

doi:10.3969/j.issn.1003-3106.2025.11.005

引用格式:彭鹏,陈慈发,张上.基于改进YOLO11n的船舶红外图像目标检测算法[J].无线电工程,2025,55(11):2174-2183.
[PENG Peng, CHEN Cifa, ZHANG Shang. Target Detection Algorithm for Ship Infrared Images Based on Improved YOLO11n[J]. Radio Engineering, 2025, 55(11): 2174-2183.]

基于改进YOLO11n的船舶红外图像目标检测算法

彭鹏^{1,2,3}, 陈慈发^{2,3,4*}, 张上^{1,2,3}

1. 三峡大学 水电工程智能视觉监测湖北省重点实验室, 湖北 宜昌 443002;
2. 三峡大学 湖北省建筑质量检测装备工程技术研究中心, 湖北 宜昌 443002;
3. 三峡大学 计算机与信息学院, 湖北 宜昌 443002;
4. 荆楚理工学院 大数据研究中心, 湖北 荆门 448001)

摘要:提出一种基于YOLO11n的船舶红外图像目标检测算法AGT-YOLO,旨在解决模型精度和召回率偏低、小目标识别困难以及复杂海况下的多尺度识别问题。通过引入改进后的GhostHGNetv2网络,增强背景区分能力;设计ASF-P2优化颈部网络,以提升对低分辨率图像和极小目标的检测能力;研究任务自适应对齐检测头(Task Adaptive Alignment Detection Head, TAADH)替换原有检测头,提升定位和分类性能;同时,融入AFGCAttention注意力机制,提高全局信息处理能力和模型的泛化能力。实验结果表明,与基准模型YOLO11n相比,AGT-YOLO的召回率提高了4.4%,平均精度均值@50(mean Average Precision at $IoU=0.5$, mAP@50)提高了3.1%,并在复杂环境下展现出较强的多尺度识别能力和鲁棒性。

关键词:目标检测;YOLO11n;红外船舶检测;多尺度模型

中图分类号:TN40;TP391.41

文献标志码:A

开放科学(资源服务)标识码(OSID):

文章编号:1003-3106(2025)11-2174-10



Target Detection Algorithm for Ship Infrared Images Based on Improved YOLO11n

PENG Peng^{1,2,3}, CHEN Cifa^{2,3,4*}, ZHANG Shang^{1,2,3}

1. Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, China Three Gorges University, Yichang 443002, China;
2. Hubei Province Engineering Technology Research Center for Construction Quality Testing Equipment, China Three Gorges University, Yichang 443002, China;
3. College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China;
4. Big Data Research Center, Jingchu University of Technology, Jingmen 448001, China)

Abstract: A ship infrared image target detection algorithm based on YOLO11n, named AGT-YOLO, is proposed to address the issues of low model accuracy and recall rate, difficulties in identifying small targets, and multi-scale recognition challenges under complex sea conditions. By introducing an improved GhostHGNetv2 network, the background discrimination capability is enhanced; the designed ASF-P2 optimized neck network improves detection capabilities for low-resolution images and very small targets; the proposed Task Adaptive Alignment Detection Head (TAADH) replaces the original detection head, enhancing localization and classification performance; meanwhile, the AFGCAttention mechanism is integrated to improve global information processing capability and the model's generalization ability. Experimental results show that compared to the baseline model YOLO11n, AGT-YOLO achieves a 4.4% increase in recall rate and a 3.1% increase in mean average precision at $IoU = 0.5$ (mAP@50), demonstrating strong multi-scale recognition capability and robustness in complex environments.

Keywords: object detection; YOLO11n; infrared ship inspection; multi-scale model

0 引言

随着我国经济高速增长与国际形势演变,海洋产业快速发展,海上作业规模持续扩大,推动舰船检测技术成为航海监管、渔业管理与国防安全的核心

收稿日期:2025-07-17

基金项目:湖北省大学生创新创业训练计划(202311075047);湖北省中央指导地方科技发展项目(2024BSB002)

Foundation Item: Hubei Province Innovation and Entrepreneurship Training Program for College Students (202311075047); Hubei Provincial Central Guidance Local Science and Technology Development Project (2024BSB002)

支撑技术^[1]。该技术不仅影响海上救援效率,更对民生服务与军事战略实施具有深远意义^[2]。

红外探测技术凭借远作用距离与强抗干扰性,在船舶检测中占据重要地位,但小目标识别仍是技术瓶颈。远距离船舶因弱特征特性易受海杂波及背景噪声干扰,尤其在复杂海天背景下,目标边缘模糊导致漏检误判风险加剧。此外,红外图像固有的低分辨率、高背景复杂度及多源噪声问题进一步制约检测精度提升^[3]。

针对上述问题,近年来众多学者开始探索深度学习在船舶红外图像目标检测中的应用。例如 Li 等^[4]以 YOLOv3 为基础,通过数据增强、锚框优化及浅层特征融合,提高微弱信号检出率;Ye 等^[5]则在 YOLOv5 中新增高分辨率特征层,并嵌入 TA(Target Attention)模块抑制复杂背景噪声,实现对小目标更精准定位。然而,Ye 等^[5]方法的注意力机制过度聚焦于小尺度区域,抑制大目标特征表达,导致大目标精度下降;且网络对场景光照、云层等环境变化敏感,小目标抗干扰能力不足。王海群等^[6]在 YOLOv8 算法中引进双向信息流和可学习权重加强网络以提高特征表达能力,通过双线性插值统一不同尺度的池化结果,并进行拼接和归一化的方法,增强了网络的多尺度特征提取能力,但仍然面临着抗噪声干扰能力不足的问题。

YOLO11n 是由 Ultralytics 团队于 2024 年 9 月 30 日发布的新一代目标检测和图像分割模型,相较于

前代 YOLOv8 模型引入 C3k2 模块和 C2PSA 注意力机制,动态调整特征提取,增强小目标检测能力。颈部部分采用了改进的路径聚合网络-特征金字塔网络++(Path Aggregation Network-Feature Pyramid Network++, PAN-FPN++)结构,增强了特征的表达能力和层次感。采用自适应损失函数(VFLoss+)和动态锚框机制分别提升了对目标方向和位置的识别精度。

