



研究与开发

基于多通道轻量化的自动调制识别网络

万进华¹, 尚俊娜¹, 张华娣^{1,2}

(1. 杭州电子科技大学通信工程学院, 浙江 杭州 310018;

2. 中国电子科技集团公司第三十六研究所, 浙江 嘉兴 314033)

摘要: 自动调制识别技术在无线通信领域具有十分重要的作用。现有的自动调制识别模型在识别精度上表现出色, 但大多数方法在参数量与模型性能之间难以实现理想的平衡。针对该问题, 设计了一种多通道融合的轻量化调制识别 (multi-channel lightweight modulation recognition, MCLMR) 网络。MCLMR 网络将幅度、相位、频率以及零中心归一化瞬时幅度的谱密度最大值作为输入。使用可分离卷积模块巧妙地组合 4 个输入, 从而深入挖掘这 4 个输入的空间相关性。设计了基于时间衰落多头自注意力 (multi-head self-attention, MHSA) 机制结合门控循环单元 (gated recurrent unit, GRU) 的 GRU-MHSA (gated recurrent unit-multi-head self-attention) 模块进一步提取时间相关性。可分离卷积模块与 GRU-MHSA 模块的结合在空间维度与时间维度提取信号特征。在基准 RML2016.10a 数据集上的仿真结果表明, 所提方法优于其他 9 种典型方法。在 2~18 dB 信噪比下平均识别精度达到 92.39%, 最高识别精度达到 93.36%, 这说明 MCLMR 不仅参数量少, 计算复杂度低, 在识别精度上也表现出色。

关键词: 调制识别; 轻量化网络; 多模块融合; 多头自注意力机制

中图分类号: TN92

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2025012

An automatic modulation recognition network based on multi-channel lightweighting

WAN Jinhua¹, SHANG Junna¹, ZHANG Huadi^{1,2}

1. School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China

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

Abstract: Automatic modulation recognition technology plays an important role in the field of wireless communication. Existing automatic modulation recognition models perform well in recognition accuracy, but most methods have difficulty in achieving an ideal balance between the number of parameters and model performance. To solve this problem, a multi-channel lightweight modulation recognition (MCLMR) network was designed. The amplitude, phase, frequency, and maximum value of the spectrum density of the zero-centered normalized instantaneous amplitude are used as inputs. A separable convolution module is used to combine the 4 inputs, thus deeply挖掘 the spatial correlation of the 4 inputs. A multi-head self-attention (MHSA) mechanism combined with a gated recurrent unit (GRU) is used to further extract temporal correlation. The combination of the separable convolution module and the GRU-MHSA module extracts signal features in both spatial and temporal dimensions. Simulation results on the RML2016.10a dataset show that the proposed method is superior to other 9 typical methods. Under a signal-to-noise ratio of 2~18 dB, the average recognition accuracy reaches 92.39%, and the maximum recognition accuracy reaches 93.36%, indicating that MCLMR not only has a small number of parameters and low computational complexity, but also performs well in recognition accuracy.

收稿日期: 2024-09-27; 修回日期: 2024-11-29

通信作者: 尚俊娜, 87071712@qq.com

基金项目: 浙江省教育厅一般科研项目 (No.Y202455360); 浙江省“十四五”第二批研究生省级教学改革项目 (No.JGCG2024202)

Foundation Items: Scientific Research Fund of Zhejiang Provincial Department of Education (No.Y202455360); The Second Batch of Graduate Education and Teaching Reform Projects at the Provincial Level in Zhejiang Province during the 14th Five-Year Plan Period (No.JGCG2024202)



quency, and maximum spectral density of the zero center normalized instantaneous amplitude were taken as inputs by the MCLMR network. A separable convolution module was used to cleverly combine four inputs to dig deeper into the spatial correlation of the four inputs. The gated recurrent unit-multi-head self-attention (GRU-MHSA) module multi-head self-attention (MHSA) based on time fading and gated recurrent unit (GRU). to further extract the temporal correlation. Signal features in spatial dimension and time dimension were extracted by the combination of separable convolution module and GRU-MHSA module. Simulation results on the benchmark dataset RML2016.10a show that the proposed method is superior to other 9 typical methods. At 2~18 dB signal-to-noise ratio, the average recognition accuracy reaches 92.39% and the highest recognition accuracy reaches 93.36%. This shows that MCLMR not only has a small number of parameters and low computational complexity, but also has excellent performance in recognition accuracy.

Key words: modulation recognition, lightweight network, multi-module integration, multi-head attention mechanism

0 引言

近年来,随着信息化时代的迅猛发展,通信技术在无线通信领域^[1]的重要性日益凸显。自动调制识别技术作为通信领域的一项关键技术,在民用等领域扮演着至关重要的角色。在传统的自动调制识别领域,主要识别方法分为两种:一种是基于似然函数的自动调制识别 (automatic modulation recognition, AMR) 算法^[2-3],另一种是基于特征的自动调制识别算法^[4-5]。基于似然函数的AMR算法通过最小化误分类概率来实现最优的调制估计,然而这种方法计算较为复杂,且其适用性受到一定限制。基于特征的AMR算法则是分析信号的不同特征来区分不同的调制信号,通过提取高阶谱特征、频谱特征、功率谱特征等来对信号进行识别。随着人工智能的快速发展,深度学习在图像分类与语音识别领域取得了突破性的成果。基于这一发现,越来越多的人开始将深度学习与调制识别进行融合。文献[6]构建了一种适用于调制识别的卷积神经网络,其结果远优于传统方法。文献[7]对卷积长短期记忆全连接深度神经网络进行改进并将其应用于自动调制识别领域。递归神经网络也在自动调制识别领域中得到了应用。文献[8]采用长短期记忆网络,从输入信号中提取时间特征,以实现更精准的分类。文献[9]选择信号的星座图作为输入,取得了

不错的效果。文献[10]选择长短期记忆网络的轻量级变体门控循环单元 (gated recurrent unit, GRU) 来减少频率偏移和频谱扩展。文献[11]结合卷积神经网络的空间特征提取能力与循环神经网络的时间序列处理能力,提出特殊的模型CNN-LSTM (convolutional neural network - long short-term memory) 来处理对时间依赖的信号。文献[12]将信号的幅度与相位作为输入,提出一种双输入网络。文献[13]将文献[12]中的双输入网络与CNN-LSTM结合,以提高识别精度。文献[14]提出的网络在多通道环境下展现出良好的鲁棒性。文献[15]提出结合两个卷积神经网络的数据驱动的融合模型。一个在IQ原始数据集上进行训练,另一个在星座图数据集上进行训练,他们提出的联合识别模型提供了一种新型的解决方案。文献[16]提出一种时空多通道学习框架。该框架以I、Q、IQ作为输入,将一维卷积、二维卷积与长短期记忆网络进行结合,在高信噪比下得到了很好的识别率。文献[17]采用了多通道输入,将原始信号的实部、虚部、幅度、相位以及傅里叶变换之后的实部与虚部同时输入模型中,随后结合多头自注意力机制与双向门控循环单元构建了多尺度神经网络 (multi-scale neural network, MSNN),该模型取得了超过93%的准确率,但模型参数量与计算复杂度也随之增加,导致计算资源增多。文献[18]提出

了一种轻量化网络,该网络在RML2016.10a数据集上取得了90%的准确率,尽管此模型可以降低复杂度,但也导致其准确率下降。文献[19]运用了残差网络(residual network, ResNet)模块与GRU模块,将原始数据并行输入ResNet与GRU模型中,得到了不错的效果,但其计算复杂度也随之增加。文献[20]将长短期记忆网络与两层ResNet进行顺序连接,在RML2016.10b数据集上进行实验,在18 dB时获得了92%的准确率,但其参数量也随之增加。文献[21]在极少的调制样本信息情况下,通过音频信号Urban-Sound8K作为源域进行训练,然后以一些调制信号样本为目标域进行微调,实验证明其得到了不错的效果。文献[22]改进了重叠低截获波形识别任务处理框架,将自注意力机制添加到分类器中,结合强大的表示能力和多样化的实例表示提高识别性能。

