

# 基于回归模型与注意力的轻量化 SAR 舰船检测模型

李丽圆<sup>1,3</sup>, 李潇雁<sup>2</sup>, 胡琰悦<sup>1</sup>, 苏晓锋<sup>1\*</sup>, 陈凡胜<sup>1,2\*</sup>

- (1. 中国科学院智能红外感知重点实验室 中国科学院上海技术物理研究所, 上海 200083;
2. 国科大杭州高等研究院, 浙江 杭州 310024;
3. 中国科学院大学, 北京 100049)

**摘要:**合成孔径雷达(SAR)具有不受云层干扰、可全天时、全天候对地观测的特点,基于SAR图像的舰船检测已广泛用于海洋搜救、港口侦察、领海防御等民用或军用领域。然而,与大型舰船相比,像素点少、对比度低的小型舰船存在漏检率高的问题,并且速度和精度之间的平衡成为舰船检测算法天基应用的难点。针对以上问题,本文提出了一种基于YOLOv5s模型改进的舰船检测轻量化模型(ImShips)。首先,针对船体大小差异导致的漏检问题,采取在网络底部使用感受野较小的标准卷积,提升了模型对小规模舰船空间信息的获取能力。同时,在网络顶部设计了放大感受野的扩张卷积,保留了更多的语义特征,有利于大目标的特征提取。接着,提出将轻量级的通道注意力机制应用于YOLOv5的骨干网络和特征融合网络,通过对提取到的特征按重要性分配权重,实现纹理信息的筛选。最后,在下采样时采取深度可分离卷积代替标准卷积,减少了模型参数的数量,进一步提高了模型的推理速度。实验结果表明,在SAR舰船检测SSDD和ISSID数据集中,改进后的ImShips模型在保证精度的同时,所需的浮点计算数比YOLOv5s模型减少了45.61%,检测速度提高了8.31%。ImShips模型网络规模小、检测速度快,在实时天基舰船检测中具有较大的应用潜力。

**关键词:**舰船检测; YOLO回归模型; 通道注意力机制; 轻量化

中图分类号: TP751.1

文献标识码: A

## The research on lightweight SAR ship detection method based on regression model and attention

LI Li-Yuan<sup>1,3</sup>, LI Xiao-Yan<sup>2</sup>, HU Zhuo-Yue<sup>1</sup>, SU Xiao-Feng<sup>1\*</sup>, CHEN Fan-Sheng<sup>1,2\*</sup>

- (1. Key Laboratory of Intelligent Infrared Perception, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;
2. Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China;
3. University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** Synthetic aperture radar (SAR) has the advantages of all-sky and all-weather earth observation without cloud interference. Ship detection based on SAR images has been widely used in civil and military fields, including maritime search and rescue, port reconnaissance, territorial sea defense. However, different from large ships, the misdetection rate of small ships with fewer pixels and lower contrast is high. And it is difficult to balance speed and accuracy during on-orbit ship detection. To solve the above problems, an improved lightweight ships detection method (ImShips) based on YOLOv5s is proposed. Firstly, the standard convolution with small receptive field is adopted at the bottom of the baseline to obtain spatial information about small ships. And the dilated convolution with enlarged receptive field is added at the top of the baseline to preserve more semantic features, which is conducive to extract large targets feature. Then, a lightweight channel attention mechanism is applied to the backbone and neck of YOLOv5. And the weight is allocated to filter more important texture information. Finally, the depth-wise separable convolution is adopted to replace the standard convolution during down-sampling to reduce the number of parameters and improve the inference speed.

收稿日期: 2021-09-26, 修回日期: 2021-10-20

Received date: 2021-09-26, Revised date: 2021-10-20

基金项目: 国家自然科学基金项目(61975222)和中国科学院地球微卫星热红外光谱仪项目(XDA19010102)

Foundation items: Supported by National Natural Science Foundation of China(61975222) and CASEarth Minisatellite Thermal Infrared Spectrometer Project(XDA19010102)

作者简介(Biography): 李丽圆(1995—), 女(蒙古族), 内蒙古自治区呼和浩特市人, 在读博士, 主要研究领域为基于深度学习的目标检测. E-mail: liliyuan@mail.sitp.ac.cn

\*通讯作者(Corresponding author): E-mail: cfs@mail.sitp.ac.cn

Compared with YOLOv5s model, the experimental results show that ImShips achieve an increase in AP, while the FLOPs are reduced by 45.61%, and the speed is increased by 8.31% in SSDD and ISSID datasets. The speed and accuracy of ImShips model are improved effectively on sea surface object detection. The proposed method has great application potential for on-orbit ship detection.

