

复杂背景下 SAR 图像船舶目标检测算法研究

樊海玮, 史 双, 蔺 琪, 孙 欢, 秦佳杰

(长安大学 信息工程学院, 陕西 西安 710064)

摘 要:针对复杂背景 SAR 图像船舶目标检测易受地物干扰影响,导致模型检测率低的问题,提出将结合通道和空间的双注意力机制 CBAM 引入目标检测网络;将膨胀卷积和 concat 特征融合技术应用于目标检测网络中提升模型对小尺寸目标的鲁棒性;为了进一步提高模型的检测速度,使用轻量级 MobileNet 作为基础特征提取网络;同时采用一个新的二分类损失函数使模型训练能够对难易样本设置不同的权重。最后,通过在构建的复杂背景 SAR 图像船舶目标检测数据集 SDATA 上进行实验,实验结果表明该算法在复杂背景 SAR 船舶目标检测中其平均检测精度与综合评价指标 F1-score 值分别为 88.9% 和 91.2%,检测速度达 42.1 fps,从而验证了该模型不仅能够有效提升复杂背景 SAR 图像船舶目标的检测精度,而且在一定程度上提高了目标的检测速度。

关键词:深度学习;目标检测;SAR 图像;双注意力机制;特征融合

中图分类号:TP312

文献标识码:A

文章编号:1673-629X(2021)10-0049-07

doi:10.3969/j.issn.1673-629X.2021.10.009

Research on Ship Target Detection Algorithm in Complex Background SAR Image

FAN Hai-wei, SHI Shuang, LIN Qi, SUN Huan, QIN Jia-jie

(School of Information Engineering, Chang'an University, Xi'an 710064, China)

Abstract: Aiming at the problem that ship detection in complex background SAR images is easily affected by ground objects, which leads to low detection rate of model, we propose to introduce CBAM, which combines channel and space mechanism, into target detection network. The expansion convolution and concatenation feature fusion technology are applied to the target detection network to improve the robustness of model to small targets. In order to further improve the detection speed of model, the lightweight MobileNet is used, and a new two class loss function is used to make the model training set different weights for difficult and easy samples. Finally, experiments are carried out on SDATA, a ship target detection data set with complex background SAR images. It is showed that the mean average precision and comprehensive evaluation index F1-score of the proposed algorithm are 88.9% and 91.2% respectively, and the detection speed is 42.1 fps. It is verified that the proposed model can not only effectively improve the detection accuracy of ship targets in SAR images with complex background, but also improve the detection speed to a certain extent.

Key words: deep learning; target detection; SAR images; dual attention mechanism; feature fusion

0 引 言

SAR 图像船舶目标检测过程中通常存在港口以及近海岸等复杂背景下的船舶目标检测情况,由于船舶紧邻港口受岸上建筑的干扰较为严重,因而容易造成漏检误检的状况。为避免复杂背景下陆地区域对 SAR 图像船舶目标检测的影响,文中提出将注意力机制思想引入目标检测算法中。目前所研究的基于卷积神经网络的复杂背景 SAR 图像船舶目标检测算法多以牺牲检测速度为代价提升目标检测精度,故而在

检测模型大、网络参数多的问题,对于实时性要求较高的应用场景如海难救援、紧急军事部署等都具有一定的局限性。

注意力机制在深度学习领域的应用从本质上与人类的视觉选择性机制类似,是从众多信息中选择出对当前任务最重要的信息。Xu K 等人在 2015 年根据注意力机制关注区域选择的不同,将注意力机制分为软注意力和硬注意力^[1]。其中软注意力机制在注意力分配时,对于输入的每一个区域给出一个 $[0, 1]$ 范围之

收稿日期:2020-08-04

修回日期:2020-12-04

基金项目:陕西高等教育教学改革研究项目(19BY032);长安大学教育教学改革研究项目(300103302403)

作者简介:樊海玮(1974-),男,副教授,CCF 会员(B4120M),研究方向为软件系统设计、深度学习;史 双(1995-),女,硕士,研究方向为深度学习、智能交通。

间的概率值,然后对其进行加权,对于特征图的每个区域都给予关注,只是每个区域的权重会因关注程度的不同而变化,因此其具有参数化,可微化性能。近年来,不断有学者提出将软注意力机制应用于卷积神经网络端到端的模型训练^[2-3]。