本文基于 YOLO11n 基础模型进行改进,针对 Ye 等^[5]研究中出现多尺度目标识别的问题,提出 TA-ADH,通过任务对齐机制联合优化定位与分类分支,利用动态卷积核增强多尺度特征适配性。针对王海群等^[6]研究出现的复杂海况特征提取不足的问题,采用改进的 GhostHGNetv2^[7]网络,在降低参数数量的同时增强背景区分能力,同时融合 AFGCAttention^[8]模块提高全局特征聚合效率。对于船舶红外图像检测小目标问题,Li 等^[4]也做了针对性的研究,但对于小目标的检测精度仍有提升空间,本文设计 ASF-P2 结构^[9],结合自适应通道筛选策略,使小目标检测精度得以提升。因此,本文提出多尺度船舶红外检测算法 AGT-YOLO(ASF-P2,GhostHGNet,TAADH)。

1 改进的 YOLO11n 目标检测算法

红外船舶检测因图像分辨率低、海面杂波干扰及船舶尺度跨度大,原始 YOLO11 特征提取不足、误检高,难以直接应用,为此提出 AGT-YOLO。AGT-YOLO 的网络结构如图 1 所示。

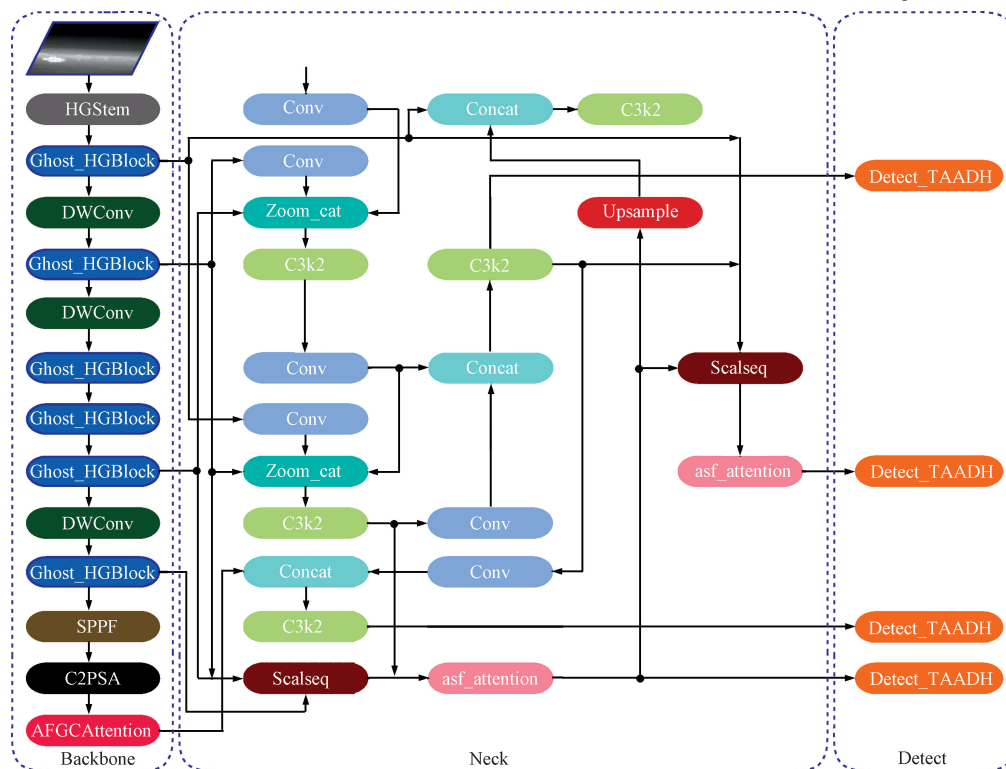


图 1 AGT-YOLO 网络结构

Fig. 1 Structure of AGT-YOLO network

首先,用改进的 GhostHGNetv2 网络代替原有特征提取网络,解决复杂背景场景下红外目标的检测,提高模型检测的性能,并降低其参数量。其次,引入 ASF-YOLO 的思想,添加 P2 特征层,用于解决船舶红外图像小目标因背景干扰以及分辨率限制导致识别困难的问题。然后,运用本文提出的检测头 TA-ADH,动态融合多尺度特征,强化对不同尺寸目标的感知。最后,采用 AFGCAttention 捕获全局上下文,提升整体模型一致性,使模型在新数据上保持鲁棒性与高泛化能力。

1.1 改进的特征提取网络 GhostHGNetv2

针对船舶红外图像中复杂背景干扰及目标特征弱化问题,本研究提出改进型 GhostHGNetv2 网络。

该网络融合 HGNetv2 架构与 Ghost 卷积技术,通过 Ghost_HGBlocks 模块优化特征提取过程,增强复杂背景下目标与背景的区别能力。

如图 2 所示,GhostHGNetv2 基于 HGNetv2 的四阶段层级结构构建,其核心单元为 HGBlock。网络输入端采用 HGStem 模块,通过深度可分离卷积(DW-Conv)实现可学习下采样层(LDS Layer)高效特征预处理。其中,LDS 层摒弃传统固定池化操作,采用动态权重机制自适应调整下采样策略,以保留更多有效特征信息^[10]。Stage 2~Stage 4 均集成 LDS 层,结合 Ghost_HGBlock 的双路径特征复用机制——主路径执行标准卷积,副路径生成轻量化特征图——在降低计算复杂度的同时提升特征表达能力。

