

# 一种改进的可解释 SAR 图像识别网络

李鹏, 冯存前\*, 胡晓伟  
(空军工程大学防空反导学院, 西安, 710051)

**摘要** SAR-BagNet 模型是一种应用于合成孔径雷达(SAR)图像识别的可解释深度学习模型。为了使 SAR-BagNet 模型在具有可解释性的同时具有较高的识别精度,以 SAR-BagNet 模型为基础,在模型框架中加入了空间注意力和坐标注意力机制,并在 MSTAR 实测数据集上进行了实验。实验结果表明,空间注意力和坐标注意力机制增强了 SAR-BagNet 模型的全局信息获取能力,能够在不降低其可解释性的基础上,有效提高模型的识别精度和决策合理性。

**关键词** 深度学习; SAR 图像识别; 模型可解释性; 注意力机制

**DOI** 10.3969/j.issn.2097-1915.2023.04.008

**中图分类号** TN957 **文献标志码** A **文章编号** 2097-1915(2023)04-0049-07

## An Improved Interpretable SAR Image Recognition Network

LI Peng, FENG Cunqian\*, HU Xiaowei

(Air Defense and Antimissile School, Air Force Engineering University, Xi'an, 710051, China)

**Abstract** The SAR-BagNet model is an interpretable deep learning model used for Synthetic Aperture Radar (SAR) image recognition. In order to maintain the interpretability of the SAR-BagNet model while also achieving high recognition accuracy, this paper uses the SAR-BagNet model as a foundation and incorporates spatial attention and coordinate attention mechanisms into the model framework. Experimental results on the MSTAR dataset demonstrate that the spatial attention and coordinate attention mechanisms enhance the SAR-BagNet model's ability to acquire global information. This enhancement effectively improves the model's recognition accuracy and decision rationality without compromising its interpretability.

**Key words** deep learning; SAR image recognition; model interpretability; attention mechanism

合成孔径雷达(SAR)成像是一种通过雷达与成像目标之间的相对位移产生等效合成天线阵列的成像技术<sup>[1]</sup>。SAR 成像技术受外界天气影响较小,具有一定的地表穿透能力,在军事目标识别<sup>[2]</sup>、城市规划<sup>[3]</sup>、环境监测<sup>[4]</sup>、自然灾害评估<sup>[5]</sup>等方面得到了广泛应用,因此,对 SAR 图像的准确解译也越来越

重要。SAR 图像解译通常包括图像分割、目标检测和识别,其中 SAR 图像目标识别被认为是最具挑战性的任务<sup>[6]</sup>。合成孔径雷达自动目标识别(SAR-ATR)技术是近年来研究的广泛应用于 SAR 图像识别的技术<sup>[7]</sup>。传统的 SAR-ATR 方法大多是基于丰富的理论模型和专家知识设计的。这些方法具有

收稿日期: 2022-12-29

基金项目: 陕西省自然科学基金(2023-JC-YB-491)

作者简介: 李鹏(1998-),男,河南驻马店人,博士生,研究方向为 SAR 图像识别。E-mail: lipeng980404@163.com

通信作者: 冯存前(1975-),男,陕西富平人,教授,博士生导师,研究方向为目标探测与识别。E-mail: fengcunqian@sina.com

**引用格式:** 李鹏,冯存前,胡晓伟.一种改进的可解释 SAR 图像识别网络[J].空军工程大学学报,2023,24(4):49-55. LI Peng, FENG Cunqian, HU Xiaowei. An Improved Interpretable SAR Image Recognition Network[J]. Journal of Air Force Engineering University, 2023, 24(4):49-55.

较强的可解释性,但人工设计特征需要较高的专业领域知识,设计过程耗时费力且不能充分利用 SAR 图像的特征<sup>[8]</sup>。传统的 SAR-ATR 算法包括基于散射中心模型(SCM)的方法<sup>[9]</sup>、基于主成分分析(PCA)的方法<sup>[10]</sup>、稀疏表示方法<sup>[11]</sup>以及多特征融合方法<sup>[12]</sup>。

近年来,随着人工智能技术的快速发展,基于深度学习(DL)的 SAR-ATR 逐渐成为该领域的主要方法。在 SAR 图像目标识别中,CNN 也取得了许多显著的成就。文献[2]在 MSTAR 数据集上使用 CNN 进行目标识别,获得了比支持向量机(SVM)更好的精度。文献[13]提出了一种结合 CNN 和 SVM 的 SAR 目标自动识别方法。文献[14]设计了一种用于地面车辆 SAR 目标识别的大尺度批量归一化 CNN (LM-BN-CNN),与传统 CNN 结构相比具有更好的收敛速度和更高的识别精度。