1 算法思想

卷积神经网络的快速发展使 SAR 船舶目标检测技术不断提升,目前已有学者将深度卷积神经网络应用到 SAR 船舶目标检测问题上,深度卷积神经网络强大的特征提取能力使得在 SAR 船舶目标检测领域取得了较好的检测结果^[4-6]。特别是深度学习算法不受场景限制,在进行复杂场景 SAR 船舶目标检测时无需进行海陆分割,只需进行数据标注便能够准确学习目标特征。然而,目前所研究的 SAR 图像船舶目标检测普遍存在模型复杂,参数量多的问题。为提高检测精度,不少学者提出进一步增加网络深度,但同时影响了检测速度,在一定程度上限制了 SAR 图像目标检测领域的发展^[7]。

文中结合注意力机制思想对卷积神经网络的特征进行改善以提升模型的检测性能,尤其是对尺寸较小的 SAR 船舶目标及复杂背景下的 SAR 船舶目标。考虑到基于卷积神经网络的两阶段检测算法内存占用大、检测速度慢等缺点,文中选取以一阶段目标检测算法 SSD^[8]为基础,设计出一个新的单阶段目标检测模型 DASSN (dual attention SAR ShipNet)。该模型将 SSD 网络的基础网络 VGG16^[9]替换为利用深度可分离卷积的 MobileNet 网络,有效降低了模型参数,从而提升了模型检测速度,与此同时在模型中引入了空间注意力 (spatial attention, SA) 以及通道注意力 (channel attention, CA) 机制形成双注意力机制模块,增强模型对复杂背景下的 SAR 船舶目标以及小尺寸 SAR 船舶目标检测的鲁棒性;继而将针对 SAR 船舶目标检测问题看为二分类问题的特点,采用了一种适用

于 SAR 船舶目标检测的改进损失函数。

2 基于 DASSN 的复杂背景 SAR 船舶检测算法

目前,深度学习网络模型在 SAR 图像船舶目标检测中的直接应用往往达不到所预期的效果,其原因在于现有的目标检测网络多是基于光学图像的应用。其一,与光学图像相比 SAR 图像较难获得,因而导致在模型训练过程中容易出现数据不平衡的问题;其二, SAR 船舶目标与光学图像目标相比往往尺寸较小,且容易受到杂波及噪声等影响导致图像的分辨程度不高,人眼很难对其进行分类处理;其三,在 SAR 船舶目标检测过程中只涉及船舶和非船舶的二分类问题,而现有的应用于光学图像目标检测网络多是基于多分类问题的。

另一个值得注意的问题是,目前检测性能较好的网络往往模型较大,参数量较多,即使一阶段网络模型的提出已经在较大程度上提升了模型的检测速度,但如果要达到检测精度更高,检测速度更快的效果,仍需要不断对网络模型进行改进和更新。因此,文中提出在一阶段目标检测网络 SSD 的基础上进行改进,生成适合复杂背景和小尺寸 SAR 船舶目标检测的网络模型。

2.1 模型构建

针对复杂背景 SAR 船舶目标检测中检测效果易受地物干扰导致模型检测准确率低的问题,提出将结合通道和空间的双注意力机制引入到目标检测网络中,同时考虑到 SAR 船舶目标尺寸普遍较小的问题,将特征融合技术应用于目标检测网络中来提升模型对小尺寸目标的鲁棒性。为进一步提高模型的检测速度,文中将原 SSD 检测模型的基础网络替换为轻量级的 MobileNet 网络,所提模型 DASSN 网络结构如图 1 所示。

