•计算机科学与技术•

DOI:10.12454/j.jsuese.202600045

本刊网刊

基于双域带状注意力的SAR图像目标检测方法

徐鹏飞[1],张博祥[2],李艺洋[2],李　炜[1*]

(1.四川大学 空天科学与工程学院,四川 成都 610065;2.中国航空工业集团有限公司 成都飞机设计研究所,四川 成都 610032)

摘　要:针对合成孔径雷达(SAR)图像在复杂背景下的目标检测模型检测精度不足、易过拟合、泛化性弱等问题,本文提出一种基于双域带状注意力的SAR目标检测(DSA–YOLO)算法。该方法充分考虑SAR成像的物理特性,通过将双域带状注意力嵌入特征提取模块,实现对空间分布特征与空间变化率特征的协同提取,从而有效提升SAR目标检测的精度。本文所提出的双域带状注意力模块在空间域中通过条带注意力增强对结构纹理与边缘信息的敏感性,在频域中通过带状滤波机制突出目标的散射与变化特征,从而提升模型对关键特征的感知能力与噪声抑制性能。此外,本文设计了改进的多尺度特征融合策略,使得模型在不同尺度目标检测中具备更强的适应性与鲁棒性。在数据集MSAR–1.0和SAR–Aircraft–1.0上的实验结果表明,本文所提出的双域带状注意力模型对SAR图像多类别目标检测和细粒度目标检测的性能均优于现有主流方法。相较于作为基线模型的YOLOv12模型,本文所提出模型的参数量仅增加了1.68×10^{6},在数据集MSAR–1.0上的浮点运算次数仅增加了5.10×10^{8},平均精度均值(mAP@0.5:0.95)提高了2.52%;在数据集SAR–Aircraft–1.0上的浮点运算次数仅增加了5.94×10^{9},平均精度均值提高了4.31%。实验结果充分验证了本文所提出的DSA–YOLO模型在SAR图像目标检测任务中的有效性与泛化能力。

关键词:合成孔径雷达图像;目标检测;双域带状注意力;带状滤波

中图分类号:TN957.52;TP183　　**文献标志码:**A　　**文章编号:**2096-3246(2026)06-0363-12

合成孔径雷达(SAR)作为一种全天时、全天候的主动式微波成像系统,凭借其穿透云雾与黑暗的能力,在军事侦察、海洋监测和灾害评估等领域发挥着不可替代的作用[1–2]。然而,SAR成像本质上是基于电磁波散射与相干叠加的物理过程,其图像呈现出的乘性相干斑噪声、几何畸变以及独特的动态范围,使得从复杂的SAR数据中实现高精度、高鲁棒的目标检测,成为遥感信息处理领域极具挑战性的前沿课题。

近年来,目标检测在军事和民用领域的重要性不断提升。然而,与光学成像不同,SAR成像属于微波主动成像,其成像机理基于电磁波的散射与相干叠加,与依赖可见光反射的光学图像有本质区别。SAR图像具有分散散射、相干斑噪声强、目标与背景灰度对比度低等特点,导致小目标特征易被噪声掩盖、多尺度目标特征提取不充分,而将传统针对光学图像而设计的目标检测模型直接应用于SAR图像,往往难以有效捕捉关键特征,甚至引入伪特征,存在明显的性能瓶颈。因此,完全套用光学图像处理方法对SAR图像进行目标检测存在一定的弊端,亟须针对SAR的物理成像特性进行结构性优化。

随着近年来深度学习方法的飞速发展,SAR图像目标检测和识别领域也同样取得了显著的进展[3]。研究者普遍借鉴光学图像处理领域的成功经验,将卷积神经网络[4–5]和Transformer[6–7]等视觉架构直接迁移到SAR图像识别任务中。这类方法通常将SAR图像视作灰度光学影像进行处理,通过提取空间纹理、边缘和形状特征实现目标检测与分类。尽管这种基于光学视觉范式的思路显著提升了检测精度和模型的泛化能力,但多数方法仍停留在光学图像的空间特征建模框架内,未能充分挖掘SAR成像所特有的频域信息、散射特性与极化特征,从而限制了其在复杂场景下的识别能力。

收稿日期:2026–01–14　　**修回日期:**2026–03–19　　**网络出版日期:**2026–04–13

基金项目:四川省重点实验室开创基金项目(SSAI202514)

作者简介:徐鹏飞(1994—),男,硕士.研究方向:SAR目标检测.E-mail:xpf13673501563@163.com

* **通信作者:**李　炜,副研究员,E-mail:li.wei@scu.edu.cn

针对上述问题，本文在YOLOv12[8]目标检测框架中引入了一种双域带状注意力机制，以更好地适应SAR成像的物理特性。与光学图像不同，SAR成像过程与频域信息高度耦合，目标的结构与散射特征在频率空间中表现得更加明显，尤其是小目标的微弱散射特征在频域中具有独特的高频分布规律。基于这一特性，本文方法在增添空间条带注意力机制提取特征的基础上，进一步采用频率带状注意力机制对SAR图像进行滤波变换来提取特征，通过对频域分量的精细化解耦，区分由相干斑噪声引发的伪高频与由目标散射中心（如金属角反射器）产生的真高频。这种设计在物理层面实现了信号增强与噪声抑制的统一，进而实现对SAR图像空间特征和空间变化率特征的同步提取，同时有效抑制噪声干扰，为后续目标检测任务提供更具物理一致性、更高质量的特征表示。所设计的双域注意力模块能够在空间特征与空间变化率特征之间实现高效融合，使模型能够同时从空间分布与空间变化率两个维度深度挖掘SAR目标的显著特征，对于小目标而言，可有效增强其微弱散射特征的表征能力；对于多尺度目标而言，可实现不同尺度特征的分层提取与融合，从而显著提升了在低对比度、强噪声干扰场景下的鲁棒性与特征区分能力。与仅依赖空间特征提取的光学图像方法相比，该机制更符合SAR微波成像的物理规律，能够有效克服直接套用光学方法所带来的信息丢失问题，突破性能瓶颈。

1 相关研究现状

在SAR图像的复杂背景下，小目标特征不易提取，分散点更离散，检测难度大。随着深度学习的推进，出现了一系列SAR目标检测算法，主要分为基于视觉显著性的方法、基于多尺度特征融合的方法、基于注意力机制的方法三大类。近年来，研究逐渐向结合SAR物理成像特性、轻量化高鲁棒性、适配最新检测框架的方向发展。

在基于视觉显著性的目标检测方法中，部分研究人员依据SAR图像中的目标通常显示为强散射亮斑的特点，提出将像素亮度作为视觉显著特征，并以此为依据在SAR图像中快速定位感兴趣的候选目标区域。王思雨等[9]采用了一种显著性预检测算法，实现了SAR图像中多尺度快速粗检候选目标的功能，进一步设计并调优了一个浅层的卷积神经网络，以此实现了对候选目标的精确检测。在此基础上，Diao等[10]将基于位置回归的卷积神经网络与SAR图像中目标的显著特征相结合，采用基于恒虚警率（CFAR）的目标预定位算法，提取飞机目标候选区域，并基于Fast R–CNN网络进行改进，在TerraSAR–X卫星采集的3.0 m分辨率数据集上的实验结果表明，所提方法的检测率相比原始Fast R–CNN网络有显著提升。但此类方法对低对比度小目标的显著性特征提取能力不足，易出现漏检问题。

除了基于视觉显著性的SAR目标检测方法外，当下也有大量基于多尺度特征融合及视觉注意力机制的SAR目标检测方法研究。Dai等[11]提出DenoDet模型，将注意力机制作为可变形多子空间特征去噪模块，有效抑制了SAR图像的相干斑噪声，在多类别SAR目标检测中实现了检测精度的提升，但该模型参数量较大，计算效率有待优化。Wang等[12]提出MSAD-Net多级检测网络，针对SAR飞机小目标设计了多阶段特征融合结构，提升了小目标特征的逐层增强能力，但该模型对多尺度目标的自适应检测能力不足。Huang等[13]设计了散射增强与特征融合网络，将SAR目标的电磁散射特性与深度特征融合，增强了模型的物理可解释性，但其仅针对飞机目标设计，泛化性较弱。Chen等[14]提出YOLO–SAD高效检测网络，基于YOLO框架进行轻量化改进，提升了SAR飞机检测的实时性，但该模型未充分挖掘SAR图像的频域信息，在复杂背景下的抗干扰能力有限。上述最新研究虽各有创新，但仍存在未充分结合SAR频域特征、对小目标与多尺度目标的协同检测能力不足，以及与最新目标检测框架融合不深入等问题，为本研究提供了改进空间。

