

## 注意力机制与雷达图像智能处理:进展与展望

陈思伟\*, 戴林裕, 王欣超

(国防科技大学 电子科学学院, 湖南 长沙 410073)

**摘要:**为系统总结注意力机制与雷达图像智能处理方向的研究进展,本文梳理了注意力机制的发展脉络,并详细介绍了以通道注意力、空间注意力、自注意力和混合注意力等为代表的典型注意力机制模型。在此基础上,从雷达图像预处理、目标检测、图像分割、目标识别等任务出发,阐述了各类注意力机制在雷达图像智能处理中的创新应用。为了验证注意力机制的应用性能,基于雷达图像目标检测应用对不同注意力机制进行了性能对比分析。最后,从可解释性、高效注意力机制设计、多模态融合下的注意力机制优化、基础模型中的注意力机制设计四个方面进行了未来研究方向的讨论和展望。

**关键词:**注意力机制;雷达图像智能处理;深度学习

中图分类号:TN957.52 文献标志码:A 文章编号:1001-2486(2026)03-036-16

## Attention mechanism and intelligent processing of radar images: progress and prospects

CHEN Siwei\*, DAI Linyu, WANG Xinchao

(College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China)

**Abstract:** To systematically summarize the research progress of attention mechanisms and radar image intelligent processing, this paper reviewed the development context of attention mechanisms and elaborated on typical models, represented by channel attention, spatial attention, self-attention, and hybrid attention. On this basis, the innovative applications of various attention mechanisms in radar image intelligent processing were discussed from the perspectives of tasks such as radar image preprocessing, target detection, image segmentation, and target recognition. To verify the application performance of attention mechanisms, a comparative analysis of different attention mechanisms was conducted on the basis of radar image target detection tasks. Future research directions were discussed from four aspects: attention mechanism interpretability, efficient attention mechanism design, attention mechanism optimization for multimodal fusion, and attention mechanism design in foundation models.

**Keywords:** attention mechanism; intelligent processing of radar images; deep learning

合成孔径雷达 (synthetic aperture radar, SAR) 通过方位向合成孔径和距离向脉冲压缩处理能够获得感兴趣区域的高分辨二维雷达图像,直观呈现观测场景的几何结构与散射特性。作为一种主动式微波成像遥感装置, SAR 具备穿透性强、不受光照与气象条件限制、全天时全天候工作的独特优势,在军事侦察、灾害评估、海洋权益维护等关键领域发挥着不可替代的作用<sup>[1-3]</sup>。

随着深度学习技术的飞速发展,因其强大的自动特征提取能力,被广泛应用于雷达图像智能处理,推动该领域快速发展。然而,雷达图像固有

的相干斑现象<sup>[4]</sup>会导致目标特征模糊、纹理细节失真,加之复杂场景下目标与背景散射特性耦合,显著增加了目标智能识别等的难度。

人类视觉系统的核心优势之一在于能够自适应聚焦关键区域、屏蔽无关干扰信息,从而高效完成目标识别等任务<sup>[5]</sup>。受此启发,注意力机制被引入神经网络模型设计,通过自适应权重分配机制,使模型优先关注对任务重要的信息区域并抑制冗余干扰,为提升图像智能处理性能提供了有效途径<sup>[5-6]</sup>。近年来,注意力机制已在雷达图像预处理、目标检测、图像分割、目标识别等任

收稿日期:2026-01-19

基金项目:国家自然科学基金资助项目(U24B20189,62122091);湖南省科技创新领军人才资助项目(2024RC1040)

\*第一作者:陈思伟(1984—),男,四川泸州人,教授,博士,博士生导师,E-mail:chenswnudt@163.com

引用格式:陈思伟,戴林裕,王欣超. 注意力机制与雷达图像智能处理:进展与展望[J]. 国防科技大学学报, 2026, 48(3): 36-51.

Citation: CHEN S W, DAI L Y, WANG X C. Attention mechanism and intelligent processing of radar images: progress and prospects[J].

Journal of National University of Defense Technology, 2026, 48(3): 36-51.

务<sup>[7-12]</sup>中得到广泛应用。研究人员提出了一系列结合注意力机制的雷达图像智能处理算法,并取得应用性能的显著提升。

因此,及时和系统总结注意力机制与雷达图像智能处理的研究成果,深入分析各类方法的核心原理和应用性能,探讨当前面临的主要挑战并展望未来研究方向,对于推动该领域的持续创新与应用深化,具有重要的学术价值和实践意义。基于此,本文聚焦注意力机制与雷达图像智能处理,旨在系统性总结相关技术的基本原理、发展脉络与研究进展,为后续研究提供参考。

## 1 注意力机制的基本概念、发展脉络与典型结构

### 1.1 基本概念

注意力机制作为深度学习领域的重要技术之一<sup>[13]</sup>,其设计灵感源于对人类视觉认知系统的仿生模拟<sup>[5-6]</sup>。在自然感知过程中,受限于大脑有限的信息处理能力,人类并非对所有输入信息进行均等化处理,而是进化出“选择性关注”的高效认知策略——通过动态聚焦场景中的关键信息(如目标轮廓、核心特征),同时主动抑制冗余干扰信息(如无关背景、次要细节),实现有限认知资源下对高价值信息的快速筛选与精准提取,这一机制是人类高效应对复杂环境感知任务的核心基础。

受此启发,研究人员将注意力机制引入深度学习模型设计,其核心逻辑是通过自适应权重分配,对输入特征图中的不同元素赋予差异化重要性权重<sup>[14]</sup>。具体而言,模型可根据任务需求,自动强化与任务紧密相关的关键特征响应,同时弱化冗余干扰信息的影响<sup>[13]</sup>,从而在提升模型学习效率的同时,增强对核心特征的捕捉能力。更为重要的是,注意力机制能够显式建模输入数据与

输出结果之间的关联路径,使模型的决策过程在一定程度上可追溯和可解释,有效缓解了传统深度学习模型的“黑箱”特性<sup>[15]</sup>。

### 1.2 发展脉络

在过去10余年中,注意力机制在自然语言处理、计算机视觉、大模型等领域发挥着越来越重要的作用。图1简要梳理了注意力机制的主要演进节点<sup>[16-26]</sup>,其发展历程可划分为四个主要阶段<sup>[27]</sup>:

1) 第一阶段始于循环注意力模型<sup>[16]</sup>(recurrent attention model, RAM),该工作将深度神经网络与注意力机制结合,通过策略梯度优化框架,以端到端方式循环预测输入数据中的关键区域,并动态更新网络参数,为后续注意力机制的工程化应用提供了思路。受RAM启发,后续多项研究在视觉识别、目标定位等任务中采用了类似的循环注意力策略<sup>[28-29]</sup>,进一步验证了“选择性关注”在提升模型效率与精度方面的有效性。但此类方法高度依赖序列建模,存在并行计算效率较低的局限性。

2) 第二阶段实现了从“循环式关注”到“显式化特征选择”的核心突破,模型可直接定位判别性特征区域。Jaderberg等<sup>[17]</sup>提出的空间变换网络(spatial transformer networks, STN)通过学习仿射变换参数,实现了特征空间对齐与关键区域筛选。随后的可变形卷积网络<sup>[30-31]</sup>(deformable convolutional networks, DCNs)通过在卷积操作中引入自适应偏移量,使卷积核能够动态聚焦于目标相关的关键空间位置,改善了传统卷积的固定感受野限制,提升了模型对不规则目标和复杂场景的适应性。该阶段的主要标志是注意力机制与具体任务的结构化结合,通过显式建模关键特征的空间分布,为后续注意力机制的模块化设计提供了参考。

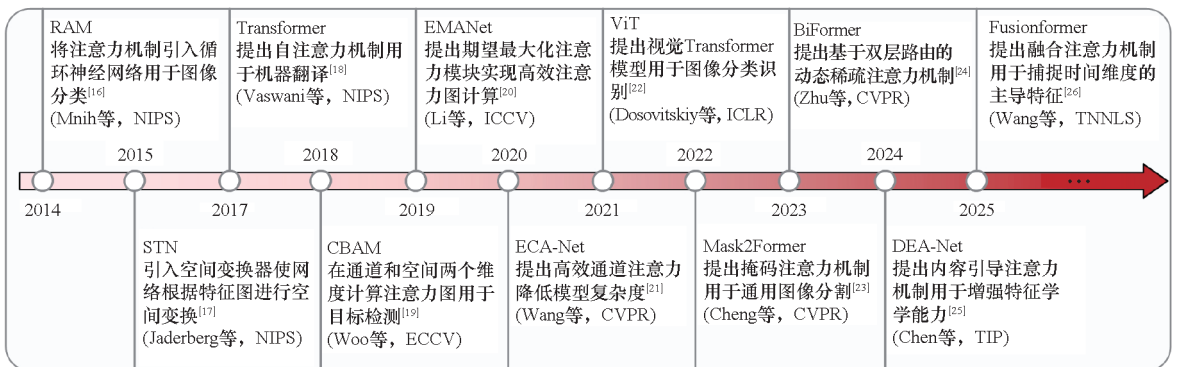


图1 注意力机制发展历程

Fig. 1 Development history of attention mechanisms

3) 第三阶段始于压缩-激发网络<sup>[32]</sup> (squeeze-and-excitation networks, SENet), 其提出了一种新颖的通道注意力机制, 使得特征建模维度从“空间域”拓展至“通道域”, 实现了对特征通道重要性的隐式自适应学习。SENet 通过“压缩-激发”操作, 自适应地校准各特征通道的权重, 其简洁高效的设计使其在图像分类等任务中取得了较大的性能提升。此后, 该方向涌现一系列代表性改进工作: 卷积块注意力模块<sup>[19]</sup> (convolutional block attention module, CBAM) 融合了通道与空间双重注意力, 实现了双维度的特征校准; 高效通道注意力网络<sup>[21]</sup> (efficient channel attention network, ECA-Net) 则简化通道交互方式, 在保持性能的同时降低了计算开销。该阶段的主要创新在于注意力机制从“显式空间定位”转向“隐式特征校准”, 提升了特征表征的判别能力与泛化性。