**Key words:** ship detection, YOLO regression model, channel attention, lightweight

## 引言

随着空间遥感技术的飞速发展,从星载传感器获取的高分辨率和大规模遥感图像不断丰富,促进了遥感图像广泛的应用,例如自然灾害评估,城市规划,交通管理,环境监测,和海上目标的探测和识别。其中,舰船检测在民用领域,例如航行安全、渔业管理、船舶救助、海洋监测等,和军事领域,例如国防建设、航行安全、港口监视等具有重要价值,遥感图像中的自动舰船检测已经引起了越来越多的兴趣。

舰船检测包括船体检测和尾流检测。在某些卫星图像中,尾迹比船体更明显。但是,当舰船停泊在港口,或者船只尾迹被云遮挡时,尾迹并非一直存在,因此船体检测应用场景更广。早在2002欧盟海洋变换空间的检测与分类(Detection and Classification of Marine Transform Space, DECLIMS)和法国研究与发展研究所(Institute for Research and Development, IRD)合作利用光学图像进行舰船目标检测,自此,基于卫星图像的舰船目标检测和分类应用系统开始出现。现有船体目标检测技术主要依赖于可见光、红外和合成孔径雷达(SAR)成像技术。其中,可见光拥有高图像分辨率,在遥感舰船检测方面有丰富的研究成果,但可见光成像基于光线的反射,在光照不足情况下,难以看清并识别目标,在全天候检测能力方面存在不足。红外热成像通过接收目标本身和背景之间的红外辐射差而工作,被动接收辐射,隐蔽性好,安全性更强。但是红外热成像的数据集较少,且分辨率低。合成孔径雷达可以在不同的光照条件和天气下全天时、全天候工作,探测距离远,可远距离作业去提供高分辨率图像。由于光学图像和SAR图像的成像机制不同,基于SAR图像的舰船检测易受海浪、岛屿、陆地等回波、射频及大气噪声等海面杂波虚警干扰。

传统的基于先验知识的SAR舰船检测方法首先采用二值化、阈值分割或形态学等方法分割目标与背景区域,接着提取与识别几何和纹理特征。在船只目标几何特征明显的场景下,基于几何特征例

如:面积、长宽比、紧凑度、船头形状等的分割方法效果很好。但当舰船被云雾遮挡或靠岸时,基于几何特征的分割方法并不适用,需要加入纹理特征。利用统计、几何方法、模型法和信号处理法等提取能量、惯量、熵、相关性等纹理特征,可以表达图像整体特征,具有旋转不变性以及抗干扰能力。但是,纹理特征受分辨率影响大,不同分辨率情况下,纹理特征偏差大。基于经典算子的方法主要包括目标候选区域提取和目标精细判别两部分,需要手工设计特征并分类。候选区域提取的方法通过使用描述纹理的局部二进制模式,描述特征的方向梯度直方图(Histogram of Oriented Gradients, HOG)、尺度不变特征变换(Scale-Invariant Feature Transform, SIFT)或角点提取(Harris)等手动提取特征,以降低背景对检测的干扰。接着使用支持向量机、极端学习机、K最邻近法、线性判别分析或自适应增强等算法对目标精细分类。上述传统算法针对简单场景下的目标背景分割取得了成功。但是,由于其手工特征的建立过分依赖专家经验,因此其泛化能力较弱,当舰船处于复杂的天气条件下时会遇到瓶颈。

自2014年深度学习技术火热发展,目标检测算法也从基于手工特征的传统方法转向了基于深度神经网络的检测技术,深度学习凭借强大的自动提取特征能力被广泛应用于SAR影像中的舰船检测。Zou等人<sup>[1]</sup>提出了改进地生成对抗网络从而生成多尺度高分辨率SAR舰船图像,与原始小样本结合训练YOLOv3模型<sup>[2]</sup>,实现小样本下的高精度舰船检测。Chen等人<sup>[3]</sup>提出了一种将特征金字塔模型(Feature Pyramid Networks, FPN)<sup>[4]</sup>嵌入到传统的区域提议网络中,并映射到新的特征空间以进行对象识别。随后,基于形状相似距离度量的K均值聚类算法用于优化FPN。上述两阶段方法通常可以达到更高的检测精度,但它们的检测计算速度往往比一阶段方法慢。

You Only Look Once (YOLO)系列是经典的一阶段目标检测算法,将图像划分为网格系统的端到

端的目标检测算法,可将目标检测问题转化为回归问题。Wang等人<sup>[5]</sup>提出了一种基于YOLOv3的新型检测模型,在主干网中添加了扩张卷积和残差连接以增强感受野和浅层信息,并融合了改进的FPN以获取更多层级信息,该模型对SAR图像检测的平均精度(Average Precision, AP)提高了6.5%。Zhao等人<sup>[6]</sup>提出了结合空间和通道注意力的金字塔模型,模型的AP相比YOLOv3提高了5.6%。但是特征金字塔网络使用金字塔来表示不同尺度的语义信息,提高了目标检测的准确性,但是网络模型复杂,参数量巨大,不能满足检测实时性的要求。Tang等人<sup>[7]</sup>首先使用噪声级分类器和SAR目标潜在区域提取模块,然后基于YOLOv5(暂无论文,利用公开的代码进行研究)进行识别。虽然取得了不错的检测精度,但是上述方法由多个独立的模块组成,存在人工预处理和后续处理。

基于YOLO系列的目标检测算法可分为基于锚(Anchor based)、无锚(Anchor free)和两者融合类。基于Anchor based检测算法会在特征图上密集生成大量的锚框,在特征图中每个节点存储的是以该节点为中心的数个子图的特征,根据节点周围预设的子图的特征计算分类,该方法适用于密集目标的检测。而Anchor free检测算法通过在图片中检测关键点的方法来设定锚框,根据锚框网格特征计算分类,大大减少了计算量。考虑到检测实时性的要求,我们的模型基于Anchor free的YOLOv5算法。

