

DOI: 10.13382/j.jemi.B2508725

基于多尺度特征交叉融合的自动调制识别方法^{*}

于 龙¹ 胡 悦¹ 范建华¹ 郭 辉¹ 胡锦超²

(1. 国防科技大学第六十三研究所 南京 210007; 2. 宁夏大学信息工程学院 银川 750021)

摘 要:针对现有基于深度学习的调制识别方法因特征提取和利用不充分从而导致识别准确率有限的难题,提出了一种基于多尺度特征交叉融合的自动调制识别方法。自动调制识别方法基于深度学习模型实现,首先,将 IQ 调制信号拆分为 I 信号和 Q 信号,形成模型的两路输入;接着,设计了多尺度特征提取模块,有效抑制信号噪声,在此基础上提取信号的多尺度特征用于调制方式识别;然后,设计了基于交叉注意力的特征融合模块,利用交叉注意力机制对 I 信号和 Q 信号提取的多尺度特征进行深度融合;最后,输出调制识别结果。基于调制识别领域的基准数据集 RadioML2016.10A 和 RadioML2016.10B 的实验表明,相较于现有主流的基于深度学习的识别模型,在 RadioML2016.10A 和 RadioML2016.10B 数据集上,该方法在全信噪比区间上的平均识别准确率分别为 57.17% 和 62.18%,提升了 4.92% 和 4.16% 以上,在低信噪比条件下的表现也优于现有主流识别模型,在高信噪比条件下的识别准确率分别提升了 2.74% 和 1.43% 以上。实验结果表明,所提出的方法有效提升了信号自动调制识别的准确率,克服了现有方法对噪声敏感、特征利用不充分的缺陷。

关键词: 调制方式识别;深度学习;多尺度特征;交叉注意力;信号处理

中图分类号: TN911.3 **文献标识码:** A **国家标准学科分类代码:** 510.4010

Automatic modulation recognition methodology based on multi-scale features cross fusion

Yu Long¹ Hu Yue¹ Fan Jianhua¹ Guo Hui¹ Hu Jinchao²

(1. The 63rd Research Institute, National University of Defense Technology, Nanjing 210007, China;

2. School of Information Engineering, Ningxia University, Yinchuan 750021, China)

Abstract: An automatic modulation recognition methodology based on multi-scale features cross fusion is proposed to address the problem of limited recognition accuracy caused by insufficient feature extraction and utilization in existing deep learning-based modulation recognition methods. The automatic modulation recognition methodology proposed in this paper is based on deep learning model. Firstly, the IQ modulation signal is split into I signal and Q signal to form two inputs for the model. Subsequently, a multi-scale features extraction module is designed to effectively suppress signal noise, and on this basis, multi-scale features of the signal were extracted for modulation recognition. Then, a feature fusion module based on cross attention was designed to deeply fuse the multi-scale features extracted from the I signal and Q signal using the cross attention mechanism. Finally, the modulation recognition results were output. The experiments were conducted based on the benchmark dataset RadioML2016.10A and RadioML2016.10B in the field of modulation recognition. Compared to existing mainstream deep learning-based recognition models, the average recognition accuracy of the proposed methodology in the full signal-to-noise ratio (SNR) range is 57.17% and 62.18% respectively on RadioML2016.10A and RadioML2016.10B dataset, representing an improvement of more than 4.92% and 4.16%. The performance under low SNR conditions is also better than existing mainstream recognition models, and the recognition accuracy under high SNR conditions has been improved by 2.74% and 1.43% respectively. The experimental results show that the proposed methodology effectively improves the accuracy of signal automatic modulation recognition and overcomes the shortcomings of existing methods, such as sensitivity to noise and insufficient feature utilization.

Keywords: modulation recognition; deep learning; multi-scale features; cross attention; signal processing

收稿日期: 2025-09-05 Received Date: 2025-09-05

^{*} 基金项目: 国家自然科学基金联合基金 (U24B20172)、国家自然科学基金青年基金 (52502519)、宁夏自然科学基金优秀青年项目 (2024AAC05017) 资助

0 引言

近年来,通信技术的飞速发展,尤其是5G、6G等技术的突破与应用,极大地便利了人们的生活,同时也促进了智能驾驶、低空经济、智慧医疗等领域的进一步发展^[1-2],但随之也带来了用频设备增多、频谱资源紧张等问题。认知无线电基于对电磁环境的有效感知,并结合实际的任务需求,可以实现对无线频谱资源的高效管控,能够有效提高频谱资源的利用率^[3-4]。其中,自动调制识别通过相关算法识别通信信号尤其是非合作信号的调制方式,是认知无线电的重要基础,也是完成通信信号解调、信息获取的关键前提,在无线频谱管控、信号识别、电子对抗等民用和军用领域有着广泛的应用前景,有助于提升对电磁空间的认知和理解^[5]。

现有的调制识别方法主要可以分为传统方法和学习方法,传统方法又可以分为基于似然函数(likelihood-based, LB)的方法和基于特征(feature based, FB)的方法^[6]。具体的, LB方法主要是基于统计思想,通过计算通信信号的似然函数值来实现调制方式的识别。此类方法主要依托贝叶斯概率框架,有着较为完备的理论基础,在理论上可以实现调制识别的最优解,常用的方法有平均似然比检测(average likelihood ratio test, ALRT)^[7]、混合似然比检测(hybrid likelihood ratio test, HLRT)^[8]以及广义似然比检测(generalized likelihood ratio test, GLRT)^[9]等,但此类方法大多需要知道信号的各类先验信息,如信道模型、信号均值、方差等,因此存在先验信息依赖、计算复杂度较高等问题。不同于LB方法, FB方法首先对信号的特征进行提取,进而以提取的特征为输入,利用分类算法对调制方式进行识别。提取的特征通常包括信号的时域、频域、能量域等,同时,常见的机器学习分类算法有决策树(decision tree, DT)、支持向量机(support vector machine, SVM)、K最近邻(K-nearest neighbors, KNN)、人工神经网络(artificial neural network, ANN)等^[10-12]。从原理不难看出,此类方法的准确率极度依赖提取的特征,而提取的特征主要基于人工经验设计,因此一方面比较费时,另一方面对噪声比较敏感,导致此类方法依然存在一定不足。近年来,深度学习逐渐在图像、信号等模态信息的处理上展现出优异的效果,是自动化智能化不可或缺的基础,因此也有学者将深度学习引入到信号的调制方式识别当中^[13-16]。常见的深度学习网络包括卷积神经网络(convolutional neural network, CNN)、长短时记忆网络(long short-term memory, LSTM)等。文献[15]提出了一种多级图神经网络,从通信信号和被视为图中节点的不同信号之间的相关性信息中提取特征,实现了比现有图神经网络更准确的调制方式识别;