4) 第四阶段始于自注意力机制。2017 年, Vaswani 等<sup>[18]</sup> 提出的 Transformer 架构标志着注意力机制进入自注意力主导的新阶段。Transformer 最初在自然语言处理领域引起广泛关注<sup>[18, 33-34]</sup>, 成为后续大规模语言模型的核心架构。此外, Wang 等<sup>[35]</sup> 提出的非局部网络 (non-local network, NLN) 将自注意力机制引入计算机视觉领域, 有效提升了视频理解、目标检测等任务中对长距离依赖特征的捕捉能力。2021 年, Dosovitskiy 等提出的视觉 Transformer<sup>[22]</sup> (vision Transformer, ViT) 无须依赖卷积模块, 仅通过将图像划分为图元序列并进行自注意力建模, 就能在图像分类任务中取得与主流卷积神经网络 (convolutional neural network, CNN) 相当的性能。

另外, 期望最大化注意力网络<sup>[20]</sup> (expectation-maximization attention network, EMANet)、十字交叉网络<sup>[36]</sup> (criss-cross network, CCNet) 和上下文 Transformer 网络<sup>[37]</sup> (contextual Transformer network, CoT) 等一系列网络模型分别从计算效率优化、长距离依赖建模、注意力交互方式等角度对自注意力机制进行改进, 在速度、精度与泛化能力上实现了有效提升。近年来, 大量深度自注意力网络模型<sup>[22-26, 38-40]</sup> 相继出现, 进一步展现其在复杂场景建模、多模态融合等任务中的优势。随着模型架构的持续优化与算力的大幅提升, 基于自注意力的模型成为深度学习领域更具通用性和强表征能力的核心架构, 为各类智能任务的进一步发展提供技术底座。

### 1.3 典型的注意力机制模型

基于注意力机制模型原理, 可将图像智能处

理中主流的注意力机制划分为四大类: 通道注意力、空间注意力、自注意力和混合注意力。下面介绍这些注意力机制的核心原理和典型架构。

#### 1.3.1 通道注意力

通道注意力机制的核心在于通过自适应学习特征图各通道的重要性权重, 实现对任务相关通道的特征增强与冗余通道的信息抑制, 进而优化特征表征空间的判别能力与鲁棒性。SENet<sup>[32]</sup> 作为通道注意力领域的开创性工作, 构建了简洁且高效的“压缩-激发”架构, 为后续各类通道注意力模型的设计提供了范式, 其模型结构如图 2 所示, 包括一个压缩模块和一个激发模块。其中, 压缩模块通过全局平均池化, 将各通道的空间信息压缩为标量统计量, 实现通道级全局信息的聚合。该操作将二维空间特征映射为一维通道描述向量, 把输入特征图  $X_{in}$  转换为大小为  $1 \times 1 \times C$  的全局特征描述子, 从而完成通道级全局信息的提取与表征。激发模块则通过全连接层与非线性激活函数, 建模通道间的非线性依赖关系, 生成用于特征校准的注意力权重向量。该模块通过全连接层将压缩向量映射到低维空间以降低计算复杂度, 通过整流线性单元 (rectified linear unit, ReLU) 激活函数引入非线性变换以增强模型表达能力, 再经过第二个全连接层映射回原通道维度, 最终由 Sigmoid 激活函数输出归一化的注意力权重, 实现对原始特征图的通道级自适应校准, 表达式<sup>[32]</sup> 为:

$$V_{att} = \sigma \{ W_2 \delta [ W_1 \cdot \text{GAP}(X_{in}) ] \} \quad (1)$$

其中,  $V_{att}$  表示注意力向量,  $X_{in}$  表示输入特征,  $W_1$  和  $W_2$  表示全连接层的权重矩阵,  $\sigma$  和  $\delta$  分别表示 Sigmoid 和 ReLU 激活函数,  $\text{GAP}(\cdot)$  表示全局平均池化。

随后, 将生成的注意力向量与原始特征图逐通道相乘, 增强高权重通道、抑制低权重通道, 表达式为:

$$X_{out} = V_{att} \odot X_{in} \quad (2)$$

其中,  $X_{out}$  表示输出特征,  $\odot$  表示逐通道相乘。

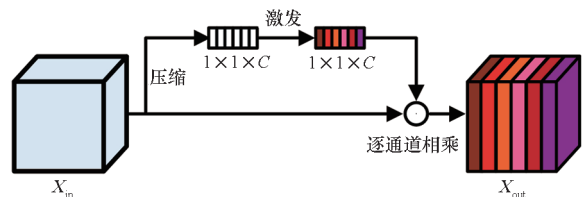


图 2 通道注意力机制示意图<sup>[32]</sup>

Fig. 2 Schematic of channel attention mechanism<sup>[32]</sup>

尽管 SENet 凭借简洁的架构与优异的通道校

准效果,成为通道注意力领域的里程碑工作,但其自身仍存在不足。比如,在压缩模块中,全局平均池化过于简单,该操作仅能提取通道内特征的一阶统计均值,面对雷达图像这类含强杂波、多尺度目标的复杂数据时,难以有效聚合全局上下文信息与局部关键特征;在激发模块中,全连接层虽能刻画非线性关联,但是增加了模型参数量与计算复杂度。针对上述不足,后续研究使用不同的策略对 SENet 进行了改进。例如,全局二阶池化网络<sup>[41]</sup> (global second-order pooling network, GSoP-Net)提升了压缩模块输出质量;ECA-Net<sup>[21]</sup>优化了激发模块以降低复杂度;基于风格的重校准模块<sup>[42]</sup> (style-based recalibration module, SRM)同时对压缩模块和激发模块进行了优化;混合局部通道注意力机制<sup>[43]</sup> (mixed local channel attention, MLCA)同时利用通道、局部通道和全局通道信息,并对检测效果、速度和模型参数量之间的平衡进行了设计。

### 1.3.2 空间注意力

空间注意力本质是面向特征空间维度的自适应区域选择与资源分配机制,其核心目标是将有限的计算与特征表达资源聚焦于目标所在的关键空间区域。

Mnih 等提出的 RAM<sup>[16]</sup>是空间注意力的开创性工作,该模型融合了循环神经网络<sup>[44]</sup> (recurrent neural network, RNN)和强化学习<sup>[45]</sup> (reinforcement learning, RL)技术,通过迭代式感知交互策略自主定位目标关键区域,并依托策略梯度优化实现端到端的空间区域选择,为后续空间注意力的学习范式提供了重要参考。除 RAM 外,STN<sup>[17]</sup>、感受野注意力卷积<sup>[46]</sup> (receptive field attention convolution, RFACConv)、聚合激励网络<sup>[47]</sup> (gather-excite network, GENet)和 NLN<sup>[35]</sup>代表了不同类型的空间注意力方法。STN 属于显式空间变换与区域预测方法,通过独立子网络学习仿射变换参数,直接对输入特征执行空间对齐、裁剪与尺度缩放操作,可精准显式地提取目标区域特征。RFACConv 将空间注意力与卷积运算深度耦合,在动态扩展网络感受野范围的同时,实现对关键特征区域的聚焦感知,进而强化模型对复杂场景中目标的特征捕捉能力。GENet 属于软掩码驱动的隐式空间选择方法,通过子网络生成像素级空间软掩码,对不同空间位置赋予自适应的重要性权重,无须显式裁剪操作即可完成关键信息增强,具备更好的泛化性与灵活性。NLN 代表自注意力关联方法,通过建模像素间的全局依赖关系,自适

应捕捉长距离空间关联信息,打破了传统局部感受野的限制,能够有效挖掘跨区域的关键特征关联。典型的空注意力机制如图 3 所示。对于给定的输入特征图  $X_{in}$ ,首先沿通道轴应用平均池化和最大池化操作,将两种池化结果拼接以构建包含丰富空间结构信息的特征描述符,随后通过卷积层和激活函数来生成空间注意力图  $X_{out}$ ,其表达式<sup>[19]</sup>为:

$$X_{out} = \sigma \{ f_{\theta} [ \text{AvgPool}(X_{in}); \text{MaxPool}(X_{in}) ] \} \quad (3)$$

其中,  $f_{\theta}$  表示卷积操作,  $\text{AvgPool}(\cdot)$  和  $\text{MaxPool}(\cdot)$  分别表示平均池化和最大池化操作。

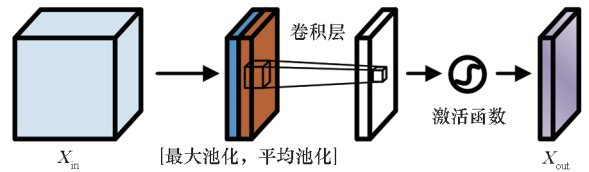


图3 空间注意力机制示意图<sup>[19]</sup>

Fig. 3 Schematic of spatial attention mechanism<sup>[19]</sup>

### 1.3.3 自注意力

自注意力机制的创新在于通过直接建模单一数据序列内部元素间的全局依赖关系,摆脱了传统卷积局部感受野与循环神经网络序列顺序的约束,能够自适应生成具备全局上下文感知能力的特征表征,为非局部、长距离关联特征挖掘提供了通用且高效的技术方案。在文本序列、图像分块序列等不同形式的数据结构中,每个元素通过计算与其他所有元素的相关性权重,动态聚合全局有效信息,形成自身的增强特征表示。该相关性建模方式无须依赖人工预设的局部窗口或固定序列先后顺序,可自主挖掘元素间隐藏的非线性关联与远距离依赖,适用于挖掘非局部、非线性的复杂依赖关系。2017年,Transformer<sup>[18]</sup>模型将自注意力机制应用于机器翻译任务,其通过“查询-键-值”的三重向量交互模式,高效建模文本序列的长距离语义依赖,在当时取得了最优的翻译性能。Transformer 架构如图 4 所示。自注意力机制可表示<sup>[18]</sup>为:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax} \left( \frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} \right) \mathbf{V} \quad (4)$$