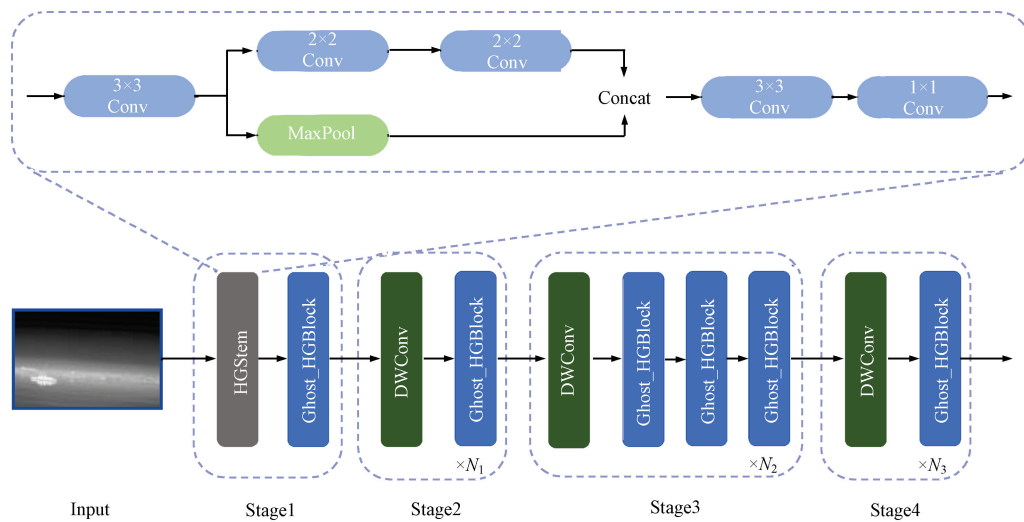


图 2 改进后的 GhostHGNetv2 网络结构

Fig. 2 Network structure of improved GhostHGNetv2

针对红外图像背景噪声敏感及目标尺度多样性问题,网络在 Stage 3 重点优化 P4/16 特征层处理流程。通过堆叠 Ghost_HGBlock 模块进行多层次特征融合,结合通道注意力机制动态筛选关键特征,有效抑制海面杂波干扰并增强小目标边缘信息。此外,全网络采用 DWConv 技术替代标准卷积,进一步压缩模型参数量,兼顾检测精度与实时性需求。该设计通过层级化特征提取与动态自适应机制,显著提升了复杂海况下多尺度船舶目标的检测鲁棒性。

在 GhostHGNetv2 网络中,HGStem 模块的结构见图 2。该过程由以下公式表示:

$$Y_1 = \text{MaxPool}_{2 \times 2}^l \lfloor F_{3 \times 3}^2(X) \rfloor, \quad (1)$$

$$Y_2 = F_{2 \times 2}^1 \{ F_{2 \times 2}^1 \lfloor F_{3 \times 3}^2(X) \rfloor \}, \quad (2)$$

$$Y = F_{1 \times 1}^1 \{ F_{3 \times 3}^2 \lfloor \text{Concat}(Y_1, Y_2) \rfloor \}, \quad (3)$$

式中: X 表示输入的三通道图像数据, Y 表示模块输出的特征图, $F_{3 \times 3}^2$ 、 $F_{2 \times 2}^1$ 、 $F_{1 \times 1}^1$ 分别表示卷积核大小为 3、2、1 的一般卷积操作,卷积步长分别为 2、1、1; $\text{MaxPool}_{2 \times 2}^l$ 表示窗口大小为 2×2,步幅为 1 的最大池化操作。

此外,另一个网络 GhostNet 在特征提取上表现优异,主要利用主、副通道。主通道捕捉主要图像信息,而副通道针对细节,卷积层较少,作为正则化机制防过拟合。副通道采用深度可分离卷积降低模型复杂度。最终,两通道输出结合形成特征表示。

Ghost 卷积过程可以表示如下:

$$Y_1 = F_{1 \times 1}^1(X), \quad (4)$$

$$Y = \text{Concat}[Y_1, D_{5 \times 5}^1(Y_1)], \quad (5)$$

式中: $D_{5 \times 5}^1$ 表示一个深度卷积操作,其内核大小为

5×5,步幅为 1,组数等于通道数。

Ghost_HGBlock 融合 GhostNet 和 Ghost 卷积的思想,将 Ghost 卷积模块整合到 HGNetv2 骨干网的 HGBlock 中。采用 Ghost 卷积模块替换 HGBlock Light 卷积模块,以实现更好的特征提取,同时也使模型更轻、参数更少、计算需求更少。GhostHG-Block 模块的原理及其流程如图 3 所示。计算过

程如下:

$$\begin{cases} Y_1 = G(X), \\ Y_2 = G(Y_1), \\ \vdots \\ Y_n = G(Y_{n-1}), \end{cases} \quad (6)$$

$$Y = F_{1 \times 1}^1 \{ F_{1 \times 1}^1 [\text{Concat}(Y_1, Y_2, \dots, Y_n)] \}, \quad (7)$$