文献[16]设计了一种基于复值卷积和频率全局滤波单元的混合神经网络,有效利用了信号不同域的特征。

将深度学习技术应用在自动调制识别领域,关键在于如何有效提取特征和高效利用提取的特征。特征提取是基于深度学习的自动调制识别技术不可缺少的部分,除了常规的空间特征或时间特征提取外,许多学者通过提取多尺度特征来进一步丰富特征信息,从而提升识别效果^[17-18]。然而,多尺度卷积的应用会在一定程度上增加模型参数量,且现有方法多存在较为复杂的数据预处理过程。此外,在特征提取的基础上,许多学者考虑引入特征融合模块来高效利用提取的特征,如引入注意力机制来融合特征^[19-20]。然而,现有方法大多会对原始信号进行维度上的变换,如转成二维图,或先进行频域等特征的提取,再利用注意力机制来进行融合,这增加了网络模型的复杂度。

综上,基于学习的调制方式识别方法,可以直接处理信号并自动提取与识别任务相关的特征,无需人工手动设计提取特征,展现出了更良好的发展前景。虽然基于学习的方法取得了一定的成功,但其性能依然有待加强,尤其是现有深度学习方法在对信号特征的提取和利用方面仍不充分,限制了识别准确率的进一步提升,导致结果仍然不能令人满意。对此,为进一步提升基于深度学习的调制方式识别方法的性能,本文从特征提取与特征融合两个维度出发,构建深度学习模型,研究基于多尺度特征交叉融合的自动调制识别方法。首先,将IQ调制信号拆分为I信号和Q信号,共两路信号,作为本文提出的深度学习网络的输入。IQ调制信号是同相(in-phase)信号(即I信号)和正交(quadrature)信号(即Q信号)的组合,广泛应用于无线通信;其次,为了实现高质量的特征提取,抑制复杂环境下的信号噪声,考虑到现有主流的小波变换方法存在的阈值选择困难等问题,设计了多尺度卷积模块,在模拟小波变换的多分辨率分析能力的同时,也更好解决了其阈值难以选取等问题,从而实现对噪声的有效抑制,并在此基础上进一步实现了信号多尺度特征的有效提取;接着,为了更加充分利用I信号和Q信号包含的不同信息,并挖掘两者之间的内在关联关系,设计了基于交叉注意力机制的特征融合模块,以融合I信号和Q信号经上述多尺度特征提取模块后得到的不同特征;最后,交叉融合后的信号特征通过全连接层和Softmax层,输出信号调制方式的识别结果。本文提出的基于多尺度特征交叉融合的自动调制识别方法,直接对原始IQ信号进行处理,利用多尺度空洞卷积,在不增加参数量的同时提升感受野范围,从而高效提取多个尺度特征,并利用交叉注意力机制有效融合分别从I信号和Q信号提取的特征,避免了复杂繁琐的数据预处理过程,克服了现有方法对噪声敏感、信号特征信息利用不充分

的局限,有效提升了调制识别的准确性。

1 方法

1.1 方法概述

本文基于深度学习理论方法,从特征提取与特征融合两个维度出发,构建了基于多尺度特征交叉融合的深度神经网络模型,以更好实现信号的调制方式识别。网络模型的整体架构如图 1 所示。网络的输入为 I 信号和 Q 信号,两路输入信号分别经过多尺度特征提取模块后,通过基于交叉注意力机制的特征融合模块进行融合,融合形成的特征先后经过全连接层和 Softmax 层,输出最终的调制识别结果。

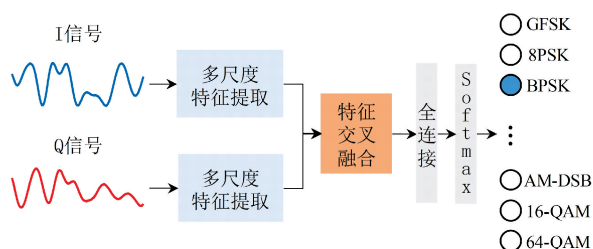


图 1 本文提出的深度学习网络架构

Fig. 1 The architecture of proposed deep learning network

1.2 多尺度特征提取模块

接收的信号中除了信号本身以外,还会夹杂着噪声或干扰等。信号和噪声的功率比值也称为信噪比,在高信噪比条件下,信号中的噪声较小,对信号的影响较低,而低信噪比条件下,信号几乎淹没在噪声中,很难挖掘出有效的信息。对此,如何在不同信噪比条件下,充分提取有效可靠的特征,是当前深度学习网络亟待解决的问题。传统的小波变换方法凭借其多分辨率分析能力,是一种有效的信号处理手段,可以较好地抑制信号中的噪声,并提取可靠的特征,然而其阈值需要手动选取,面向不同信噪比条件时自适应能力较差。对此,基于小波变换的多分辨率思想,本文设计了多尺度特征提取模块,一方面模仿小波变换的多分辨率分析能力,实现对信号噪声的抑制,另一方面,在此基础上更好地提取信号的多尺度特征信息,结构如图 2 所示。

由图 1 可知, I 信号和 Q 信号均分别接入多尺度特征提取模块,此处以 I 信号为例进行分析。由图 2 可知, I 信号的输入维度为 1×128 , 多尺度特征提取模块主要由残差块与多尺度空洞卷积两部分组成。首先,输入的 I 信号经过两个残差块结构。残差块是一种经典的深度学习网络结构,通过引入跳跃连接来解决深度学习中易出现的网络退化问题。本文中,每个残差块由两个卷积核为 1×3 的常规卷积层组成,每个卷积层后面接批量归一

