

# 基于多尺度特征融合提取的调制样式识别算法

杨嘉豪<sup>1,2</sup> 张东坡<sup>\*2</sup> 何劲<sup>1</sup>

(1. 上海交通大学上海市北斗导航与位置服务重点实验室, 上海 200240;  
2. 中国电子科技集团公司第36研究所, 浙江嘉兴 314033)

**摘要:** 调制样式识别技术是通信信号侦察的重要组成部分, 是对未知通信信号进行分类识别、信息提取的重要前提。现有的基于深度学习调制样式识别方法在低信噪比情况下特征提取能力较差, 针对这一问题, 本文提出了一个基于多尺度特征融合提取的信号调制样式识别算法。算法利用由多个不同大小的卷积核构成的多尺度卷积模块提取信号的多尺度特征, 并通过卷积层融合特征, 提取出信号调制样式信息的关键特征, 随后通过由多头自注意力机制构成的全局特征提取模块提取信号的全局特征, 并通过平均池化层和全连接层实现调制样式分类。为了优化网络参数与运算复杂度, 本文提出利用组卷积方式代替普通卷积简化模型。实验结果表明, 在RadioML2016.10a数据集上, 所提方法可以准确识别各个调制类型, 在高信噪比下大部分调制类型的识别准确率高于95%; 相较于现有的神经网络识别方法, 所提方法的识别率提升了1.47%~7.26%, 且在较低信噪比下(-6 dB~0 dB)识别率提升了4.73%~9.09%, 体现出更好的抗噪性能; 与利用普通卷积方式构建网络相比, 采用组卷积方式可以在识别性能相差不大的情况下将网络参数量及运算量分别减小38.9%和54.9%。同时设计了消融实验验证所提算法各个模块对识别性能的提升。实验结果验证了所提算法在分类精度和抗噪性能方面的有效性。

**关键词:** 调制样式识别; 卷积神经网络; 多头注意力机制; 组卷积

**中图分类号:** TN975 **文献标识码:** A **DOI:** 10.12466/xhcl.2025.03.007

**引用格式:** 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. DOI: 10.12466/xhcl.2025.03.007.

**Reference format:** YANG Jiahao, ZHANG Dongpo, HE Jin. Modulation recognition algorithm based on multi-scale feature fusion extraction[J]. Journal of Signal Processing, 2025, 41(3): 494-503. DOI: 10.12466/xhcl.2025.03.007.

## Modulation Recognition Algorithm Based on Multi-Scale Feature Fusion Extraction

YANG Jiahao<sup>1,2</sup> ZHANG Dongpo<sup>\*2</sup> HE Jin<sup>1</sup>

(1. Shanghai Key Laboratory of Navigation and Location-based Services, Shanghai Jiao Tong University, Shanghai 200240, China;

2. The 36th Research Institute of China Electronics Technology Group Corporation, Jiaxing, Zhejiang 314033, China)

**Abstract:** Modulation recognition technology is a crucial component of communication signal reconnaissance, serving as an essential prerequisite for classifying and identifying unknown communication signals and extracting information. Existing deep learning-based modulation recognition methods have poor feature extraction capabilities under low signal-to-noise ratio (SNR) conditions. To address this issue, this study proposes a signal modulation recognition algorithm based on multi-scale feature fusion extraction. The algorithm uses a multi-scale convolution module composed of multiple convolu-

收稿日期: 2024-06-12; 修回日期: 2024-08-27

\*通信作者: 张东坡 65866151@qq.com \*Corresponding Author: ZHANG Dongpo, 65866151@qq.com

基金项目: 国家自然科学基金重点项目(61934008)

Foundation Item: The State Key Program of National Natural Science Foundation of China (61934008)

tion kernels of different sizes to extract multi-scale features of the signal and fuses these features through convolutional layers to extract key features of the signal's modulation information. Then, a global feature extraction module composed of a multi-head attention mechanism is used to extract global features of the signal, and modulation recognition is achieved through average pooling layers and fully connected layers. To optimize the network parameters and computational complexity, this study proposes replacing the standard convolution with group convolution to simplify the model. Results from experiments on the RadioML2016.10a dataset show that the proposed method can accurately identify various modulation types, with recognition accuracy exceeding 95% for most modulation types under high SNR conditions. Compared to existing neural network recognition methods, the proposed method improves the recognition rates by 1.47% to 7.26%. Under lower SNR conditions (-6 to 0 dB), it achieves an improvement of 4.73% to 9.09%, demonstrating better noise resistance. Additionally, using group convolution instead of standard convolution reduces the network parameters and computational load by 38.9% and 54.9%, respectively, with minimal performance difference. An ablation study was designed to verify the performance improvement of each module in the proposed algorithm. Experimental results validate the effectiveness of the proposed algorithm in terms of recognition accuracy and noise resistance.

**Key words:** modulation recognition; convolutional neural network; multi-head attention mechanism; group convolution

## 1 引言

通信信号侦查作为电子战的重要组成部分,在现代军事与信息化环境中发挥着关键作用。随着通信和网络技术的发展,电磁环境日益复杂,如何对未知通信信号进行分类识别、信息提取是未来电子战装备需要应对的主要挑战之一。调制样式识别是开展解调、译码及信息提取的重要前提,也是通信信号侦查的重要组成部分,被广泛应用在情报侦查、电子干扰等领域,具有重要的军事价值与研究意义。调制样式识别方法主要分为基于似然(Likelihood-Based, LB)方法<sup>[1]</sup>与基于特征提取(Feature-Based, FB)方法<sup>[2-4]</sup>两大类。基于似然法依赖信号和信道的先验知识构建似然函数,计算复杂且在复杂环境下识别正确率会迅速下降。而基于特征法则需要人工提取的特征,通过提取信号的高阶矩<sup>[5]</sup>、循环谱<sup>[6]</sup>、循环统计量<sup>[7]</sup>等特征,再利用前馈神经网络、支持向量机<sup>[5]</sup>等分类器进行分类。基于似然法虽然运算简单但需要大量人工经验提取特征,且对具有不同传输信道和附加噪声的复杂场景适应能力差,同时在少样本情况下无法有效提取合适的特征用于调制识别。