当SAR图像中舰船实际尺寸较小时,仅将舰船显示一个亮点,导致像素点少、对比度低的小型舰船存在漏检率高的问题。并且速度和精度之间的平衡成为舰船检测算法天基应用的难点。为了获得更高的检测精度和更快的识别速度,本文提出了一种改进的SAR舰船检测模型(ImShips)。针对船体大小差异导致的漏检问题,使用了不同大小感受野的扩张卷积(Dilated Conv)和轻量化的通道注意力机制(SElayer)去筛选更重要的舰船特征。为了满足星载实时性检测的要求,结合了深度可分离卷积(DWConv)对模型进行压缩,大大减少了模型复杂度与推测时间。

## 1 海面舰船检测轻量化模型

### 1.1 回归模型概述

由卷积神经网络的性能优良的驱动,基于深度学习的检测模型已经成为解决舰船检测问题的新方法。根据是否生成提议区域进行分类,基于深度

学习的检测方法可以大致分为:一阶段和两阶段的检测方法。两阶段的经典检测方法包括:R-CNN<sup>[8]</sup>, Fast R-CNN<sup>[9]</sup>和Faster R-CNN<sup>[10]</sup>。一阶段的经典检测方法包括:SSD<sup>[11]</sup>, RetinaNet<sup>[12]</sup>和YOLO系列。两阶段检测模型可提供较高的定位精度,而一阶段检测模型在速度方面具有绝对优势。因此,为了保证识别的实时效果,我们采用一阶段法对舰船进行检测。

YOLO系列模型在深度学习目标检测领域的准确率和速度综合表现最好。YOLOv5在YOLOv4<sup>[13]</sup>的基础上,结合EfficientDet<sup>[14]</sup>中的复合模型缩放方法,可以动态平衡模型参数的精度、速度和数量。在实时海洋观察和及时舰船救援的一些实际应用中,提高舰船检测速度与提高检测精度同等重要。为了保证模型在边缘计算设备上部署时的实时推理速度,我们的模型基于计算机视觉实时应用首选YOLO系列,实现了针对SAR图像快速多尺度舰船目标检测。所提出的总体框架如图1所示。首先输入端采取了Mosaic数据增强、自适应锚框计算、自适应图片缩放,丰富了数据集,使网络的鲁棒性更好。骨干网(Backbone)中使用了Focus结构、Ghost-Bottleneck结构和CSP结构<sup>[15]</sup>,以及针对不同尺度目标设计的Dilated Conv模块和SElayer模块,在第1.2和1.3节中进行了详细介绍。Focus结构中,首先对640×640×3的输入图片采用切片操作,输出320×320×12的特征图,再经过一次32个卷积核的卷积操作,最终变成320×320×32的特征图。特征融合网络(Neck)中,利用FPN<sup>[4]</sup>、PAN<sup>[13]</sup>和改进的上采样结构DWConv模块在修改后的网络不同阶段提取多尺度信息,轻量化处理的DWConv模块在1.4节中进行了详细介绍。

### 1.2 扩张卷积Dilated Conv模块

池化操作增大了感受野但是降低了分辨率,导致检测精度低。而扩张卷积<sup>[16]</sup>在不降低空间维度的前提下增大了相应的感受野指数,提高感受野,即空间上融合更多特征融合,可以提取多尺度空间信息。因此,在网络顶部,原YOLOv5s模型中的标准卷积被替换为扩张卷积层。扩张卷积公式如(1),其中, $F$ 为离散函数, $*$ 为扩张卷积, $K$ 为离散滤波器, $s$ 为步长, $l$ 为空洞因子, $t$ 为自变量, $F(s)$ 为步长为 $s$ 的离散卷积, $K(t)$ 为自变量为 $t$ 的离散滤波器。扩张卷积感受野公式如(2),其中, $RF_{i-1}$ 为上一层的感受野, $k$ 为卷积尺寸。膨胀的卷积核将原