针对上述提到的问题,本文提出了一种多通道融合的轻量化调制识别(multi-channel lightweight modulation recognition, MCLMR)网络模型。

本文主要工作与贡献如下。

(1) 本文设计了一种轻量化网络模型MCLMR,其输入数据分别为信号的瞬时幅度、瞬时相位、瞬时频率以及一个关键特征——零中心归一化瞬时幅度的谱密度最大值。3种时域特征与1种频域特征进行结合,提供更全面的信号描述。这一频域的关键特征的引入,显著缓解了信号之间的混淆问题。

(2) 本文设计了可分离卷积模块与GRU-MHSA模块,可分离卷积模块通过巧妙地组合4种特征,在空间维度上提取不同特征之间互补的空间相关性。GRU-MHSA模块结合了GRU与MHSA,进一步提取时间相关性。可分离卷积模块与GRU-MHSA模块的结合从空间与时间维度提取信号特征。其中GRU负责捕捉序列的时间动态变化,而MHSA则引导模型聚焦于序列中的

关键信息,从而提高了模型的分类准确性。

(3) MCLMR在两个公开数据集上验证了其合理性。实验结果表明,MCLMR在RML2016.10a数据集和RML2016.10b数据集上均取得了优异成绩。与文献[17]中的MSNN模型对比,MCLMR能达到与之相当的性能,并且参数量仅为MSNN的10.41%。与其他经典模型相比,MCLMR不仅参数量少,而且具有更高的识别率。

1 调制信号模型及预处理

1.1 调制信号模型

本文调制信号模型是基于单输入单输出的通信系统模型,发送端发送符序列经过调制变换并通过通信信道到达接收端,接收信号的计算式为:

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

其中, t 表示时间, $s(t)$ 为发送端发送的符合序列, $f(\cdot)$ 为调制函数, $h(t)$ 为通过信号的冲激响应,*表示卷积操作, $n(t)$ 为高斯白噪声, $r(t)$ 为接收机接收到的信号。

本文离散信号 $r(n)$ 是由接收信号经过采样速率转换得到的。离散信号 $r(n)$ 由同相分量与正交分量组成,计算式为:

$$r(n)=r_1(n)+jr_Q(n) \quad (2)$$

其中, j 表示虚数。

1.2 信号预处理

1.2.1 瞬时特征

瞬时特征能够有效地提取信号的变换特征,幅度特征的提取可以通过瞬时幅度实现,这一过程通常涉及将实际信号进行希尔伯特变换,然后取其模值。所使用数据集在已经包含I(同相分量)和Q(正交分量)的情况下,可以省略希尔伯特变换步骤。瞬时幅度的计算式为:

$$A(t)=\sqrt{x_I^2(t)+x_Q^2(t)} \quad (3)$$

其中, $x_I^2(t)$ 表示同相分量, $x_Q^2(t)$ 表示正交分量。



相位特征的提取可以通过瞬时相位实现, 从原数据集中得到同相分量与正交分量后, 利用反正切函数 $\arctan(\cdot)$ 即可得到瞬时相位, 瞬时相位的计算式为:

$$P(t) = \arctan \frac{x_Q(t)}{x_I(t)} \quad (4)$$

频率特征的提取则可以通过瞬时频率实现, 瞬时频率是瞬时相位的导数, 所以瞬时频率的计算式为:

$$F(t) = \frac{d}{dt} \arctan \frac{x_Q(t)}{x_I(t)} \quad (5)$$

其中, dt 是对时间求导, $\arctan \frac{x_Q(t)}{x_I(t)}$ 是对瞬时相位求导。

上述3种时域特征可以区分不同调制方式的信号强度分布、相位分布以及频谱分布。但对于混淆信号 16QAM 与 64QAM 区分程度却并不理想, 16QAM 与 64QAM 作为调幅调相信号, 不仅具有幅度成分, 而且具有相位成分。因此本文引入一种频域特征: 零中心归一化瞬时幅度的谱密度最大值。它能提供额外的频域幅度信息。将时域相位信息与频域幅度信息融合, 使模型能够获得时域与频域的信息, 增加输入特征的多样性, 缓解信号的混淆问题, 提高模型识别性能。

1.2.2 关键特征

文献[23]提出了一种瞬时特征 $\gamma_{\max}$, 通过使用该特征对信号进行区分, 得到不错的效果。因此本文增加这个关键特征来区分易混淆的信号。

零中心归一化瞬时幅度的谱密度最大值的计算式为:

$$\gamma_{\max} = \max \frac{|\text{DFT}(A_{\text{cn}}(t))|^2}{N_s} \quad (6)$$

其中, $A_{\text{cn}}(t)$ 为零中心归一化包络, N_s 为傅里叶变换的点数, $\text{DFT}(\cdot)$ 表示离散傅里叶变换, $A_{\text{cn}}(t)$

的计算式为:

$$A_{\text{cn}}(t) = \frac{\text{abs}(A(t))}{\text{mean}(\text{abs}(A(t))) - 1} \quad (7)$$

其中, $\text{abs}(\cdot)$ 表示绝对值, $\text{mean}(\cdot)$ 表示平均值。

本文选取瞬时幅度、瞬时相位、瞬时频率以及零中心归一化瞬时幅度的谱密度最大值作为特征输入网络的原因有如下3个。

(1) 在调制过程中, 在 RML2016.10a 数据集中, 调制主要是对信号进行幅度变换、相位变换、频率变换或两种变换叠加。因此, 将上述4种特征输入网络, 理论上可以适应任何信号。

(2) 选取4种一维特征作为输入可以大大减少网络参数量, 且提取一维特征的方法易于实现。

(3) 本文将3种时域特征与1种频域特征同时输入网络中, 使网络能够同时捕捉时域特征与频域特征, 增加输入特征的多样性, 从而缓解信号的混淆问题, 提高模型识别性能。

2 多模态融合模型

在自动调制识别领域, 现有的调制分类方法往往不能充分挖掘不同特征间的互补潜力。为此, 本文提出了轻量化的多模态模型。首先对输入进行3种时域变换与1种频域特征, 时域变换获得输入在不同方向的特征区分易识别信号。频域特征结合时域特征来区分混淆信号。可分离卷积模块通过巧妙地组合4种特征, 在空间维度上提取时域特征与频域特征之间互补的空间相关性, GRU-MHSA 结合 GRU 与 MHSA, 进一步提取时间相关性。本文提出的模型是专为 AMR 任务量身定制的。通过结合时域频域特征和深度学习技术, 该模型能够高效地识别不同调制方式, 显著提升 AMR 任务的准确性和鲁棒性。整体网络模型框架如图1所示, 模型参数结构见表1。

