http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0853

基于YOLOv11的舰船关键部位检测算法

王瑶，曹司磊*，刘明刚，王颖颖
(海军航空大学，烟台 264001)

摘　　要：在当前及未来海上维权及海域监测活动中研究如何迅速准确获取舰船关键部位信息，有效判读海域实时情况，具有非常重要的实际意义。基于深度学习的目标检测算法能够区分多类型目标信息，但针对舰船关键部位检测算法的研究较少，且对多通道特征图间的关联性分析不足，对尺度不同的舰船关键部位目标的特征利用率较低，容易错失重要特征信息而使检测精度降低。对此，提出一种基于YOLOv11的舰船关键部位检测算法，提升对关键部位的精准高效检测水平。对当前舰船关键部位检测存在的重难点问题进行深入剖析，结合现有算法进行对比研究；根据实际问题提出舰船关键部位检测的YOLOv11目标检测算法的改进方案；构建可见光舰船关键部位数据集，通过设计对比实验及消融实验，充分证明了所提算法对舰船关键部位检测的精准性及时效性，能够为实现复杂海洋环境中对多类型舰船关键要害部位的精细区分、有效监测海上舰船信息提供可行技术支持和理论支撑。

关　键　词：舰船目标；关键部位；YOLOv11；注意力机制；特征融合
中图分类号：E919
文献标志码：A　　**文章编号**：1001-5965(2026)11-3969-13

随着科学技术发展和材料设备的革新，智能化、自主化、精确化已成为海上斗争形式发展的主流趋势。海洋环境复杂多变，多类型舰船密集分布，有效监测海域实时信息、掌握舰船动态是维护海洋安全的重要途径。当前对舰船目标的检测方式没有区分舰船的关键部位，如雷达、驾驶舱、天线、船首，而舰船姿态、关键部位位置严重影响监测设备对舰船目标的精准判读。且海上舰船目标种类繁多、形态各异，在复杂海洋环境中对多类型舰船关键部位的精细区分，要能够有效区分背景与舰船信息，要求有效应对舰船移动所产生的几何形变对检测带来的影响。传统的海上舰船检测算法大多需要人为设置特征，算法复杂程度高、实时性不强、鲁棒性低。随着计算机技术、人工智能技术等新兴技术的进步革新和在多领域的深入应用，基于深度学习的目标检测算法逐渐显现出特有优势，且已广泛应用于对舰船的检测识别工作中。在复杂场景中，使用基于深度学习的目标检测技术能够大幅提升精度及效率，使海上舰船目标检测识别算法迅速从传统算法向深度学习领域发展。

文献[1]通过聚合密集注意力特征法获得高分辨率特征图像，并通过密集连接和迭代融合方法增强模型泛化性，有效提升了网络对舰船目标的检测精度；文献[2]使用高帽及低帽变换的方法对输入图像进行处理，并使用阈值分割法和归并算法获取候选区域，基于深度学习的分类思想检测舰船目标，相比传统算法，网络检测目标的精度及速度提升显著；文献[3]通过在候选框中获取目标候选特征，改进了YOLO网络，提升了分辨率高的卫星图像中舰船目标的检测精度。文献[4]将坐标注意力

收稿日期：2024-12-03；**录用日期**：2025-07-17；**网络出版时间**：2025-09-04 15:16
网络出版地址：link.cnki.net/urlid/11.2625.V.20250904.1419.001
基金项目：青年面上专项基金(ZR2024QE034)
* **通信作者**. E-mail：caosileimail@163.com

引用格式：王瑶，曹司磊，刘明刚，等. 基于YOLOv11的舰船关键部位检测算法[J]. 北京航空航天大学学报，2026，52(11)：3969-3981.
Wang Y，Cao S L，Liu M G，et al. Detection algorithm of ship key parts based on YOLOv11[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(11)：3969-3981 (in Chinese).

(coordinate attention, CA)机制引入YOLOx模型中,并在网络的IM-YOLO-S颈部结构中应用双向特征金字塔网络(bidirectional feature pyramid network, BiFPN)算法,提升了网络对小尺度舰船目标的检测性能;文献[5]基于合成孔径雷达(synthetic aperture radar, SAR)图像构建检测多尺度舰船目标的跨连接特征金字塔网络(cross connected feature pyramid network, CCFPN),对深层及浅层特征图进行融合,并使用通道拼接方式提升了特征图中的信息量,运用多路空洞卷积增强对浅层目标特征的提取能力,能准确检测复杂背景中的多尺度舰船目标;文献[6]在YOLOv3-tiny中应用了特征映射方法,使模型的预测层能更多提取关键目标特征,提升了小尺寸舰船分布密集区域的目标检测精度。

当前多种目标检测算法能够实现不同场景下的舰船目标检测任务,但对于未来日益复杂海战场环境及精准打击需求,仍存在以下几点不足:

1)将舰船目标看作整体进行检测识别,未对关键部位进行精准检测,没有充分利用细粒度特征,不能提供精准的舰船目标指引信息;

2)海上舰船目标种类多,形态各异,现有算法在对多通道特征图间的关联性分析薄弱,对不同尺度舰船目标的特征利用率不高,易丢失关键特征信息而降低检测精度;

3)反舰设备雷达/红外制导体制受自然环境及电磁干扰影响严重,在实现反舰设备精准目标选择方面存在一定技术瓶颈;使用红外/雷达图像进行舰船目标检测不仅易受到多种干扰还会丢失舰船类型、关键部位等重要信息。当前针对舰船目标的数据集较少,尤其是可见光图像数据集,对舰船关键部位的分类数据极为缺乏。

针对当前舰船目标检测算法的不足及图像采集过程中目标特性变化剧烈的问题,本文提出基于YOLOv11的可见光舰船图像关键部位检测算法。具体如下:①针对海上航行舰船目标尺寸由于运动造成几何变化而影响精准检测的问题,利用第4代可变形卷积(deformable convolutionsv4, DCN v 4)重构的新模块替换YOLOv11主干网络中C3k2原来的bottleneck块,无需增加太多模型的训练参数,能够更准确地估计移动舰船目标空间形变。②针对舰船关键部位属于小目标检测范畴、检测难度大的问题,利用可变核卷积(arbitrary kernel convolutional, AKConv)改进跨阶段部分网络模块(volumetric grid spatial cross stage partial network block, VoVGSCSP)结构,构建AKConv跨阶段部分网络模块(arbitrary kernel convolutional-volumetric grid spatial cross stage partial network block, AVB),再和BiFPN结合构建新型颈部网络Bi-AVB-FPN,以丰富多尺度特征融合,增强特征语义信息,帮助模型更准确地定位关键舰船部位的目标区域。③针对边缘设备资源普遍有限且检测算法适用性低的问题,在Neck部分引入了轻量级分组混洗卷积(grouped spatial convolution, GSConv)模块,以减少计算量和内存损失。通过改进不仅可以实现更高的检测精度,特别是在小尺度目标检测方面,还可以降低功耗和能耗。④针对SAR图像、红外图像获取易受复杂电磁环境及自然环境干扰影响的问题,自建可见光舰船关键部位数据集。

2024年提出的YOLOv11目标检测网络属于一阶目标检测网络,模型性能得到显著提升,相对其他基于深度学习的目标检测算法表现出精度及速度方面的突出优势,且可见光图像具有特征突出、信息多样、直观易读的特性,受温度、风力等自然环境及干扰影响较小,包含丰富的光谱信息,能够较好抵抗烟幕等遮挡干扰影响,有利于低空、超低空图像信息全面完整表达,研究基于YOLOv11的可见光舰船图像关键部位检测能够更好地适应复杂环境,准确检测识别多类型关键目标信息。本文通过自建可见光数据集训练并测试改进的要害部位检测网络性能,证明本文算法能够实现对海上舰船关键部位精准实时的检测,网络具有更高的准确性和实时性,对复杂海洋环境适应能力更强。

