



吉林大学学报(工学版)  
*Journal of Jilin University(Engineering and Technology Edition)*  
ISSN 1671-5497,CN 22-1341/T

## 《吉林大学学报(工学版)》网络首发论文

题目: 基于注意力机制改进的无锚框 SAR 图像舰船检测模型  
作者: 高云龙, 任明, 吴川, 高文  
DOI: 10.13229/j.cnki.jdxbgxb20221367  
收稿日期: 2022-10-25  
网络首发日期: 2022-12-16  
引用格式: 高云龙, 任明, 吴川, 高文. 基于注意力机制改进的无锚框 SAR 图像舰船检测模型[J/OL]. 吉林大学学报(工学版).  
<https://doi.org/10.13229/j.cnki.jdxbgxb20221367>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

# 基于注意力机制改进的无锚框 SAR 图像舰船检测模型

高云龙, 任明, 吴川, 高文

(中国科学院 长春光学精密机械与物理研究所, 长春 130033)

**摘要:** 为提升模型对 SAR 图像多尺度舰船目标的检测能力, 保证检测网络的实时性, 提出基于注意力机制改进的无锚框 SAR 图像舰船检测模型。在 YOLOX 网络特征金字塔处嵌入空洞注意力模块, 调节感受野与多尺度融合的关系, 强化特征的表示能力。在检测头部设计中心性预测分支, 对锚点的分类得分进行加权处理, 调整模型的损失函数, 优化检测结果。在数据集 SSDD 上进行对比实验, 本文提出的模型优于主流的深度网络检测模型, 精度达到 94.73%, 且在检测精度和检测速度中取得最佳平衡。

**关键词:** 计算机视觉; 舰船目标检测; 空洞卷积; 注意力机制; 无锚框

**中图分类号:** TP391 **文献标志码:** A

**DOI:** 10.13229/j.cnki.jdxbgxb20221367

## An improved anchor-free model based on attention mechanism for ship detection in SAR images

GAO Yun-long, REN Ming, WU Chuan, GAO Wen

(Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Science, Changchun 130033, China)

**Abstract:** In order to improve the detection capability of detectors for multiscale ships in SAR images and ensure the real-time performance of the detection networks, an improved anchor-free model based on attention mechanism for ship detection in SAR images is proposed. On the basic framework of the off-the-shelf YOLOX, a lightweight dilated convolutional attention module (DCAM) is embedded in front of feature pyramid network (FPN) to adjust the relationship between receptive field and multiscale fusion, and strengthen the representation ability of features. The detection head is redesigned by introducing the center-ness prediction branch, which can weight the classification scores of the anchor points, in the meantime, the loss function of the proposed model is also revised to optimize the final detection performance. Through the comparative experiments on dataset SSDD, the proposed model in this paper is superior to the mainstream deep learning detection models, with an accuracy of 94.73%, and achieves the best trade-off between detection accuracy and detection speed.

**Key words:** computer vision; ship detection; dilated convolution; attention mechanism; anchor-free

## 0 引言

合成孔径雷达<sup>[1]</sup> (SAR) 是一种主动式微波遥

感技术传感器, 相比于光学传感器, 具有对目标实施全天候、全天候、多角度、远距离监测能力, 在海洋救援、海洋执法等民用领域, 以及海洋实时监

**收稿日期:** 2022-10-25.

**基金项目:** 国家自然科学基金项目 (61401425); 吉林省科技发展计划重点研发项目 (2022021146GX).

**作者简介:** 高云龙 (1993-), 男, 助理研究员, 博士. 研究方向: 目标检测, 深度学习. E-mail: gaoyl15@mails.jlu.edu.cn

**通信作者:** 高文 (1985-), 女, 副研究员, 博士, 硕士生导师. 研究方向: 目标检测, 深度学习. E-mail: gaowen@ciomp.ac.cn

测检测等军事领域都有广泛的应用空间<sup>[2]-[6]</sup>。随着合成孔径雷达高分辨率成像技术的不断发展,大量 SAR 图像可用于海上舰船检测, SAR 舰船目标检测成为当前研究热点,在对海域实时检测中发挥着重要作用。因此,研究 SAR 舰船目标检测算法具有重要意义。

传统的舰船检测方法主要依赖于统计图像像素分析,大多数是基于阈值的方法。其中恒虚警率检测算法(CFAR)<sup>[7]</sup>,已成为 SAR 图像目标检测的经典方法,广泛应用于实际的舰船目标检测系统。CFAR 通过对背景杂波进行统计建模,确定区分舰船目标与背景的阈值。基于阈值的算法通常具有较高的拟合度精度,但往往需要大量的计算资源估计阈值参数,实时性较差,因此近年来基于阈值的方法研究主要集中在关注海杂波建模精度以及计算复杂性之间的权衡。此外,不同形状、方向的舰船目标和复杂场景也限制了统一检测模型的构建,因此,它们的泛化能力较弱,检测性能也远不能令人满意。