上述基于 CNN 的模型虽然在 SAR 图像识别中能够实现较高的识别性能和计算效率,但这些模型通常被视为“黑盒子”,其固有的识别机制缺乏逻辑分析和数学解释<sup>[15]</sup>。在这种情况下,识别结果不像传统的 SAR 图像目标识别方法那样可靠且具有说服力,这限制了 CNN 模型在实际中的应用,特别是在军事领域和重大灾害预测等一些需要高可靠性的特殊场景下。近年来,为了更好地解释 CNN 的机制,人们提出了几种本身具有可解释性的 CNN 模型框架。自解释神经网络<sup>[16]</sup>通过正则化分别获得分类中至关重要的两个概念和每个概念的相关性,然后将它们结合起来进行预测。基于案例的可解释模型<sup>[17]</sup>学习代表类别的原型,并通过将给定图像的特征与所学习原型的相似性进行比较来做出决策。BagNet 模型<sup>[18]</sup>将 CNN 的高性能和灵活性与词袋模型(BoF)的可解释性结合起来,根据局部特征出现的次数而不是图像特征的空间关系对图像进行分类。然而,上述可解释模型都是基于光学图像,在 SAR 图像的分类中并不能取得令人满意的结果。

基于以上分析,在之前的工作中,为了使基于 CNN 的 SAR 图像分类模型具有更好的可解释性,我们根据 SAR 图像的特点,提出了用于 SAR 图像分类且具有可解释的 SAR-BagNet 模型<sup>[19]</sup>。但是由于其感受野较小,导致了全局信息的丢失,因此在识别精度上还达不到传统的卷积神经网络(如 Res-Net-18)的水平。

本文基于注意力机制对 SAR-BagNet 网络做出改进,在不降低其可解释性的基础上通过注意力机制来让模型获得全局信息。实验结果表明,增加注

意力机制的 SAR-BagNet 在识别精度和决策合理性上优于原来的 SAR-BagNet 网络。

## 1 SAR-BagNet 网络原理

SAR-BagNet 使用全局平均池化层(GAP)和全连接层作为线性分类器连接在最后一层卷积层之后来生成特定类别的热图。GAP 的作用是将最后一层卷积层的特征图的每个通道都压缩成一个单像素值  $P_k$ ,然后连接到全连接层,得到最终的分类分数  $S_c$ 。此时,单像素值  $P_k$  可以表示为:

$$P_k = \sum_i \sum_j x_{ij}^k \quad (1)$$

式中: $x_{ij}^k$  表示最后一层卷积层的第  $k$  个通道特征层在坐标  $(i, j)$  处的值。最后每一类的分类分数可以表示为:

$$S_c = \sum_k \omega_c^k P_k \quad (2)$$

式中: $\omega_c^k$  为全连接层的权值,表示从第  $k$  个通道特征层所对应的单像素值  $P_k$  在类别  $c$  的权重。然后,将全连接层的权值与最后一个卷积层的特征图相乘得到热图。类别  $c$  所对应的热图在坐标  $(i, j)$  处的像素值  $H_{ij}^c$  可以表示为:

$$H_{ij}^c = \sum_k \omega_c^k x_{ij}^k \quad (3)$$

从式(3)可以看出,热图的分辨率与最后一层卷积层的特征图的分辨率一致,想要获得一张清晰的热图,这就需要增大特征图的尺寸。而特征图的大小受到感受野和全局步长的影响。感受野和全局步长是卷积神经网络中的基本概念。CNN 的特征图中每个单元值取决于卷积网络输入的一个区域,这个区域的大小就是 CNN 模型的感受野大小,如图 1 所示。假设一个 CNN 有  $n$  个卷积层,且卷积层之间没有池化层,则最后一层卷积层的特征图中,每个单元对应的输入图像上的感受野的计算公式为<sup>[20]</sup>:

$$R_n = R_{n-1} + (f_n - 1) \prod_{i=1}^{n-1} s_i \quad (R_0 = 1) \quad (4)$$

式中: $R_n$  表示第  $n$  层感受野的大小; $f_n$  表示第  $n$  个卷积核的大小; $s_i$  代表第  $i$  个卷积层对应的步长。