1 舰船关键部位检测算法

1.1 关键部位检测算法

目前常用的关键部位检测算法可分为自上而下(Top-Down)和端到端2种。

Top-Down算法先检测目标信息,再搜索目标内部关键部位信息,主流算法有CPN(corner proposal network)、Hourglass算法、CPM(convolutional pose machibes)算法、HRNet(high resolution network)算法等。

Top-Down算法主要处理步骤为:①通过基于深度学习的目标检测算法如快速区域卷积神经网络(fast region-based convolutional neural networks, Faster R-CNN)、YOLO等检测目标类别信息,并根据目标边界框位置进行裁剪并缩放至指定尺寸;②将第1步获得的区域输入经过特定训练的用于关键部位检测的网络,如HRNet、Stacked Hourglass Network、掩码区域卷积神经网络(mask region-based convolutional neural networks, Mask R-CNN)关键点分支等进行处理,生成热力图,每张热力图对应一

个关键部位中心点, 热力图中数值最高的位置表示最有可能是关键部位中心点的位置, 再通过对热力图解码获取中心点具体坐标值; ③对获取的坐标做非极大值抑制(non-maximum suppression, NMS)等后续处理, 降低冗余, 提升检测准确度, 也可能采取反向传播等矫正操作, 以调整关键部位中心点位置。

Top-Down 算法利用目标检测算法获取目标位置信息, 简化了关键要害部位检测的后续流程, 并针对多目标情况能够进行单独处理, 该处理方式能更加准确地获取关键部位中心点位置信息, Top-Down 算法流程如图 1 所示。但该算法计算复杂度较高, 训练时间长, 检测速度较慢, 不适合对实时性要求较高的场景; 且该算法非常依赖目标检测结果的准确性, 易产生较大误差。

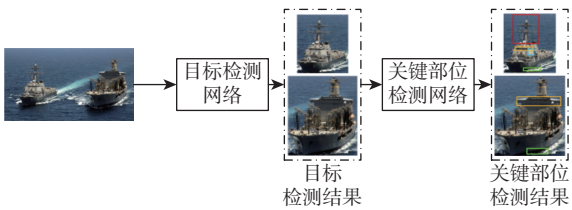


图 1　自上而下的关键部位检测算法

Fig. 1　Top-down key parts detection algorithm

端到端的算法利用一个统一网络框架对目标及关键部位进行检测, 并在图中用框展示出对应区域, 如图 2 所示。这类算法的网络结构较为简洁, 但求解过程中容易忽视目标和关键部位之间的联系, 在应对特征不显著情况及检测小目标问题中表现不佳。



图 2　端到端的关键部位检测算法

Fig. 2　End-to-end key location detection algorithm

1.2　舰船目标检测算法

舰船目标及要害部位检测算法可分为传统的基于人工设计特征的算法和基于深度学习的算法。传统的海上舰船目标检测算法众多, 使用较为广泛的有针对 SAR 图像的恒虚警率法[7]、基于自适应 Canny 算子对舰船红外图像进行边缘检测的算法[8]、对普通光学舰船目标利用小波变化进行检测的算法等。传统的海上舰船目标检测算法图像获取难度大, 大多需要人为设定目标特征, 工作量大, 设置特征也较为困难, 且受多种干扰因素影响, 检测精度不高。

2006 年, Hinton 等在《科学》杂志提出“多隐层神经网络”使卷积神经网络(convolutional neural network, CNN)崛起[9]。深度学习诞生后, 以其优越特性使目标检测技术由传统的机器学习迅速向基于深度学习的方式转变, 并广泛应用于军事目标侦查监控、智能安防系统、舰船目标检测[10]等领域。在检测海上舰船目标方面, 相比传统算法, 深度 CNN 有效利用大量数据, 通过成熟训练, 能够自动提取目标信息, 得到泛化性高的目标特征, 能够较好地应对形状变化、背景、光照、遮挡等影响, 鲁棒性强。基于深度学习的目标检测算法分为一阶目标检测算法和二阶目标检测算法。二阶目标检测算法将目标检测问题划分为 2 个步骤: ①针对输入图像生成候选区; ②对候选区作分类及位置回归。代表算法有 R-CNN、Fast R-CNN、Faster R-CNN 等。一阶目标检测算法把检测流程简化为端至端的回归, 无需候选区, 仅在卷积网络中完成目标特征提取、分类和回归工作, 极大提升了网络检测速度, 更适用于实时应用场景。典型算法有单次多框检测器(single shot multibox detetor, SSD)、YOLO 等。

1.3　研究难点问题

目前, 关键部位检测算法在实现对舰船关键部位的精准检测方面普遍存在以下几点问题:

1) 处理关键部位可变形能力不足。海上舰船分布密集, 舰船目标类型众多, 不同类型舰船的关键部位数量及分布存在差异, 且在舰船目标运动过程中, 受位移、姿态变换影响及图像采集设备性能、运动等因素干扰, 获取到的图像信息中关键部位易出现几何变化, 现有的检测算法对该种差异难以作出灵活调整。

2) 模型复杂度与效率的平衡关系难以协调。在网络提升目标检测精度的同时, 往往伴随着模型复杂度和计算量的增加, 难以适用于实时监测环境。

3) 对遮挡重叠目标及小目标的检测精度不高。对于目标存在遮挡、重叠情况及小目标检测方面性能不足, 网络敏感度不高。舰船的关键部位属于小目标物体, 只占舰船整体的较小部分, 尤其在存在遮挡重叠情况下, 网络能够获取的特征信息有限, 且不同要害部位之间也会对检测结果产生干扰。在复杂海洋环境中, 背景信息与目标信息的分布会存在不平衡性, 这些都对舰船关键部位定位产生干扰。

4) 深浅层特征信息融合欠佳。CNN 通过下采样和池化操作提取深浅层特征信息, 浅层特征包含丰富的目标位置及细节信息而缺少高层语义信息, 深层特征包含更多抽象信息和整体信息, 是检测关键部位特征的关键。现有算法对深浅层特征信息融合欠佳, 需要更好地融合不同层的特征信息, 以

提升网络检测效果。

针对当前算法适用场景、存在问题及提升对舰船关键部位精准检测的需要，本文以最新的YOLOv11目标检测算法为基础，提出针对复杂海洋环境中舰船关键部位检测算法，提升网络检测性能。

2 本文改进方案

为实现对舰船目标关键部位检测，结合海域监控实时性需求，选用YOLOv11算法进行改进。YOLOv11算法通过大幅减少参数极大提升了网络效率及运行速度，对比YOLOv10算法减少了2%的推理时间，成为实时任务场景的最佳选择。在COCO数据集中，YOLOv11算法具有更高的平均精度均值(mean average precision, mAP)，对比YOLOv8算法的参数量减少22%，网络更为轻量，且很好地平衡了速度与精度的关系，YOLOv11算法的总体设计使其具备捕捉复杂图像信息的优势，在困难环境(光线不足或场景杂乱)中的目标检测能力更强。YOLOv11网络能够在手机、云服务器或物联网设备等边缘设备高效部署运行，具有极强的灵活性，能够适用不同任务场景及工作环境。

针对复杂环境中动态舰船目标关键要害部位检测的重难点问题，为进一步提升检测的准确性、高效性及对复杂环境的适应能力，对YOLOv11算法作出如下改进。

2.1 C3k2_DCNv4结构

在多类型关键要害部位目标的舰船检测中，检测目标大多是由不同角度拍摄成像的非规则的形状，而普通卷积的卷积核通常是规整的方形结构，难以精准覆盖目标区域。为进一步提升对多尺度形态舰船要害部位检测精度，引用DCNv4[11]。使用DCNv4卷积重构的新模块替换YOLOv11主干网络中C3k2原来的Bottleneck块，以更准确地估计动态舰船目标的空间形变，提升网络对由于运动造成的舰船关键部位几何变化的适应能力。