近年来,随着深度学习技术的崛起,其在数据处理能力上展现出了强大的优势<sup>[8]</sup>。在通信领域,深度学习技术被引入到调制识别问题中,取得了令人瞩目的成就。O'SHEA等人<sup>[9-10]</sup>首次引入了一种基于卷积神经网络(Convolutional Neural Networks, CNN)的自动调制识别模型,相比于传统仅依赖专家经验手工提取特征的方法,获得了更高的识别性能。随着深度学习的发展,越来越多的研究

人员尝试用不同的深度学习方法进行信号调制样式识别研究。文献[11]通过CNN对信号提取特征,同时修改卷积核的维数提高调制识别的准确率。文献[12]将调制识别网络从CNN架构转移到循环神经网络(Recurrent Neural Network, RNN)架构,提出基于长短时记忆(Long Short-Term Memory, LSTM)网络调制识别模型。结果表明,基于LSTM的模型可以准确地捕获信号不同持续时间内时序序列之间的关系,提高调制识别精度。文献[13]提出联合CNN架构和RNN架构网络模型,利用CNN与LSTM分别对信号提取特征实现对数字调制信号的分类,相比于单流提取特征方式获得了更高的识别精度。CNN网络提取信号序列的空间特征,RNN网络提取信号序列中的时序关系,但两者都在信号局部进行特征提取,无法实现的全局信息交互,导致识别正确率无法进一步提高。

文献[14]提出基于Transformer模型的调制识别网络模型,对信号进行切片后输入模型进行特征提取;文献[15]通过设计帧级嵌入模块和优化帧长度,增强了Transformer的全局建模能力,同时引入双分支门线性单元方案,减小模型大小并提升性能。Transformer模型主要通过注意力机制<sup>[16]</sup>利用信号中长距离相关性实现对信号的全局特征提取,但其缺乏提取局部特征的能力,调制识别效果也无法进一步提高。

上述网络模型在高信噪比下表现出良好的特征提取能力和较高的识别正确率,然而在低信噪比下提取信号特征能力不足,识别正确率不高。此外,传统深度神经网络模型结构复杂,难以在识别精度和运算量之间取得平衡,这些因素限制了其在

特定场景下的应用。本文结合 CNN 提取局部特征和注意力机制提取全局特征的优点,提出一个端对端的多尺度特征融合提取(Multi-scale Feature Fusion Extraction, MFFE)网络模型。模型通过多个不同大小的卷积核对原始信号提取不同尺度的局部特征<sup>[17]</sup>,并通过三层卷积层对局部特征进行融合,随后基于多头注意力(Multi-Head Attention, MHA)机制<sup>[16]</sup>设计了全局特征提取模块,通过该模块进行全局特征提取,增强在低信噪比下对信号的表征能力,提高识别正确率。同时,采用组卷积(Group Convolution)代替普通卷积以简化模型,减少网络参数和运算复杂度。所提模型可以有效提取信号的局部信息和全局信息,提高调制识别的正确率,同时平衡模型所需的参数及计算量。

## 2 问题描述

调制识别系统结构如图1所示,信号接收端接收或截获到信号,经采样后获得信号离散采样序列,对采样数据进行预处理后,经过调制识别网络对信号进行特征提取并获得调制分类。本文考虑单通道情况下的信号接收过程,接收信号一般表示为:

$$r(t) = x(t) * h(t) + n(t) \quad (1)$$

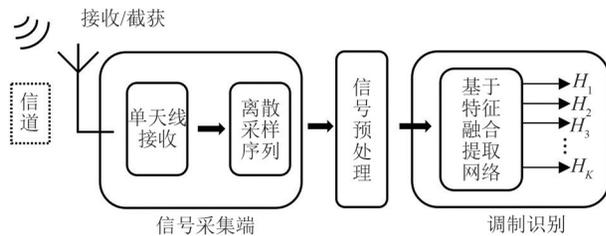


图1 调制识别系统结构图

Fig. 1 Architecture diagram of the modulation recognition system

式中, $r(t)$ 表示接收信号, $x(t)$ 表示发射信号, $h(t)$ 表示信道响应, $n(t)$ 表示噪声,\*表示卷积操作。发射信号 $x(t)$ 表示为:

$$x(t) = A(t) \cos(2\pi f_c t + \phi(t)) \quad (2)$$

其中 $f_c$ 为载波频率, $A(t)$ 为对载波的幅度调制, $\phi(t)$ 为对载波的相位调制。实际中通常采用IQ调制来提高频谱利用率,任意幅度、相位调制过程都可以看作是对两个正交分量的幅度调制,基带调制信号为:

$$z(t) = I(t) + jQ(t) \quad (3)$$

实际应用中,需要对接收信号进行采样并进行下变频和滤波等预处理,获得离散的I、Q序列。

## 3 基于MFFE的调制样式分类方法

调制识别任务可以分为两个部分:特征提取和分类。前者是建立样本空间到特征空间的映射,以提取不同调制类别间的特征;后者是建立特征空间到调制类别空间的映射,实现调制类型的识别。调制识别任务的关键是构建一个映射,获得能良好表征不同调制类型的特征,以提高识别准确率。本文提出基于MFFE网络结构提取信号的完备特征并实现调制类型分类,网络主要由多尺度卷积(Multi-Scale Convolution, MSC)模块、CNN特征融合模块和全局特征提取(Global Feature Extraction, GFE)模块组成,网络结构如图2所示。首先,通过多尺度卷积模块从原始IQ数据提取信号的多尺度特征。随后,经过CNN特征融合模块将这些多尺度特征进行融合,并通过全局特征提取模块提取信号的全局特征。最后,对全局特征进行平均池化(Global Average Pooling, GAP)操作,通过全连接层并输入Softmax函数,获取分类结果。

