

# 基于注意力机制特征融合网络的SAR图像飞机目标快速检测

赵 琰, 赵凌君, 匡纲要

(国防科技大学电子信息系统与复杂电磁环境效应国家重点实验室, 湖南长沙 410073)

**摘要:** 针对合成孔径雷达(Synthetic Aperture Radar, SAR)图像中飞机目标散射点离散化程度高, 周围背景干扰复杂, 现有算法对飞机浅层语义特征表征能力弱等问题, 本文提出了基于注意力特征融合网络(Attention Feature Fusion Network, AFFN)的SAR图像飞机目标检测算法. 通过引入瓶颈注意力模块(Bottleneck Attention Module, BAM), 本文在AFFN中构建了包含注意力双向特征融合模块(Attention Bidirectional Feature Fusion Module, ABFFM)与注意力传输连接模块(Attention Transfer Connection Block, ATCB)的注意力特征融合策略并合理优化了网络结构, 提升了算法对飞机离散化散射点浅层语义特征的提取与判别. 基于自建的Gaofen-3与TerraSAR-X卫星图像混合飞机目标实测数据集, 实验对AFFN与基于深度学习的通用目标检测以及SAR图像特定目标检测算法进行了比较, 其结果验证了AFFN对SAR图像飞机目标检测的准确性与高效性.

**关键词:** 注意力机制; 特征融合; 飞机目标快速检测; SAR图像; 卷积神经网络

中图分类号: TN391

文献标识码: A

文章编号: 0372-2112(2021)09-1665-10

电子学报URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20200486

## Attention Feature Fusion Network for Rapid Aircraft Detection in SAR Images

ZHAO Yan, ZHAO Ling-jun, KUANG Gang-yao

(State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System, National University of Defense Technology, Changsha, Hunan 410073, China)

**Abstract:** Aiming at the problems of high discretization of aircraft's backscattering points, complex background interference of surroundings in Synthetic Aperture Radar (SAR) images and weak representation of shallow semantic features of aircraft by existing algorithms, an Attention Feature Fusion Network (AFFN) was proposed for aircraft detection in SAR images. By introducing Bottleneck Attention Module (BAM), this article constructed an attention feature fusion strategy consisting of Attention Bidirectional Feature Fusion Module (ABFFM) and Attention Transfer Connection Block (ATCB) in AFFN, and rationally optimized the network structure so as to strengthen the abilities of extracting and discriminating shallow semantic features of aircraft. Based on a self-built Gaofen-3 and TerraSAR-X mixed aircraft dataset, AFFN was compared with several CNN-based general object detection methods and methods designed for specific objects in SAR images. The experimental results illustrated the accuracy and effectiveness of our method for aircraft detection in SAR images.

**Key words:** attention feature fusion network (AFFN); feature fusion; rapid aircraft detection; synthetic aperture radar (SAR) images; convolutional neural network (CNN)

### 1 引言

合成孔径雷达(Synthetic Aperture Radar, SAR)因具备全天时、全天候的对地观测能力, 被广泛应用于军事与民用领域, 相应的解译技术也取得了迅速发展<sup>[1-4]</sup>.

飞机目标具有高价值、时敏性等特点, 对其准确、快速的检测可为军事决策提供有力的信息支撑. 当前, SAR图像飞机目标检测算法可分为基于灰度统计建模的传统算法与基于语义特征的机器学习算法. 传统算法<sup>[5-8]</sup>

首先通过对 SAR 图像灰度特征统计建模以获取目标潜在位置,再根据目标几何、散射等特征设计鉴别算子,对飞机目标进一步鉴别.然而,由于飞机目标姿态敏感性强、散射点分布离散化程度高且易受周围环境干扰,目标特征提取的偏差易导致算法性能大幅降低,且复杂的控制参量也使该类算法对复杂场景的适应性较低.基于语义特征的机器学习算法以基于卷积神经网络(Convolutional Neural Network, CNN)的算法为主.文献[9, 10]分别基于视觉显著性与 Canny 算子从 SAR 图像中截取目标切片,再以训练后的 CNN 对切片中目标二次鉴别.文献[11]首先以恒虚警(Constant False Alarm Rate, CFAR)检测器筛选出潜在的飞机区域,再对该区域中的飞机进行鉴别与位置回归.基于 RetinaNet<sup>[12]</sup>网络,文献[13]设计了多分支空洞卷积模块并利用卷积注意力模块对提取的语义特征进一步优选.文献[14]设计了基于部件的飞机目标检测算法.得益于对目标语义信息的强大表征,基于 CNN 的算法较传统算法显著提升了对 SAR 图像中飞机目标的检测能力.由于飞机在 SAR 图像中姿态敏感性强,散射点分布高度离散化(如图 1 所示),浅层特征图富含目标局部、细节多种特征.然而,以上基于 CNN 的算法主要关注对目标高层语义信息的提取,而对目标浅层信息的挖掘、利用较为有限,这弱化了算法对目标的准确判别.

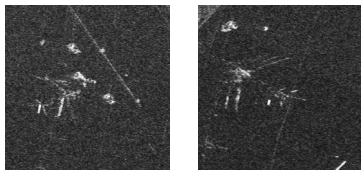


图1 SAR 图像飞机散射点分布离散化

针对上述问题,本文提出了一种基于注意力机制特征融合网络(Attention Feature Fusion Network, AFFN)的端到端的 SAR 图像飞机目标检测算法.算法核心为由注意力双向特征融合模块(Attention Bidirectional Feature Fusion Module, ABFFM)与注意力传输连接模块(Attention Transfer Connection Block, ATCB)构成的注意力特征融合策略.其中,ABFFM 可将低、中、高层语义特征充分融合并优选,强化对飞机多层语义特征的提取与判别.ATCB 主要对浅层语义信息进行优选,并对周围环境干扰进行抑制.基于自建的 Gaofen-3 与 TerraSAR-X 卫星图像混合飞机目标实测数据集,实验将 AFFN 与基于深度学习的通用目标检测以及 SAR 图像特定目标检测算法进行了对比,其结果验证了 AFFN 对 SAR 图像飞机目标检测的准确性与高效性.