YOLOv11中应用Bottleneck模块[12]由1×1及3×3的卷积层组成，完成对输入数据的特征提取和降维操作，模块结构如图3所示。

Bottleneck结构使用1×1卷积层降低了特征图维度，但卷积层数量增加，经过多个卷积层和变换维度操作，有可能会出现总体计算量增加的情况，特别在资源受限的任务场景中会提升训练和部署模型的难度；Bottleneck结构使用1×1卷积层执行降维、升维操作，在反向传播时，梯度信息通过这些卷积层会产生较多损失，影响训练模型的效果。

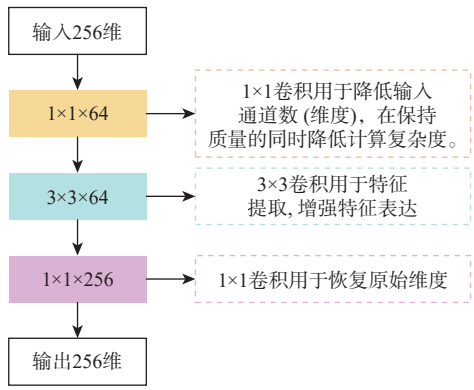


图3 Bottleneck模块

Fig. 3 Bottleneck module

2024年提出的DCNv4是一种高效动态稀疏操作算子，在提取特征图时，与前期版本(如DCNv3、DCNv2、DCN)相比，提供了更广泛的特征级采样。DCNv4集成了一种调制机制，该机制不仅可以学习样本的偏移量，还可以对学习到的特征幅度进行调制，这使DCNv4能够控制其样本空间和相对影响[13]。

DCNv4能够显著提升网络收敛及处理速度，在图像分类、语义和实例分割、图像生成等任务中性能优势显著。DCNv4主要优势有：①摒弃空间聚合中的softmax归一化，DCNv4使用动态自适应窗口，能够根据输入特征计算权重，取消权重归一化的限制，增强了表达能力及动态特性；②改进内存访问，使用共享的聚合权重和采样偏移，使单线程能同步处理同组中的多个通道，避免了非必要的读取内存和运算等操作，加快了网络收敛及处理速度；③DCNv4引入向量化加载，能够一次读取多个通道数据，降低访问内存的指令数，进一步提高了访问内存的效率。

由图4可知，注意力机制与DCNv3通过有限的动态权值(0,1)聚合空间特征；注意力机制的采样点集即窗口是相同的，而DCNv3针对每个位置应用专用窗口。卷积的聚合权值无界(−∞,+∞)，范围更加灵活，对每个位置应用专业滑动窗口，但窗口形状及聚合权值与输入无关。DCNv4结合了卷积与DCNv3的优点，使用自适应的聚合窗口和无界值范围的动态聚合权值。DCNv4中通过一个线程处理同一组内的多个通道，这些通道共享聚合权重和采样偏移量，该方式能够减少读取内存及计算双线性插值系数等工作，且能够合并多个内存访问指令。图4中使用星形符号表示查询像素；使用离散小方框标识参与聚合计算的响应像素采样位置，每个响应像素会分配对应的有界及无限聚合权重，权值大小通过不同颜色表示。

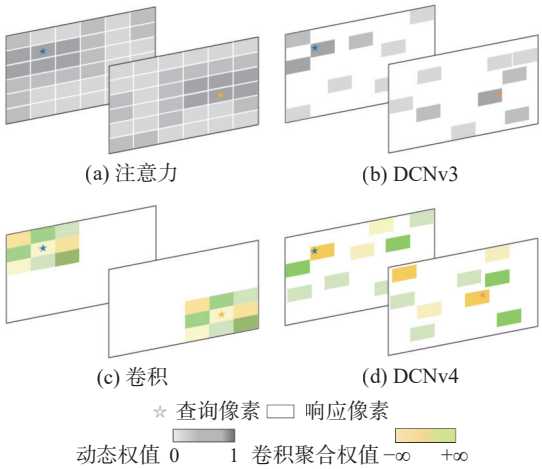


图 4　网络性能对比

Fig. 4　Comparison of network performance

如表 1 所示, 配备了 DCNv4 的 FlashInternImage Backbone 在提升速率的同时, 最终性能也优于原始版本 InterImage Backbone。

表 1 中, AP(S)、AP(M)、AP(L) 分别表示不同尺寸包围框对应的检测精度。

通过引入 DCNv4 改进 YOLOv11 模型的主干网络, 可以自适应地改变卷积核的形状, 尽可能地覆盖舰船关键区域, 从而获得更丰富的特征信息, 实现更精准的舰船关键部位的特征提取。

2.2　多尺度特征金字塔网络

对舰船键关键部位属于小尺寸目标, 检测难度大的问题, 需要增强模型的颈部网络以丰富特征信息的融合, 更有助于定位舰船关键部位的目标区域。因此, 本实验将 AKConv 引入 VoVGSCSP 模块中, 用 AKConv 作为双分支卷积 Concat 操作后得到新型卷积, 替换普通卷积 Conv, 构建新的模块 AVB, 将颈部网络的前两层 C3k2 模块替换, 以进一步提升模型精度。

表 1　DCNv4 性能对比[13]

Table 1　Performance comparison of DCNv4[13]

模型	参数量	帧率/(帧·s^{-1})		AP/%	AP(S)/%	AP(M)/%	AP(L)/%
		FP32	FP16				
Swin-T	48.2×10^6	37	50	51.1	34.3	53.9	66.1
ConvNext-T	48.5×10^6	39	54	51.0	33.5	53.9	65.8
InternImage-T	48.7×10^6	33	37	53.9	37.7	57.7	68.7
FlashInterImage-T	48.5×10^6	39	48	57.7	37.9	58.6	69.8

2.2.1　BiFPN

BiFPN[14] 为特征金字塔网络(feature pyramid network, FPN)的拓展, 在金字塔层级内增加双向互联, 使网络中的信息能同时从上至下和从下至上传输, 网络结构对比如图 5 所示, P_n 表示金字塔特征层下采样倍率。

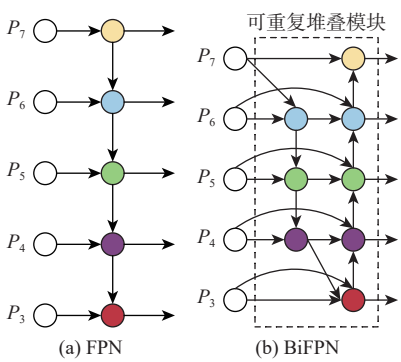


图 5　网络结构对比

Fig. 5　Structure comparison of network

BiFPN 网络以 ResNet 等 CNN 为骨干网络, 从其中多层提取目标特征组成特征金字塔, 在金字塔每层级间应用双向连接, 加强了低层高层间特征流通, 同时双向连接能从上下方向整合金字塔中各层信息以有效获取多尺度特征信息。BiFPN 还应用了特征加权融合方式, 融合不同层间特征, 能在训练中得到权重, 以保证最优的整合特征。

海上舰船目标广泛分布, 类型多样, 获取图像中包含多种尺寸、不同形态的舰船目标, BiFPN 的作用模式能更好获取多尺度的特征表达, 有效增强了网络处理不同复杂度和尺寸目标的能力, 对目标大小存在显著差异的场景十分重要。

