

DOI:10.20079/j.issn.1001-893x.240712001

利用多尺度卷积注意力的宽带信号稀疏检测方法*

龚安¹, 张静蕾¹, 郭兰图², 赵晓蕾², 刘玉超²

(1. 中国石油大学(华东)青岛软件学院、计算机科学与技术学院, 山东 青岛 266580;
2. 中国电波传播研究所, 山东 青岛 266107)

摘要: 宽带侦察场景下, 虽然信号检测识别准确率高, 但计算资源消耗过大的问题亟待解决。为此, 提出了一种基于多尺度卷积注意力的稀疏检测方法 (Multi-scale Convolution Attention Sparse Detection, MSCA-S)。该方法结合信号时频图的先验知识, 通过建模信号在时间轴上的远距离依赖关系并抑制频率轴的无关干扰, 设计了多尺度水平卷积注意力机制 (Multi-scale Horizontal Convolution Attention, MSHCA), 联合提取信号的多维特征, 有效提升检测识别精度, 并通过水平卷积降低模型计算复杂度。基于 MSHCA, 构建了层次化堆叠的宽带信号检测方法, 利用稀疏特征参数进一步减少计算资源需求。在频谱范围为 2.5 MHz 的青岛实采及仿真宽带信号数据集上进行实验, MSCA-S 在不同信噪比下的平均检测精度达 95.6%, 相比频率敏感宽带信号检测方法、基于 Swin-Transformer 的协议信号识别方法和基于 101 层残差网络的信号检测方法, 精度分别提升了 0.05%、2.94% 和 6.14%, 计算量分别降低了 1.53×10^{10} 、 1.79×10^{10} 和 4.59×10^{10} 。

关键词: 宽带信号检测识别; 注意力机制; 多尺度卷积; 稀疏算法

开放科学(资源服务)标识码(OSID):



微信扫描二维码
听独家语音释文
与作者在线交流
享本刊专属服务

中图分类号: TN911.7 文献标志码: A 文章编号: 1001-893X(2025)11-1737-10

A Sparse Detection Method for Broadband Signals by Utilizing Multi-scale Convolutional Attention

GONG An¹, ZHANG Jinglei¹, GUO Lantu², ZHAO Xiaolei², LIU Yuchao²

(1. Qingdao Institute of Software, College of Computer Science and Technology, China University of Petroleum (East China),
Qingdao 266508, China; 2. China Radio Wave Propagation Research Institute, Qingdao 266107, China)

Abstract: In broadband reconnaissance scenarios, achieving high signal detection accuracy often entails significant computational costs. To address this, a multi-scale convolution attention sparse detection (MSCA-S) method is proposed, which incorporates prior knowledge of signal spectrograms by capturing long-range temporal dependencies and suppressing irrelevant frequency-domain interference. MSCA-S introduces a multi-scale horizontal convolution attention (MSHCA) mechanism that jointly extracts multi-dimensional signal features, enhancing detection accuracy while reducing computational complexity through horizontal convolution. Building on MSHCA, a hierarchically stacked broadband signal detection framework is developed, and sparse feature parameters are used to further optimize computational efficiency. MSCA-S is evaluated on a real-world and simulated broadband signal dataset (2.5 MHz spectrum) collected in Qingdao, achieving an average detection accuracy of 95.6% across varying signal-to-noise ratios. Compared with the frequency-sensitive signal detector, the Swin-Transformer-based protocol recognition method, and the Res-101 detection method, MSCA-S improves accuracy by 0.05%, 2.94%, and 6.14%, respectively, while reducing computational costs by 1.53×10^{10} , 1.79×10^{10} , and 4.59×10^{10} , respectively.

Key words: broadband signal detection and recognition; attention mechanisms; multi-scale convolution; sparse algorithms

* 收稿日期: 2024-07-12; 修回日期: 2024-10-31

基金项目: 国家自然科学基金重点项目 (U20B2038)

通信作者: 赵晓蕾 Email: zhaoxiaoleiqz@126.com

0 引言

宽带信号的精准检测识别是军用和民用通信领域中不可或缺的组成部分,也是非合作方开展目标信号获取、电子对抗等频管任务的重要基础^[1-2]。宽带信号检测识别是指在宽带接收条件下,检测出多路信号的存在性和特征参数,并识别各信号类型^[3]。传统的信号检测识别方法需要领域专家直接设计采样信号的特征,不仅费时费力,而且识别精度低,通用性差。

计算机视觉的目标检测任务是确定图像中所有感兴趣物体的位置和类别。宽带信号时频图水平轴表示时间,垂直轴表示频率,颜色深浅表示信号强度。与传统的检测方法不同,采用目标检测的方式对信号时频图进行分析,并结合深度学习来解决信号检测问题的方法为信号检测识别提供了创新思路^[4-6],越来越广泛地应用于宽带信号的精准检测识别领域。深度学习在射频领域的应用主要集中在从宽带捕获信号后的窄带信号识别任务上^[7-9]。对于信号种类多、收发量大且存在短期突发信号的非合作频管宽带场景,基于深度学习的多信号检测识别正面临严峻挑战。

现如今,已有学者将深度学习应用在宽带信号检测识别领域。Zha 等人^[10]使用 SSD (Single Shot MultiBox Detector) 和多输出卷积神经网络 (Convolutional Neural Network, CNN) 分别实现了多信号检测和识别。Yuan 等人^[11]利用 CNN 多信号感知模块从候选信号中提取特征,实现了莫尔斯信号的检测识别。Li 等人^[12]通过改进的 YOLOv3 目标检测网络处理信号时频图,实现了宽带多信号的检测识别。Guo 等人^[13]使用基于 YOLOv5 的混合信号目标检测模型,实现宽带混叠信号的检测识别。然而,上述算法均将信号时频图当作普通图像处理,没有考虑特有的物理意义,导致在宽带环境下对信号检测识别出现虚报误判,准确率不高。