## 2 注意力机制特征融合网络

### 2.1 网络整体结构

AFFN 的结构如图 2 所示.在自底向上的前向通路中,以 VGG-16 网络从输入 SAR 图像中提取基础特征,并抽取网络 conv4\_3、conv5\_3 以及 conv7 卷积层的输出进行融合,其特征图空间尺寸分别为输入图像的 8、16、32 倍下采样.随后,这三层特征图被送入锚点细化模块以滤除负样本锚点框,并对正样本锚点框的位置初步修正.在构建特征金字塔时,由于 conv7 层的语义特征与浅层特征差异较大,仅采用由多层卷积构成的传输层构建 Feat\_3 特征图,其尺寸为 10×10 像素.为强化对飞机散射点浅层语义信息的判别,通过引入瓶颈注意力模块(BAM),AFFN 设计了由 ABFFM 与 ATCB 构成的注意力特征融合策略.其中,ABFFM 可将 conv4\_3、conv5\_3 以及 Feat\_3 特征图融合并提纯,得到尺寸为 20×20 像素的精细化中层特征图 Feat\_2. ATCB 则可融合并优选 Feat\_2 与 conv4\_3 层特征,得到尺寸为 40×40 像素的浅层特征图 Feat\_1.采用目标检测模块在 Feat\_1、Feat\_2 以及 Feat\_3 精细化特征图上对目标进行鉴别与位置预测,并经非极大值抑制等后处理操作,可得飞机目标最终检测结果.算法以交叉熵(Cross Entropy, CE)与 smooth-L1 损失分别优化锚点细化模块(Anchor Refinement Module, ARM)与目标检测模块(Object Detection Module, ODM)中的鉴别与位置回归子网络,并在线困难样本挖掘(Online Hard Example Mining, OHEM)技术,提升算法对困难样本的关注度.

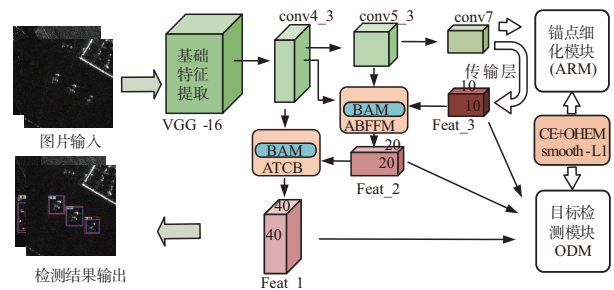


图2 注意力特征融合网络整体结构

AFFN 与 RefineDet 算法的相似之处在于两者均构建了精细化特征金字塔,采用 ARM 与 ODM 对目标多级鉴别与位置预测,以相同的损失函数对算法优化.然而, AFFN 中基于 BAM 模块所特别设计的注意力特征融合策略,可有效提升算法对 SAR 图像飞机目标散射点浅层语义信息的挖掘与利用.

### 2.2 注意力特征融合策略

#### 2.2.1 瓶颈注意力模块

瓶颈注意力模块(BAM)由并行的通道与空间注意力分支组成,如图 3 所示.

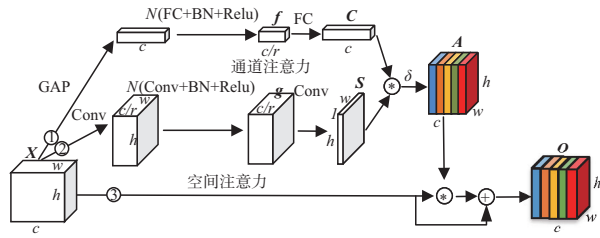


图3 BAM结构

(1)通道注意力. 由于特征图不同通道对SAR图像信息表征各异,通道数为 $c$ 的输入特征图 $X$ 经全局平均池化(Global Avg Pooling, GAP)以及 $N$ 组“全连接层(FC)-批归一化层(BN)-Relu激活函数”的映射,可得通道数为 $c/r$ 的编码后特征 $f$ . 该特征经全连接层(FC)的解码可得与输入特征 $X$ 具有相同通道数的通道注意力权重 $C$ . (2)空间注意力. 由于SAR图像飞机目标的散射点分布较为离散且周围地物干扰较强, $X$ 经卷积(convolution)以及 $N$ 组“卷积层(Conv)-批归一化层(BN)-Relu激活函数”的映射,可得通道数为 $c/r$ 的空间注意力特征 $g$ . 该特征经卷积映射,可得与 $X$ 具有相同空间尺寸,通道数为1的空间注意力权重 $S$ . 最终, $C$ 与 $S$ 相乘并经Sigmoid函数( $\delta$ )的非线性映射,可得联合注意力权重 $A$ . 以权重 $A$ 与 $X$ 逐元素相乘并构建残差连接,可得加权后输出特征图 $O$ .

2.2.2 注意力双向特征融合模块

在构建中层精细化特征图 Feat\_2 时,引入了 conv4\_3 浅层特征图,设计了图4所示的注意力双向特征融合模块(ABFFM).

ABFFM 由特征降采样分支1、特征变换分支2以及上采样分支3组成. 分支1、2中的BN层可对不同尺度

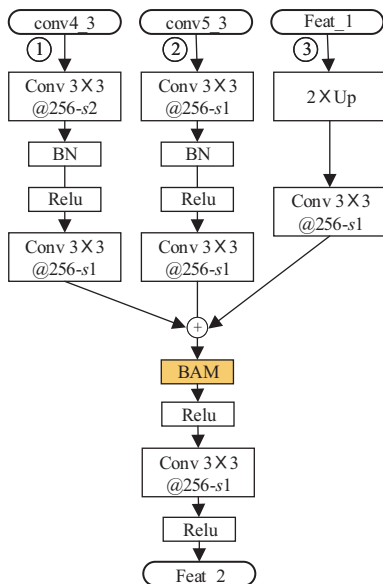


图4 ABFFM结构

的特征归一化处理,强化特征融合的稳定性并加速网络收敛. 分支3以双线性插值与卷积结合代替反卷积操作对 Feat\_3 特征图上采样,可一定程度避免高维特征的网格化分布. conv4\_3、conv5\_3、Feat\_1 特征分别经这三路分支并逐元素相加,可得通道数为256维的精细化融合特征. 为进一步凸显目标关键信息,抑制背景干扰,ABFFM 引入了BAM(图4黄色框所示,其具体原理见2.2.1小节)对特征进行优选. 最终经卷积层、激活函数对优选后特征的映射,可得 Feat\_2 精细化特征图.