全局步长表示卷积神经网络在输入图像上的等效步长,见图 1。卷积神经网络的全局步长  $S_g$  等于所有卷积层的步幅的乘积:

$$S_g = \prod_{i=1}^n s_i \quad (5)$$

从式(3)可以看出,热图的分辨率与最后一个卷积层的特征图的分辨率相同。第  $n$  个卷积层的特征图大小  $Q_n$  的计算公式为:

$$Q_n = \left\lfloor \frac{W - RF_n + 2P}{S_g} + 1 \right\rfloor \quad (6)$$

式中: $W$ 表示输入图像的大小; $P$ 表示卷积神经网络对图像的填充数;符号 $\lfloor \cdot \rfloor$ 表示向下取整。

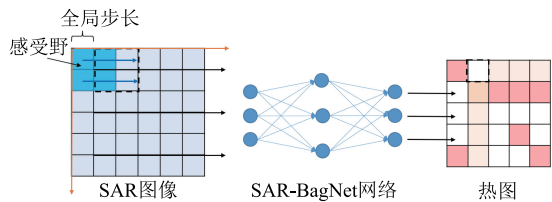


图1 SAR-BagNet原理示意图

在SAR-BagNet模型中,为了能够得到清晰的热图,设置每一层卷积层的参数 $P=0, s_i=1$ ,同时控制感受野的大小。这样设置参数不仅保证了生成的热图具有较高的分辨率,同时又能准确判断输入SAR图像中每一小块区域对决策结果的影响,如图1所示。在图1中,对于一个输入大小为 $6 \times 6$ 的SAR图像,设置感受野为 $2 \times 2$ ,全局步长为1,则可以得到大小为 $5 \times 5$ 的精细热图。图中白色、橙色和粉色方块显示不同的激活强度,按顺序增加,热图中的每一个像素值,都能在SAR图像中找到一块与之对应的区域,如图1中的虚线方框所示。SAR-BagNet在决策过程中,可以生成每个类别所对应的热图,之后将每个热图的像素值相加来表示输入图像与不同类别的匹配值(即分类分数 $S_c$ )<sup>[19]</sup>。与事后

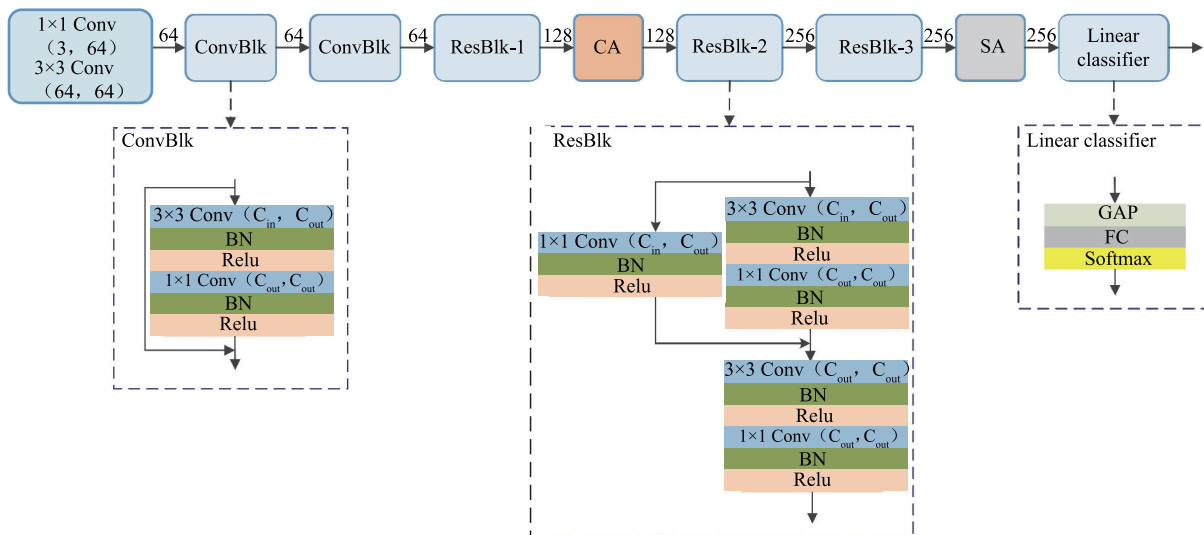


图2 本文的模型结构

在图2中,浅蓝色模块为初始的SAR-BagNet模型,CA(coordinate attention)和SA(spatial attention)分别表示本文添加的坐标注意力模块和空间注意力模块。模型的输入部分为2个卷积层,每个卷积层后分别有1个归一化层(BN)和Relu激活层。在每一个模块后面的数字为输出通道数。

