

一种面向 SAR 图像多尺度舰船目标的检测算法

宋浩生, 甘精伟, 虞华, 王琳, 秘璐然

(中国电子科技集团公司第54研究所, 石家庄 050081)

摘要: 在合成孔径雷达图像的舰船目标检测任务中, 不同目标的尺度多样性给检测算法带来了巨大挑战; 为了解决多尺度舰船目标检测难题, 提出了一种 BAPT-YOLOv8n 算法, 该算法以 YOLOv8n 为基础框架, 通过引入卷积块注意力模块重构颈部金字塔网络, 提升了对多层次特征的融合能力与对多尺度目标的特征提取能力; 此外, 采用基于 Transformer 的检测头结构, 进一步提高特征表示能力和上下文信息利用能力, 从而改善了小目标的检测效果; 在 HRSID 数据集和 SSDD 数据集上的对比实验表明, 所提算法在检测精度上分别达到 93.6% 与 98.9%, 优于其他对比算法; 消融实验进一步验证了算法中各改进部分的有效性, 表明该算法能够更好地适应多尺度舰船目标检测问题。

关键词: 合成孔径雷达; 舰船检测; 深度学习; 注意力金字塔; Transformer

Detection Algorithm for Multi-scale Ship Targets in SAR Images

SONG Haosheng, GAN Jingwei, YU Hua, WANG Lin, MI Luran

(The 54th Research Institute of CETC, Shijiazhuang 050081, China)

Abstract: In ship target detection in synthetic aperture radar (SAR) images, different targets with diverse scales have brought significant challenges to detection algorithms. To address these issues, a BAPT-YOLOv8n algorithm is proposed. Taken the YOLOv8n as a framework, convolutional block attention modules are introduced to reconstruct the neck pyramid network, and to enhance the fusion capability of multi-level features and the feature extraction capability for multi-scale targets. Additionally, the Transformer-based detection head structure is used to further improve the abilities of feature representation and context utilization, thereby enhancing the detection performance of small targets. Comparative experiments on the HRSID and SSDD datasets show that, compared with other algorithms, the proposed algorithm achieves the detection accuracy of 93.6% and 98.9%, respectively. The ablation experiments further validate the effectiveness of the improved algorithm, the results show that this algorithm can better adapt multi-scale ship target detection tasks.

Keywords: SAR; ship detection; deep learning; attention pyramid; Transformer

0 引言

合成孔径雷达 (SAR, synthetic aperture radar) 是一种高分辨率成像雷达^[1-4]。它能够在全天时、全天候的条件下工作, 有效地识别伪装和穿透掩盖物, 检测隐藏的物体^[5]。随着星载和机载合成孔径雷达的快速发展, 合成孔径雷达已广泛应用于民用和军事领域, 如 TerraSAR-X、RADARSAT-2、Sentinel-1 等。近年来, 随着海洋活动的日益频繁, 舰船检测在军事侦察、海洋管理以及海洋科学研究等领域发挥着至关重要的作用。

利用 SAR 图像进行海上舰船目标检测也成为了 SAR 应用的一个热点研究课题。SAR 图像中的舰船分布包含海上场景和近岸场景两种典型情况, 存在不同类型的复杂背景, 包括港口、岛屿、水坝或者其他与舰船类似的物体, 以及由舰船尾流和波浪引起的不同海面杂波。同时, SAR 图像具有复杂的尺度变化, 这也成为影响舰船检测精度的主要因素之一。

目前, 提出了许多方法来检测 SAR 图像舰船目标。传统的 SAR 图像舰船检测方法仅提取简单的底层特征, 设计舰船尾迹、纹理、尺度等特征来检测舰船, 需要选

收稿日期: 2024-07-04; 修回日期: 2024-08-22。

作者简介: 宋浩生 (1999-), 男, 学士, 在读研究生。

通讯作者: 甘精伟 (1964-), 男, 研究员。

引用格式: 宋浩生, 甘精伟, 虞华, 等. 一种面向 SAR 图像多尺度舰船目标的检测算法[J]. 计算机测量与控制, 2025, 33(1): 211-217, 225.