针对宽带信号检测识别准确率不高的问题,朱等人^[14]提出一种基于 Swin Transformer^[15]的短波信号识别算法,显著提升了识别准确率;耿等人^[16]提出了基于频率敏感注意力的特定信号检测方法 FSSDet,该方法考虑了信号时频图的物理意义,能够捕获远程依赖特性,从而取得了较高的检测识别准

准确率。FSSDet 的研究证明,增强模型在时间轴上的远程建模能力可有效提升宽带信号检测识别效果。虽然上述研究通过 Transformer 等模型的强大特征提取能力提高了准确率,但这些模型参数量大,需要大量数据支撑,对计算资源的需求也很高。目前这些问题并未得到有效解决。

针对现有宽带信号检测识别方法难以同时兼顾高准确率和低计算资源的问题,本文提出一种基于多尺度卷积注意力的稀疏检测方法 (Multi-scale Convolution Attention Sparse Detection, MSCA-S)。该方法适用于非合作宽带频管场景,利用目标检测的方式对宽带信号时频图进行信号定位和分类,进而实现信号的检测识别。为确保较高的检测识别准确率,本文参考文献^[16]的思路,考虑到信号时频图在时间轴上有较大的跨度和远程依赖关系、频率轴上却呈现局域性的先验知识,对现有的卷积注意力模型^[17]进行改进,提出多尺度水平卷积注意力机制 (Multi-scale Horizontal Convolution Attention, MSHCA)。该机制在时间轴方向进行卷积注意力计算,一方面可以充分捕获时间轴方向的相关性;另一方面限制频率轴的感受野,减少干扰和相邻信号的影响。由于宽带信号识别是一项密集识别任务,多尺度信息交互可以适应不同信号的差异。基于此,本文将多尺度水平卷积注意力机制的构建块按照分层结构堆叠,设计出多尺度水平卷积注意力的特征提取网络 (Multi-scale Horizontal Convolutional Attention Network, MSHConvNet),扩大感受野以获得信号的层次化特征表达。

具体来说,本文主要贡献如下:

1) 设计了多尺度水平卷积注意力机制 (MSHCA),并基于此构建了 MSHConvNet。该网络能够更有效地提取时频图中不同的信号特征,提升非合作频管场景下宽带信号的识别准确率。

2) 以 Sparse R-CNN^[18] 为目标检测框架,将 MSHConvNet 作为特征提取骨干网络,搭建了一种轻量化、可学习的端到端宽带信号检测识别方法 MSCA-S,提高了运行效率。

3) 在包含 12 类典型信号的自建宽带数据集上与现有经典方法进行对比实验验证,结果表明 MSCA-S 在 2.5 MHz 频谱范围数据上有更高平均识

别准确率的同时具备更低计算量需求。

1 系统模型

MSCA-S 是基于 Sparse R-CNN 设计的一种端到端的纯稀疏信号检测识别方法。该方法使用 N 个可学习候选框和对应的可学习候选特征取代以往目标检测器^[19-20]中 $H \times W \times k$ 的密集候选框和密集特征交互机制,减少参数量并加快模型收敛速度。

如图 1 所示, MSCA-S 的总体框架主要由三部分组成: MSHACovNet、动态实例交互头 (Dynamic Instance Interactive Head, DIIH) 以及分类和定位的预测头。

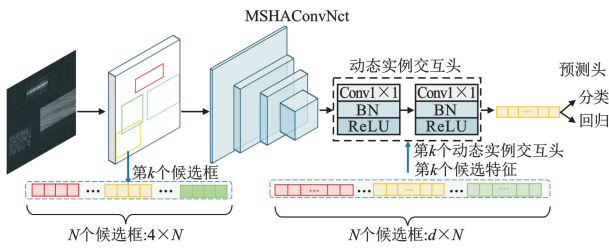


图 1 MSCA-S 框架

MSCA-S 检测识别流程: 首先, 对原始 IQ 数据进行预处理; 然后, 对预处理后的宽带信号时频图进行回归和分类, 获取信号位置参数和类别; 最后, 将位置参数转换为信号的中心频率、带宽等特征参数结果。

MSCA-S 以信号时频图、 N 个候选框和候选特征的代表参数为输入, 候选框和候选特征在训练过程通过反向传播更新。可学习的候选框 ($4 \times N$) 由

0~1 的四维参数构成, 用于粗略地表示归一化后时频图信号的中心坐标、高度、宽度。可学习的候选特征 ($d \times N$) 是一个用于细致描述信号位置和形态的 256 维隐向量。由于可学习特性, MSCA-S 在实验过程中表现出较低的初始化敏感度, 这使其在不同初始条件下均能保持较高的性能稳定性, 从而显著提升模型的灵活度和鲁棒性。

MSCA-S 处理信号时频图时, MSHACovNet 利用 RoIAlign 操作为 N 个候选框在宽带信号时频图中提取多尺度特征图。每个 RoI 特征 f_i 与对应的候选特征 p_i 在 DIIH 中进行交互, 过滤掉无效的候选框, 生成最终特征。DIIH 是由 2 个连续的 1×1 卷积组成, 并通过 ReLU 激活函数处理。随后输入到各自的专用头中根据候选特征进行信号定位和分类, 输出信号类别和位置的四维归一化向量。

为适应信号时频图和普通图像的差异, 除了对特征参数、网络结构等进行调整外, 本文主要的改进是设计了特征提取网络 MSHACovNet。

1.1 多尺度水平卷积注意力的特征提取网络

如图 2 所示, MSHACovNet 采用 4 个阶段的分层结构, 获得层级化特征表达。每个阶段由一个下采样模块和 L_i 个堆叠的多尺度水平卷积注意力构建块 (Multi-scale Horizontal Convolution Attention block, MSHCA block) 组成, 实现更大的感受野和更全面的特征表示。具体的 MSHCA block 结构在 1.3 中详细说明。实验证明在 MSHACovNet 中使用批量归一化 (Batch Normalization, BN) 比层归一化对分割性能的增益更大。

