2	Teacher	96.2	70.5
	Student	94.5	66.9
	Distill-Student	95.9	69.8
9:1	Teacher	96.6	71.0
	Student	94.9	67.4
	Distill-Student	96.4	70.6

表 4 消融实验

Tab. 4 Ablation experiments

LW-Net	Pruning	KD	mAP ₅₀ /%	mAP _{50~95} /%	Param /M	Flops /G
—	—	—	95.2	66.9	7.6	18.6
√	—	—	95.5	67.8	4.7	14.3
√	√	—	94.9	67.4	0.9	3.5
√	√	√	96.4	70.6	0.9	3.5

下降了 0.4%。最后,通过知识蒸馏使得剪枝后的学生模型学习教师模型的特征,模型精度提高了 1.5%。

4.3.5 与不同检测模型对比

为验证 YOLOv8n-MFLW 模型的有效性,本文将其与主流深度学习检测模型 SSD512,Faster RCNN 及 YOLO 系列进行对比实验,实验结果如表 5 所示。与改进前多光谱模型相比,YOLOv8n-MFLW 模型的 mAP_{50} 和 $mAP_{50\sim 95}$ 分别提

高 1.2% 和 3.7%,参数量、计算量和内存占用量分别减少 88.1%,81.2% 和 82.8%。与单谱段检测模型 YOLOv8n 相比, mAP_{50} 和 $mAP_{50\sim 95}$ 分别提高 8.0% 和 12.7%,参数量、计算量和内存占用量分别减少 70.0%,57.3%,61.6%。与典型双模态检测模型 YOLOv8-CFT 相比, mAP_{50} 和 $mAP_{50\sim 95}$ 分别提高 1.8% 和 2.1%,参数量、计算量和内存占用量分别减少 93.9%,91.1%,93.1%。

表 5 不同模型的性能对比
Tab. 5 Performance of different models

Spectral Band	Model	$mAP_{50}/\%$	$mAP_{50\sim 95}/\%$	Param/M	Flops/G	MemFootprint /M
B3	SSD512	84.3	53.8	24.0	87.8	92.1
B3	Faster RCNN	86.6	56.5	136.8	200.9	521.9
B3	YOLOv5n	86.9	56.8	1.9	4.5	5.0
B3	YOLOv7tiny	87.4	57.2	6.0	13.0	12.3
B3	YOLOv8n	88.4	58.3	3.0	8.2	6.0
B3	YOLOv8s	90.0	60.5	11.1	28.6	22.5
B3	YOLOv10n	87.9	57.6	2.7	8.4	5.7
B1& B3	YOLOv5-ACDF	91.5	63.4	8.0	12.7	16.1
B1& B3	YOLOv7-CMF	92.2	64.3	16.6	50.9	33.4
B1& B3	YOLOv8-CFT	91.7	63.8	11.5	28.8	23.1
B1& B3	Ours	93.5	65.9	0.7	2.5	1.6
Fusion	YOLOv8n-MF	95.2	66.9	7.6	18.6	13.4
Fusion	Ours	96.4	70.6	0.9	3.5	2.3

注:B3 is single-spectrum detection; B1 & B3 are two-spectrum detection; Fusion is four-spectrum detection.

4.3.6 特征图可视化

为了更直观地观察模型在舰船目标识别能力方面的提升,选取 A,B,C 三组海上红外舰船图像,A 组为强太阳光谱干扰下的舰船图像,B 组为岸岛干扰下的多目标舰船图像,C 组图像不仅有前两类干扰源,还受到海杂波干扰。本文使用梯度加权映射 Gard-CAM^[22]生成热力图,图中高亮区域表示对模型检测结果影响较大的特征,单谱段、多光谱和轻量化多光谱模型的热力图如图 10 所示。改进后的模型能更好地降低噪声的影响,增强对舰船目标特征的感知。

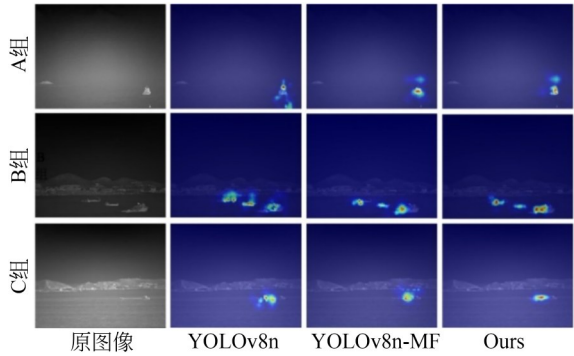


图 10 热力图对比
Fig. 10 Heatmap comparison

4.3.7 改进前后识别性能分析

单谱段、多光谱和改进后的轻量化多光谱模型在不同场景下的舰船目标检测结果如图 11 所示。由 A 组的对比实验可知,单谱段红外检测容易受到太阳光谱和海面噪声的影响,造成漏检和误检。由 B 组实验可知,原算法易将岛屿和渔船