2025年提出的YOLOv12以注意力为中心重构了网络结构，设计了A2C2f高效特征融合模块与C3k2轻量化特征提取模块，在保持实时性的同时，实现了注意力机制与特征提取的深度融合，相比前代YOLO模型，在参数量、计算效率与检测精度上实现了三重提升。但目前YOLOv12在SAR图像检测领域的应用仍有很大的提升空间，其原始模块未针对SAR图像的相干斑噪声、小目标微弱散射、多尺度目标特征差异进行适配，直接应用会出现特征提取不充分以及检测精度下降等问题，这也是本文基于YOLOv12进行改进的核心动机。Nguyen等[15]提出的SSD–YOLOv12将单词多边框检测器（SSD）的多尺度锚框分配策略与YOLOv12的注意力特征提取架构结合，通过浅层特征图匹配小锚框，既保留了SSD对小目标的高检出率，又通过YOLOv12的区域注意力模块扩大了感受野，解决了传统SSD对遮挡目标、复杂背景目标特征提取不足的问题，在小目标、多尺度目标检测场景中性能优于单一SSD或原生YOLOv12，但其采用的SSD边界框回归精度低于YOLO系列的CIoU/DIoU损失。

现有注意力机制在SAR检测中的应用多集中于空间域，少数结合了信道注意力，尚未有研究将频域注意力与空间注意力结合形成双域协同的注意力机制，且未与最新的YOLOv12框架融合。本文首次将双域带状注意力（DSA）嵌入YOLOv12框架，实现空间域与频域域的特征协同提取，弥补了现有研究的不足。

2　模型结构设计

2.1　整体架构

本文提出了一种针对SAR图像目标检测任务的多域注意力检测（DSA–YOLO）模型，其整体结构如图1所示。图1中，H为图像高度，W为图像宽度。

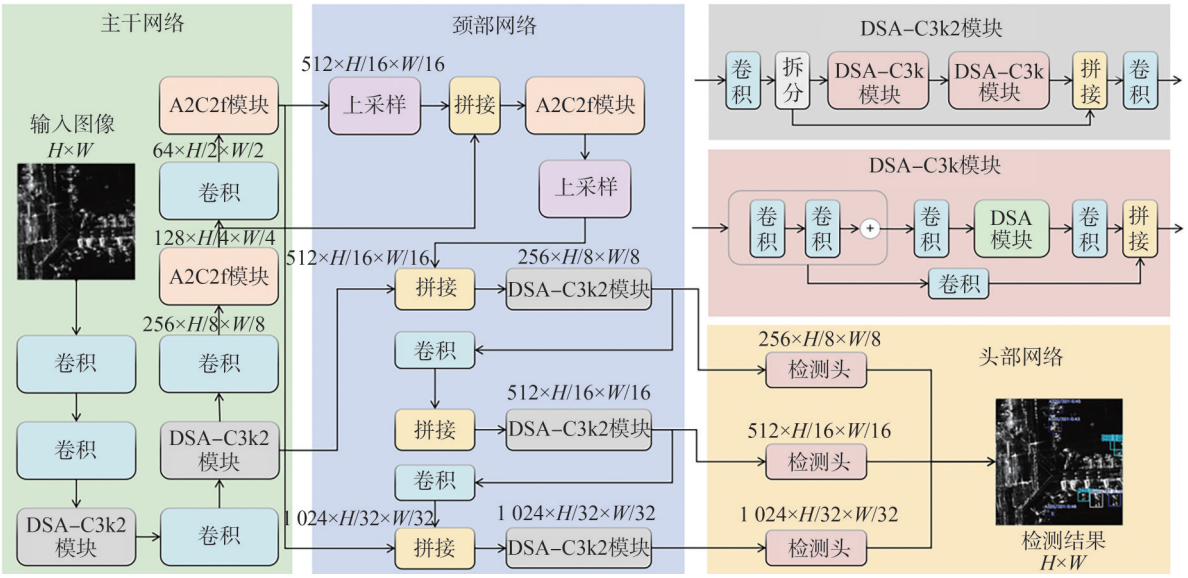


图1　DSA–YOLO总体模型构架

Fig. 1　Overall model architecture of DSA–YOLO

DSA–YOLO模型主要由主干网络、颈部网络和头部网络3个部分组成。首先，图像从输入端进入主干网络，经由多层卷积层、两个DSA–C3k2模块以及两个A2C2f模块，实现多层次特征提取与融合。其次，由主干网络提取出的特征进一步进入颈部网络，颈部网络由拼接层、上采样层、卷积层、A2C2f模块以及3个DSA–C3k2模块组成，对主干网络提取到的特征进行融合和调整，通过上采样以及拼接操作，将不同层次的特征信息进行整合，增强特征的表达。最后，由颈部网络融合及调整后的特征进入头部网络中，头部网络由3个检测层组成，负责最终的目标检测任务，并输出检测到的目标类别和位置信息。3个检测层分别负责检测不同尺寸的目标，综合3个检测层的结果能够有效提升模型整体的目标检测性能。

2.2　双域带状注意力机制的嵌入设计

DSA–YOLO的核心创新在于将DSA嵌入到网络的关键位置。经过深入的实验探索，本文确定将C3k2模块作为DSA嵌入的最佳载体。通过在C3k2模块中集成DSA模块，模型能够在保持计算效率的同时，实现对空间分布特征与频域变化特征的深度聚合，从而提高模型在小目标、多尺度、遮挡、噪声等情形下的目标检测精度。

图1中，DSA–C3k模块由数个卷积核大小为3×3、步长为1、通道数为256的卷积层和一个DSA模块拼接而成。DSA–C3k2模块由两个卷积层、两个DSA–C3k模块，以及一个按1∶1比例进行拆分的拆分层与一个按第二维度进行直接拼接的拼接层拼接而成。输入特征首先经过1层卷积核大小为3×3、步长为1、通道数为256的卷积层，然后特征被拆分为两个通道数为128的特征，一部分特征经过两个DSA–C3k模块，与另一部分未经进一步处理的特征共同被输入到一个按第二维度进行直接拼接的拼接层，最后再经过1层卷积核大小为3×3、步长为1、通道数为256的卷积层后输出。

针对原YOLOv12模型中的C3k2模块在SAR图像空间细节捕捉与空间变化率特征挖掘上的不足，本文对该模块的分支结构与特征交互方式进行了重构：在C3k模块的主分支中，嵌入双域带状注意力模块得到DSA–C3k模块，该模块通过空间域带状注意力分支对SAR图像中目标的边缘轮廓、纹理细节等空间特征进行精准定位与增强，同时利用频率域带状注意力分支对图像频域内的高频细节（如目标局部特征）与低频全局信息（如背景环境分布）进行分层提取；输出经由DSA–C3k模块拼接并增加特征自适应调整层得到DSA–C3k2模块，实现主分支输出的双域特征与原始输入特征的高效融合，有效避免特征信息丢失。通过这

一改进，模型能够同时从空间维度与频率维度深度挖掘SAR图像中的目标特征，显著提升了对低对比度、强噪声干扰下小目标与多尺度目标的特征表征能力。

3 双域带状注意力机制原理

3.1 空间带状注意力机制

空间带状注意力(SSA)[16]单元由垂直和水平带状注意力组成，其结构示意图如图2所示。SSA通过水平和垂直方向的条带特征聚合，能够有效增强SAR目标的边缘轮廓、结构纹理等空间特征，对多尺度目标而言，可通过不同条带长度适配不同尺度的空间特征分布。