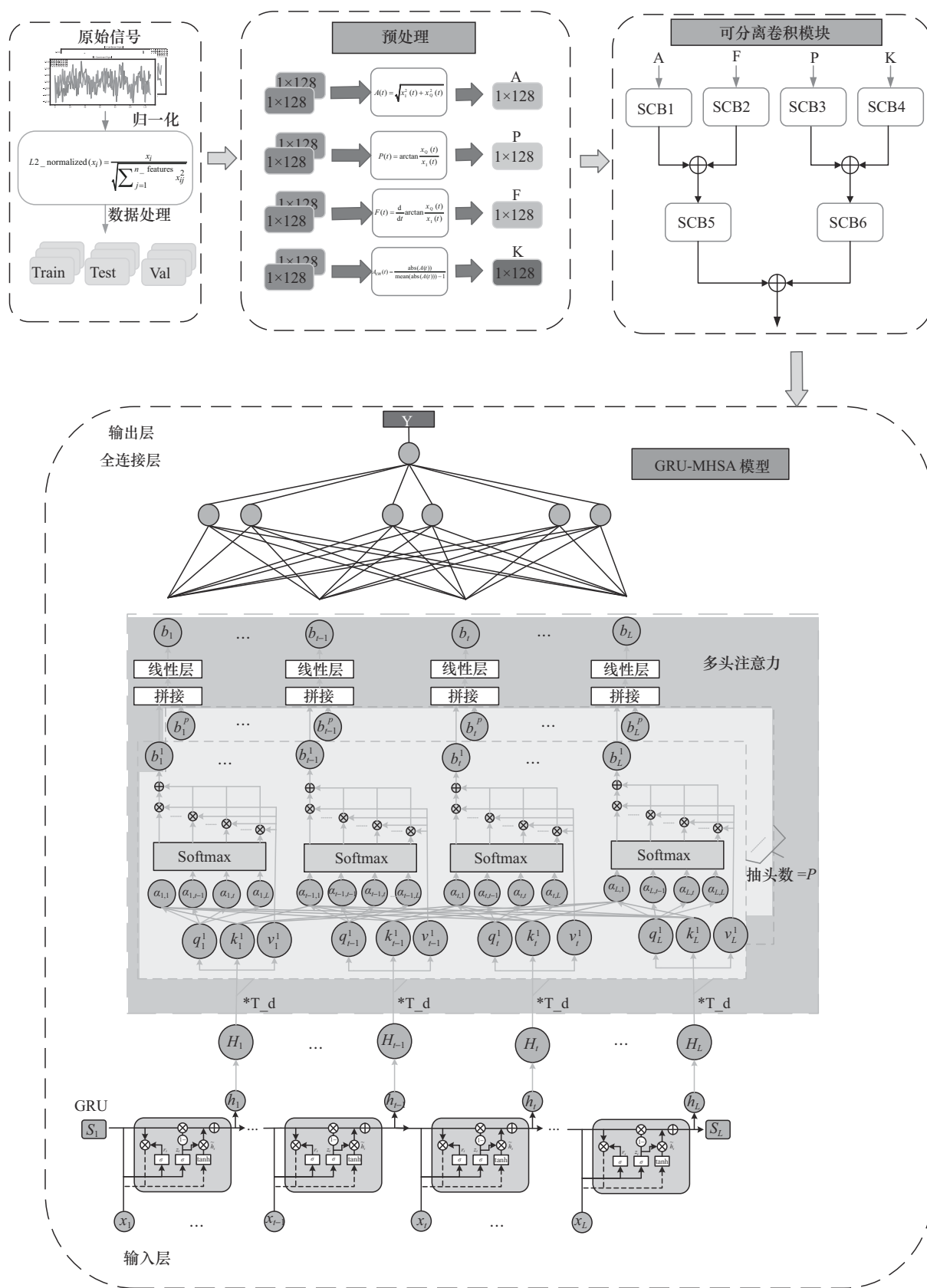


图1 整体网络模型框架



表1 模型参数结构

层	输出形状	参数
SCB1	(128,50)	108
SCB2	(128,50)	108
SCB3	(128,50)	108
SCB4	(128,50)	108
SCB5	(128,50)	5 850
SCB6	(128,50)	5 850
GRU	(128)	88 320
MHSA	(128)	24 961
FC	(11)	1 419

2.1 可分离卷积模块

本文采用可分离卷积替代常规卷积来减少多输入带来的计算复杂度增加的问题,通过可分离卷积减少模型参数,加快计算速度。在RML2016.10a数据集上常规卷积与可分离卷积性能对比见表2。可分离卷积模块如图2所示,可分离卷积模块由多个可分离卷积块(separable convolution block, SCB)组成。可分离卷积块结构如图3所示,每个卷积块由一个一维可分离卷积层、最大池化层与Dropout层组成。可分离卷积层中的卷积核是50个大小为(1,8)的卷积核,激活函数为ReLU,最大池化层对特征图进行下采样,池化窗口大小为2,步幅为1,Dropout层随机丢弃50%的神经元,防止模型过度拟合。

表2 常规卷积与可分离卷积性能对比

卷积层	2~18 dB 准确率	参数	计算时间/ms
常规卷积	91.88%	196 599	25
可分离卷积	92.39%	126 831	24

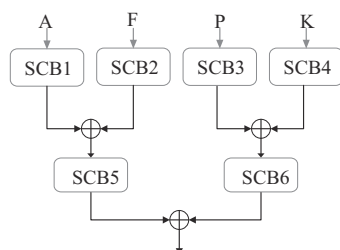


图2 可分离卷积模块

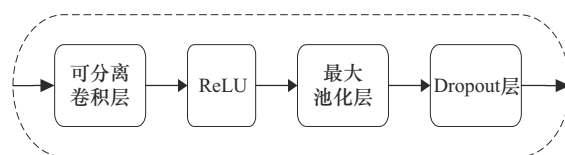


图3 可分离卷积块结构

可分离卷积模块振幅(A)输入SCB1块得到的特征图与频率(F)输入SCB2块得到的特征图进行拼接,相位(P)输入SCB3块得到的特征图与零中心归一化瞬时幅度谱密度最大值(K)输入SCB4块得到的特征图进行拼接。最后SCB5块与SCB6块提取拼接后的特征图并将特征图进行拼接后输入GRU。多流输入设计能够在不同尺度上捕捉特征之间的互补信息,可分离卷积模块结构设计的原因有如下两个。

(1) 零中心归一化瞬时幅度谱密度最大值在计算时依赖于信号的瞬时幅度,并且关注的是频谱中最强的成分(即最大值),而这个最大值通常对应于信号中最强的频率成分。因此它是一个包含幅度信息和频率信息的综合特征,因此将 K 与 A 或 F 结合后会导致信息冗余,降低模型的学习效率。本文设计基于信息互补性和减少信息冗余的考虑,最大限度提升模型性能。

(2) P 与 K 能够同时提供时域相位信息与频域相位信息,这对混淆信号的区分有较好的效果,在RML2016.10a数据集上不同组合方式的性能对比见表3。

表3 不同组合方式的性能对比

组合方式	2~18 dB 准确率	最高准确率
A 与 F , K 与 P	92.39%	93.39%
A 与 P , K 与 F	92.02%	93.11%
A 与 K , P 与 F	91.74%	92.95%

2.2 GRU模块

为了提升模型在时间维度的提取能力,本文引入了GRU模块。GRU通过其简化的门控机制,在保持优异的长期依赖处理能力的同时,显著减少了模型参数量。

GRU 结构如图 4 所示, GRU 单元由两个门组成: 重置门与更新门。首先通过重置门控来得到重置之后的数据。\$x_t\$ 表示当前时刻的输入信息, \$h_{t-1}\$ 则是前一时刻的隐藏状态, \$[h_{t-1}, x_t]\$ 表示 \$h_{t-1}\$ 和 \$x_t\$ 拼接在一起, \$\sigma(\cdot)\$ 将输入压缩到 0~1 之间。\$z_t\$ 是更新门, \$r_t\$ 是重置门的输出, 控制 \$h_{t-1}\$ 有多少信息用于计算当前时刻的候选隐藏状态, 重置门表示为式 (8)。\$\tilde{h}_t\$ 为候选隐藏状态, \$\tilde{h}_t\$ 的计算式如式 (10), \$r_t * h_{t-1}\$ 表示 \$r_t\$ 与 \$h_{t-1}\$ 逐元素相乘, \$\tanh\$ 将输入压缩到 -1~1 之间。随后更新门 \$h_{t-1}\$ 控制有多少信息传递到当前时刻的隐藏状态, 计算式如式 (9) 所示。\$h_t\$ 为当前时刻的隐藏状态, \$h_t\$ 由 \$h_{t-1}\$ 和 \$\tilde{h}_t\$ 加权组合得到, 如式 (11) 所示。最后当前时刻的隐藏状态 \$h_t\$ 作为下一时刻的输入。

在此过程中, \$\sigma(\cdot)\$ 代表 sigmoid 函数, \$W_r\$ 与 \$W_z\$ 为权重矩阵, \$*\$ 表示逐元素乘法, \$W\$ 为模型参数。具体计算式如下。

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \quad (8)$$

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \quad (9)$$

$$\tilde{h}_t = \tanh(W[r_t * h_{t-1}, x_t]) \quad (10)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (11)$$

2.3 时间衰减的 MHSA 模块

虽然 GRU 在捕获序列中的长距离依赖性方面表现良好, 但逐步迭代处理的机制会导致 GRU 无法完全捕获长距离上下文信息。本文引入的基于时间衰落的 MHSA 模块来减少较早时间步对当前时间步的影响, 从而增强模型对长距离依赖的捕捉能力。基于时间衰落的 MHSA 模块结构如下。