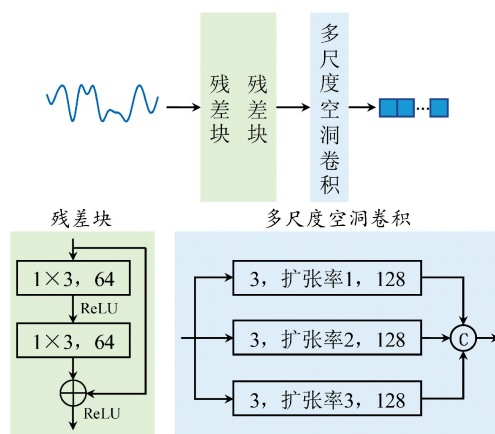


图 2 多尺度特征提取模块结构

Fig. 2 The structure of multi-scale features extraction module

化(batch normalization, BN)层和 ReLU 激活函数,最后将输出与原始输入跳跃连接。接着,经过两个残差块后输出的特征进入多尺度空洞卷积部分,多尺度空洞卷积的结构如图 2 所示,输入的特征分别经过一个卷积核为 3 的常规一维卷积(即扩张率为 1)、一个卷积核为 3、扩张率为 2 的一维空洞卷积、一个卷积核为 3、扩张率为 3 的一维空洞卷积,3 个卷积输出的结果进行通道连接。空洞卷积是一种特殊的卷积形式,可以在不增加参数的情况下,扩大感受野的范围,空洞卷积的等效卷积核大小根据下式计算:

$$k' = k + (k - 1) \times (d - 1) \quad (1)$$

式中: k 是原始的卷积核大小; d 表示扩张率; k' 表示经扩张后的等效卷积核大小。因此,本文的多尺度空洞卷积实际上由一个卷积核为 3 的常规一维卷积、一个等效卷积核为 5 的常规一维卷积以及一个等效卷积核为 7 的常规一维卷积组成,且参数量要比常规卷积更小。基于上述设置,通过多尺度空洞卷积的多分辨率特性,实现对输入信号噪声的抑制以及对多尺度特征的提取。Q 信号部分的多尺度特征提取模块网络结构与 I 信号部分的相同。此外,对于 I 信号和 Q 信号的多尺度特征提取模块,将采用权值共享的方法,在提升网络性能的同时降低网络所需的参数量。

1.3 基于交叉注意力机制的特征融合模块

本文构建的深度学习网络的输入为 IQ 调制信号,并分为了 I 信号与 Q 信号形成两路输入。实际上, I 信号与 Q 信号虽然模态相同,但共同作用于自动调制识别的结果,因此,需要考虑如何对上述提取的 I 信号和 Q 信号的多尺度特征进行有效融合。现有的融合方法,如数值相加或通道连接,都较为简单粗暴,无法深度融合并利用不同信息之间的特征。近年来,深度学习中的注意力机制可以使网络关注更加应该关注的特征或信息,更适用于

多个特征的融合。其中,交叉注意力是深度学习中一种典型的注意力机制,主要用于关注两个或多个不同序列或不同模态数据之间的关联关系,其公式如下:

$$\text{CrossAttention}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^T}{\sqrt{d}}\right)\boldsymbol{V} \quad (2)$$

式中: $\boldsymbol{Q}$ 、 $\boldsymbol{K}$ 、 $\boldsymbol{V}$ 分别代表查询向量、键向量和值向量; d 表示键向量的维度。一般而言,交叉注意力机制中的查询向量来自于一路输入或某一模态信息,键向量和值向量则来自于另一路输入或另一模态信息。虽然 I 信号和 Q 信号都属于同一种信号模态,但考虑到两者是共同作用于调制识别结果的,因此,I 信号和 Q 信号可以认为是调制方式识别任务中的两路不同输入,适用于交叉注意力机制。基于上述分析,为了更有效地融合 I 信号和 Q 信号提取的多尺度特征,充分利用两路输入信息来分析调制方式识别的情况,本文设计了基于交叉自注意力机制的特征融合模块,如图 3 所示。

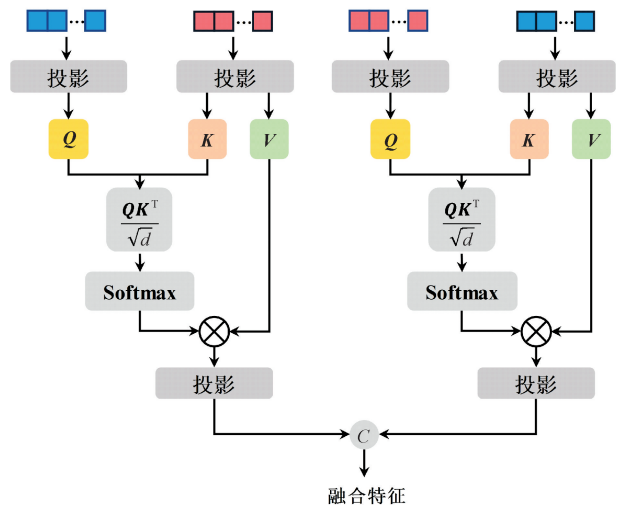


图 3 基于交叉注意力机制的特征融合模块

Fig. 3 Feature fusion module based on cross attention mechanism

考虑到在本文提出的自动调制识别方法中,用于输入的 I 信号和 Q 信号并不存在主次之分,因此,为了更好地进行特征融合,将分别通过两次交叉注意力机制得到特征融合的结果,并利用通道连接进行最终融合。一方面将 I 信号提取的特征经过投影变换后,作为交叉注意力机制中的查询向量,Q 信号提取的特征经过投影变换后,作为交叉注意力机制中的键向量和值向量;另一方面,同步将 Q 信号提取的特征经过投影变换后,作为交叉注意力机制中的查询向量,I 信号提取的特征经过投影变换后,作为交叉注意力机制中的键向量和值向量。通过上述操作,各得到一组输出特征,再利用通道连接将两组输出特征进行融合。融合后的特征经过全连接层和 Softmax 层,最终得到识别的调制方式类型。