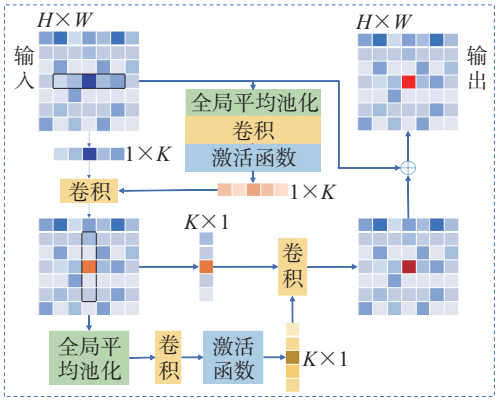


图 2 空间带状注意力(SSA)结构示意图

Fig. 2 Schematic diagram of the spatial strip attention structure

空间带状注意力的具体计算方法如下：对于给定输入张量$\boldsymbol{X}\in\mathbb{R}^{C\times H\times W}$($C$为输入图像的通道数)，首先，通过一个轻量级的分支产生注意力权重$\boldsymbol{A}$，该分支由全局平均池化(GAP)、卷积核大小为1×1的卷积(简称1×1卷积)以及Sigmoid函数组成。注意力权重的生成过程可以表示为：

$$\boldsymbol{A}=\sigma(\mathrm{Conv}_{1\times1}(\mathrm{GAP}(\mathrm{X})))\in\mathbb{R}^{K} \tag{1}$$

式中，GAP(·)为全局平均池化，$\mathrm{Conv}_{1\times1}(\cdot)$为1×1卷积，$\sigma(\cdot)$为Sigmoid函数，$K$为生成带状的长度。然后，将注意力权重$\boldsymbol{A}$与输入张量$\boldsymbol{X}$进行一维卷积，得到细化后的特征，该过程可表示为：

$$\hat{\boldsymbol{X}}_{C,H,W}=\sum_{k=0}^{[\frac{K}{2}]-1}\boldsymbol{A}_k\boldsymbol{X}_{C,H,W-[\frac{K}{2}]+k} \tag{2}$$

可将水平空间带状注意力算子概括为$\hat{\boldsymbol{X}}=\mathrm{S}_K^{\mathrm{H}}(\boldsymbol{X})$。同理，垂直注意力算子可表示为$\mathrm{S}_K^{\mathrm{V}}(\cdot)$，从而可以将SSA的整体计算过程表示为：

$$\boldsymbol{X}_K^{\mathrm{SSA}}=\mathrm{SSA}_K(\boldsymbol{X})=\mathrm{S}_K^{\mathrm{V}}(\mathrm{S}_K^{\mathrm{H}}(\boldsymbol{X})) \tag{3}$$

式中，$\boldsymbol{X}_K^{\mathrm{SSA}}$为SSA单元计算模块$\mathrm{SSA}_K(\cdot)$的输出。

SSA通过水平和垂直注意力算子分别在水平和垂直两个方向上进行信息整合，隐式地扩大了网络的感受野，提升了模型目标检测的性能。

3.2 频率带状注意力机制

本文采用频率带状注意力(FSA)[16]单元来提取图像的空间变化率信息，其结构示意图如图3所示。图3中，$\boldsymbol{W}_{\mathrm{l}}^{\mathrm{V}}$和$\boldsymbol{W}_{\mathrm{h}}^{\mathrm{V}}$分别为用于调制低频和高频分量的水平注意力参数，$\boldsymbol{W}_{\mathrm{l}}^{\mathrm{H}}$和$\boldsymbol{W}_{\mathrm{h}}^{\mathrm{H}}$分别为用于调制低频和高频分量的垂直注意力参数。

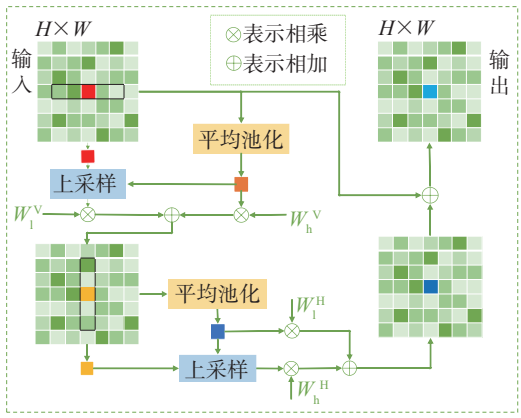


图 3 频率带状注意力(FSA)结构示意图

Fig. 3 Schematic diagram of the frequency strip attention (FSA) structure

FSA通过带状平均池化分离图像的低频与高频分量，能够有效突出SAR目标的散射特征与变化特征，对小目标而言，可有效提取其微弱的高频散射特征，抑制背景噪声的伪高频干扰；对多尺度目标而言，可实现不同尺度目标的频域特征分层提取。

频率带状注意力的具体计算方法如下：对于给定输入张量$\boldsymbol{X}\in\mathbb{R}^{C\times H\times W}$，首先利用条形平均池化作为均值滤波器，获得每个带状区域$\boldsymbol{S}\in\mathbb{R}^{1\times K}$的低频分量$\boldsymbol{X}^{\mathrm{l}}$。然后，通过从输入特征$\boldsymbol{X}$中去除所得低频分量，得到与低频分量互补的高频分量$\boldsymbol{X}^{\mathrm{h}}$。接着，将信道可学习注意力参数施加到不同的频率分量上，使网络能够有选择地关注信息频率。FSA的参数较小，可以通过反向传播直接更新。最后，通过将细化后的频率相加得到水平算子的输出。上述过程可以表示为：

$$\boldsymbol{X}^{\mathrm{l}}=\mathrm{AP}(\boldsymbol{S}) \tag{4}$$

$$\boldsymbol{X}^{\mathrm{h}}=\boldsymbol{X}-\boldsymbol{X}^{\mathrm{l}} \tag{5}$$

$$\hat{\boldsymbol{X}}=\boldsymbol{W}_{\mathrm{l}}^{\mathrm{V}}\boldsymbol{X}^{\mathrm{l}}+\boldsymbol{W}_{\mathrm{h}}^{\mathrm{V}}\boldsymbol{X}^{\mathrm{h}} \tag{6}$$

式(4)～(6)中，AP(·)表示平均池化过程。上述水平注意力算子可以概括为$\hat{\boldsymbol{X}}=\mathrm{F}_K^{\mathrm{H}}(\mathbf{X})$。类似的，垂直注意力算子可表示为$\mathrm{F}_K^{\mathrm{V}}(\cdot)$。故FSA整体计算公式可表示为：

$$\boldsymbol{X}_K^{\mathrm{FSA}}=\mathrm{FSA}_K(\boldsymbol{X})=\mathrm{F}_K^{\mathrm{V}}(\mathrm{F}_K^{\mathrm{H}}(\boldsymbol{X})) \tag{7}$$

式中，$\boldsymbol{X}_K^{\mathrm{FSA}}$为FSA单元计算$\mathrm{FSA}_K(\cdot)$的输出。

3.3 双域协同机制

DSA基于双域学习理论，通过空间和频率两个域的信息融合，增强图像特征表示。在空间域，SSA利用简单卷积分支生成注意力权重，聚合相邻位信息，提

取SAR图像的空间特征；在频率域，FSA通过带状平均池化分离频率成分，提取SAR图像的空间变化率特征。此外，DSA应用不同带状大小进行多尺度学习，从而能够更好地提取不同尺寸的特征。

图4为双域带状注意力（DSA）结构示意图。如图4所示，DSA的具体结构如下：输入特征首先会经过一层卷积核为3×3且步长为1的卷积对通道数量进行适当调整。然后，将调整后的特征按照1∶1的比例进行拆分，并分别输入到条带长度$k_1=5$、$k_2=9$的两个FSA模块和条带长度$k_1=5$、$k_2=9$的两个SSA模块中。接着，将经由两个FSA模块和两个SSA模块处理后的特征进行拼接。拼接后的特征结合了FSA模块提取的SAR图像空间变化率特征与SSA模块提取的SAR图像空间特征。最后，将拼接后的特征输入到一个卷积核大小为3×3且步长为1的卷积层中进行进一步精炼，调整通道维度的特征表示，使特征更具代表性，为后续的处理做准备。经由卷积层提炼后的特征即为DSA模块的输出，该输出包含了丰富的空间和空间变化率信息，能够有效提取图像中小目标与多尺度目标的特征。