在所有的卷积层中,为了控制感受野的大小,都只使用了 $1 \times 1$ 和 $3 \times 3$ 的小卷积核,且步长都设置为

解释的方法不同,SAR-BagNet模型在决策之前就已经生成了对应的热图,然后根据热图进行决策。因此,SAR-BagNet属于事前可解释模型,且可视化的热图是忠于决策结果的。

然而,虽然SAR-BagNet得到了精细的热图,增强了模型的可解释性,但是由于感受野较小,损失了输入图像的全局信息,导致了分类准确性的下降。在文献[19]中,SAR-BagNet对于MSTAR数据集10类车辆目标的识别精度为98.25%,而传统的CNN网络如ResNet-18,其识别精度为99.05%。这主要是由于ResNet-18具有较大的感受野,可以有效提取输入图像的局部和全局信息,信息获取能力强。由于在SAR-BagNet中,较大的感受野会导致其可解释性变差,因为无法准确判定输入图像中的某个区域对识别结果的影响大小。为了缓解SAR-BagNet模型的识别精度和可解释性之间的矛盾,本文引入了注意力机制,在不降低其可解释性的基础上增加模型的全局信息获取能力。

## 2 模型结构和注意力机制

### 2.1 模型的总体框架

本文改进的模型结构总体框架如图2所示。

1. 模型的线性分类器是由全局平均池化层(GAP)、全连接层(FC)和Softmax层构成。对于一个SAR图像大小为 $100 \times 100$ 的输入,由式(6)可知,最后一层卷积层输出的特征图大小为 $256 \times 82 \times 82$ 。GAP层可将 $256 \times 82 \times 82$ 的特征图降维为 $256 \times 1 \times 1$ 维的向量,然后将此向量与全连接层的 $256 \times 10$ 的权值矩阵相乘得到分类向量,其中全连接层的偏置矩阵设置为 $\mathbf{0}$ ,这样得到结果的过程只涉及乘法和加法运算。

虽然 softmax 函数为非线性函数,但是分类结果已经由 GAP 层和 FC 层决定,因此分类器是线性的。在得到每一类别对应热图的过程中,只需要将 FC 层每种类别对应的权值与每个通道的特征图相乘,之后将每个通道的特征图相加求和,即可得到每个类别对应的热图,如式(3)所示。

SAR-BagNet 模型的总体架构与 ResNet-18 相似,都采用了 18 层的卷积层以及残差连接。残差连接结构可以减缓深度学习模型训练过程中的退化问题,有利于梯度的反向传播。此外,由于 SAR 图像数据难获取,因此 SAR 图像分类数据集一般都是小数据集,采用 18 层的卷积层的小型神经网络可以缓解训练过程中的过拟合问题。

## 2.2 坐标注意力

侯齐斌等人在文献[21]中提出了坐标注意力模块,坐标注意力通过精确的位置信息对通道关系和空间位置关系进行编码,它不仅捕获跨通道信息,还捕获方向感知和位置敏感信息,这有助于模型更准确地定位和识别感兴趣的对象。具体操作可以分为坐标信息嵌入和坐标注意力生成 2 个步骤。在坐标信息嵌入的过程中,全局平均池化方法通常用于通道注意编码空间信息的全局编码。然而,由于它将全局空间信息压缩到一个单元中,导致难以保存位置信息。因此,坐标注意力分别沿高度  $H$  和宽度  $W$  方向对特征图进行一维的池化操作。具体来说,给定输入的  $C \times H \times W$  的特征图  $X$ ,首先使用尺寸为  $(H, 1)$  或  $(1, W)$  的平均池化核分别沿着水平坐标  $x$  和垂直坐标  $y$  对每个通道进行编码。因此,高度为  $h$  的第  $c$  通道的输出可以表示为:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (7)$$

同样,宽度为  $w$  的第  $c$  通道的输出可以写成:

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (8)$$

式(7)和式(8)分别沿 2 个空间方向聚合特征,得到 1 对方向感知的特征图,这 2 种变换可以使坐标注意力模块捕捉到沿着一个空间方向的长期依赖关系,并保存沿着另一个空间方向的精确位置信息,这有助于模型捕捉到感兴趣的区域。