### 3.1 信号多尺度特征提取

首先通过多尺度卷积模块提取信号不同尺度下的局部特征,多尺度卷积模块特征提取示意图如图3所示。不同调制类型的信号可能在不同频率范围内表现出不同的特征,如高频信号可能包含快速变化的细节,低频信号需要更关注整体结构。为了有效地识别不同类型的信号,通过不同大小的卷积核对信号进行特征提取获得信号不同尺度上的特征信息。通过卷积提取信号的特征,每个尺度下特征包括四层卷积层,每层卷积层通过ReLU(Rectified Linear Unit)函数激活。多尺度卷积模块采用三个不同大小的卷积核对原始数据进行卷积提取特征,卷积核大小分别为3、7、11。每个尺度提取特征时首先将信号进行二维卷积提取信号的IQ路特征,随后通过三个一维卷积层提取信号局部特征,每层卷积输出通道数设置为48。

### 3.2 基于组卷积的网络优化

大卷积核比小卷积核需要更多的参数与计算量,使得网络变得复杂。为了保证网络性能同时平衡网络的复杂度,本文采取组卷积的方式来优化网络。组卷积将输入特征分成多个组,并对每个组进行独立的卷积操作,最后将各组的结果合并起来。对于一个卷积层,输入张量 $X$ 的大小为 $[c_{in} \ h_{in} \ w_{in}]$ ,期望输出张量 $Y$ 的大小为 $[c_{out} \ h_{out} \ w_{out}]$ ,则卷积核张量的大小为 $[c_{in} \ c_{out} \ k_h \ k_w]$ 。为方便理解,假设输入输出张量的通道数都为4,同时令 $X = [x_1, x_2, x_3, x_4]$ ,