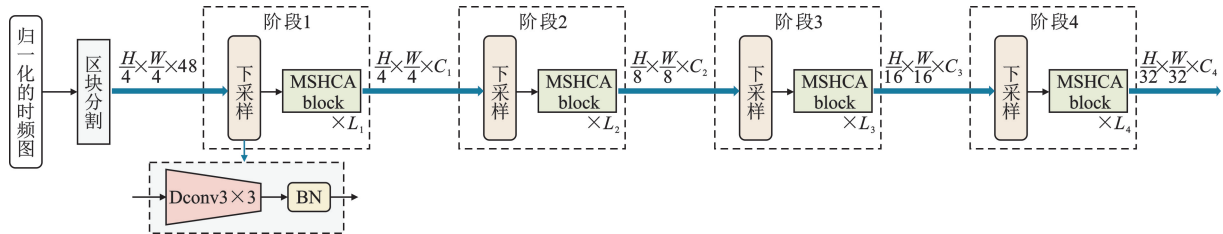


图 2 MSHACovNet 结构

特征提取时, 块分割模块将宽带信号时频图分割成 4×4 尺寸的图像块, 分层结构的 4 个阶段空间分辨率依次递减, 特征图尺寸依次为 $\frac{H}{4} \times \frac{W}{4} \times C_1$ 、 $\frac{H}{8} \times \frac{W}{8} \times C_2$ 、 $\frac{H}{16} \times \frac{W}{16} \times C_3$ 和 $\frac{H}{32} \times \frac{W}{32} \times C_4$, 其中, H 和 W 分别为

输入图像的高和宽, C_i 为第 i 个阶段输出通道数。

本文每个阶段的通道数 $[C_1, C_2, C_3, C_4] = [64, 128, 320, 512]$, 每个阶段堆叠 MSHCA block 数 $[L_1, L_2, L_3, L_4] = [3, 3, 5, 3]$ 。

1.2 $1 \times k$ 水平条形卷积

卷积核通常为方形, 而信号时频图中大部分信号呈现窄条状, 如图 3(a) 所示, 形状不匹配导致资源浪费, 在频轴上也引入了无关干扰。

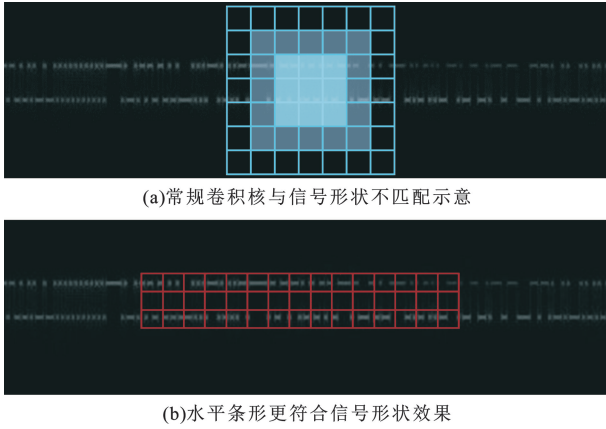


图3 卷积核与信号适配情况的对比

受全局卷积神经网络^[21]启发,以 $k = 15$ 为例,要获取 15×15 感受野,用 1×15 的水平卷积核串联 15×1 的垂直卷积核的参数量仅为直接使用 15×15 卷积参数量的 13%。

为了模型的轻量化并获得模型更好的非线性表达效果,结合信号时频图的依赖关系主要体现在时间轴方向上,本文在 MSHCA 机制中采用了 $1 \times k$ 的水平条形滤波器, $k \in \{5, 11, 15, 21\}$,以限制信号频率轴的感受野。如图 3(b)所示,条形滤波器的感受野与信号形状更匹配,可以缓解因感受野与信号特

征不匹配所导致的特征提取效果差。

1.3 多尺度水平卷积注意力机制

本文提出 MSHCA 机制,将信号检测问题转换为目标检测问题,同时区分了信号时频图和普通图像的差异。该机制仅使用不同尺度水平条形滤波器并联的方式来扩张卷积,适合挖掘信号时频图的物理意义。该机制的核心模块如图 4(a)所示,主要由 3 个部分组成: 3×3 的深层可分离卷积用于汇集局部信息;不同尺寸的水平卷积用来捕获差异化邻近背景特征表示; 1×1 卷积核的卷积用来建模融合不同通道之间的关系。

信号持续时间差异导致时间轴跨度变化较大,MSHCA 机制旨在加强时间轴的关联关系,聚焦重点信息,抑制频率轴的无关干扰,跨时间轴提取特征。

本文进一步设计基于 MSHCA 机制的多尺度水平卷积注意力构建块 MSHCA block,其结构如图 4(b)所示。其中, \oplus 表示 Add 运算, \otimes 表示按元素进行的矩阵乘法运算。整个构建块采用残差网络结构, BN 层的加入缓解了梯度消失和梯度爆炸问题。激活函数使用的是高斯误差线性单元激活函数 (Gaussian error Linear Unit, GeLU),它可以保留概率性和输入的依赖性。

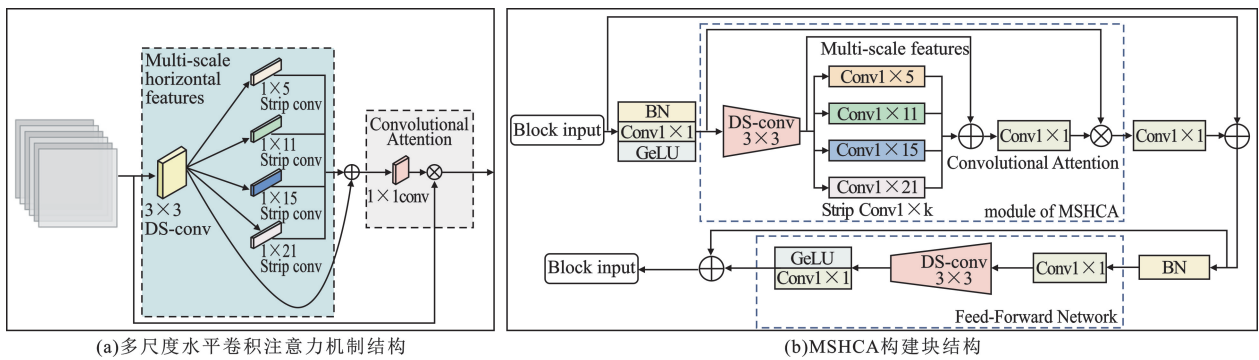


图4 MSHCA 模型结构

构建块首先通过 BN、 1×1 卷积和激活函数进行卷积操作,形式如下:

$$F = \text{GeLU}(\text{Conv}_{1 \times 1}(\text{BN}(Z_i))) \quad (1)$$

式中: Z_i 为第 i 个 MSHCA block 的输入; BN 为批归一化操作; $\text{Conv}_{1 \times 1}$ 是 1×1 卷积操作。