择目标特征并确定特征向量维数,依赖于人工参与。因此,传统算法通常只针对特定场景,不同场景需要进行重构,导致泛化性差、检测效率低且虚警率较高。其中恒虚警率检测算法(CFAR, constant false alarm rate)^[6]及其改进算法^[7-9]是研究和应用最为广泛的。然而,建立一个准确的CFAR算法统计模型是一项非常耗费人力的任务,并且这些模型容易受到波浪和环境变化的影响,并不适用于对海岸、港口、暗礁附近的舰船进行检测。

随着人工智能的快速发展,基于卷积神经网络(CNN, convolutional neural network)的方法被许多研究者应用于SAR图像中的舰船检测。在早期阶段,卷积神经网络被用于舰船检测的各个部分,如陆地掩蔽、感兴趣区域提取和舰船目标识别。同时,一些基于CNN的舰船检测方法也陆续出现并且在高分辨率合成孔径雷达图像中展现出优于传统方法的检测性能。现阶段,基于卷积神经网络的舰船检测算法大多采用端到端模式,涌现出了一系列的研究成果。Li等人^[10]提出了一种改进的Faster R-CNN方法,并提供了SAR舰船检测数据集(SSDD)来训练和测试模型。Jiao等人^[11]将FPN和Faster R-CNN结合来检测多尺度检测。在文献[12]中,作者通过将卷积块注意力模块(CBAM, convolutional block attention module)模块引入到FPN结构中构建出密集注意金字塔网络,来提高算法对多尺度舰船的检测性能。文献[13]提出了一种用于SAR舰船检测的多维域深度学习网络,利用空间和频域的互补特征在复杂背景下表现良好。随着Transformer在自然语言处理邻域优秀表现,研究者们开始探索其在计算机视觉领域的应用潜力。端到端基于Transformer的检测器(DETR, detection transformer)^[14-17]由于其精简的结构和无需手工制作的组件,受到了学术界的广泛关注。

经过验证,现有的合成孔径雷达舰船检测方法更注重从单一和突出的特征信息中发现突出物体,如海上场

景中的大型舰船。在实际进行多尺度舰船目标检测时,由于目标尺度跨度大,同一算法很难兼顾不同尺度舰船目标检测效果。对于较小目标,尾流或其他干扰物体被误识别为舰船或小型舰船目标的情况较为普遍,这进一步限制了多尺度舰船目标检测的性能。

1 BAPT-YOLOv8n 算法设计思路

针对SAR图像多尺度舰船目标检测面临的多尺度平衡和小目标检测困难等问题,本文设计实现了一种BAPT-YOLOv8n网络,用以提升SAR图像多尺度舰船目标检测的效果。该网络以YOLOv8n为基准框架,在颈部金字塔网络中引入卷积块注意力模块,提高多层次特征的融合能力,并采用基于Transformer结构的检测头,进一步提升多尺度目标检测能力。该网络能够充分利用多层次特征信息,并增强小目标特征表征,能够有效提升SAR图像多尺度船舶目标检测效果。

BAPT-YOLOv8n框架如图1所示。BAPT-YOLOv8n网络由骨干、颈部和头部3个部分组成。其中,骨干部分采用了经典的CSPDarkNet-53网络,用于提取图像中丰富的特征信息;颈部则采用本文所提的新型注意力金字塔结构来提升网络的多尺度目标特征融合能力;头部通过引入Transformer结构检测头来提升对算法的特征表示能力和上下文信息利用能力。通过颈部和头部的专门设计,保证了网络在SAR图像多尺度船舶目标的检测能力。

2 BAPT-YOLOv8n 算法介绍

与YOLOv8n算法相比,BAPT-YOLOv8n算法在继承优点的同时,构建了新型的注意力金字塔结构,并采用Transformer结构检测头替换原检测头结构。下面将重点深入讲解算法的每个部分。

2.1 骨干部分

骨干部分在目标检测算法中扮演着至关重要的角色,其任务是从输入图像中提取多尺度、高层次的特征表示,这些特征对于后续的目标定位与分类至关重要。

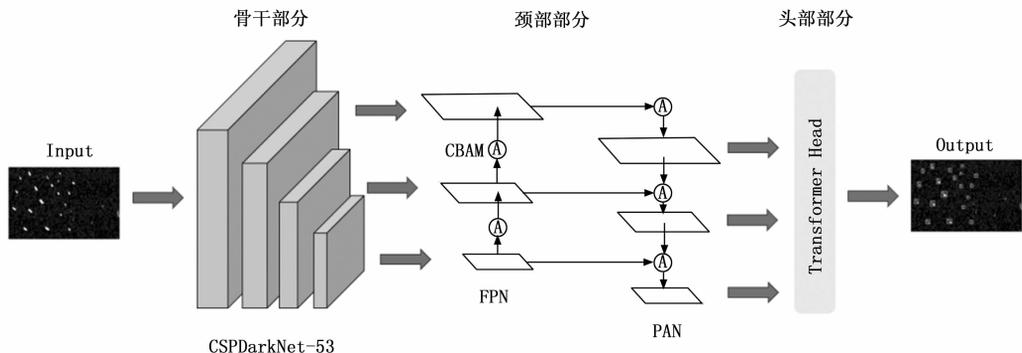


图1 BAPT-YOLOv8n 算法框架图

因此, 骨干部分的设计和性能直接决定了整个目标检测算法的准确性和效率。本文采用的骨干网络是基于 DarkNet-53 的改进版本 CSPDarkNet-53, 该网络引入了跨阶段局部连接结构, 以提升网络的学习能力和特征表达能力。具体而言, CSPDarkNet-53 由 52 个卷积层和 1 个全连接层组成, 其详细结构见图 2。