其中: $\mathbf{Q}$ 为查询向量,表示当前关注的词; $\mathbf{K}$ 为键向量,表示序列中所有词的身份标识; $\mathbf{V}$ 为值向量,包含每个词的实际信息; $d_k$ 为键向量的维度, $\sqrt{d_k}$ 用于缩放点积,避免维度过高导致的数值溢出或梯度消失问题,确保训练过程的稳定性。

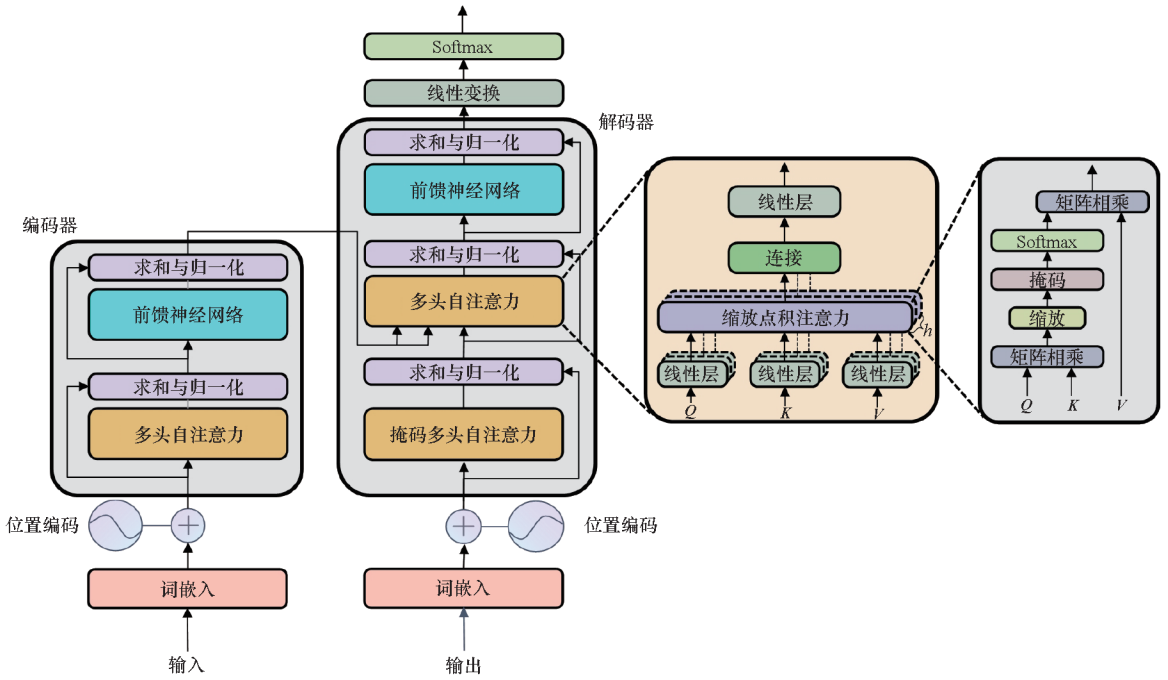


图 4 Transformer 架构示意图<sup>[18]</sup>

Fig. 4 Schematic of Transformer architecture<sup>[18]</sup>

随着技术发展,自注意力机制的应用领域迅速从自然语言处理拓展至计算机视觉领域。2021 年, Dosovitskiy 等提出的 ViT 模型<sup>[22]</sup> 通过将图像均匀切分为固定尺寸的图元并线性映射为序列特征,直接迁移 Transformer 自注意力架构完成图像分类,验证了自注意力模型在图像分类等任务中的有效性。ViT 的成功表明:自注意力机制能够有效捕捉图像像素间的全局空间关联,为图像处理任务提供新的技术路径,也为其在雷达遥感图像等复杂场景处理中的应用奠定了基础。

### 1.3.4 混合注意力

混合注意力机制的设计理念是突破单一注意力机制的建模局限,通过有机融合不同类别的注意力模型(如通道注意力、空间注意力、自注意力等),实现各机制间的优势互补,进而更全面、精准地挖掘数据内部的复杂特征关联,为各类复杂视觉任务提供更强的特征学习与表征能力。2017 年, Chen 等<sup>[48]</sup> 提出空间 - 通道注意力卷积网络 (spatial and channel-wise attentions in CNN, SCA-CNN), 将空间注意力与通道注意力同步融入图像描述任务,借助多层特征图动态调节语句生成所需的上下文信息,开启了多维度注意力融合的研究思路。Woo 等<sup>[19]</sup> 提出了 CBAM, 通过串行结合通道注意力与空间注意力,让模型从通道重要性区分、空间位置聚焦两个维度对特征进行精细化筛选与增强,既保留了通道注意力对全局特征

的校准能力,又兼顾空间注意力对局部关键区域的精准捕捉优势。

在此之后,研究人员提出了全局注意力机制<sup>[49]</sup> (global attention mechanism, GAM)、空间通道协同注意力机制<sup>[50]</sup> (spatial and channel synergistic attention, SCSA) 等混合注意力机制。GAM 重点解决传统混合注意力信息损耗大、跨维度交互弱的问题,兼顾信息完整性与通道 - 空间交互效果。SCSA 则摒弃简单叠加思路,专注于挖掘空间与通道注意力的内在协同效应,通过双向引导优化进一步提升特征校准精度,推动混合注意力机制持续完善。

## 2 注意力机制在雷达图像智能处理中的发展与应用

当前,雷达图像处理正朝着高分辨率、智能化、多源融合方向加速发展。注意力机制作为深度学习领域的重要技术,凭借其“选择性关注关键信息、抑制冗余干扰”的特点,在雷达图像智能处理中发挥着重要作用。下文将系统总结不同类型注意力机制在雷达图像智能处理任务中的应用原理和技术发展情况。

### 2.1 通道注意力与雷达图像处理

雷达图像的多通道特征通常受到冗余信息干扰,导致目标判别性特征被掩盖。通道注意力机制通过自适应分配特征通道权重,能够针对性强

化对目标检测、分类、识别等任务至关重要的判别性信息,同时抑制无关或低质量通道的干扰,从而提升模型的鲁棒性与任务精度。本小节介绍通道注意力机制在雷达图像智能处理中的典型应用与技术创新。

Dai 等<sup>[51]</sup>利用通道注意力机制进行极化雷达图像超分辨率重建,将模糊核宽度等退化先验信息经多层感知机转化为权重向量,以逐通道乘法方式对极化雷达图像特征图进行动态加权,使网

络能自适应不同退化形式,从而提高超分辨率重建性能。Su 等<sup>[52]</sup>提出全频域通道注意力网络 (complete frequency channel attention network, CFCANet) 以提升雷达图像场景分类精度,如图 5 所示。该方法改变 SENet 仅依赖全局平均池化的单一信息聚合方式,引入二维离散余弦变换对特征图进行频域分解,通过为每个通道分配完整低频分量,充分挖掘利用了集中于低频区域的目标有效信息。

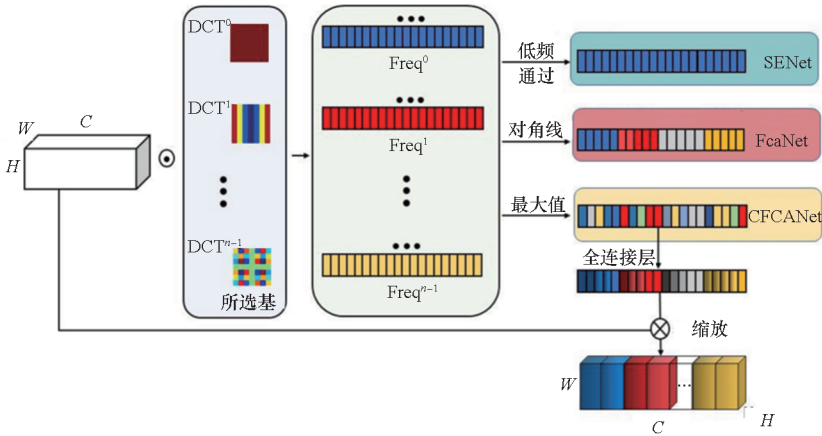


图 5 全频域通道注意力示意图<sup>[52]</sup>

Fig. 5 Schematic of full-frequency domain channel attention<sup>[52]</sup>

Lang 等<sup>[53]</sup>提出轻量化级联多域注意力网络 (lightweight cascaded multidomain attention network, LW-CMDANet),聚焦雷达图像自动目标识别,融合了离散余弦变换 (discrete cosine transform, DCT) 与离散小波变换 (discrete wavelet transform, DWT) 的域特征提取优势。LW-CMDANet 包含多频谱注意力模块和多分辨率频谱注意力模块。其中,多频谱注意力模块将输入特征图按通道分组并施加不同 DCT 频率分量,经逆 DCT 重构生成注意力权重,同时捕捉目标全域轮廓与局部细节;多分辨率频谱注意力模块则将特征分解为低频、水平高频及垂直高频分量,通过“压缩-激发”操作动态校准各子带权重,强化对目标方位角变化敏感的高频纹理特征,实现轻量化与高精度的平衡。

Li 等<sup>[54]</sup>设计了混合卷积通道注意力 (mixed convolution channel attention, MCCA) 模块用于雷达图像舰船检测,通过交替执行二维卷积与一维卷积操作,分别建立单个通道与所有通道的全局关联性,以及相邻通道的局部依赖,形成双路权重融合机制,能够有效强化关键通道、抑制无效通道。Liu 等<sup>[55]</sup>提出双通道注意力模块用于雷达图像相

干斑滤波,如图 6 所示,通道注意力模块自适应地缩放每个通道特征,像素注意力模块使网络能够充分关注信息量更丰富的特征区域。该双通道注意力模块使网络聚焦富含纹理信息的有效通道,从而提高滤波性能。