1.4 损失函数

考虑到本文研究的调制方式识别在本质上属于多分类的问题,因此,本文方法的损失函数采用了经典的分类交叉熵(categorical cross-entropy, CCE)损失函数:

$$Loss = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{ij} \log(y_{preij}) \quad (3)$$

式中: N 代表样本数量; C 代表待分类的数量; y_{ij} 表示实际的类别结果; y_{preij} 表示网络模型预测得到的类别结果。

2 实验与结果分析

2.1 数据集与实验设置

为了验证本文提出的方法的性能,采用自动调制识别领域的基准数据集 RadioML2016. 10A 和 RadioML2016. 10B 来进行验证^[21]。RadioML2016. 10A 数据集包括 3 种模拟调制方式和 8 种数字调制方式,共计 11 种常用的调制方式,信号的信噪比范围在 -20 ~ 18 dB,间隔为 2 dB,每种调制方式、每个信噪比数据共 1 000 条,数据形式为 IQ 数据。RadioML2016. 10B 数据集在 RadioML2016. 10A 数据集的基础上移除了 AM-SSB 调制模式,保持了信噪比和样本维度的配置,并进一步扩充了数据集样本数量,数据总量达到了 1 200 000。数据集的主要参数如表 1 所示。

表 1 数据集参数

Table 1 Parameters of dataset

数据集	数据集属性	属性参数
RadioML2016. 10A	调制方式类型	8PSK、BPSK、CPFSK、GFSK、4-PAM、AM-DSB、AM-SSB、16-QAM、64-QAM、QPSK、WBFM
	信号信噪比/dB	-20 : 2 : 18
	样本数量	220 000
	样本维度	2×128
RadioML2016. 10B	调制方式类型	8PSK、BPSK、CPFSK、GFSK、4-PAM、AM-DSB、16-QAM、64-QAM、QPSK、WBFM
	信号信噪比/dB	-20 : 2 : 18
	样本数量	1 200 000
	样本维度	2×128

本文采用的计算平台的操作系统为 Ubuntu16. 04, CPU 为 i5-7300HQ,内存为 16 G,显卡为 NVIDIA GeForce RTX 3060。深度学习模型基于 Tensorflow 架构和 Keras 应用程序接口实现。为了体现本文提出方法的有效性,分别与现有的 CNN1^[22]、CNN2^[23]、ResNet^[24]、DenseNet^[25]、MCNET^[26]等多种基于深度学习的调制方式识别方法进行对比。对于每一种方法,数据集集中的训练

集、验证集与测试集均按照 6:2:2 的比例进行划分,并采用随机抽取的方法生成,且所有方法均采用相同的训练集、验证集与测试集,每一种方法均采用 Adam 方法进行优化, β_1 和 β_2 分别设置为 0.9 和 0.999。每种深度神经网络的初始学习率设置为 0.001,训练的轮数设置为 100 次,当验证精度连续 5 轮没有提升时,则将学习率变为原来的 50%,且如果连续 50 轮验证精度没有提升,则停止训练,防止出现过拟合。

2.2 结果分析

1) 评价指标

信号自动调制识别领域中常用的评价指标为识别准确率和混淆矩阵^[27],本文也将采用这两类指标对上述方法进行评价。

由于自动调制识别本质上属于分类问题,因此识别准确率 (accuracy) 指标最能体现其性能,其计算公式如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

式中: TP 表示分类正确的正类信号; FP 表示错误分类为正类的负类信号; FN 表示错误分类为负类的正类信号; TN 表示分类正确的负类信号。由于调制识别分类任务与信噪比和调制类型等参数相关,因此本文中的识别准确率为平均识别准确率。具体的,平均识别准确率是指在给定的信噪比条件下,如某一信噪比或信噪比区间,各个调制方式或所有调制方式识别准确率的平均值。

混淆矩阵用于说明不同调制方式的识别情况。混淆矩阵的纵轴表示信号调制方式的真实类别,横轴则表示不同深度学习方法预测的信号调制方式类别。因此,混淆矩阵的对角线表示本方法预测的调制识别结果是正确的,其余部分则表示预测的结果是错误的。通过混淆矩阵,可以清晰地展示调制识别方法的预测结果及不同调制方式的识别准确率。

2) 实验结果与分析

首先,对本文方法在不同信噪比下不同调制方式的平均识别准确率进行分析,图 4 和 5 所示分别为 RadioML2016. 10A 和 RadioML2016. 10B 数据集上的结果。对于 RadioML2016. 10A 数据集,在 11 类调制方式中,除了 AM-SSB 调制方式,其余调制方式的识别准确率均随着信噪比的增大而提升,如 8PSK、AM-SSB、BPSK、4-PAM、64-QAM、QPSK、CPFSK、GFSK、AM-DSB 等调制方式在高信噪比条件下识别准确率基本都可以达到 90% 以上,其中 CPFSK、GFSK、AM-DSB 等调制方式的识别准确率可以达到 100%;但是,也可以看到,16-QAM 和 WBFM 两种调制方式,即使在高信噪比条件下,识别准确率依然较低。对于 RadioML2016. 10B 数据集,可以看到,由于数据集中去除了 AM-SSB 调制方式,剩余的调制方式的识

别准确率也基本是随着信噪比的增加而提高,其中,AM-DSB、BPSK、CPFSK、GFSK 和 4-PAM 等调制方式的识别准确率在高信噪比条件下均接近或达到 100%;同时,与 RadioML2016. 10A 数据集的结果类似,16-QAM 和 WBFM 调制方式的识别效果依然不佳,这有可能是因为本文提出的深度学习网络难以有效地从 IQ 信号中提取与这两类调制方式相关的特征,从而导致在两个数据集上都没有取得较好的结果。此外,可以看到,由于数据集样本数量的提升,RadioML2016. 10B 数据集上的识别结果要好于 RadioML2016. 10A 数据集上的识别结果,这也是深度学习的特点决定的,即更多的数据可以让模型更好地学习到特征,从而更好识别调制方式。