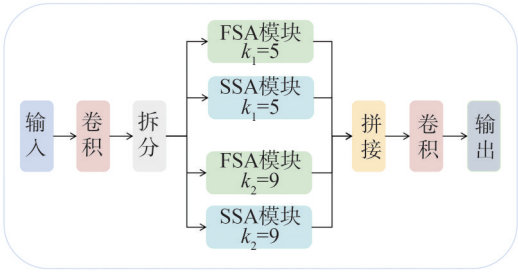


图4　双域带状注意力（DSA）结构示意图

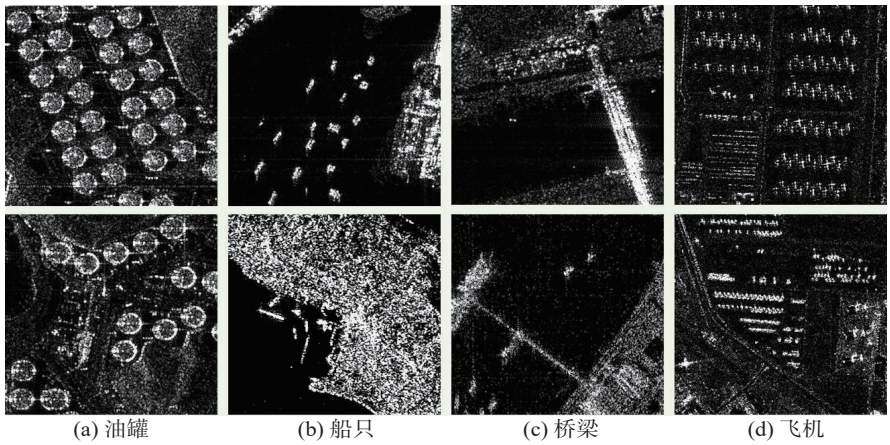
Fig. 4　Schematic diagram of the dual-domain strip attention (DSA) structure

在SAR图像的复杂背景下，小目标通常仅表现为少数离散的强散射中心，极易被相干斑噪声淹没。传统的方形感受野在提取小目标特征时往往会引入过多的背景杂波，导致目标信息在网络深层被稀释。而DSA通过应用不同带状大小进行多尺度学习，不仅可以适应目标的尺寸变化，其独特的结构对小目标检测也具有显著的增强作用。一方面，在空间域中，SSA模块狭长的带状感受野能够在捕捉小目标局部结构的同时，最大限度地减少正交方向上背景杂波的引入，防止小目标微弱的像素特征被局部噪声稀释；另一方面，在图像的频率域中，由于SAR图像的小目标通常表现为孤立的强散射点（即典型的高频分量），FSA模块通过精细分离高低频信息，能够针对性地放大和保护这些高频散射特征，避免其在特征下采样过程中被低频背景信息平滑掉。双域协同机制从根本上提升了模型在复杂场景下对微弱小目标的感知与定位能力。

4　实　验

4.1　数据说明

本文使用的大规模多类SAR目标检测数据集–1.0（MSAR–1.0）[17]是采用海丝一号卫星和高分三号卫星数据切片得到的大型数据集。该数据集场景包括机场、港口、近岸、岛屿、远海、城区等；类型包括飞机、油罐、桥梁和船只4类目标，由1 851架桥梁，39 858条船只，12 319个油罐和6 368架飞机组成，包含大量小目标和多尺度目标。MSAR–1.0数据集能够有效验证不同模型对SAR图像中不同种类目标的检测能力。图5为MSAR–1.0数据集部分切片样例。

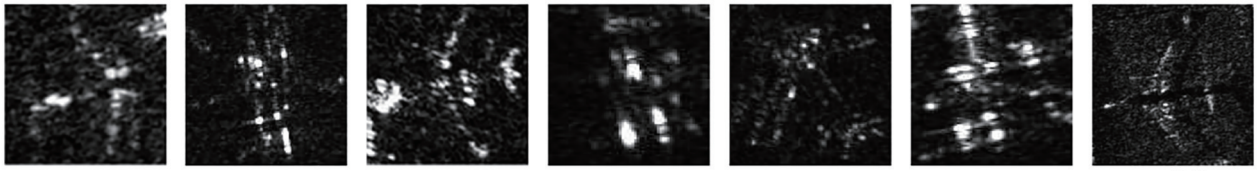


(a) 油罐　(b) 船只　(c) 桥梁　(d) 飞机

图5　MSAR–1.0数据集样本示例

Fig. 5　Sample examples of the MSAR–1.0 dataset

本文所用数据集SAR–AIRcraft–1.0[18]是专门用于SAR图像中飞机目标检测的高质量公共数据集，该数据集在遥感图像目标识别领域有着重要的学术价值以及实际应用潜力。该数据集采用尖端的SAR成像技术捕获，包括4 368张精心挑选的7分类飞机图像，能够有效验证不同方法模型对细粒度SAR图像目标的检测能力。图6为数据集SAR–AIRcraft–1.0中不同类别SAR飞机样本示例。



(a) A220 (b) A320/321 (c) A330 (d) ARJ21 (e) Boeing737 (f) Boeing787 (g) 其他

图 6 不同型号飞机 SAR 图像样本示例

Fig. 6 Examples of SAR image samples for different aircraft models

4.2 评估指标

为定量评估模型算法在 SAR 目标检测中的性能，采用了 3 种标准目标检测指标，包括参数量、浮点运算次数（FLOPs）和平均精度均值（mAP），本文中使用的平均精度均值指标为 mAP@0.5:0.95，即在不同 IoU 阈值（从 0.50 到 0.95，步长为 0.05）上的平均 mAP。

4.3 实验环境及参数设置

为保证评估的一致性与公平性，所有方法均采用标准化的数据集预处理流程。实验环境的硬件配置如下：CPU 型号为 Intel (R) Core (TM) i7–9700，基础频率为 3.00 GHz，GPU 型号为 NVIDIA GeForce RTX 3090 Ti。软件环境如下：操作系统为 Windows11，编程语言为 Python 3.12，CUDA 版本为 v12.1。实验在 GPU 上进行训练，统一参数配置情况如下：初始学习率均为 0.01，训练批次大小设置均为 32，迭代次数设置均为 200，训练均使用 Adam 方法进行优化。

IOU 阈值设置：本实验中所有模型的检测框匹配 IOU 阈值均设置为 0.50 至 0.95，mAP@0.5:0.95 的计算采用 IOU 阈值从 0.50 到 0.95，以 0.05 为步长的平均精度均值，符合目标检测领域的通用评价标准。

针对数据集 MSAR–1.0，图像输入尺寸统一调整为 256 像素×256 像素，并针对 256 像素×256 像素的输入图像计算模型的浮点运算次数。

针对数据集 SAR–AIRcraft–1.0，图像输入尺寸统一调整为 512 像素×512 像素，并针对 512 像素×512 像素的输入图像计算模型的浮点运算次数。

4.4 实验结果

4.4.1 对照实验结果

为验证所提出方法对 SAR 图像中不同种类目标的检测能力，本文在数据集 MSAR–1.0 上进行了对照实验，各模型的运行结果如表 1 所示。

从表 1 可以看出，本文提出的 DSA–YOLO 模型相比于 YOLOv12 模型，参数量仅增加了 1.68×10^6，浮点运算次数仅增加了 5.1×10^8，性能上仅有略微下降，在所有对比的模型中仅次于 YOLOv12 模型，但在平均精度方面却有显著提升。在表 1 所对比的方法模型中，DSA–YOLO 模型的平均精度均值最高，即该方法针对不同类别目标的综合检测性能最佳，比 mAP@0.5:0.95 第 2 高的 DenoDet 模型提高了 1.07%，比 YOLOv12 模型提高了 2.52%。表 1 的数据表明，DSA–YOLO 模型在略微增加模型参数量和浮点运算次数的基础上，显著提高了模型的目标检测性能，证明了本文所提方法在多类别 SAR 图像目标检测任务中的有效性。