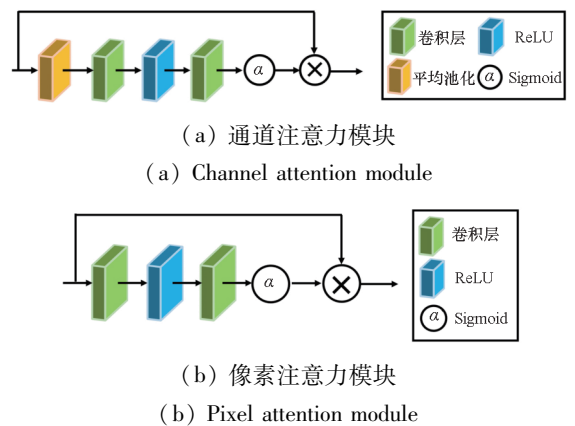


图 6 双通道注意力模块<sup>[55]</sup>

Fig. 6 Dual-channel attention module<sup>[55]</sup>

Ma 等<sup>[56]</sup>将通道注意力嵌入特征提取网络,增强对雷达图像舰船目标关键特征的捕捉能力,提升复杂近岸场景下的舰船检测性能。Wang 等<sup>[57]</sup>提出非局部通道注意力网络 (non-local

channel attention network, NLCA Net) 用于陆海目标检测与识别,通过引入空间金字塔池化(spatial pyramid pooling, SPP)捕获像素与区域间的长程依赖,缓解传统卷积的局部限制,增强对目标全局上下文信息的提取能力,降低了近岸场景下的虚警率。Yu 等<sup>[58]</sup>提出高效多分支跨通道注意力(efficient multibranch cross-channel attention, EMC<sup>2</sup>A)模块以提升雷达图像目标分类的精度与鲁棒性,其采用局部多尺度特征交互策略,先通过全局平均池化压缩多分支提取的不同感受野特征并生成通道级统计向量,再经通道混洗整合多尺度信息和一维循环卷积实现跨通道交互,最后利用 Sigmoid 激活函数实现通道间非互斥竞争,使网络动态聚焦高贡献特征通道。Zhang 等<sup>[59]</sup>提出基于图像级注意力的自适应多视角融合网络(adaptive multiview fusion network based on image attention, IA-AMF-Net),通过提取输入图像的通道特征权重向量,再经跨图像 Softmax 归一化计算各视角注意力权重,自适应评估不同视角分类特征的重要性,既突出高贡献视角特征和抑制冗余,又保持融合特征维度恒定,增强了多视角场景下的鲁棒性。E 等<sup>[60]</sup>提出卷积压缩-激发小波神经网络(convolutional squeeze-and-excitation wavelet neural network, CSWNN)用于雷达图像变化检测,其通道注意力模块先对卷积层输出特征图进行全局平均池化以生成通道描述符,再通过全连接层与 Sigmoid 函数学习通道重要性权重,对原始特征加权重构,强化对变化区域敏感的特征响应。Xu 等<sup>[9]</sup>提出多模态互补注意力模块(multimodal complementary attention module, MCAM)用于光学与雷达图像多模态语义分割,该模块不仅计算各模态自身的通道注意力权重,还引入跨模态互补注意力机制,使两种模态相互引导,利用模态一的注意力权重优化另一模态的特征响应,强化异源图像间的语义互补性。Yuan 等<sup>[61]</sup>在频域分支中嵌入卷积-通道注意力模块(convolution operations and a channel attention module, Conv-CAM),用于雷达图像小样本目标识别,通过自适应校准频域特征的通道权重,增强模型对目标关键频率成分的聚焦能力,有效缓解小样本场景下特征不足的问题。

这些方法从频域增强、通道关联建模、任务场景适配等不同角度优化通道注意力机制,验证了通道注意力机制在雷达图像智能处理任务中的有效性。

## 2.2 空间注意力与雷达图像处理

雷达图像通常受到复杂杂波干扰,目标几何特征易被掩盖,而空间注意力机制通过动态加权图像空间位置特征,能够突出目标核心结构、抑制冗余干扰,进而提升检测、分割、识别等任务的精度。本小节系统梳理空间注意力机制在雷达图像智能处理中的典型应用与技术创新。

文献[62]利用空间注意力机制进行 SAR 图像有源干扰类型识别,取得不同场景和不同干扰参数情形下更高的识别精度和更稳健的性能。Gao 等<sup>[63]</sup>借鉴人眼视觉系统的空间选择机制,构建融合自下而上与自上而下的空间注意力模型用于雷达图像目标检测。其中,自下而上阶段通过提取强度与方向显著图快速定位潜在目标区域,自上而下阶段则利用训练数据学到的最优权重对特征图加权融合,进一步强化真实目标显著性并抑制复杂背景杂波。Wang 等<sup>[64]</sup>提出多视角注意力卷积与长短时记忆网络用于自动目标识别,其空间注意力机制通过计算特征图空间权重,动态聚焦目标区域关键散射特征,同时抑制背景杂波干扰,从而有效提升特征判别性,帮助模型精准定位不同方位角下的目标结构,进而提高识别性能。

针对雷达图像多尺度舰船检测,Hu 等<sup>[65]</sup>设计局部注意力模块(local attention module, LAM)和非局部注意力模块(non-local attention module, NLAM),如图 7 所示。其中,LAM 引入可变形卷积,通过自适应学习采样点空间偏移使卷积核更适配舰船形态变化,高效捕捉多角度舰船局部特征;NLAM 则通过计算全局像素间相关性抑制海杂波与陆地干扰,增强上下文语义信息利用。二者协同提升检测精度与鲁棒性。

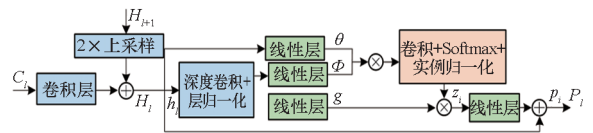


图 7 非局部注意力模块<sup>[65]</sup>

Fig. 7 Non-local attention module<sup>[65]</sup>

Li 等<sup>[66]</sup>设计的全局注意力模块(global attention module, GAM)面向舰船分割任务,其对高层特征进行全局平均池化以提取上下文信息,并生成指示目标类别的空间注意力权重。在此基础上,将权重与低层特征图相乘,强化目标区域特征响应并抑制杂波干扰,使网络聚焦舰船边缘与轮廓细节。Meng 等<sup>[67]</sup>提出的层注意力模块针对雷达图像变化检测,先对多层特征进行加权变换,

再通过 Softmax 归一化的相关矩阵捕获层间空间依赖关系,实现低层空间细节与高层语义信息在空间维度的互补增强,提升变化检测性能。为应对雷达图像目标尺度差异大、强散射点稀疏的挑战,Chang 等<sup>[68]</sup>提出自适应尺度分布注意力(adaptive scale distribution attention, ASA),利用不同膨胀率的空洞卷积捕获多尺度目标散射分布特征,使网络能自适应感知从局部细节到全局上下文的空间信息,提升目标检测性能。Dai 等<sup>[69]</sup>提出的多视角注意力(multi-view attention, ViA)机制,聚焦小样本识别任务,通过胶囊向量的高维空间关系建模,强化具有判别性的视角特异性特征,同时抑制角度敏感导致的低判别性特征干扰,结合联合动态路由聚合多视角互补空间信息,改善小样本识别性能。Shang 等<sup>[70]</sup>提出的舰船增强注意力网络(ship augmented attention network, SA<sup>2</sup>Net),通过引入与舰船形状匹配且长宽比大的矩形卷积核,增强对舰船几何特征的捕捉能力,进而提升识别性能。在跨模态与图像修复类任务中,Zhang 等<sup>[71]</sup>将空间注意力模块引入雷达辅助光学图像云去除模型,使模型能自动定位云区域并生成与云厚度相关的注意力图,在恢复云下信息的同时最大程度保留非云区域细节。Zhang 等<sup>[72]</sup>提出的多尺度空间注意力(multiscale spatial attention, MSSA)模块,利用不同窗口大小的 Transformer 块捕捉多尺度全局空间信息,增强模型对海冰不同尺度特征的提取能力,提升极地海冰与开阔水域的分类性能。Tu 等<sup>[73]</sup>则针对光学与雷达图像多模态融合云去除任务,提出相似性特征注意力(similarity feature attention, SFA),如图 8 所示,通过计算两类图像特征向量间的余弦相似度生成可见性注意力图,有效区分云覆盖与无云区域,指导光学编码器聚焦无云区域特征提取、抑制云区干扰,提高融合后图像质量。

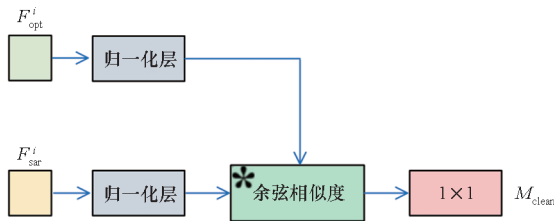


图 8 相似性特征注意力模块<sup>[73]</sup>

Fig. 8 Similarity feature attention module<sup>[73]</sup>

这些方法从目标定位、尺度适配、跨模态协同等不同角度优化空间注意力机制,适配雷达图像的成像特性与任务需求,验证了空间注意力机制

在提升目标信息提取精度、抑制复杂干扰等方面的有效性。

### 2.3 自注意力与雷达图像处理

雷达图像常蕴含分布式目标散射点关联、大范围地形结构连续性等复杂空间关联特性,而传统卷积操作受限于局部感受野,难以充分建模这些全局依赖。自注意力机制通过动态捕捉全局上下文信息,改善了局部建模的局限,有望应对上述挑战。本小节系统梳理自注意力机制在雷达图像智能处理中的典型应用与技术创新。

Cui 等<sup>[74]</sup>利用自注意力机制提升雷达图像与光学图像的匹配精度,通过为每个特征分配可学习权重,使网络自动聚焦于对配准最具判别力的关键区域,抑制纹理贫乏或冗余区域的干扰。Hu 等<sup>[75]</sup>提出图自注意力(graph self-attention, GSA)机制用于同类跨模态匹配任务,通过计算同一模态内节点间的自注意力权重,聚合局部邻域信息以增强特征判别力,有效挖掘模态内结构上下文信息,使模型自适应聚焦配准关键区域。针对雷达图像变化检测,Dong 等<sup>[76]</sup>提出多尺度自注意力深度聚类(multiscale self-attention deep clustering, MSDC)模块,如图 9 所示。MSDC 通过计算不同尺度特征图之间的关联权重,自适应聚合全局上下文信息,使网络聚焦于对变化敏感的关键区域,有效区分真实变化与伪变化。