2.2.3 注意力传输连接模块

在 RefineDet 算法中,conv4\_3 与 Feat\_2 层特征经图5所示的传输连接模块(TCB)可得浅层特征图 Feat\_1. 然而,由于缺乏特征优选机制,Feat\_1 中可能包含较多环境散射信息,对目标自身特征造成干扰. 为此,论文将BAM引入到原始TCB模块中,并以“上采样-卷积”代替反卷积操作并引入BN层对特征归一化处理,构建了注意力传输连接模块(ATCB). ATCB的结构如图6所示,对原始TCB的改进以黄色框标出.

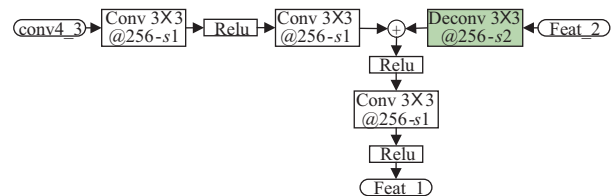


图5 TCB结构

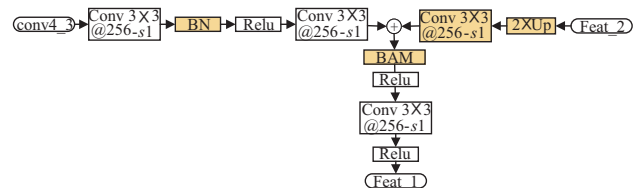


图6 ATCB结构

2.3 其他模块

由于飞机目标在SAR图像中占比较小,较低重叠度的锚点框与真值易导致正负锚点框数比例失衡,且SAR图像飞机的姿态敏感度较高,直接对目标最终位置进行预测易使结果与真值偏差较大. 为平衡算法的检测精度与速度,AFFN以 RefineDet 的锚点细化模块(ARM)和目标检测模块(ODM)对飞机目标进行多级鉴别与位置预测. 在AFFN中,conv4\_3、conv5\_3以及conv7特征图被送入ARM以滤除大量负样本,并对正样本锚点框的位置初步回归,Feat\_1、Feat\_2与Feat\_3特征图则被送入ODM以对目标进行精细化的位置回归与鉴别.

2.4 损失函数

AFFN 采用交叉熵损失并结合在线困难样本挖掘

技术来优化 ARM 与 ODM 的鉴别分支. 交叉熵损失函数如式(1)所示, 其中,  $N$  为一个训练批次中的切片数,  $y_i$  与  $p_i$  分别为第  $i$  个切片中飞机目标的真值以及网络对该切片中飞机目标的预测输出. 采用 smooth-L1 损失优化 ARM 与 ODM 的位置回归分支. smooth-L1 损失如式(2)所示,  $x$  为网络预测值与真值的偏差.

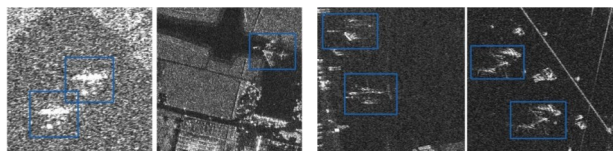
$$L_{\text{cls}} = -\frac{1}{N} \sum_i y_i \log(p_i), \quad i = 1, 2, 3, \dots, N \quad (1)$$

$$\text{smooth-L1} = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (2)$$

### 3 实验与分析

#### 3.1 实验数据集

由于目前尚无公开的 SAR 图像飞机目标检测数据集, 因此本文采用自建数据集对 AFFN 进行了评测. 该数据集包含来自 Gaofen-3 与 TerraSAR-X 卫星的 119 景覆盖全球 23 座机场的原始 L2 级 SAR 图像. 其中, 像元分辨率约为 0.5m 与 1.0m 的图像数分别有 72 景与 47 景. 经判读专家的标注并对图像裁剪, 共获取了 2136 张含飞机目标的像素尺寸为 320×320 的切片图像, 共包含机身长度约在 25m 到 70m 之间的飞机目标 3776 架. 论文以 7:3 的比例将数据集划分为训练集与测试集, 并采用对比度变换、旋转、镜像对称、尺度扩展等技术对训练集进行扩增. 图 7 展示了该数据集的部分切片以及相应的飞机目标真值. 其中(a)组切片来自 Gaofen-3 号卫星图像, (b)组切片来自 TerraSAR-X 卫星图像.



(a)来自 Gaofen-3 的图像切片 (b)来自 TerraSAR-X 的图像切片

图 7 自建 SAR 图像飞机数据集, 蓝色框为目标真值框

#### 3.2 网络超参数设置

算法自身超参数以及网络训练超参数设置如表 1 所示. 为稳定网络训练并防止梯度爆炸, 采用 Warmup 技术在前 4 个周期对算法学习率动态增加. 网络每隔 10 个周期对模型参数进行保存, 并在第 100 个周期将学习率衰减为先前的 0.1 倍, 以进一步对模型进行调优. 算法采用 Pytorch 1.5.0 深度学习框架搭建, 在配有 CUDA10.2 计算库的机器上完成测试.

#### 3.3 评价指标

为充分评测模型性能, 实验采用精确率( $P$ )、召回率( $R$ )、 $F1$  分数以及平均精度(AP)指标来衡量算法检测性能. 采用帧率(FPS)、模型参数量(Params)、浮点运

表 1 网络超参数设置

网络	锚点框尺寸	32, 64, 128
	锚点框比例	1, 1.414, 0.707
	$r$ (BAM)	16
训练	迭代总次数(Epochs)	200
	小批次(Minibatch)	12
	初始学习率	5e-4
	优化器	Adam
	权重衰减值	1e-4

算量(FLOPs)来衡量算法运算速度、资源占用量以及计算复杂度. 此外, 实验绘制了不同算法的准确率-召回率(Precision-Recall, PR)曲线并展示了飞机目标检测结果图. 精确率与召回率如式(3)、(4)所定义, 其中 TP、FP、GT 分别代表算法正确、错误检测以及真值目标数.

$$P = \text{TP}/(\text{TP} + \text{FP}) \quad (3)$$

$$R = \text{TP}/\text{GT} \quad (4)$$

$F1$  是  $P$  与  $R$  的调和评价指标, 如式(5)定义.

$$F1 = 2PR/(P + R) \quad (5)$$

AP 为精确率-召回率曲线(PR)与坐标轴所围面积值, 如式(6)定义.

$$\text{AP} = \int_0^1 P(r) dr \quad (6)$$

FPS 衡量了算法的检测速度, 如式(7)所定义, 其中  $t$  为算法处理每张图像的耗时.