在坐标注意力生成的过程中,首先对式(7)和式(8)的结果进行连接(Concat)操作,得到  $C \times 1 \times (H + W)$  维的特征图,然后经过卷积核大小为  $1 \times 1$ ,通道数为  $C/r$  的卷积层压缩通道,之后进行归一化操作(BN)和非线性激活(Non-Linner)。连接 2 个特征图的目的是为了进行统一的归一化操作, $r$  为缩

减因子。可以表示为:

$$f = \delta(F_1([\mathbf{z}^h, \mathbf{z}^w])) \quad (9)$$

式中:  $[\ ]$  表示连接操作;  $\delta$  表示归一化和非线性激活函数;  $F_1$  为  $1 \times 1$  的卷积层;  $f \in R^{C/r \times (H+W)}$ 。

将  $f$  分为 2 个独立的特征  $f^h \in R^{C/r \times H}$  和  $f^w \in R^{C/r \times W}$ 。使用另外 2 个卷积层和 Sigmoid 函数进行特征转化,使其维度与输入一致:

$$g^h = \sigma(F_h(f^h)), g^w = \sigma(F_w(f^w)) \quad (10)$$

式中:  $F_h$  和  $F_w$  表示 2 个不同的  $1 \times 1$  卷积层,其中卷积层的输出通道数为  $C$ ,  $\sigma$  表示 Sigmoid 激活层。最后,坐标注意力模块的输出可以写成:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (11)$$

这样通过坐标注意力模块对特征图的处理,特征图的每个权重都包含了通道间信息、横向空间信息和纵向空间信息,能够帮助 SAR-BagNet 网络更准确地定位目标信息,增强模型的全局信息提取能力。坐标注意力模块的详细结构如图 3 所示。

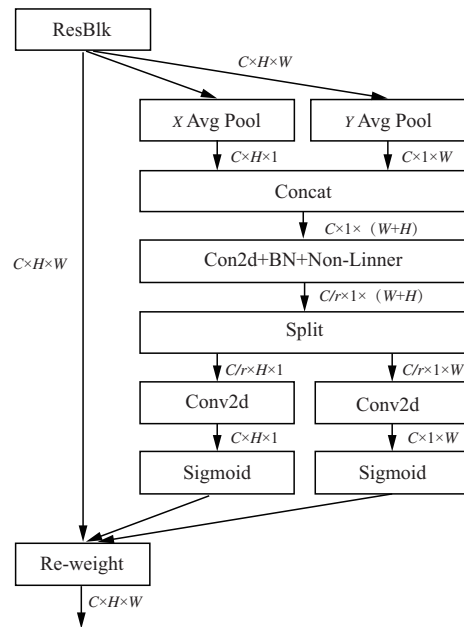


图 3 坐标注意力模块

## 2.3 空间注意力

在卷积神经网络中,输入图像的各个区域对识别结果的贡献并不一样,与任务相关的区域才需要关心。空间注意力模块就是寻找输入图像中重要的部位进行处理。空间注意力实现起来较为简单,首先对  $C \times H \times W$  维的特征图在通道维进行平均池化和最大池化,并将两部分的特征图进行拼接,得到  $2 \times H \times W$  的特征图,然后将特征图经过一个卷积层和 Sigmoid 激活层得到  $H \times W$  维的空间注意力。将初始的特征图与空间注意力相乘,即可让网络关注重要特征并抑制不必要的特征。空间注意力的生成过程可以表示为<sup>[22]</sup>:

$$S = \sigma(F_7(\text{Avgpool}(X); \text{Maxpool}(X))) \quad (12)$$

式中: Avgpool 和 Maxpool 分别为全局平均池化和最大池化; S 为生成的空间注意力;  $F_7$  为  $7 \times 7$  的卷积层, 其输出通道数为 1, 具体流程如图 4 所示。

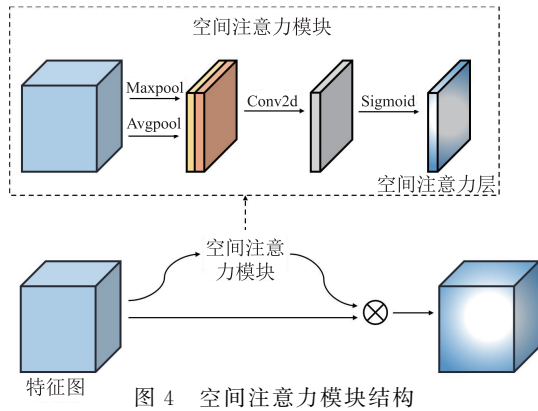


图 4 空间注意力模块结构

### 3 实验

#### 3.1 实验数据集与参数设置

本文采用的数据集为 SAR 图像目标识别领域常用的 MSTAR 数据集, 由美国国防高级研究计划署(DARPA)所公布的实测 SAR 地面静止军事车辆目标数据。采集该数据集的传感器为高分辨率的聚束式 SAR, 该雷达的分辨率为  $0.3 \text{ m} \times 0.3 \text{ m}$ , 工作波段为 X 波段, 采用 HH 极化方式。