近年来,深度学习算法在 SAR 图像舰船目标检测任务中得到了广泛的应用,神经网络能够在大量标注数据集驱动下自动学习适用于目标检测的特征,取得了优于传统算法的检测性能。基于深度网络的检测算法主要分为基于锚框(anchor-based)和无锚框(anchor-free)的算法。基于锚框的算法包括为双阶段算法和单阶段算法。双阶段算法如 Faster-RCNN<sup>[8]</sup>具有较高的检测精度,但计算量大实时性低;单阶段算法如 RetinaNet<sup>[9]</sup>,YOLOv3<sup>[10]</sup>等通过提取的特征直接回归目标的坐标信息,提高了检测效率。Kang 等<sup>[11]</sup>将 CFAR 算法和 Faster-RCNN 算法结合,提出一种新的双阶段检测网络,利用 CFAR 算法生成的滑动窗口作为候选区域检测小型舰船目标,以获得更好的检测性能。Yang 等<sup>[12]</sup>设计了一种对尺度变化和背景干扰具有较强鲁棒性的单阶段舰船检测器,并引入了坐标注意力模块,以获得更具代表性的语义特征,从而准确定位和区分舰船目标。Miao 等<sup>[13]</sup>提出了一种改进的轻量级 RetinaNet 模型,通过将主干的浅层卷积网络替换为 Ghost 模块,并减少深层卷积节点的数量,显著减少了网络的运算复杂度,同时保持了模型的健壮性和检测舰船目标的能力。基于锚框的检测算法在精度的提高上取得了一定的进展,但是锚框的固定尺寸导致模型在不同数据集上的泛化能力降低,同时锚框的设置增加了大量的计算成本,而且在模型训练中会导致大量的负样本,加剧了正负样本不平衡

问题。

基于无锚框的算法摆脱了对初始锚框的依赖,在回归阶段直接预测边界框的关键点位置或者特征图锚点到舰船目标边界框的位置信息,不再受限于锚框对舰船目标尺度和长宽比的限制,对多尺度、多形态舰船目标检测具有更好的鲁棒性。Sun<sup>[14]</sup>使用 FCOS 作为基础网络,并引入基于注意力机制的类别定位模块,通过生成引导向量来提高复杂场景中的定位性能。Jiang<sup>[15]</sup>提出了一种有效的无锚框轻型探测器,以提高对小型舰船的聚焦能力,结果表明,与传统模型相比,该探测器可以检测近海和离海舰船,具有更高的精度。Gao<sup>[16]</sup>通过密集连接和迭代融合多尺度特征,引入了一种具有密集注意特征聚合机制的无锚卷积网络,实验结果证明了对多尺度舰船检测的有效性。基于无锚框的算法虽然提升了检测器的检测速度,但对于多尺度目标检测,尤其是中、小目标检测结果不稳定,需进一步提升算法检测性能。

SAR 图像中舰船目标密集且尺度跨度较大,小尺度目标很容易淹没在复杂背景和噪声中,不利于特征的提取。卷积神经网络中,每一层卷积网络都在前一层信息的基础上提取更加抽象的特征,实现对上一层滤波器的聚合。较浅的卷积层学习到的是高分辨率、低语义、无法被有效识别的低层特征;较深的卷积层学习到的是粗糙、高度语义化、有助于识别的抽象特征。以往目标检测算法只在后期卷积层输出的特征图上进行预测,目前检测算法倾向于如何充分利用各个卷积层所学习到的特征来提升检测性能。本文拟尝试研究特征融合的机理,并引入注意力机制,改善融合方法,提高模型多尺度检测性能。

针对多尺度目标检测问题,本文提出了一种基于注意力机制改进的无锚框 SAR 图像舰船检测模型 DCAM-YOLOX。在 YOLOX<sup>[17]</sup>网络特征金字塔处嵌入空洞注意力模块 DCAM (dilated convolutional attention module),调节感受野与多尺度融合的关系,强化特征的代表能力。在检测头部设计中心性预测分支(Center-ness),对锚点的分类 Classification (Cls) 得分进行加权处理,调整模型的损失函数,优化检测结果。本文的主要贡献如下:

(1) 在 YOLOX 检测网络基础上,本文提出了基于注意力机制改进的无锚框 SAR 图像舰船检测模型,该模型在检测精度和检测速度中取得最佳平衡。

(2) 引入空洞注意力模块丰富特征的语义信息，保持参数个数不变的情况下增大了卷积核的感受野，让每个卷积输出都包含较大范围的上下文信息。

(3) 增加中心性预测分支，约束过多低质量且离目标中心较远的检测框。中心性与边界框回归 Regression (Reg) 分支共享同一组网络参数，在不引入额外参数的情况下优化检测结果。

## 1 模型的构建

本文提出的整体方案和网络架构如图 1 所示。网络主要分为三个组成部分：主干特征提取网络 (FEN)，特征金字塔 (FPN) 和检测头部 (Detection Head)。首先，将调整大小的 SAR 图像 ( $640 \times 640$ )

