

一种基于贝叶斯网络的微波雷达和图像融合与分类算法

李庆元, 杜磊, 姜渊伟

电科云(北京)科技有限公司, 北京

收稿日期: 2023年12月9日; 录用日期: 2024年1月15日; 发布日期: 2024年1月22日

摘要

近年来, 深度学习技术的进步在合成孔径雷达(SAR)自动目标识别(ATR)技术中取得了出色的表现。然而, 由于斑点噪声的干扰, SAR图像的分类任务仍然具有挑战性。为了解决这个问题, 本研究提出了一种多尺度局部-全局特征融合网络(MFN), 该网络集成了卷积神经网络(CNN)和Transformer网络。所提出的网络包括三个分支: CovNeXt-SimAM分支, Swin Transformer分支和多尺度特征融合分支。CovNeXt-SimAM分支在不同的尺度上提取SAR图像的局部纹理细节特征。通过将SimAM注意机制结合到CNN块中, 从空间和通道注意角度增强了模型的特征提取能力。此外, Swin Transformer分支用于提取不同尺度下的SAR图像全局语义信息。最后, 多尺度特征融合分支用于融合局部特征和全局语义信息。此外, 为了解决由于经验确定的模型超参数问题而导致模型精度和效率较低的问题, 采用贝叶斯超参数优化算法确定了最佳的模型超参数。该研究提出的模型在MSTAR数据集上, 标准工作条件(SOCs)和扩展工作条件(EOCs)下, 对SAR车辆目标分别取得了99.26%和94.27%的平均识别准确率。与基准模型相比, 识别准确率分别提高了12.74%和25.26%。结果表明, 贝叶斯-MFN降低了SAR图像之间的类间距离, 导致更紧凑的分类特征和更少的斑点噪声干扰。与其他主流模型相比, 贝叶斯-MFN模型展现出最佳的分类性能。

关键词

合成孔径雷达(SAR), 斑点噪声, 贝叶斯超参数优化算法

A Microwave Radar and Image Fusion and Classification Algorithm Based on Bayesian Networks

Qingyuan Li, Lei Du, Yuanwei Lou

Dianke Cloud (Beijing) Technology Company, Beijing

Received: Dec. 9th, 2023; accepted: Jan. 15th, 2024; published: Jan. 22nd, 2024

文章引用: 李庆元, 杜磊, 姜渊伟. 一种基于贝叶斯网络的微波雷达和图像融合与分类算法[J]. 图像与信号处理, 2024, 13(1): 47-58. DOI: 10.12677/jisp.2024.131005

Abstract

In recent years, the advancement of deep learning technology has led to excellent performance in synthetic aperture radar (SAR) automatic target recognition (ATR) technology. However, due to the interference of speckle noise, the task of classifying SAR images remains challenging. To address this issue, a multi-scale local-global feature fusion network (MFN) integrating a convolution neural network (CNN) and a transformer network was proposed in this study. The proposed network comprises three branches: a CovNeXt-SimAM branch, a Swin Transformer branch, and a multi-scale feature fusion branch. The CovNeXt-SimAM branch extracts local texture detail features of the SAR images at different scales. By incorporating the SimAM attention mechanism to the CNN block, the feature extraction capability of the model was enhanced from the perspective of spatial and channel attention. Additionally, the Swin Transformer branch was employed to extract SAR image global semantic information at different scales. Finally, the multi-scale feature fusion branch was used to fuse local features and global semantic information. Moreover, to overcome the problem of poor accuracy and inefficiency of the model due to empirically determined model hyperparameters, the Bayesian hyperparameter optimization algorithm was used to determine the optimal model hyperparameters. The model proposed in this study achieved average recognition accuracies of 99.26% and 94.27% for SAR vehicle targets under standard operating conditions (SOCs) and extended operating conditions (EOCs), respectively, on the MSTAR dataset. Compared with the baseline model, the recognition accuracy has been improved by 12.74% and 25.26%, respectively. The results demonstrated that Bayes-MFN reduces the inter-class distance of the SAR images, resulting in more compact classification features and less interference from speckle noise. Compared with other mainstream models, the Bayes-MFN model exhibited the best classification performance.

Keywords

Synthetic Aperture Radar (SAR), Speckle Noise, Bayesian Hyperparameter Optimization Algorithm

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

作为一种主动式微波遥感探测平台,合成孔径雷达(SAR)具有全天候、全天监测和一定的穿透能力的优点,因此在军事和民用领域被广泛使用[1]。与可以直接被人眼解读的光学图像不同, SAR 图像包含丰富的目标信息。然而,由于 SAR 的独特成像机制, SAR 图像解读的任务变得具有挑战性[2]。此外,获取和标记 SAR 图像很困难,需要大量人力和物力资源。因此, SAR 自动目标识别(SAR ATR)近年来成为研究的焦点。由于斑点噪声的存在, SAR 图像具有特定的细粒度特征,表现为大的和小的类间差异。因此, SAR 图像分类的任务仍然具有挑战性。

随着深度学习理论的发展,深度神经网络已广泛应用于各个领域[3]。在 SAR ATR 领域,与基于模板匹配[4]和模型匹配[5]的方法相比,基于深度神经网络的方法具有自动特征提取的优势。其中,最具代表性的方法是基于卷积神经网络(CNN)。CNN 以层次方式收集局部特征,获得强大的图像表示,并可以充分提取 SAR 图像的局部纹理细节。2022 年, Liu 等人提出了一种纯卷积神经网络,名为 ConvNeXt,它已成为最具代表性的 CNN,相比传统的 CNN, ConvNeXt 算法更简洁,参数更少[6]。然而,受到卷积