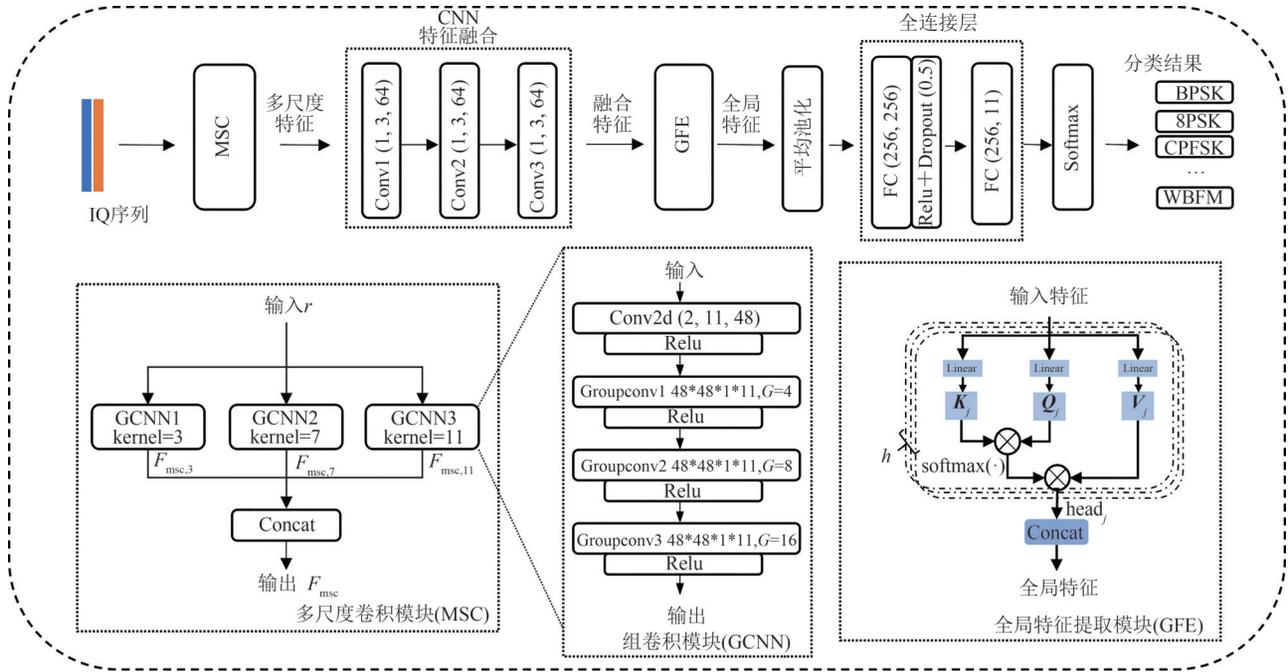


图2 MFFE网络结构图

Fig. 2 Network structure of multi-scale feature fusion extraction

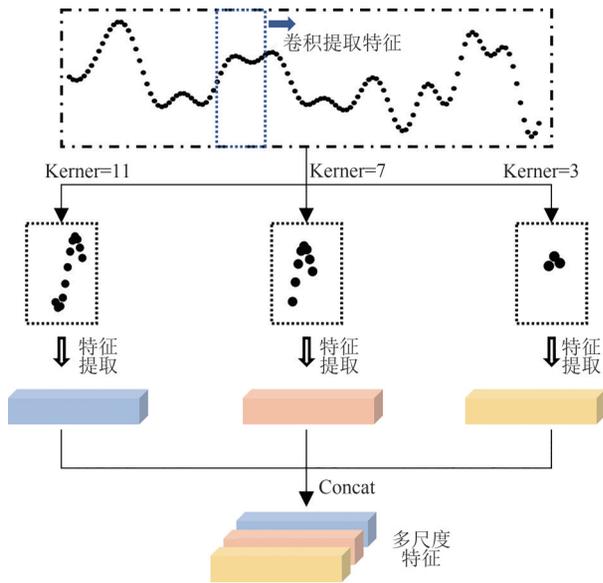


图3 多尺度卷积模块特征提取示意图

Fig. 3 Schematic of feature extraction in the multi-scale convolution module

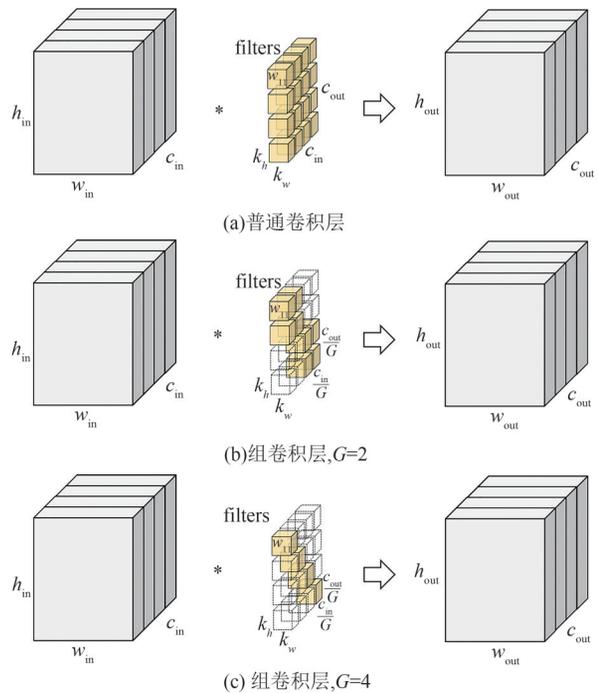


图4 分组卷积层与普通卷积层对比示意图

Fig. 4 Comparison between group convolution layer and convolution layer

$x_i \in \mathbb{R}^{h_{in} \times w_{in}}$  及  $Y=[y_1, y_2, y_3, y_4], y_i \in \mathbb{R}^{h_{out} \times w_{out}}$ , 卷积示意图如图4(a)所示,对于普通卷积输出  $Y$  可以表示为:

$$\begin{aligned}
 y_1 &= x_1 * w_{11} + x_2 * w_{12} + x_3 * w_{13} + x_4 * w_{14} \\
 y_2 &= x_1 * w_{21} + x_2 * w_{22} + x_3 * w_{23} + x_4 * w_{24} \\
 y_3 &= x_1 * w_{31} + x_2 * w_{32} + x_3 * w_{33} + x_4 * w_{34} \\
 y_4 &= x_1 * w_{41} + x_2 * w_{42} + x_3 * w_{43} + x_4 * w_{44}
 \end{aligned} \quad (4)$$

式中,  $*$  表示卷积神经网络中的卷积操作。对于图4(b)和(c)中,分别进行组卷积操作,根据输入张量和输出张量的通道数分别为2组和4组。当

$G=2$ 时,组卷积输出 $Y$ 可以表示为:

$$\begin{aligned} y_1 &= x1 * w_{11} + x2 * w_{12} \\ y_2 &= x1 * w_{21} + x2 * w_{22} \\ y_3 &= x3 * w_{33} + x4 * w_{34} \\ y_4 &= x3 * w_{43} + x4 * w_{44} \end{aligned} \quad (5)$$

当 $G=4$ 时,组卷积输出 $Y$ 可以表示为:

$$\begin{aligned} y_1 &= x1 * w_{11} \\ y_2 &= x2 * w_{22} \\ y_3 &= x3 * w_{33} \\ y_4 &= x4 * w_{44} \end{aligned} \quad (6)$$

在相同的输入输出大小下,组卷积所需的卷积核个数与运算量为普通卷积的 $1/G$ 。本文通过组卷积的方式减小网络的参数大小与运算复杂度,在提取多尺度特征中,对原始信号进行二维卷积操作提取信号的IQ路特征后,分别利用 $G=4$ 、 $G=8$ 、 $G=16$ 的组卷积代替普通卷积进行特征提取,通过逐层增加 $G$ 值,逐步降低计算复杂度和参数数量,同时保证在较浅层特征提取中保持较高的特征表示能力,而在更深层中逐步减少计算负担。将不同尺度下的特征连接起来,得到多尺度特征:

$$F_{\text{msc}} = \text{Concat}(F_{\text{msc},3}, F_{\text{msc},7}, F_{\text{msc},11}) \quad (7)$$

其中 $F_{\text{msc},3}$ 、 $F_{\text{msc},7}$ 、 $F_{\text{msc},11}$ 分别表示通过3、7、11大小的卷积核获得的局部特征。对多尺度特征通过三个卷积层,进一步处理和组合不同尺度的特征,使得模型能够综合利用多尺度信息,得到融合特征,每层卷积输出通道数设置为64。

### 3.3 基于多头注意力机制的全局特征提取

利用全局特征提取模块从融合特征中提取全局关键特征,本文采用多头自注意力机制进行全局特征提取。MHA机制能够在整个信号序列范围内整合信息,丰富特征表示的多样性和适应性,从而提升模型在复杂信号处理中的灵活性和准确性。同时MHA通过并行计算显著缩短推理时间,提升模型的计算效率。

MHA机制对输入特征进行不同的线性变换,生成查询向量(Query Vector) $\mathbf{Q} \in \mathbb{R}^{d_q \times d_q}$ 、键向量(Key Vector) $\mathbf{K} \in \mathbb{R}^{d_k \times d_k}$ 和值向量(Value Vector) $\mathbf{V} \in \mathbb{R}^{d_v \times d_v}$ , $d_q$ 、 $d_k$ 和 $d_v$ 分别代表查询向量的维度、键向量的维度和值向量的维度,这三个向量包含了原始特征的全部信息。令 $d_q = d_k = d_v = \frac{D}{h}$ , $h$ 为头的个数,本文设置为4。在形式上,上述过程可以描述为:

$$\begin{cases} \mathbf{Q} = FW_q \\ \mathbf{K} = FW_k \\ \mathbf{V} = FW_v \end{cases} \quad (8)$$

其中, $F$ 表示输入特征, $W_q$ 、 $W_k$ 和 $W_v$ 是可学习的参数,用于进行线性变换的权重矩阵。通过使用查询向量 $\mathbf{Q}$ 和键向量 $\mathbf{K}$ 之间的相关性,计算包含全局信息的自注意力矩阵映射。随后,将自注意力矩阵映射与相应的值向量 $\mathbf{V}$ 加权,获得输出矩阵:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (9)$$