自注意力机制是一种机器学习技术, 它可以捕捉输入序列中不同位置之间的相关性和重要性。自注意力机制通过将查询 (Query)、键 (Key) 和值 (Value) 向量映射到输出。具体步骤如下, 输入序列 \$\Theta = [\theta_1, \theta_2, \dots, \theta_L]^T \in \mathbf{R}^{L \times d_m}\$, \$\Theta\$ 表示向量, \$L\$ 是序列长度, 分别与 3 个可训练矩阵 \$W_q\$、\$W_k\$ 与 \$W_v\$ 相乘来获得查询、键和值向量:

$$\begin{cases} q_t = \theta_t \times W_q \\ k_t = \theta_t \times W_k \\ v_t = \theta_t \times W_v \\ t = 1, 2, \dots, L \end{cases} \quad (12)$$

随后, 为了调节模型对近期和远期输入的关注度, 使得模型能够在处理时序数据时更好地捕捉时间依赖性, 本文引入了时间衰减因子 \$T_d\$ (time_decay), 时间衰减率 decay_rate 是一个可训练的变量。在每个时间步 \$t\$, \$T_d\$ 计算式为:

$$\text{time_decay}_t = \text{decay_rate}^t \quad (13)$$

decay_rate 由模型初始化设置为 0.9。decay_rate 在训练过程中通过反向传播算法进行更新, 其更新步骤如下, 首先计算损失函数对 \$T_d\$ 的梯度:

$$\frac{\partial \text{Loss}}{\partial \text{time_decay}_t} \quad (14)$$

然后使用链式法则来计算梯度:

$$\frac{\partial \text{Loss}}{\partial \text{decay_rate}} = \sum_{t=1}^T \frac{\partial \text{Loss}}{\partial \text{time_decay}_t} \cdot \frac{\partial \text{time_decay}_t}{\partial \text{decay_rate}} \quad (15)$$

其中, \$T\$ 是时间步总数。

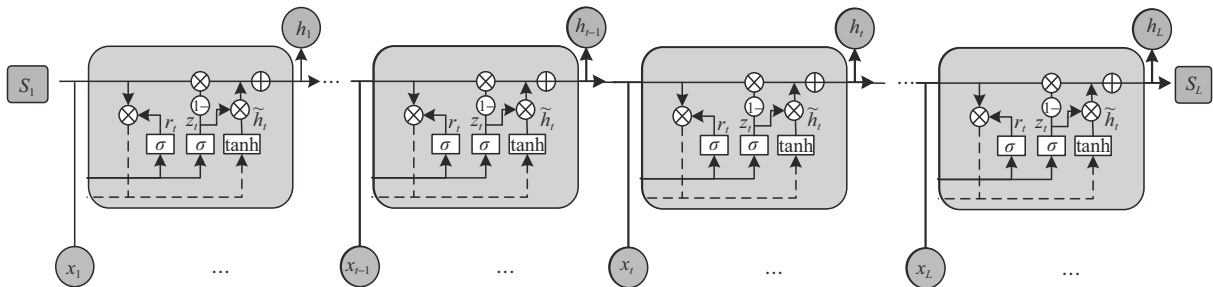


图4 GRU 结构



$$\frac{\partial \text{time_decay}_t}{\partial \text{decay_rate}} = t \cdot \text{decay_rate}^{t-1} \quad (16)$$

最后使用梯度下降法更新 decay_rate:

$$\text{decay_rate} \leftarrow \text{decay_rate} - \eta \cdot \frac{\partial \text{Loss}}{\partial \text{decay_rate}} \quad (17)$$

其中, η 是学习率, 更新 decay_rate, 每个时间步内, time_decay 也会得到更新, 相应的查询、键和值向量也会得到更新:

$$\begin{cases} q_t = q_t * T_d \\ k_t = k_t * T_d \\ v_t = v_t * T_d \\ t = 1, 2, \dots, L \end{cases} \quad (18)$$

基于梯度更新的时间衰减因子可以根据当前训练过程的梯度信息来自动调整时间衰减因子。模型可以根据当前任务的需求动态调整对历史数据的依赖, 从而更好地捕捉时间序列。

自注意力机制的中心概念引入权重分数来衡量不同元素的重要性。通过计算查询 q 和键 k 之间的相似度来确定值 v 的注意力权重分布。注意力得分 $\alpha_{i,j}$ 可以表示为:

$$\alpha_{i,j} = \frac{q_i \cdot k_j^T}{\sqrt{d_k}} \quad (19)$$

其中, i 和 j 分别表示第 i 个和第 j 个时间步长, d_k 是比例因子。最后, 注意力输出 b_i 通过对 v 中的所有元素执行加权求和来获得:

$$b_i = \sum_{j=1}^{j=L} \text{Soft max}(\alpha_{i,j}) v_j \quad (20)$$

此外, 向量 q 、 k 、 v 被打包到矩阵 Q 、 K 、 V 中用于矩阵乘法计算, 其可以被描述为等式, 自注意力机制可以被描述为:

$$\begin{cases} Q = [q_1, q_2, \dots, q_L]^T \\ K = [k_1, k_2, \dots, k_L]^T \\ V = [v_1, v_2, \dots, v_L]^T \end{cases} \quad (21)$$

$$B = \text{Soft max}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (22)$$

通过引入多个独立的注意力头, MHSA 能够学习不同的权重分布和关键信息。时间衰落的多头注意力机制如图 5 所示。具体步骤如下, 首先通过线性变换将输入序列映射到多个头部的 Q 、

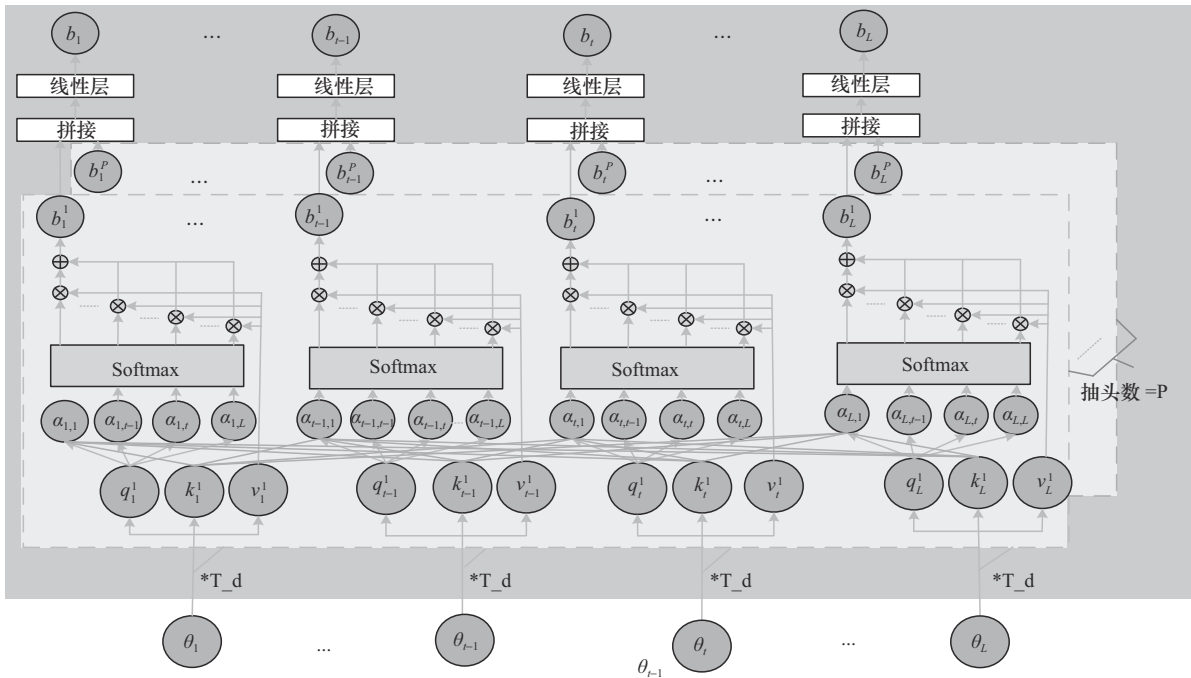


图5 时间衰落的多头注意力机制

K 和 V , 然后在每个头上执行自注意力机制。对于第 l 个头, 该过程可以表示为:

$$\begin{cases} Q^l = \theta \times W_l^Q \\ K^l = \theta \times W_l^K \\ V^l = \theta \times W_l^V \end{cases} \quad (23)$$

$$B^l = \text{Softmax}\left(\frac{Q^l (K^l)^T}{\sqrt{d_k}}\right) V^l$$

沿着特征维度连接来自所有头的输出, 并对组合矩阵进行线性变换, 以获得 MHSA 机制的

结果:

$$B_{\text{MHSA}} = \text{concat}(B^1, B^2, \dots, B^P) W^O \quad (24)$$

其中, P 是头的数目, $W^O \in R^{d_m \times d_m}$ 是用于线性投影的矩阵。

2.4 GRU-MHSA 模型图

GRU-MHSA 模型的网络结构如图 6 所示, 预测模型由 GRU 层、MHSA 层、全连接层与输出层组成。首先, 从可分离卷积模块出来的子序列以滑动窗口的方式被送到 GRU 层。GRU 能够提

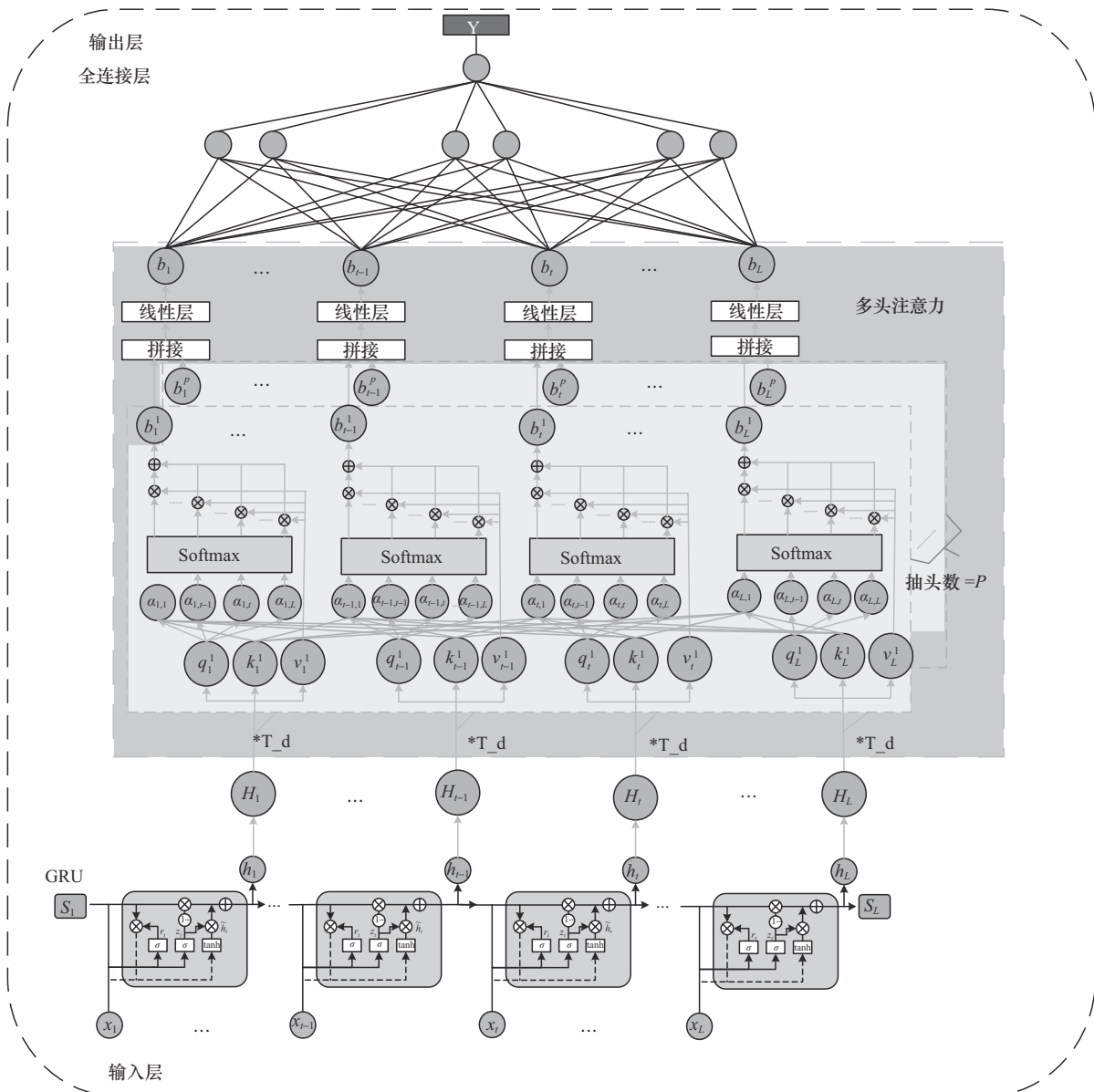


图 6 GRU-MHSA 模型的网络结构



取输入序列的上下文信息与依赖关系，并将提取的内容编码为隐藏状态序列。然后GRU隐藏状态序列被送到MHSA层，由多个自注意力头处理。自注意力机制通过计算每个位置与其他位置之间的注意力权重，直接对整个序列中不同位置之间的依赖关系进行建模和关联。每个自注意力头可以通过学习不同的注意力权重分布来关注序列中的不同特征。

最后，MHSA层的输出被馈送到全连接层，用于进一步融合和处理特征，全连接层的输出被馈送到输出层，生成最终的预测结果。通过将GRU与MHSA相结合，该模型能够捕捉输入与输出之间的依赖关系，提取更丰富的特征信息，从而提高预测精度。

3 实验与分析

本文使用的是公共RML2016.10a数据集与RML2016.10b数据集。RML2016.10a数据集分别为I通道的数据与Q通道的数据，对每个样本标记了信噪比与调制方式，每个样本由128个采样点组成， 2×128 个向量输入模型。数据集中包含11种不同的调制方式，包括{BPSK、QPSK、8PSK、AM-SSB、AM-DSB、CPFSK、GFSK、PAM4、16QAM、64QAM、WBFM}。调制方式的信噪比（SNR）范围为-20 dB到18 dB，每个信噪比下的每种调制信号样本数量为1 000，样本总量为220 000。而在RML2016.10b数据集则比RML2016.10a数据集少AM-SSB类型，且每个信噪比下的每种调制信号增加到6 000个样本，共1 200 000个样本。这两个数据集均包含加性高斯白噪声、中心频率偏移、多径衰落等。

3.1 损失函数设置

本文使用交叉熵损失函数来进行验证调试，给定两个概率分布 p 与 g ，通过 g 表示 p 的交叉熵如式（25）所示：

$$H(p, g) = - \sum p(x) \log g(x) \quad (25)$$

其中， p 是真实结果的概率分布， g 是预测结果的概率分布。

本文使用了3种损失函数作为对比，RML2016.10a数据集不同损失下的准确率见表4。

表4 RML2016.10a数据集不同损失下的准确率

损失函数	2~18 dB 准确率
交叉熵损失	92.39%
焦点损失	91.51%
合页损失	91.92%

交叉熵损失函数是一种用于度量预测概率 y 与样本真实类别 y_i 之间的差别的衡量指标，它能够为每个预测结果提供平滑且有效的梯度信号，有助于模型参数的优化，在分类问题中发挥关键作用，可以显著提高模型分类的准确率。相比较其他损失函数，交叉熵损失函数在训练模型时候更容易获得稳定的收敛，如式（26）所示：

$$L = - \frac{1}{N} \sum_{i=1}^N [y_i \ln \hat{y}_i + (1 - y_i) \ln (1 - \hat{y}_i)] \quad (26)$$

其中， N 为样本总数， y_i 是第 i 个样本的真实类别标签， $\hat{y}_i$ 是模型给出的第 i 个样本的估计值，损失的值越小，表示模型的预测结果与真实结果之间差异越小，模型性能越好。