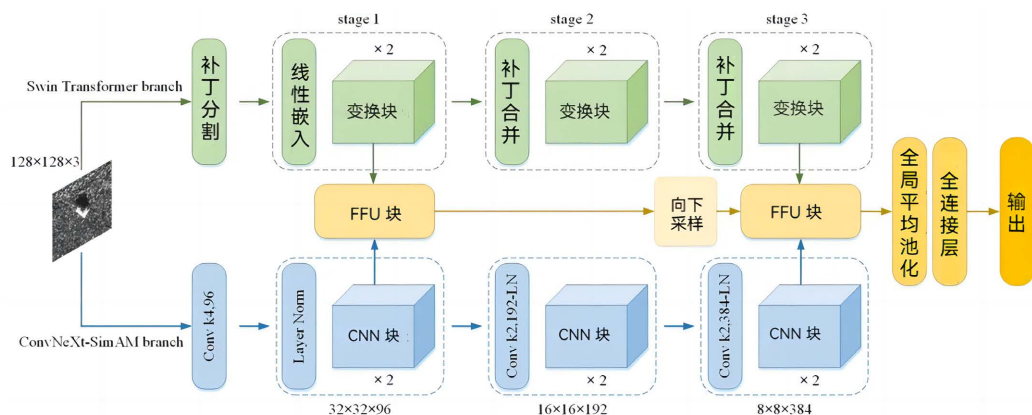
核大小的限制, CNN 捕捉图像全局特征的能力较弱。因此, 难以获得 SAR 图像元素之间的长距离关系, 导致忽略了细粒度图像部分之间的语义信息关系[7]。

随着 transformers [8] 的出现, 卷积神经网络在计算机视觉领域的垄断地位正在逐渐被打破。与 CNN 的架构不同, transformer 架构使用自注意机制来进行全局关系建模。2020 年, Dosovitskiy [9] 等人提出了用于图像分类的 Vision Transformer (ViT) 模型。ViT 模型是首个应用于计算机视觉领域的纯 Transformer 模型, 通过将图像分割成多个块并将其编码为序列向量, 解决了 Transformer 在图像领域的输入问题。这种方法在图像分类领域取得了出色的性能。然而, ViT 模型具有大量的参数, 计算成本高, 对输入图像的大小有一定限制。为解决 ViT 模型的缺点, Liu 等人提出了一种具有滑动窗口操作和分层设计的 Swin Transformer 模型, 具有更高的效率[10]。然而, 由于 SAR 图像子类别之间的细微差异, Swin Transformer 模型提取 SAR 图像的局部细节特征能力较弱。模型对全局语义信息的依赖限制了其表示能力, 而可区分的细节特征对于 SAR 图像分类任务至关重要。

SAR 图像具有明显的细粒度特征, 因此有必要进行全局建模以获取 SAR 图像的长距离依赖关系。然而, SAR 图像具有低分辨率和模糊的边缘纹理细节, 当仅对其进行全局建模时, 很难将 SAR 目标与背景区分开来。因此, 有必要获取它们的局部特征以增强其捕捉微妙差异的能力。为了充分利用 SAR 图像的细节多样性, 抑制斑点噪声, 改善 SAR 图像之间的小类差异问题, 本文提出了一种名为 MFN 的三支多尺度特征融合网络用于 SAR 图像分类。MFN 模型结合了 CNN 和 Swin Transformer 架构的优点。受 ConvNeXt 和 Swin Transformer 模型的启发, MFN 包括三个分支: ConvNeXt-SimAM、Swin Transformer 和多尺度特征融合。多尺度特征融合分支可以从不同层次全面挖掘浅层细节特征和高级语义信息, 并通过特征融合单元(FFU)融合 SAR 图像的局部特征和全局特征, 增强 SAR 目标分类模型的代表能力。最后, 使用贝叶斯超参数优化算法来优化模型的超参数, 提高模型的性能。

2. 方法

2.1. ConvNeXt-SimAM 分支



注: 绿色部分代表 Swin Transformer 分支, 蓝色部分代表 CovNeXt-SimAM 分支。蓝色部分代表 CovNeXt-SimAM 分支。黄色部分代表多尺度特征融合分支。

Figure 1. Structure of the multi-frequency network model

图 1. 多频网络模型结构图

为了有效地获取不同尺度下合成孔径雷达车辆目标的局部细节特征和全局语义表示, 我们结合了 ConvNeXt 和 Swin Transformer 模型的优点, 设计了一个名为 MFN 的分类模型, 该模型集成了不同尺度

上的局部 - 全局特征。MFN 模型的结构图如图 1 所示。这里提出的 MFN 模型包含三个分支：ConvNeXt-SimAM、Swin Transformer 和多尺度特征融合。ConvNeXt-SimAM 和 Swin Transformer 分支可以分为三个阶段，在不同层次上构建特征图。这三个阶段对应于 4 倍、8 倍和 16 倍下采样。此外，特征图的高度和宽度减小，通道的数量增加，在同一阶段的特征维度保持不变。最后，不同尺度上的局部和全局特征由多尺度特征融合分支进行融合和分类。