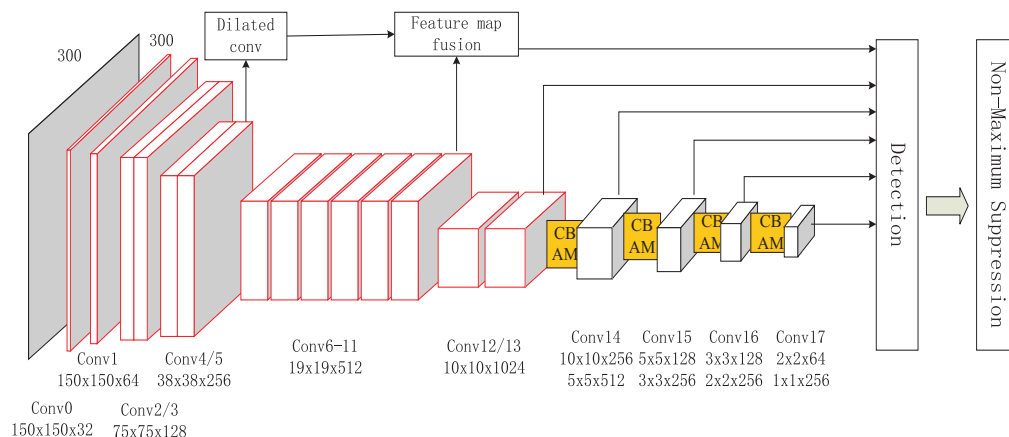


图 1 DASSN 目标检测网络结构

模型的输入定义为 300×300 , 网络模型中前 13 层卷积层为 MobileNet 基础网络。该目标检测模型在 MobileNet 网络后加入 8 个卷积层网络 (conv14_1/2, conv15_1/2, conv16_1/2, conv17_1/2), 并选取其中 4 层作为最终的目标检测特征图, 整个网络总共选择 6 层用于目标检测。所提 DASSN 模型分别在原 conv13、conv14_2、conv15_2、conv16_2 卷积层基础上引入 CBAM (convolutional block attention model) 双注意力机制得到相同尺寸的特征图用于目标检测。值得注意的是, 所提网络引入膨胀卷积 (dilated conv) 先对 conv5 进行处理, 然后与 conv11 进行特征融合后作为最终用于目标检测的特征图。

2.2 基础网络模块

所提目标检测网络选用轻量化 MobileNet 作为基础网络, 该网络的最大特点在于将标准卷积替换为深度可分离卷积 (depthwise separable convolution), 通过引入宽度乘数 (width multiplier) 和分辨率乘数 (resolution multiplier) 两个超参数来减少参数量和计算量。

深度可分离卷积可看作将一个标准卷积分为深度卷积 (D-conv) 和 1×1 的点卷积 (P-conv)。其中, 深度

卷积只卷积输入的一个通道, 与标准卷积相比能够大幅减少网络参数量, 点卷积为传统的卷积运算, 但由于其卷积核的尺寸为 1×1 , 相比于标准卷积的大尺寸卷积核具有较少的参数量。

2.3 双注意力机制模块

文中提出将通道注意力机制与空间注意力机制结合生成的双注意力模块 CBAM 应用到目标检测网络中, CBAM 是一种专门为卷积神经网络所设计的简单有效的注意力模块, 其网络结构如图 2(a) 所示, 分别从通道和空间两个方面生成卷积网络特征图的 attention map, 然后将 attention map 与输入特征图相乘进行特征自适应学习。CBAM 作为一个轻量级的模块能够比较方便地融入卷积网络实现端到端的训练。

(1) 通道注意力机制 (CA)。

模型在特征提取过程中并不是每个卷积核都能起关键性的作用, 甚至某些卷积核处理会对模型造成负面影响, 进而影响模型的检测性能。为了提高模型的检测精度, 文中引入通道注意力机制, 使得网络模型能够有效关注到重要通道, 而忽略甚至抑制负面通道, 其网络结构如图 2(b) 所示。

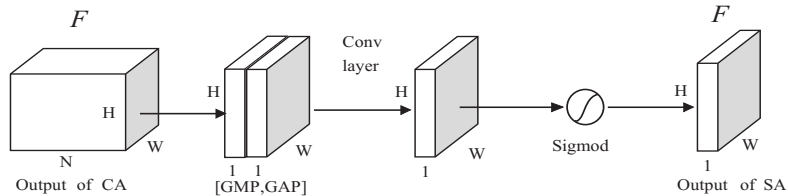
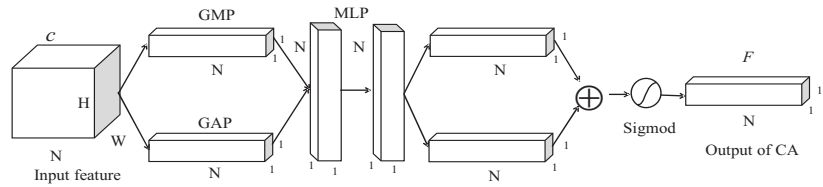
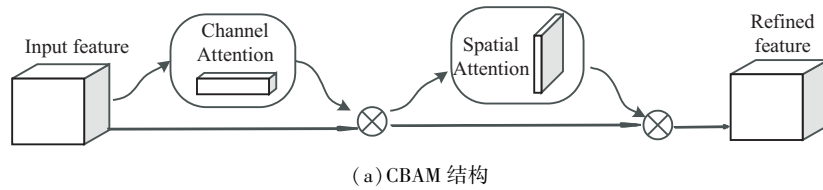