$$\text{FPS} = 1/t \quad (7)$$

模型参数量(Params)为网络含参网络层的参数总量, 其衡量了模型的空间资源占用量, 如式(8)所定义. 其中,  $K_l$ 、 $C_{l-1}$ 、 $C_l$  分别为卷积层尺寸、输入与输出通道数.

$$\text{Params} = \sum_{l=1}^D K_l^2 \times C_{l-1} \times C_l \quad (8)$$

浮点运算量(FLOPs)衡量了模型浮点运算次数, 反映了模型计算复杂度, 如式(9)所定义, 其中  $D$ 、 $M_l$ 、 $K_l$ 、 $C_{l-1}$ 、 $C_l$  分别代表网络总层数, 第  $l$  层输出特征图的边长, 卷积核尺寸以及输入以及输出通道数.

$$\text{FLOPs} = \sum_{l=1}^D M_l^2 \times K_l^2 \times C_{l-1} \times C_l \quad (9)$$

#### 3.4 结果与分析

实验将 AFFN 与基于 CNN 的通用以及 SAR 图像特定目标检测算法进行了对比, 结果如表 2 所示. 对比通用目标检测算法, 在检测精度方面, 除 AFFN 的检测精确率( $P$ )较 LRF 低, 为 0.856 外, 算法的召回率( $R$ )、 $F1$  分数以及 AP 值均高于其他对比算法, 分别为 0.873、0.864 以及 0.871. 在检测速度与模型参数方面, AFFN 的测试速度为 91FPS, 模型参数量以及浮点运算量分别为 32.190M 与 39.215G, 较所对比的两阶段算法、单阶

表2 不同算法对SAR图像飞机目标测试结果

算法类别	网络名称	<i>P</i>	<i>R</i>	<i>F1</i>	AP	Params(M)	FLOPs (G)	FPS
通用目标 两阶段算法	Faster-RCNN <sup>[16]</sup>	0.736	0.790	0.762	0.749	136.689	149.228	14
	Cascade-RCNN <sup>[17]</sup>	0.707	0.733	0.720	0.697	304.076	179.934	8
	FPN <sup>[18]</sup>	0.778	0.810	0.794	0.814	120.690	134.150	16
通用目标 单阶段算法	SSD <sup>[19]</sup>	0.798	0.778	0.788	0.820	<b>23.801</b>	<b>30.570</b>	<b>250</b>
	RFB <sup>[20]</sup>	0.854	0.764	0.806	0.838	31.471	34.422	100
	LRF <sup>[21]</sup>	<b>0.899</b>	0.740	0.812	0.809	65.774	48.114	91
	YOLOv3 <sup>[22]</sup>	0.502	0.791	0.614	0.630	61.529	32.687	15
	RefineDet <sup>[15]</sup>	0.856	0.806	0.830	0.843	33.915	37.426	90
	RetinaNet <sup>[12]</sup>	0.837	0.764	0.799	0.789	135.366	111.994	28
SAR图像 特定目标 检测算法	PADN <sup>[13]</sup>	0.820	0.777	0.798	0.806	102.163	123.562	33
	DAPN <sup>[23]</sup>	0.789	0.830	0.809	0.831	128.230	167.733	12
	文献[9]	0.520	0.315	0.395	0.189	4.314	0.328	0.5
	文献[10]	0.544	0.238	0.331	0.132	6.158	11.661	0.5
	文献[11]	0.706	0.235	0.353	0.200	134.309	37.279	0.4
	文献[14]	0.644	0.807	0.716	0.741	102.357	97.549	2
本文算法	AFFN	0.856	<b>0.873</b>	<b>0.864</b>	<b>0.871</b>	32.190	39.215	91

段算法(如 LRF, RetinaNet 等)性能更优;对比 SAR 图像特定目标检测算法, AFFN 较基于深度学习的算法(如 PADN、DAPN)以及深度与传统特征相结合的算法(如文献[9~11, 14])优势更为显著。

由此可得如下结论:

(1)注意力特征融合策略强化了算法对飞机离散化散射点多维特征的表征与判别能力。对比通用目标检测算法, AFFN 的 ABFFM 模块在构建 Feat\_2 时引入了目标细粒度特征的浅层语义信息并对融合特征后优选。ATCB 可突出目标浅层语义特征, 并抑制浅层特征图中的周围地物干扰。对比 SAR 图像特定目标检测算法, AFFN 较文献[9~11]可提取目标多维度语义信息, 较 PADN、DAPN 更加关注目标浅层特征。虽然文献[14]以飞机目标组件提取与匹配对目标精细化鉴别, 但多样化的飞机散射点分布与复杂的背景干扰一定程度降低了算法的泛化能力。

(2)合理的特征金字塔结构可有效平衡不同维度语义信息, 提升算法对多尺度目标的表征。由于飞机自身尺寸以及像元分辨率的差异, SAR 图像飞机目标具有一定的多尺度特性。相较于 FPN、DAPN、YOLOv3 复杂的金字塔网络结构, AFFN 构建了仅包含 Feat\_1、Feat\_2 与 Feat\_3 的特征金字塔, 使目标不同维度语义信息更加均衡。虽然 Faster-RCNN、Cascade-RCNN 采用多尺度锚点框策略对目标进行回归, 但其仅采用了骨干网络的高层特征图, 对浅层语义特征并未充分利用。SSD、RFB 缺乏自顶向下的特征金字塔结构, 导致目标浅层与高层语义信息的关联、互补能力较差。

(3)目标多级检测策略可有效滤除负样本, 提升算

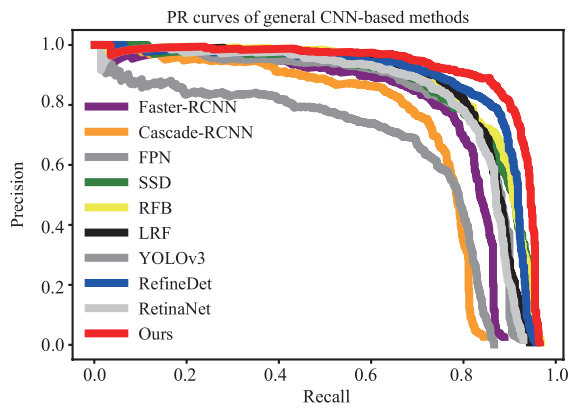
法处理效率与检测精度。相较于其他方法, AFFN 通过引入 ARM 与 ODM, 构建了目标多级检测机制。ARM 对负样本锚点框的预筛选和对正样本锚点框的初步回归, 可降低 ODM 对目标的搜索空间, 提升了算法对目标位置回归与鉴别的准确性。