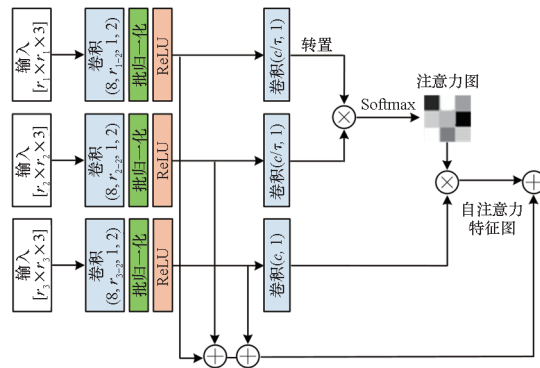


图 9 多尺度自注意力深度聚类模块<sup>[76]</sup>

Fig. 9 Multi-scale self-attention deep clustering module<sup>[76]</sup>

Peng 等<sup>[77]</sup>则在频域内利用自注意力机制分别建模高频细节与低频全局特征的依赖关系。其中,高频分支通过局部窗口注意力提取边缘等细节信息,低频分支借助全局注意力捕捉图像整体结构,同时结合平均池化降低计算复杂度,保留弱目标纹理特征,避免传统方法过度平滑导致的细节丢失。在小样本雷达图像目

标分类任务中, Yang 等<sup>[78]</sup> 提出混合损失图注意力网络(mixed loss graph attention network, MGA-Net), 其每一层自注意力机制通过可学习的特征更新注意力权重分配, 引导节点特征在信息聚合过程中强化类内相似性、弱化类间干扰, 缓解传统图神经网络的过平滑问题, 通过保留节点自身特征与加权相邻特征的拼接, 使深层网络保持特征判别力。

Zhu 等<sup>[79]</sup> 通过 Swin Transformer 模块建模雷达图像中冰川与冰架前沿的长距离空间依赖关系, 其窗口自注意力机制不仅增强了网络对全局上下文信息的感知能力, 还通过跨窗口信息交互提升了边界区域的判别精度。Li 等<sup>[80]</sup> 提出自注意力卷积网络(self-attention convolutional neural network, SACNN) 用于雷达图像目标检测, 通过捕获杂波的长程空间相关性, 使网络同时建模目标的局部稀疏性和杂波的全局连续性, 在图像重建过程中通过动态权重分配强化目标区域显著性和抑制杂波干扰, 从而提升输出图像的信杂比和特征保真度。Saleh 等<sup>[81]</sup> 在维度交互注意力(dimension-wise interactive attention, DIA) 模块中引入自注意力机制以提升雷达图像洪水检测性能。不同于传统自注意力在二维空间的高复杂度计算, DIA 将注意力分解为通道-通道、通道-宽度和通道-高度三个维度分别建模, 使计算复杂度降至线性级别的同时保留全局建模能力, 既降低了模型计算负担, 又增强了对洪水边界、细小水体变化等关键特征的感知能力。在跨模态图像修复与增强任务中, Pan 等<sup>[82]</sup> 提出混合动态残差自注意力网络(hybrid dynamic residual self-attention network, HDRSA-Net) 用于雷达图像辅助的光学图像云影去除。其引入的自注意力机制在通道维度上建模全局上下文关系, 通过计算特征通道间的交叉协方差, 降低计算复杂度并保留全局结构信息, 使网络能在高分辨率遥感图像中自适应筛选和融合高信息量特征, 避免冗余干扰, 提升云和阴影区域重建的精度。Patnaik 等<sup>[83]</sup> 将自注意力机制用于雷达图像上色任务, 通过建模全局像素间的依赖关系, 有效捕捉河流、植被、城市边界等关键地物特征的空间结构, 相比传统 CNN 的局部感知, 自注意力通过计算图像任意两个位置的相关性, 使模型在上色过程中保持大范围空间一致性, 避免因纹理模糊导致的色彩失真。

这些方法从跨模态匹配、变化检测、目标分类与检测、图像修复等不同任务出发, 通过灵活设计

自注意力的建模维度、计算方式与融合策略, 充分适配雷达图像的全局关联特性, 有效提升了雷达图像智能处理性能。

### 2.4 混合注意力与雷达图像处理

雷达图像兼具复杂的空间结构与多通道互补物理散射信息, 单一注意力机制难以全面捕捉关键信息并抑制杂波, 而混合注意力机制通过有机融合不同维度注意力的优势, 能精准聚焦任务相关信息, 有效抑制杂波干扰。本小节系统梳理混合注意力机制在雷达图像智能处理中的典型应用与技术创新。

Li 等<sup>[10]</sup> 提出双分支关联聚合图注意力网络(dual-branch correlation aggregation graph attention network, DCA-Net) 用于简缩极化逆合成孔径雷达(inverse synthetic aperture radar, ISAR) 空间目标部件识别。DCA-Net 利用图注意力机制的动态权重分配机制强化了部件间的拓扑关联, 提高了空间目标部件识别性能。Pei 等<sup>[84]</sup> 在极化 ISAR 空间目标极化信息重构网络中引入 CBAM, 动态生成注意力权重, 使网络优先关注卫星部件的强散射区域与结构边缘, 抑制背景杂波及无效区域的干扰, 提升了极化信息重构精度。Deng 等<sup>[85]</sup> 利用 CBAM 自适应筛选多尺度特征中的关键极化信息, 提升了双极化 SAR 到全极化 SAR 的数据重建精度与物理一致性。Bao 等<sup>[86]</sup> 设计稀疏注意力引导的细粒度金字塔模块, 通过融合空间与通道注意力机制缓解雷达图像细粒度飞机检测中“类间差异小、类内差异大”的核心挑战。其中, 空间注意力帮助模型在复杂背景中定位飞机关键局部结构, 通道注意力则动态增强对散射特性敏感的特征通道, 二者协同从多尺度特征中自适应提取并聚合最具判别力的细粒度信息, 提升不同型号飞机的检测精度与区分能力。Yang 等<sup>[87]</sup> 提出三重注意力(triplet attention, TriAtt) 机制用于视频雷达图像多阴影目标跟踪, 如图 10 所示。TriAtt 通过三个并行分支分别捕获通道与高度、宽度维度的跨维度交互特征, 强化阴影在帧间

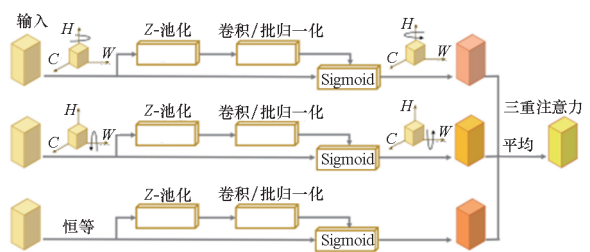


图 10 三重注意力机制<sup>[87]</sup>

Fig. 10 Triple attention mechanism<sup>[87]</sup>

微弱且稀疏的特征一致性,在不显著增加参数的前提下,使网络每帧都能自动聚焦对匹配与关联最具判别力的局部结构,抑制背景杂波干扰,降低目标漏检率和误检率。

Yan 等<sup>[88]</sup>提出多级空间和通道混合注意力(multilevel spatial and channel-wise attention, MSCA)机制用于视频雷达图像地面动目标检测,该机制先利用空间注意力在浅层特征图上强化亮线边缘与位置等关键浅层信息,再通过通道注意力在中层对动目标散射特征最敏感的通道进行动态加权,抑制杂波与断裂亮线带来的虚警。Zhang 等<sup>[89]</sup>提出多级注意力孪生网络用于光学与雷达图像关键点检测,借助混合注意力机制分别在浅层和深层提取空间细节与语义信息,缓解成像机理差异导致的非线性辐射差异和几何畸变问题,提升异源图像匹配精度。针对双极化雷达图像舰船分类任务,Xie 等<sup>[90]</sup>融合了 ConvNeXt 注意力模块与通道-空间注意力模块(channel spatial attention module, CSAM)。其中,ConvNeXt 注意力模块在特征提取阶段引入多头自注意力,使网络聚焦舰船结构中类别相关的关键区域,抑制复杂杂波干扰;CSAM 则在融合阶段同时建模通道间相关性和空间上下文信息,自适应增强对分类有益的极化特征通道与空间位置,实现更准确的舰船识别。

Fan 等<sup>[91]</sup>将通道注意力残差块(channel attention residual block, CARBlock)与全局注意力块(global attention block, GABlock)结合用于雷达图像干扰抑制。其中,CARBlock 沿通道维度自适应加权,强化目标回波特征并抑制干扰,避免传统方法因模型失配导致的信号损失;GABlock 则进一步建模全局频点相关性,修复特征分离过程中可能被削弱的目标细节,降低回波失真。Imani<sup>[92]</sup>融合自注意力与交叉注意力机制用于极化雷达图像分类。其中,自注意力建模极化特征内部的全局依赖关系;交叉注意力则挖掘极化散射机制与上下文信息之间的深层关联,使网络在训练样本有限的情况下仍能准确捕捉地物类别的复杂特征,提升分类精度。He 等<sup>[93]</sup>结合多头自注意力模块(multi-head self-attention block, MSAB)和交叉注意力模块(cross-attention block, CAB)用于雷达图像辅助的光学图像云雾去除任务。MSAB 建模全局上下文关系,增强厚云区域细节恢复能力;CAB 通过对齐并融合共享特征与独立特征,缓解模态异构带来的负面影响,避免直接堆叠特征导致的信息混淆,提升跨模态信息互

补效果。

这些方法通过灵活组合空间注意力、通道注意力、自注意力等不同机制,针对雷达图像智能处理的各类核心任务实现了优势互补,验证了混合注意力机制在雷达图像智能处理任务中的有效性。