表 1 在 MSAR–1.0 数据集上与当前最优方法的对比

Tab. 1 Comparison with SOTA methods on the MSAR–1.0 dataset

模型	参数量/10^6	FLOPs/10^9	mAP@0.5:0.95/%				
			总体	船只	飞机	桥梁	油罐
Faster R–CNN[19]	41.36	26.13	55.69	80.50	9.09	69.70	63.47
ATSS[20]	32.12	12.89	68.00	88.85	44.06	75.73	63.37
Reppoints[21]	32.12	12.88	67.18	89.87	41.05	74.91	62.92
DDOD[22]	32.27	29.02	60.80	86.59	63.16	71.42	22.04
GFL[23]	32.27	13.08	67.24	89.21	42.16	74.05	63.54
PAA[24]	32.12	12.89	63.37	90.29	22.63	77.01	63.55
TOOD[25]	32.03	12.62	66.58	89.04	42.78	71.00	63.49
VFNet[26]	32.72	12.09	65.31	87.94	38.80	71.03	63.49
FCOS[27]	32.12	12.89	67.52	89.17	43.72	73.92	63.28
DenoDet[11]	34.23	12.89	69.90	89.40	47.87	78.92	63.41
YOLOv12[8]	25.58	6.32	68.45	84.41	41.59	66.95	80.69
DSA–YOLO（本文模型）	27.26	6.83	70.97	86.23	42.50	73.83	81.38

表2为本文提出模型与当前最优的一系列目标检测模型在SAR–AIRcraft–1.0数据集上检测结果的对比情况。从表2可以看出，YOLOv12模型有着最小的参数量和最低的浮点运算次数，而本文的DSA–YOLO模型仅次于YOLOv12模型。相比于YOLOv12模型，本文的DSA–YOLO模型参数量增加了 1.68×10^6，但仍比其余模型低 4.77×10^6 以上，浮点运算次数增加了 1.94×10^9，但仍比其余模型低 5.15×10^9 以上。从表2还可以看出，DSA–YOLO模型有着最高的平均精度均值，相比于YOLOv12模型提高了4.31%，比mAP@0.5:0.95第2高的DenoDet模型高了1.02%，表明DSA–YOLO在细粒度的SAR目标检测任务中仍有着优越的检测性能。在7分类的SAR飞机目标检测任务中，本文的DSA–YOLO模型对其中4类目标的检测结果均有着最高的平均精度，充分表明了本文的DSA–YOLO模型针对细粒度SAR图像目标检测任务的有效性。

表 2　在 SAR–Aircraft–1.0 数据集上与当前最优方法的对比

Tab. 2　Comparison with SOTA methods on the SAR–Aircraft–1.0 dataset

模型	参数量/10^6	FLOPs/10^9	mAP@0.5:0.95/%							
			总体	A220	A320	A330	ARJ21	B737	B787	其他
Faster R–CNN[19]	41.36	64.71	64.71	56.96	88.09	96.99	52.78	35.74	59.08	63.32
ATSS[20]	32.12	51.58	66.01	58.57	88.60	96.57	60.14	37.79	59.18	61.25
Reppoints[21]	32.12	51.57	64.11	58.03	93.03	97.02	61.90	36.41	41.74	60.61
DDOD[22]	36.82	48.50	67.13	64.27	89.75	86.28	61.28	41.50	59.73	67.10
GFL[23]	32.27	32.27	66.90	59.90	87.34	92.88	62.53	46.16	61.82	57.64
PAA[24]	32.12	51.58	66.79	66.26	89.56	96.41	62.20	35.90	54.10	63.12
TOOD[25]	32.03	50.53	62.66	53.94	82.20	91.86	60.51	35.36	53.81	60.95
VFNet[26]	32.72	48.39	66.17	60.94	90.91	95.25	61.29	39.41	56.46	58.93
FCOS[27]	32.12	51.58	62.63	62.27	93.19	63.65	63.10	40.06	52.37	63.79
DenoDet[11]	34.23	48.53	68.60	64.82	90.54	93.82	59.58	46.32	56.04	69.56
YOLOv12[8]	25.58	25.18	65.31	62.65	62.60	62.83	66.19	69.29	70.10	63.51
DSA–YOLO（本文模型）	27.26	27.12	69.62	68.97	70.98	70.55	67.89	73.60	70.60	64.76

4.4.2　消融实验结果

为验证DSA模块中SSA单元和FSA单元条带长度选择的合理性，本文设计了不同条带长度（3、5、7、9、11）的消融实验，在MSAR–1.0数据集与SAR–Aircraft–1.0数据集上分别验证条带长度 k_1 和 k_2 对模型性能的影响，实验结果如表3所示。实验中，SSA与FSA模块采用相同的条带长度组合，且保持其他实验参数一致。从表3可以看出，随着条带长度 k_1 和 k_2 的增大，两组实验结果中的mAP@0.5:0.95值均呈现先增大后减小的趋势，同时，不同条带长度组合对参数量及浮点运算次数的影响较小，可以忽略不计。当条带组合为 $k_1=5$ 且 $k_2=9$ 时，两组实验结果中的mAP@0.5:0.95值同时达到最大，即条带长度为5和9的组合能够更好地提取多尺度目标及小目标的特征，从而达到最佳的目标检测精度。

表 3　条带长度消融实验结果

Tab. 3　Ablation experiment results of different strip lengths

k_1	k_2	MSAR–1.0数据集			SAR–Aircraft–1.0数据集		
		参数量/10^6	FLOPs/10^9	mAP@0.5:0.95/%	参数量/10^6	FLOPs/10^9	mAP@0.5:0.95/%
3	3	27.25	6.82	69.51	27.25	27.11	68.37
3	5	27.25	6.82	69.88	27.25	27.11	68.53
3	7	27.25	6.83	70.28	27.25	27.11	68.79
3	9	27.26	6.83	70.28	27.26	27.12	69.02
3	11	27.26	6.83	69.53	27.26	27.12	68.87
5	5	27.26	6.83	69.87	27.26	27.12	68.63
5	7	27.26	6.83	70.34	27.26	27.12	69.25
5	9	27.26	6.83	70.97	27.26	27.12	69.62
5	11	27.26	6.83	70.24	27.26	27.12	69.31
7	7	27.26	6.83	69.93	27.26	27.12	69.01
7	9	27.27	6.83	70.55	27.27	27.13	68.92
7	11	27.27	6.83	70.21	27.27	27.13	68.88
9	9	27.27	6.83	70.02	27.27	27.13	68.73
9	11	27.27	6.83	70.30	27.27	27.13	68.78
11	11	27.28	6.83	70.06	27.28	27.14	68.64

更进一步，本文以YOLOv12模型为消融实验的基线，在SAR–Aircraft–1.0数据集上探究DSA1模块（位于主干网络的DSA模块）和DSA2模块（位于颈部网络的DSA模块）对模型的影响，实验结果如表4所示。从表4可以看出，相比于YOLOv12模型，仅加入DSA1模块后的模型的平均精度均值提高了2.07%，除ARJ21类别目标外，其余各类目标检测的平均精度均有所提高；仅加入DSA2模块后的模型的平均精度均值提高了3.10%，各细粒度分类目标检测的平均精度均有所提高；而同时加入DSA1模块和DSA2模块的模型的平均精度均值提高了4.31%，相较于仅加入DSA1模块和

仅加入DSA2模块的提升更为显著。通过比较，可以发现同时加入DSA1和DSA2模块后的模型在A320/321及A330两类目标上的检测平均精度的提升尤为显著，同时其余类别目标的检测平均精度相较YOLOv12模型也均有提升。以上结果充分证明了本文所提出模型的有效性。

表 4　模块消融实验结果

Tab. 4　Ablation experiment results of modules

DSA1	DSA2	参数量/10^6	FLOPs/10^9	mAP@0.5:0.95/%							
				总体	A220	A320	A330	ARJ21	B737	B787	其他
×	×	25.58	25.18	65.31	62.65	62.60	62.83	66.19	69.29	70.10	63.51
√	×	25.74	26.81	67.38	67.72	65.85	64.30	65.86	71.19	70.68	66.00
×	√	27.10	25.59	68.41	68.06	66.01	67.85	67.41	71.87	70.92	66.80
√	√	27.26	27.12	69.62	68.97	70.98	70.55	67.89	73.60	70.60	64.76