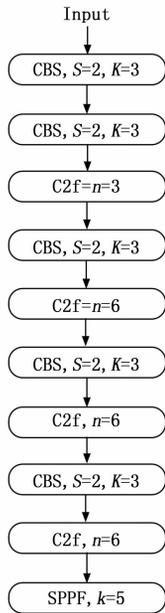


图 2 CSPDarkNet-53

其中, CBS 由 3 个部分组成, 分别是一层 CNN、一层 BatchNorm 与 SiLU 激活函数, 通过这种组成在提取的过程中增强网络的表达能力和学习能力, 从而提升算法在 SAR 舰船检测任务中的性能和效果。网络还通过 C2f 模块将特征提取过程分为压缩阶段和扩展阶段。在压缩阶段, 通过卷积操作对输入特征图进行降维, 以减少计算复杂度和内存消耗; 在扩展阶段, 通过卷积和上采样操作将特征图恢复到原始尺寸, 并进行特征融合, 以保持信息的完整性和丰富性。此外, 引入残差连接进一步增强了网络的特征提取能力, 有效缓解了梯度消失问题, 提高了模型的训练稳定性。这种设计既保证了特征提取的有效性, 又降低了计算复杂度, 使得模型在处理大规模数据时能够更加高效。除此之外, CSPDarkNet-53 通过引入 SPPF 模块, 能够有效处理不同尺寸的输入图像, 而无需事先进行尺寸调整。SPPF 通过对不同尺度的特征进行池化操作, 能够捕获图像中的多尺度信息。相比于传统的 SPP 方法, SPPF 在保持相似性能的同时, 进一步提升了处理速度和效率, 这主要得益于 SPPF 对池化操作的优化和并行化处理。

因此, CSPDarkNet-53 作为 BAPT-YOLOv8n 算法的骨干部分, 在保持较高精度的同时有效降低了计算复杂度和内存消耗, 使其适用于 SAR 图像舰船检测任务,

尤其在实时性要求较高的场景下展现出优异的性能。

2.2 颈部部分

颈部部分是连接骨干网络和头部网络的中间部分, 主要负责进一步加工和优化骨干网络提取的特征信息。YOLOv8n 算法颈部部分采用的是 PAN-FPN 网络, 然而, 当前的 PAN-FPN 在应对复杂场景和多样性目标时存在一些局限性。首先, 它主要依赖于简单的上采样和下采样操作, 这种方式在处理高分辨率图像或目标尺寸差异较大的场景时可能不够精细, 导致一些细节信息的丢失, 从而影响目标检测的准确性。另外, PAN-FPN 在融合不同尺度特征图时通常采用相同的权重分配方式, 没有考虑到不同尺度特征图在目标检测任务中的重要性差异。这种静态权重分配可能会限制算法在多尺度目标检测任务中的性能表现, 因为不同尺度的特征对于不同大小或类型的目标具有不同的信息贡献。

在进行图像特征提取时, 高层次的特征图往往具有较大的接受域, 语义信息较多, 空间信息较少, 适合检测大型舰船。而低层次特征图空间信息较多, 特征较浅, 接收域较小, 适合小尺度舰船检测。因此, 如何有效地将语义信息与空间信息相结合是多尺度舰船检测的关键。此外, 需要在海量特征中提取特定尺度的显著信息, 以突出不同尺度的舰船目标, 消除复杂背景的干扰, 从而更好地检测 SAR 图像中各种场景下的多尺度舰船目标。CBAM 注意力机制可以突出特定尺度的显著特征图, 并通过多尺度特征图进行加权, 根据待检测舰船的尺度自适应地提取合适的尺度特征。

本文在 YOLOv8n 颈部结构基础上对网络进行改进, 将 CBAM 注意力机制通过横向与上下连接嵌入到颈部网络中, 用以提升算法对不同尺度和位置的特征关注, 并优化特征融合过程。构建出新型注意力金字塔网络结构作为算法的颈部部分, 如图 3 所示。改进后的颈部部分可以更好的提取 SAR 图像中包含语义、空间信息的多尺度和显著特征来检测多尺度舰船。