### 3 雷达图像目标检测中的注意力机制性能对比

#### 3.1 雷达图像目标检测

随着深度学习的发展,以 CNN 为代表的目标检测算法被广泛应用于雷达图像智能处理领域,显著提升了目标检测的性能。当前雷达图像目标检测算法主要分为两类:双阶段检测器(two-stage detectors)和单阶段检测器(one-stage detectors)。双阶段检测器以 Fast R-CNN<sup>[94]</sup>为代表,通过生成候选区域,对每个区域进行分类与回归。该类方法检测精度高,尤其适用于小目标和密集场景,在雷达图像舰船检测、车辆检测等任务中表现优异,但存在计算成本高、训练复杂的缺点。单阶段检测器以 YOLO (you only look once) 系列<sup>[95]</sup>为代表,可直接预测目标的位置与类别。然而,标准 CNN 结构在处理雷达图像时仍存在局限性,其局部感受野难以建模长距离关系,且对感兴趣区域(如强散射点、目标轮廓)关注不足。因此,通过引入注意力机制来引导目标检测模型聚焦于判别性更强的区域,可进一步提升检测性能。近年来,多种注意力模块(如 SE、CBAM、Transformer 多头注意力机制等)已被应用于雷达图像目标检测任务,并取得提升。

在本文雷达图像目标检测的注意力机制性能对比实验中,综合考量算法的性能、训练复杂度、模型可扩展性等因素,选取 YOLO 系列的 YOLOv11s<sup>[96]</sup>作为实验的检测算法。在注意力机制选取方面,实验选取 GAM<sup>[49]</sup>、RFACConv<sup>[46]</sup>、CoT<sup>[37]</sup>、SCSA<sup>[50]</sup>和 MLCA<sup>[43]</sup>5 种注意力机制进行性能对比。

#### 3.2 数据集与评估指标

该实验环境为 Windows 10 操作系统和 Pytorch 深度学习框架,计算机配置为 Intel Xeon Gold 5315Y@3.2 GHz 处理器,256 GB 内存,四块显存为 48 GB 的 NVIDIA GeForce RTX A6000 显卡。

为对比不同注意力机制在雷达图像目标检测任务中的性能增益,选取舰船目标检测数据集 HRSID<sup>[97]</sup>和飞机目标检测数据集 SAR-AIRCRAFT-

1.0<sup>[98]</sup>进行实验。HRSID 数据集中包含 5 604 张舰船目标图像, SAR-AIRcraft-1.0 数据集包含 5 类飞机目标共 4 368 张图像。训练集和测试集的划分方式与原文献<sup>[97-98]</sup>保持一致,其中 HRSID 为 6.5 : 3.5, SAR-AIRcraft-1.0 为 8 : 2, 实验数据信息如表 1 所示, 数据示例如图 11 所示。

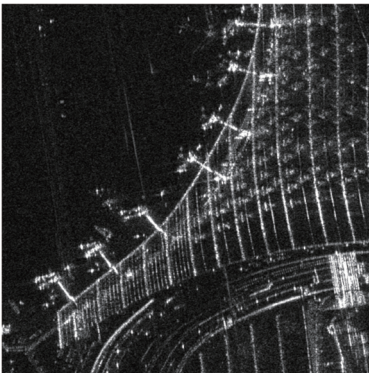
表 1 实验数据集信息

Tab. 1 Experimental dataset information

数据集	发布时间	样本数量	波段
SAR-AIRcraft-1.0	2023	4 368	C
HRSID	2020	5 604	C/X
数据集	极化方式	分辨率/m	训练测试比
SAR-AIRcraft-1.0	单极化	1	8 : 2
HRSID	单极化	0.5 ~ 3	6.5 : 3.5



(a) HRSID 舰船目标  
(a) HRSID ship target



(b) SAR-AIRcraft-1.0 飞机目标  
(b) SAR-AIRcraft-1.0 aircraft target

图 11 SAR 目标检测数据集样本示例

Fig. 11 Sample examples of SAR target detection dataset

对比实验采用的评价指标包括准确率  $P$ 、召回率  $R$ 、平均精度均值  $mAP@0.5$  和  $mAP@0.5 : 0.95$ , 表达式<sup>[97-98]</sup>分别为:

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$mAP = \frac{\int_0^1 P(R) dR}{n} \quad (7)$$

其中:  $TP$  为正确检测出的目标数;  $FN$  为漏检目标数;  $FP$  为错检目标数;  $n$  为目标类别数;  $P(R)$  为准确率和召回率的曲线;  $mAP$  取交并比 (intersection over union, IoU) 为 0.5 和 0.5 ~ 0.95 两种情况下数值。

### 3.3 对比实验结果分析

实验选取 YOLOv11s 作为性能对比的基准, 并在 YOLOv11s 的基础上分别添加所选取的 5 种注意力机制。所有实验均采用统一的超参数设置: 迭代次数为 100; 初始学习率为  $1 \times 10^{-3}$ ; Batch Size 为 16; 输入包含舰船目标图像的尺寸为 800 像素  $\times$  800 像素, 包含飞机目标图像的尺寸为 1 024 像素  $\times$  1 024 像素; 优化器采用 AdamW。数据增强方式均采用随机翻转和拼接, 随机因子为 0.5。表 2 展示了注意力模块添加位置、参数量、推理时的计算量与实时性能。

表 2 注意力模块的实时性能评估

Tab. 2 Real-time performance evaluation of attention modules

注意力机制类型	注意力模块位置	参数量	计算量/GFLOPs	推理时间/ms
基准		$9.4 \times 10^6$	21.3	167.0
GAM	骨干网络	$11.2 \times 10^6$	22.7	151.9
RFACConv	骨干网络	$10.9 \times 10^6$	28.6	185.2
CoT	骨干网络	$11.8 \times 10^6$	24.6	154.5
SCSA	颈部网络	$9.4 \times 10^6$	21.3	141.7
MLCA	颈部网络	$9.4 \times 10^6$	21.4	152.0

表 3 和表 4 对比了不同注意力机制在舰船及飞机目标检测任务中各项性能指标的绝对值增益。在舰船目标检测中, RFACConv 在四项指标中的提升均为最高, 分别为 2.69%、1.35%、2.56% 和 3.02%。在飞机目标检测实验中, GAM 带来的准确率提升最高, 为 4.26%; MLCA 带来的召回率提升最高, 为 3.25%; CoT 带来的  $mAP@0.5$  和  $mAP@0.5 : 0.95$  提升最高, 分别为 1.84%、6.95%。

表3 雷达图像舰船目标检测任务中不同注意力机制性能对比

Tab.3 Performance comparison of different attention mechanisms for ship target detection in radar images

注意力机制类型	准确率 $P/\%$	召回率 $R/\%$	mAP@0.5/ $\%$	mAP@0.5;0.95/ $\%$
基准	88.13	79.63	87.55	60.60
GAM	90.47(+2.34)	78.91(-0.72)	89.27(+1.72)	61.19(+0.59)
RFACConv	<b>90.82(+2.69)</b>	<b>80.98(+1.35)</b>	<b>90.11(+2.56)</b>	<b>63.62(+3.02)</b>
CoT	89.27(+1.14)	80.86(+1.23)	89.59(+2.04)	62.69(+2.09)
SCSA	90.22(+2.09)	80.42(+0.79)	89.20(+1.65)	62.20(+1.60)
MLCA	88.60(+0.47)	80.80(+1.17)	89.00(+1.45)	61.40(+0.80)
平均提升	1.75	0.76	1.88	1.62

表4 雷达图像飞机目标检测任务中不同注意力机制性能对比

Tab.4 Performance comparison of different attention mechanisms for aircraft target detection in radar images

注意力机制类型	准确率 $P/\%$	召回率 $R/\%$	mAP@0.5/ $\%$	mAP@0.5;0.95/ $\%$
基准	82.47	81.24	84.14	56.14
GAM	<b>86.73(+4.26)</b>	82.22(+0.98)	85.95(+1.81)	60.88(+4.74)
RFACConv	85.60(+3.13)	81.91(+0.67)	85.88(+1.74)	61.56(+5.42)
CoT	85.78(+3.31)	82.47(+1.23)	<b>85.98(+1.84)</b>	<b>63.09(+6.95)</b>
SCSA	84.53(+2.06)	80.38(-0.86)	84.22(+0.08)	59.97(+3.83)
MLCA	81.98(-0.49)	<b>84.49(+3.25)</b>	84.34(+0.20)	60.00(+3.86)
平均提升	2.45	1.05	1.13	4.96

综上,卷积注意力最大程度上保留了目标结构空间信息,在舰船目标检测中 RFACConv 均取得最高增益。而飞机目标密集排列且背景复杂,不同注意力机制在不同指标上各有优势:GAM 有效提升准确率,说明其通过建模长距离依赖和全局上下文信息,有效抑制了虚警;MLCA 提升召回率,表明其对局部细节和边缘结构的增强有助于减少漏检;CoT 在 mAP@0.5;0.95 指标上提升达 6.95%,说明其在多尺度 IoU 阈值下显著提升了模型的定位精度。

#### 4 总结与展望

本文系统总结了注意力机制在雷达图像处理领域的应用,首先分析了注意力机制的四阶段发展脉络。在此基础上,以雷达图像智能处理任务为纲,综述了各类注意力机制在不同任务场景下的创新应用与具体实现。最后,开展了雷达图像目标检测实验,对 5 种代表性注意力机制进行了性能对比与分析。接下来对注意力机制研究进行展望。

1)注意力机制可解释性:注意力机制为雷达图像智能处理提供了“选择性关注”的核心能力,

但现有研究大多仅聚焦性能提升,可解释性研究仍处于起步阶段。未来应建立注意力机制与雷达散射物理机理的关联,突破数据驱动黑箱局限,通过量化分析注意力权重与目标电磁散射特性、几何结构的对应关系,揭示注意力聚焦关键区域与通道的物理合理性,实现模型决策过程可追溯、可验证,为军事侦察、定量遥感等高精度应用提供理论支撑。同时,可构建面向雷达图像特性的可解释性评估体系,设计专用评价指标,形成雷达图像的可解释性分析方法。