2.2.2　AKConv

AKConv[15] 可改变核卷积通过不规则卷积核函数提取特征, 能够为多种变化的目标提供拥有任意采样尺寸和形状的卷积核, 填补了传统卷积的不足。在 COCO2017 及 VisDrone-DET2021 数据集上进行目标检测实验及对比实验, AKConv 显现出更优秀的性能, 具体如下。

1) 样本形状灵活。对比常规卷积固定的样本形状, AKConv 允许卷积有不同大小和形状, 能够灵活适应形变目标, 极大提升了网络的检测准确度。

2) 参数量及运算效率。AKConv 通过改变卷积尺寸能够降低参数量及运算量, 从而提升模型运

算效率及性能。

3）优化网络性能。作为轻量级的模型，AKConv 能够代替常规卷积以降低模型的参数量及计算开销，且能够保持模型性能，同时提升网络泛化能力及效率。

4）可扩展性。AKConv 能够扩展为更大尺寸的内核，为改进网络性能提供更多选项，能够适应多种场景需求，具备广泛的适用性。

图 6 为尺寸为 5 的卷积初始样本，AKConv 能够设计不同的初始采样形状实现任意形状的采样。

卷积核内采样点数量不同的 AKConv 在 COCO-2017、VOC 7+12 及 Vis-Drone-DET2021 数据集中的表现如表 2 所示，mAP 为平均精度，mAP50 表示交并比为 0.5 时的 mAP。结果充分体现了 AKConv 的优势。

AKConv 结构如图 7 所示，其中，C 为通道数，H 为特征图高度，W 为特征图宽度，$2N$ 为二维采样点总数。

受舰船目标航行及图像获取设备运动影响，会造成获取目标的几何变化而影响网络检测效果，通过引入 AKConv，为动态舰船目标检测提供多种尺寸和形状的卷积核，提升对动态变化场景的适应能力，增强检测准确性。

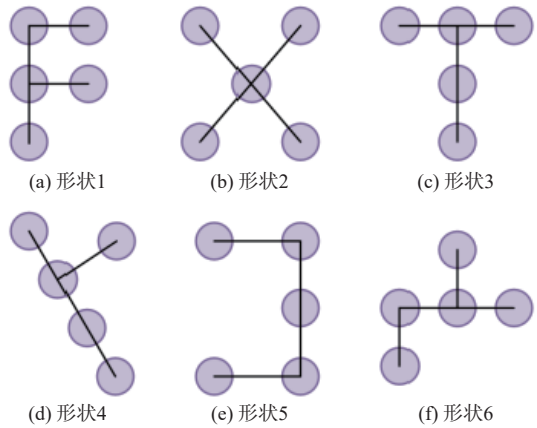


图 6 采样形状

Fig. 6 Sample shape

2.2.3 VoVGSCSP 模块

经过卷积操作获取的图像低层信息包含目标更丰富的细节特征，而高层信息则含有更多的语义信息及空间特征，将不同层级信息有效融合，实现跨层连接互通，能够进一步提升对海上多类型动态舰船目标的检测精度。

VoVGSCSP 模块[16]将输入特征映射至体素化网络中来获取更多空间中的局部特征信息，并引用跨阶段局部网络（cross stage partial network,

表 2 引入不同数量 AKConv 的网络性能对比[15]

Table 2 Performance comparison of networks with different numbers of AKConv introduced[15]

模型	AKConv卷积数量	精确率/%	召回率/%	mAP50/%	mAP/%	浮点运算量/次	参数量
基线模型		77.3	69.8	76.4	50.2	13.2×10^9	6.06×10^6
YOLOv7-tiny	3	80.1	68.4	76.1	50.3	12.4×10^9	5.56×10^6
	4	78.2	70.3	76.2	50.7	12.4×10^9	5.66×10^6
	5	77.0	71.1	76.5	50.8	12.6×10^9	5.75×10^6
	6	79.6	69.9	76.9	51.0	12.9×10^9	5.85×10^6
	8	78.6	70.1	76.7	51.2	13.4×10^9	6.04×10^6
	9	81.0	69.3	76.7	51.3	13.7×10^9	6.14×10^6

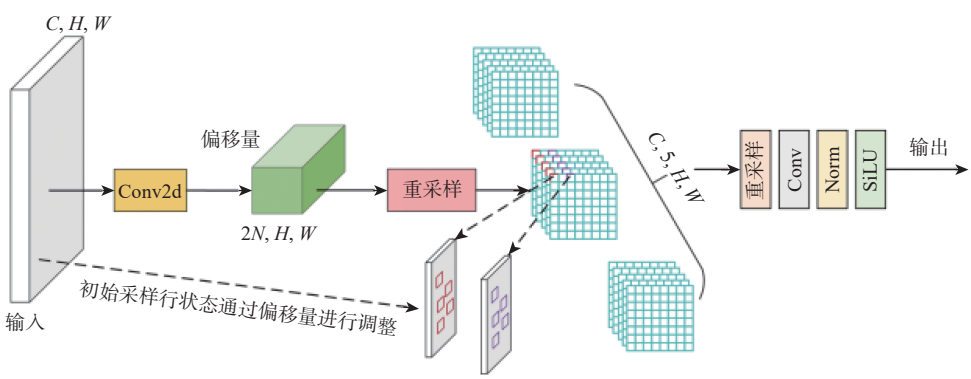


图 7 AKConv 结构

Fig. 7 Structure of AKConv

CSPNet)思想, 利用跨阶段局部连接的方法融合和分离特征以维持特征表达的丰富性并降低计算冗余, 提高网络效率。VoVGSCSP 能高效融合多个不同层次的特征信息, 提高特征表达能力, 更好地平衡模型网络性能及运算效率间的平衡, 更适用于实时场景。模块结构如图 8 所示。

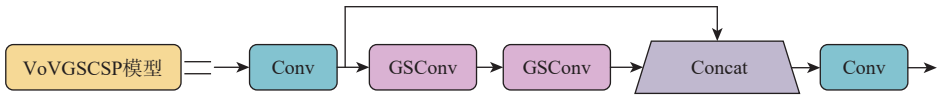


图 8　VoVGSCSP 结构

Fig. 8　Structure of VoVGSCSP

使用 AKConv 替换 VoVGSCSP 模块的普通卷积, 构建新的模块 AVB 模块, 并使用 AVB 模块替换颈部网络的前 2 层 C3k2 模块, 能够在提升对动态多类型舰船目标检测精度的同时, 维持较好的运算速率, 符合实时性及精准检测需求。

2.3　GSConv 结构

针对边缘设备资源普遍有限且检测算法适用性低的问题, 在 Neck 部分引入了轻量级 GSConv 模块[17]。该模块采用了分组卷积和空间卷积相结合的优化策略, 以减少计算量和内存损失。

为加快 CNN 的运算过程, 输入 CNN 的图像基本都要在骨干网络中进行此类转换: 空间信息向通道逐步传递。但每次对特征图作空间(宽度、高度)压缩及通道扩展时都会产生部分语义信息的丢失。通道密集卷积(channel dense convolution, SC)能够最大程度保留每个通道的隐藏连接, 而通道稀疏卷积(channel sparse convolution, DSC)彻底断开了这些连接。GSConv 则以更低的时间复杂度, 最大可能地保存这些连接。卷积计算的时间复杂度由浮点运算量定义。SC、DSC、GSConv 在不考虑偏置情况下的时间复杂度 T 为

$$T_{\mathrm{SC}} \sim O(W \times H \times K_1 \times K_2 \times C_1 \times C_2) \tag{1}$$

$$T_{\mathrm{DSC}} \sim O(W \times H \times K_1 \times K_2 \times 1 \times C_2) \tag{2}$$

$$T_{\mathrm{GSConv}} \sim O\left(W \times H \times K_1 \times K_2 \times \frac{C_2}{2} \times (C_1 + 1)\right) \tag{3}$$