输入到主干特征提取网络中，提取三个不同分支不同尺度的特征图，特征尺寸分别为  $80 \times 80$ 、 $40 \times 40$  和  $20 \times 20$ 。然后，将多尺度特征图传递到空洞注意力模块以生成相应的显著特征图 (P3、P4 和 P5)。空洞注意力模块集成了空洞卷积、通道和空间注意力，在强化舰船目标与复杂背景差异性的同时，提升了目标特征的代表能力。此外，利用 BiFPN 网络构建的特征金字塔进一步实现了多尺度显著特征的融合，使得融合特征 (P3'、P4' 和 P5') 既含有便于分类的抽象语义信息，又包括适于回归定位的图像细节。本文将 BiFPN 网络的输入和输出的数量调整为 3 级，以 2 为因子进行上采样和下采样操作，并通过级联合并相同空间分辨率的特征图。最后，使用具有两个平行分支的解耦头 (Decoupled Head) 来预测边界框、中心性和分类的三维张量检测结果。

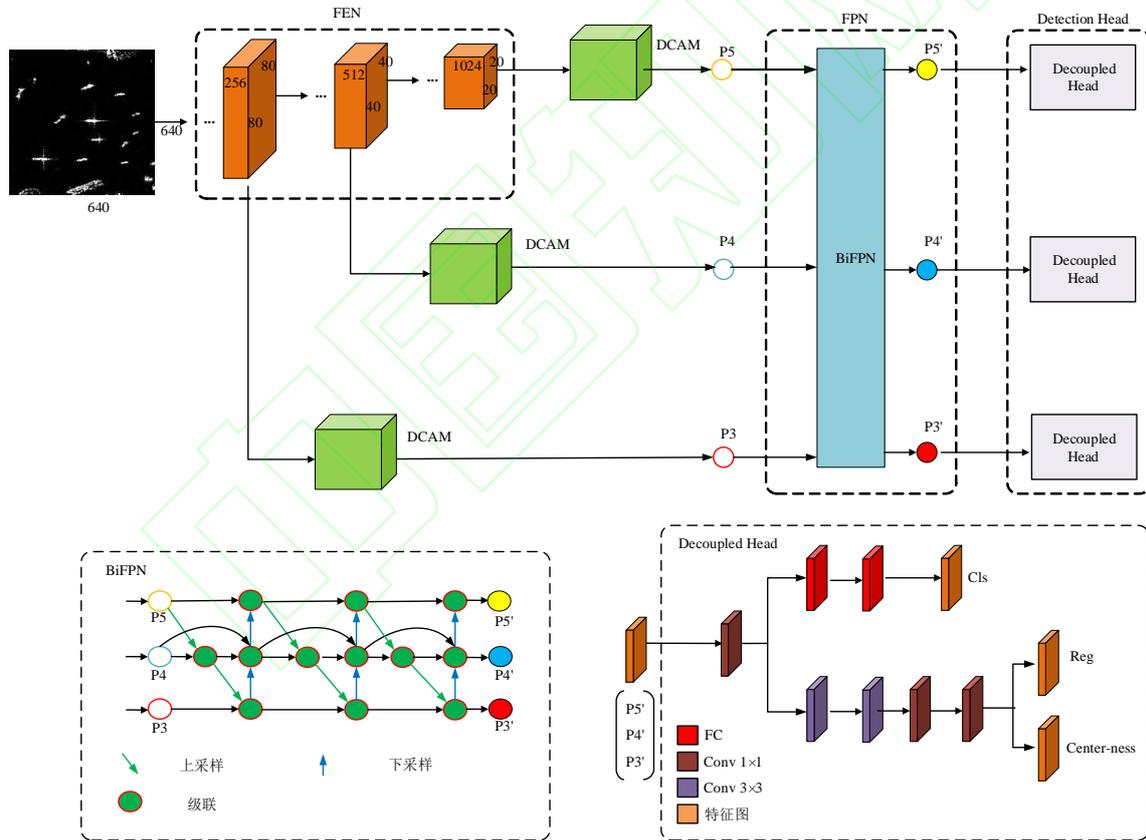


图 1 本文提出的检测网络结构图

Fig.1 The architecture of the proposed detection network

### 1.1 空洞注意力模块

不同尺度目标的检测性能受网络感受野的影响。合理增加卷积层的感受野可使检测网络提取更为丰富的上下文信息，强化特征的代表能力。注意

力机制更关注于图像的局部信息，细化输入映射。通道注意力机制关注有效通道，抑制无关通道，提高网络中信息的流动效率；空间注意力机制能够有效地关注舰船目标，抑制图像中其它干扰信息。本

文提出的空洞注意力模块包括 3 个子网络：空洞卷积块（Dilated Block）、通道注意力（Channel Attention）机制和空间注意力（Spatial Attention）机制。空洞注意力模块的网络结构如图 2 所示。