为了充分提取合成孔径雷达车辆目标的局部细节特征，我们通过堆叠 CNN 块和下采样操作构建了 ConvNeXt-SimAM 分支，并将整个分支划分为三个阶段，如图 1 所示。为了与 Swin Transformer 分支中每个阶段的特征图大小保持一致，ConvNeXt-SimAM 分支中每个阶段的特征图的高度和宽度通过下采样卷积操作减小到前一阶段的一半，同时通道数增加一倍。此外，分别从第 1 阶段和第 3 阶段提取浅层局部细粒度信息和深层粗粒度信息，并将其发送到多尺度特征融合分支，该分支可以融合不同尺度的局部特征，增强模型的特征表示能力。CNN 块的架构如图 2(a)所示。受 ConvNeXt 的启发，CNN 块由一系列组卷积、 7×7 卷积核大小、层归一化(LN)、 1×1 卷积、GELU 激活函数、SimAM 注意力模块和 1×1 卷积组成。为了引入额外的可学习参数，我们在 CNN 块的第二个卷积层(conv k1, $4 \times \text{dim}$)后插入了 SimAM 注意力模块，增强了 CNN 块的局部特征提取能力，以充分提取合成孔径雷达车辆的细节特征并抑制斑点噪声。与现有的通道注意力或空间注意力机制不同，SimAM 模块可以从通道和空间维度的特征图中推导出三维注意力权重，而不添加额外的参数，这灵活而有效地增强了卷积模块的表示能力。Yang 等人[11]基于神经学的发现为特征图上的每个神经元定义了能量函数，而输入特征是 $X \in \mathbb{R}^C \times H \times W$ ，因此每个通道具有 $H \times W$ 个能量函数。通过最小化能量函数来评估每个神经元的重要性，并赋予对影响更大的神经元更高的权重，从而增强提取局部特征的能力并有效抑制噪声干扰。因此，最小化能量函数后的公式如式(1)所示：

$$e_i^* = \frac{4(\sigma^- + \lambda)}{(t-u)^2 + 2\sigma^2 + 2\lambda}$$

其中， t 代表 X 中的单通道的目标神经元，即像素值； μ 和 σ 分别代表通道特征图上所有神经元的像素值的均值和方差； λ 代表正则化系数。神经元 t 的能量越低，其重要性越高。

经过 SimAM 注意机制处理后，输入特征 X 的最终输出结果如下：

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{E}\right) \otimes X$$

在 ConvNeXt 块中，首先进行组卷积 $f_{d7 \times 7}$ ，卷积核大小为 7×7 ，通道数为 dim ，其中组数等于通道数。使用组卷积可以减少模型的计算量。随后使用层归一化，然后使用卷积操作 $f_{1 \times 1}$ ，卷积核大小为 1×1 ，通道数为 $4 \times \text{dim}$ 来升级维度。最后，使用 GELU 激活函数和卷积操作 $f_{1 \times 1}$ ，卷积核大小为 1×1 ，通道数为 dim 进行降维。CNN 块的过程如下：

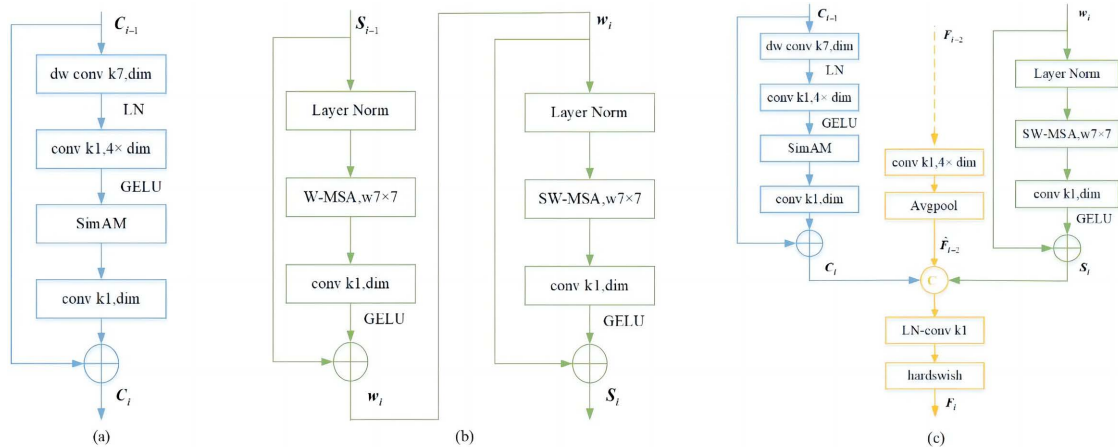
$$C_i = f^{1 \times 1} \left(\text{SimAM} \left(\text{GELU} \left(f^{1 \times 1} \left(\text{LN} \left(f^{d7 \times 7} (C_{i-1}) \right) \right) \right) \right) \right) + C_{i-1}$$

其中 C_i 表示 CNN 块的输出特征，LN 是一个层归一化层。

2.2. Swin Transformer 分支

Swin Transformer 分支的工作流程如图 1 所示。通过三个阶段，Swin Transformer 分支构建了不同大小的特征图。在第一阶段，特征图首先通过补丁划分和线性嵌入层进行分割，并在通道方向上进行平铺。然后通过补丁合并层进行降采样，并增加通道维度以构建多阶层次结构。每个阶段都堆叠了 N 个 Trans

块, 其结构如图 2(b)所示。由于 Trans 块中包含 W-MSA 和 SW-MSA 两个结构, 通常都是成对使用的, 所以 N 取偶数。为了避免模型参数数量的大幅增加, N 的值取为 2。



注: (a) 卷积神经网络模块架构; (b) 两个连续的转换模块架构; (c) 第 3 阶段中的特征融合过程

Figure 2. Diagrams of different processes

图 2. 不同过程的示意图

引入 Swin Transformer 分支可以增强 ConvNeXt-SimAM 分支捕捉全局语义信息的能力。如图 1 所示, 输入图像经过第一阶段后, 特征图的大小变为 $32 \times 32 \times 96$ 。随后, 通过补丁合并操作对特征图进行降采样。特征图的高度和宽度减半, 通道数加倍。Trans 块的结构如图 2(b)所示。首先, 特征图 S_{i-1} 通过层归一化层进入窗口多头自注意力(W-MSA)模块, 在每个窗口中计算自注意力, 大大降低计算复杂度。然后, 通过卷积层(卷积核大小为 1×1)和 GELU 激活函数进行处理。最后, 在每个模块后应用残差连接。接下来的模块中引入了移位窗口多头自注意力(SW-MSA)模块, 其它层保持不变。SW-MSA 的引入可以实现不同窗口之间的交互作用。Trans 块的计算过程如下所示:

$$w_i = GELU\left(f^{1 \times 1}\left(W-MSA\left(LN\left(S_{i-1}\right)\right)\right)\right) + S_{i-1}$$

$$S_i = GELU\left(f^{1 \times 1}\left(SW-MSA\left(LN\left(w_i\right)\right)\right)\right) + w_i$$