式中: W 和 H 分别为输出特征图宽度和高度; K_1、K_2 为卷积核大小; C_1、C_2 分别为输入及输出特征图的通道数。

在 PASCAL VOC 2007+12 数据集中对比 SC、DSC、ShuffleNet、GhostNet 及 GSConv 卷积模块对模型性能的作用意义如表 3 所示, 加粗显示最优结果。

表 3　卷积模块对比[17]

Table 3　Comparison of convolution modules[17]

卷积模块	mAP50/%	mAP95/%	浮点运算量/次
SC	58.9	35.8	4.2×10^9
DSC	56.9	34.6	$\mathbf{3.3\times10^9}$
ShuffleNet	56.5	34.1	$\mathbf{3.3\times10^9}$
GhostNet	59.3	37.0	3.8×10^9
GSConv	**59.8**	**38.1**	3.7×10^9

GSConv 还可以根据实际硬件性能动态分配计算资源, 以平衡计算成本和检测性能, 不仅可以实现更高的检测精度, 而且可以降低功耗和能耗。特别是在小尺寸特征图方面, GSConv 卷积产生的计算负担比标准卷积低, 这在一定程度上提高了对小目标的检测精度。

GSConv 结构如图 9 所示。结合本节改进措施, 对 YOLOv11 网络结构进行改进, 网络总体结构模型如图 10 所示。

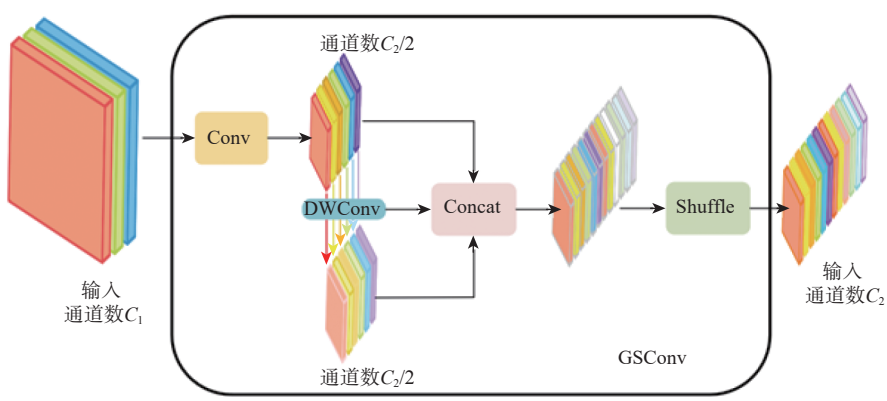


图 9　GSConv 结构

Fig. 9　Structure of GSConv

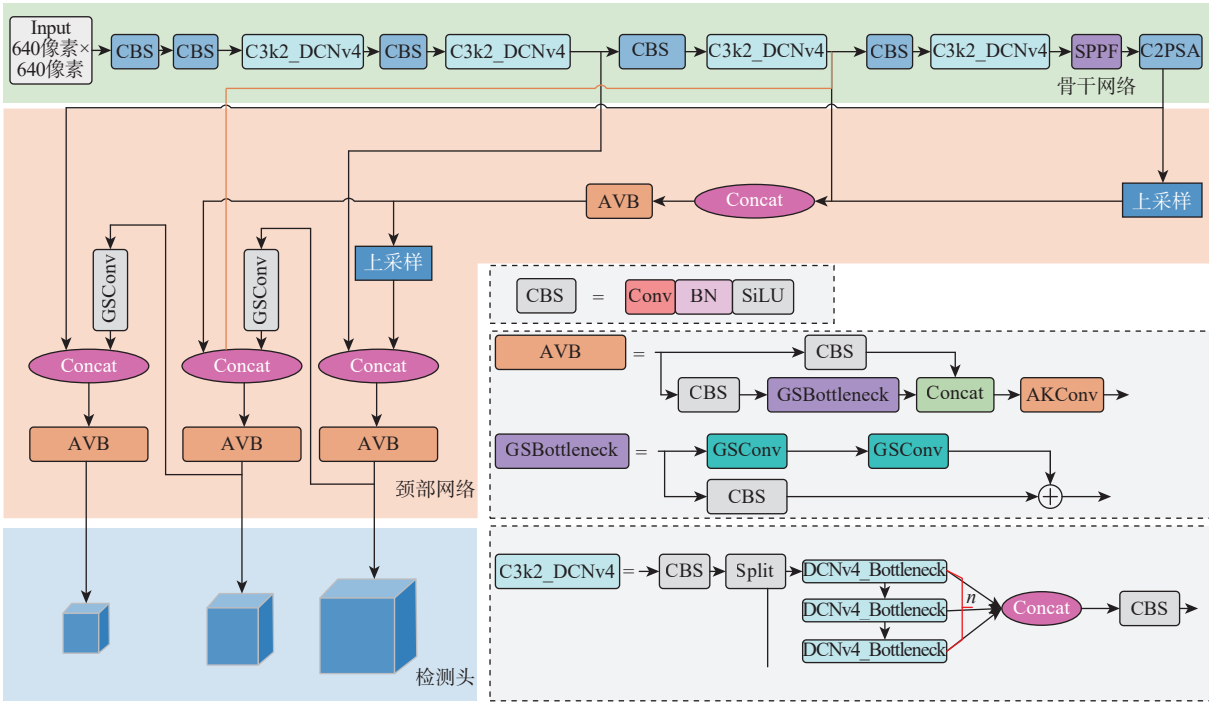


图 10 本文改进的 YOLOv11 网络结构

Fig. 10 Network structure of the improved YOLOv11

3 数据集构建

3.1 数据集组成

目前针对舰船目标的数据集多为 SAR 图像和红外图像，且主要针对舰船主体，尚没有公开的针对舰船关键部位可见光数据集。为贴合实际海洋环境及图像采集设备低空、超低空飞行视角，探索利用本文算法实现对舰船及其关键部位的精准检测识别，通过网络收集、已有数据集节选、实景拍摄等方式选取包含可见光舰船正视及侧视图像，数据集中包含多种类型舰船，存在遮挡、小目标场景，共 5 589 张图像。将舰船关键部位分为驾驶舱、天线桅杆、雷达、水线及卫通。包含关键部位图像数量如表 4 所示。

表 4 数据集结构

Table 4 Structure of dataset

要害部位	数量/张
驾驶舱	6 945
天线桅杆	7 197
雷达	3 237
水线	8 158
卫通	4 137

数据集中部分图像如图 11 所示。数据集采用 PASCAL VOC 格式，使用 labelImg 工具标注。在标注重要目标时，由于检测框是矩形的，难免会框选到无关信息，尤其是标注水线检测框时，过于狭长的结构会增加检测难度。因此，在框选水线时，特



图 11 数据集部分图像

Fig. 11 Dataset partial images

别注意框定接近水平且受环境噪声干扰较弱的位置，在减少标注框过于狭长问题的同时，避免了大面积框选舰船主体信息；对于驾驶舱、天线桅杆、船首等关键部位以目标边缘为基准来标注。

图像标注如图 12 所示。图像格式为.jpg，为加速网络训练过程，将图像像素大小统一为 640 像素×640 像素，命名为 000001 格式，并对图像进行标注。将数据集中 80% 图像划分为训练集，20% 划分为测试集。

3.2 实验条件及评价指标

3.2.1 实验条件

实验在 Ubuntu18.04 操作系统中完成，使用 Pytorch 1.10 深度学习框架，cuda11.1+cudnn8.0.4+opencv4.5.8 环境；硬件环境为：Intel(R) Xeon(R) Platinum i9-13900k, Nvidia GeForce RTX 4 090；使用 python 语言实现编程操作。