图 7 为基线模型 YOLOv12 与本文提出的 DSA–YOLO 模型在训练过程中分别在训练集与验证集上的框回归损失函数曲线、分类损失函数曲线和分布焦点损失函数曲线。从图 7 可以看出，无论是在训练集还是验证集上，DSA–YOLO 模型的框回归损失函数、分类损失函数以及分布焦点损失函数的收敛速度均快于 YOLOv12 模型，该结果也表明了与 YOLOv12 模型相比，本文提出的 DSA–YOLO 模型性能更加优越。

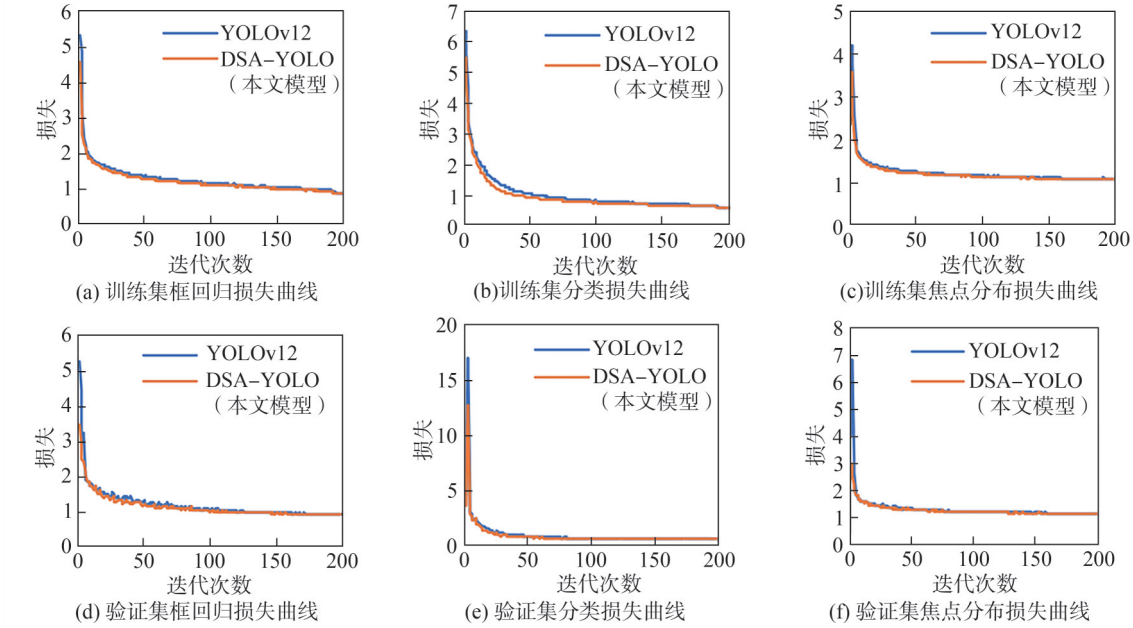


图 7　基线模型 YOLOv12 与本文模型 DAS–YOLO 损失函数曲线

Fig. 7　Loss function curves of baseline model YOLOv12 and proposed model DAS–YOLO

表 5 为 YOLOv12 与本文所提出的 DSA–YOLO 模型在 SAR–Aircraft–1.0 数据集上的精确率、召回率、虚警率、误检率、漏检率结果对照。从表 5 可以看出，DSA–YOLO 模型相较于 YOLOv12 模型，精确率提高了 3.28%，召回率提升了 2.55%，虚警率降低了 2.76%，误检率降低了 0.52%，漏检率降低了 2.55%，即本文 DSA–YOLO 模型在精确率、召回率、虚警率、误检率、漏检率 5 项检测指标上均优于 YOLOv12 模型。

表 5　模型检测结果对照表

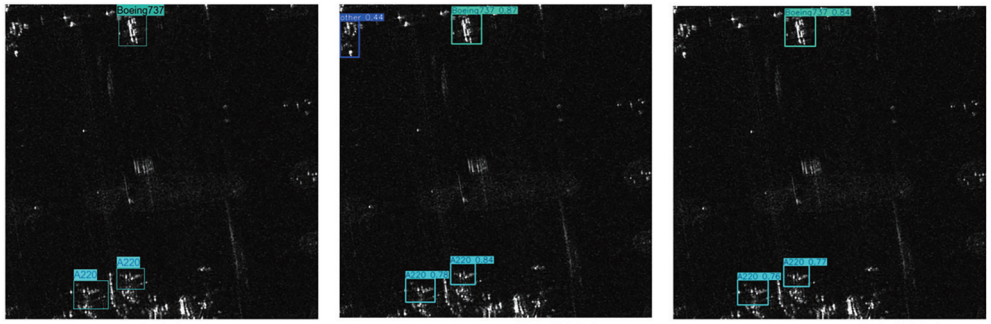
Tab. 5　Comparison table of model detection results

模型	精确率/%	召回率/%	虚警率/%	误检率/%	漏检率/%
YOLOv12	84.13	91.08	14.25	1.62	8.92
DSA–YOLO（本文模型）	87.41	93.63	11.49	1.10	6.37

图 8～10 为基线模型 YOLOv12 与本文所提出模型 DSA–YOLO 在测试集上虚警、漏检、误检的部分测试结果与真实标签的比较情况。

图 8 展示了虚警情形，即 YOLOv12 错误地将背景杂波检测为目标（图 8(b)），而此时 DSA–YOLO 的检测结果（图 8(c)）与真实标签（图 8(a)）相同。从表 5 中结果可知，本文提出的 DSA–YOLO 模型相较于 YOLOv12 模型，拥有更高的精确率与更低的虚警率，结合图 8 中情形，可知 DSA–YOLO 模型能够更精确地检测出目标，同时能够降低将背景杂波误判为目标的频率。

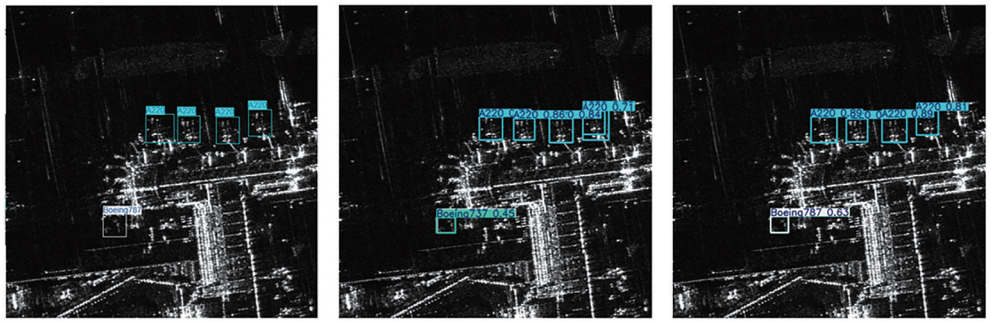
图 9 展示了误检情形，即 YOLOv12 错误地将目标 Boeing787 检测为 Boeing737（图 9(b)），而此时 DSA–YOLO 能够正确检测所有目标（图 9(c)），与真实



(a) 真实标签　(b) YOLOv12检测结果　(c)DSA-YOLO检测结果

图 8　虚警情形示意图

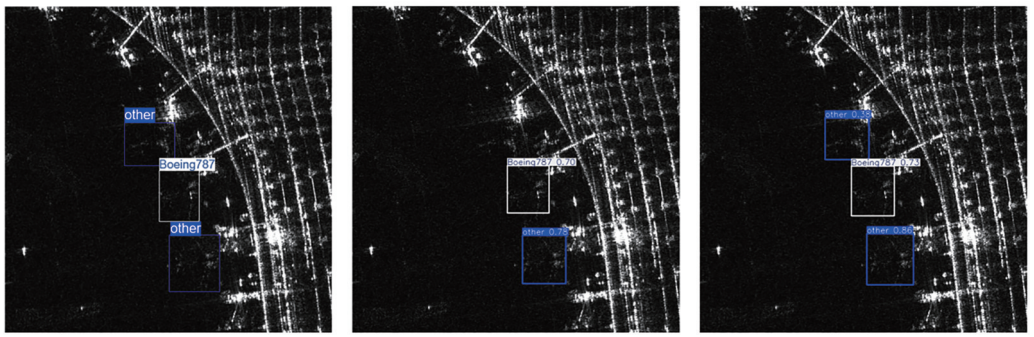
Fig. 8　Schematic diagram of false alarm scenarios