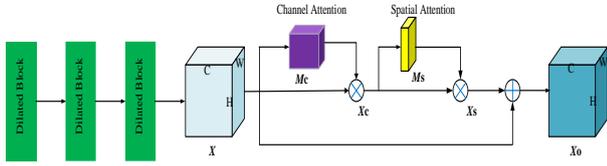


图 2 空洞注意力模块结构图

Fig.2 The architecture of DCAM

子网络具体的网络结构如图 3 所示。空洞卷积块由 3 层卷积网络组成，卷积核大小分别为  $1 \times 1$ 、 $3 \times 3$  和  $1 \times 1$ ，每层卷积后连接一个批标准化处理（BN）及 ReLU 激活函数，加速网络的快速收敛。对于卷积核大小为  $3 \times 3$  的中间层卷积网络，我们为不同尺度的分支设置了不同的扩张系数  $d$ ，此时等效的卷积核大小为  $3+2(d-1)$ ，扩大感受野的同时不引入额外的成本。

对于输入特征图  $X \in \mathbb{R}^{C \times H \times W}$ ，通道注意力自适应学习各通道的权重值，生成权重张量  $M_c \in \mathbb{R}^{C \times 1 \times 1}$ 。首先，使用平均池化和最大池化沿空间维度  $H \times W$  压缩输入，以基于全局信息获得两个通道张量。然后，通过元素求和合并这两个张量，并将和向量  $s$  输出到由两个全连接（FC）层组成的网络。为了降低通道注意力的复杂性，将第一个 FC 层的节点大小设置为  $C/r$ ，其中  $r$  是缩减率。最后，在网络的末端使用 sigmoid 函数作为简单的选通机制，以获得取值区间为  $(0,1)$  的缩放张量  $M_c$ 。

对于输入特征图  $X \in \mathbb{R}^{C \times H \times W}$ ，空间注意力自适应学习空间维度的权重值，生成权重张量  $M_s \in \mathbb{R}^{1 \times H \times W}$ 。首先，使用平均池化和最大池化沿通道维度  $C$  压缩输入，生成  $1 \times H \times W$  的 2 个中间张量。然后，将级联的张量传递到一个空洞卷积块中，进一步扩大感受野并有效利用上下文信息。最后，在网络的末端使用 sigmoid 函数作为简单的选通机制，以获得取值区间为  $(0,1)$  的缩放张量  $M_s$ 。

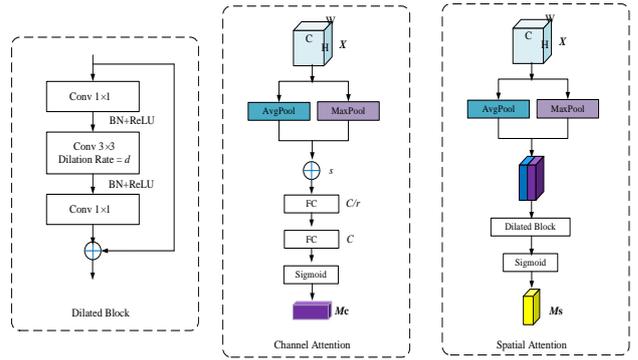


图 3 子网络结构图

Fig.3 The architecture of sub-networks

空洞卷积块使检测网络能够调节不同尺度分支的感受野大小，通道注意力机制和空间注意力机制分别生成权重张量，细化输入映射。空洞注意力模块的整个过程如公式（1-3）所示。

$$X_c = M_c \otimes X \quad (1)$$

$$X_s = M_s \otimes X_c \quad (2)$$

$$X_o = X_s + X \quad (3)$$

### 1.2 中心性预测

中心性预测分支实现了维度为  $H \times W$  的特征图各锚点 *Center-ness* 值的预测，用以表示该特征点到真实框的距离，中心性预测分支输出的维度为  $H \times W \times 1$ 。如图 4 所示，假设特征图上某点回归分支预测值为  $l, r, t, b$ ，则该特征点的 *Center-ness* 值为：

$$Center-ness = \sqrt{\frac{\min(l,r)}{\max(l,r)} \times \frac{\min(t,b)}{\max(t,b)}} \quad (4)$$

由公式（4）可知，特征点距离真实框中心越近，*Center-ness* 值越大，反之则越小。



图 4 中心性预测

Fig.4 Center-ness prediction

在模型训练时采用二元交叉熵损失<sup>[18]</sup>(BCE Loss) 对该分支进行调参，约束过多低质量且离目标中心较远的预测框。在测试时，将中心性得分乘上特征点的分类得分，过滤具有低 IoU 高置信度的预测框，优化检测性能。