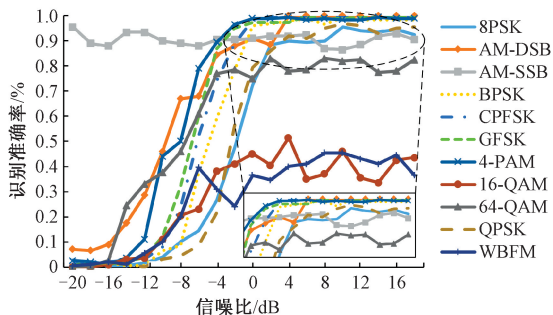


图 4 本文方法在 RadioML2016. 10A 数据集上不同信噪比条件下各调制方式识别准确率

Fig. 4 The recognition accuracy of various modulation modes using proposed method under different SNR conditions on RadioML2016. 10A dataset

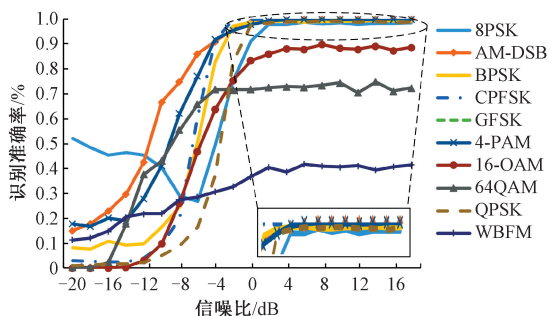


图 5 本文方法在 RadioML2016. 10B 数据集上不同信噪比条件下各调制方式识别准确率

Fig. 5 The recognition accuracy of various modulation modes using proposed method under different SNR conditions on RadioML2016. 10B dataset

图 6 和 7 所示分别为本文方法以及对比方法在不同数据集上不同信噪比条件下的平均识别准确率。可以看到,无论是在 RadioML2016. 10A 还是 RadioML2016. 10B 数据集上,本文方法在整体信噪比区间上的识别准确率

基本都优于其他方法,尤其在低信噪比区间范围内,本文方法的优势更加明显,这也表明了本文方法的特征提取更为充分,融合手段更为合理。当然,本文方法在低信噪比区间,尤其是-10 dB 以下的区间时,虽然基本优于其他方法,但并未展现出明显优势,这表明本文方法在面向低信噪比信号时,仍要考虑如何实现对噪声的进一步抑制。

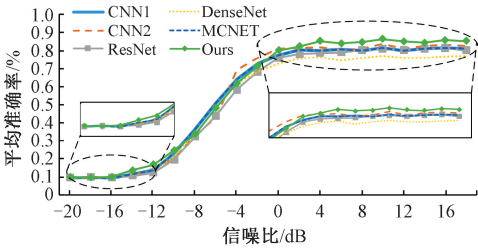


图6 不同方法在在 RadioML2016. 10A 数据集上不同信噪比条件下的识别准确率

Fig. 6 The recognition accuracy of different methods under different SNR conditions on RadioML2016. 10A dataset

表2 不同方法在 RadioML2016. 10A 数据集上不同信噪比条件下的识别准确率

Table 2 The recognition accuracy of different methods under different SNR conditions on RadioML2016. 10A dataset (%)											
方法	-20 dB	-16 dB	-12 dB	-8 dB	-4 dB	0 dB	4 dB	8 dB	12 dB	16 dB	18 dB
CNN1	9.09	9.68	14.37	37.86	65.45	76.18	77.95	77.91	78.95	78.95	79.77
CNN2	9.82	9.77	14.45	33.32	68.86	81.23	82.00	81.14	81.05	83.18	82.77
ResNet	9.59	9.59	12.27	32.22	58.23	75.91	78.59	80.36	80.41	81.41	80.14
DenseNet	9.55	9.91	13.77	34.05	61.50	72.72	76.36	75.91	76.05	76.73	76.95
MCNET	9.32	9.23	13.86	36.41	63.77	77.45	80.05	80.00	80.23	81.45	81.05
本文	9.91	10.00	16.73	33.73	61.68	80.14	85.32	84.73	85.09	85.86	85.32

表3 不同方法在 RadioML2016. 10B 数据集上不同信噪比条件下的识别准确率

Table 3 The recognition accuracy of different methods under different SNR conditions on RadioML2016. 10B dataset (%)											
方法	-20 dB	-16 dB	-12 dB	-8 dB	-4 dB	0 dB	4 dB	8 dB	12 dB	16 dB	18 dB
CNN1	10.93	11.34	18.12	41.24	66.35	80.68	83.41	83.48	83.85	82.87	83.53
CNN2	10.61	11.46	15.73	36.78	64.23	84.10	88.32	88.63	88.33	87.66	88.64
ResNet	10.61	11.60	18.76	36.81	59.38	81.53	85.53	85.96	86.18	85.78	86.23
DenseNet	10.48	11.23	16.86	37.36	63.13	79.91	84.84	85.82	85.77	85.37	85.61
MCNET	10.23	11.35	17.40	38.26	66.85	83.17	84.94	85.33	85.13	85.07	84.57
本文	11.07	12.39	21.38	40.06	70.70	87.42	89.43	89.92	89.41	89.42	89.74

识别准确率。表4 中信噪比区间包括了[-20 dB, 0 dB]、[2 dB, 18 dB]以及全信噪比区间。可以看到,首先,在[-20 dB, 0 dB]的低信噪比区间上,本文方法在 RadioML2016. 10A 数据集上的平均准确率为 34.52%,在 RadioML2016. 10B 数据集上的平均准确率为 39.74%,均优于其他对比方法;同时,在 0 dB 以上的信噪比区间上,

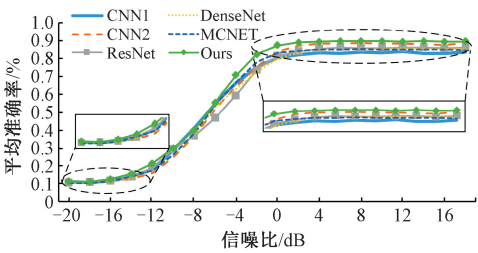


图7 不同方法在 RadioML2016. 10B 数据集上不同信噪比条件下的识别准确率

Fig. 7 The recognition accuracy of different methods under different SNR conditions on RadioML2016. 10B dataset