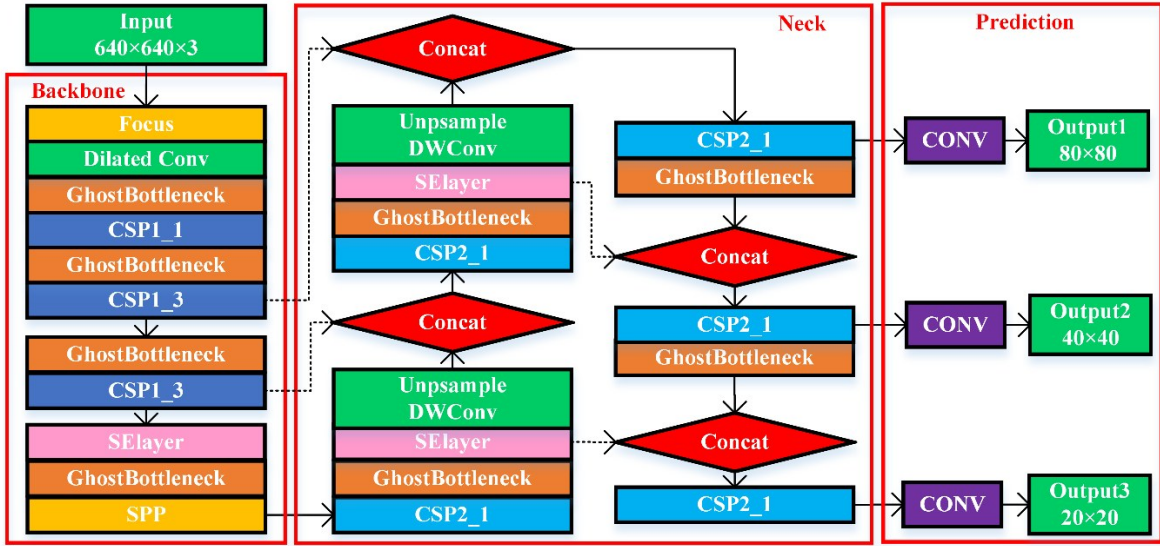


图1 轻量化回归模型ImShips

Fig. 1 Lightweight regression ImShips model

卷积核不占用的位置填充为0,膨胀后的卷积核尺寸计算公式如(3)。

我们的骨干网络由各种膨胀率的卷积层的几个分支组成,有利于提取具有大纵横比的舰船特征并通过其全局依赖性增强局部特征。骨干网络底部感受野小,可以保留更多的图像细节,有利于小目标的特征提取。网络顶部感受野大,保留了较多的语义特征,有利于大目标的特征提取,有效提高了ImShips模型对不同尺度舰船的特征提取效果。

$$(F * K)(s + lt) = \sum_{s+lt=p} F(s)K(t), \quad (1)$$

$$RF_i = RF_{i-1} + (k - 1) \times s, \quad (2)$$

$$\text{kernel\_dilation} = \text{dilation} * (\text{kernel} - 1) + 1. \quad (3)$$

### 1.3 通道注意力SElayer模块

如图2中 $h$ 至 $h+1$ 层卷积神经网络所示,每个神经元可以被看作一个滤波器,前一层输入的特征图 $w_i$ 和 $w_j$ 通过滤波器和窗口滑动,计算输出的新特征图 $F_i$ 和 $F_j$ 。卷积是对局部区域进行空间和通道间的特征融合,而普通的卷积操作默认对输入特征图的所有通道进行相加融合,并没有考虑通道间的关系。如公式(4),\*表示标准卷积操作, $F_u: X \rightarrow U$ 为卷积映射, $X$ 为输入,输出为 $U = [u_1, u_1, \dots, u_c] \in R^{H \times W \times C}$ ,卷积核为 $V = [v_1, v_1, \dots, v_c]$ , $v_c$ 表示第 $c$ 个卷积核, $v_c^s$ 表示 $s$ 个通道的2-D卷积核。

软注意机制<sup>[17]</sup>起源于人类视觉系统,用于定位目标最显著的特征,消除计算机视觉任务的冗余,

包含空间注意<sup>[18]</sup>、通道注意和混合模块<sup>[19]</sup>。在空间注意模块中,将图像中的空间信息转化为掩模,然后对其进行评分,提取关键信息。但是,空间注意模块对每个通道中的图像特征的利用是均等的,因此忽略了通道信息,仅局限于原始图像特征提取阶段。在混合模块同时对通道注意和空间注意进行评分,但是所需计算量较大。2017年,SENet<sup>[20]</sup>关注了通道之间的关系,可以自动学习到不同通道特征的重要程度,该模型赢得了ImageNet大型视觉识别挑战赛(ILSVRC2017)的冠军。我们考虑到YOLOv5s不能很好地解决小型舰船的物体检测问题,尤其是在场景有干扰或阴天有雾的情况下。受注意力机制可以增强重要信息的启发,我们利用SElayer模块改进骨干网络,从而抑制环境引起的干扰,进一步提高网络对小目标的检测能力。第一步,利用全局平均池化操作将某个通道的空间特征压缩为全局特征,得到向量 $Z$ ,如公式(5)。值得一提的是第二步,为了降低模型复杂度,在选择不同通道的权重值时,首先使用一个全连接层以降维系数 $r$ 对向量 $W_1$ 降维,然后使用ReLU函数激活,接着使用全连接层恢复原始的维度,最后利用sigmoid函数 $\sigma$ 激活,如公式(6)。第三步将权重值与原始输入 $U$ 相乘,如公式(7)。轻量级的SElayer模块提升了模型对通道特征的敏感性,应用在ImShips的骨干网络和特征融合网络中,在只增加了较少的计算量的情况下,增强了对小规模舰船空间信息的获取和检测性能。