输入图像像素大小为 640 像素×640 像素，初始学习率为 0.01，采用随机梯度下降更新学习率，

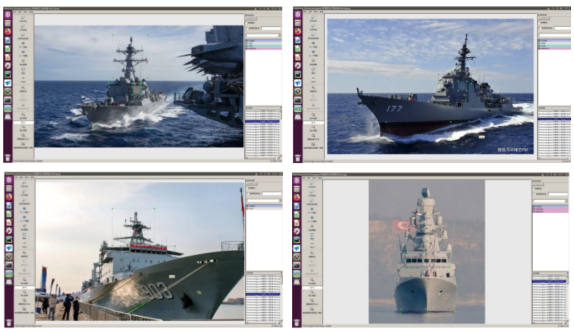


图 12　数据集标注

Fig. 12　Dataset annotation

动量为 0.937, 权重延迟为 0.000 5。在训练过程中使用 Mosaic 数据增强的方式, 每次读取 4 张图像, 分别进行翻转、缩放等操作后拼接, 使检测背景更加丰富。设置标签平滑为 0.01, 防止模型过拟合, 增加模型的泛化性。所有的模型皆训练 500 个 epoch, batchsize 设置为 32, worker 线程数 16。

3.2.2　评价指标

TP(true positive): 准确判定为正样本的样本, 记为 T_P; TN(true negatives): 准确判定为负样本的样本, 记为 T_N; FP(false positive): 负样本判定为正样本, 记为 F_P; FN(false negative): 正样本被判定为负样本, 记为 F_N; Q 为查询的总次数。

1) 准确率 P 为全部被检测是正的样本真正属于该类的比, 表达式为

$$P=\frac{T_P}{T_P+F_P} \tag{4}$$

2) 召回率 R 为目标被检测为正类的数量占所有被检测为目标类总数的比例, 表达式为

$$R=\frac{T_P}{T_P+F_N} \tag{5}$$

3) A_P 为准确率-召回率曲线(PR 曲线)与横轴之间的图形面积, 曲线横轴为 R 值, 表示分类器涵盖正样本的能力, 纵轴为 P 值, 表示对正类的预测精度, PR 曲线表示了两值间的变化情况, 分为连续的 PR 曲线和离散的 PR 曲线 2 种, 表达式为

$$A_P=\begin{cases}\int_0^1 PR\mathrm{d}r\\ \sum_{k=1}^{n}P(k)\Delta r(k)\end{cases} \tag{6}$$

式中: $P(k)$为第 k 个采样点对应的精确率; $\Delta r(k)$为相邻 2 个采样点之间召回率的变化量。

4) mAP 为所有查询结果排序的 AP 的平均值, 即预测正确的部分占预测结果的比例。表达式为

$$m_{AP}=\sum_{q=1}^{Q}A_P(q)/Q \tag{7}$$

5) 检测时间, 单位 ms。数值越小则检测单张图像时间越短, 算法速度越快。

3.3　实验结果分析

使用本文自建可见光舰船要害部位数据集训练优化的基于 YOLOv11 的舰船要害部位检测算法, 训练集与测试集比例为 8:2。训练过程中使用马赛克数据增强方式, 在网络收敛后, 得到成熟的网络模型。

本文算法损失函数曲线如图 13 所示, loss 函数为无量纲指标, 数值大小表征模型预测结果与真实标签的偏差程度。可以得出, 网络经过充分的迭代训练, 损失值平缓下降, 最终趋于稳定至收敛, 最后收敛的损失函数值较小, 没有过拟合、欠拟合问题; 在损失函数下降过程中, 没有出现大幅度的波动情况, 证明本文算法网络模型设计合理, 规模深度及数据集构建完善, 使网络得到充分训练和平稳收敛。

本文算法 PR 曲线如图 14 所示。不同类型目标检测精度如表 5 所示。

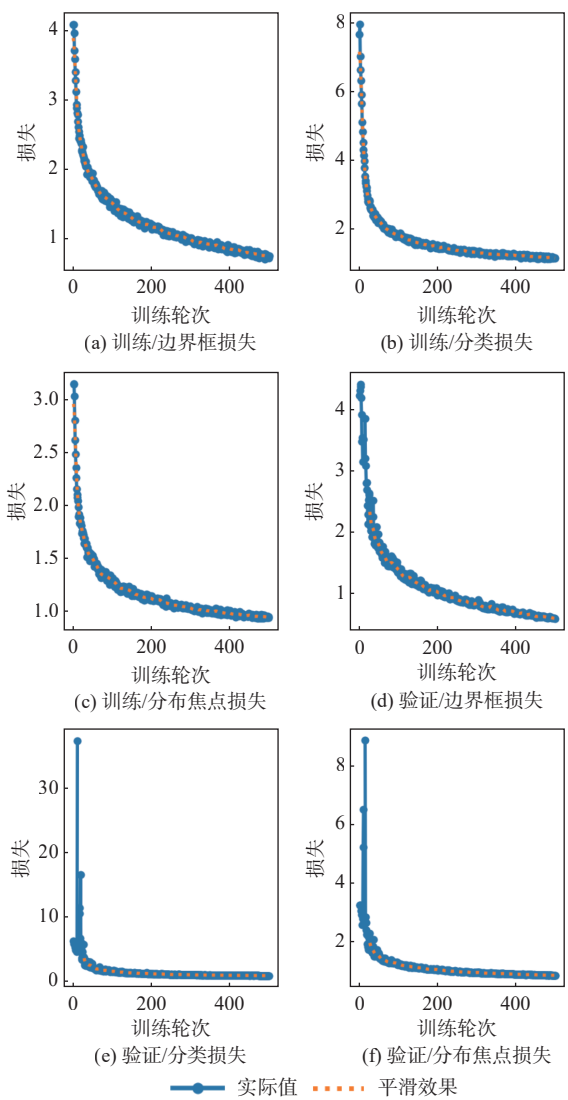


图 13　本文算法损失函数曲线

Fig. 13　Proposed algorithm loss function curves

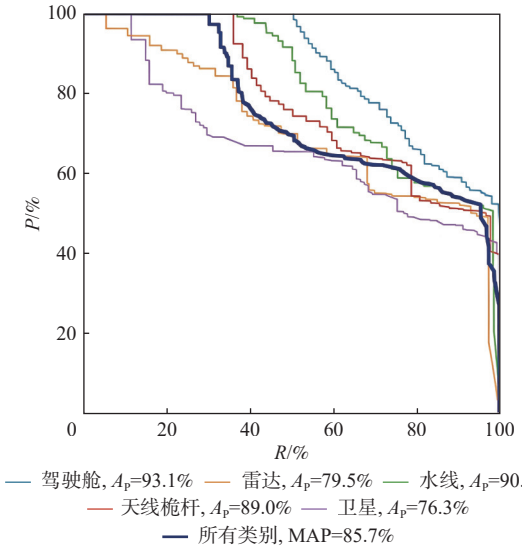


图 14 本文算法 *PR* 曲线

Fig. 14 Proposed algorithm *PR* curves

表 5 要害部位检测精度

Table 5 Detection accuracy of critical parts

类型	AP/%
驾驶舱	93.1
雷达	79.5
水线	90.5
天线桅杆	89.0
卫通	76.3

由各类要害部位检测结果统计知，基于 YOLOv11 改进的本文要害部位检测算法对多类型舰船要害部位检测精度达到 85.7%，检测精度高，检测时间为 1.1 ms，能够满足作战实时性要求。驾驶舱检测精度达到 93.1%，检测精度最高，主要由于驾驶舱特征较为突出且相对其他部位像素占比大，易于提取特征；水线检测精度达到 90.5%，在标注水线时，减小了矩形包围框大小，尽量只框选水线部分而减少背景因素的选取，提升了网络训练的鲁棒性，对水线的检测准确度高。本文算法网络对雷达、卫通的检测精度分别为 79.0% 和 76.3%，体现了该网络能够较为准确地获取像素占比较少的关键部位信息。本文算法网络的检测效果如图 15 所示，目标关键部位使用不同颜色框选展示。