表2 和 3 分别为在不同数据集中,本文方法与对比方法在不同信噪比条件下的平均识别准确率,加粗代表各类方法中的最优结果。由表 2、3 可以看到,在所有的数据集上,本文方法在绝大多数信噪比条件下均取得了最优的识别结果,表明本文方法更加有效。

此外,表 4 为不同方法在不同信噪比区间上的平均

本文方法在 RadioML2016. 10A 数据集上的平均准确率为 84.84%,相较于其他对比方法,提升了 2.74%~8.70%,在 RadioML2016. 10B 数据集上的平均准确率为 89.61%,相较于其他对比方法,提升了 1.43%~6.43%;此外,在全信噪比区间上,本文方法在 RadioML2016. 10A 数据集上的平均准确率为 57.17%,相比于其他方法,平

均准确率最高提升了 4.92%,在 RadioML2016. 10B 数据集上的平均准确率为 62.18%,相较于其他方法,最高提升了 4.16%。综上,本文方法在低、高信噪比以及全信噪比区间条件下,都取得了最好的识别准确率。

表 4 不同方法在不同信噪比区间上的平均识别准确率

Table 4 The average recognition accuracy of different methods on different SNR intervals (%)				
方法	数据集	[-20 dB,0 dB]	[2 dB,18 dB]	[-20 dB,18 dB]
CNN1	RadioML2016. 10A	34.26	78.59	54.20
CNN2		34.27	82.10	55.79
ResNet		31.76	79.98	53.46
DenseNet		32.70	76.14	52.25
MCNET		34.32	80.73	55.20
本文		34.52	84.84	57.17
CNN1	RadioML2016. 10B	37.44	83.18	58.02
CNN2		36.14	88.18	59.56
ResNet		35.80	85.81	58.31
DenseNet		35.78	85.20	58.02
MCNET		37.22	85.01	58.72
本文		39.74	89.61	62.18

进一步的,本文设计了消融实验来验证提出的网络模型中不同模块的作用。Ours-Multiscale 表示模型具有多尺度特征提取模块,特征融合模块利用通道连接代替,Ours-Cross 表示模型具有交叉注意力特征融合模块,多尺度特征提取模块用标准的卷积块代替,表 5 为在不同数据集上的消融实验结果。可以看到,在 RadioML2016. 10A 数据集和 RadioML2016. 10B 数据集上,本文方法的准确率均取得了最好的结果。与 Ours-Multiscale 和 Ours-Cross 模型相比,本文方法在 RadioML2016. 10A 数据集上分别提升了 1.18% 和 5.62%,在 RadioML2016. 10B 数据集上别提升了 7.93%,表明了多尺度特征提取模块和特征融合模块的有效性。可以看到,Ours-Multiscale 模型效果不如本文的模型,表明在特征提取后,合理的特征融合可以进一步提升模型效果;同时,Ours-Multiscale 模型的效果要好于 Ours-Cross 模型,这可能是因为 Ours-Cross 模型虽然有特征融合机制,但之前的特征提取部分效果不佳,所以融合效果也相应不好。因此,精心设计特征提取和特征融合模块,才能更好地提升模型的效果。

表 5 不同数据集上的消融实验结果

Table 5 Ablation experiments on different datasets		
方法	RadioML2016. 10A	RadioML2016. 10A
Ours-Multiscale	55.99	58.90
Ours-Cross	51.55	54.25
本文	57.17	62.18

此外,图 8 和 9 所示分别为 RadioML2016. 10A 和 RadioML2016. 10B 数据集上,本文方法与各对比方法在 4 dB 时的混淆矩阵,通过混淆矩阵可以看到不同方法在不同调制方式识别上的具体情况。可以看到,在

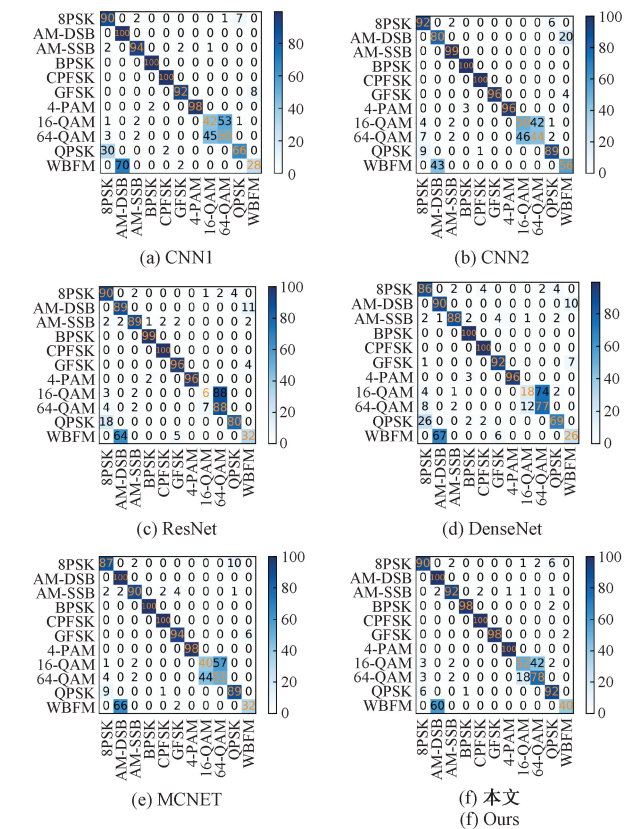


图 8 2016. 10A 数据集上不同方法在 4 dB 时混淆矩阵
Fig. 8 Confusion matrices of different methods at 4 dB on RadioML2016. 10A dataset

RadioML2016. 10A 数据集上,本文方法在 AM-DSB、AM-SSB、BPSK、CPFSK、GFSK、4-PAM、QPSK 等调制方式上均具备较好的识别效果,识别准确率均达到 90% 以上。其他方法存在较多问题,如 CNN1、ResNet、DenseNet、