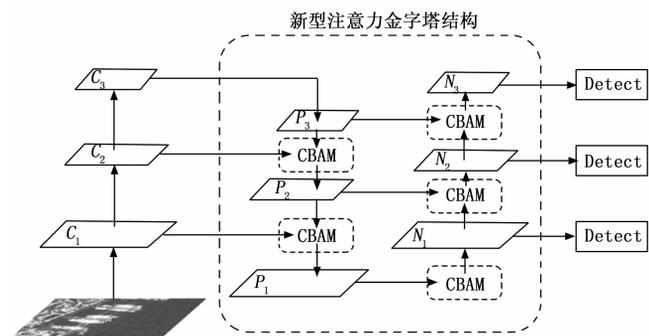


图 3 新型注意力金字塔网络

具体来说, 在骨干部分选用最后 3 个卷积阶段生成的特征图 (C_1, C_2, C_3) 作为注意力金字塔的输入。

在注意力金字塔的自顶向下网络中，通过对高层特征图进行上采样，得到语义信息更多、分辨率更高的特征图，并通过横向连接将骨干部分生成的特征图和自顶向下路径中具有相同空间大小的特征图合并，再经过 CBAM 注意力机制生成融合特征图 (P_1, P_2, P_3)。同理，在注意力金字塔的自底向上网络中，通过下采样、横向连接与 CBAM 注意力机制，生成最终的融合特征图 (N_1, N_2, N_3)，并送入 Transformer 检测头中。

注意力金字塔结构中的关键组件 CBAM 注意力机制由通道注意和空间注意组成，分别沿两个维度获得注意力图。如图 4 所示。

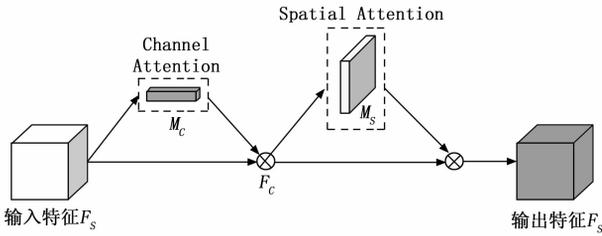


图 4 CBAM 注意力机制

将输入特征图 $F \in R^{C \times H \times W}$ 乘以通道注意力图 $M_c \in R^{C \times 1 \times 1}$ 和空间注意力图 $M_s \in R^{C \times H \times W}$ ，得到一个显著特征图。CBAM 的总体流程如下：

$$F_c = M_c(F) \otimes F \quad (1)$$

$$F_s = M_s(F_c) \otimes F_c \quad (2)$$

其中： $F_c \in R^{C \times H \times W}$ 是通过通道注意力后得到的特征图， \otimes 表示逐元素乘法。

通道注意力模块主要关注输入特征图的具体重要内容。它同时通过最大池化运算和平均池化运算来聚合空间信息，然后通过多层感知器 (MLP) 来减少参数。计算过程总结如下：

$$M_c(F) = \sigma \{ MLP[avgPool(F)] + MLP[maxPool(F)] \} \quad (3)$$

其中：AvgPool 和 MaxPool 分别表示平均池化和最大池化操作， σ 为 sigmoid 函数。

空间注意模块则是关注显著特征的具体位置。它首先分别沿着通道维度应用最大池化运算和平均池化运算，并将结果在通道维拼接起来。然后，利用 7×7 的卷积层对拼接的特征图生成空间注意力图。空间注意力的计算方法如下：

$$M_s(F_c) = \sigma \left\{ C_{7 \times 7} \left[\begin{matrix} avgPool(F) \\ maxPool(F) \end{matrix} \right] \right\} \quad (4)$$

其中： $C_{7 \times 7}$ 表示 7×7 卷积层。

CBAM 通过通道注意力和空间注意力依次提高了输入特征图的代表能力。通道注意力集中于全局，空间注意力集中于局部信息。在本文中，CBAM 将特征金字塔中的特征图紧密相连，强调了被探测舰船不同尺度

对应的特定尺度特征图的重要信息，可以有效地消除 SAR 图像各种场景中的虚警，例如近岸场景中的地物和近岸场景中类似小型舰船的散斑噪声。

2.3 头部部分

头部部分负责处理颈部结构传递的特征，并生成最终的目标检测结果，包括目标的类别、位置和置信度等。

传统 YOLOv8n 的无锚框检测头存在以下问题：在处理小目标时表现不佳，小目标的细节和位置往往无法得到精确的捕捉，导致小目标的检测精度较低；SAR 图像中存在大量的背景杂波和海况变化，而 YOLOv8n 的无锚框方法对此类干扰较为敏感，从而影响了目标检测的性能；SAR 图像中的舰船目标通常会出现较多的重叠情况，这可能导致 NMS 算法难以有效地去除重叠的检测框，且小尺度的舰船目标在像素级别上的差异可能不明显，容易受到 NMS 算法的影响而被错误地抑制。