式中:G 表示前文提到的 Ghost 卷积操作。

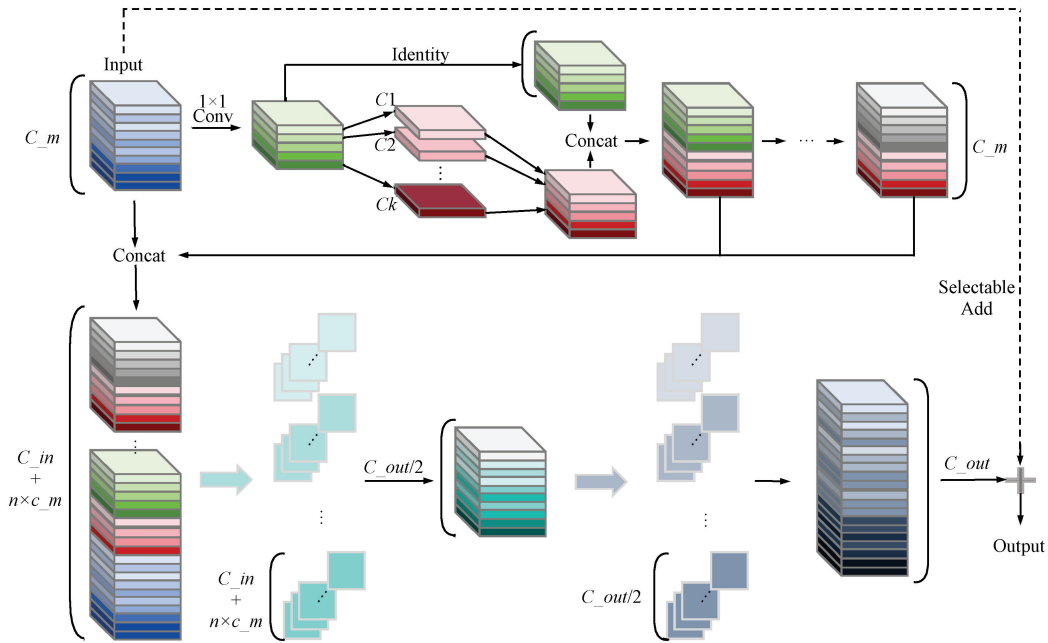


图 3 Ghost_Block 模块结构

Fig. 3 Structure of Ghost_Block module

1.2 改进的颈部结构 ASF-P2

ASF-YOLO 是 Kang 等^[9]提出的 YOLO 框架,融合了尺度和空间特征,适用于快速精确的分割任务。其颈部设计包括尺度序列特征融合 (Scale Sequence Feature Fusion, SSFF) 模块、时间融合编码器 (Temporal Fusion Encoder, TFE) 模块和通道与位置注意机制 (Channel and Position Attention Mechanism, CPAM) 模块。该设计使模型能自适应关注不同尺度小目标的相关通道和空间位置。本文运用卷积和 Zoom_cat 实现 TFE 功能,设计 Scalseq 融合不同尺度特征,避免冗余计算,并引入 asf_attention 提取代表性特征信息。

1.2.1 尺度序列特征融合模块 Scalseq

船舶目标尺度跨度大,传统 FPN 虽融合多尺度特征,但仅靠简单拼接/求和,忽视层间深层关联,导致小目标漏检、大目标误检频发,检测精度与鲁棒性不足,亟需强化特征交互。通过设计 Scalseq 模块,

将 P2~P5 同宽高比特征按尺度顺序排列,串联深层语义与浅层细节;随后用标准差递增的高斯核依次卷积,递进融合跨层信息^[11-12],如下所示:

$$F_{\sigma}(i, j) = \sum_u \sum_v f(i-u, i-v) \times G_{\sigma}(u, v), \quad (8)$$

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \quad (9)$$

F_{σ} 利用递增标准差 σ 的二维高斯卷积平滑特征图 f 生成多尺度表示,将其水平堆叠后,通过三维卷积 (3D-CNN) 从尺度序列中提取特征 (如图 4 所示)。基于能提供全面准确信息的 P3 级特征图设计 Scalseq 模块:首先调整最高级特征图的通道数和空间维度与 P3 匹配,然后扩展张量形状并沿深度维度连接,最后用 3D 卷积、批处理归一化和 ReLU^[13] 完成特征提取。拼接的特征图构成尺度序列。为增强小目标识别能力,额外添加 P2 级 Scalseq,提升模型多尺度特征提取能力。

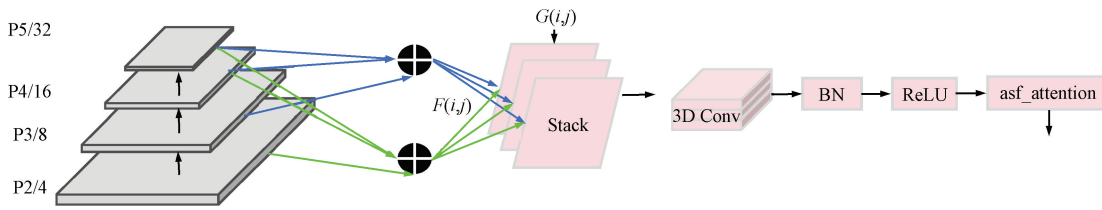


图4 Scalseq 模块结构
Fig. 4 Structure of Scalseq module

1.2.2 通道与位置注意力机制 asf_attention

asf_attention 机制专为小目标检测设计,融合了通道注意力和位置注意力网络。通道注意力网络接收 PANet 后的特征图作为输入 1,这些特征图包含卷积提取的详细信息。不同于 SENet^[14]中的全局平均池化和降维全连接层,提出的新机制采用一维

卷积核,不降维地捕捉每个通道及其 k 个邻近通道的局部交互,提高通道注意力的预测效果。位置注意力网络接收通道注意力网络与 Scalseq 输出的叠加作为输入 2,首先按宽度和高度分割输入特征图,在 pw 和 ph 轴上进行特征编码,然后合并生成输出关键位置信息,模块的整体结构如图 5 所示。

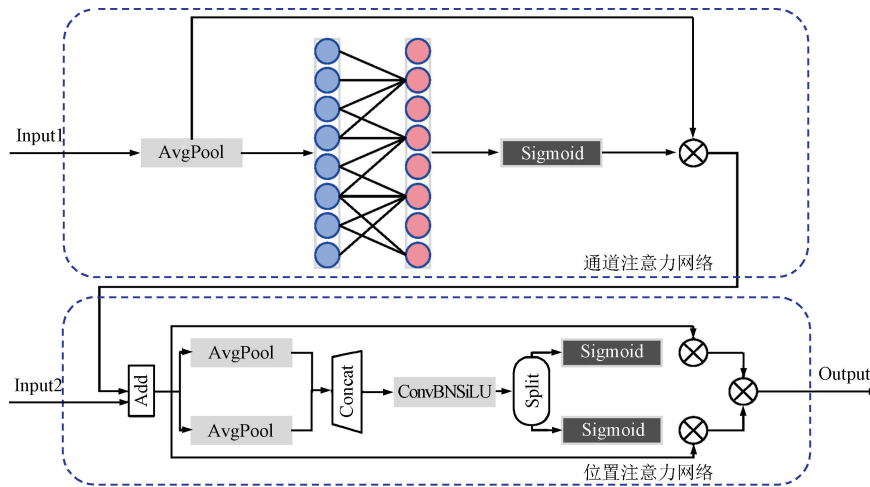


图5 asf_attention 模块结构
Fig. 5 Structure of asf_attention module

1.3 TAADH

现有 YOLO 系列算法的解耦头采用分类与定位分支独立设计,虽简化了任务学习流程,却导致两大核心问题:其一,分类任务依赖语义特征,而定位任务需精准空间信息,二者特征共享易引发冲突,尤其在复杂海况下加剧误检与漏检;其二,固定尺度感受野难以适配船舶红外图像中的目标尺度剧烈变化特性(如近岸大船与远距离小艇并存场景),且多分支结构增加了模型存储与计算负担。这些局限性在红外船舶检测任务中尤为突出,需通过特征交互机制与归一化策略优化实现突破。