由于 SAR 图像为灰度图, 本文将 SAR 图像转换为伪 RGB 图像(将单通道灰度图复制为三通道图像), 并在预处理阶段对图像进行了标准化。为了增强图像的多样性, 防止过拟合, 还加入了随机亮度调节处理。训练集和测试集的划分如表 1 所示。

在训练过程中, 网络优化器为 Adam, 初始学习率  $L_R = 1 \times 10^{-3}$ , 迭代次数 Epoch = 200, 批处理大小 Batchsize = 64。本文所有的实验操作均在基于 Python 版本为 3.7 的 Pytorch 平台上实施。

表 1 训练集和测试集的 SAR 图像划分

类别	训练集		测试集	
	角度/(°)	数量	角度/(°)	数量
2S1	17	299	15	274
BMP2	17	233	15	195
BRDM2	17	298	15	274
BTR60	17	256	15	195
BTR70	17	233	15	196
D7	17	299	15	274
T62	17	299	15	273
T72	17	232	15	196
ZIL131	17	299	15	274
ZSU23/4	17	299	15	274

#### 3.2 识别精度对比

本文对比添加的模块和初始的 SAR-BagNet 在测试集上的识别精度, 在训练过程中的对于 SAR 图像的预处理都保持一致, 对比结果如表 2 所示。

实验中, 加入了初始的 SAR-BagNet 模型对 MSTAR 测试集的识别精度只有 98.25%, 这低于传统的 ResNet-18 的识别精度。在加入坐标注意力 CA 之后, 识别精度达到了 98.80%, 主要是因为坐标注意力能够使模型捕获通道和位置信息, 增强模型的信息提取能力。而在加入空间注意力之后, SAR-BagNet 模型的识别精度也有所提升, 达到了 98.64%, 说明空间注意力模块能够使得模型更关注 SAR 图像中对分类结果有重要影响的区域, 而忽略与目标分类无关的区域, 这样有利于模型对分类信息的正确提取, 减少无关信息和错误信息对模型的影响。

而当 SAR-BagNet 模型同时加入坐标注意力和空间注意力之后, 模型的识别精度有了进一步的提高, 达到了 99.18%, 超过了传统的卷积神经网络 ResNet-18 的识别精度, 说明这 2 个模块可以使模型有效提取到全局信息, 并利用其关注输入 SAR 图像中对识别区域有重要影响的区域。

与 ResNet-18 网络精度对比的原因主要是 ResNet-18 网络结构与 SAR-BagNet 类似, 都有 18 层的卷积层且采用残差结构。然而, ResNet-18 虽然具有较高的识别精度, 但是缺乏可解释性。

表 2 不同模型的识别精度

模型	识别精度/%
ResNet-18	99.05
SAR-BagNet	98.25
SAR-BagNet + CA	98.80
SAR-BagNet + SA	98.64
SAR-BagNet + CA + SA	99.18

#### 3.3 可解释性对比分析

为了说明注意力模块有助于模型关注到 SAR 图像中的目标区域, 增加模型的全局信息获取能力, 本节对比了本文改进的模型与初始 SAR-BagNet 模型生成的热图, 见图 5。此外, 还加入了 ResNet-18 网络运用类激活映射(CAM)的方法生成的热图, 因为 CAM 的方法与文中方法类似, 都是使用特征图与类别的权值相乘得到热图。但是 CAM 生成的热图无法准确反映 SAR 图像中的重要区域。