其中 w_i 是 W-MSA 层的输出特征, S_i 是 Trans 块的输出特征。

多尺度特征融合分支为提高 SAR 图像分类模型的识别精度, 有必要融合不同尺度上的局部特征和全局特征。同时, 为尽可能保留局部特征和全局特征, 实现了一个多尺度特征融合分支。多尺度特征融合分支能够根据输入特征自适应地融合局部特征、不同尺度下的全局表示以及前一级融合的语义信息。多尺度特征融合分支的工作流程如图 1 所示。我们通过多尺度特征融合分支捕捉了第 1 阶段的浅层局部和全局特征以及第 3 阶段的深层局部和全局语义信息。最后, 融合路径的特征信息逐渐填补语义差距, 并通过连接全局平均池化层和全连接层得到最终输出。

如图 3 所示, 多尺度特征融合分支由 FFU 和下采样操作组成。其中, FFU 包括 LN 层、卷积核尺寸为 1×1 的卷积层以及硬切线激活函数。第 3 阶段的具体过程如图 2(c)所示。首先, 通过卷积核尺寸为 1×1 和平均池化层(Avgpool), 将前一阶段的 FFU 输出特征 \widehat{F}_{i-2} 向下采样, 然后调整通道维度以便后续的融合操作。通过局部特征 C_i 和来自 ConvNeXt-SimAM 和 Swin Transformer 分支的全局特征 S_i , 经过 LN 层进行归一化后, 利用卷积核尺寸为 1×1 和硬切线激活函数进行融合。最后, 本阶段融合的特征 F_i 有效捕

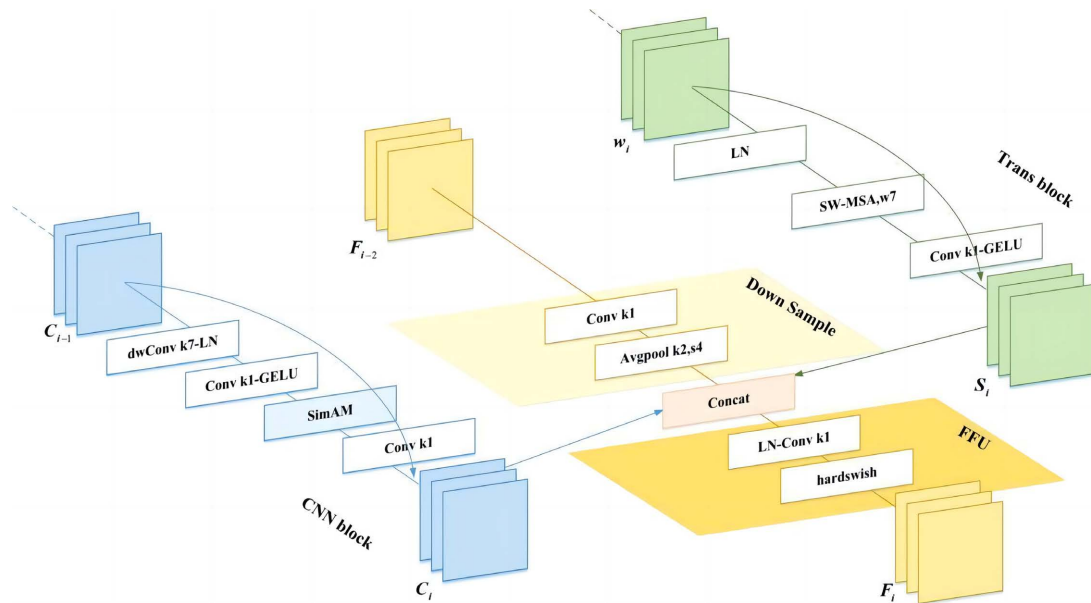
捉了 SAR 图像的局部特征和全局语义信息。得到的特征 F_i 随后通过全局平均池化和全连接层进行分类:

$$\widehat{F}_{i-2} = \text{Avgpool}(f^{1 \times 1}(F_{i-2}))$$

$$F_i = \text{hardswish}\left(f^{1 \times 1}\left(\text{LN}\left(\text{Concat}\left(\widehat{F}_{i-2}, S_i, C_i\right)\right)\right)\right)$$

$$O = \text{FC}(\text{glo_avg_pool}(F_i))$$

其中, O 表示模型的最终输出, FC 表示全连接层, glo_avg_pool 表示全局平均池化。



注: 绿色部分代表 Trans 块, 蓝色部分代表 CNN 块结构, 黄色部分是通过下采样操作从前一个 FFU 的输出特征到下一个 FFU 的过程。

Figure 3. Implementation details of the CNN block, Trans block, and FFU block in stage 3
图 3. 第 3 阶段的 CNN 块、Trans 块和 FFU 块的实现细节