随后执行 MSHCA 机制。MSHCA 模块的输入是卷积操作的输出 F ,可直接作为注意力权重进行重新赋权,此过程可以描述为以下形式:

$$\text{Atten} = \text{Conv}_{1 \times 1}(\sum_{i=0}^5 \text{Scale}_i(\text{DS_Conv}(F))) \quad (2)$$

$$y_{\text{MSHCA}} = \text{Atten} \otimes F \quad (3)$$

式中: F 为 MSHCA 模块的输入特征,同时也是卷积操作的输出; DS_Conv 表示 3×3 的深度可分离卷积,由逐深度卷积和逐点卷积组成; $\text{Scale}_i, i \in \{1, 2, 3, 4\}$ 表示第 i 个多尺度分支,分别是 1×5 、 1×11 、 1×15 、 1×21 的水平条形卷积, Scale_0 是 DS_Conv(F) 的本身连接; \otimes 是按元素进行的矩阵乘法运算,利用多尺度卷积特征通过按元素乘法来实现空间注意力; Atten 是注意力机制的映射; y_{MSHCA} 为注意力机制的输出。

最终,MSHCA 模块的输出依次经过 1×1 卷积层、BN 层和前馈网络(Feed-Forward Network, FFN)。在 FFN 中,结构依次为 1×1 的卷积层、 3×3 的深度可分离卷积、GeLU 激活函数和 1×1 卷积。FFN 结构使特征提取方法的表达能力更强,能更好地表示信号与周围区域之间的作用关系。该过程的表现形式如下:

$$\hat{\mathbf{Z}}_i = \text{Conv}_{1 \times 1}(\mathbf{y}_{\text{MSHCA}}) + \mathbf{Z}_i \quad (4)$$

$$\mathbf{Z}_{i+1} = \text{FFN}(\text{BN}(\hat{\mathbf{Z}}_i)) + \hat{\mathbf{Z}}_i \quad (5)$$

式中: $\mathbf{y}_{\text{MSHCA}}$ 为注意力机制的输出; \mathbf{Z}_i 和 \mathbf{Z}_{i+1} 分别为第 i 个 MSHCA 构建块的输入和输出;FFN 的运算如图 4(b) 中所示意的 Feed-Forward Network 结构;+ 运算表示残差结构。

1.4 损失函数

在分类和坐标预测集合中使用集合预测损失,产生一个最优的二分匹配。

匹配损失如下:

$$\mathbf{L} = \lambda_{\text{cls}} \times \mathbf{L}_{\text{cls}} + \lambda_{\text{L1}} \times \mathbf{L}_{\text{L1}} + \lambda_{\text{giou}} \times \mathbf{L}_{\text{giou}} + \lambda \|\mathbf{w}\|_2^2 \quad (6)$$

式中: \mathbf{L}_{cls} 为预测类别和真实类别的损失; \mathbf{L}_{L1} 和 \mathbf{L}_{giou} 分别是预测框和真实框的归一化 4 维向量之间的 L1 损失和 IoU 损失; λ_{cls} 、 λ_{L1} 、 λ_{giou} 为各分量的权重系数; $\lambda \|\mathbf{w}\|_2^2$ 是 L2 正则化项。训练损失函数和匹配损失相同,最终所有损失是训练批次的总和。

综上所述,MSCA-S 通过引入了水平条形卷积和稀疏检测框架,其计算量和参数量在理论上远低于基于 Transformer 和其他基于 CNN 的检测方法,而且 MSHCA 机制结合层次堆叠结构对 MSCA-S 的信号特征提取进行了适应性改进。MSCA-S 在理论上可以提高时间轴方向的远程建模能力,从而提高了检测识别准确率。

2 实验与结果分析

2.1 数据集和预处理

2.1.1 数据集

由于缺乏模拟宽带环境的公开信号数据集,本文根据多信道信号的特点,设计了一种模拟宽带场景的信号数据集构建方法,用于训练和验证 MSCA-S 的有效性。信号数据是由实采和仿真相结合获取的,确保对实测和仿真信号均有验证。具体来说,1Q 信号采样率为 2.5 MHz,仿真信号的信噪比范围在 $-10 \sim 20$ dB 之间,步长为 2 dB。最终生成 12 类常

用的典型窄带信号,信号参数如表 1 所示。

表 1 信号参数表

信号调制类型	带宽/kHz	生成方式	采集地点
8FSK	150	实采	青岛
32QAM	400	实采	青岛
0.3GMSK	27	实采	青岛
2FSK	40	实采	青岛
PPM	70	实采	青岛
$\Pi/4$ QPSK	100	实采	青岛
FM	16	实采	青岛
4PSK	48	仿真	无需采集
QPSK	100	仿真	无需采集
16QAM	387	仿真	无需采集
8PSK	24	仿真	无需采集
4FSK	3	仿真	无需采集

仿真信号的数据标签是已知的,而实采信号数据则使用了窄带收集设备和双滑窗的能量检测算法,完成对信号中心频率、带宽和起止时间的自动标注。由于是检验非合作场景下侦察信号的任务,发射源的信息不是必要标签。