### 1.3 损失函数

本文提出的检测模型的损失函数如公式(5)所示。损失函数主要包括三部分：分类损失  $L_{Cls}$ 、边界框回归损失  $L_{Reg}$  和中心性预测损失  $L_{Center-ness}$ 。其中， $(x, y)$  表示特征点的位置坐标， $\mathbf{P}, \mathbf{t}, \mathbf{c}$  分别表示模型的分类预测值，边界框回归预测值和中心性预测值， $\mathbf{C}^*, \mathbf{t}^*, \mathbf{c}^*$  分别表示分类真实值，边界框真实值和中心性真实值。 $N_{pos}$  表示正样本个数， $\lambda$  为权重系数。 $L_{Cls}$  采用 Focal Loss<sup>[19]</sup>， $L_{Reg}$  表示 GIoU 损失<sup>[20]</sup>， $L_{Center-ness}$  为二元交叉熵损失。

$$L(\{\mathbf{P}_{x,y}\}, \{\mathbf{t}_{x,y}\}) = \frac{1}{N_{pos}} \sum_{x,y} L_{Cls}(\mathbf{P}_{x,y}, \mathbf{C}_{x,y}^*) + \frac{\lambda}{N_{pos}} \sum_{x,y} [C_{x,y}^* \geq 1] L_{Reg}(\mathbf{t}_{x,y}, \mathbf{t}_{x,y}^*) + \frac{1}{N_{pos}} \sum_{x,y} L_{Center-ness}(\mathbf{c}_{x,y}, \mathbf{c}_{x,y}^*) \quad (5)$$

$$\sum_{x,y} [C_{x,y}^* \geq 1] = \begin{cases} 1, & C_{x,y}^* \geq 1 \\ 0, & otherwise \end{cases} \quad (6)$$

## 2 数据集与模型实现

### 2.1 SAR 数据集

本文使用的是中国科学院航天信息研究所数字地球实验室发布的用于舰船探测的 SAR 数据集。SSDD<sup>[21]</sup> 由 102 张高分-3 号<sup>[22]</sup> 图像和 108 张哨兵-1 号<sup>[23]</sup> 图像生成。对于高分-3 号，图像的分辨率分别为 3m、5m、8m 和 10m，成像模式分别为条带图 (UFS)、精细条带图 1 (FSI)、全极化 1 (QPSI)、完全极化 2 (QPSII) 和精细条带地图 2 (FSII)。哨兵-1 号成像模式包括 S3 Strip Map (SM)、S6 SM 和 IW 模式。

SSDD 共有 43819 张舰船图片和 59535 个舰船目标。每个图像的像素为  $256 \times 256$ 。舰船目标的标记格式与 Pascal VOC<sup>[24]</sup> 类似。SSDD 舰船目标的分布统计如表 1 所示，其中“Size”、“Min (Pixel)”和

“Max (Pixel)”分别表示舰船尺度类别、最小舰船像素个数和最大舰船像素个数。“Number”表示舰船目标的总数，“Percentage”表示舰船占整个数据集目标的百分比。从表 1 中，我们可以看到 SSDD 具有以下特征：图片中包含多尺度舰船目标，且尺寸范围变化较大；小型舰船和中型舰船占目标总数的很大比例。所有这些特征都给舰船检测带来了困难，并对舰船检测模型的性能提出了更高的要求。在实验中，本文按照 7:2:1 的比率随机分割训练集、验证集和测试集。训练集和验证集用于训练模型，测试集用于测试模型。

表 1 SSDD 舰船目标分布统计

Table 1 The statistical distribution of ships in SSDD

Size	Min (Pixel)	Max (Pixel)	Number	Percentage
Small Ship	4×6	32×32	35695	59.96%
Medium Ship	32×32	96×96	23660	39.74%
Large Ship	96×96	207×109	180	0.30%

### 2.2 评价指标

为了定量评估模型的检测性能，我们采用了四个广泛使用的标准，即精确性 (Precision)、召回率 (Recall)、平均精度 (AP) 和 F1 值，具体定义见公式 (7)-(10)。此外，为了更好地描述模型的多尺度目标检测性能，本文根据表 1 将 SSDD 数据集划分为大型、中型和小型舰船目标子数据集，相应的检测精度指标分别为 APL, APM 和 APS。具体的评价指标详见表 2。

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$AP = \int_0^1 P(R) dR \quad (9)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (10)$$

其中， $TP$ ， $FP$ ， $FN$  分别表示检测结果中 True Positives, False Positives 以及 False Negatives 的个数。

表 2 评价指标

Table 2 Evaluation metrics

指标	注释
----	----

AP50	AP (IoU=0.5)
AP75	AP (IoU=0.75)
APS	AP (Small Ship)
APM	AP (Medium Ship)
APL	AP (Large Ship)

### 2.3 模型实现

本文的模型、相关实验均基于 Python 语言及 Pytorch 架构实现，运行在 Ubuntu 16.04，GeForce-GTX 1080ti 内存为 11GB 的 GPU 平台上。本文使用在 ImageNet 数据集上预训练的权值初始化模型，并采用端到端的模式以随机梯度下降策略优化模型参数。不同尺度分支（特征图尺度为  $80 \times 80$ 、 $40 \times 40$  和  $20 \times 20$ ）空洞卷积的扩张系数  $d$  分别设置为 1,2,3。初始学习率设置为 0.0001，缩减率  $r$  设置为 8，权值系数  $\lambda$  设置为 5，迭代次数为 2000。