图 2 双注意力机制模块结构示意图

从图 2(b) 中可以看出, 注意力机制的输入特征维度为 $N \times H \times W$, 其中 N 为输入特征的通道数, H 为输入特征图的高, W 为特征图的宽。通道注意力可表示为式 (1)。

$$F_c = \text{sig} \{ \text{MLP} [\text{GAP} (F)] + \text{MLP} [\text{GMP} (F)] \} \quad (1)$$

其中, F 表示输入特征图; GAP 表示全局平均池化; GMP 表示全局最大池化; MLP 表示多层感知机; sig 表示 sigmoid 激活函数,

经过注意力机制处理后得到一个 N 维向量 F_c , 其表达式如下:

$$F_c = (\alpha_1, \alpha_2, \dots, \alpha_N)^T \quad (2)$$

其中, $\alpha_{i(i=1,2,\dots,N)}$ 表示第 i 个通道的重要等级,最后将该向量与输入相乘得到 $N \times H \times W$ 的最终输出,该输出能够有效关注重要通道,抑制非重要通道的影响。

(2) 空间注意力机制(SA)。

SAR 图像在空间各处存在不同价值的信息,例如在复杂场景下停靠于港口的船舶目标,船舶才是重点关注的对象,而港口设备不是需要关注的对象。空间注意力机制的引入能够有效关注到船舶目标而抑制图像中其他非重要信息,从而进一步提高检测精度。

文中采用将空间注意力机制添加在通道注意力机制的输出端,即空间注意力机制的输入为通道注意力机制的输出。所应用的空间注意力机制结构如下所述:先在输入特征图的基础上使用最大池化和平均池化进而得到两个不同的特征描述图,然后通过 concatenation 方式将两个特征描述合并,接着使用卷积操作生成空间注意力特征图,最后经过 sigmoid 函数得到最后空间注意力机制的输出。

假设空间注意力机制的输出表示为 F_s , 其计算过程可表示如下:

$$F_s = \text{sig}\{f^{7 \times 7}[\text{GAP}(F'), \text{GMP}(F')]\} \quad (3)$$

其中, F' 为输入特征图;GAP 和 GMP 与通道注意力机制中的表示意义一致,分别表示全局平均池化和全局最大池化; $f^{7 \times 7}$ 表示 7×7 的卷积操作;sig 仍为 sigmoid 激活函数。

注意力机制处理后得到的输出 F_s 为一个维度为 $1 \times W \times H$ 的矩阵,表示为:

$$F = \begin{pmatrix} \psi_{1,1} & \psi_{1,2} & \cdots & \psi_{1,w} \\ \psi_{2,1} & \psi_{2,2} & \cdots & \psi_{2,w} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{H,1} & \psi_{H,2} & \cdots & \psi_{H,w} \end{pmatrix} \quad (4)$$

其中, $\psi_{i,j}$ ($i = 1, 2, \dots, H, j = 1, 2, \dots, W$) 表示空间坐标为 (i, j) 位置信息的重要等级,最终的输出为该矩阵与输入特征图相乘,空间注意力机制的引入能够成功关注到空间的有效信息。

2.4 多尺度特征融合模块

为了进一步提升模型整体的检测精度,增强算法对小尺寸目标的鲁棒性,同时考虑到基础特征提取层具有大尺寸、高分辨率的特性,能够很好反映目标的位置和信息,故而所提算法将 conv5 引入目标检测网络。conv5 的特征图尺寸为 38×38 ,为能够从 conv5 获得更多的语义信息,文中采用膨胀卷积对 conv5 层特征图进行下采样,卷积核为 3×3 ,滑动步长 $\text{Stride} = 2$, $\text{pad} = 2$,膨胀系数 $\text{dilation} = 2$ 。经过膨胀卷积后 conv5 的特征图尺寸为 $19 \times 19 \times 512$,然后将该特征图与 conv11 进行特征融合,获得最终用于目标检测的特征图。特征