为了模拟宽带信道环境,将不同时间、地点、频率范围采集的信号数据作为仿真的背景数据,并在仿真过程中加入了随机频偏和相偏,进一步添加了 Chirp 探测干扰和加性噪声等^[22]。

对于这 12 种窄带信号,本文随机将 $K(1 \leq K \leq 12)$ 种经过去直流、归一化的窄带信号通过上采样和频率搬移映射至背景数据的时间和频率上。信号持续时间在 $1 \sim 6$ s 之间,不同信号的中心频率随机分布在 $100 \sim 2\,500$ kHz 之间,信号间的频差为 10 kHz 以上。将上述操作得到的宽带信号切分为 20 s 的信号片段,总共得到 77 805 条数据。

2.1.2 预处理

本实验使用短时傅里叶变换(Short-Time Fourier Transform, STFT)对原始 IQ 数据进行预处理。将信号分割为多个互相重叠的时间窗口,并对每个窗口进行傅里叶变换,得到各窗口的频谱。随后将所有窗口的频谱组合起来再进行取模值、取对数处理,得到信号时频图。最后归一化处理,得到 $65\,536 \times 2\,663$ 的宽带信号时频数据。

信号数据预处理的具体设置参数如表 2 所示。与一般频扫设备等相关仪器的 kHz 数量级相比,本文预处理使用的频率分辨率为 40 Hz,能有效地发现并识别短时突发信号。

表 2 预处理参数表

参数名称	具体设置
STFT 点数	65 536
窗类型	hann
窗长	63 356
窗重叠率/%	70
时间分辨率/ms	8
频率分辨率/Hz	40

2.2 实验环境和参数设置

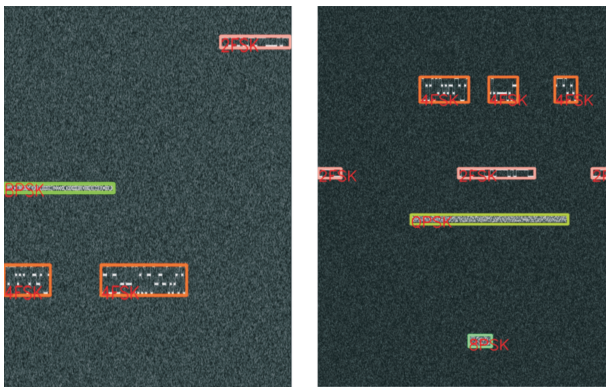
本文检测方法直接对宽带信号时频图进行检测识别,按照 7 : 2 : 1 的比例划分训练集、测试集、验证集。训练的批大小设为 32,轮次设为 34,初始学习率设为 0.000 1,dropout 设为 0.3,使用 adam 优化器,使用多阶段预热(warm-up)学习策略。本文实验采用 65 536×2 663 尺寸的信号时频图数据集进行验证和测试。特别地,为了实验效率,在性能效率对比实验中使用的是 800×512 尺寸信号时频图。本文实验所用的具体参数如表 3 所示。

表 3 实验环境参数

名称	参数
操作系统	Ubuntu 20.04.5 LTS
CPU	Intel(R) Xeon(R) Gold 6348 CPU @ 2.60 GHz×2
GPU	NVIDIA A100-PCIE-40GB×4
内存/GB	768
编译环境	PyCharm-Professional-2023.1 x64
Anaconda 版本	Anaconda23.1.0(Python 3.10)
深度学习框架	Torch2.0.1

2.3 检测结果转换

MSCA-S 信号检测时不仅会输出信号类型,还会通过反向传播算法学习到信号检测框的 4 维归一化表示(中心点坐标、高度、宽度),检测结果如图 5 所示。最后需要对 4 维向量进行结果转换,得到信号特征。



(a)检测结果示例1

(b)检测结果示例2

图 5 MSCA-S 检测结果

信号的中心频率 f_{MSCA-S} 可通过以下公式计算:

$$f_{MSCA-S} = F_{start} + (1-f) \times (F_{start} - F_{end}) \quad (7)$$

带宽 BW_{MSCA-S} 的计算公式为

$$BW_{MSCA-S} = (F_{start} - F_{end}) \times h \quad (8)$$

起始时间 t_{start} 的计算公式为

$$t_{start} = T_{start} + (T_{start} - T_{end}) \times (t-w \div 2) \quad (9)$$

终止时间 t_{end} 则通过以下公式计算:

$$t_{end} = T_{start} + (T_{start} - T_{end}) \times (t+w \div 2) \quad (10)$$

式(7)~(10)中:(t, f, w, h)为某一个信号检测后的归一化结果, t, f 分别表示时间轴和频率轴上的中心点归一化坐标, w, h 分别为信号在时间轴方向和频率轴方向的归一化长度; F_{start} 和 F_{end} 分别表示时频图的起止频率; T_{start} 和 T_{end} 分别为时频图的起止时间。

2.4 评价指标

本实验采用了目标检测任务中普遍使用的平均精度均值(mean Average Precision, mAP)作为评价指标,具体计算公式如下:

$$AP = \sum_{i=1}^{n-1} (r_{i+1} - r_i) p_{interp}(r_{i+1}) \quad (11)$$

$$mAP = \frac{\sum_{i=1}^K AP_i}{K} \quad (12)$$

式中: r 为召回率; p_{interp} 为下一个召回率与当前召回率之间的最大准确率;AP是召回率对应准确率的平均值; AP_i 为第*i*类的AP值; K 为总类别数。

除 mAP 之外,本文还设计了更适合信号检测任务的信号误差指标,误差越小检测结果越准确。中心频率的相对误差 f_{err} 定义如下:

$$f_{err} = \begin{cases} \frac{|f_{MSCA-S} - f_{real}|}{BW_{real}}, & \text{if } |f_{MSCA-S} - f_{real}| \leq BW_{real} \\ 1, & \text{else} \end{cases} \quad (13)$$

带宽误差 BW_{err} 定义如下:

$$BW_{err} = 1 - \frac{BW_{real} \cap BW_{MSCA-S}}{BW_{real} \cup BW_{MSCA-S}} \quad (14)$$