(a) 真实标签　(b) YOLOv12检测结果　(c) DSA-YOLO检测结果

图 9　误检情形示意图

Fig. 9　Schematic diagram of misclassification scenarios



(a) 真实标签　(b) YOLOv12检测结果　(c) DSA-YOLO检测结果

图 10　漏检情形示意图

Fig. 10　Schematic diagram of missed detection scenarios

标签（图 9(a)）相同。从表 5 可知，本文提出的 DSA-YOLO 模型相较于 YOLOv12 模型，拥有更高的精确率和更低的误检率，结合图 9 中情形，可知 DSA-YOLO 模型能够更精确地检测出目标，减少对目标的误报。

图 10 展示了漏检情形，即 YOLOv12 未能完整检测出所有存在的目标，仅检测到了部分目标（图 10(b)），而此时 DSA-YOLO 能够完整地检测出所有存在目标（图 10(c)），与真实标签（图 10(a)）相同。从表 5 可知，本文提出的 DSA-YOLO 模型相较于 YOLOv12 模型，拥有更高的召回率和更低的漏检率，结合图 10 中情形，可知本文的 DSA-YOLO 模型能够更准确地检测出目标，减少对目标的漏报。

综合定性与定量结果可以得出，本文提出的 DSA-YOLO 模型在复杂 SAR 图像场景下的目标检测性能显著优于 YOLOv12 模型，充分证实了本文模型的有效性。

5　结　论

针对 SAR 图像存在相干斑噪声干扰、目标与背景灰度对比度低、复杂场景下小目标特征易被掩盖、多尺度目标特征提取不充分等问题，本文在 YOLOv12 模型框架中融入双域带状注意力机制，通过对模型核心特征提取单元 C3k2 模块进行深度改进，构建出更适配 SAR 图像特性的特征提取网络，使模型能够更好地提取 SAR 图像中的空间特征与空间变化率特征，并将其有效融合，从而显著提高了针对 SAR 图像小目标与

多尺度目标的检测精度。与目标检测领域中领先的一系列模型相比，本文提出的DSA-YOLO模型表现出更加优越的性能，并进一步提高了在具有复杂环境的SAR图像中检测目标的性能，在SAR图像目标检测任务中有着良好的应用前景。

当前SAR目标检测模型多为特定任务设计，换个任务或场景就需重新训练，且过度依赖监督学习和大量标注样本。后续研究将进一步扩大无标注训练样本规模，优化自监督学习框架，设计适配SAR图像散斑噪声、离散目标外观等特性的主干网络。同时，进一步将具有明确物理意义的电磁散射特征融入模型，增强模型的可解释性。

参考文献：

[1] Singh P,Diwakar M,Shankar A,et al.A review on SAR image and its despeckling[J]. Archives of Computational Methods in Engineering,2021,28(7):4633–4653.

[2] Yasir M, Wan Jianhua, Xu Mingming, et al. Ship detection based on deep learning using SAR imagery: A systematic literature review[J].Soft Computing,2023,27(1):63–84.

[3] Luo Ru,Zhao Lingjun,He Qishan,et al.Intelligent technology for aircraft detection and recognition through SAR imagery: Advancements and prospects[J]. Journal of Radars,2024,13(2):307–330.[罗汝,赵凌君,何奇山,等.SAR图像飞机目标智能检测识别技术研究进展与展望[J].雷达学报,2024,13(2):307–330.]

[4] Wang Rufei, Xu Fanyun, Pei Jifang, et al. An improved faster R–CNN based on MSER decision criterion for SAR image ship detection in harbor[C]//Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium.Yokohama:IEEE,2019:1322–1325.

[5] Ke Xiao, Zhang Xiaoling, Zhang Tianwen, et al. SAR ship detection based on an improved faster R–CNN using deformable convolution[C]//Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS.Brussels:IEEE,2021:3565–3568.

[6] Ke Xiao,Zhang Xiaoling,Zhang Tianwen,et al.Sar ship detection based on swin transformer and feature enhancement feature pyramid network[C]//Proceedings of the 2022 IEEE International Geoscience and Remote Sensing Symposium.Kuala Lumpur:IEEE,2022:2163–2166.

[7] Ze Zhenghua, Han Jianghong. Transformer-in-transformer network with a complex-valued mixup module for SAR automatic target recognition[C]//Proceedings of the 2024 IEEE International Conference on Signal,Information and Data Processing(ICSIDP).Zhuhai:IEEE,2024:1–5.

[8] Tian Yunjie, Ye Qixiang, David Doermann. YOLOv12: Attention-centric real time object detectors[EB/OL].[2026–01–01].https://arxiv.org/abs/2502.12524.

[9] Wang Siyu, Gao Xin, Sun Hao, 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.[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.]

[10] Diao Wenhui,Dou Fangzheng,Fu Kun,et al.Aircraft detection in sar images using saliency based location regression network[C]//Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium.Valencia: IEEE,2018:2334–2337.

[11] Dai Yimian, Zou Minrui, Li Yuxuan, et al. DenoDet: Attention as deformable multisubspace feature denoising for target detection in SAR images[J]. IEEE Transactions on Aerospace and Electronic Systems,2025,61(2):4729–4743.

[12] Wang Xi, Xu Wei, Huang Pingping, et al. MSADNet: Multi-stage SAR aircraft target detection network[J]. IEEE Geoscience and Remote Sensing Letters,2025,22:4002105.

[13] Huang Bocheng, Zhang Tao, Quan Sinong, et al. Scattering enhancement and feature fusion network for aircraft detection in SAR images[J].IEEE Transactions on Circuits and Systems for Video Technology,2025,35(2):1936–1950.

[14] Chen Junyi,Shen Yanyun,Liang Yinyu,et al.YOLO–SAD: An efficient SAR aircraft detection network[J]. Applied Sciences,2024,14(7):3025.

[15] Nguyen P T,Cao L V.SSD–YOLOv12: An improved YOLOv12n model for ship detection using SAR data[J].Neural Computing and Applications,2026,38(3):40.

[16] Cui Yuning,Knoll A.Dual-domain strip attention for image restoration[J].Neural Networks,2024,171:429–439.

[17] 陈杰,黄志祥,夏润繁,等.大规模多类SAR目标检测数据集–1.0[DS/OL].雷达学报(2022)[2026–01–01].https://radars.ac.cn/web/data/getData dataType=MSAR.

[18] Wang Zhirui,Kang Yuzhuo,Zeng Xuan,et al.SAR–AIRcraft–1.0: High-resolution SAR aircraft detection and recognition dataset[J].Journal of Radars,2023,12(4):906–922.

[19] Ren Shaoqing, He Kaiming, 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.

[20] Zhang Shifeng, Chi Cheng, Yao Yongqiang, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR).Seattle:IEEE,2020:9756–9765.

[21] Li Yibing, Li Zifan, Ye Fang, et al. Reppoints-based multi-scale task enhancement network and sample assignment

method for oriented object detection[J]. IEEE Geoscience and Remote Sensing Letters,2023,20:6009305.

[22] Chen Zehui, Yang Chenhongyi, Chang Jiahao, et al. DDOD: Dive deeper into the disentanglement of object detector[J]. IEEE Transactions on Multimedia,2024,26:284–298.

[23] Li Xiang, Lv Chengqi, Wang Wenhai, et al. Generalized focal loss: Towards efficient representation learning for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2023,45(3):3139–3153.