融合方式采用级联融合(concatenation fusion)方式,其具有计算量小、精度高的特点,更加适用于跳层连接融合。

2.5 模型损失函数

原 SSD 网络的损失函数为位置误差与置信度误差的加权和,位置损失是预测框与真实框参数之间的平滑 L1 损失,置信损失为 softmax 损失。对于 SAR 图像船舶目标检测,置信损失转化为 softmax 对船舶以及“其他”两类目标的损失函数,可用标准交叉熵损失函数表示,见公式(5),定位损失函数依然遵循原 SSD 网络的计算方式。

$$L_{\text{conf}} = -\log(p_i) \quad (5)$$

其中, p_i 指标签为船舶类的模型估计概率,但该方法存在不会区分难易样本的问题。在复杂背景 SAR 船舶目标检测样本中,陆地区域中通常会存在与船舶目标相似的干扰物影响 SAR 船舶目标的检测性能,由于 SAR 图像特殊的成像机制导致 SAR 船舶图像中通常存在不同严重程度的斑点噪声,同时对于密集型的多只船舶目标可能会存在相互连接,遮挡等问题,从而使得目标检测模型很难准确检测出目标。为避免复杂场景对 SAR 图像船舶目标检测带来的干扰,文中采用一种新的二分类损失函数作为 SAR 船舶目标检测的置信损失函数,见式(6)。

$$L_{\text{conf}} = -(1 - p_i)^\lambda \log(p_i) \quad (6)$$

为调节简单样本的权重,文中在原始二分类损失函数的基础上加入新的参数 λ ,当 $\lambda = 0$ 时,可表示为式(5),当 p_i 的取值愈接近于 0 时, $-\log(p_i)$ 的取值愈趋近于无穷大,表明该样本属于困难样本且容易错误分类,则当前样本的分类损失函数值将会被赋予较大的权重;当 p_i 取值愈接近 1 时,表明当前样本分类错误的概率越小,样本属于简单样本且容易正确分类,同时将会被给予较小的权重。通过对比分析,将文中所提模型置信损失函数中的样本权重因子 λ 设置为 3,权重因子 λ 的引入能够有效减少简单样本在模型训练过程中所占的比例。

3 实验

3.1 实验数据集及模型训练

为验证所提模型在复杂 SAR 图像中船舶目标的检测情况,文中构建了一个复杂背景下的 SAR 船舶目标检测数据集。该数据集采用目前使用较成熟的 SAR 船舶目标检测数据集 SSDD,SSDD 中共含有 1 160 幅图像,平均每幅图像包含 2.12 条船舶^[5]。实验中通过对原 SSDD 数据集中复杂背景 SAR 图像采用随机裁剪、图像旋转、翻转、平移、大小缩放等方式得到复杂背景 SAR 图像共计 1 150 幅,基于此,另行搜集了复杂背

景 SAR 图像 900 幅,并选取其中部分图像对其利用传统数据增强技术进行数据扩充达 1 200 幅,所搜集的复杂背景 SAR 图像均来自于目前国内所发布的数据,且多基于 Sentinel-1 SAR 数据^[10]。所有数据均按照 VOC 数据格式进行标注,共计 3 250 幅复杂背景 SAR 船舶图像,命名为 SDATA(SAR data),最终按 7:2:1 的比例将数据集分为训练集、验证集及测试集。

模型训练过程中选用 Adam 优化器进行网络参数迭代更新,共训练 200 epochs, batch_size 设置为 16,采用 Poly 机制动态调整学习率。

在模型训练过程中,正负样本的 IOU 阈值设定为 0.5,当 $\text{IOU} > 0.5$ 时,为正样本;当 $\text{IOU} \leq 0.5$ 时,为负样本。图 3 展示了 DASSN 模型与 MobileNet-SSD 网络模型在文中数据集训练过程中 loss 函数值随 epoch 的变化情况,可以看出,相较于 MobileNet-SSD 所提网络模型 DASSN 的收敛速度相对较快,而且 DASSN 网络模型的损失函数值变化曲线相对比较平缓。不难发现两种模型均在 160 epochs 左右损失函数值变化幅度较