为增强算法对多尺度目标检测的能力，本文引入 Transformer 结构的检测头替换 YOLOv8n 的检测头。Transformer 结构具有强大的特征提取能力和自注意力机制，能够捕捉丰富的上下文信息和长距离依赖关系，能够增强对舰船目标的检测能力并提高定位精度。这种检测头结构也可以在避免使用 NMS 后处理繁琐操作的同时更好的解决重叠框导致的小目标检测失准问题。

如图 5 所示，Transformer 结构检测头由自注意力机制 (Self-Attention) 和跨注意力机制 (Cross-Attention) 以及前向传播网络组成。

不同于 YOLOv8n 检测头的无锚框机制进行检测，Transformer 检测头基于对象查询 (Object Query) 进行检测。对象查询是一组可学习的参数，每个检测目标由一个对象查询表示。所有的对象查询经过 Self-Attention 模块后与新型注意力金字塔输出的三层目标检测层共同作用于 Cross-Attention 模块，通过前向传播优化并映射到分类分数和边界框。

随着检测头结构的替换，需要相应地引入查询策略并调整算法的损失函数。

查询选择 (Query Selection) 的作用是从注意力金字塔输出的特征序列中选择固定数量的特征作为对象查询。本文所使用的查询选择策略是 IoU 感知的查询选择。该策略确保模型在选择预测框时不仅考虑分类分数，还综合考虑了 IoU 分数，即预测框与真实框之间的重叠度量。通过这种方式，还能增强算法对未知尺度和形状舰船目标的泛化能力。使得算法在检测多尺度舰船目标时，都能保持较高的精度。

在模型训练中，训练效果与损失函数的设计紧密相关。损失函数的设计不仅仅是为了量化模型预测与实际

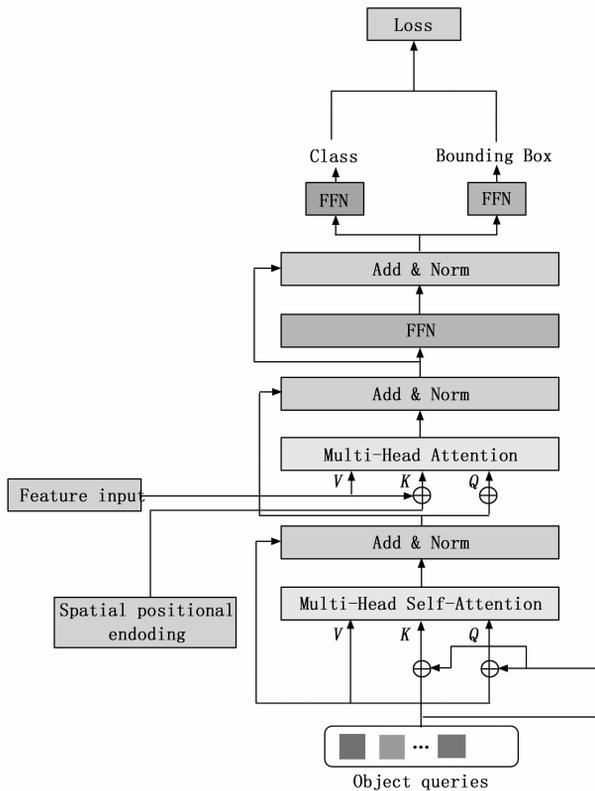


图 5 基于 Transformer 的检测头结构图

标签之间的差异,更重要的是能够有效引导模型在训练过程中朝着最小化预测误差的方向迭代优化。恰当选择的损失函数不仅可以加速模型的收敛速度,还能显著提高模型在未见过数据上的泛化能力和准确性。因此,查询选择策略调整后,损失函数 $L(\hat{y}, y)$ 也应进行更新:

$$L(\hat{y}, y) = L_{\text{box}}(\hat{b}, b) + L_{\text{cls}}(\hat{c}, \hat{b}, y, b) = L_{\text{box}}(\hat{b}, b) + L_{\text{cls}}(\hat{c}, c, IoU) \quad (5)$$

其中: $\hat{y} = \{\hat{c}, \hat{b}\}$, $y = \{c, b\}$ 分别表示预测值和真实值 c 和 b 分别表示类别和边界框。 IoU 引入分类损失函数以实现正样本的分类和定位一致性约束。

3 实验与结果分析

3.1 数据集

为验证 BAPT-YOLOv8n 算法的效果和性能,本文选用 HRSID 和 SSDD 这两个不同分辨率范围的 SAR 多尺度舰船检测数据集进行性能测试和消融实验。