(4)端到端的处理流程与精简的网络结构降低了模型参数数量与计算复杂度, 保证了算法的检测速度。与通用算法相比, AFFN 并未含有两阶段算法典型的特征池化操作(如 ROI Pooling、ROI Wrapping、ROI Aligning), 避免了池化特征后处理操作导致的计算负担。此外, AFFN 合理设计的模块与网络结构, 使其具备不逊于单阶段算法的检测速度与准确性; 与 SAR 图像特定算法相比, AFFN 较文献[9~11, 14]具备端到端的处理能力, 较 PADN、DAPN 的网络结构更精简, 模型参数数量更少、计算复杂度更低。

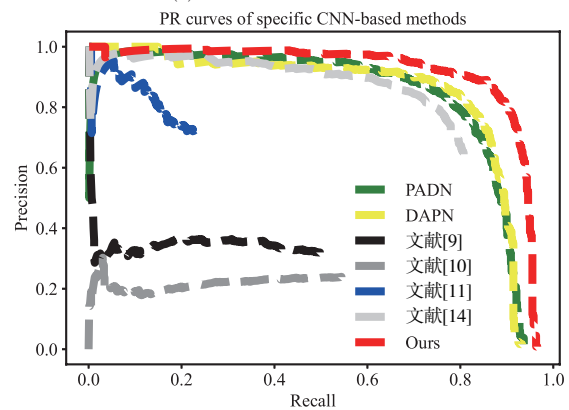
图 8(a)、(b)两图分别为 AFFN 与通用目标检测算法以及与 SAR 图像特定算法检测结果的精确率-召回率曲线对比图。不难看出, 随召回率提升(算法检测到真实飞机数增加), AFFN(红色曲线)性能较其他算法变化更稳定, 这也表明了 AFFN 较其他算法对目标与虚警的判别能力更强。

图 9~11 展示了不同算法对三种典型场景中飞机目标的检测结果, 其中蓝色框为目标真值, 红色框为算法的目标预测值, 不难看出:

(1)AFFN 可有效对散射点强度较弱且离散化程度高的飞机目标准确检测。在图 9 中, 飞机强散射点分布较为离散, 整体散射强度较弱, 在机翼等部位呈现多次散射且周围环境(如廊桥)散射干扰较强。在不同算法