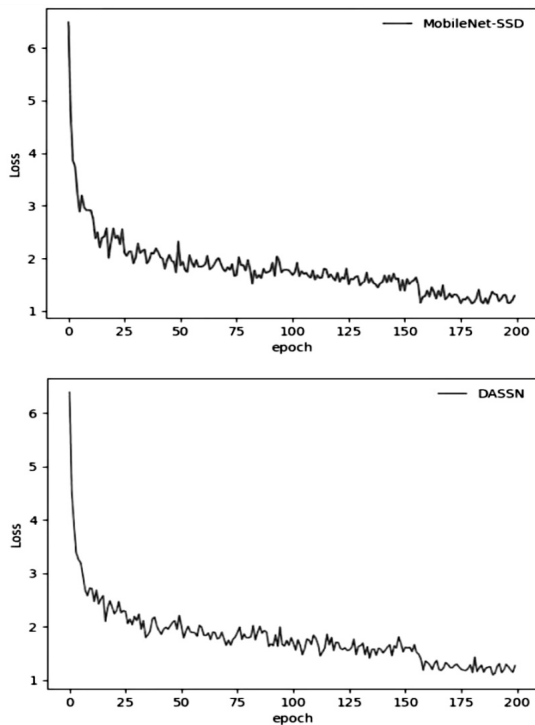


图 3 两种网络损失函数

表 1 不同改进方式对 DASSN 模型检测性能的影响

实验	模型	precision	recall	F1-score
1	MobileNet-SSD	0.784	0.843	0.812
2	MobileNet-SSD+CA	0.832	0.878	0.854
3	MobileNet-SSD+CBAM	0.883	0.912	0.897
4	文中模型	0.891	0.935	0.912

(2) 模型改进前后实验对比。

DASSN 网络模型与原 SSD 网络模型一致,其输入

为明显,而在 160 epochs 之后,DASSN 模型的损失函数变化幅度明显减小,说明此时模型性能也相对稳定,表明两种模型在文中数据集上的训练迭代次数应选取在 175 epochs 左右可使得模型相对比较稳定。

3.2 实验结果及分析

(1) 模型性能分析。

实验过程中置信阈值的选取将会对模型检测性能产生不同的影响,为了解 DASSN 模型在不同置信阈值下的检测性能,分别在不同置信阈值条件下,对模型的综合指标 F1-score 参数进行对比分析。

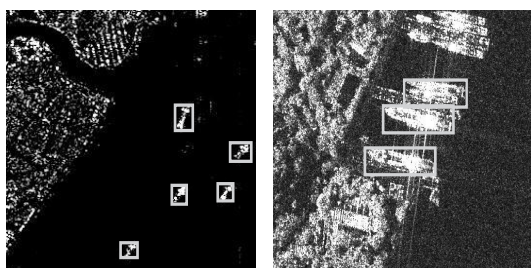
在该模型中,当置信阈值取 0.2 时,模型 F1-score 取得最大值 0.912,当置信阈值不断增大的时候模型综合评价指标 F1-score 的取值不断变小,表示模型的检测性能也在不断下降,故而为取得最好的检测性能,最终将 DASSN 模型的置信阈值设定为 0.2。

为进一步验证文中所提改进方式对网络模型检测性能的影响,分别对模型改进过程中引入不同改进技术所生成的四个不同模型进行实验分析。

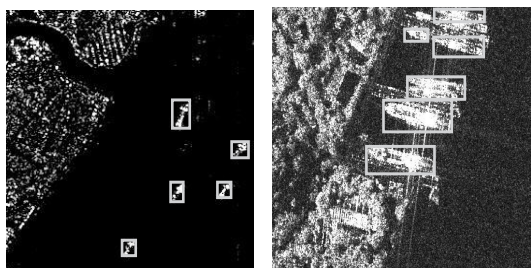
从表 1 分析可得,模型改进方式的提出对模型的检测性能均有所改善。MobileNet-SSD 网络由于其轻量化的特点,在一定程度上损失了检测精度,因而其检测准确率只达到了 78.4%。在加入通道注意力机制后,模型的检测准确率相较于 MobileNet-SSD 网络提升了 4.8 个百分点。在引入空间注意力机制后,由于模型的双注意力机制使得模型检测准确率有了较大的提升,相较于最初的 MobileNet-SSD 网络其检测准确率提升了 9.9 个百分点,其 F1-score 值也增长至 0.897,相较于模型二提升了 4.3 个百分点,表明引入双注意力机制的检测效果远比引入单通道注意力机制更为明显。实验表明文中最终模型检测性能最优,其检测精度达到 0.891,相较于模型三检测精度有所提升,表明引入的特征融合技术能够提升对 SAR 船舶目标的检测性能,相较于原 MobileNet-SSD 网络其检测精度提升了 10.7 个百分点,F1-score 值也增长了 10.0 个百分点,表明文中网络模型对复杂背景 SAR 船舶目标检测性能有了很大的提升。

图像尺寸均为 300×300 。为了进一步体现文中网络对复杂背景 SAR 船舶目标的检测性能,实验在所提数据

集 SDATA 上对 SSD300 网络及 DASSN 网络进行对比。图 4 分别为 SSD300 网络及 DASSN 网络在复杂背景 SAR 图像船舶目标检测的结果示意图。