HRSID 数据集由 Wei 等人发布^[18]。该数据集图像来自 Sentinel-1 和 TerraSAR-X 卫星,包括不同分辨率的 SAR 图像、极化、海况、海域和沿海港口,如开阔海域、近岸区域及复杂港口环境。图像总数为 5 604 张,统一尺寸为 800×800 ,分辨率范围为 $0.5 \sim 3$ m。训练集和测试集根据作者所提供的配置文件进行划分,训练数据为 3 642 个,测试数据为 454 个。

SSDD 数据集由 Li 等人^[10]构建。它的图像来自 Sentinel-1、TerraSAR-X 和 RadarSat-2 卫星,这些数据有 HH、HV、VV 和 VH 四种极化模式,包含不同环境下的多尺度舰船。图像总数为 1 160 张,分辨率从 1 到 15 米不等。平均每张图像有 2.12 艘船。平均图像大小为 500×500 。训练集和测试集的划分为 $8:2$ 。将数字以 1 和 9 结尾的图像划分为测试集,其他图像划分为训练集。

3.2 评价指标

本文使用精度 (P)、召回率 (R) 和平均精度 (mAP) 来评估模型的检测性能。其中,精度 (P) 表示为算法检测到的正例样本在所有预测为正例样本中的比例。在舰船检测中,高精度意味着模型在识别舰船时,其预测为舰船的样本中,真实舰船的比例很高。然而,仅仅追求高精度可能导致模型过于保守,错过一些真实的舰船。精度的定义是:

$$P = \frac{TP}{TP + FP} \quad (6)$$

其中: TP 代表真阳性,即被模型预测为正类的正样本; FP 代表假阳性,即被模型预测为正类的负样本。

召回率 (R) 表示为检测到的正例样本在总实际正例样本中的比例。在舰船检测的背景下,高召回率意味着模型能够发现数据集中绝大多数的舰船,即使这可能导致一些非舰船对象被错误地识别为舰船,召回的定义是:

$$R = \frac{TP}{TP + FN} \quad (7)$$

其中: FN 代表假阴性,即被模型预测为负类的正样本。

平均精度 (mAP) 用来评价数据集中不同舰船检测方法的性能。在舰船检测中, mAP 是衡量不同检测方法整体性能的重要指标。由于 mAP 综合考虑了模型的精度和召回率,因此它更适合用于比较不同检测算法或模型在同一数据集上的表现, mAP 定义为:

$$mAP = \int_0^1 P(R) dR \quad (8)$$

其中: $P(R)$ 是准确度-召回率的曲线,展示了不同召回率水平下模型所能达到的精度。这条曲线直观地反映了模型在不同应用场景下的表现。

3.3 对比实验

基于 HRSID、SSDD 两个不同分辨率范围的多尺度舰船数据集,对 BAPT-YOLOv8n 算法与 Faster R-CNN^[19]、DCN^[20]、YOLOv5n^[21]、YOLOv7n^[22]、YOLOv8n^[23]、DAPN^[12]、ARPN^[24]、BL-Net^[25] 以及 YOLO 的改进算法^[26-27] 进行对比实验,结果如表 1~2 所示。

表 1 HRSID 数据集对比实验结果

方法	$mAP_{50}/\%$	$P/\%$	$R/\%$
Faster R-CNN	80.7	81.5	82
DCN	82.1	81.5	83.5
YOLOv5n	92.4	90.9	85.3
YOLOv7n	89.6	91.5	80.6
YOLOv8n	91.3	89.8	83.5
DAPN	81.8	80.5	83.4
ARPN	81.8	85.7	83.8
BL-Net	88.7	91.6	89.7
YOLO-SRBD	92.4	91.3	86.7
SEFRNet	93.3	94.7	87.7
Ours	93.6	92.7	90.4

表 2 SSDD 数据集对比实验结果

方法	$mAP_{50}/\%$	$P/\%$	$R/\%$
Faster R-CNN	88.9	89.7	87.8
DCN	92.3	86.2	93.0
YOLOv5n	96.3	95.7	94.2
YOLOv7n	93.7	91.1	84.9
YOLOv8n	95.1	92.5	93.6
DAPN	90.6	85.8	91.4
ARPN	89.9	85.4	90.6
BL-Net	95.3	91.3	96.1
YOLO-SRBD	98.3	97.9	96.0
SEFRNet	96.1	95.9	95.0
Ours	98.9	96.8	97.4