(a) 与通用算法曲线对比



(b) 与SAR图像特定算法曲线对比

图8 AFFN与不同算法检测结果的准确率-召回率(PR)曲线对比图

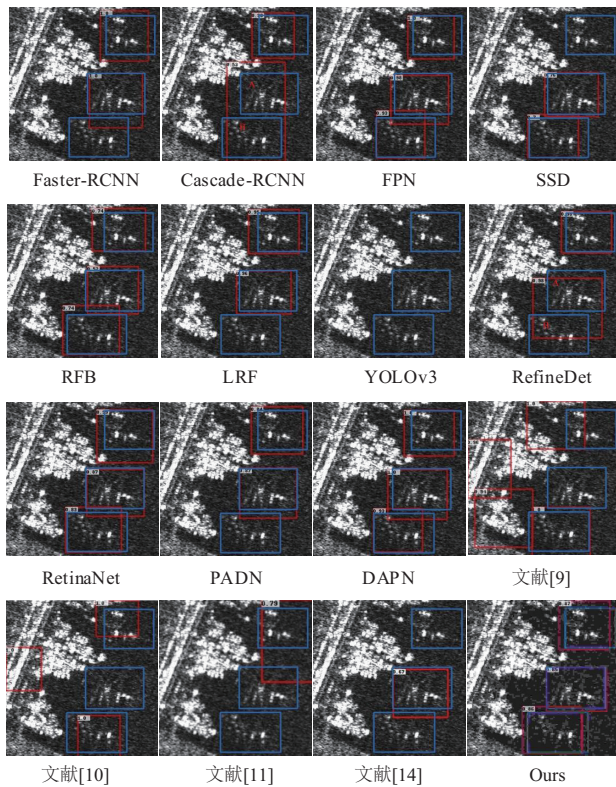


图9 不同算法对散射点强度较弱且离散化程度高的飞机检测结果

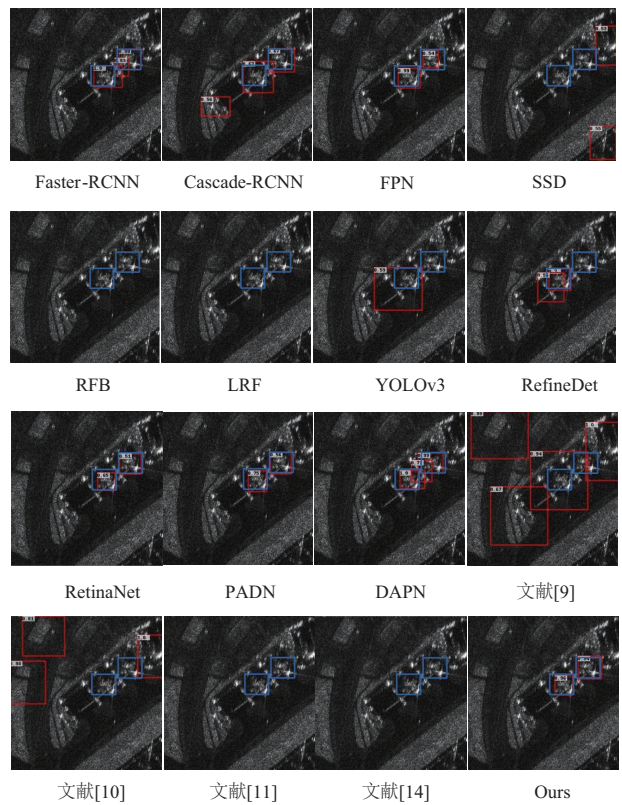


图10 不同算法对尺寸较小且自身散射结构区分度低的飞机检测结果

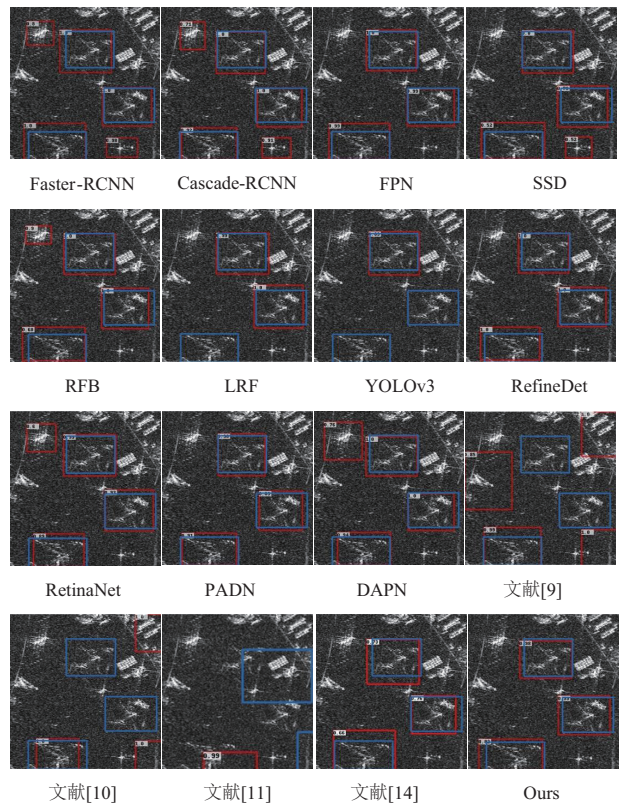


图11 不同算法对背景干扰复杂多样的飞机检测结果

的检测结果中, Faster-RCNN、SSD、LRF、YOLOv3、文献[9~11, 14]以及PADN算法均将部分飞机漏检, Cascade-RCNN与RefineDet算法未将并排停放的飞机目标准确区分(如图9飞机A与B)。然而, AFFN可将这三架飞机准确检测且预测值与目标真值偏差较小。这表明合理设计的注意力特征融合策略, 优化的网络结构以及目标多级检测策略有效提升了对复杂环境条件下目标检测的准确性。

(2) AFFN可有效对尺寸较小且自身散射结构区分度低的飞机准确检测。在图10中, 飞机目标具有尺寸小、主要部件辨识度弱的特点。由于对目标浅层特征提取能力有限且周围环境与飞机的相似度较高, 较多对比算法将周围地物误检为目标(如Faster-RCNN、Cascade-RCNN、DAPN、SSD、文献[9, 10]等)。PADN、RetinaNet以及FPN算法虽可将目标正确检测, 但目标置信度整体较低。然而, AFFN不仅将所有飞机正确检测且对目标预测的置信度较高。

(3) AFFN可有效将SAR图像环境干扰与飞机准确判别。在图11中, 飞机周围地物成像复杂多样, 地物与飞机部件成像相似度较高。在不同算法检测结果中, Faster-RCNN、Cascade-RCNN、DAPN、SSD、RFB、RetinaNet以及文献[9~11]算法均将地物误检为飞机, 部分飞机被LRF与YOLOv3算法漏检。然而, AFFN可对目标准确检测, 且较PADN、RefineDet、文献[14]对目标的检测置信度最高。

## 4 讨论

### 4.1 注意力特征融合策略探究

实验将AFFN的ATCB替换为TCB, ABFFM替换为TCB, ATCB与ABFFM均替换为TCB, 并将调整的算法记为AFFN-ATCB、AFFN-ABFFM以及AFFN(TCB)。表3为以RefineDet为基准, 调整后AFFN算法对SAR图像飞机目标的测试结果。

表3 注意力特征融合网络对算法性能的影响

网络名称	$P$	$R$	$F1$	AP	Params (M)	FLOPs (G)	FPS
AFFN	0.856	0.873	0.864	0.871	32.190	39.215	91
AFFN-ATCB	0.850	0.869	0.859	0.870	32.172	39.201	95
AFFN-ABFFM	0.842	0.886	0.863	0.871	30.402	38.504	95
AFFN(TCB)	0.840	0.857	0.848	0.865	30.385	38.490	100
RefineDet	0.856	0.806	0.830	0.843	33.915	37.426	90

对比AFFN-ATCB、AFFN-ABFFM以及AFFN(TCB)三种算法的检测结果可得, ATCB与ABFFM均可提升算法对飞机目标特征的提取与判别。当某一模块缺失时, 另一模块可发挥作用, 稳定算法性能, 当两者均缺失时, 算法性能下降显著。对比AFFN、AFFN(TCB)、

RefineDet结果可得, 根据目标散射点特征而优化设计的网络结构可平衡不同维度语义信息, 对算法检测性能的提升也至关重要。此外, AFFN、AFFN-ATCB、AFFN-ABFFM、AFFN(TCB)、RefineDet算法对三种不同场景中飞机目标检测结果如图12至图14所示, 不难看出:

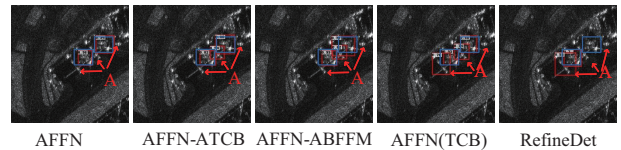


图12 AFFN对尺寸较小且自身散射结构区分度低的飞机检测结果

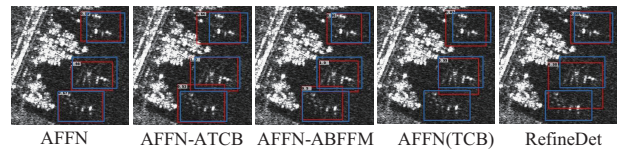


图13 AFFN对散射点强度较弱且离散化程度高的飞机检测结果

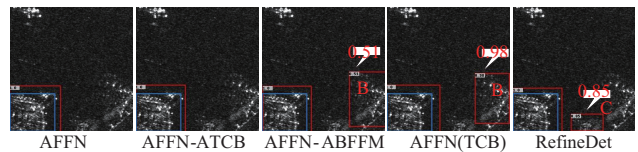


图14 AFFN对散射点关联度低且分布离散的飞机检测结果

(1) 浅层语义信息对SAR图像飞机的准确判别至关重要, 注意力特征融合策略可有效对浅层特征图中目标几何、纹理信息进行表征, 提升了算法对飞机目标的判别能力。

(2) ATCB与ABFFM均可提升算法对目标的特征提取与判别能力, 两者优势互补, 共同提升了AFFN对SAR图像飞机目标的检测能力。

(3) 合理的网络结构设计可为注意力特征融合策略赋能增效。AFFN平衡了高、中、低不同层的CNN语义特征, 使注意力特征融合策略对目标多层次信息的表征更加均衡。

### 4.2 注意力机制探究

图15与图16分别为Squeeze-and-Excitation(SE)<sup>[24]</sup>与Convolutional Block Attention Module(CBAM)<sup>[25]</sup>, 是除了BAM外的两种主流注意力机制。其中, SE仅以通道注意力机制对输入特征 $X$ 加权, CBAM则以串行的通道与空间注意力机制对输入特征 $X$ 加权。实验将以BAM、SE、CBAM为注意力模块的AFFN分别命名为AFFN(BAM)、AFFN(SE)、AFFN(CBAM)并以不含任何注意力机制的AFFN作为基准, 命名为AFFN(Basic)。