(a) SSD300 网络检测结果



(b) DASSN 网络检测结果

图 4 两种网络检测结果对比

为了清晰地对两种网络结构的性能进行对比,分别从平均检测精度、查准率、查全率及 F1-score 四个性能指标对两种模型进行比较。

表 2 模型改进前后性能对比

网络	mAP/%	precision	recall	F1-score
SSD300	80.1	0.822	0.854	0.838
DASSN	88.9	0.891	0.935	0.912

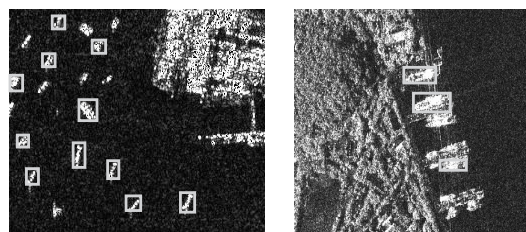
从表 2 可以看出,改进后的 DASSN 模型在平均检测精度上相较于 SSD300 网络提升了 8.8 个百分点,模型查准率与查全率分别提升了 6.9 和 8.1 个百分点,同时综合评价指标 F1-score 也提升了 7.4 个百分点,表明文中所提模型在复杂背景 SAR 图像船舶目标检测上能够较大程度地提高检测性能。

(3) 与其他目标检测算法对比。

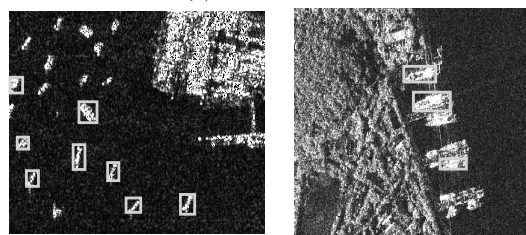
为验证文中所提算法的有效性,选取目前常用的深度学习目标检测算法 Faster RCNN^[11]、YOLOv3^[12]及 RetinaNet^[13],对包括文中算法在内的四种目标检测算法均在所提复杂背景 SAR 船舶图像数据集上进行实验。

图 5 分别展示了四种不同模型对两幅 SAR 图像的检测情况,两幅图像均具有复杂的背景且会对 SAR 船舶目标检测进行干扰,可以看出除所提模型外,RetinaNet 网络具有相对较好的检测性能,但其检测效果相对于所提模型仍具有一定的空间。YOLOv3 作为 YOLOv2^[14]算法的升级版,其检测精度及检测速度都得到了一定程度的提高,但对尺寸较小的密集型船舶

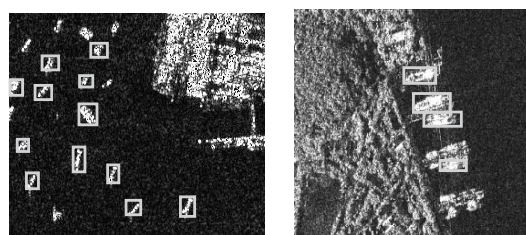
目标存在许多漏检情况,并且对紧密连接的船舶检测效果也不理想。所提模型虽然在检测过程中存在部分漏检的情况,但相较于其他网络的检测性能均具有明显的提升。对于密集连接的海岸船舶,由于其存在遮挡粘连等因素导致模型检测困难,但文中模型仍能够表现出较好的检测性能。



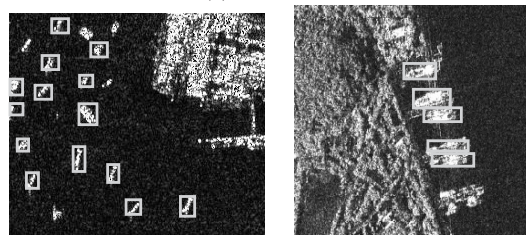
(a) Faster RCNN



(b) YOLOv3



(c) RetinaNet



(d) DASSN

图 5 不同检测模型的检测结果对比

通过查全率、查准率、F1-score 以及检测速度四个方面对各个算法进行比较,其检测结果如表 3 所示。

表 3 不同算法的检测性能对比

算法	precision	recall	F1-score	FPS
Faster RCNN	0.817	0.866	0.841	5.3
YOLOv3	0.804	0.841	0.822	37.3
RetinaNet	0.831	0.873	0.851	16.5
文中算法	0.899	0.935	0.912	42.1