## 3 实验设计与结果分析

### 3.1 空洞注意力模块有效性分析

首先，本文设计消融实验验证空洞卷积块的个数对检测性能的影响。实验结果如图 5 所示。由图 5 可知，当空洞卷积块的个数为 3 时，模型取得最优的 AP50 和 AP75 值。当感受野过小时，模型未能提取丰富的上下文信息；当感受野过大时，卷积运算并不能保证所有的像素都得到计算，造成卷积核的不连续，不利于小型舰船目标的检测。

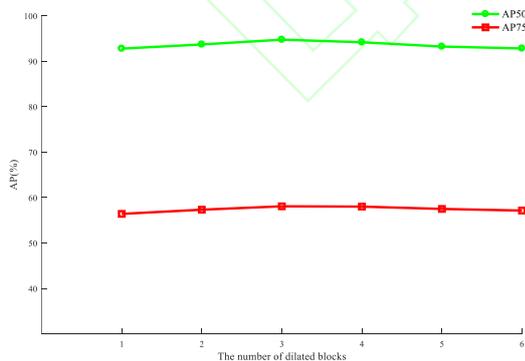


图 5 空洞卷积块个数对检测性能的影响

Fig.5 The affect using different number of dilated blocks

为了进一步验证空洞注意力模块的有效性，本文在其它子网络保持一致的情况下，统计了该模块

对模型检测性能的影响，实验结果如表 3 所示。由实验结果可知，引入空洞注意力模块在评价指标上均取得更好的检测结果。该模块集成了空洞卷积和注意力机制，在扩大感受野的同时，使得提取的特征更关注于显著区域，强化了目标与背景的差异。此外，本文通过热力图可视化 DCAM 模块输出特征的空间响应，有助于对卷积神经网络的决策过程进行调试，实验结果如图 6 所示。其中，第一列为边界框真值，第二列为空洞注意力模块输出特征热力图，第三列为不包含空洞注意力模块输出特征热力图。蓝色表示该目标区域激活值最小，红色表示该目标区域激活值最大。通过对比可知，DCAM 强化了目标特征的代表能力，在一定程度上避免了复杂背景对舰船检测结果的干扰，实现了舰船目标与背景的有效定位；相反，未引入 DCAM 模块的神经网络则误提取较多的干扰特征。图 7 为模型是否包含空洞注意力模块的检测对比图，其中，第一列为边界框真值，第二列为包含空洞注意力模块的模型，第三列为不包含空洞注意力模块的模型。绿色、红色、蓝色和黄色矩形框分别表示 ground truth, detection target, false alarm 和 missing target。通过对比可知，不包含空洞注意力模块存在更多的 false alarm 和 missing target 现象，而引入空洞注意力模块后，模型抗复杂背景的干扰能力增强，提升了多目标的检测效果。

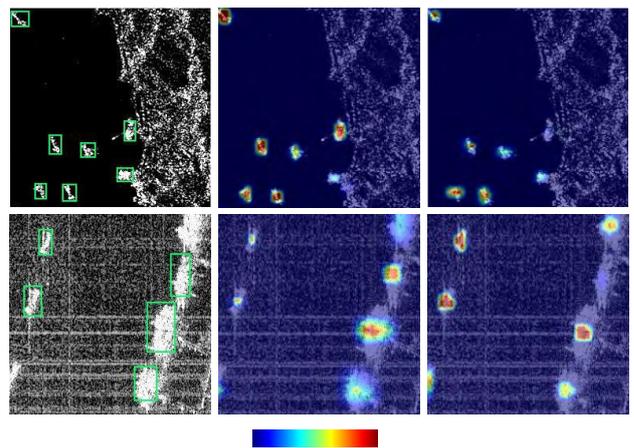


图 6 特征可视化对比

Fig.6 The comparison of feature vision



不包含DCAM	91.06%	92.14%	88.85%	0.90	57.59%	62.18%	60.23%	0.61
包含DCAM	<b>94.73%</b>	<b>94.02%</b>	<b>90.28%</b>	<b>0.92</b>	<b>58.85%</b>	<b>64.04%</b>	<b>63.46%</b>	<b>0.64</b>

表 4 特征金字塔对检测结果的影响

Table 4 The impact of FPN to detection results

模型	IoU=0.5				IoU=0.75				FPS
	AP50	Precision	Recall	F1	AP75	Precision	Recall	F1	
DCAM-YOLOX + FPN	91.48%	91.25%	87.06%	0.89	55.31%	61.41%	59.37%	0.60	68
DCAM-YOLOX + PAN	92.69%	92.00%	87.96%	0.90	56.70%	62.02%	61.63%	0.62	65
DCAM-YOLOX + 5-level BiFPN	94.57%	93.91%	90.11%	0.92	58.24%	63.36%	63.09%	0.63	57
DCAM-YOLOX + 3-level BiFPN	<b>94.73%</b>	<b>94.02%</b>	<b>90.28%</b>	<b>0.92</b>	<b>58.85%</b>	<b>64.04%</b>	<b>63.46%</b>	<b>0.64</b>	<b>60</b>