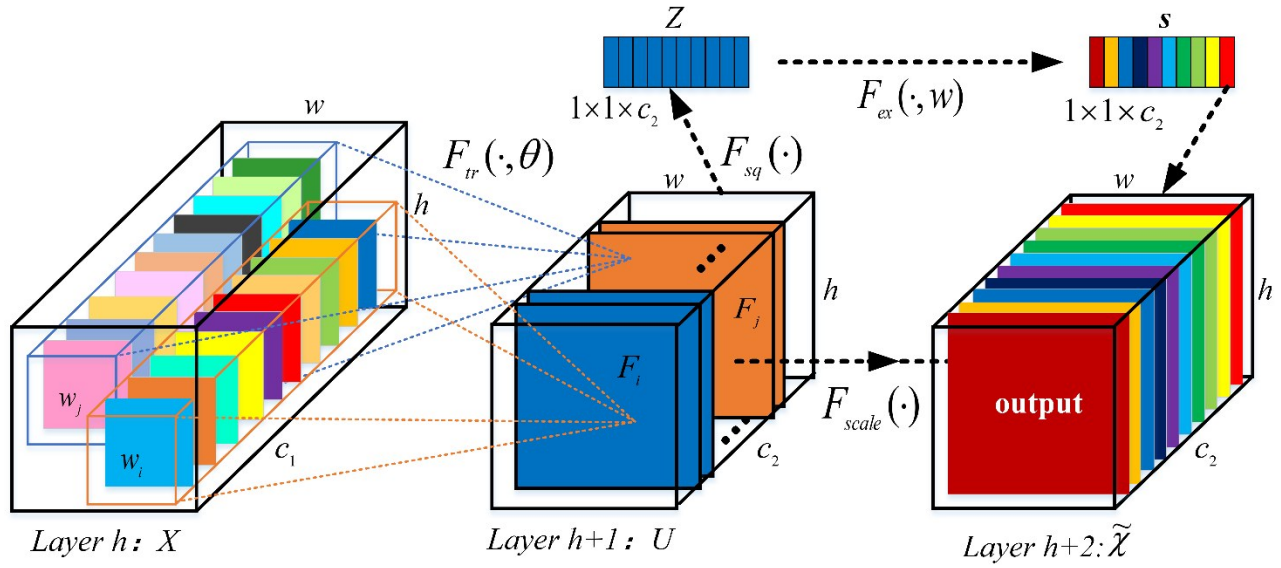


图2 SLayer模块

Fig. 2 SLayer module

$$u_c = v_c * X = \sum_{s=1}^{c'} v_c^s * X^s, X \in R^{H \times W \times c'}, \quad (4)$$

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i,j), z \in R^c, \quad (5)$$

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 ReLu(W_1 z)),$$

$$W_1 \in R^{\frac{c}{r} \times c}, W_2 \in R^{c \times \frac{c}{r}}, \quad (6)$$

$$\tilde{\chi}_c = F_{scale}(u_c, s_c) = s_c \cdot u_c. \quad (7)$$

#### 1.4 深度可分离卷积 DWConv 模块

为了应对地面基站数量、存储空间和计算资源有限,平衡星载舰船检测的精度和速度,我们提出了一种基于轻量级回归算法的星载图像舰船检测的框架。基于 MobileNet 网络<sup>[21]</sup>中的组卷积和深度可分离卷积对通道进行分组,有效减少了卷积计算量,我们利用深度可分离卷积代替标准卷积去减少参数。假设输入特征图尺寸为  $C_{in} \times H \times W$ , 卷积核大小为  $C_{in} \times K \times K$ , 输出通道数为  $C_{out}$ 。DWConv 模块的第一步是:保持输入输出特征图通道数不变的情况下,对于输入特征图的每个通道,通过一个尺寸为  $K \times K$  的逐深度卷积核,这一过程所需的计算量为  $C_{in} \times K \times K$ 。第二步是:通过尺寸为  $C_{in} \times 1 \times 1$  的逐点卷积核实现通道数改变,所需计算量为  $C_{in} \times 1 \times 1 \times C_{out}$ 。因此,深度可分离卷积与标准卷积的计算量比值如公式(8),本文使用的卷积核大小为  $3 \times 3$ ,因此深度可分离卷积的计算量约为标准卷积的九分之一到八分之一。随着要提取的特征越来越多,使用 DWConv 模块代替标准卷积(Conv)

大大节省了计算资源,有利于提高检测速度。

$$\frac{Param\_DWConv}{Param\_Conv} = \frac{C_{in} \times K \times K + C_{in} \times 1 \times 1 \times C_{out}}{C_{in} \times K \times K \times C_{out}} = \frac{1}{C_{out}} + \frac{1}{K \times K} \quad (8)$$

## 2 实验分析与讨论

### 2.1 数据

该实验采用了公开的 SAR 图像舰船目标检测数据集(SAR Ship Detection Dataset, SSDD<sup>[22]</sup>)和高分辨率 SAR 图像 ISSID<sup>[23]</sup>数据集。SSDD 数据集包含 2456 个舰船,舰船目标长或者宽度所占图像尺寸的比例在 0.04~0.24 范围内,比模式分析、统计建模和计算学习(Pattern Analysis, Statistical Modelling and Computational Learning, PASCAL VOC)竞赛数据集的 0.2~0.9 要小很多,包含小型船只。ISSID 数据集包含 5498 张 1m、3m、5m 和 7m 分辨率的 SAR 图像和 15136 艘舰船。

### 2.2 评价函数

准确率(Pr.)和召回率(Re.)是模型性能两个不同维度的度量,所需的混淆矩阵如表 1。准确率表示预测结果中,预测为正样本的样本中,正确预测为正样本的概率,如公式(9)。召回率表示在原始样本的正样本中,最后被正确预测为正样本的概率,如公式(10)。平均精准度(AP)是 P-R 曲线下围成的面积,使用积分来进行计算,如公式(11)。参