2)高效注意力机制设计:高效注意力机制是实现雷达图像智能处理实时化、轻量化与高精度的关键支撑,未来可围绕性能提升与资源优化的动态平衡展开深入研究。针对雷达图像高分辨率、多通道、大尺度等特点,可采用稀疏化建模、维度分解、参数共享等策略设计轻量化注意力架构,在保留全局特征关联建模能力的同时降低计算复杂度与内存开销,满足机载、星载等资源受限平台的实时处理需求。此外,还需强化注意力机制与雷达成像物理先验的深度融合,将散射机理、相干斑特性、极化通道相关性等知识嵌入权重计算过程,抑制无效特征的冗余计算,实现物理引导下的

高效特征聚焦。

3) 多模态融合下的注意力机制优化: 在雷达与光学、红外、激光等多模态数据融合处理中, 现有注意力机制普遍存在模态权重分配固化、跨模态特征关联挖掘不充分等问题, 严重制约了融合性能的提升。未来研究可面向多模态融合任务开展专用注意力机制的深度优化。一方面是突破静态权重分配的局限, 构建基于模态质量评估的动态注意力调节模块, 根据各模态数据可靠性实时自适应分配权重。另一方面可探索跨模态语义关联注意力, 通过搭建模态间语义映射网络, 挖掘雷达散射特征、光学目标轮廓、红外热辐射与目标材质间的内在关联, 使注意力不仅聚焦单模态显著区域, 更能依托多模态语义一致性增强目标关键特征表达, 为全天候、复杂环境下的雷达多模态智能处理提供稳定高效的技术支撑。

4) 基础模型中的注意力机制设计: 近年来, 研究者正积极探索适用于雷达图像的基础模型架构<sup>[99]</sup>。但由于雷达数据具备的独特物理特性, 直接迁移基于自注意力的通用视觉 Transformer 模型, 极易产生表征偏差与语义鸿沟等问题。未来的注意力机制设计可突破仅依赖空间或谱域特征的建模方式, 深度融合雷达物理先验知识, 实现从“空间主导”向“物理信息融合”的范式转变。同时, 考虑到雷达图像具有高空间分辨率、宽覆盖范围的特点, 其计算与内存开销通常巨大, 因此构建适配超高分、大幅宽雷达图像的专用注意力架构, 是推动相关基础模型走向实际工程应用的关键。

## 参考文献 (References)

- [1] CHEN S W, WANG X S, XIAO S P, et al. Target scattering mechanism in polarimetric synthetic aperture radar [M]. Singapore: Springer Singapore, 2018.
- [2] CHEN S W. Imaging radar polarimetric rotation domain interpretation: theory and application [M]. Boca Raton: CRC Press, 2024.
- [3] CUMMING I G, WONG F H. Digital processing of synthetic aperture radar data [M]. Boston, MA: Artech House, 2005.
- [4] DAI L Y, CHEN S W. Context2Context: a zero-shot SAR image speckle filter [J]. IEEE Transactions on Geoscience and Remote Sensing, 2026, 64: 5204911.
- [5] OLSHAUSEN B A, ANDERSON C H, VAN ESSEN D C. A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information [J]. Journal of Neuroscience, 1993, 13(11): 4700–4719.
- [6] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254–1259.
- [7] KO J, LEE S. SAR image despeckling using continuous attention module [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 3–19.
- [8] TIAN Y, ZHOU Z, CUI Z Y, et al. Scene adaptive SAR incremental target detection via context-aware attention and Gaussian-box similarity metric [J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 5205217.
- [9] XU D D, QIAN J, FENG H, et al. Semantic segmentation of multimodal optical and SAR images with multiscale attention network [J]. IEEE Geoscience and Remote Sensing Letters, 2025, 22: 6007605.
- [10] LI M D, XIAO S P, CHEN S W. Compact polarimetric ISAR space target components recognition with dual-branch correlation aggregation graph attention network [J]. IEEE Transactions on Geoscience and Remote Sensing, 2025, 63: 5107917.
- [11] LIU X C, SUN Z X, DING H W, et al. GAFF: global attention feature flow network for optical and SAR image registration under geometric transformations [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2026, 19: 4629–4646.
- [12] DONG F B, YIN Q, ZHANG J, et al. Polarimetric SAR salt crust classification via autoencoded and attention-enhanced feature representation [J]. Remote Sensing, 2026, 18(1): 164.
- [13] NIU Z Y, ZHONG G Q, YU H. A review on the attention mechanism of deep learning [J]. Neurocomputing, 2021, 452: 48–62.
- [14] LV H X, CHEN J L, PAN T Y, et al. Attention mechanism in intelligent fault diagnosis of machinery: a review of technique and application [J]. Measurement, 2022, 199: 111594.
- [15] GAO Y, RUAN Y J. Interpretable deep learning model for building energy consumption prediction based on attention mechanism [J]. Energy and Buildings, 2021, 252: 111379.
- [16] MNIH V, HEES N, GRAVES A, et al. Recurrent models of visual attention [C]//Proceedings of the 28th International Conference on Neural Information Processing Systems, 2014: 2204–2212.
- [17] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks [C]//Proceedings of the 29th International Conference on Neural Information Processing Systems, 2015.
- [18] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017.
- [19] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Proceedings of the 15th European Conference on Computer Vision, 2018: 3–19.
- [20] LI X, ZHONG Z S, WU J L, et al. Expectation-maximization attention networks for semantic segmentation [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 9166–9175.
- [21] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: efficient channel attention for deep convolutional neural networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 11531–11539.
- [22] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16 × 16 words: transformers for image recognition at scale [C]//Proceedings of the International Conference on Learning Representations (ICLR), 2021.

- [23] CHENG B, MISRA I, SCHWING A G, et al. Masked-attention mask transformer for universal image segmentation [ C ]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition ( CVPR ), 2022: 1280 – 1289.
- [24] ZHU L, WANG X, KE Z, et al. BiFormer: vision transformer with bi-level routing attention [ C ]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition ( CVPR ), 2023: 10323 – 10333.
- [25] CHEN Z, HE Z, LU Z M. DEA-Net: single image dehazing based on detail-enhanced convolution and content-guided attention [ J ]. IEEE Transactions on Image Processing, 2024, 33: 1002 – 1015.
- [26] WANG C, WANG Z, DONG H, et al. Fusionformer: a novel adversarial Transformer utilizing fusion attention for multivariate anomaly detection [ J ]. IEEE Transactions on Neural Networks and Learning Systems, 2025, 36(8): 14479 – 14492.
- [27] GUO M H, XU T X, LIU J J, et al. Attention mechanisms in computer vision: a survey [ J ]. Computational Visual Media, 2022, 8(3): 331 – 368.
- [28] GREGOR K, DANIELKA I, GRAVES A, et al. DRAW: a recurrent neural network for image generation [ C ]// Proceedings of 32nd International Conference on Machine Learning, 2015.
- [29] XU K, BA J L, KIROUS R, et al. Show, attend and tell: neural image caption generation with visual attention [ C ]// Proceedings of the 32nd International Conference on Machine Learning, 2015: 2048 – 2057.
- [30] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks [ C ]// Proceedings of the IEEE International Conference on Computer Vision ( ICCV ), 2017: 764 – 773.
- [31] ZHU X Z, HU H, LIN S, et al. Deformable ConvNets V2: more deformable, better results [ C ]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition ( CVPR ), 2019: 9300 – 9308.
- [32] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [ C ]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7132 – 7141.
- [33] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [ C ]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019: 4171 – 4186.
- [34] YANG Z L, DAI Z H, YANG Y M, et al. XLNet: Generalized autoregressive pretraining for language understanding [ C ]// Proceedings of the 33rd Conference on Neural Information Processing Systems ( NeurIPS 2019 ), 2019.
- [35] WANG X L, GIRSHICK R, GUPTA A, et al. Non-local neural networks [ C ]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 7794 – 7803.
- [36] HUANG Z L, WANG X G, WEI Y C, et al. CCNet: criss-cross attention for semantic segmentation [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(6): 6896 – 6908.
- [37] LI Y H, YAO T, PAN Y W, et al. Contextual Transformer networks for visual recognition [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(2): 1489 – 1500.
- [38] GUO M H, CAI J X, LIU Z N, et al. PCT: point cloud transformer [ J ]. Computational Visual Media, 2021, 7(2): 187 – 199.
- [39] YUAN L, CHEN Y P, WANG T, et al. Tokens-to-token ViT: training vision transformers from scratch on ImageNet [ C ]// Proceedings of the IEEE/CVF International Conference on Computer Vision ( ICCV ), 2021: 538 – 547.
- [40] YUAN L, HOU Q B, JIANG Z H, et al. VOLO: vision outlooker for visual recognition [ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(5): 6575 – 6586.
- [41] GAO Z L, XIE J T, WANG Q L, et al. Global second-order pooling convolutional networks [ C ]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition ( CVPR ), 2019: 3019 – 3028.
- [42] LEE H, KIM H E, NAM H. SRM: a style-based recalibration module for convolutional neural networks [ C ]// Proceedings of the IEEE/CVF International Conference on Computer Vision ( ICCV ), 2019: 1854 – 1862.
- [43] WAN D H, LU R S, SHEN S Y, et al. Mixed local channel attention for object detection [ J ]. Engineering Applications of Artificial Intelligence, 2023, 123: 106442.
- [44] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [ J ]. Neural Computation, 1997, 9(8): 1735 – 1780.
- [45] SUTTON R S, MCALLESTER D, SINGH S, et al. Policy gradient methods for reinforcement learning with function approximation [ C ]// Proceedings of the 13th International Conference on Neural Information Processing Systems, 1999: 1057 – 1063.
- [46] ZHANG X, LIU C, YANG D G, et al. RFACConv: receptive-field attention convolution for improving convolutional neural networks [ EB/OL ]. ( 2023 – 04 – 06 ) [ 2026 – 01 – 17 ]. <https://arxiv.org/abs/2304.03198>.
- [47] HU J, SHEN L, ALBANIE S, et al. Gather-excite: exploiting feature context in convolutional neural networks [ C ]// Proceedings of the 32nd Conference on Neural Information Processing Systems ( NeurIPS 2018 ), 2018.
- [48] CHEN L, ZHANG H W, XIAO J, et al. SCA-CNN: spatial and channel-wise attention in convolutional networks for image captioning [ C ]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition ( CVPR ), 2017: 6298 – 6306.
- [49] LIU Y C, SHAO Z R, HOFFMANN N. Global attention mechanism: retain information to enhance channel-spatial interactions [ EB/OL ]. ( 2021 – 12 – 10 ) [ 2026 – 01 – 17 ]. <https://arxiv.org/abs/2112.05561>.
- [50] SI Y Z, XU H Y, ZHU X Z, et al. SCSA: exploring the synergistic effects between spatial and channel attention [ J ]. Neurocomputing, 2025, 634: 129866.
- [51] DAI L Y, LI M D, CHEN S W. PolSAR image super-resolution with polarimetric degradation modulation network [ J ]. IEEE Transactions on Aerospace and Electronic Systems, 2025, 61(6): 17938 – 17954.
- [52] SU B, LIU J, SU X, et al. CFCANet: a complete frequency channel attention network for SAR image scene classification [ J ]. IEEE Journal of Selected Topics in Applied