由实验结果可以看出，无论在 HRSID 数据集还是 SSDD 数据集中，BAPT-YOLOv8n 算法的检测效果均优于基线 YOLO 系列算法和其他经典 SAR 图像目标检测算法。与 YOLO 的改进算法相比，虽然 SEFRNet 和 YOLO-SRBD 分别在 HRSID 和 SSDD 数据集中的 P 值高于所提出算法。但是单方面追求高 P 值可能导致算法过于保守而错过一些真实的舰船。这使得两个算法都在另一个数据集中的表现出现下降，且在两个数据集中， mAP 均低于 BAPT-YOLOv8n 算法。这说明 BAPT-YOLOv8n 算法具有更好的 SAR 图像多尺度舰船检测性能。

令人意外的是，实验中 YOLOv5n 在检测性能方面优于 YOLOv7n 和 YOLOv8n。这一优越性能的背后，我们推测可能与 YOLOv5 的网络结构相关。具体而言，YOLOv5n 的网络结构在特征提取阶段展现出了对 SAR 图像特点的出色适应性，这为其在检测任务中的卓越表现奠定了基础。此外，我们注意到 YOLOv5n 在设计上的轻量级特性也是其性能优越的一个重要因素。在本次实验所使用的数据集环境下，YOLOv5n 凭借其轻量级的架构，能够更有效地收敛，因此得到较好的性能表现。

算法的部分效果图如图 6 所示。其中左侧的图片为原图，中间的图片为真实标签，右侧的图片为 BAPT-YOLOv8n 算法的推理结果。由结果可以看出，在分别仅有小型、中型、大型舰船的 3 种条件下，本文提出的算法能够精准地识别出 SAR 图像中的舰船目标；即使在存在小型与中型舰船混合的海上场景中，尽管存在较大的海上杂波干扰，算法仍能够识别图像中所有的舰船目标；在同时存在中型与大型舰船的近岸场景中，尽管陆地存在较大的干扰，算法仍能够有效识别出舰船目标；在大、中、小 3 种尺度舰船同时存在时，BAPT-YOLOv8n 算法依旧可以识别到所有舰船目标。说明本文提出的算法在面对复杂多变的海上和近岸场景时，不仅在单一尺度舰船目标识别方面表现出色，对多尺度舰船目标依然能够保持高效的识别能力，更好地适应多尺度舰船目标检测问题。

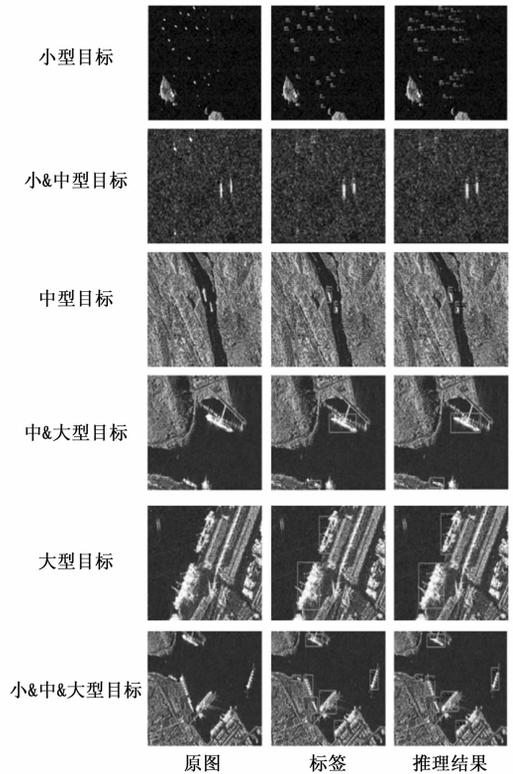


图 6 提出的算法效果图

3.4 消融实验

为了验证不同的改进方法对模型的影响，本次研究采用了逐步添加改进方法的消融实验，并在 HRSID 与 SSDD 数据集上进行了性能比较。 M_1 表示使用新型注意力金字塔颈部结构， M_2 表示使用基于 Transformer 结构的检测头。实验结构如表 3、表 4 所示。

实验结果表明，在 HRSID 数据集上，YOLOv8n 算法 mAP_{50} 为 91.3%。仅将新型注意力金字塔颈部结构替换 YOLOv8n 的颈部结构后， mAP_{50} 提升了 1 个百

分点, 达到 92.3%。仅将基于 Transformer 结构的检测头替换 YOLOv8n 的检测头后, mAP_{50} 提升了 1.3 个百分点, 达到了 92.6%。当同时使用两种改进方案时, mAP_{50} 提升了 2.3 个百分点, 达到了 93.6%。