MCNET 等方法在 16-QAM、64-QAM、WBFM 等调制方式上识别错误率较高, CNN2 在 16-QAM、64-QAM、WBFM 等调制方式的识别上存在不足。在 RadioML2016. 10B 数据集上, 本文方法在 8PSK、AM-DSB、BPSK、CPFSK、GFSK、4-PAM、QPSK 上均取得了接近或达到 100% 的识别准确率, 其他方法则与 RadioML2016. 10A 数据集的结果一样, 存在较多问题。与其他方法对比, 本文方法的混淆矩阵对角线更加清晰, 表明了本文方法更加可靠有效。但同时也需要看到, 本文方法在 16-QAM 和 WBFM 等调制方式的识别上尚存在不足, 存在一定的混淆现象。这也是后续的提升方向。

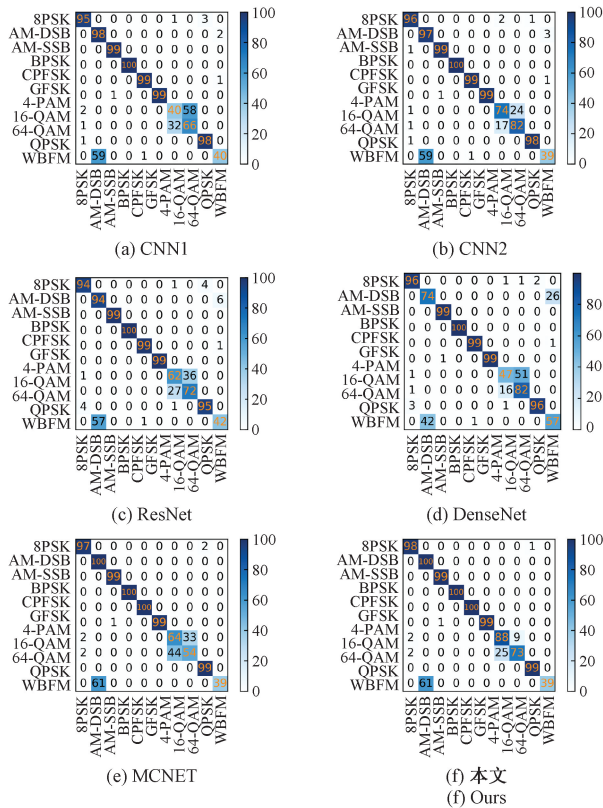


图9 2016. 10B 数据集不同方法在 4 dB 时混淆矩阵

Fig. 9 Confusion matrices of different methods at 4 dB on RadioML2016. 10B dataset

3 结论

面向自动调制识别任务, 为更加充分提取信号特征并更充分利用, 本文基于深度学习思想, 从特征提取与特征融合两个维度入手, 分别设计了多尺度特征提取模块与基于交叉注意力机制的特征融合模块, 提出了一种基于多尺度特征交叉融合的自动调制识别方法。基于调制识别领域基准数据进行了实验, 在 RadioML2016. 10A 数据集上的准确率达到 57. 17%, 相较于现有主流的深度学习

学习方法提升了 4. 92% 以上, 在 RadioML2016. 10B 数据集上的准确率达到 62. 18%, 相较于现有主流的深度学习学习方法提升了 4. 16% 以上, 并且本文方法在不同数据集上不同的信噪比区间都表现出了更加优异的性能, 证明了本文方法的有效性。此外, 如何进一步提升复杂环境下尤其是极低信噪比条件下的自动调制识别准确率, 也是未来仍要研究的方向之一。

参考文献

- [1] 彭艺, 唐剑, 杨青青, 等. 基于强化学习的应急无人机通信中继选择策略[J]. 电子测量与仪器学报, 2022, 36(7): 9-15.
- PENG Y, TANG J, YANG Q Q, et al. Relay selection strategy for emergency UAV communication based on reinforcement learning [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(7): 9-15.
- [2] 朱浩, 倪锐峰. 基于点云稀疏语义特征的智能网联汽车协同感知配准算法[J]. 仪器仪表学报, 2023, 44(10): 314-324.
- ZHU H, NI R F. A cooperative perception registration algorithm for intelligent and connected vehicles based on sparse semantic features [J]. Chinese Journal of Scientific Instrument, 2023, 44(10): 314-324.
- [3] 王诗, 严洁, 蓝浩, 等. 用户决策在异构认知网络接入中的算法研究[J]. 电子测量与仪器学报, 2024, 38(10): 224-234.
- WANG SH, YAN J, LAN H, et al. Access selection algorithm based on user-driven for heterogeneous cognitive radio networks [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(10): 224-234.
- [4] QU Y P, LU ZH L, ZENG R, et al. Enhancing automatic modulation recognition through robust global feature extraction[J]. IEEE Transactions on Vehicular Technology, 2025, 74(3): 4192-4207.
- [5] 童子滔, 张治中, 张涛, 等. 基于零样本学习和自编码器的调制信号识别研究[J]. 电子测量技术, 2024, 47(14): 1-9.
- TONG Z T, ZHANG ZH ZH, ZHANG T, et al. Zero shot learning and autoencoder based modulation signal recognition [J]. Electronic Measurement Technology, 2024, 47(14): 1-9.
- [6] 陶志勇, 闫明豪, 刘影, 等. 基于 AG-CNN 的轻量级调制识别方法[J]. 电子测量与仪器学报, 2022, 36(4): 241-249.