起止时间误差 T_{err} 定义如下:

$$T_{err} = 1 - \frac{T_{real} \cap T_{MSCA-S}}{T_{real} \cup T_{MSCA-S}} \quad (15)$$

式(13)~(15)中: f_{real} 、 BW_{real} 和 T_{real} 分别为信号的真实频率、真实带宽和真实时间段; f_{MSCA-S} 、 BW_{MSCA-S} 、 T_{MSCA-S} 分别为 MSCA-S 检测出该信号的中心频率、带宽和检测结果转换的信号时间段。

2.5 实验结果与分析

2.5.1 不同信噪比下各检测方法的对比实验

将本文的检测方法与目标检测任务中具有代表

性的基于 Swin Transformer 的 Swin-T 检测方法^[14]、信号检测识别任务中具有代表性的基于 FSA Transformer 的检测方法 FSSDet^[16] 和基于卷积神经网络的经典检测方法 Res-101^[23] 进行对比实验。为控制变量,以上对照组的检测方法都仅保留核心特征提取网络,检测框架均与 MSCA-S 保持一致。

2.5.1.1 不同信噪比下各检测方法的 mAP

各检测方法在信噪比为 -10~20 dB 范围内信号类型识别 mAP 的对比结果如图 6 所示,验证了方法在不同信噪比环境的识别准确率。mAP 计算所采用的 IoU 为 0.5~0.95,步进为 0.05。

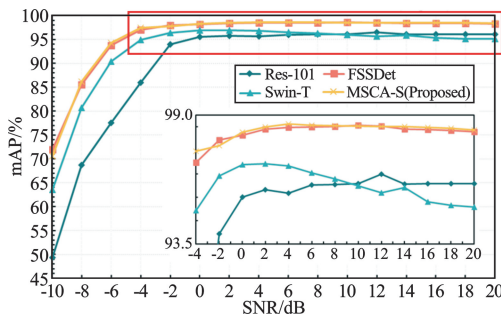


图 6 不同信噪比下各检测方法 mAP 对比

通过对比图可知, MSCA-S 在信噪比为 -10~20 dB 的 mAP 平均为 95.6%, 分别比 Res-101、

Swin-T、FSSDet 提高 6.14%、2.94%、0.05%。MSCA-S 在信噪比为 0~20 dB 的 mAP 平均为 98.49%, 分别比 Res-101、Swin-T、FSSDet 提高了 2.51%、2.44%、0.08%。MSCA-S 在不同信噪比条件下均比 Swin-T、Res-101 的检测识别结果提高 2% 以上,低信噪比尤为明显。而对于同样使用加强时域远程建模能力思想的 FSSDet 来说,除了信噪比为 -10 dB、-2 dB、10 dB、12 dB 这 4 个场景外, MSCA-S 的 mAP 均高于 FSSDet, 总体来说两者相差 1% 左右。

从实验结果分析,在识别 2.5 MHz 宽带信号数据方面,本文提出的 MSCA-S 效果整体略优于其他信号检测方法,证明了加强时域远程建模能力的 MSHCA 机制在提高检测识别准确率上有很好的效果。

2.5.1.2 不同信噪比下各检测方法的误差

各检测方法对于信噪比大于 10 dB 的识别准确率很高且相差不大,因此重点设计 -10~10 dB 信噪比场景下各检测方法的平均中心频率误差、平均带宽误差和平均起止时间误差对比实验,计算误差所使用的 IoU 范围为 0.75~0.95。信号误差对比情况如图 7 所示。

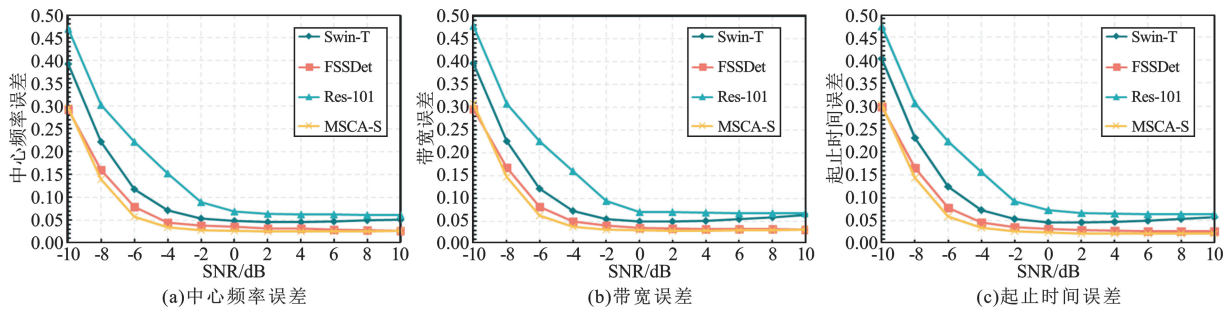


图 7 不同信噪比下各检测方法的误差对比

由实验结果可知, MSCA-S 与 Swin-T 和 Res-101 相比,中心频率误差、带宽误差、起止时间误差依次降低了 0.021 以上、0.019 以上、0.022 以上。与 FSSDet 相比,在信噪比为 -10 dB 时, MSCA-S 误差略高于 FSSDet, 差距在 0.01 以内;在 -8~10 dB 信噪比场景下, MSCA-S 对信号特征参数的检测误差均低于 FSSDet。MSCA-S 的误差在 -8~4 dB 之间降低明显,可达到相差 0.022;在信噪比为 2、4 dB 时, MSCA-S 检测误差最低,中心频率误差、带宽误差、起止时间误差依次是 0.025、0.029、0.023。

综上, MSCA-S 在信噪比大于 -10 dB 时,检测的中心频率、起止时间和带宽均比其他检测方法准确。从实验结果分析,由于充分考虑了信号时频图的先验知识, MSCA-S 在检测识别宽带信号方面进行了适应性改进,比其他检测方法更适合提取宽带信号的特征,因此降低了宽带信号检测误差。