在图 5 中, SAR-BagNet 模型分别添加了不同的模块生成了不同的热图, 图的下方表示增加的模

块。热图中红色区域表示对识别结果有积极影响的区域,而蓝色区域表示对识别结果有消极影响的区域,颜色越深,表示影响越大。从图 5 中可以看出,CAM 生成的热图仅仅能够反映 SAR 图像中较为重要的大致区域,无法精确区分 SAR 图像中的具体细节。在 SAR-BagNet 生成的热图中,可以精准判断出 SAR 目标的哪些部位对识别结果有重要影响。在第 1 幅图对应的热图中,可以看到 SAR 中车辆目标的边缘区域显示为深红色,表明目标的边缘特征作为了模型的分类特征。在第 3 幅图中,SAR 图像

中车辆目标中的阴影边缘为深红色,表明模型把阴影边缘作为识别特征对图像进行分类。

在增加了 SA 和 CA 后,可以明显看到:在第 1 幅图中,相比于初始的模型,热图中的红色区域几乎覆盖了整个车辆目标,表明模型将整体的车辆信息作为分类特征而不仅仅是车辆的边缘信息。在第 3 幅图中,红色区域为整体的阴影区域而不仅仅是阴影边缘。这表明在增加了 CA 和 SA 模块之后,模型利用了更多的全局信息而不仅仅是局部信息进行分类,这也解释了模型的识别精度提高的原因。

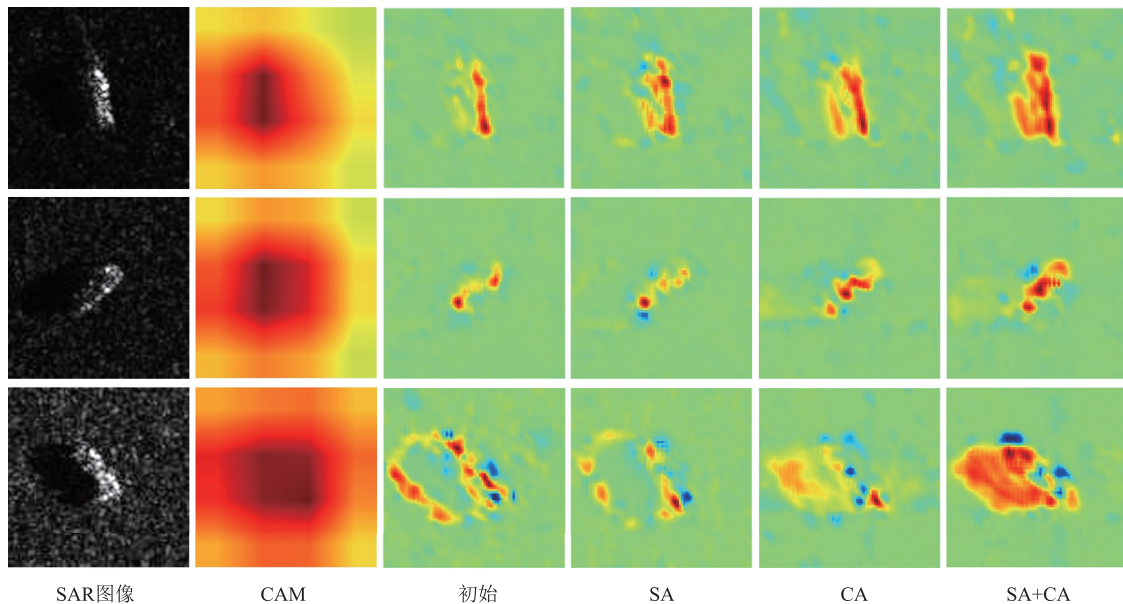


图 5 不同模型生成的热图比较

## 4 结语

本文在 SAR-BagNet 这一事前可解释性模型的框架上加入了坐标注意力和空间注意力模块,提高了模型的识别精度以及全局信息的获取能力,但是由于两种模块的加入,模型的参数数量会增多。消融实验结果表明,同时利用坐标注意力和空间注意力这两种模块对于模型识别精度和决策合理性的提升比仅利用单一模块的效果更好。在计算机视觉领域,虽然现在很多的深度学习模型利用注意力机制去提高模型的识别和检测精度,但是在事前可解释性模型以及可解释性领域的应用还比较少。

可解释性作为深度学习领域一项重要的研究课题,目前仍然处于探索阶段,需要进一步开展相关研究。本文只是在 SAR-BagNet 这一具有事前可解释性结构的模型中讨论了注意力机制对于模型的影响,因此得出的结论还比较局限。在下一步的工作中,将扩展研究对象,得出更加具有代表性的结论。