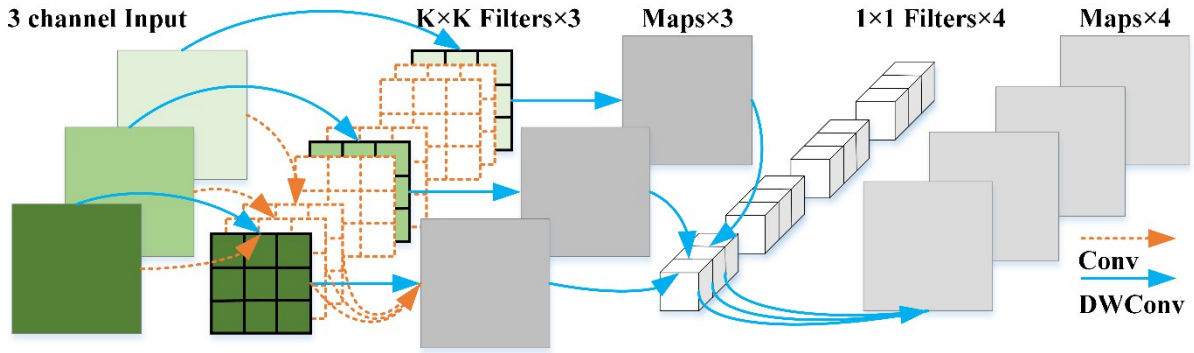


图3 DWConv 模块

Fig. 3 DWConv module

数量(Para.)、浮点计算数(FLOPs)衡量卷积网络的复杂度,无偏差时如公式(12-13)。帧频(Frame)、推测时间( $T_{mean}$ )衡量网络速度。综上,准确率、召回率、平均精度和帧频越高,参数量、浮点计算数和推测时间越低,模型性能越好。

表1 混淆矩阵

Table 1 Confusion matrix

真实分类	实验预测分类	
	船体	背景
船体	TP	FN
背景	FP	TN

其中,  $TP$  (True Positive) 为真阳性,  $TN$  (True Negative) 为真阴性,  $FP$  (False Positive) 为假阳性,  $FN$  (False Negative) 为假阴性。  $k_H \times k_W \times C_{in}$  为卷积核大小,  $C_{out}$  为输出通道数,  $g$  为分组卷积组数,  $C_{out} \times H_{out} \times W_{out}$  为输出数据点。

$$Precision = \frac{TP}{TP + FP}, \quad (9)$$

$$Recall = \frac{TP}{TP + FN}, \quad (10)$$

$$AP = \int_0^1 p(r) dr, \quad (11)$$

$$Para. = k_H \times k_W \times C_{in} / g \times C_{out}, \quad (12)$$

$$FLOPs = (2 \times k_H \times k_W \times C_{in} / g - 1) \times C_{out} \times H_{out} \times W_{out} \quad (13)$$

### 2.3 实验结果与分析

我们的实验基于台式计算机,使用 NVIDIA-smi 470.57.02, 驱动版本为 470.57.02, CUDA 版本为 11.4, NVIDIA RTX 3070 GPU, Pytorch 框架。我们利用 SGD 优化器来更新网络权重。网络的初始学习率设置为前 100 次迭代的 0.01, 后 100 次迭代

的 0.001。优化器的重量衰减为 0.001, 动量为 0.98。

如表 2 所示, ImShips 模型的网络层数为 390, 网络参数为 414 万, 每秒可处理帧数为 119.18 帧。我们的模型虽然加入了一些网络层数, 但是网络参数比 YOLOv5s 减小了 43.13%, 比 YOLOv5x 减小了 95.28%。如表 3 所示, 通过比较, 我们发现 YOLOv5m、YOLOv5l 和 YOLOv5x 随着网络参数加深, 准确率、召回率和平均精度都在下降, 我们分析其检测效果并不理想的原因如下。首先, 训练 YOLOv5 模型时, 图片尺寸设置为 S 并不会影响任意尺寸图像的检测, 这个数值设置的目的是使输入图像先被重新调整成大小为  $S \times S$  的特征图, 满足检测网络结构, 最后再重新调整成原始图像尺寸进行显示。我们综合考虑了 YOLOv5m、YOLOv5l 和 YOLOv5x 所需算力与 GPU 运算能力的限制, 训练上述模型时, 图片尺寸 (IS) 被设置为 320, 批处理大小 (BS) 为 8。而针对轻量化模型 ImShips 和 YOLOv5s 训练时, 在台式计算机运算能力可以保证训练的情况下, 我们设置图片尺寸为 640, 批处理大小为 32。随着图片尺寸增大, 纹理和上下文信息等判别性特征会被更好地捕捉。其次, 在一定范围内, 批处理大小越大, 其确定的下降方向越准, 引起训练震荡越小, 可能也对模型训练产生了影响。最后, 当 SAR 图像中舰船实际尺寸较小时, 可能仅将舰船显示为 SAR 图像中的一个亮点, 在通过 CNN 模型时, 深度卷积层可能会丢失小船有用的特征信息, 从而导致漏检和误报。因此, YOLOv5m、YOLOv5l 和 YOLOv5x 并不适用于 SAR 舰船检测。

我们的模型和 YOLOv5s 模型表现出了针对 SAR 图像舰船检测的优越性。与 YOLOv5s 相比, ImShips 模型的准确率提高了 0.65%, 召回率提高了



表2 模型参数

Table 2 Model parameter

Model	Layers	Para. /Million	Frame
ImShips	390	4.14	119.18
YOLOv5s	283	7.28	109.27
YOLOv5m	391	21.38	85.10
YOLOv5l	499	47.06	74.77
YOLOv5x	607	87.78	54.07