2.5.1.3 各检测方法的性能

对于处理 800×512 尺寸的信号时频图,各检测器对于参数量、计算量、深度学习模型推理时间、每秒帧数的对比如表 4 所示。

表 4 各检测方法性能对比表

检测方法	参数量/ 10^6	FLOPs/ 10^9	推理时间/ms	每秒帧数
Res-101 ^[23]	125.2	97.7	100.4	9.9
Swin-T ^[14]	109.6	69.9	80.3	12.3
FSSDet ^[16]	103.7	67.1	82.0	12.5
MSCA-S	99.1	51.8	37.7	26.6

由实验结果可知, MSCA-S 对于这 4 个指标均优于其他模型, 与对照组中表现最好的方法相比, 参数量和计算量分别比 FSSDet 降低了 4.6×10^6 和 1.53×10^{10} , 推理时间比 Swin-T 降低了 52.6 ms, 每秒帧数比 FSSDet 增多了 14.1。

实验表明, 本文提出的 MSCA-S 降低了计算量, 提高了推理速度, 在低资源情况下更有优势。

2.5.2 不同信号类型的识别效果

MSCA-S 在 4 dB 信噪比条件下表现最佳, 识别准确率达到 98.63%。图 8 所示的混淆矩阵进一步揭示 4 dB 信噪比下 MSCA-S 在不同调制信号上的识别性能。结果表明, MSCA-S 对 FM 和 0.3GMSK 信号的识别准确率极高, 几乎不存在误分类。然而, 8FSK 信号的召回率最低, 这可能是由于其频率特征与 2FSK 信号存在一定程度的重叠, 导致部分 8FSK 信号被误识别为 2FSK 信号。此外, 观察到 2FSK 的误报率最高, 暗示该类别信号特征可能与其他类别的噪声信号较为相似。因此, 后续研究将着重改进 2FSK 信号的识别性能以及区分 8FSK 和 2FSK 信号

的能力。

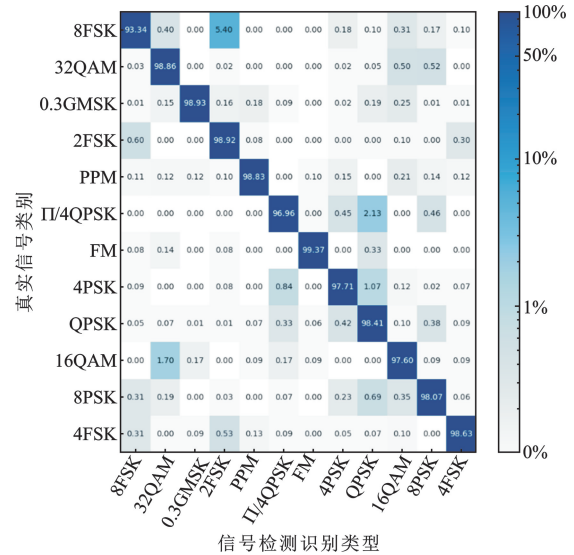


图 8 MSCA-S 识别混淆矩阵 (SNR=4 dB)

2.5.3 消融实验

在不同信噪比下设计消融实验, 验证 MSHCA 机制的有效性和通道数 C_i 、构建块数 L_i 对检测方法的影响。通过改变 MSHConvNet 中 4 个层次阶段的结构、通道数、构建块数设计 4 种网络结构, 消融实验以第三阶段为例来反映效果。

具体网络结构如表 5 所示。其中, C 为普通带有残差结构的 5×5 卷积, M 是 MSHCA 构建块。这 4 种模型在不同信噪比下的 mAP 如图 9 所示。

表 5 消融实验中特征提取骨干网络结构设计

模型名称	各阶段网络结构	$[C_1, C_2, C_3, C_4]$	$[L_1, L_2, L_3, L_4]$	设置描述
MSHConvNet-I	[M, M, C, M]	[64, 128, 320, 512]	[3, 3, 5, 3]	第三阶段 MSHCA 替换为 CNN
MSHConvNet-II	[M, M, M, M]	[64, 128, 320, 512]	[3, 3, 12, 3]	增加第三阶段构建块数
MSHConvNet-III	[M, M, M, M]	[64, 128, 256, 512]	[3, 3, 5, 3]	减少第三阶段的通道数
MSHConvNet-IV	[M, M, M, M]	[64, 128, 320, 512]	[3, 3, 5, 3]	本文默认结构

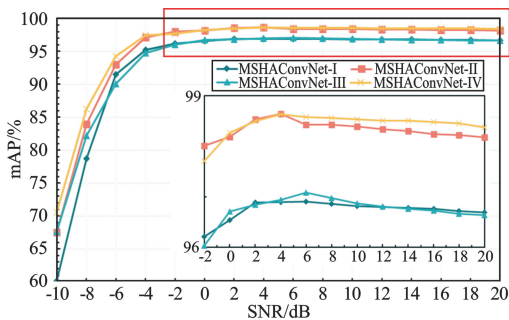


图 9 不同信噪比下消融实验的 mAP 对比

通过对比 MSHConvNet-I 与 MSHConvNet-IV 的性能可以发现, 在 4 个阶段均引入 MSHCA 机制可提升识别准确率 1.5% 以上, 尤其是在低信噪比的情况下, 提升幅度可达 10% 左右。这体现窄条形卷积相较于标准卷积在处理时频图特征上的优势。MSHCA 机制通过不同尺寸的水平卷积有效捕获信号在时间轴上的长期依赖关系, 并滤除频率轴上的无关信息, 即使在低信噪比条件下依然保持显著的性能提升。因此, 本文所提机制能够有效地提取信

号特征,从而提高 MSCA-S 识别准确率。

MSHAConvNet-III 与 MSHAConvNet-IV 的对比结果显示,适当增加通道数量可以进一步提升准确率 1% 以上。增加通道数能够在每一层中提取和表示更丰富的特征,尤其是对于复杂的信号特征,更高的特征维度有助于捕获更细微的信号差异,从而更好地分离信号不同模式,提高识别准确率。