## 参考文献

- [1] 庄钊文,王雪松,黎湘,等. 雷达目标识别[M]. 北京: 高等教育出版社,2015.
- [2] WANG Y P, ZHANG Y B, Qu H Q, et al. Target Detection and Recognition Based on Convolutional Neural Network for SAR Image[C]//2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). [S.l.]:IEEE, 2018: 1-5.
- [3] QUAN S, XIONG B, XIANG D, et al. Eigenvalue-based Urban Area Extraction Using Polarimetric SAR data[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(2): 458-471.
- [4] CAI J, JIA H, LIU G, et al. An Accurate Geocoding Method for GB-SAR Images Based on Solution Space Search and Its Application in Landslide Monitoring [J]. Remote Sensing, 2021, 13(5): 832.
- [5] CAO H, ZHANG H, WANG C, et al. Operational Flood Detection Using Sentinel-1 SAR Data over

- Large Areas[J]. *Water*, 2019, 11(4): 786.
- [6] NOVAK L M, OWIRKA G J, NETISHEN C M. Performance of a High-Resolution Polarimetric SAR Automatic Target Recognition System [J]. *Lincoln Laboratory Journal*, 1993, 6(1): 11-24.
- [7] BLACKNELL D, GRIFFITHS H. Radar Automatic Target Recognition (ATR) and Non-Cooperative Target Recognition (NCTR) [Z/OL]. 2013. DOI: 10. 1049/ PBRA033E.
- [8] FENG Z, ZHU M, STANKOVIC L, et al. Self-matching CAM: A Novel Accurate Visual Explanation of Cnns for SAR Image Interpretation [J]. *Remote Sensing*, 2021, 13(9): 1772.
- [9] 丁柏圆, 文贡坚, 余连生, 等. 属性散射中心匹配及其在SAR目标识别中的应用 [J]. *雷达学报*, 2017, 6(2): 157-166.
- [10] 李翔, 张静, 吴凌华, 等. 一种基于改进核主成分分析的SAR图像识别方法研究 [J]. *海军航空工程学院学报*, 2009, 24(3): 307-312.
- [11] 徐波. 基于稀疏表示的SAR图像目标识别算法及仿真 [J]. *自动化与仪器仪表*, 2018(6): 176-179.
- [12] 徐英, 谷雨, 彭冬亮. 基于稀疏表示和多特征融合的SAR图像识别 [J]. *指挥信息系统与技术*, 2020, 11(3): 29-35.
- [13] WU T D, YEN Y, WANG J H, et al. Automatic Target Recognition in SAR Images Based on a Combination of CNN and SVM [C]//2020 International Workshop on Electromagnetics: Applications and Student Innovation Competition (IWEM). [S. l.]: IEEE, 2020: 1-2.
- [14] ZHOU F, WANG L, BAI X, et al. SAR ATR of Ground Vehicles Based on LM-BN-CNN [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(12): 7282-7293.
- [15] LIPTON Z C. The Mythos of Model Interpretability: In Machine Learning, the Concept of Interpretability is Both Important and Slippery [J]. *Queue*, 2018, 16(3): 31-57.
- [16] ALVAREZ M D, JAAKKOLA T. Towards Robust Interpretability with Self-Explaining Neural Networks [Z]. *Advances in Neural Information Processing Systems*, 2018. DOI: 10. 48550/ arxiv. 186. 07538.
- [17] CHEN C, LI O, TAO D, et al. This Looks Like That: Deep Learning for Interpretable Image Recognition [Z]. 2018. DOI: 10. 48550/ arxiv. 186. 10574.
- [18] BRENDEL W, BETHGE M. Approximating Cnns with Bag-of-Local-Features Models Works Surprisingly Well on Imagenet [Z]. *ArXiv Preprint arXiv:1904. 00760*, 2019.
- [19] LI P, FENG C, HU X, et al. SAR-BagNet: An Ant-hoc Interpretable Recognition Model Based on Deep Network for SAR Image [J]. *Remote Sensing*, 2022, 14(9): 2150.
- [20] DUMOULIN V, VISIN F. A Guide to Convolution Arithmetic for Deep Learning [Z]. *ArXiv Preprint ArXiv:1603. 07285*, 2016.
- [21] HOU Q, ZHOU D, FENG J. Coordinate Attention for Efficient Mobile Network Design [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2021: 13713-13722. DOI: 10. 48550/ arxiv. 2013. 02907.
- [22] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional Block Attention Module [C]//Proceedings of the European Conference on Computer Vision. [S. l.]: ECCV, 2018: 3-19.

(编辑:徐敏)