基于贝叶斯超参数优化算法的多重因素网络(Bayes-MFN)由于神经网络模型的复杂性,手动参数调整可能耗时费力,而且由于该模型常被视为黑盒方法,也可能无法找到最佳超参数组合。要建立最优的神经网络模型,优化超参数尤为重要。与传统的网格搜索和随机搜索方法相比,贝叶斯优化算法的算法框架是连续的。也就是说,贝叶斯优化方法通过使用先验知识来近似未知函数的后验分布,从而调整最佳超参数组合,减少大量不必要的计算。由于贝叶斯优化方法能够在有限的函数采样值下获得复杂目标函数的最优值,因此在许多领域中备受青睐。超参数可以分为模型训练相关的参数和模型结构相关的参数两组。由于所提出模型的设计旨在以更少的可学习参数实现更好的分类性能,模型结构相关的超参数是固定的。与模型训练相关的超参数的适当选择可以使神经网络学习更快且性能更好。在训练过程中,批量大小影响模型的泛化性能,学习率(lr)决定模型的收敛状态[12]。此外,热身和衰减(wd)超参数与深度神经网络模型的稳定性密切相关。适当的 epoch 大小可以帮助避免模型的过拟合或欠拟合。因此,有必要针对这四个超参数执行贝叶斯优化,将损失函数作为目标函数。本研究采用了 Python 中 HyperOpt 包提供的贝叶斯优化算法,以更少的步骤实现超参数优化[13]。在本文中,选择了树状帕尔泽估计器(TPE)搜索算法[14]进行高效的超参数优化。优化超参数的过程如下:(1) 定义一个最小化的目标函数,(2) 定

义待优化的超参数和搜索空间, 以及(3) 选择搜索算法[15]。采用 TPE 算法通过最小化损失函数来找到最佳超参数组合。

3. 实验和结果

3.1. MSTAR 数据集

为了评估提出的模型的性能, 我们使用了移动和静止目标获取与识别(MSTAR)数据集[16], 该数据集由美国国防高级研究计划局和空军实验室于上世纪 90 年代中期发布。MSTAR 数据集中的合成孔径雷达(SAR)军用车辆是由分辨率为 0.3 米的高分辨率聚焦 SAR 采集的, 工作在 X 频段和 HH 偏振模式下。采集条件可分为标准操作条件(SOCs)和扩展操作条件(EOCs) [17]。SOC 指的是训练集和测试集样本的低次角度和外观略有不同, 但样本的序列号和目标配置相同的情况。相反, EOC 指的是训练集和测试集在俯仰角度、目标定义和其他因素上有显著差异的情况, 使之成为比 SOC 更复杂的分类任务。MSTAR 数据集包含了各种尺寸的 SAR 图像, 包括 128×128 、 158×158 和 177×178 像素。为了保持目标尺寸的一致性, 原始图像在不影响目标的情况下均匀剪裁为 128×128 像素。在 SOC 分类中, 共有 10 种车辆目标, 包括 BMP2 (步兵战斗车)、BRDM2 (装甲侦察车)、BTR60 (装甲运输车)、BTR70 (装甲运输车)、D7 (推土机)、T62 (坦克)、T72 (坦克)、ZIL131 (货车)、ZSU234 (自行式防空炮)和 2S1 (自行式榴弹炮)。训练和测试样本分别是在 17° 和 15° 俯仰角下拍摄的图像。图 4 显示了 MSTAR 数据集中不同车辆目标的光学图像和 SAR 图像。从图中可以明显看出, SAR 图像具有低分辨率和模糊的目标边缘信息, 使得每个车辆目标在视觉上难以区分。

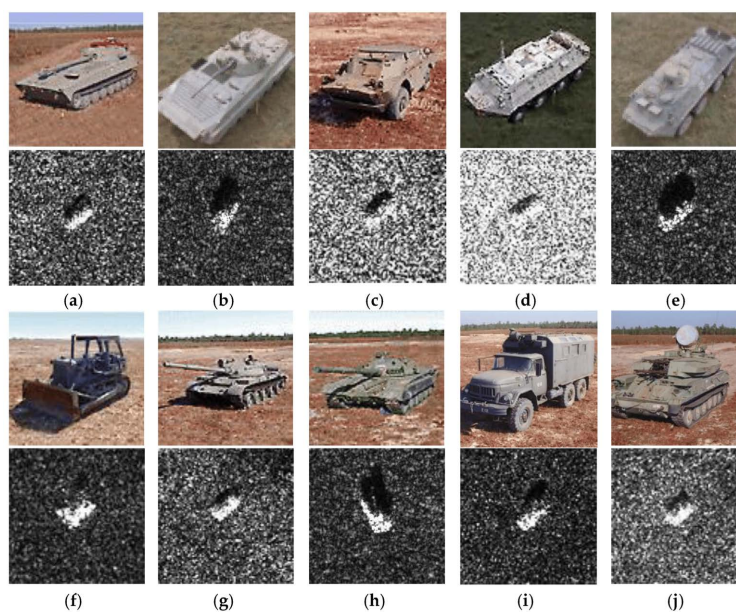


Figure 4. Ten class vehicle targets for optical images (top) and SAR images (bottom)

图 4. 针对光学图像的十类车辆目标(上方)和 SAR 图像(下方)

根据 EOC 分类, 训练集和测试集之间存在很大差异, 如不同的俯仰角; 然而, 数据集中的目标类别比 SOC 更相似。因此, 分类任务更复杂, 模型的鲁棒性比 SOC 下更为重要。根据之前的研究[18], 本研究关注三种地面铰接式车辆目标: T72、BRDM2 和 ZSU234。训练和测试样本分别采集于 17° 和 30° 的俯仰角。

3.2. 实验设置

表 1 显示了 SAR 军车数据集上 MFN 模型的架构细节。分辨率为 128×128 的输入图像同时通过 CovNeXt-SimAM 和 Swin Transformer 分支进行处理。每个分支分为三个阶段，以获取不同尺度上的局部和全局特征。