由图 15 知，本文算法能够有效检测出图像中舰船目标的关键要害部位，对于多角度舰船目标均表现出较好的检测性能，检测精准度高，标注框位置准确，对像素占比较小的关键要害部位仍能做到有效检测提取目标信息；特别是在面对恶劣海况、舰船关键部位受到遮挡干扰的情况下，仍能准确检测得到关键部位信息，基本无漏检情况。

为进一步验证本文算法的性能优势，并应对在



图 15 本文算法检测效果

Fig. 15 The proposed algorithm detection effect

反舰设备实际作战环境中对要害部位检测精度、速度的要求及硬件设备容量的限制，在相同实验环境及参数设置下分别对 SSD 算法[18]、CenterNet-Hourglass 算法[19]、YOLOv8n 算法[20]、YOLOv10n 算法[21]、YOLOv11n 算法[22]进行训练，迭代次数为 500 epoch，输入图像分辨率不变，使用相同的训练方式，对网络性能进行测试，统计网络基本性能参数如表 6 所示，最优结果加粗显示。

本文提出的基于 YOLOv11 的关键部位检测算法检测精度为 85.7%，对比 SSD 算法、CenterNet-Hourglass 算法、YOLOv8n 算法、YOLOv10n 算法、YOLOv11n 算法分别提升 11.4%、8.1%、6.0%、7.6%、4.9%。本文算法检测精度有明显提升；检测单张图像时间为 1.1 ms，明显快于 SSD 算法、CenterNet-Hourglass 算法，对比 YOLO 系列算法中的 YOLOv10n 算法、YOLOv11n 算法也分别提升 0.4 ms、0.3 ms，网络实时性更强。且本文算法参数量对比其他算法最少，能更好适应硬件平台资源受限的应用场景。

不同算法的检测效果图如图 16 所示。可以看出，本文算法对 5 类关键要害部位的检测漏检率更低，基本能完整检测出图像中存在的要害部位，对比检测精度也明显高于其他对比算法。

3.4 消融实验

本文针对复杂海洋环境中运动舰船目标关键部位检测，对 YOLOv11 算法提出多个模块的改进，为验证每个模块作用价值，设计消融实验验证各改进方案对检测结果的影响。实验环境及实验参数设置与对比实验相同，对比算法性能如表 7 所示，其中，最优结果加粗显示。

由表 7 可知，未进行改进的 YOLOv11 算法检测精度为 80.8%，检测单幅图像的时间为 1.4 ms，应

表 6　关键部位检测对比实验

Table 6　Comparative experiment of key parts detection

算法	精确率/%	召回率/%	mAP50/%	参数量	时间/ms
SSD[18]	60.2	78.9	74.3	164 765 132	3.4
CenterNet-Hourglass[19]	61.9	80.1	77.6	95 432 146	3.3
YOLOv8n[20]	62.7	86.3	79.7	3 006 623	1.0
YOLOv10n[21]	64.2	80.5	78.1	2 696 366	1.5
YOLOv11n[22]	65.3	82.1	80.8	2 583 127	1.4
本文算法	**70.6**	**87.7**	**85.7**	**2 437 559**	**1.1**



(a) YOLOv8n[20]



(b) YOLOv10n[21]



(c) 本文算法

图 16　不同检测算法效果图

Fig. 16　Renderings of different detection algorithms

用 C3k2_DCNv4 结构后, 算法平均检测精度提升 3.1%, 检测时间减少 0.4 ms, 这是因为在 C3k2_DCNv4 结构中, 利用基于 DCNv4 重构的新模块替换了 YOLOv11 主干网络 C3k2 中原来的 Bottleneck 块, 能够提升网络对抗运动产生的几何形变干扰情况, 从而增强网络对运动舰船目标空间形变更精准的估计能力, 能有效提高网络适应海上船舶目标运动造成的几何变化的能力, 更准确地估计移动船舶目标空间形变, 且使用 DCNv4 能精简训练过程, 减少模型训练参数, 因此, 应用 C3k2_DCNv4 结构后, 网络的检测精度及检测效率均有显著提升。

应用 Bi-AVB-FPN 后, 算法检测精度进一步提升 1.3%, 检测时间降低 0.1 ms, Bi-AVB-FPN 网络能够有效增强模型的颈部网络, 丰富目标特征信息的融合, 从而提升网络对小目标的检测能力, 更有助于定位舰船关键要害部位的目标区域; Bi-AVB-FPN 网络中, 引用 CSPNet 思想, 利用跨阶段局部连接的方法融合和分离特征, 在维持特征表达的丰富性的同时有效降低了计算冗余, 减少了模型的参数量, 加快了检测速率。

最后在模型的 Neck 部分引入了轻量级 GSConv 模块, 以减少计算量和内存损失, 平衡计算成本和检测性能的关系, 算法参数量得到进一步降低, 同时检测精度提升了 0.5%, 检测时间为 1.1 ms, 增加

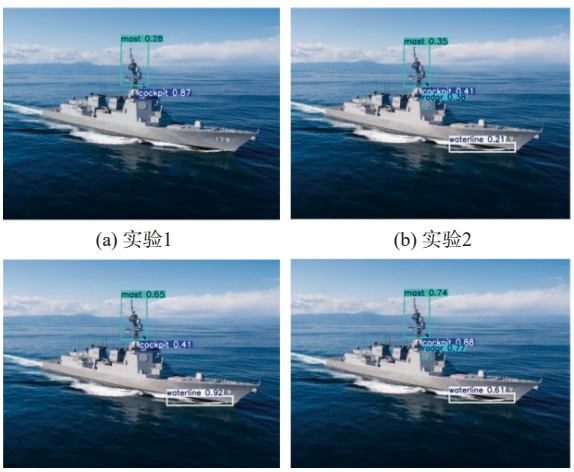
表 7　消融实验结果对比

Table 7　Comparison of ablation experiment results

C3k2_DCNv4	Bi-AVB-FPN	GSConv	精确率/%	召回率/%	mAP50/%	参数量	时间/ms
×	×	×	65.3	82.1	80.8	2 583 127	1.4
√	×	×	68.6	87.4	83.9	2 571 439	1.0
√	√	×	70.3	86.7	85.2	2 527 319	0.9
√	√	√	**70.6**	**87.7**	**85.7**	**2 437 559**	**1.1**

了 0.2 ms，主要是因为对算法进行改进后，在一定程度上增加了模型的复杂性，所以检测时间有所增加，但整体网络模型参数量有效降低，能够更好适应硬件条件受限的场景，算法检测精度对比未改进时提升 4.9%，对关键要害部位小目标检测性能得到提升，能够精准检测多尺度、多角度要害部位信息，且检测时间减少了 0.3 ms，进一步提升了网络检测速率。

在消融实验中，对 YOLOv11 算法不同模块改进的检测效果如图 17 所示。可以看到，经过不断改进的算法检测效果不断提升，检测精准度更高，漏检率更低，最终得到的优化网络能够完整检测出船舶的关键部位，且检测准确度及包围框的尺寸位置都更贴近实际目标特征。



(a) 实验1　(b) 实验2

(c) 实验3　(d) 实验4

图 17　不同模块改进检测效果

Fig. 17　Detection renderings of different improved modules

4　结　论

1）本文以提升对海上舰船关键部位精准实时检测为目的，针对现有算法动态目标尺度变化鲁棒性不强、检测精度及速度难以平衡的问题，提出以 YOLOv11 为基础的舰船关键部位检测算法，自建可见光数据集对网络进行训练，并设计对比实验和消融实验，证明优化的网络模型在复杂海洋环境中对动态舰船目标关键部位精准实时检测的能力。

2）本文算法对比改进前，检测精度由 80.8% 提升至 85.7%，提升 4.9%，检测单幅图像时间由 1.4 ms 减少至 1.1 ms，且算法参数量减少，有效简化了算法模型，为搭载到 NVIDIA Jetson TX2 等嵌入式平台构建任务式计算机，提供了轻量化的便利条件。

下一步将对网络在实际场景中的应用、提升复杂环境适应性开展研究，考虑结合红外图像探索全天时海上信息的获取判读。

参考文献（References）

[1] Gao F, He Y S, Wang J, et al. Anchor-free convolutional network with dense attention feature aggregation for ship detection in SAR images[J]. Remote Sensing, 2020, 12(16): 2619.