3.2 实施细节

本文训练集、验证集与测试集的比例设置为6:2:2，在每个信噪比的每个调制方式的样本中随机选择，在训练过程中，采用了提前学习率衰减策略以优化模型性能。初始学习率设定为0.001，若验证集的损失在连续5个周期内未见改善，则将学习率降低50%。若验证集损失在连续50个周期内持续未降低，则终止训练并保存损失最小的模型权重。此外，将批处理大小固定为512，所有实验均在TensorFlow框架下进行。

3.3 消融实验

在相同的实验环境下, 本文对提出的模型进行了消融实验以评估实验中每个变量对实验的影响。不同框架下的准确率曲线如图7所示, 在0~18 dB信噪比范围内, 本文模型MCLMR的平均识别率高于其他模型。RML2016.10a数据集不同框架下准确率见表5, 在引入关键特征后, MCLMR的准确率在2~18 dB信噪比范围内比缺少关键特征MCLMR-A提高了1%以上。在缺少注意力机制的情况下, 模型MCLMR-B的准确率比MCLMR低1%, 这说明了多头注意力机制在配合GRU之后可以更有效地捕捉长距离依赖关系, 从而捕捉到更多信息。在缺少多头注意力机制与关键特征后, MCLMR-C表现的结果最差。MCLMR-D(缺乏时间衰减因子的GRU-MHSA)在没有时间衰减因子的帮助下准确率明显不如MCLMR, 在缺乏时间衰减因子的情况下, 多头注意力机制无法很好地配合GRU捕获长距离上下文信息。

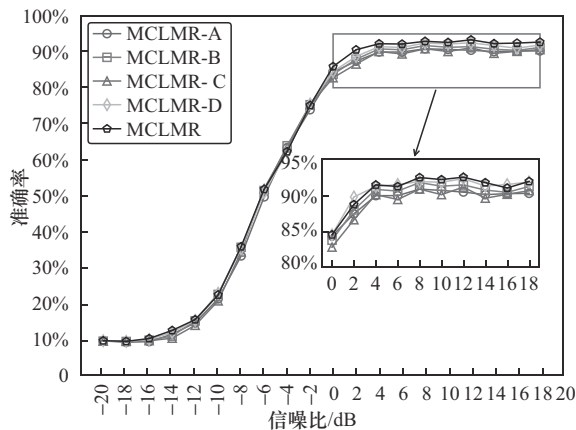


图7 不同框架下的准确率曲线

表5 RML2016.10a数据集不同框架下准确率

准确率	2 dB	4 dB	6 dB	8 dB	10 dB	12 dB	14 dB	16 dB	18 dB
MCLMR-A	87.41%	89.95%	89.95%	90.86%	90.68%	90.45%	90.09%	90.23%	90.27%
MCLMR-B	87.95%	90.82%	90.59%	91.77%	91.27%	91.45%	90.73%	90.41%	91.23%
MCLMR-C	86.59%	90.09%	89.41%	90.86%	90.09%	91.14%	89.59%	90.14%	90.64%
MCLMR-D	88.71%	91.45%	91.23%	92.48%	92.19%	92.53%	91.77%	91.02%	91.94%
MCLMR	90.59%	92.27%	92.18%	92.95%	92.64%	93.36%	92.32%	92.45%	92.73%

为了进一步验证所选关键特征解决混淆信号的有效性, 本文分别对比了MCLMR-A与MCLMR在RML2016.10a数据集中6 dB与18 dB下的混淆矩阵, SNR=6 dB如图8所示, SNR=18 dB如图9所示。

在图8中, 当信噪比为6 dB时, MCLMR中16QAM信号的识别率比MCLMR-A提高了7%, 64QAM识别率提高13.5%。

在图9中, 当信噪比为18 dB时, MCLMR中16QAM信号的识别率比MCLMR-A提高了12%, 64QAM识别率提高了8%。

在两组对比数据中, MCLMR在加入关键特征后, 网络可以额外捕捉频域特征, 增强了特征表达能力。

3.4 对比实验

本节给出了不同模型情况下MCLMR与其他9种现有模型之间的准确率比较, 不同模型的识别准确率比较如图10所示, 在公共RML2016.10a数据集中, 低信噪比下所有模型的性能急剧下降, 因为低信噪比意味着信号强度相对于噪声较弱, 噪声会掩盖信号中的有用信息, 信号中的有用信息被噪声淹没, 模型难以从中提取出有效特征, 导致模型性能下降。MCLMR在高信噪比2~18 dB下的平均识别准确率比其他模型高出1%~10%。在高信噪比12 dB时, 最大准确率为93.36%。在公共RML2016.10b数据集上进行了同样的实验, 不同模型的识别准确率比较如图11所示, MCLMR性能也超越了其他模型。