Table 1. Specific parameters of MFN model

表 1. MFN 模型的具体参数

阶段	输出尺寸	卷积 NeXt-SimAM 分支	多尺度特征融合分支	Swin Transformer 分支
-	32×32 , 96	$4 \times 4, 96$	-	$4 \times 4, 96$
阶段一	32×32 , 96	$d7 \times 7, 961 \times 1, 3841 \times 1, 96$ $\times 2d7 \times 7, 961 \times 1, 3841 \times 1, 96 \times 2$	$\rightarrow 1 \times 1, 96 \leftarrow$ $\rightarrow 1 \times 1, 96 \leftarrow$	Window size = 7×7 head = $3, 1 \times 1, 96 \times 2$ 2window size = 7×7 head = $3, 1 \times 1, 96 \times 2$
阶段二	16×16 , 192	$2 \times 2, 192$ $d7 \times 7, 1921 \times 1, 7681 \times 1, 192 \times 2d7$ $\times 7, 1921 \times 1, 7681 \times 1, 192 \times 2$	$1 \times 1, 384$ Avgpool k2, s4	Patch merging Window size = 7×7 head = $6, 1 \times 1, 192 \times 2$ window size = 7×7 head = $6, 1 \times 1, 192 \times 2$
阶段三	$8 \times 8, 384$	$2 \times 2, 384$ $d7 \times 7, 3841 \times 1, 15361 \times 1, 384 \times 2d7$ $\times 7, 3841 \times 1, 15361 \times 1, 384 \times 2$	$\rightarrow 1 \times 1, 384 \leftarrow$ $\rightarrow 1 \times 1, 384 \leftarrow$	Patch merging Window size = 7×7 head = $12, 1 \times 1, 384 \times 2$ window size = 7×7 head = $12, 1 \times 1, 384 \times 2$
分类器	$1 \times 1, 1$	-	global average pooling $1 \times 1, \text{numclass}$	-

注： $d7 \times 7, 96$ 这样的表达形式是输出的简写形式，其中 d 代表组卷积， 7×7 代表卷积核大小，96 代表通道数。Avgpool k2, s4 表示采用 2×2 的池化窗口大小和步长为 4 进行的平均池化操作。

随后，通过多尺度特征融合分支获得包含丰富 SAR 特征信息的融合特征。在全局平均池化之后，SAR 特征通过全连接层进行分类。

3.3. MSTAR 实验结果与分析

3.3.1. 贝叶斯超参数优化结果

采用贝叶斯方法对 MFN 的四个超参数进行优化：epoch、学习率(lr)、预热和衰减(wd)超参数以及批量大小。SOC 和 EOC 实验中贝叶斯优化的结果如表 2 所示。下一个实验将使用最优的超参数集对 SAR 图像进行分类。

Table 2. Optimization results of Bayesian hyperparameters in SOC and EOC experiments using the MSTAR dataset

表 2. 在 MSTAR 数据集上使用贝叶斯超参数进行 SOC 和 EOC 实验的优化结果

超参数	SOC	EOC
时间	180	180
批量大小	16	8
学习率	0.0039	0.00016
衰减	0.0335	0.0356

3.3.2. SOC 实验结果

在 SOC 实验中，表 3 展示了关于 10 种 SAR 车辆目标的详细信息。在训练集和测试集中， 17° 和 15° 俯仰角度下，分别有 2747 个和 2452 个 SAR 车辆数据样本。

Table 3. Detailed information on vehicle targets under SOC in the MSTAR dataset
表 3. MSTAR 数据集中 SOC 下车辆目标的详细信息

目标	训练		测试	
	倾角度数	数量	倾角度数	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
ZSU_23_4	17°	299	15°	274

在 SOC 实验中, 所提出模型对于 10 种车辆目标的识别结果的混淆矩阵如图 5 所示。测试集中的数据基本上集中在对角线上, 表明这些目标被正确分类到它们的类别中, 整体识别准确率达到到了 99.26%。在 SOC 实验中, 所提出模型的平均精确度、召回率和 F1 分数分别达到了 99.18%、99.29%和 99.23%。2S1、BRDM2 和 BTR70 三种类型的精确度达到了 100%。混淆矩阵显示, ZIL131 的精确度是最低的, 一些 ZIL131 的数据被错误地识别为 BRDM2 和 BTR60, 可能是因为 ZIL131、BRDM2 和 BTR60 作为卡车具有较高的相似性。总体而言, 在 SOC 实验中, 所提出的模型展现了良好的识别性能, 有效验证了其有效性。

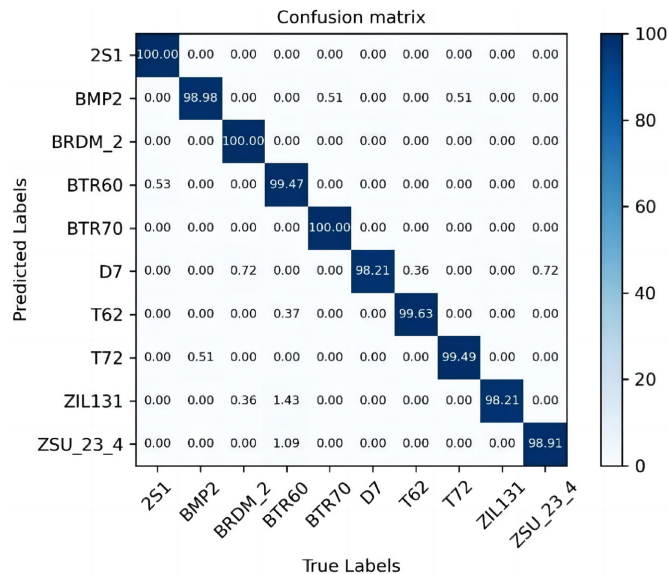


Figure 5. Ten categories of SAR vehicle target confusion matrix under SOC
图 5. SAR (合成孔径雷达)车辆目标混淆矩阵的十个类别