式中, $\sqrt{d_k}$ 为缩放因子。对查询向量 $\mathbf{Q}$ 、键向量 $\mathbf{K}$ 和值向量 $\mathbf{V}$ 分别进行独立的投影,每个向量进行 $h$ 次投影,每次使用不同的参数。将这些投影的输出连接在一起,获得最终的输出:

$$\text{head}_j = \text{Attention}(\mathbf{Q}_j, \mathbf{K}_j, \mathbf{V}_j) \quad (10)$$

$$\text{MultiHead} = \text{Concat}(\text{head}_1, \dots, \text{head}_h) \quad (11)$$

通过多头注意力机制,网络可以从不同的子空间中提取特征,并在全局范围内整合信息,增强对信号的特征能力。

### 3.4 分类及训练

将式(11)提取的全局特征输入到平均池化层中对每个维度进行平均池化得到一维数据。平均池化操作将全局特征的每个维度进行平均,平滑特征并降低特征维度,减少模型所需的计算复杂度和参数量,同时防止模型过拟合。对平均池化后的一维数据通过全连接方式连接到分类器,包括两个全连接层,第一个全连接层通过Relu函数激活,同时通过dropout函数随机令50%的神经元失活,第二个全连接层通过Softmax函数进行分类输出,其可以表示为:

$$P(y = k|\mathbf{r}) = \frac{\exp(z_k)}{\sum_{i=1}^K \exp(z_i)} \quad (12)$$

式中, $K$ 表示分类总数, $P(y = k|\mathbf{r})$ 表示第 $k$ 类的预测概率, $z$ 表示全连接层输出向量。Softmax函数将模型的原始输出转换为表示概率分布的形式,使得模型能够输出每个类别的概率值。计算出不同类别的概率后,利用交叉熵损失函数计算预测误差,其可以表示为:

$$\mathcal{L}_{\text{ce}}(\mathbf{W}) = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K y_{i,k} \log P(y = k|\mathbf{r}_i) \quad (13)$$

其中, $y_{i,k}$ 为信号调制类型的标签, $\mathbf{W}$ 为网络模型的参数。当损失函数最小时网络训练完成,利用最终

训练完成的网络参数预测调制类型：

$$\hat{y} = \arg \max_k P(y = k | \mathbf{r}) \quad (14)$$

## 4 实验结果与分析

### 4.1 数据集选择及训练参数设置

本文采用 RadioML2016.10a 数据集<sup>[18]</sup>，该数据集共包含 8 种数字调制以及 3 种模拟调制，每个信号包括 I、Q 两路数据。为模拟实际信号环境，数据集考虑了频偏、时钟抖动等因素，使用选择性衰落模型 (Selective Fading Model) 来模拟莱斯信道 (Rician) 和瑞利信道 (Rayleigh) 的衰落过程，同时还包含了加性高斯白噪声。实验数据详细参数如表 1 所示。

实验中根据不同的调制类型和信噪比将数据集划分为训练集 (60%)、验证集 (20%) 和测试集 (20%)。训练过程中，设置批次大小 (batch size) 为 256，初始学习率为 0.001。如果在 10 轮内验证集损失函数没有下降，将学习率衰减为原来的 0.4。训练过程以验证集损失函数为优化目标，总训练次数为 200 次。当训练集损失函数达到最小时保存模型。

### 4.2 结果分析

各个调制类型在 -8 dB、-4 dB、0 dB 及 10 dB 信噪比下的混淆矩阵，如图 5 所示。在 -8 dB 信噪比条件下，整体识别率较低，在 -4 dB 信噪比下各个类型的信号识别率开始上升，但 WBFM、QPSK 和 8PSK 调制类型的识别率偏低，表明噪声对信号特征干扰较大。随着信噪比的提高，识别性能显著改善，在 0 dB 时，绝大部分调制类型的识别率均超过 90%，在 10 dB 时，多数调制类型的识别率超过 95%，但 WBFM 调制识别正确率依旧没有改善。从不同信

噪比下的混淆矩阵分析，本文所提模型随着信噪比的增加，识别正确率逐渐提升，在高信噪比下表现出高精度识别能力，表明本文所提模型可以有效识别信号调制样式。

分析 WBFM 信号误判的原因，在数据集中 AM-DSB 信号和 WBFM 信号为模拟语音信号采集，在采集中存在静音期。静音期内的信号中无传输信息，两类信号仅含载波信息，使得模型难以从中提取有效特征，导致 WBFM 信号在静音期的识别率下降。为进一步提升算法识别效果，对网络模型分类结果中误判信号通过专家特征进行分类判决。分析混淆矩阵，模型预测为 AM-DSB 的信号主要包括非静音期 AM-DSB 信号和静音期的 AM-DSB、WBFM 信号，静音期信号不传输信息且不包含信号的调制信息，需要将静音期信号剔除。非静音期 AM-DSB 信号为非恒包络调制信号，静音期信号包络恒定，利用零中心归一化瞬时幅度谱密度的最大值  $\gamma_{\max}$  并设置阈值进行恒包络调制和非恒包络调制的判断，可以表示为：

$$\gamma_{\max} = \max | \text{DFT}(A_n(i) - 1) |^2 \quad (15)$$

其中， $A_n(i) = A(i)/m_a$ ， $m_a = \frac{1}{N} \sum_{i=1}^N A(i)$  为瞬时幅度  $A(i)$  的平均值。通过设置阈值  $\gamma_{th}$ ，可以区分非恒包络的非静音期 AM-DSB 信号和恒包络的静音期信号。对模型中预测为 AM-DSB 标签的信号进行静音期信号剔除，18 dB 下混淆矩阵如图 6 所示。

从混淆矩阵中可以发现，通过专家特征进行分类判决后，AM-DSB 标签中仅保留非静音期信号，静音期信号被剔除，混淆矩阵中两类信号不再混淆，改善了算法识别性能。

表 1 实验数据集详细参数

Tab. 1 Parameters of the experimental dataset

| 参数      | 描述                                                               |
|---------|------------------------------------------------------------------|
| 调制样式    | BPSK、QPSK、8PSK、PAM4<br>QAM64、AM-DSB、AM-SSB、CPFSK、GFSK、WBFM、QAM16 |
| 样本数     | 220000                                                           |
| 采样率     | 200 kHz                                                          |
| 信号长度    | 128                                                              |
| 每符号采样数  | 8                                                                |
| 信噪比     | -20 dB~18 dB                                                     |
| 最大载波偏移  | 500 Hz                                                           |
| 最大采样率偏移 | 50 Hz                                                            |