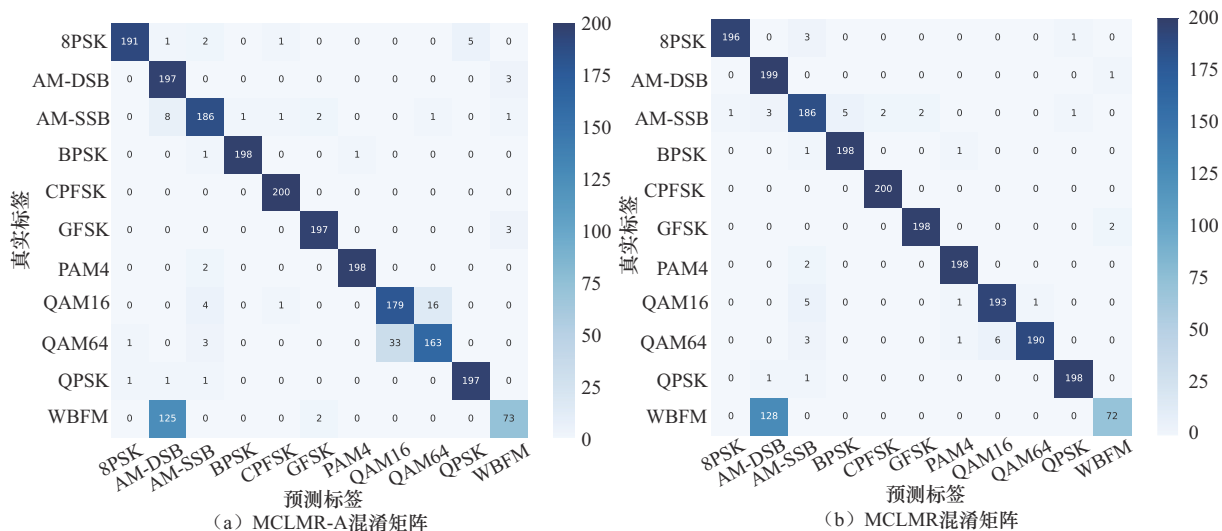


图8 SNR=6 dB

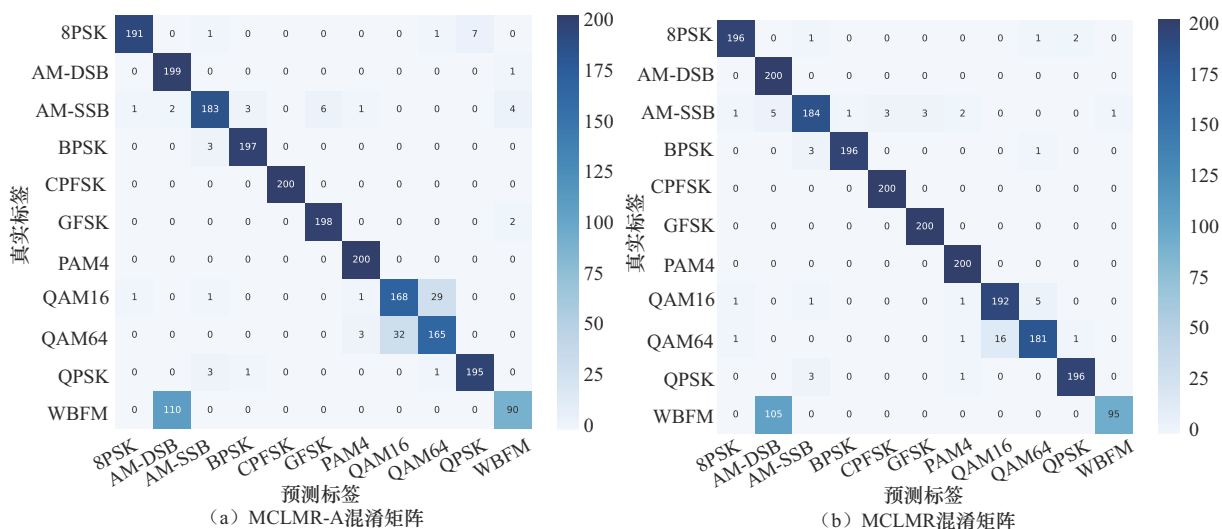


图9 SNR=18 dB

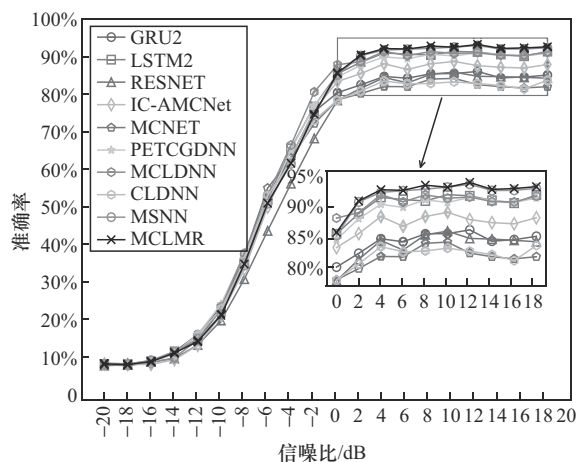


图10 不同模型的识别准确率比较

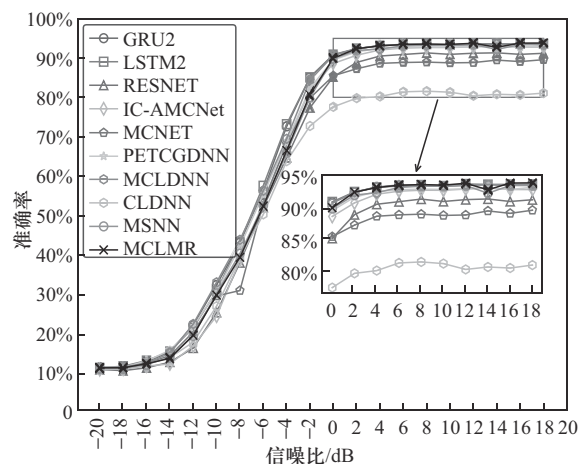


图11 不同模型的识别准确率比较

SNR=6 dB 如图 12 所示, 当信噪比为 6 dB 时, MCLMR 中 16QAM 与 64QAM 整体识别率比其他模型提高了 6.75% 以上。

SNR=18 dB 如图 13 所示, 当信噪比为 18 dB 时, MCLMR 中 16QAM 与 64QAM 整体识别率比其他模型提高了 5.25% 以上。

现有模型对于 16QAM 与 64QAM 的分类还存在混淆的情况, 可能是因为没有考虑 16QAM 信号与 64QAM 信号在频域上的幅度分布特性存在差异。而 MCLMR 模型在使用频域关键特征的情况后, 考虑了频域的幅度信息, 有效缓解了这两

个信号的混淆。

3.5 收敛性分析

RML2016.10a 数据集上 MCLMR 模型的损失函数收敛曲线如图 14 所示, 初始阶段损失函数的损失较高, 这是因为模型的权重在初始化时是随机设置的。由于初始权重是随机的, 模型在最初阶段的表现通常很差, 损失很高。随着训练轮次增加, 损失值逐渐下降, 在第 75 轮左右损失函数呈现出明显收敛的趋势, 随后损失值趋于平稳。在第 125 轮左右, 损失值变化很小, 这表明模型已经在训练集上达到最优拟合。