表3 模型评价

Table 3 Model evaluation

Model	IS	BS	Pr. (%)	Re. (%)	AP (%)	GFLOPs	$T_{mean}$
ImShips	640	32	97.18	98.94	98.68	9.3	8.39
YOLOv5s			96.55	97.91	97.46	17.1	9.15
YOLOv5m			89.10	93.35	91.70	51.4	11.75
YOLOv5l	320	8	86.44	93.90	91.54	115.6	13.37
YOLOv5x			84.07	92.48	89.60	219.0	18.49

1.05%, AP 提高了 1.25%, 所需的浮点计算数减少了 45.61%, 推断时间减少了 8.31%, 平均推断时间检测的图片大小大约为  $400 \times 300$ , 如表 3 所示。ImShips 模型有效提高了海面物体检测任务的速度和准确度。

如图 4(a), 我们的模型在对比度高的简单背景影像中, 检测效果非常好, 不受较大的海面杂波影

响。如图 4(b-j), ImShips 模型可以解决针对复杂背景 SAR 图像中, 包括靠岸区域、河道、港口小目标和多目标舰船检测。如图 4(k) 体现出了我们的模型针对舰船目标尺寸差异过大的问题, 检测性能表现良好。针对复杂背景目标检测, 传统方法要进行海陆分割才能进行检测, 而 ImShips 模型针对 SAR 图像中多个小型舰船目标的检测, 在简单场景和复杂场景中, 包括舰船处于大片海域、靠岸区域、河道、港口, 检测性能依旧优越。但是我们的模型也存在少量的虚警, 我们分析了图 4(g) 右上角的陆地虚警, 考虑到是图像输入尺寸太小, 限制了网络检测的能力。图 4(k) 最上方置信度为 0.30 的框中, 包含了两个紧邻的船只只被检测为一个船只, 我们分析可能的原因有数据集包含的舰船类型不够丰富, 以及网络特征提取能力还需增强, 我们的模型还需要提高针对紧邻目标的分辨能力。

### 3 结语

实验结果表明, 我们的模型针对 SAR 舰船检测 SSDD 和 ISSID 数据集具有优越性, 改进后的 ImShips 模型在比 YOLOv5s 模型的权重数减少了 45.61%, 检测速度提高了 8.31%, 平均精准度提高了 1.25%。ImShips 模型网络规模更小, 实时性更好, 有效提高了海面物体检测任务的速度和准确度。通过大量

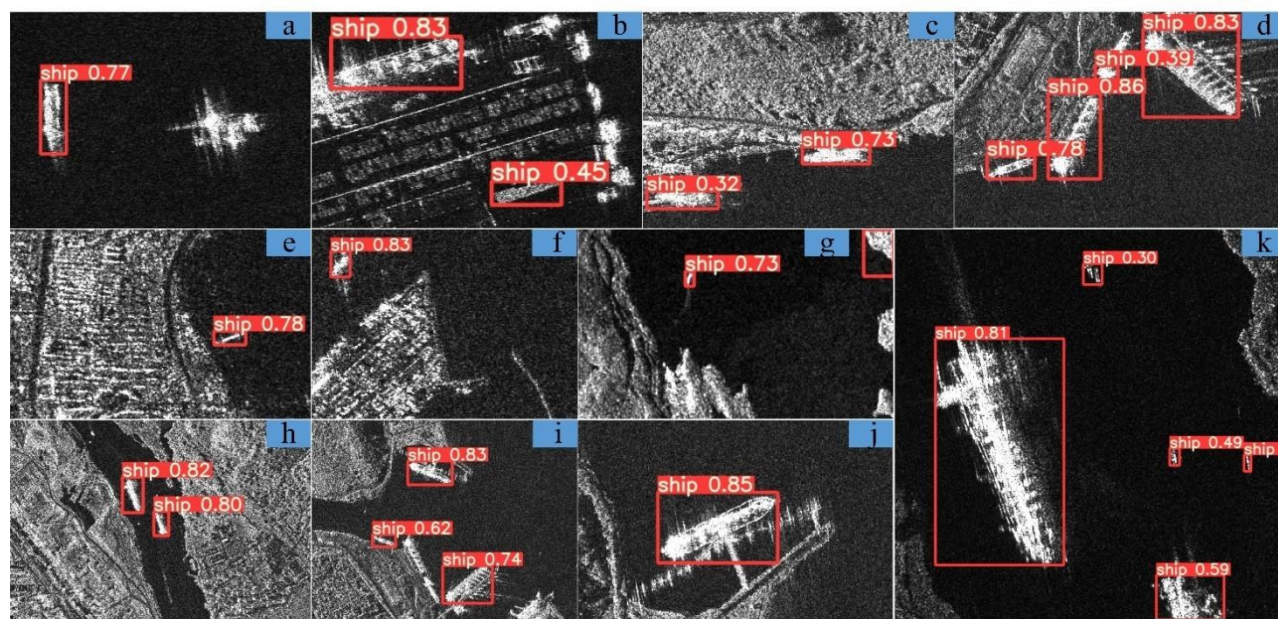


图4 针对简单背景:(a)大片海域,和针对复杂背景:(b-g)靠岸区域,(h)河道,(i,j)港口,(k)多尺度船体大小的SAR图像回归预测结果

Fig. 4 Regression prediction results of SAR images with simple background: (a) wide sea, and complex background: (b-g) shoreline, (h) river channel, (i, j) port, (k) different hull sizes