虽然 MSHAConvNet-II 与 MSHAConvNet-IV 的整体性能相近(差异在 1% 以内),但 MSHAConvNet-IV 的结果更稳定。这表明,当网络深度和参数量达到合理的平衡时,简单堆叠更多的构建块对识别性能的提升有限,尤其是在训练数据不充足的情况下,更深的网络模型反而容易过拟合。

3 结束语

本文针对非合作频段场景中的宽带信号检测识别问题,提出了一种基于多尺度卷积注意力的稀疏检测方法。该方法结合信号时频图的物理特性,设计了多尺度水平卷积注意力机制,有效减少了模型参数量和计算量,同时保持了 Transformer 等大型模型的特征提取性能。实验结果表明,MSCA-S 在频谱范围为 2.5 MHz 的宽带环境中,无需子带切分即可实现实时检测识别,对短时突发信号也表现出良好性能。研究结果揭示了多尺度卷积注意力机制在宽带信号检测中的普遍适用性和增强远程建模能力的有效性。与已发表的研究相比,MSCA-S 在保持高精度的同时,显著减少了计算量,具有更高的实用价值。

尽管 MSCA-S 在大多数情况下表现优异,但在处理未见的新信号时,检测识别性能仍有待提高。未来研究将致力于完善无监督的新信号发现机制,使其能够在侦察过程中及时定位并预警未知信号的频率、带宽和最后检测时间等关键信息。

参考文献:

[1] 查雄,许漫坤,彭华,等. 基于深度残差网络的特定协议信号识别[J]. 电子学报,2019,47(7):1532-1537.
 [2] O'SHEA T J,ROY T,CLANCY T C. Over-the-air deep learning based radio signal classification [J]. IEEE Journal of Selected Topics in Signal Processing,2018,12(1):168-179.

[3] 耿频永. 基于深度学习的短波特定信号检测识别 [D]. 济南:山东大学,2023.
 [4] 王瑞东,王世练,张炜,等. 基于自注意力机制的干扰信号检测识别 [J]. 电讯技术,2023,63(6):790-797.
 [5] 魏新龙,李立,温驰,等. 基于模型驱动深度学习的 OTFS 检测方法 [J]. 电讯技术,2024,64(8):1181-1186.
 [6] 刘人玮,李天昀,张浩庭,等. 基于 YOLOX 的跳频信号检测 [J]. 电讯技术,2023,63(7):933-940.
 [7] 韩刚涛,马瑞鹏,吴迪,等. 基于时频图切割的宽带信号智能检测与识别 [J]. 郑州大学学报(工学版),2023,44(3):42-49.
 [8] 易冬,马瑞鹏,胡涛,等. 基于谱图分解的宽带通信信号智能检测算法 [J]. 强激光与粒子束,2023,35(6):154-162.
 [9] 殷赞,王超杰,程子恒,等. 一种基于注意力机制卷积神经网络模型的自动调制识别算法 [J]. 电波科学学报,2023,38(5):773-779.
 [10] ZHA X, PENG H, QIN X, et al. A deep learning framework for signal detection and modulation classification [J]. Sensors,2019,19(18):40-42.
 [11] YUAN Y,SUN Z H,WEI Z H, et al. DeepMorse: a deep convolutional learning method for blind morse signal detection in wideband wireless spectrum [J]. IEEE Access,2019,7:80577-80587.
 [12] LI R D, HU J H, LI S Q, et al. Blind detection of communication signals based on improved YOLO3 [C] // 2021 6th International Conference on Intelligent Computing and Signal Processing. Xi'an: IEEE, 2021: 424-429.
 [13] GUO S R, YAO J L, WU P F, et al. Blind detection of broadband signal based on weighted bi-directional feature pyramid network [J]. Sensors,2023,23(3):1-6.
 [14] 朱政宇,陈鹏飞,王梓暄,等. 基于 Swin-Transformer 的短波协议信号识别 [J]. 通信学报,2022,43(11):127-135.
 [15] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: hierarchical vision Transformer using shifted windows [C] // 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021: 9992-10002.
 [16] 耿频永,曹叶文,赵晓蕾,等. 基于频率敏感注意力的短波宽带特定信号检测 [J]. 数据采集与处理,2023,38(1):63-73.
 [17] GUO M H, LU C Z, HOU Q, et al. SegNext: rethinking convolutional attention design for semantic segmentation [J]. Advances in Neural Information Processing Systems, 2022, 35: 1140-1156.
 [18] SUN P Z, ZHANG R F, JIANG Y, et al. Sparse R-CNN: end-to-end object detection with learnable proposals [C] // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 14449-14458.

- [19] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [20] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C]//The 16th European Conference Computer Vision. Cham: Springer, 2020: 213-229.
- [21] PENG C, ZHANG X Y, YU G, et al. Large kernel matters—improve semantic segmentation by global convolutional network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 1743-1751.
- [22] O'SHEA T J, WEST N E. Radio machine learning dataset generation with GNU radio [C]//The 6th GNU Radio Conference. Colorado: GNU Radio, 2016: 1-6.
- [23] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE

Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.

作者简介:

龚安男, 1971年生于四川巴中, 2001年获硕士学位, 现为副教授、硕士生导师, 主要研究方向为大数据智能处理技术。

张静蕾 女, 2000年生于山东泰安, 2022年获学士学位, 现为硕士研究生, 主要研究方向为信号检测识别。

郭兰图 男, 1982年生于山东鱼台, 2007年获硕士学位, 现为研究员, 主要研究方向为频谱管理和复杂电磁环境建模领域技术。

赵晓蕾 男, 1992年生于山东青岛, 2018年获硕士学位, 现为工程师, 主要研究方向为智能信号处理、模式识别、机器学习等。

刘玉超 男, 1984年生于山东聊城, 2010年获硕士学位, 现为高级工程师, 主要研究方向为电磁频谱管理和复杂电磁环境建模领域技术。