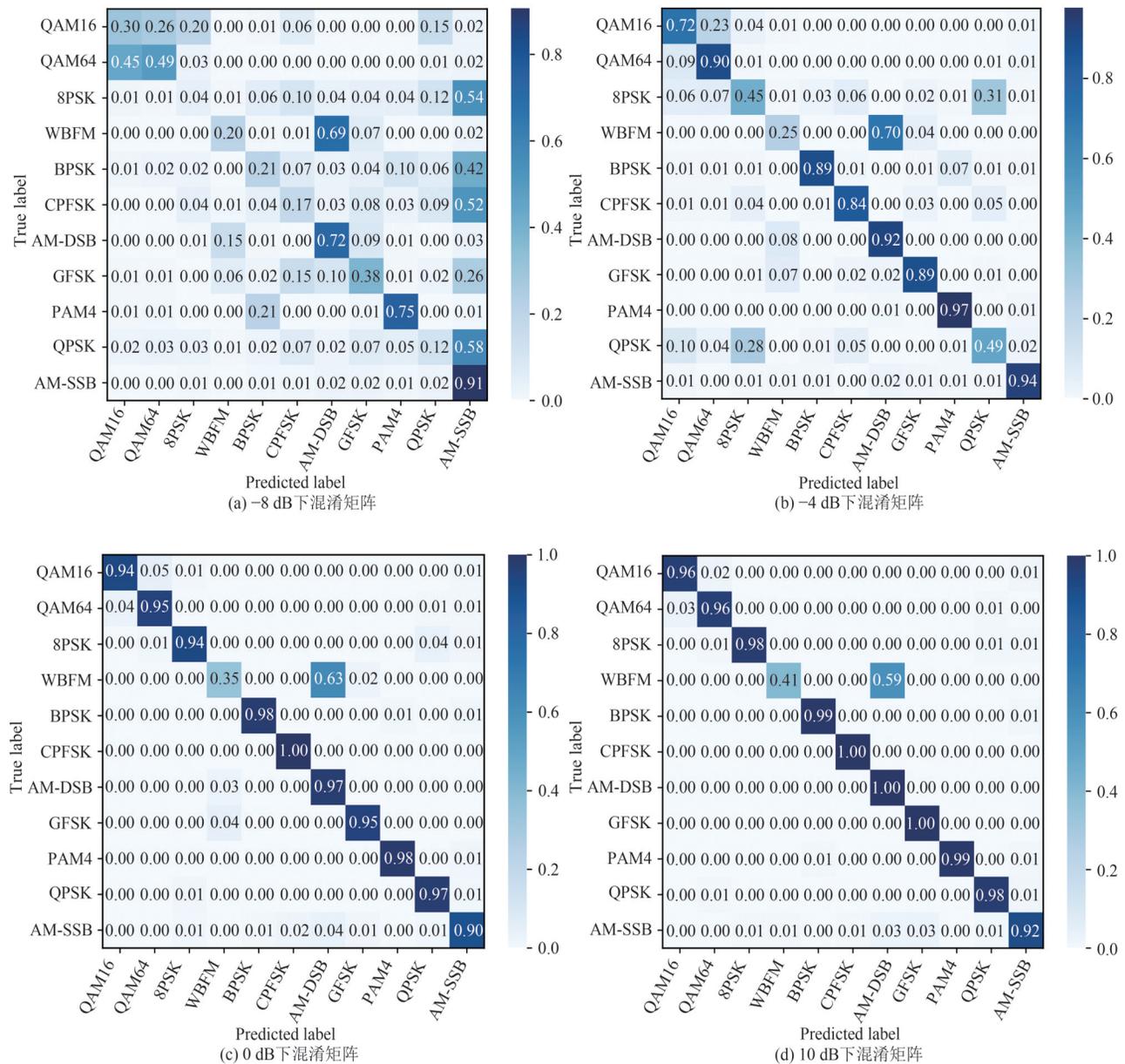


图5 -8 dB、-4 dB、0 dB及10 dB下混淆矩阵

Fig. 5 Confusion matrices at -8, -4, 0, and 10 dB

将所提网络模型与其他调制识别网络进行比较,包括由CNN网络模型(CNN<sup>[7]</sup>),LSTM网络模型(LSTM<sup>[12]</sup>),CNN与LSTM混合模型(MCLDNN<sup>[19]</sup>、PET-CGDNN<sup>[20]</sup>)以及基于Transformer结构的网络(MCformer<sup>[14]</sup>)。仿真实验结果如图7所示,在信噪比低于-8 dB时,所有模型的识别准确率都较低,但随着信噪比的提高,各个模型的识别准确率逐渐上升。在-6 dB到0 dB的范围内,本文所提模型的识别率明显高于其他模型,表现出更好的抗噪能力。当信噪比大于2 dB时,所有

模型的识别准确率都显著提高,并达到相对稳定的高识别率。图中插入的放大图显示了在高信噪比条件下不同模型的识别准确率差异。在信噪比高于2 dB时,本文所提模型的识别准确率始终保持在较高水平,并在大多数情况下略高于其他模型。同时,本文所提模型在高信噪比范围内的识别准确率始终高于92%,并且波动较小,表现出更好的识别性能。