表 5 中心性预测对检测结果的影响

Table 5 The impact of center-ness to detection results

模型	IoU=0.5				IoU=0.75			
	AP50	APL	APM	APS	AP75	APL	APM	APS
无中心性预测	92.91%	82.32%	94.72%	90.88%	57.60%	40.02%	74.63%	48.65%
与分类共享的中心性预测	93.89%	83.00%	96.26%	91.38%	58.15%	42.79%	75.04%	49.56%
与边界框回归共享的中心性预测	<b>94.73%</b>	<b>83.07%</b>	<b>96.70%</b>	<b>92.96%</b>	<b>58.85%</b>	<b>43.81%</b>	<b>75.29%</b>	<b>50.79%</b>

表 6 各模型的检测结果

Table 6 The detection results of detectors on SSDD

模型	IoU=0.5				IoU=0.75				FPS
	AP50	APL	APM	APS	AP75	APL	APM	APS	
RetinaNet	85.70%	81.27%	96.20%	85.58%	41.52%	39.59%	64.18%	40.25%	39
CenterNet	84.19%	15.68%	89.46%	79.74%	32.91%	4.23%	44.77%	26.14%	78
Faster-RCNN	83.80%	63.53%	94.57%	69.23%	21.83%	40.01%	42.06%	5.59%	16
YOLOv3	90.98%	61.79%	95.96%	90.72%	48.15%	21.18%	62.65%	39.25%	61
YOLOv4	93.69%	74.80%	96.42%	91.28%	50.42%	25.64%	64.67%	40.00%	50
YOLOX	91.56%	63.95%	94.03%	88.78%	56.69%	38.39%	65.49%	48.78%	<b>95</b>
DCAM-YOLOX	<b>94.73%</b>	<b>83.07%</b>	<b>96.70%</b>	<b>92.96%</b>	<b>58.85%</b>	<b>43.81%</b>	<b>75.29%</b>	<b>50.79%</b>	60

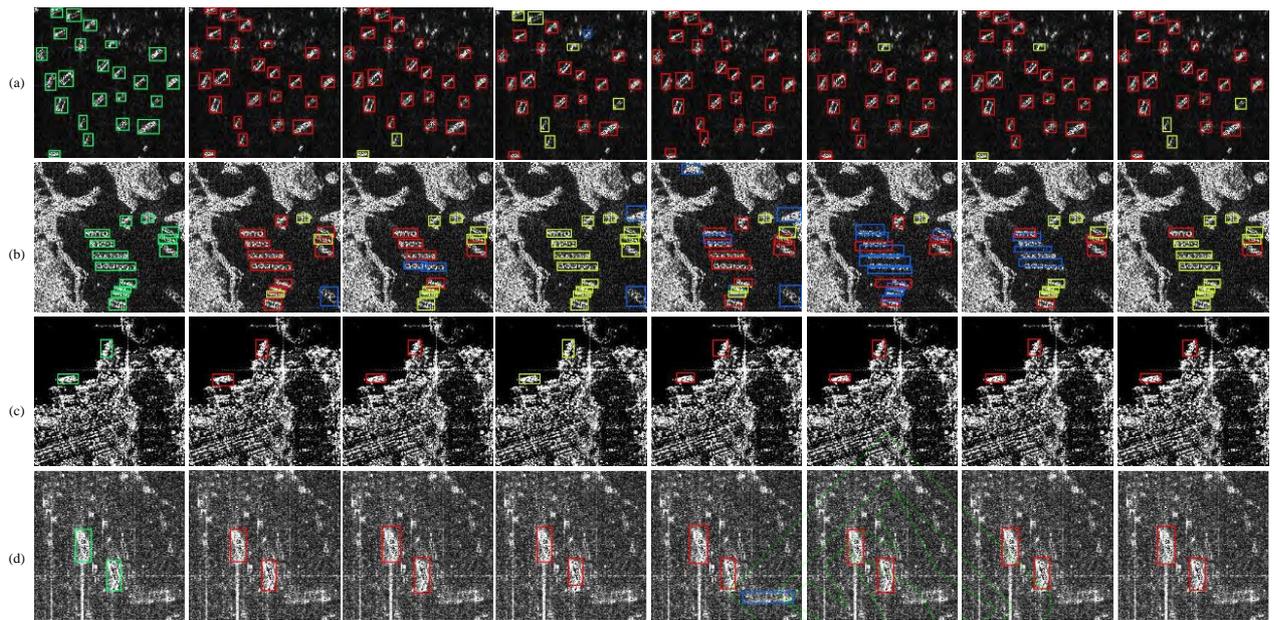


图 8 各模型检测结果对比

Fig.8 The comparison results of detectors

## 4 结论

本文提出了一种基于注意力机制改进的无锚框 SAR 图像舰船检测模型。根据空洞注意力模块调整不同尺度分支的卷积感受野，充分提取上下文信息；同时利用通道和空间注意力机制细化输入映射，提高网络中信息流动效率。设计中心性预测分支，在不增加网络参数的情况下优化特征点的边界框预测，提高定位精度。在 SSDD 数据集进行的实验结果表明，与现有主流的深度网络模型相比，该模型具有一定的优越性。

综上所述，DCAM-YOLOX 适用于 SAR 图像舰船目标的检测，我们相信，本文提出的模型将有助于舰船检测在未来研究中的应用。

## 参考文献

[1] Mao C, Huang L, Xiao Y, et al. Target Recognition of SAR Image Based on CN-GAN and CNN in Complex Environment [J]. IEEE Access, 2021, 9: 39608-39617.  
 [2] Wang C, Bi F, Zhang W, et al. An Intensity-Space Domain CFAR Method for Ship Detection in HR SAR Images [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(4): 529-533.