3.3.3. EOC 实验结果

对于 EOC 实验, 表 4 中展示了三种类型的 SAR (合成孔径雷达)车辆关节目标的详细信息。在训练和测试数据集中, 分别有 17 度和 30 度俯仰角的 SAR 车辆数据样本 896 个和 384 个。由于 SAR 数据样本较少, 目标类别之间的相似性较高, 俯仰角之间的差异较大, 因此识别任务比 SOC 实验更具挑战性, 需要更高的模型性能。

Table 4. Detailed information on vehicle targets under EOC in the MSTAR dataset
表 4. MSTAR 数据集中 EOC 下车辆目标的详细信息

目标	训练		测试	
	倾角度数	数量	倾角度数	数量
T72	17°	299	30°	133
ZSU_23_4	17°	299	30°	118
BRDM2	17°	298	30°	133

在 EOC 实验中, 所提出模型的混淆矩阵如图 6 所示, 用于展示三种类型车辆目标的识别结果。其识别准确率略低于 10 种目标的分类准确率, 但仍达到 94.27%。在 EOC 实验中, 所提出模型的平均精确率、召回率和 F1 得分分别达到 94.39%、94.40%和 94.24%。结果表明, 所提出的模型具有良好的泛化性和鲁棒性。

Table 5. Experimental results of MSTAR dataset ablation
表 5. MSTAR 数据集脱敏实验的实验结果

方法	准确度	
	SOC	EOC
ConvNeXt 分支	86.52%	69.01%
+Swin Transformer 分支	95.13%	92.71%
+SimAM	96.21%	92.97%
+贝叶斯优化	99.26%	94.27%

3.4. 脱敏实验

表 5 呈现了进行脱敏实验的结果, 以评估每个组成部分对提出的模型在 MSTAR 数据集上性能的个别影响。评估过程包括从本地路径(ConvNeXt 分支)开始; 添加 Swin Transformer 分支、SimAM 注意力模块和贝叶斯优化超参数; 最后形成我们的模型。

ConvNeXt 分支在 SOC 和 EOC 下的准确率分别为 86.52%和 69.01%。添加 Swin Transformer 分支后, 准确率分别提高了 8.61%和 23.7%, 表明仅考虑局部特征是不足的。局部细节和全局语义特征的融合显著提高了模型的性能, 验证了提出的方法。在 ConvNeXt 分支中增加 SimAM 注意力模块后, SOC 和 EOC 下的准确率分别提高了 1.08%和 0.26%, 表明注意力机制可以提升模型的特征提取和表达能力。

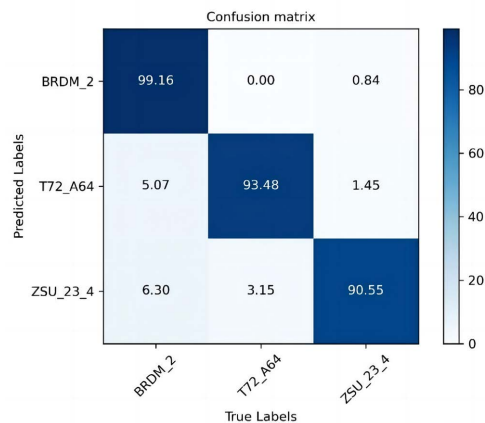


Figure 6. Three categories of SAR vehicle target confusion matrix in EOC experiments
图 6. EOC 实验中 SAR(合成孔径雷达)车辆目标混淆矩阵的三个类别

Table 6. Comparison between Bayesian optimization hyperparameters and empirical selection
表 6. 贝叶斯优化超参数和经验选择的比较

超参数	贝叶斯优化后的超参数		实证选择的超参数	
	SOC	EOC	SOC	EOC
时间	180	180	150	150
批量大小	16	8	16	16
学习率	0.0039	0.00016	1×10^{-4}	1×10^{-4}
衰减	0.0335	0.0356	0.05	0.05

最后, 使用贝叶斯优化算法来优化模型的超参数, SOC 和 EOC 实验的准确率分别提高了 3.05%和 1.3%。此外, 本研究比较了贝叶斯超参数优化算法和经验选择超参数对模型分类性能的影响。在 SOC 和 EOC 下, 贝叶斯算法优化的超参数集和经验选择的超参数集的结果如表 6 所示。此外, 使用经验选择的超参数集和贝叶斯超参数优化算法优化的超参数集对 MFN 模型进行对比的结果如表 7 所示。

Table 7. Comparison of model performance before and after using the Bayesian optimization algorithm in the MSTAR dataset
表 7. MSTAR 数据集中使用贝叶斯优化算法前后模型性能的比较

	SOC				EOC			
	准确性	精确度	召回率	F1-Score	准确性	精确度	召回率	F1-Score
MFN	96.21%	95.95%	96.18%	95.99%	92.97%	93.01%	93.19%	92.93%
Bayes-MFN	99.26%	99.18%	99.29%	99.23%	94.27%	94.39%	94.40%	94.24%

如表 7 结果表明, 使用贝叶斯优化算法有效地提高了模型在 SOC 和 EOC 下的分类性能, 涵盖了所有四个评估指标。此外, 在模型优化过程中, 贝叶斯超参数优化算法显著节省了时间和资源。

4. 结论

1) 在 SAR 自动目标识别技术中, 散斑噪声会对分类模型的性能产生负面影响。为了解决这个问题, 提出了一种 MFN 方法, 通过并行化方法来保留 SAR 局部细节特征和全局语义信息, 以增强 SAR 纹理细节特征, 并改善散斑噪声引起的类内差异大和类间距离小等问题。