表4展示了不同注意力配置的AFFN对SAR图像飞机的测试结果。不难看出, AFFN(BAM)较AFFN(CBAM)与AFFN(SE)对AFFN(Basic)的提升更为显

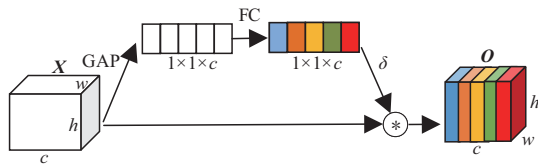


图 15 SE 结构

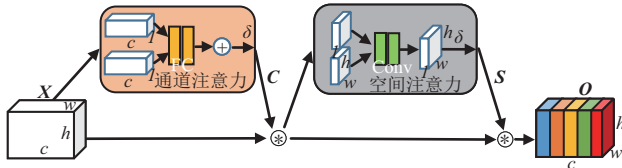


图 16 CBAM 结构

著. 此外, 由于算法主体结构的相似性, 这四种算法的模型参数量、浮点计算量以及检测速度差异较小. 虽然 CBAM 的参数量更少, 但其串行化的注意力机制降低了计算并行度, 限制了运算速度.

表 4 不同注意力机制对算法性能的影响

网络名称	P	R	F1	AP	Params (M)	FLOPs (G)	FPS
AFFN(BAM)	0.856	0.873	0.864	0.871	32.190	39.215	91
AFFN(SE)	0.835	0.865	0.850	0.868	32.189	39.197	98
AFFN(CBAM)	0.846	0.872	0.859	0.867	32.172	39.198	85
AFFN(Basic)	0.856	0.853	0.854	0.865	32.155	39.197	100

图 17~19 展示了 AFFN(BAM)、AFFN(CBAM)、AFFN(SE) 以及 AFFN(Basic) 对三种不同场景中飞机的检测结果. 在图 17 中, 飞机尺寸较小且散射结构区分度低, AFFN(SE) 与 AFFN(CBAM) 均将周围地物误检为目标, AFFN(Basic) 将目标漏检; 在图 18 中, 飞机部件散射点关联度低且分布离散. 由于缺乏对目标全局特征的

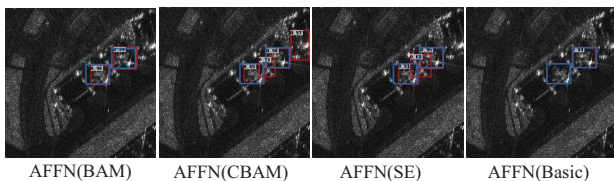


图 17 注意力机制对尺寸较小且散射结构区分度低的飞机检测的影响

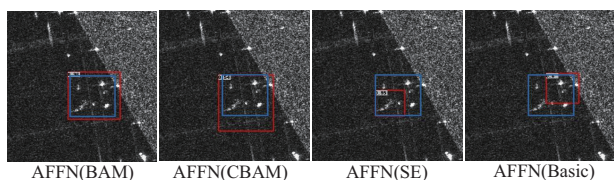


图 18 注意力机制对散射点关联度低且分布离散的飞机检测的影响

有效感知, AFFN(SE) 与 AFFN(Basic) 的预测框均小于目标真值. AFFN(CBAM) 对目标的位置预测与真值存

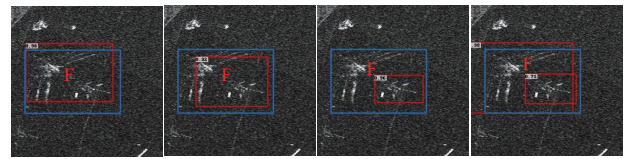


图 19 注意力机制对多次散射成像的飞机检测的影响

在一定差异, 且对目标预测的置信度较低; 在图 19 中, 飞机机翼与尾翼处存在多次散射, 且飞机特定部位(F 区域) 散射点较弱. AFFN(SE) 与 AFFN(Basic) 将目标尾翼误判为飞机或对目标的预测较真值偏差较大, 且 AFFN(CBAM) 对目标的预测框较真值存在一定偏差. 对比采用 BAM、CBAM、SE 三种不同注意力机制的 AFFN 对 SAR 图像飞机的检测结果可得, BAM 中并行化的通道与空间注意力较 CBAM、SE 对目标离散化散射点具有更强的感知与关联能力, 对目标结构、位置预测的完整性与准确性也更优.

### 5 总结

针对 SAR 图像飞机目标散射点离散化程度高, 目标周围环境干扰严重, 当前算法检测效率低等问题, 本文提出了注意力特征融合网络(AFFN). 通过引入瓶颈注意力模块(BAM), 论文在 AFFN 中设计了由注意力双向特征融合(ABFFM)与注意力传输连接(ATCB)模块构成的注意力特征融合策略, 构建了包含低、中、高维特征的精细化特征金字塔, 强化了对 SAR 图像飞机散射点浅层语义信息的感知与综合利用能力. 基于自建的 Gaofen-3 与 TerraSAR-X 混合飞机目标实测数据集, 实验将 AFFN 与基于 CNN 的通用目标检测算法以及 SAR 图像特定目标检测算法进行了比较, 结果表明 AFFN 对复杂环境中的飞机目标检测准确性更高, 检测速度更快. 此外, 对注意力特征融合策略和 BAM 模块的消融实验进一步验证了其对 SAR 图像飞机目标浅层特征提取与判别的重要作用. 未来, 我们将深入结合目标散射特征与网络结构设计, 进一步提升算法对 SAR 图像飞机目标的检测能力.