针对上述问题对检测头进行改进优化,结合任务对齐的单阶段目标检测器(Task-aligned One-stage Object Detection, TOOD)算法^[15]的思想,设计

TAADH 模块,并提出一个改进型组归一化(Group Normalization, GN-Conv)卷积模块,GN-Conv 以组归一化替代批量归一化层(Batch Normalization, BN)^[16]。GN^[17]创新性地通过通道分组统计量计算,有效缓解对批量规模约束,在动态计算图框架及极小批量(如抛处理规模(Batch Size) ≤ 4)场景下仍保持稳定性能。实验表明,相较于 BN,GN 在模型精度、跨任务泛化能力等维度均展现出显著优势。

TAADH 架构采用一种特征提取器,旨在集成来自多个卷积层的特征信息,从而学习并融合出表征任务交互关系的联合特征。在定位分支中,该联合特征被输入到 DCNv2(可变形卷积网络 V2)模块中,直接用于生成其执行特征对齐所需的偏移量和

掩码参数。另一方面,分类分支则利用联合特征驱动一个动态特征选择机制,该机制由 1×1 卷积层和

3×3 卷积层组合构成(其具体结构如图 6 所示),用以筛选与分类任务最相关的特征。

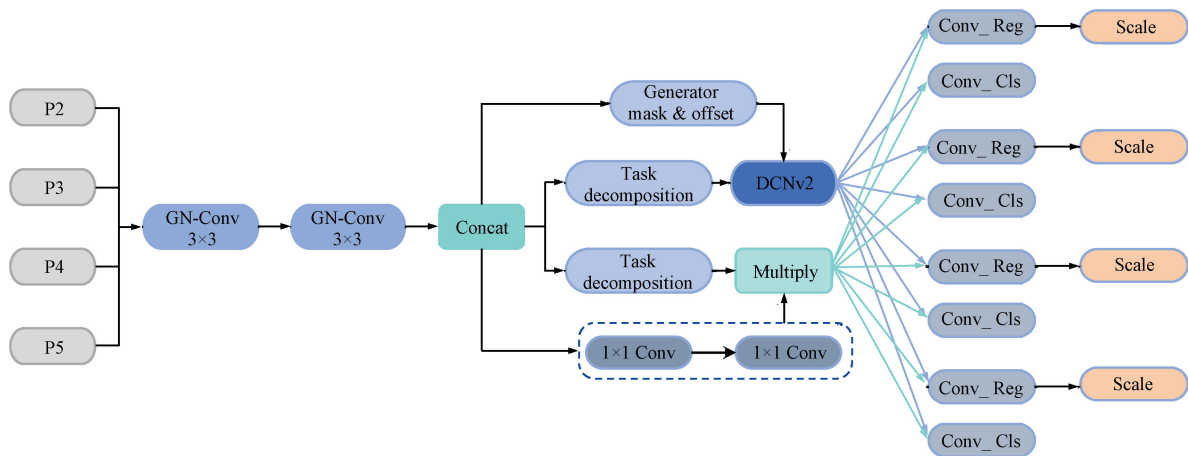


图 6 TAADH 模块结构

Fig. 6 Structure of TAADH module

鉴于定位与分类任务的目标差异,二者共享同一组特征时易引发特征干扰,各自所需的特征关注点也存在显著区别。为此,本文设计了一种任务解耦模块(结构如图 7 所示),基于跨层级任务交互特征动态推导特征权重系数,以促进定位与分类任务的特征解耦。其中,权重系数 w 由层级间交互特征计算生成,用于建模层级之间的关联特性。

softmax 注意力和线性注意力相结合,使其在计算复杂度上具有线性特性,降低模型的计算成本。通过代理向量 A 从 K 和 V 中聚合信息,然后将信息广播回 Q ,这种方式使得模型能够有效地处理全局信息,从而提升模型的整体性能。

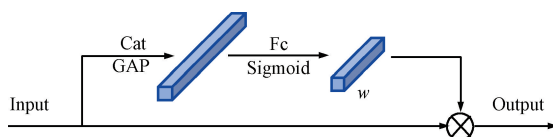


图 7 任务分解模块结构

Fig. 7 Structure of task decomposition module

文献[18]研究证实,GN 可显著提升检测头在目标定位与分类任务中的性能表现。为进一步优化模型效率,本研究引入卷积权重共享机制,通过复用跨层卷积核减少冗余参数。针对多尺度目标检测的固有挑战,设计特征尺度自适应调节层(Scale 层),通过可学习的缩放因子对各层级特征进行动态校准,以适配不同检测头对目标尺度的差异化需求。

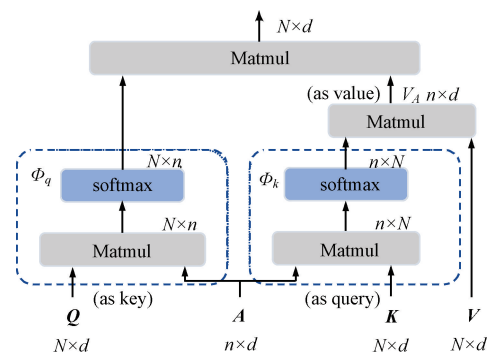


图 8 AFGC 模块结构

Fig. 8 Structure of AFGC module

借鉴 TOOD 思想,通过在检测头上实现定制任务对齐结构,避免传统独立分类和定位分支间的交互缺失。TAADH 提取任务交互联合特征,DCNv2 和交互特征生成 offset/mask 定位,动态选择分类,动态优化提升检测。

该模块继承了 softmax 和线性注意力的优点。为了更好地利用位置信息,以最大限度地发挥其潜力,模块中有一个经过精心设计的代理偏差,以提高代理的注意力。具体地说,在注意力计算中引入代理偏差,是受相对位置偏置(Relative Position Bias, RPB)^[19]的启发,即:

$$O^i = \sigma(QA^T + B_2) \sigma(AK^T + B_1) V, \quad (10)$$

式中: $B_1 \in \mathbb{R}^{n \times n}$, $B_2 \in \mathbb{R}^{N \times n}$ 表示原始代理偏差。在代理模型设计中,为提高参数利用效率,每个代理偏差均由 3 个正交的偏差分量(如系统偏差、随机偏差和模型结构偏差)复合生成。该方法避免直接将原