识别为舰船,改进后的算法能在复杂背景下精准检测出舰船目标。由 C 组对比实验可知,改进后的模型在多种干扰下的检测有着更高的置信度。本文提出的轻量化多光谱检测模型有着更好的鲁棒性和泛化能力,能够适应复杂的海战检测任务。

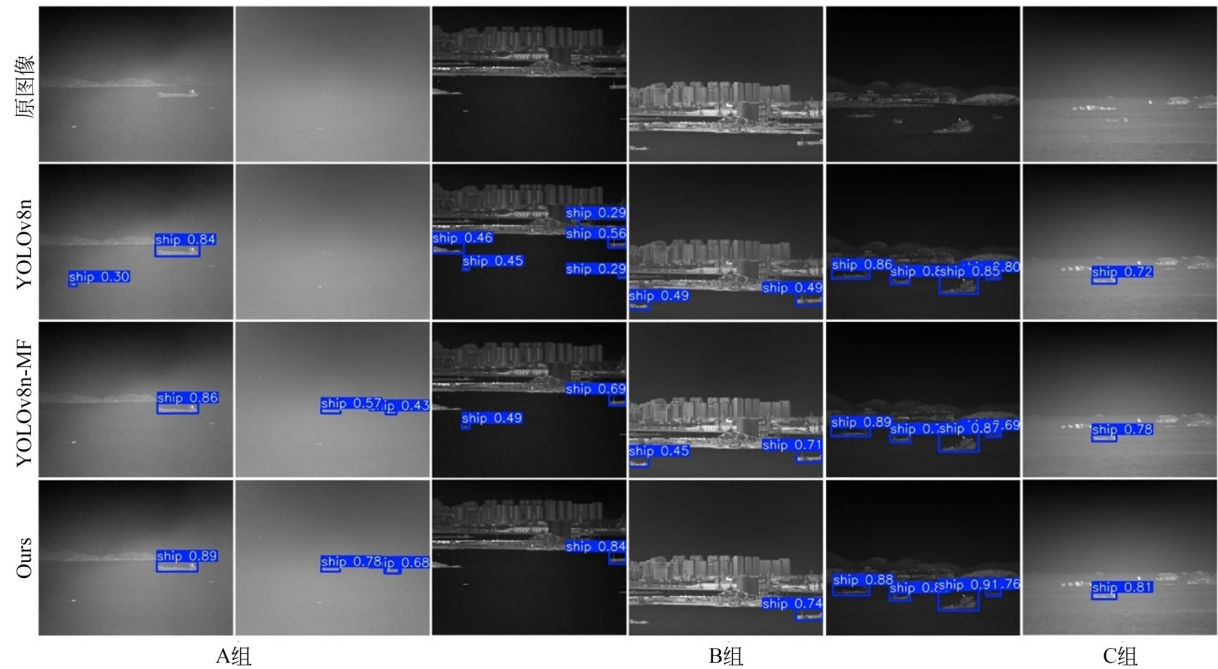


图 11 不同模型检测效果对比
Fig. 11 Comparison of different model detection effects

5 结 论

针对基于深度学习的多光谱融合检测模型体积大,难以部署的问题,本文在 YOLOv8n-MF 的基础上对其进行轻量化,提出了一种轻量化红外多光谱融合舰船目标检测模型 YOLOv8n-MFLW,引入轻量型网络结构对模型进行精度无损的初步压缩,然后,对模型进行结构化剪枝,以损失少量精度的代价,大幅减少模型参数量和计算量。最后,通过知识蒸馏提高剪枝后模型的检测精度。经过实验验证,相较于原模型,改进后的轻量化多光谱融合红外舰船目标检测算法的

检测精度可以达到 96.4%,提高了 1.2%,模型参数量、计算量及内存占用量仅为 0.9 MB, 3.5 GFlops 和 2.3 MB,分别减少 88.1%, 81.2% 和 82.8%。证明本文提出的模型体积小、精度高,具有良好的检测性能,能够胜任复杂环境下的舰船目标检测任务。

作者贡献声明:

赵佳乐:多光谱轻量化检测方法的提出,论文构思和撰写;
姜树理:研究指导和论文审核;
林超:实验的设计及数据整理和分析。

参考文献:

[1] 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.
QIU R C, LOU S L, LI T J, *et al.* Research on de-

- tection of ship target at sea based on multi-spectral infrared images[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(3): 698. (in Chinese)
- [2] ZHANG L, LIU Z Y, ZHANG S F, *et al.* Cross-modality interactive attention network for multispectral pedestrian detection [J]. *Information Fusion*, 2019, 50: 20-29.
- [3] ZHANG H, FROMONT E, LEFEVRE S, *et al.* Guided attentive feature fusion for multispectral pedestrian detection[C]. 2021 *IEEE Winter Conference on Applications of Computer Vision (WACV)*. January 3-8, 2021. Waikoloa, HI, USA. IEEE, 2021: 72-80.
- [4] ZHAO R, HAO J T, HUO H. Research on multi-modal pedestrian detection and tracking algorithm based on deep learning [J]. *Future Internet*, 2024, 16(6): 194.
- [5] DUNDAR A, JIN J, MARTINI B, *et al.* Embedded streaming deep neural networks accelerator with applications[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 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 Multibox Detector* [M]. Computer Vision - ECCV 2016. Cham: 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]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 779-788.
- [9] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 6517-6525.
- [10] GE Z, LIU S T, WANG F, *et al.* YOLOX: Exceeding YOLO Series in 2021 [EB/OL]. 2021: 2107.08430. <https://arxiv.org/abs/2107.08430v2>.
- [11] WANG C Y, BOCHKOVSKIY A, LIAO H M. 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)*. June 17-24, 2023, Vancouver, BC, Canada. IEEE, 2023: 7464-7475.
- [12] REIS D, KUPEC J, HONG J, *et al.* Real-time flying object detection with YOLOv8 [EB/OL]. 2023: 2305.09972. <https://arxiv.org/abs/2305.09972v2>.
- [13] FEI X, GUO M Y, LI Y. ACDF-YOLO: Attentive and cross-differential fusion network for multi-modal remote sensing object detection [J]. *Remote Sensing*, 2024, 16(18): 3532.
- [14] HOU Z Q, LI X Y, YANG C, *et al.* Dual-branch network object detection algorithm based on dual-modality fusion of visible and infrared images [J]. *Multimedia Systems*, 2024, 30(6): 333.
- [15] FANG Q Y, HAN D P, WANG Z K. Cross-modality fusion transformer for multispectral object detection [EB/OL]. 2021: 2111.00273. <https://arxiv.org/abs/2111.00273v4>.
- [16] 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.
- ZHAN W, ZHANG J, GUO M, *et al.* Multi-spectral infrared ship image multi-level fusion method [J]. *Laser & Infrared*, 2019, 49(2): 240-245. (in Chinese)
- [17] ZHAO Y A, LV W Y, XU S L, *et al.* DETRs Beat YOLOs on real-time object detection [C]. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 16-22, 2024, Seattle, WA, USA. IEEE, 2024: 16965-16974.
- [18] LI H L, LI J, WEI H B, *et al.* Slim-neck by GSConv: a lightweight-design for real-time detector architectures [J]. *Journal of Real-Time Image Processing*, 2024, 21(3): 62.
- [19] FANG G F, MA X Y, SONG M L, *et al.* DepGraph: towards any structural pruning [C]. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 17-24, 2023, Vancouver, BC, Canada. IEEE, 2023: 16091-16101.
- [20] LEE J, PARK S, MO S, *et al.* Layer-adaptive

- sparsity for the magnitude-based pruning [EB/OL]. 2020: 2010.07611. <https://arxiv.org/abs/2010.07611v2>.
- [21] SHU C Y, LIU Y F, GAO J F, *et al.* Channel-wise knowledge distillation for dense prediction [C]. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 10-17, 2021, *Montreal, QC, Canada*. IEEE, 2021: 5291-5300.
- [22] SELVARAJU R R, COGSWELL M, DAS A, *et al.* Grad-CAM: visual explanations from deep networks *via* gradient-based localization[J]. *International Journal of Computer Vision*, 2020, 128 (2): 336-359.

作者简介:



赵佳乐(2001—),男,山东济宁人,硕士研究生,现就读于烟台大学,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,山东临沂人,博士,副教授,硕士生导师,主要从事光电探测、光电目标识别、机器视觉等方面的研究。Email: shulilou@sina.com