### 参考文献

[1] 张嘉峰, 杨子渊, 张鹏, 等. Fisher 纹理分布下基于匹配滤波的极化 SAR 图像 CFAR 检测方法[J]. 电子学报, 2019, 47(12): 2533 - 2543.  
Zhang J F, Yang Z Y, Zhang P, et al. A new CFAR detection method of polarimetric SAR imagery based on matched filter under fisher texture[J]. Acta Electronica Sinica, 2019, 47(12): 2533 - 2543.(in Chinese)

[2] 王彬, 王国宇. 基于广义 Gamma 分布的高分辨率 SAR 图像海岸线检测[J]. 电子学报, 2018, 46(4): 827 - 833.

- Wang B, Wang G Y. A coastline detection method in high-resolution SAR images based on the generalized gamma distribution[J]. *Acta Electronica Sinica*, 2018, 46(4): 827 – 833.(in Chinese)
- [3] 魏志强, 毕海霞, 刘霞. 基于深度卷积神经网络的图上半监督极化SAR图像分类算法[J]. *电子学报*, 2020, 48(1): 66 – 74.
- Wei Z Q, Bi H X, Liu X. A graph-based semi-supervised PolSAR image classification method using deep convolutional neural networks[J]. *Acta Electronica Sinica*, 2020, 48(1): 66 – 74.(in Chinese)
- [4] Zhao Y, Zhao L J, Xiong B L, 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.
- [5] Hu H, Huang L Q, Yu W X. Aircraft detection for HR SAR images in non-homogeneous background using GGMD-based modeling[J]. *Chinese Journal of Electronics*, 2019, 28(6): 1271 – 1280.
- [6] Zhang H, Hao M X, Zhang C, et al. Aircraft target detection algorithm based on high resolution spaceborne SAR imagery[A] *MIPPR 2017: Remote Sensing Image Processing, Geographic Information Systems, and Other Applications*[C]. Xiangyang, China: SPIE, 2018: 193 – 198.
- [7] Tan Y H, Li Q Y, Li Y S, et al. Aircraft detection in high-resolution SAR images based on a gradient textural saliency map[J]. *Sensors*, 2015, 15(9): 23071 – 23094.
- [8] Dou F, Diao W, Sun X, et al. Aircraft recognition in high resolution SAR images using saliency map and scattering structure features[A] 2016 IEEE International Geoscience and Remote Sensing Symposium(IGARSS) [C]. Beijing, China: IEEE, 2016: 1575 – 1578.
- [9] 王思雨, 高鑫, 孙皓, 等. 基于卷积神经网络的高分辨率SAR图像飞机目标检测方法[J]. *雷达学报*, 2017, 6(2): 195 – 203.
- Wang S Y, Gao X, Sun H, et al. An aircraft detection method based on convolutional neural networks in high-resolution SAR images[J]. *Journal of Radars*, 2017, 6(2): 195 – 203.(in Chinese)
- [10] 郭倩, 王海鹏, 徐丰. 星载合成孔径雷达图像的飞机目标检测[J]. *上海航天*, 2018, 35(6): 57 – 64.
- Guo Q, Wang H P, Xu F. Aircraft target detection from spaceborne synthetic aperture radar image[J]. *Aerospace Shanghai*, 2018, 35(6): 57 – 64.(in Chinese)
- [11] Diao W H, Dou F Z, Fu K, et al. Aircraft detection in sar images using saliency based location regression network [A] 2018 IEEE International Geoscience and Remote Sensing Symposium [C]. Valencia, Spain: IEEE, 2018. 2334 – 2337.
- [12] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[A]. 2017 IEEE International Conference on Computer Vision (ICCV) [C]. Venice, Italy: IEEE, 2017: 2999 – 3007.
- [13] Zhao Y, Zhao L J, Li C Y, et al. Pyramid attention dilated network for aircraft detection in SAR images[J]. *IEEE Geoscience and Remote Sensing Letters*, 2021, 18(4): 662 – 666.
- [14] He C, Tu M X, Xiong D H, et al. A component-based multi-layer parallel network for airplane detection in SAR imagery[J]. *Remote Sensing*, 2018, 10(7): 1016.
- [15] Zhang S, Wen L, Bian X, et al. Single-shot refinement neural network for object detection[A]. *Proceedings of the IEEE conference on computer vision and pattern recognition*[C]. Salt Lake City, USA: IEEE, 2018: 4203 – 4212.
- [16] Ren S Q, He K M, 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.
- [17] Cai Z W, Vasconcelos N. Cascade R-CNN: Delving into high quality object detection[A]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Salt Lake City, USA: IEEE, 2018. 6154 – 6162.
- [18] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[A]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Honolulu, Hawaii: IEEE, 2017. 936 – 944.
- [19] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[A]. *European Conference on Computer Vision (ECCV)* [C]. Amsterdam, Netherlands: Springer, 2016: 21 – 37.
- [20] Liu S T, Huang D, Wang Y H. Receptive field block net for accurate and fast object detection[A]. *Proceedings of the European Conference on Computer Vision (ECCV)* [C]. Munich Germany: Springer, 2018: 385 – 400.
- [21] Wang T C, Anwer R M, Cholakkal H, et al. Learning rich features at high-speed for single-shot object detection[A]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV) [C]. Seoul, Korea: IEEE, 2019. 1971 – 1980.
- [22] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement[EB/OL]. <https://arxiv.org/abs/1804.02767>, 2018.

- [23] Cui Z Y, Li Q, Cao Z J, et al. Dense attention pyramid networks for multi-scale ship detection in SAR images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11): 8983 – 8997.
- [24] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [A]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Salt Lake City, USA: IEEE, 2018. 7232 – 7141.
- [25] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[A]. Computer Vision-ECCV 2018 [C]. Munich, Germany: IEEE, 2018.3 – 19.

#### 作者简介



赵 琰 男, 1996年出生于陕西西安, 国防科技大学电子科学学院在读博士研究生, 研究方向为遥感信息处理, 合成孔径雷达目标自动识别.

E-mail: zy34731@qq.com



匡纲要 男, 1966年出生于湖南衡阳, 国防科技大学电子科学学院电子信息系统与复杂电磁环境效应(CEMEE)国家重点实验室教授, 博士生导师, 研究方向为遥感图像智能解译、SAR图像目标检测与识别.

E-mail: kuangmeng@hotmail.com



赵凌君(通信作者) 女, 1981年出生于安徽安庆, 国防科技大学电子科学学院副教授, 研究方向为遥感信息处理, 合成孔径雷达目标自动识别.

E-mail: zhaolingjund@126.com