文献调研,我们对星载SAR舰船探测所面临的挑战做了以下总结:1)陆海分割不正确,在陆海交界处容易产生过多的误报。2)SAR图像中复杂的背景,例如噪声,云层和礁石等,很容易干扰大小不同的舰船目标。3)船只的外观和大小各异,小型船只难以被准确地定位和检测。4)在资源有限的星载实时应用中,算法的检测性能和计算复杂度难以平衡硬件资源的存储空间。5)数据集较少,标注工作量巨大,尤其是被旋转边框标注的图像。由此,我们接下来的工作将会着重于制作旋转框数据集,或者标注所需工作量更大的语义分割数据集,并合理部署轻量化模型到硬件平台。

## References

- [1] Zou L, Zhang H, Wang C, *et al.* MW-ACGAN: Generating Multiscale High-Resolution SAR Images for Ship Detection[J]. *Sensors*, 2020, **20**, 6673.
- [2] Redmon Joseph, Farhadi Ali. YOLOv3: An Incremental Improvement [C]. *Computer Vision and Pattern Recognition*, April 2018.
- [3] Chen P, Li Y, Zhou H, *et al.* Detection of Small Ship Objects Using Anchor Boxes Cluster and Feature Pyramid Network Model for SAR Imagery [J]. *Mar. Sci. Eng.* 2020, **8**, 112.
- [4] Lin T, Dollár P, Girshick R, *et al.* Feature Pyramid Networks for Object Detection [C]. *Computer Vision and Pattern Recognition*, 2017, . **936**–944
- [5] Wang Jingpu, Lin Youquan, Guo Jie, *et al.* SSS-YOLO: towards more accurate detection for small ships in SAR image [J], *Remote Sensing Letters*, 2021, **12**:2, **93**–102
- [6] Zhao Y, Zhao L, Xiong B, *et al.* Attention Receptive Pyramid Network for Ship Detection in SAR Images [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2020, **13** **2738**–2756
- [7] Tang G, Zhuge Y, Claramunt C, *et al.* N-YOLO: A SAR Ship Detection Using Noise-Classifying and Complete-Target Extraction [J]. *Remote Sens.* 2021, **13**, 871,
- [8] Girshick R, Donahue J, Darrell T, *et al.* Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]. *Computer Vision and Pattern Recognition*, 2014, 580–587.
- [9] Girshick R. Fast R-CNN [C]. *IEEE International Conference on Computer Vision (ICCV)*, 2015, **1440**–1448
- [10] Ren S, He K, 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
- [11] Liu W, Anguelov D, Erhan D, *et al.* SSD: Single Shot MultiBox Detector [C]. *Computer Vision (ECCV) Lecture Notes in Computer Science*, 2016, . 9905,
- [12] Lin T, Goyal P, Girshick R, *et al.* Focal Loss for Dense Object Detection [C]. *IEEE International Conference on Computer Vision (ICCV)*, 2017, **2999**–3007,
- [13] Bochkovskiy A, Wang C, Liao H. YOLOv4: Optimal Speed and Accuracy of Object Detection [C]. *Computer Vision and Pattern Recognition*, 2020, .
- [14] Tan Mingxing, Pang Ruoming, Quoc V. Le. EfficientDet: Scalable and Efficient Object Detection [C]. *Computer Vision and Pattern Recognition*, 27 Jul 2020.
- [15] Han K, Wang Y, Tian Q, *et al.* GhostNet: More Features From Cheap Operations [C]. *Computer Vision and Pattern Recognition*, 2020, **1577**–1586,
- [16] Yu F, Koltun V. Multi-Scale Context Aggregation by Dilated Convolutions [C]. *International Conference on Learning Representations (ICLR)*, 2016.
- [17] Li L, Xu M, Liu H, *et al.* A Large-Scale Database and a CNN Model for Attention-Based Glaucoma Detection [J]. *IEEE Trans. Med. Imaging*, 2019, **39**, 413 – 424, .
- [18] Woo S, Park J, Lee J, *et al.* CBAM: Convolutional Block Attention Module [C]. *the European Conference on Computer Vision (ECCV)*, 2018.
- [19] Li X, Wang W, Hu X, *et al.* Selective Kernel Networks [C]. *Conference on Computer Vision and Pattern Recognition*, 2019.
- [20] Hu J, Shen L, Albanie S, *et al.* Squeeze-and-Excitation Networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020. **42**, (8):2011–2023, .
- [21] Howard A, Sandler M, Chu G, *et al.* Searching for MobileNetV3 [C]. *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, , **1314**–1324,
- [22] Yang X, Sun H, Fu K, *et al.* Automatic Ship Detection in Remote Sensing Images from Google Earth of Complex Scenes Based on Multiscale Rotation Dense Feature Pyramid Networks [J]. *Remote Sens.* 2018, **10**, 132.
- [23] Wei S, Zeng X, Qu Q, *et al.* HRSID: A High-Resolution SAR Images Dataset for Ship Detection and Instance Segmentation [J]. *IEEE Access*, 2020, , **8**: 120234–120254, .