- TAO ZH Y, YAN M H, LIU Y, et al. Lightweight modulation recognition method based on AG-CNN[J]. *Journal of Electronic Measurement and Instrumentation*, 2022, 36(4):241-249.
- [7] MANSOURI M, BAKLOUTI R, HARKAT M F, et al. Kernel generalized likelihood ratio test for fault detection of biological systems [J]. *IEEE Transactions on NanoBioscience*, 2018, 17(4):498-506.
- [8] KAZIKLI E, DULEK B, GEZICI S. Optimal joint modulation classification and symbol decoding[J]. *IEEE Transactions on Wireless Communications*, 2019, 18(5):2623-2638.
- [9] LIU J, SUN S, LIU W. One-step persymmetric GLRT for subspace signals [J]. *IEEE Transactions on Signal Processing*, 2019, 67(14):3639-3648.
- [10] ZHANG L M, LIN CH, YAN W J, et al. Real-time OFDM signal modulation classification based on deep learning and software-defined radio [J]. *IEEE Communications Letters*, 2021, 25(9):2988-2992.
- [11] WANG Y, LIU M, YANG J, et al. Data-driven deep learning for automatic modulation recognition in cognitive radios[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(4):4074-4077.
- [12] 梁猛, 赵贝. 改进粒子群优化超限学习机的调制信号识别[J]. *西安邮电大学学报*, 2023, 28(2):57-64.
- LIANG M, ZHAO B. Modulation signal recognition based on improved particle swarm optimization extreme learning machine [J]. *Journal of Xi'an University of Posts and Telecommunications*, 2023, 28(2):57-64.
- [13] CHEN K, ZHU L ZH, CHEN S, et al. Deep residual learning in modulation recognition of radar signals using higher-order spectral distribution [J]. *Measurement*, 2021, 185:1-8.
- [14] LIANG ZH, TAO M L, XIE J, et al. A radio signal recognition approach based on complex-valued CNN and self-attention mechanism [J]. *IEEE Transactions on Cognitive Communications and Networking*, 2022, 8(3):1358-1373.
- [15] KE Y, ZHANG W CH, ZHANG Y, et al. GIGNet: A graph-in-graph neural network for automatic modulation recognition [J]. *IEEE Transactions on Vehicular Technology*, 2025, 74(6):10058-10062.
- [16] LUO Q P, ZHAO M M, CHEN Z J, et al. Complex-valued convolution and frequency global filter for automatic modulation recognition [J]. *IEEE Communications Letters*, 2023, 27(7):1779-1783.
- [17] CHEN Y T, WANG X F, LAI SH M, et al. Automatic modulation classification based on multi-scale graph neural network with LSTM [C]. 2024 5th International Conference on Machine Learning and Computer Application (ICMLCA), 2024:217-223.
- [18] 杨嘉豪, 张东坡, 何劲. 基于多尺度特征融合提取的调制样式识别算法[J]. *信号处理*, 2025, 41(3):494-503.
- YANG J H, ZHANG D P, HE J. Modulation recognition algorithm based on multi-scale feature fusion extraction[J]. *Journal of Signal Processing*, 2025, 41(3):494-503.
- [19] LI A, LI Y, ZHANG Q. A synergistic cross-domain network integrating time-frequency-spatial features for modulation recognition [J]. *IEEE Communications Letters*, 2025, 29(11):2596-2600.
- [20] 周金, 李玉芝, 张徐, 等. 复杂电磁环境下的调制识别网络[J]. *计算机应用*, 2025, 45(8):2672-2682.
- ZHOU J, LI Y ZH, ZHANG X, et al. Modulation recognition network for complex electromagnetic environments [J]. *Journal of Computer Applications*, 2025, 45(8):2672-2682.
- [21] O'SHEA T J, WEST N. Radio machine learning dataset generation with GNU radio [C]. *Proceedings of the 6th GNU Radio Conference*, 2016:1-6.
- [22] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks [C]. *Engineering Applications of Neural Networks: 17th International Conference*, 2016:213-226.
- [23] TEKBIYIK K, EKTİ A R, GÖRÇİN A, et al. Robust and fast automatic modulation classification with CNN under multipath fading channels [C]. 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring). *IEEE*, 2020:1-6.
- [24] QI P H, ZHOU X Y, ZHENG SH L, et al. Automatic modulation classification based on deep residual networks with multimodal information [J]. *IEEE Transactions on Cognitive Communications and Networking*, 2020, 7(1):21-33.
- [25] SHAIK S, KIRTHIGA S. Automatic modulation classification using DenseNet [C]. 2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP), 2021:301-305.

[26] NJOKU J N, MOROCHO-CAYAMCELA M E, LIM W. CGD-Net: Efficient hybrid deep learning model for robust automatic modulation recognition [J]. IEEE Networking Letters, 2021, 3(2):47-51.

[27] FU Y Q, MA Y, FENG ZH X, et al. Hybrid-view self-supervised framework for automatic modulation recognition[J]. IEEE Internet of Things Journal, 2025, 12(6):7360-7375.

作者简介



于龙,2003 年于解放军理工大学获得学士学位,2006 年于解放军理工大学获得硕士学位,2018 年于陆军工程大学获得博士学位,现在国防科技大学第六十三研究所高级工程师,主要研究方向为无线通信抗干扰、智能计算等。

E-mail: elong025@163.com

Yu Long received his B. Sc. degree and M. Sc. degree from the PLA University of Science and Technology in 2003 and 2006, respectively, and Ph. D. degree from Army Engineering University of PLA in 2018. He is currently a senior engineer in the Sixty-Third Research Institute, National University of Defense Technology, Nanjing. His current research interests focus on anti-jamming communication techniques, intelligent computing and so on.



胡悦(通信作者),2015 年于合肥工业大学获得学士学位,2018 年于合肥工业大学获得硕士学位,2024 年于东南大学获得博士学位,现为国防科技大学第六十三研究所助理研究员,主要研究方向为多模态信息融合、智能计算等。

E-mail: hyhfut@163.com

Hu Yue (Corresponding author) received his B. Sc. degree from Hefei University of Technology in 2015, M. Sc. degree from Hefei University of Technology in 2018, and Ph. D. degree from Southeast University in 2024. He is currently a research associate in the 63rd Research Institute, National University of Defense Technology. His main research interests include multi-sensor fusion, intelligent computing and so on.



胡锦涛超,2015 年于北方民族大学获得学士学位,2018 年于长安大学获得硕士学位,2023 年于东南大学获得博士学位,现为宁夏大学副教授、硕士生导师,主要研究方向为具身智能感知与理解。

E-mail: jinchaohu@nxu.edu.cn

Hu Jinchao received his B. Sc. degree from North Minzu University in 2015, M. Sc. degree from Chang'an University in 2018, and Ph. D. degree from Southeast University in 2023. He is currently an associate professor and M. Sc. Supervisor at Ningxia University. His main research interests include embodied AI perception and understanding.