1.4 AFGCAttention 注意力机制

AFGCAttention 模块结构如图 8 所示,将传统的

始偏差项设为可训练变量,转而通过分量解耦与重组优化参数学习过程。由于线性注意力受到特征多样性不足的影响,该模块遵循并采用深度可分离卷积(Depthwise Convolution, DWC)模块来保持特征多样性。

因此,该模块由三部分组成,即纯代理注意、代理偏差和 DWC 模块。本文的模块可以公式化为:

$$O^A = \sigma(QA^T + B_2)\sigma(AK^T + B_1)V + DWC(V), \quad (11)$$

式中: $Q, K, V \in \mathbb{R}^{N \times C}$, $A \in \mathbb{R}^{N \times C}$, $B_1 \in \mathbb{R}^{n \times N}$, $B_2 \in \mathbb{R}^{N \times n}$ 。在默认设置中,代理 A 通过池化获得,即 $A = \text{合并}(Q)$ 。整体模块复杂度表示为:

$$Q = 4NC^2 + NC + 2nNC + 2NnC + k^2NC, \quad (12)$$

式中: N, n 为输入特征和代理的数量, $k = 3$ 为 DWC 的内核大小, $4NC^2$ 为预处理阶段, NC 为获取代理阶段, $2nNC + 2NnC$ 为运用代理注意力阶段, k^2NC 为 DWC 特征提取阶段。值得注意的是,该模型仅表现出 N 的线性复杂度。

该模型能够在保持较低计算复杂度的同时,提供较大感受野,从而提升模型的性能。结合代理注意力机制的这些优点,将其应用于船舶红外图像目标检测模型中,进一步提高模型对小目标和低分辨率图像的检测精度。由于代理注意力能够提供更大的感受野,可以帮助模型更好地理解目标与其上下文环境的关系,从而在复杂的海天背景下,更准确地检测和识别船舶目标,提升模型的鲁棒性。

2 实验与结果分析

2.1 实验环境与数据集

实验在 Windows 11 操作系统, GPU 为 NVIDIA GeForce RTX 4070Ti Super 显存 16 GB, PyTorch 1.2.0+CUDA 11.3 深度学习框架以及 Python 3.10.5 中实现,实验采用单卡 GPU 训练架构,批处理规模(batch size)设定为 16,输入数据尺寸规范化为 640 pixel × 640 pixel^[20]。模型共执行 300 次完整训练周期(epoch),参数优化基于随机梯度下降(Stochastic Gradient Descent, SGD)算法实现,初始学习率、动量(momentum)、权重衰减等参数均使用 YOLO11n 模型中的原始参数。

本研究构建于 infiRay 发布的海事船舶红外实测数据集,该数据集覆盖海上航道、近岸港口及沿海区域的多场景、多时段、多分辨率船舶红外图像。数据集分为 7 种船只类别,分别是邮轮、散货船、军舰、帆船、皮划艇、集装箱船、渔船,共计 8 326 张。将数据集图片按照 7 : 3 的比例随机划分成训练集和验证集进行实验。

为验证 AGT-YOLO 算法在跨场景下的泛化能力,本研究引入天津大学机器学习与数据挖掘团队构建的 VisDrone2019 基准数据集。该数据集涵盖多尺度目标分布特性(含 8 599 张航拍图像),重点聚焦小目标检测难题。实验采用分层随机抽样策略(Stratified Random Sampling),将原始数据集依据样本类别分布划分为互斥子集,其中训练集占比 70%,验证集占比 30%,以消除数据分布偏差并验证模型在未知场景下的迁移学习效能。

2.2 评价指标

实验选取目标识别任务中常见的评价指标:准确率(P)、召回率(R)、平均精度均值(mAP@0.5, mAP@0.5 : 0.95)、参数量(Para)和每秒帧数(Frames Per Second, FPS)。各项指标介绍如下。

(1) P 表示检测目标准确率,量化模型排除负样本干扰的能力,计算公式为:

$$P = \frac{TP}{TP + FP}, \quad (13)$$

式中: TP 为正确识别的正样本数量, FP 为将负样本误判为正样本的错误率。

R 表示检测目标召回率,衡量模型捕捉正样本的灵敏度,计算依赖漏检统计量,计算公式为:

$$R = \frac{TP}{TP + FN}, \quad (14)$$

式中: FN 表示本应被检出但被模型忽略的正样本。

(2) mAP 表示所有类别检出正确率的均值, mAP@0.5 表示交并比(Intersection over Union, IoU)为 0.5 时各类别平均精度(Average Precision, AP)值的算术平均,侧重宽松定位场景的评估; mAP@0.5 : 0.95 通过多阈值积分(步长 0.05),考核模型在高精度定位要求($IoU \geq 0.5$)下的综合性能,避免单一阈值评价的偏差,计算公式为:

$$mAP = \frac{1}{c} \times \sum_{i=1}^c AP_i, \quad (15)$$

式中: AP_i 表示第 i 个类别的 AP, c 表示数据集中的标签类别数。

(3) Para 表示衡量模型的规模和空间复杂度的指标。

(4) FPS 反映模型训练速度的刷新率,计算公式为:

$$FRS = \frac{1\ 000}{preprocess + inference + postprocess}, \quad (16)$$

式中: $preprocess$ 表示预处理时间, $inference$ 表示推理时间, $postprocess$ 表示后处理时间。

2.3 消融实验

为验证算法 AGT-YOLO 各个模块的有效性,在 infiRay 的船舶红外图像数据集上进行消融实验,其

中“√”表示使用该方法。以原始 YOLO11n 算法为基础模型,依次引入 GhostHGNetv2、TAADH、ASF-P2、AFGCAttention 模块。实验结果如表 1 所示。