- Earth Observations and Remote Sensing, 2021, 14: 11750 – 11763.
- [53] LANG P, FU X J, FENG C, et al. LW-CMDANet: a novel attention network for SAR automatic target recognition [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 6615 – 6630.
- [54] LI C N, LANG H T. An oriented SAR ship detector with mixed convolution channel attention module and geometric nonmaximum suppression [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 8074 – 8084.
- [55] LIU S Q, LEI Y, ZHANG L Y, et al. MRDDANet: a multiscale residual dense dual attention network for SAR image denoising [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5214213.
- [56] MA X R, HOU S L, WANG Y Y, et al. Multiscale and dense ship detection in SAR images based on key-point estimation and attention mechanism [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5221111.
- [57] WANG Z X, XIN Z H, LIAO G S, et al. Land-sea target detection and recognition in SAR image based on non-local channel attention network [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5237316.
- [58] YU X, QIAN Y, GENG Z, et al. EMC<sup>2</sup>A-net: an efficient multibranch cross-channel attention network for SAR target classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5210314.
- [59] ZHANG R L, DUAN Y Z, ZHANG J D, et al. An adaptive multiview SAR automatic target recognition network based on image attention [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 13634 – 13645.
- [60] E J H, WANG L, ZHAO C H, et al. SAR images change detection based on attention mechanism-convolutional wavelet neural network [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2025, 18: 12133 – 12149.
- [61] YUAN S L, GAO H Y, GU X Y, et al. Synthetic aperture radar target recognition with limited training data based on frequency-domain-assisted dual-stream attention hierarchical deformable convolutional networks [J]. Engineering Applications of Artificial Intelligence, 2025, 161: 112309.
- [62] 陈思伟, 崔兴超, 李铭典, 等. 基于深度 CNN 模型的 SAR 图像有源干扰类型识别方法 [J]. 雷达学报, 2022, 11(5): 897 – 908.  
CHEN S W, CUI X C, LI M D, et al. SAR image active jamming type recognition based on deep CNN model [J]. Journal of Radars, 2022, 11(5): 897 – 908. (in Chinese)
- [63] GAO F, LIU A D, LIU K, et al. A novel visual attention method for target detection from SAR images [J]. Chinese Journal of Aeronautics, 2019, 32(8): 1946 – 1958.
- [64] WANG C W, LIU X Y, PEI J F, et al. Multiview attention CNN-LSTM network for SAR automatic target recognition [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 12504 – 12513.
- [65] HU Q, HU S H, LIU S Q. BANet: a balance attention network for anchor-free ship detection in SAR images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5222212.
- [66] LI J C, GOU S P, LI R M, et al. Ship segmentation via encoder-decoder network with global attention in high-resolution SAR images [J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 4016605.
- [67] MENG D S, GAO F, DONG J Y, et al. Synthetic aperture radar image change detection via layer attention-based noise-tolerant network [J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 4026505.
- [68] CHANG H, FU X J, DONG J, et al. MLSNet: multiclass lightweight SAR detection network based on adaptive scale distribution attention [J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 4010305.
- [69] DAI Q J, ZHANG G, XUE B, et al. Capsule-guided multi-view attention network for SAR target recognition with small training set [J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 4012605.
- [70] SHANG Y Z, PU W, LIAO D L, et al. SA<sup>2</sup>Net: ship augmented attention network for ship recognition in SAR images [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16: 10036 – 10049.
- [71] ZHANG S, LI X D, ZHOU X Y, et al. Cloud removal using SAR and optical images via attention mechanism-based GAN [J]. Pattern Recognition Letters, 2023, 175: 8 – 15.
- [72] ZHANG Z, DENG G B, LUO C Y, et al. A multiscale dual attention network for the automatic classification of polar sea ice and open water based on sentinel-1 SAR images [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 5500 – 5516.
- [73] TU K L, YANG C, QING Y X, et al. Cloud removal with optical and SAR imagery via multimodal similarity attention [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2025, 226: 116 – 126.
- [74] CUI S, MA A L, ZHANG L P, et al. MAP-Net: SAR and optical image matching via image-based convolutional network with attention mechanism and spatial pyramid aggregated pooling [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1000513.
- [75] HU X, WU Y, LIU X Y, et al. Intra-and inter-modal graph attention network and contrastive learning for SAR and optical image registration [J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5220216.
- [76] DONG H H, MA W P, JIAO L C, et al. A multiscale self-attention deep clustering for change detection in SAR images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5207016.
- [77] PENG S H, DENG Y K, WANG C L. A frequency self-attention MIMO network for SAR weak targets denoising and sidelobe suppression [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2026, 19: 10741 – 10754.
- [78] YANG M J, BAI X R, WANG L, et al. Mixed loss graph attention network for few-shot SAR target classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5216613.
- [79] ZHU Q, GUO H D, ZHANG L, et al. GLA-STDeepLab: SAR enhancing glacier and ice shelf front detection using Swin-TransDeepLab with global-local attention [J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5218113.

- [80] LI M, HUO W B, WU J J, et al. SAR image reconstruction method for target detection using self-attention CNN-based deep prior learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5222714.
- [81] SALEH T, HOLAIL S, ZAHRAN M, et al. LiST-Net: enhanced flood mapping with lightweight SAR transformer network and dimension-wise attention[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5211817.
- [82] PAN J, XU J G, YU X Y, et al. HDRSA-Net: hybrid dynamic residual self-attention network for SAR-assisted optical image cloud and shadow removal[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2024, 218: 258 – 275.
- [83] PATNAIK N, RAJ R, MISRA I, et al. Demystifying SAR with attention[J]. *Expert Systems with Applications*, 2025, 276: 127182.
- [84] PEI Z J, LI M D, CHEN S W. Quad-pol ISAR data reconstruction from compact-pol mode based on polarimetric and spatial feature aggregation network[J]. *IEEE Geoscience and Remote Sensing Letters*, 2025, 22: 3503705.
- [85] DENG J W, LI M D, CHEN S W. Urban damage-level estimation with reconstructed quad-pol SAR data from dual-pol SAR mode [J]. *IEEE Geoscience and Remote Sensing Letters*, 2024, 21: 4007305.
- [86] BAO W, HU J J, HUANG M Y, et al. Detecting fine-grained airplanes in SAR images with sparse attention-guided pyramid and class-balanced data augmentation[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, 15: 8586 – 8599.
- [87] YANG X Q, SHI J, CHEN T J, et al. Fast multi-shadow tracking for video-SAR using triplet attention mechanism[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 5224212.
- [88] YAN H, XU X, JIN G D, et al. Moving targets detection for video SAR surveillance using multilevel attention network based on shallow feature module[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5200518.
- [89] ZHANG S C, FU Z T, LIU J, et al. Multilevel attention Siamese network for keypoint detection in optical and SAR images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 5404617.
- [90] XIE N S, ZHANG T, GUO W W, et al. Dual branch deep network for ship classification of dual-polarized SAR images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 5207415.
- [91] FAN W W, TIAN T, WANG S Y, et al. AMDIS-Net: a multiple attention mechanism based deep dense interference suppression network for sentinel-1 SAR data [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2025, 61(6): 16869 – 16887.
- [92] IMANI M. Attention based network for fusion of polarimetric and contextual features for polarimetric synthetic aperture radar image classification [J]. *Engineering Applications of Artificial Intelligence*, 2025, 139: 109665.
- [93] HE Z C, WANG P, ZOU Y K, et al. DADIGAN: a dual attention blocks-based disentangled iterative generative adversarial network for cloud and shadow removal on SAR and optical images[J]. *Information Fusion*, 2026, 125: 103487.
- [94] GIRSHICK R. Fast R-CNN [C]//*Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015: 1440 – 1448.
- [95] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016: 779 – 788.
- [96] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements [J/OL]. (2024 – 10 – 23) [2026 – 01 – 17]. <https://arxiv.org/abs/2410.17725>.
- [97] WEI S J, ZENG X F, QU Q Z, et al. HRSID: a high-resolution SAR images dataset for ship detection and instance segmentation [J]. *IEEE Access*, 2020, 8: 120234 – 120254.
- [98] 王智睿, 康玉卓, 曾璇, 等. SAR-AIRcraft-1.0: 高分辨率 SAR 飞机检测识别数据集[J]. *雷达学报*, 2023, 12(4): 906 – 922.  
WANG Z R, KANG Y Z, ZENG X, et al. SAR-AIRcraft-1.0: high-resolution SAR aircraft detection and recognition dataset[J]. *Journal of Radars*, 2023, 12(4): 906 – 922. (in Chinese)
- [99] 王欣超, 陈思伟. 雷达遥感基础模型研究: 进展与展望[J]. *国防科技大学学报*, 2026, 48(2): 382 – 395.  
WANG X C, CHEN S W. Research on foundation models for radar remote sensing: progress and prospects[J]. *Journal of National University of Defense Technology*, 2026, 48(2): 382 – 395. (in Chinese)