[2] 白玉, 迟文恺, 谢宝蓉, 等. 结合目标提取和深度学习的红外舰船检测[J]. 电讯技术, 2023, 63(2): 193-198.
Bai Y, Chi W K, Xie B R, et al. Infrared ship detection by combining target extraction and deep learning[J]. Telecommunication Engineering, 2023, 63(2): 193-198(in Chinese).

[3] Tang G, Liu S B, Fujino I, et al. H-YOLO: a single-shot ship detection approach based on region of interest preselected network[J]. Remote Sensing, 2020, 12(24): 4192.

[4] 刘忻伟, 朴永杰, 郑亮亮, 等. 面向航天光学遥感复杂场景图像的舰船检测[J]. 光学精密工程, 2023, 31(6): 892-904.
Liu X W, Piao Y J, Zheng L L, et al. Ship detection for complex scene images of space optical remote sensing[J]. Optics and Precision Engineering, 2023, 31(6): 892-904(in Chinese).

[5] 林鑫伟, 徐志京, 黄海. 复杂背景下的 SAR 图像多尺度舰船检测[J]. 中国航海, 2023, 46(2): 17-24.
Lin X W, Xu Z J, Huang H. Multi-scale detection of ship target against complex background out of SAR image[J]. Navigation of China, 2023, 46(2): 17-24(in Chinese).

[6] 王玺坤, 姜宏旭, 林珂玉. 基于改进型 YOLO 算法的遥感图像舰船检测[J]. 北京航空航天大学学报, 2020, 46(6): 1184-1191.
Wang X K, Jiang H X, Lin K Y. Remote sensing image ship detection based on modified YOLO algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(6): 1184-1191(in Chinese).

[7] 李双伟. 基于红外偏振成像的近岸舰船目标识别和抗干扰技术研究[D]. 烟台: 烟台大学, 2016: 12-15.
Li S W. Research on inshore ship target recognition and antijamming technology based on infrared polarization imaging[D]. Yantai: Yantai University, 2016: 12-15(in Chinese) .

[8] 乔延军. 面向户外增强现实的地理实体目标识别与检测[D]. 武汉: 武汉大学, 2017: 10-30.
Qiao Y J. Object recognition and detection of geographical entity based on outdoor augmented reality[D]. Wuhan: Wuhan University, 2017: 10-30(in Chinese) .

[9] Hinton G E, Salakhutdinov R P. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.

[10] 苏娟, 杨龙, 黄华, 等. 用于 SAR 图像小目标舰船检测的改进 SSD 算法[J]. 系统工程与电子技术, 2020, 42(5): 1026-1034.
Su J, Yang L, Huang H, et al. Improved SSD algorithm for small-sized SAR ship detection[J]. Systems Engineering and Electronics, 2020, 42(5): 1026-1034(in Chinese).

[11] Ahn S, Chang J W, Kang S J. An efficient accelerator design methodology for deformable convolutional networks[C]//Proceedings of the IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2020: 3075-3079.

[12] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[EB/OL]. (2015-12-10)[2024-11-20]. https://doi.org/10.48550/arXiv.1512.03385.

[13] Xiong Y W, Li Z Q, Chen Y T, et al. Efficient deformable ConvNets: rethinking dynamic and sparse operator for vision applications[C]//Proceedings of the IEEE/CVF Conference on Computer

Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 5652-5661.

[14] Tan M X, Pang R M, Le Q V. EfficientDet: scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 10778-10787.

[15] Zhang X, Song Y Z, Song T T, et al. AKConv: convolutional kernel with arbitrary sampled shapes and arbitrary number of parameters[EB/OL]. (2023-11-20)[2024-11-22]. https://arxiv.org/abs/2311.11587v1.

[16] Li Y K, Shen H B, Fu Y P, et al. A method of dense point cloud SLAM based on improved YOLOV8 and fused with ORB-SLAM3 to cope with dynamic environments[J]. Expert Systems with Applications, 2024, 255: 124918.

[17] Li H L, Li J, Wei H B, et al. Slim-neck by GSConv: a lightweight-design for real-time detector architectures[EB/OL]. (2024-07-02)[2024-11-22]. https://doi.org/10.48550/arXiv.2206.02424.

[18] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C]//Proceedings of the Computer Vision-ECCV 2016. Berlin: Springer, 2016: 21-37.

[19] Newell A, Yang K, Deng J. Stacked hourglass networks for human pose estimation[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016: 486-499.

[20] Yaseen M. What is YOLOv8: an in-depth exploration of the internal features of the next-generation object detector[EB/OL]. (2024-08-28)[2024-11-20]. https://arxiv.org/abs/2408.15857.

[21] Wang A, Chen H, Liu L H, et al. YOLOv10: real-time end-to-end object detection[C]//Proceedings of the Advances in Neural Information Processing Systems 37. Vancouver: NeurIPS, 2024: 107984-108011.

[22] Khanam R, Hussain M. YOLOv11: an overview of the key architectural enhancements[EB/OL]. (2024-11-23)[2024-11-25]. https://arxiv.org/abs/2408.15857.

Detection algorithm of ship key parts based on YOLOv11

WANG Yao，CAO Silei*，LIU Minggang，WANG Yingying

(Naval Aviation University，Yantai 264001，China)

Abstract：In contemporary and forthcoming maritime rights safeguarding and oceanic surveillance endeavors, it is critically important to investigate methods for swiftly and precisely acquiring data on essential components of vessels and adeptly analyzing real-time marine conditions. Despite the capability of deep learning-based object identification algorithms to identify various types of target information, there is a deficiency in research focused on algorithms for recognizing critical components of ships, and the examination of the relationship between multi-channel feature maps is inadequate. The feature utilization rate of key parts of ships with different scales is low, which may lead to the loss of important feature information and a decrease in detection accuracy. To address this issue, a ship key part detection method based on YOLOv11 is proposed to improve the accuracy and efficiency of key part detection. This manuscript initially performs a thorough examination of the critical and challenging issues in the present detection of essential ship components and engages in a comparative analysis with existing methods. Then, an improved YOLOv11 object detection network for ship key part detection is proposed based on actual problems. Ultimately, a dataset of key components for visible light ships has been established, and through the implementation of comparative and ablation experiments, it is thoroughly evidenced that the proposed algorithm developed in this study exhibits high precision and promptness in identifying ship key components, thereby offering viable technical assistance and a theoretical foundation for the meticulous differentiation of essential and critical parts across various ship types in intricate marine settings and the efficient surveillance of maritime vessel information.

Keywords：ship targets; critical parts; YOLOv11; attention mechanism; feature fusion

Received：2024-12-03；**Accepted**：2025-07-17；**Published Online**：2025-09-04 15：16

URL：link.cnki.net/urlid/11.2625.V.20250904.1419.001

Foundation item：Youth-General Program Category(ZR2024QE034)

* **Corresponding author**. E-mail：caosileimail@163.com