[24] Kim K, Lee H S. Probabilistic anchor assignment with IoU prediction for object detection[C]//Computer Vision – ECCV 2020. Cham: Springer,2020:355–371.

[25] Feng Chengjian, Zhong Yujie, Gao Yu, et al. TOOD: Task-aligned one-stage object detection[C]//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision(ICCV). Montreal: IEEE,2021:3490–3499.

[26] Zhang Haoyang, Wang Ying, Dayoub F, et al. VarifocalNet: An IoU-aware dense object detector[C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Nashville: IEEE,2021:8510–8519.

[27] Tian Zhi, Shen Chunhua, Chen Hao, et al. FCOS: A simple and strong anchor-free object detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4):1922–1933.

A SAR Image Target Detection Method Based on Dual-domain Strip Attention

XU Pengfei[1], *ZHANG Boxiang*[2], *LI Yiyang*[2], *LI Wei*[1*]

(1.School of Aeronautics and Astronautics, Sichuan University, Chengdu 610065, China;
2.Chengdu Aircraft Design & Research Institute, Aviation Industry Corporation of China, Ltd., Chengdu 610032, China,)

Abstract:

Objective Accurate target detection in synthetic aperture radar (SAR) images with complex backgrounds is a challenging task in remote sensing information processing. These images are affected by multiplicative speckle noise, low gray-level contrast between targets and backgrounds, and insufficient feature extraction for small, multi-scale targets. Most existing detection models are adapted from optical image processing frameworks, which fail to fully exploit SAR frequency-domain information and lack targeted optimization for the physical characteristics of SAR images. This leads to low detection accuracy, overfitting, and poor generalization. In addition, recent attention-centric object detection frameworks have not been effectively adapted to SAR image detection tasks, creating a gap between advanced detection algorithms and SAR application requirements. A high-precision, lightweight detection model is urgently necessary to efficiently extract spatial and frequency domain features of SAR targets and improve detection performance for small, multi-scale targets in complex scenes.

Methods A dual-domain strip attention-based SAR target detection model (DSA–YOLO) was proposed by embedding the dual-domain strip attention (DSA) mechanism in the YOLOv12 framework. First, we reconstructed the C3k2 module, the core feature-extraction unit of YOLOv12, and integrated the DSA module to form the DSA–C3k2 module, which we embedded in the model's backbone and neck networks to enable collaborative extraction of SAR image features in the spatial and frequency domains. The DSA module comprised spatial strip attention (SSA) and frequency strip attention (FSA) units, with strip lengths of 5 and 9 set for multi-scale feature learning. The SSA unit enhanced sensitivity to structural texture and edge information in SAR targets through horizontal and vertical strip attention aggregation and expanded the network receptive field to highlight targets' spatial distribution features. The FSA unit separated the low- and high-frequency components of SAR images using strip average pooling, distinguishing the false high frequency caused by speckle noise from the true high frequency generated by target scattering centers, and highlighting the scattering and variation features of targets in the frequency domain. Second, we designed an improved multi-scale feature fusion strategy that integrated the dual-domain features extracted by the DSA–C3k2 module with the original features through splicing and upsampling operations. This approach leveraged the complementary advantages of low-level positional information and high-level semantic information, improving the model's adaptability to small and multi-scale targets. Finally, the MSAR–1.0 and SAR–Aircraft–1.0 datasets were used for experimental verification; the intersection over union (IOU) threshold was set to 0.50, and mean average precision from 0.50 to 0.95 (mAP@0.5:0.95) was used as the core evaluation index. We conducted comparative experiments with mainstream detection models, ablation experiments on the DSA module, and ablation experiments on different strip lengths to assess the effectiveness and rationale of the proposed method. Experimental hardware was configured with an Intel Core i7–core processor 9700 CPU and an NVIDIA GeForce RTX 3090 Ti GPU. The model was trained with the Adam optimizer with an initial learning rate of 0.01 and a batch size of 32.

Results and Discussions On the MSAR–1.0 dataset, the proposed DSA–YOLO model had 27.26×10^6 parameters and required 6.83×10^9 floating-point operations (FLOPs), an increase of only 1.68×10^6 parameters and 0.51×10^9 FLOPs compared with the YOLOv12 baseline model. The DSA–YOLO model achieved an mAP@0.5:0.95 of 70.97%, 2.52 percentage points higher than YOLOv12 and 1.07 percentage points higher than the second-best-performing model, DenoDet. On the SAR – Aircraft – 1.0 dataset for fine-grained target detection, the DSA – YOLO model had

27.26×10^6parameters and 27.12×10^9 FLOPs, a 1.94×10^9 increase in FLOPs compared with YOLOv12. The mAP@0.5:0.95 reached 69.62%, 4.31 percentage points higher than YOLOv12 and 1.02 percentage points higher than DenoDet. The DSA−YOLO model achieved the highest average precision across the four aircraft target types, showing excellent fine-grained detection performance. Ablation experiments on the DSA module showed that the model with only the backbone network DSA module (DSA1) achieved an mAP@0.5:0.95 of 67.38%, and the model with only the neck network DSA module (DSA2) achieved an mAP@0.5:0.95 of 68.41% on the SAR−Aircraft−1.0 dataset. The model with DSA1 and DSA2 modules achieved the highest mAP@0.5:0.95 of 69.62%, confirming the synergistic effect of the DSA module across different network positions. Ablation experiments with different strip lengths (3, 5, 7, 9, 11) showed that model performance improved with increasing strip length, then declined. When strip lengths were set to 5 and 9, the model achieved a peak mAP@0.5:0.95 of 70.97% on the MSAR−1.0 dataset and 69.62% on the SAR−Aircraft−1.0 dataset. In addition, varying strip length resulted in negligible changes in the model's number of parameters and FLOPs, confirming the rationale for selecting strip lengths 5 and 9 for the DSA module. Quantitative test-set results showed that the DSA−YOLO model's false alarm rate, misdetection rate, and missed detection rate were 11.49%, 1.10%, and 6.37%, respectively, 2.76, 0.52, and 2.55 percentage points lower than those of the YOLOv12 model. The loss function curve showed that DSA−YOLO converged faster on box regression loss, classification loss, and distribution focal loss on the training and validation sets, indicating better model stability and generalization ability. The proposed model only slightly increased computational overhead while improving detection accuracy, balancing detection performance and computational efficiency, and outperforming mainstream models such as Faster R−CNN, ATSS, and CenterNet in multi-category and fine-grained SAR target detection tasks.

Conclusions The DSA−YOLO model effectively addresses insufficient feature extraction and low detection accuracy in SAR images with complex backgrounds by fusing the dual-domain strip attention mechanism with the YOLOv12 framework. The collaborative extraction of spatial and frequency-domain features fully leverages SAR's physical imaging characteristics, effectively suppresses speckle noise interference, and significantly improves the feature representation ability of small and multi-scale targets. The improved multi-scale feature fusion strategy further enhances the model's adaptability to targets of different scales, and the DSA module's lightweight design ensures real-time performance. The experimental results on two public datasets confirm that the proposed model has higher detection accuracy and lower false alarm rate, missed detection rate, and misdetection rate compared with mainstream detection models, and detection performance is significantly improved with only a small increase in parameters and computational complexity. These results provide an effective, novel solution for SAR image target detection in complex backgrounds. The idea of fusing the dual-domain attention mechanism with advanced detection frameworks can be extended to other remote sensing image processing tasks, demonstrating important theoretical significance and engineering application value. The dual-domain feature extraction method, based on the physical characteristics of SAR imaging, can guide the design of specialized detection models for remote sensing images, and the model's lightweight improvement strategy has practical value for deploying SAR detection systems on edge devices.

Key words: synthetic aperture radar images; target detection; dual-domain strip attention; strip filtering

（编辑　吴芝明）

引用格式: Xu Pengfei,Zhang Boxiang,Li Yiyang,et al.A SAR image target detection method based on dual-domain strip attention[J].Advanced Engineering Sciences,2026,58(6):363–374.[徐鹏飞,张博祥,李艺洋,等.基于双域带状注意力的SAR图像目标检测方法[J].工程科学与技术,2026,58(6):363–374.]