2) 采用贝叶斯超参数优化算法对 MFN 的超参数进行优化。实验结果表明, 所提出的模型在 MSTAR 数据集上具有特定的泛化性和有效性。同时, 与其他主流模型相比, Bayes-MFN 在 SOC 和 EOC 下取得了出色的性能。

3) 设计了并行的局部和全局特征分支, 以充分探索不同层次上的局部和全局特征。此外, 还添加了 SimAM 注意机制到局部特征模块中, 以增强网络特征的学习能力。这种增强使得网络能够更多关注有限特征信息中 SAR 图像的纹理细节, 并有效抑制散斑噪声。构建了多尺度特征融合分支, 其中包括 FFU 和下采样模块。该分支自适应地融合了不同尺度上的局部细节特征和全局语义信息。提出了基于贝叶斯超参数优化算法的三分支多尺度特征融合模型, 充分结合了 ConvNeXt 和 SwinTransformer 的优势, 在 MSTAR 数据集上以更少的参数和更高的识别准确性取得了出色的结果。

参考文献

- [1] Wang, K., Zhang, G., Leng, Y. and Leung, H. (2018) Synthetic Aperture Radar Image Generation with Deep Generative Models. *IEEE Geoscience and Remote Sensing Letters*, **16**, 912-916. <https://doi.org/10.1109/LGRS.2018.2884898>
- [2] Gao, F., Yang, Y., Wang, J., Sun, J., Yang, E. and Zhou, H. (2018) A Deep Convolutional Generative Adversarial Networks (DCGANs)-Based Semi-Supervised Method for Object Recognition in Synthetic Aperture Radar (SAR) Im-

- ages. *Remote Sensing*, **10**, 846. <https://doi.org/10.3390/rs10060846>
- [3] Wang, L., Bai, X., Xue, R. and Zhou, F. (2021) Few-Shot SAR Automatic Target Recognition Based on Conv-BiLSTM Prototypical Network. *Neurocomputing*, **443**, 235-246. <https://doi.org/10.1016/j.neucom.2021.03.037>
- [4] Novak, L.M., Owirka, G.J. and Brower, W.S. (1997) The Automatic Target-Recognition System in SAIP. *Linc.Lab.J.*, **10**, 187-202.
- [5] Hummel, R. (2000) Model-Based ATR Using Synthetic Aperture Radar. *Proceedings of the IEEE International Radar Conference*, Arlington, VA, USA, 7-12 May 2000, 856-861.
- [6] Liu, Z., Mao, H. and Wu, C.Y. (2022) A Convnet for the 2020s. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, New Orleans, LA, USA, 19-24 June 2022, 11976-11986. <https://doi.org/10.1109/CVPR52688.2022.01167>
- [7] Dong, H., Zhang, L. and Zou, B. (2021) Exploring Vision Transformers for Polarimetric SAR Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, 5219715. <https://doi.org/10.1109/TGRS.2021.3137383>
- [8] Vaswani, A., Shazeer, N. and Parmar, N. (2017) Attention Is All You Need. *Proceedings of the Advances in Neural Information Processing Systems*, Long Beach, CA, USA, 4-9 December 2017, 147-152.
- [9] Dosovitskiy, A., Beyer, L. and Kolesnikov, A. (2020) An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale. arXiv:2010.11929
- [10] Liu, Z., Lin, Y. and Cao, Y. (2021) Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Montreal, QC, Canada, 10-17 October 2021, 10012-10022. <https://doi.org/10.1109/ICCV48922.2021.00986>
- [11] Yang, L., Zhang, R.Y., Li, L. and Xie, X. (2021) Simam: A Simple, Parameter-Free Attention Module for Convolutional Neural Networks. *Proceedings of the International Conference on Machine Learning*, Virtual, 18-24 July 2021, 11863-11874.
- [12] Li, S., Wang, S., Dong, Z., Li, A., Qi, L. and Yan, C. (2022) PSBCNN: Fine-Grained Image Classification Based on Pyramid Convolution Networks and SimAM. *Proceedings of the IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Cloud and Big Data Computing, International Conference on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech)*, Falerna, Italy, 12-15 September 2022, 1-4. <https://doi.org/10.1109/DASC/PiCom/CBDCCom/Cy55231.2022.9927801>
- [13] You, H., Lu, Y. and Tang, H. (2023) Plant Disease Classification and Adversarial Attack Using SimAM-EfficientNet and GP-MI-FGSM. *Sustainability*, **15**, 1233. <https://doi.org/10.3390/su15021233>
- [14] Yu, T. and Zhu, H. (2020) Hyper-Parameter Optimization: A Review of Algorithms and Applications. arXiv:2003.05689.
- [15] Bergstra, J., Komer, B., Eliasmith, C., Yamins, D. and Cox, D.D. (2015) Hyperopt: A Python Library for Model Selection and Hyperparameter Optimization. *Computational Science & Discovery*, **8**, 014008. <https://doi.org/10.1088/1749-4699/8/1/014008>
- [16] Zhang, J., Wang, Q. and Shen, W. (2022) Hyper-Parameter Optimization of Multiple Machine Learning Algorithms for Molecular Property Prediction Using Hyperopt Library. *Chinese Journal of Chemical Engineering*, **52**, 115-125. <https://doi.org/10.1016/j.ciche.2022.04.004>
- [17] Bergstra, J., Bardenet, R. and Bengio, Y. (2011) Algorithms for Hyper-Parameter Optimization. *Proceedings of the Advances in Neural Information Processing Systems 24*, Granada, Spain, 12-15 December 2011, 241-253.
- [18] Kang, K. and Ryu, H. (2019) Predicting Types of Occupational Accidents at Construction Sites in Korea Using Random Forest Model. *Safety Science*, **120**, 226-236. <https://doi.org/10.1016/j.ssci.2019.06.034>