表 3 HRSID 数据集消融实验结果

方法	M_1	M_2	$mAP_{50}/\%$	$P/\%$	$R/\%$
YOLOv8n	×	×	91.3	89.8	83.5
	×	√	92.6	92	84.6
	√	×	92.3	91.5	83.9
	√	√	93.6	92.7	90.4

表 4 SSDD 数据集消融实验结果

方法	M_1	M_2	$mAP_{50}/\%$	$P/\%$	$R/\%$
YOLOv8n	×	×	95.1	92.5	93.6
	×	√	97.4	94.7	92.1
	√	×	98.0	96.6	94.1
	√	√	98.9	96.8	97.4

同样在 SSDD 数据集上, YOLOv8n 算法 mAP_{50} 为 95.1%。颈部结构替换为新型注意力金字塔后 mAP_{50} 提升了 2.9 个百分点, 达到了 98.0%。检测头替换为基于 Transformer 结构的检测头后, mAP_{50} 提升了 2.3 个百分点, 达到了 97.4%。同时使用两种改进方案时, mAP_{50} 提升了 3.8 个百分点, 达到了 98.9%。

因此, 这两种改进方法都增强了算法在多尺度舰船目标检测方面的性能, 并且在两个不同的数据集上表现出一致的提升效果, 进一步验证了它们在提高检测性能方面的有效性和可靠性。

4 结束语

本文提出了一种面向 SAR 图像多尺度舰船目标检测的 BAPT-YOLOv8n 算法, 该算法通过构建新型注意力金字塔颈部结构、采用基于 Transformer 结构的检测头, 显著提高了对不同尺度舰船的关注度, 增强了对多尺度舰船目标的检测能力。在 HRSID 和 SSDD 数据集上的实验结果表明, 本文提出的 BAPT-YOLOv8n 算法与其他目标检测算法相比具有更好的检测性能, 具备更好的多尺度舰船目标检测能力。在小目标舰船密集的场景中, 算法偶尔会出现少量的漏检和误报现象, 这为我们后续的优化和改进提供了方向。在后续的研究中, 我们还将探索算法应用于其他 SAR 图像的目标检测任务中。

参考文献:

[1] BALZ T, HAMMER H, AUER S. Potentials and limitations of SAR image simulators-A comparative study of three simulation approaches. [J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2015, 101: 102-109.

[2] 徐 丰, 王海鹏, 金亚秋. 深度学习在 SAR 目标识别与地

物分类中的应用 [J]. 雷达学报, 2017, 6 (2): 136-148.

[3] NOVELLO C, FORNARO G, MARTOR-ELLA M. Focused SAR image formation of moving targets based on Doppler parameter estimation [J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53 (6): 3460-3470.

[4] 王 俊, 郑 彤, 雷 鹏, 等. 深度学习在雷达中的研究综述 [J]. 雷达学报, 2018, 7 (4): 395-411.

[5] 黄钟铃, 姚西文, 韩军伟. 面向 SAR 图像解译的物理可解释深度学习技术进展与探讨 [J]. 雷达学报, 2022, 11 (1): 107-125.

[6] ROBEY F C, FUHRMANN D R, KELLY E J, et al. A CFAR adaptive matched filter detector [J]. IEEE Transactions on Aero-space and Electronic Systems, 1992, 28 (1): 208-216.

[7] 艾加秋, 曹振翔, 毛宇翔, 等. 一种复杂环境下改进的 SAR 图像双边 CFAR 舰船检测算法 [J]. 雷达学报, 2021, 10 (4): 499-515.

[8] SCHWEGMANN C P, KLEYNHANS W, SALMON B P. Manifold adaptation for constant false alarm rate ship detection in south african oceans [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8 (7): 3329-3337.

[9] QIN X, ZHOU S, ZOU H, et al. A CF-AR detection algorithm for generalized gamma distributed background in high-resolution SAR images [J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10 (4): 806-810.

[10] LI J W, QU C W, SHAO J Q. Ship detection in SAR images based on an improved faster R-CNN [C] // Beijing, China, 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR DATA), 2017: 1-6.

[11] WANG J, LU C, JIANG W. Simultaneous ship detection and orientation estimation in SAR images based on attention module and angle regression [J]. SENSORS, 2018, 18 (9): 2851.

[12] CUI Z, LI Q, CAO Z, et al. Dense attention pyramid networks for multi-scale ship detection in SAR images [J]. IEEE Transactions on Geoscience & Remote Sensing, 2019, 57 (11): 8983-8997.

[13] LI D, LIANG Q, LIU H, et al. A novel multidimensional domain deep learning network for SAR ship detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-13.

[14] KIM B, LEE J, KANG J, et al. HOTR: End-to-End Human-Object interaction detection with transformers [C] // Nashville, TN, USA: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 74-83.

(下转第 225 页)