表 1 消融实验结果

Tab. 1 Results of ablation experiments

单位:%

Alias	Improvement strategy				P	R	$mAP@0.5$	$mAP@0.5:0.95$
	GhostHGNetv2	TAADH	ASF-P2	AFGCAttention				
A	×	×	×	×	90.8	84.9	90.3	63.4
B	√	×	×	×	92.3	85.2	91.0	64.0
C	√	√	×	×	91.7	86.5	91.6	64.6
D	√	√	√	×	91.0	89.0	93.2	64.6
E	√	√	√	√	91.9	89.3	93.4	64.7

注:“√”表示应用了该模型,“×”表示未应用模型。

从表 1 可以看出,使用改进后的 GhostHGNetv2 模块替换 YOLO11n 的主干特征提取网络,准确率上升了 1.5 个百分点,提高了模型总体识别目标的准确性,面对多种类别目标的样本,检测精度都有所提高,验证了引入该模块能够提高区分背景和待检测目标差异的能力。在加入 GhostHGNetv2 方法的基础上对检测头进行定制化任务对齐设计后, $mAP@0.5:0.95$ 增加了 1.2 个百分点,验证了 TAADH 在锚框选择上精确度更高。通过引入 ASF-P2 特征融合机制优化 YOLO11 的颈部网络,实验表明该结构显著提升了多尺度特征融合效率,最终使模型在 IoU 阈值为 0.5 时的 mAP 提升至 93.2%。相较于基线模型的对比方法, mAP 指标获得 1.6 个百分点的增益,充分验证

了 ASF-P2 在增强目标检测精度方面的有效性。引入 AFGCAttention 注意力模块后,准确率、召回率和 mAP 都有小幅度的提升,验证了模型能够有效处理全局信息,以提升整体性能。

2.4 对比实验

为评估 AGT-YOLO 算法在目标检测任务上的性能,开展了一系列对比实验,选取 9 种公认目标检测架构:SSD、CenterNet、YOLOv5n、YOLOv6、YOLOv7-tiny、YOLOv8n、YOLOv9t、YOLOv10n、YOLO11n,这几种经典目标检测算法模型作为比较基准。通过这些对比实验,能够更全面地评估模型的优势和效果。基于多模型对照实验的定量分析结果(如表 2 所示),系统量化了本研究所提架构在检测精度、推理效率及泛化能力维度的综合性能优。

表 2 AGT-YOLO 与已有模型实验对比

Tab. 2 Experimental comparison between AGT-YOLO and existing models

模型	$P/\%$	$R/\%$	$Para/10^6$	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	FPS
SSD	82.0	80.5		86.8	42.8	33
CenterNet	88.1	91.2		91.7	58.8	39
YOLOv5n	90.6	84.7	2.50	90.3	62.8	135
YOLOv6	89.3	83.1	4.23	88.6	62.3	141
YOLOv7-tiny	89.2	87.0	6.03	91.0	60.4	110
YOLOv8n	90.4	85.4	3.01	90.8	63.7	135
YOLOv9t	90.8	85.5	1.97	90.1	63.7	71
YOLOv10n	88.8	83.6	2.70	89.8	62.6	112
YOLO11n	90.8	84.9	2.58	90.3	63.4	136
AGT-YOLO	91.9	89.3	2.51	93.4	64.7	121

由表 2 可知,改进算法 AGT-YOLO 较基线模型 YOLO11n, $mAP@0.5$ 提升 3.1% (90.3% → 93.4%), $mAP@0.5:0.95$ 提升 1.3% (63.4% → 64.7%),印证了多尺度特征融合机制对定位精度

的增强效果。在 FPS 指标上,也能够实现更高的帧率,从而提供更流畅的实时检测体验,满足海上救援等要求快速响应的需求。在参数量上,AGT-YOLO 算法也未劣化。总体而言,AGT-YOLO 能够

解决船舶红外目标检测方法存在的模型精度和召回率较低、小目标因背景干扰以及受分辨率限制导致识别困难、复杂海况下特征提取困难等问题。

2.5 泛化实验

对 YOLO11n 与算法 AGT-YOLO 在公共数据集 VisDrone2019 上做泛化实验,结果如表 3 所示。AGT-YOLO 与 YOLO11n 相比,mAP@0.5 从 34.7%增加到 40.5%,提升了 5.8 个百分点;mAP@0.5:0.95 从 19.9%提升为 23.8%,提高了 3.9 个百分点;P 从 45.6%提升到 51.0%,提高了 5.4 个百分点;R 从 35.6%提升到 39.7%,提高了 4.1 个百分点。以上数据表明,AGT-YOLO 算法不仅能提高复杂场景中不同分辨率下多尺度目标检测率,还能提高日常情况小目标的识别率,说明算法具有较强的通用性。

表 3 泛化实验

Tab. 3 Generalization experiment 单位:%

模型	mAP@0.5	mAP@0.5:0.95	P	R
YOLO11n	34.7	19.9	45.6	35.6
AGT-YOLO	40.5	23.8	51.0	39.7

2.6 可视化分析

采用 GradCAMPlusPlus^[21] 技术生成目标检测显著性图。算法以高阶导数重标梯度贡献,将模型聚焦区域映射为色彩梯度:深红对应核心判别特征,橙黄次之,冷蓝则指示对结果影响微弱的冗余信息,直观呈现网络关注焦点。图 9 展示了在 infiRay 船舶红外图像数据集上,YOLO11n 算法和 AGT-YOLO 算法在多尺度(图 9(a))、遮掩(图 9(b))、海天线(图 9(c))、小目标(图 9(d))等场景下的检测效果热力图,该结果加了边界框和置信度,便于分析^[22]。