[3] Wang S, Wang M, Yang S, et al. New Hierarchical Saliency Filtering for Fast Ship Detection in High-Resolution SAR Images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(1): 351-362.  
 [4] Leng X, Ji K, Xing X, et al. Area Ratio Invariant Feature Group for Ship Detection in SAR Imagery [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(7): 2376-2388.  
 [5] Sun K, Liang Y, Ma X, et al. DSDet: A Lightweight Densely Connected Sparsely Activated Detector for Ship Target Detection in High-Resolution SAR Images [J]. Remote Sensing, 2021, 13(14): 2743.  
 [6] Yang R, Pan Z, Jia X, et al. A Novel CNN-Based Detector for Ship Detection Based on Rotatable Bounding Box in SAR Images [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 1938-1958.  
 [7] Avi A, Roe D. CFAR detection algorithm for objects in sonar images [J]. IET Radar, Sonar Navigation, 2020, 14(11): 1757-1766.  
 [8] REN S, HE K, GIRSHICK R. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern

- Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [9] Lin T, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection [C]//2017 IEEE International Conference on Computer Vision. Venice : IEEE Computer Society, 2017: 2999-3007.
- [10] J. Redmon, A. Farhadi. YOLOv3: An Incremental Improvement [EB/OL]. (2018-04-08)[2022-11-24]. <https://arxiv.org/abs/1804.02767>.
- [11] Kang M, Leng X, Lin Z, et al. A modified faster R-CNN based on CFAR algorithm for SAR ship detection [J]. Remote Sensing, 2017, 1-4.
- [12] Yang X, Zhang X, Wang N, et al. A Robust One-Stage Detector for Multiscale Ship Detection With Complex Background in Massive SAR Images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-12.
- [13] Miao T, Zeng H, Yang W, et al. An Improved Lightweight RetinaNet for Ship Detection in SAR Images [J]. Remote Sensing, 2022, 15: 4667-4679.
- [14] Sun Z, Dai M, Leng X, et al. An Anchor-Free Detection Method for Ship Targets in High-Resolution SAR Images [J]. Remote Sensing, 2021, 14: 7799-7816.
- [15] Jiang Y, Li W, Liu L. R-CenterNet+: Anchor-Free Detector for Ship Detection in SAR Images [J]. Sensors, 2021, 21(17): 5693.
- [16] Gao F, He Y, 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.
- [17] Ge Z, Liu S, Wang F, et al. YOLOX: Exceeding yolo series in 2021 [EB/OL]. (2021-08-06)[2022-11-24]. <https://arxiv.org/abs/2107.08430>.
- [18] Tian Z, Shen C, Chen H, et al. FCOS: Fully convolutional one-stage object detection [C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 9626-9635.
- [19] Lin T, Goyal P, Girshick R, et al. Focal loss for dense object detection [C]//2017 IEEE International Conference on Computer Vision. Venice : IEEE Computer Society, 2017: 2980-2988.
- [20] Rezatofighi, H, Tsoi N, Gwak J, et al. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression [C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach: IEEE, 2019: 658-666.
- [21] Wang Y, Wang C, Zhang H. A SAR Dataset of Ship Detection for Deep Learning under Complex Backgrounds [J]. Remote Sensing, 2019, 11(7):765.
- [22] Huang L, Liu B, Li B, et al. OpenSARShip: A Dataset Dedicated to Sentinel-1 Ship Interpretation [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(1): 195-208.
- [23] Everingham M, Zisserman A, Williams C, et al. The 2005 PASCAL Visual Object Classes Challenge [J]. MLCW, 2005, 117-176.
- [24] Sutanto A, Kang D. A Novel Diminish Smooth L1 Loss Model with Generative Adversarial Network [C]//12th International Conference of Intelligent Human Computer Interaction, Daegu: Springer, 2020: 361-368.
- [25] Bochkovskiy A, Wang C, Liao M. YOLOv4: Optimal Speed and Accuracy of Object Detection [EB/OL]. (2020-04-23)[2022-11-24]. <https://arxiv.org/abs/2004.10934>.
- [26] Tan M, Pang R, Le Q. EfficientDet: Scalable and Efficient Object Detection [C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle: IEEE, 2020: 10778-10787.
- [27] Duan K, Bai S, Xie L, et al. CenterNet: Keypoint Triplets for Object Detection [C]//2019 IEEE/CVF International Conference on Computer Vision, Seoul: IEEE, 2019: 6568-6577.