表2给出了所提网络模型与其他网络模型具体性能。所提模型的整体平均识别率为63.34%,相比

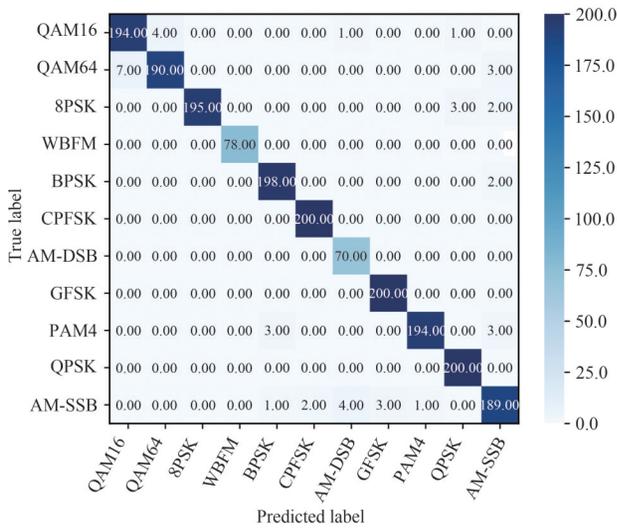


图 6 18 dB 下混淆矩阵

Fig. 6 Confusion matrices at 18 dB

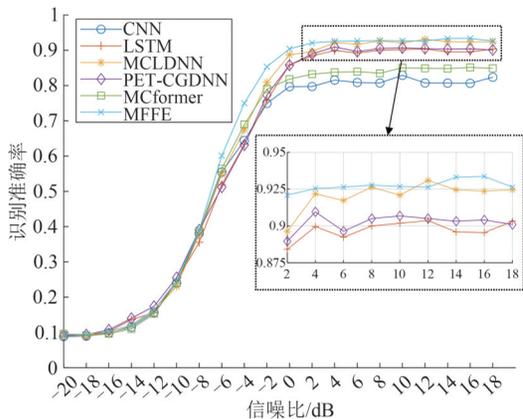


图 7 不同网络性能比较

Fig. 7 Comparison of performance among different networks

其他模型有明显提升,相比于 CNN 模型的 56.08% 高出 7.26%, 相比 MCLDNN 模型的 61.87% 高出 1.47%, 表明所提网络在不同信噪比下的识别性能

优于其他模型。在信噪比大于 2 dB 的高信噪比条件下,所提模型的平均识别率达到 92.74%,略高于 MCLDNN 的 92.07%,并且明显高于其他模型,如 CNN 的 81.13% 和 MCformer 的 84.36%,表明在较高信噪比环境下,所提模型可以保持较高的识别精度,且在所有对比模型中表现最佳。在 -6 dB 到 0 dB 的信噪比范围内,所提模型的平均识别率为 77.72%,同时 MCLDNN、MCformer 和 CNN 模型在该信噪比范围内的平均识别率分别为 72.99%、71.55% 和 68.63%,虽然所有模型在低信噪比下的识别率都有明显下降,但所提网络识别率对比模型高出 4.73%~9.09%,具有显著的性能提升,表明所提模型在较低信噪比条件下具有更好的抗噪能力和识别性能,体现出所提模型对噪声具有鲁棒性。

4.3 消融实验

上述实验表明本文所提模型具有优异的识别性能,为了分析本文所提各个特征处理模块的有效性,通过消融实验来分析模型中 MSC 模块和 GFE 模块的有效性。

对所提模型分别移除 MSC 模块、CNN 特征融合模块和 GFE 模块,识别结果如表 3 所示。通过对比可以发现在移除 MSC 模块后网络的识别精度下降了 3.33%,这表明本文所提出的 MSC 模块可以充分地捕获信号中不同频率范围内表现出的特征,这些特征可以反映出不同调制类型的差异,提升模型调制识别的性能;对比移除 CNN 特征融合模块后网络识别精度下降了 1.84%,表明 CNN 特征融合模块可以进一步处理和组合不同尺度的特征,使得模型能够综合利用多尺度信息,提高整体识别性能;对比移除 GFE 模块后网络识别精度下降了 1.43%,这表明本文提出的 GFE 模块可以能够同时关注输入特征的不同部分,在信号的全局范围内提取有用特征,提升模型的识别性能。

表 2 所提网络模型与其他网络模型识别性能

Tab. 2 Recognition performance of the proposed network model compared to other network models

| 模型        | -6 dB~0 dB 平均正确率/% | 2 dB 以上平均正确率/% | 平均正确率/% |
|-----------|--------------------|----------------|---------|
| CNN       | 68.63              | 81.13          | 56.08   |
| LSTM      | 69.22              | 89.73          | 60.11   |
| MCLDNN    | 72.99              | 92.07          | 61.87   |
| PET-CGDNN | 69.35              | 90.23          | 60.74   |
| MCformer  | 71.55              | 84.36          | 58.14   |
| MFFE      | 77.72              | 92.74          | 63.34   |

表3 消融实验结果

Tab. 3 Results of ablation experiments

| 模型          | 平均识别率/% | 最高识别率/% | 0 dB以上平均正确率/% |
|-------------|---------|---------|---------------|
| 移除MSC模块     | 60.01   | 90.45   | 89.03         |
| 移除CNN特征融合模块 | 61.50   | 91.55   | 90.62         |
| 移除GFE模块     | 61.91   | 91.86   | 91.12         |
| MFFE        | 63.34   | 93.36   | 92.51         |

为了对比所提出的组卷积层对网络参数和运算量的优化效果,进行MSC模块中组卷积层与普通卷积层替换的实验对比,并分析组卷积层模型在简化网络参数和加快运算方面的优化效果。从表4的对比结果可以看出,与其他网络模型所需资源相比,本文所需的参数和运算量要求较大,原因是为了提升模型性能引入了大的卷积核,增加了所需的存储和运算资源。当采用组卷积方式进行网络优化时,网络模型的参数量和浮点运算量分别减少了38.9%和54.9%,与此同时,优化前后网络的平均识别率基本保持不变。由此可见,本文所采用的组卷积层方式能够有效减小网络的参数量和计算复杂度,令模型可以同时兼顾识别的性能和占用的资源。

表4 组卷积层与普通卷积层对比结果

Tab. 4 Performance results of group convolutional layer versus convolutional layer

| 模型        | 参数量/K   | 浮点运算量/M | 平均正确率/% |
|-----------|---------|---------|---------|
| CNN       | 1592.38 | 13.01   | 56.08   |
| LSTM      | 201.10  | 25.82   | 60.11   |
| MCLDNN    | 406.20  | 49.15   | 61.87   |
| PET-CGDNN | 71.87   | 8.32    | 60.74   |
| MCformer  | 47.98   | 48.64   | 58.14   |
| MFFE-CNN  | 318.99  | 28.92   | 63.41   |
| MFFE-GCNN | 195.00  | 13.05   | 63.34   |