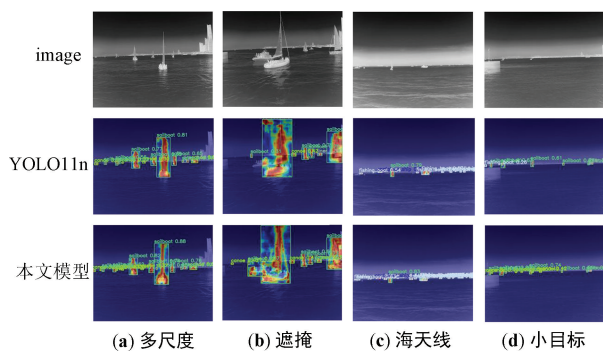


图 9 船舶检测热力图

Fig. 9 Heatmap of ship inspection

在多尺度目标检测中,基线算法并未完全检测出小尺寸目标,而 AGT-YOLO 算法能够比较准确地检测出各个尺度的目标。在出现遮掩场景的情况

下,基线算法并未检测出靠近被遮掩的目标,而 AGT-YOLO 算法能准确地识别出目标的重叠部分。在海天线干扰的背景下,基线算法对小目标的检测出现了漏检的现象,而 AGT-YOLO 算法能够准确地识别出这些小目标。在针对小目标检测下,基线算法出现了大量漏检还有极个别错检现象,而 AGT-YOLO 算法出现漏检的概率极低,没有错检现象的出现。对比热力图分析可知,AGT-YOLO 算法检测效果更好,能够适应复杂场景下多尺度、多目标的船舶红外图像检测。

3 结束语

针对船舶红外图像目标检测精度低、召回率低、小目标识别难及多尺度识别问题,提出的基于 YOLO11n 的 AGT-YOLO 改进算法,采用增强 GhostHGNetv2 网络作为主干特征提取网络,提升背景区分能力。同时,设计 ASF-P2 优化颈部网络,自研 TAADH,并融入 AFGCAttention 机制,以提高特征提取、小目标检测、定位和分类性能,以及多尺度目标识别和模型泛化能力。实验结果显示,与 YOLO11n 相比,AGT-YOLO 的召回率提高了 4.4%,平均精度 mAP@50 提高了 3.1%,且在泛化实验中,mAP@0.5、mAP@0.5:0.95 分别提升了 5.8、3.9 个百分点。总之,AGT-YOLO 算法在船舶红外目标检测上表现优异,证明了本文算法的可行性。



参考文献

- [1] 王斌. 海面目标红外检测方法研究[D]. 大连:大连海事大学,2018.
- [2] 古静. 基于深度学习的红外船舶目标检测方法研究[D]. 镇江:江苏科技大学,2023.
- [3] 王勇,王柏容,慕东东. 红外船舶检测的研究现状及展望[J]. 大连海事大学学报,2023,49(4):103-115.
- [4] LI H, DENG L B, YANG C, et al. Enhanced YOLO v3 Tiny Network for Real-time Ship Detection from Visual Image[J]. IEEE Access, 2021, 9: 16692-16706.
- [5] 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.
- [6] 王海群,魏培旭,解浩龙,等. 基于改进 YOLOv8 的红外船舶检测[J]. 电光与控制,2025,32(1):61-67.
- [7] DING Q X, LI W C, XU C C, et al. GMS-YOLO: An Algorithm for Multi-scale Object Detection in Complex Environments in Confined Compartments[J]. Sensors, 2024, 24(17): 5789.

- [8] HAN D, YE T, HAN Y, et al. Agent Attention: On the Integration of Softmax and Linear Attention [EB/OL]. (2023-12-14) [2024-11-26]. <https://arxiv.org/abs/2312.08874>.
- [9] KANG M, TING C M, TING F F, et al. ASF-YOLO: A Novel YOLO Model with Attentional Scale Sequence Fusion for Cell Instance Segmentation [EB/OL]. (2023-12-11) [2024-11-26]. <https://arxiv.org/abs/2312.06458>.
- [10] YANG X, ZHANG Y F, LI T, et al. Image Super-resolution Based on the Down-sampling Iterative Module and Deep CNN [J]. *Circuits, Systems, and Signal Processing*, 2021, 40: 3437-3455.
- [11] PARK H J, KANG J W, KIM B G. ssFPN: Scale Sequence (S^2) Feature-based Feature Pyramid Network for Object Detection [J]. *Sensors*, 2023, 23(9): 4432.
- [12] TRAN D, BOURDEV L, FERGUS R, et al. Learning Spatiotemporal Features with 3D Convolutional Networks [C] // 2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015: 4489-4497.
- [13] ELFWING S, UCHIBE E, DOYA K. Sigmoid-weighted Linear Units for Neural Network Function Approximation in Reinforcement Learning [J]. *Neural Networks*, 2018, 107: 3-11.
- [14] JIN X, XIE Y P, WEI X S, et al. Delving Deep into Spatial Pooling for Squeeze-and-Excitation Networks [J]. *Pattern Recognition*, 2022, 121: 108159.
- [15] FENG C J, ZHONG Y J, GAO Y, et al. Tood: Task-aligned One-stage Object Detection [C] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 3490-3499.
- [16] IOFFES, SZEGEDY C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift [EB/OL]. (2015-02-11) [2024-11-26]. <https://arxiv.org/abs/1502.03167>.
- [17] WU Y X, HE K M. Group Normalization [C] // Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer, 2018: 3-19.
- [18] TIAN Z, SHEN C H, CHEN H, et al. FCOS: A Simple and Strong Anchor-free Object Detector [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 44(4): 1922-1933.
- [19] SHAW P, USZKOREIT J, VASWANI A. Self-attention with Relative Position Representations [EB/OL]. (2018-03-06) [2024-11-26]. <https://arxiv.org/abs/1803.02155>.
- [20] 杨慧聪, 陈慈发, 张上. 基于通道剪枝的 YOLOv8n 印刷电路板缺陷检测 [J]. *无线电工程*, 2025, 55(2): 271-280.
- [21] CHATTOPADHAY A, SARKAR A, HOWLADER P, et al. Grad-CAM++: Generalized Gradient-based Visual Explanations for Deep Convolutional Networks [C] // 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Tahoe: IEEE, 2018: 839-847.
- [22] 贾星宇, 李大鹏. 基于 YOLO11 的无人机航拍图像小目标检测算法 [J]. *无线电工程*, 2025, 55(8): 1560-1570.

作者简介

彭 鹏 男, (2001—), 硕士研究生。主要研究方向: 图像处理、计算机视觉。

(*通信作者) 陈慈发 男, (1967—), 教授。主要研究方向: 嵌入式系统、物联网、计算机测控系统。

张 上 男, (1979—), 博士, 副教授。主要研究方向: 物联网、计算机应用、图像处理。