从表 3 可以看出,Faster RCNN 网络具有相对较高的检测精度,但明显可以看出其检测速度具有较大的劣势,其 FPS 仅达到 5.3;而 RetinaNet 相对 Faster RCNN 与 YOLOv3 具有较好的检测性能,其查准率较

一阶段目标检测网络 YOLOv3 提升了 2.9 个百分点,但其检测速度相对 YOLOv3 有了明显的降低,相对于 Faster RCNN 其查准率提升了 1.4 个百分点,检测速度 FPS 也提升了 11.2,但相较于文中所提模型其查准率降低了 6.8 个百分点,文中算法的 F1-score 值比其他三种算法分别提升了 7.1、9.0 及 6.1 个百分点,表明所提模型较其他三种模型在复杂背景 SAR 船舶目标检测数据集中表现最佳,同时值得注意的是所提模型的检测速度相较于速度最快的 YOLOv3 模型也提升了 4.8 个百分点,表明所提算法在复杂背景 SAR 船舶目标检测上无论是检测精度还是检测速度都得到了一定的提升,对将深度学习算法应用于实际的 SAR 船舶目标检测环境具有一定的推动作用。

4 结束语

针对复杂场景下 SAR 图像船舶目标检测问题,提出了一种具有较高检测速度和检测精度的 SAR 图像船舶目标检测算法,该算法具有较少的参数量,相对于原 SSD 网络模型更加轻量化。算法为解决复杂场景易造成干扰的问题,将结合通道注意力与空间注意力机制的双注意力模块 CBAM 引入网络模型,同时考虑到 SAR 图像中船舶目标多数尺寸较小,故而文中算法引入基于膨胀卷积的特征融合技术,在不增大输入特征图的基础上使用膨胀卷积增大感受野达到丰富特征数据的目的。最终,通过 3 种实验对比,验证了文中所提模型对复杂背景下 SAR 船舶目标检测的有效性,同时也证明了文中算法对实时性 SAR 应用领域具有一定的现实意义。

参考文献:

- [1] XU K, BA J, KIROS R, et al. Show, attend and tell: neural image caption generation with visual attention[C]//Proceedings of 2015 international conference on machine learning. Lille, France: JMLR. org, 2015: 2048–2057.
- [2] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks[C]//Advances in neural information processing systems. Montreal, QC, Canada: [s. n.], 2015: 2017–2025.
- [3] CHEN L C, YANG Y, WANG J, et al. Attention to scale: scale-aware semantic image segmentation[C]//Proceedings of 2016 IEEE conference on computer vision and pattern recognition. New York, NY: IEEE, 2016: 3640–3649.
- [4] KANG M, JI K, LENG X, et al. Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection[J]. Remote Sensing, 2017, 9(8): 860–874.
- [5] 李健伟, 曲长文, 彭书娟, 等. 基于卷积神经网络的 SAR 图像舰船目标检测[J]. 系统工程与电子技术, 2018, 40(9): 1953–1959.
- [6] JIAO J, ZHANG Y, SUN H, et al. A densely connected end-to-end neural network for multiscale and multiscale SAR ship detection[J]. IEEE Access, 2018, 6(1): 20881–20892.
- [7] 李健伟, 曲长文, 彭书娟, 等. 基于生成对抗网络和线上难例挖掘的 SAR 图像舰船目标检测[J]. 电子与信息学报, 2019, 41(1): 143–149.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Computer vision – ECCV 2016. Amsterdam, Netherlands: Springer, 2016: 21–37.
- [9] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//Proceedings of international conference on learning representations. San Diego, CA, United states: ICLR, 2015: 1–14.
- [10] WANG Y Y, WANG C, ZHANG H. Combining a single shot multibox detector with transfer learning for ship detection using sentinel-1 SAR images[J]. Remote Sensing Letters, 2018, 9(8): 780–788.
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//International conference on neural information processing systems. Istanbul, Turkey: MIT Press, 2015: 1137–1149.
- [12] REDMON J, FARHADI A. YOLOv3: an incremental improvement[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City, UT, United States: IEEE, 2018.
- [13] LIN T, GOYAL P, GIRSHICK R B, et al. Focal loss for dense object detection[C]//IEEE international conference on computer vision. New York, NY: IEEE, 2017: 2999–3007.
- [14] REDMON J, FARHADI A. YOLOv4: better, faster, stronger[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu, HI, United States: IEEE, 2017: 6517–6525.