## 5 结论

本文提出了一个基于多尺度特征融合提取网络模型用于通信信号调制样式智能识别,所提出的网络模型可以提取出信号的整体特征与快速变化特征,并通过多头自注意力机制提取信号的全局特征,最终实现调制样式的分类。实验结果表明,在RadioML2016.10a数据集上,所提算法在高信噪比条件下均表现出优异的识别性能;在-6 dB至0 dB的低信噪比范围内表现出更好的抗噪性能,识别率

相较于现有方法提升了4.73%至9.09%。同时在网络中利用组卷积的形式优化网络参数与运算量,简化了网络模型。本文提出的算法不仅提高了调制样式识别的精度,而且在高噪声环境下表现出较强的鲁棒性,显示出在实际通信信号侦查应用中的潜力。

## 参考文献

- [1] HAMEED F, DOBRE O A, POPESCU D C. On the likelihood-based approach to modulation classification[J]. IEEE Transactions on Wireless Communications, 2009, 8(12): 5884-5892.
- [2] MOSER E, MORAN M K, HILLEN E, et al. Automatic modulation classification via instantaneous features[C]//2015 National Aerospace and Electronics Conference (NAECON). Dayton, OH, USA. IEEE, 2015: 218-223.
- [3] HUANG Sai, LIN Chunsheng, XU Wenjun, et al. Identification of active attacks in Internet of Things: Joint model-and data-driven automatic modulation classification approach[J]. IEEE Internet of Things Journal, 2021, 8(3): 2051-2065.
- [4] WANG Yu, GUI Guan, OHTSUKI T, et al. Multi-task learning for generalized automatic modulation classification under non-Gaussian noise with varying SNR conditions[J]. IEEE Transactions on Wireless Communications, 2021, 20(6): 3587-3596.
- [5] CORUK R B, GOKDOGAN B Y, BENZAGHTA M, et al. On the classification of modulation schemes using higher order statistics and support vector machines[J]. Wireless Personal Communications, 2022, 126(2): 1363-1381.
- [6] ZHANG Xiaolin, SUN Jianting, ZHANG Xiaotong. Automatic modulation classification based on novel feature extraction algorithms[J]. IEEE Access, 2020, 8: 16362-16371.
- [7] MA Jitong, HU Mutian, WANG Tianyu, et al. Automatic modulation classification in impulsive noise: Hyperbolic-tangent cyclic spectrum and multibranch attention shuffle

- network[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 5501613.
- [8] MOHAMMED A, KORA R. A comprehensive review on ensemble deep learning: Opportunities and challenges [J]. Journal of King Saud University-Computer and Information Sciences, 2023, 35(2): 757-774.
- [9] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks[M]//JAYNE C, ILIADIS L, eds. Engineering Applications of Neural Networks. Cham: Springer International Publishing, 2016: 213-226.
- [10] 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.
- [11] OIKONOMOU T K, EVGENIDIS N G, NIXARLIDIS D G, et al. CNN-based automatic modulation classification under phase imperfections[J]. IEEE Wireless Communications Letters, 2024, 13(5): 1508-1512.
- [12] RAJENDRAN S, MEERT W, GIUSTINIANO D, et al. Deep learning models for wireless signal classification with distributed low-cost spectrum sensors[J]. IEEE Transactions on Cognitive Communications and Networking, 2018, 4(3): 433-445.
- [13] PARMAR A, A D K, CHOUHAN A, et al. Dual-stream CNN-BiLSTM model with attention layer for automatic modulation classification [C]//2023 15th International Conference on COMMunication Systems & NETWORKS (COMSNETS). Bangalore, India. IEEE, 2023: 603-608.
- [14] HAMIDI-RAD S, JAIN S. MCformer: A transformer based deep neural network for automatic modulation classification [C]//2021 IEEE Global Communications Conference (GLOBECOM). ACM, 2021: 1-6.
- [15] CHEN Yantao, DONG Binhong, LIU Cuiting, et al. Abandon locality: Frame-wise embedding aided transformer for automatic modulation recognition[J]. IEEE Communications Letters, 2023, 27(1): 327-331.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Advances in Neural Information Processing Systems. Long Beach, CA, USA: Curran Associates, Inc., 2017: 6000-6010.
- [17] ZHANG Hao, ZHOU Fuhui, WU Qihui, et al. A novel automatic modulation classification scheme based on multi-scale networks [J]. IEEE Transactions on Cognitive Communications and Networking, 2022, 8(1): 97-110.
- [18] O'SHEA T, WEST N E. Radio machine learning dataset generation with GNU radio [C]//Proceedings of the GNU Radio Conference. Boulder, CO, USA: The GNU Radio Foundation, Inc., 2016: 1-1.
- [19] XU Jialang, LUO Chunbo, PARR G, et al. A spatio-temporal multi-channel learning framework for automatic modulation recognition [J]. IEEE Wireless Communications Letters, 2020, 9(10): 1629-1632.
- [20] ZHANG Fuxin, LUO Chunbo, XU Jialang, et al. An efficient deep learning model for automatic modulation recognition based on parameter estimation and transformation [J]. IEEE Communications Letters, 2021, 25(10): 3287-3290.

#### 作者简介



杨嘉豪 男,2000年生,江西瑞金人。上海交通大学硕士在读,主要研究方向为信号检测与识别、深度学习。  
E-mail: yangjiahao\_0601@163.com



张东坡 男,1975年生,山东曹县人。中国电子科技集团公司第36研究所副总工程师,主要研究方向为软件无线电技术和通信信号分析与处理技术。  
E-mail: 65866151@qq.com



何 劲 男,1980年生,上海人。上海交通大学副研究员,主要研究方向为先进电子防御技术、实时信号处理系统设计与开发。  
E-mail: jinhe@sjtu.edu.cn

(责任编辑:刘建新)