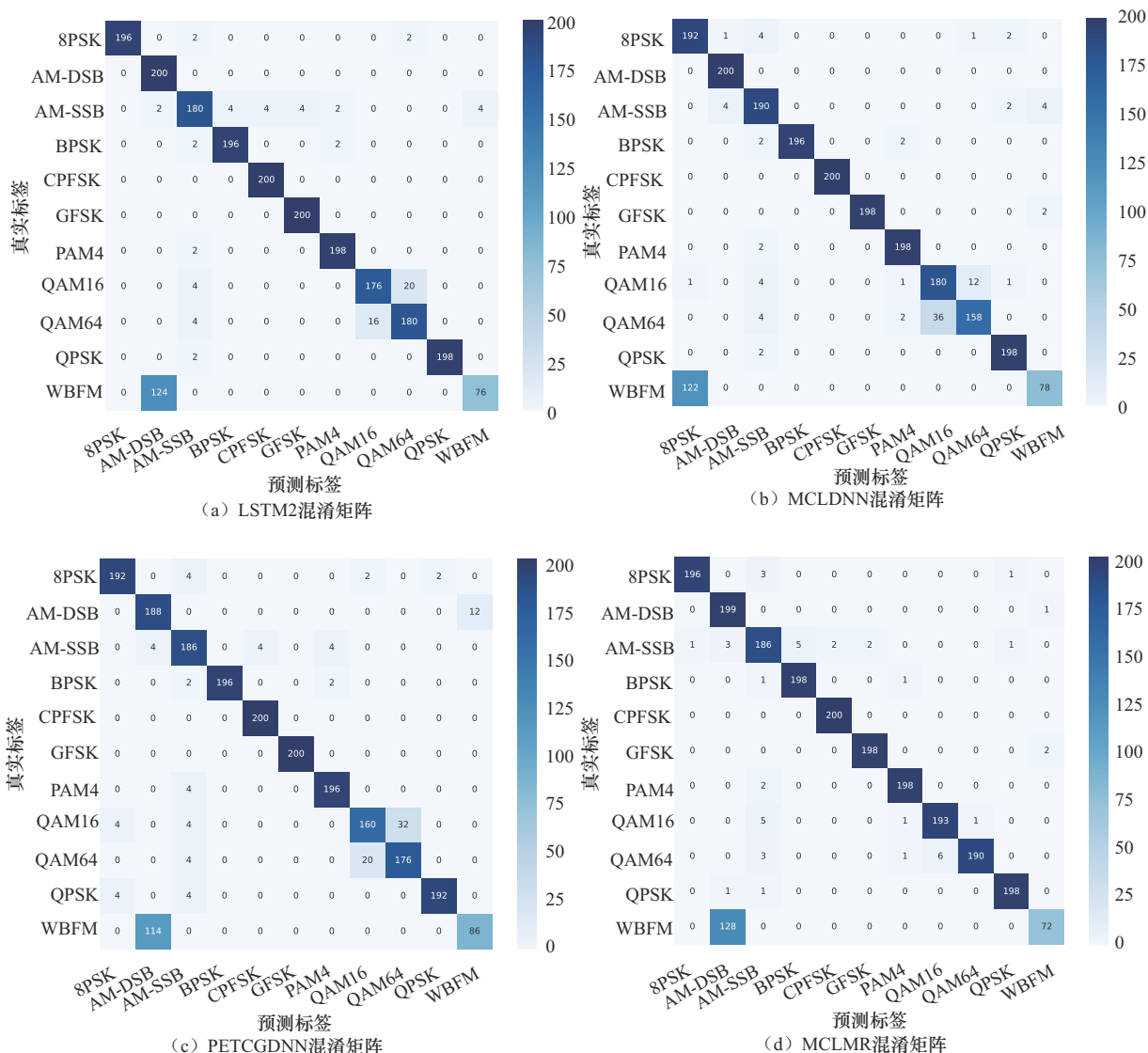


图 12 SNR=6 dB

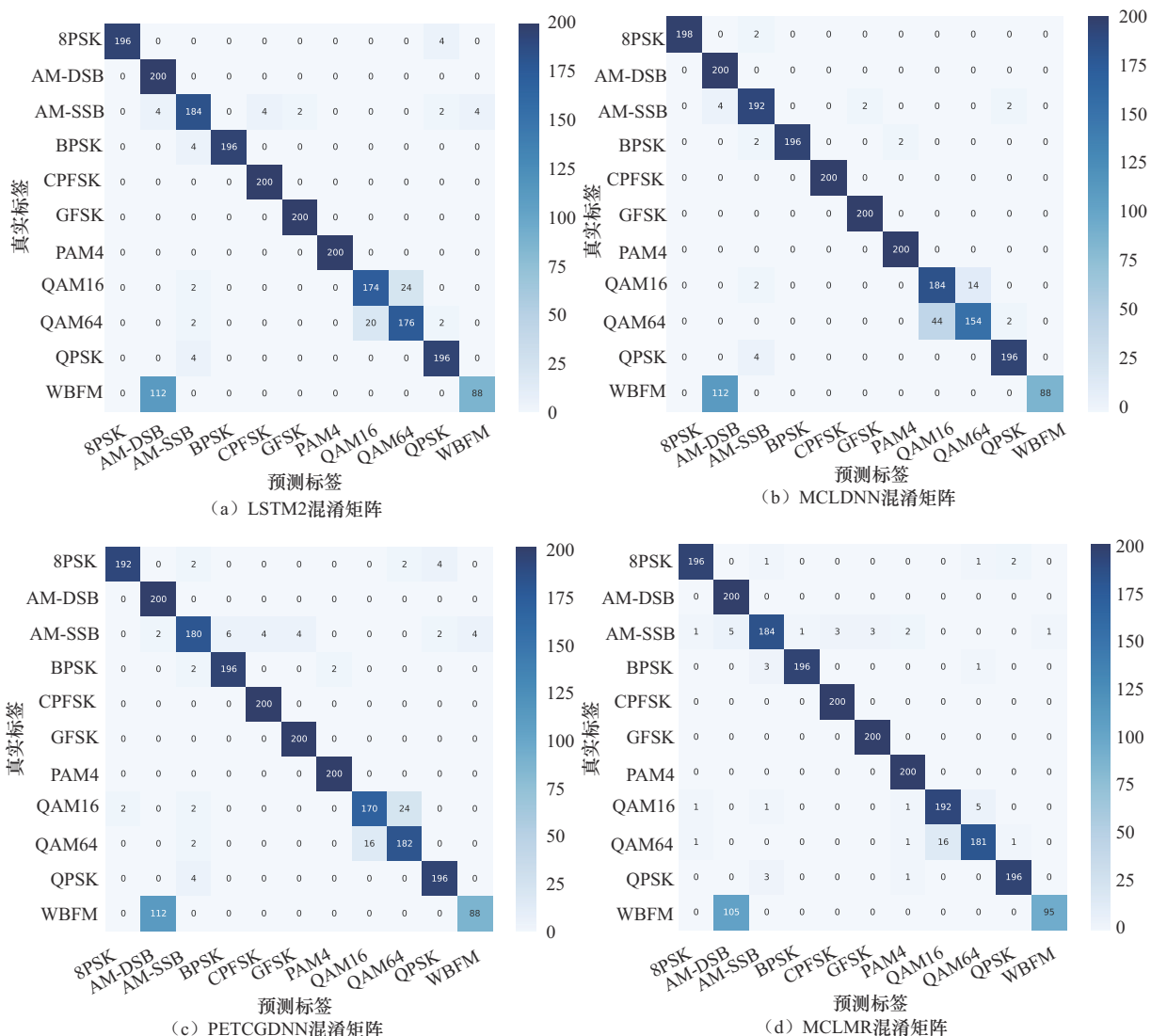


图13 SNR=18 dB

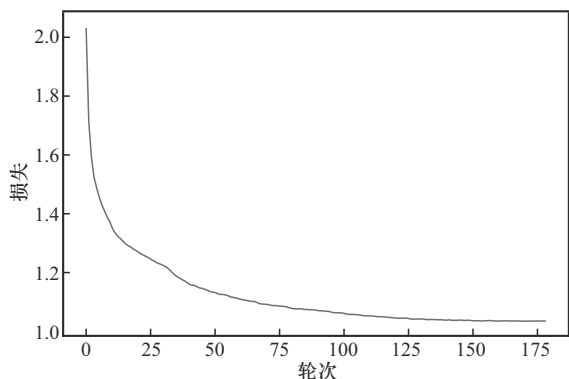


图14 RML2016.10a数据集上MCLMR模型的损失函数收敛曲线

3.6 复杂度对比分析

GRU2^[24]、LSTM2^[25]、RESNET^[26]与IC-

AMCNet^[27]等9种神经网络都被用于自动调制识别。本文通过对比不同模型参数量与对单个输入实例进行推理（预测）所需的时间来对比分析模型的复杂度。RML2016.10a数据集复杂性分析的比较见表6，列出了每个模型的模型参数，计算时间和识别精度。本文设计的MCLMR模型参数量共 1.2×10^5 ，比最轻量化的模型略大，但在2~18 dB信噪比下，本文模型平均准确率比PETCGDNN高2%。与MSNN相比，本文模型平均准确率高0.24%，但是参数量只有MSNN的10.41%。

相较于其他模型,本文模型的参数量最小,且准确率也是最高的。从计算时间来看,本文模型对单个输入实例进行推理所需的时间较为居中。主要原因是本文设计的模型是多输入模型,需要对单个输入进行多重变换得到所需特征,且在每个卷积层后都进行了池化和Dropout操作。而其他模型则是直接输入原始信号,无须进行其他操作。

表6 RML2016.10a数据集复杂性分析的比较

模型	参数/ $\times 10^6$	2~18 dB 准确率	计算时间/ms
GRU2	0.15	84.96%	14
LSTM2	0.2	90.73%	17
RESNET	3.1	84.51%	33
IC-AMCNet	1.26	87.60%	7
MCNET	0.12	82.57%	61
PETCGDNN	0.07	90.22%	9
CLDNN	0.51	82.23%	20
MCLDNN	0.4	91.20%	24
MSNN	1.2	92.15%	21
MCLMR	0.12	92.39%	24

3.7 模型的局限性

为了验证MCLMR的鲁棒性,在RML2016.10a数据集上进行了实验。在高信噪比下,WBFM信号与AM-DSB信号之间仍然存在混淆是因为这两种调制信号存在静默周期,静默周期会引入不连续性和时间上的不均匀分布,从而导致模型很难区分这两种调制信号。

4 结束语

本文提出了一种更为轻量级的模型架构,通过引入可分离卷积模块将瞬时幅度、瞬时相位、瞬时频率以及零中心归一化瞬时幅度的谱密度最大值巧妙地组合,从时域以及频域捕捉信号特征。本文结合了GRU与MHSA,这种独特的组合不仅使得模型能够更为高效地捕捉信号中的长距离依赖关系,还进一步提升了模型的整体性

能。通过在多个不同数据集上的广泛实验验证,MCLMR不仅在识别精度上表现优异,更在参数量和计算速度上展现出显著的优势,为实际应用场景中的高效信号处理与识别提供了强有力的支持。

参考文献:

- [1] DOBRE O A, ABDI A, BAR-NESS Y, et al. Survey of automatic modulation classification techniques: classical approaches and new trends[J]. IET Communications, 2007, 1(2): 137-156.
- [2] WEI W, MENDEL J M. Maximum-likelihood classification for digital amplitude-phase modulations[J]. IEEE Transactions on Communications, 2000, 48(2): 189-193.
- [3] SU W, XU J L, ZHOU M C. Real-time modulation classification based on maximum likelihood[J]. IEEE Communications Letters, 2008, 12(11): 801-803.
- [4] MARCHAND P, LE MARTRET C, LACOUME J L. Classification of linear modulations by a combination of different orders cyclic cumulants[C]//Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics. Piscataway: IEEE Press, 1997: 47-51.
- [5] GHASEMZADEH P, BANERJEE S, HEMPEL M, et al. Accuracy analysis of feature-based automatic modulation classification with blind modulation detection[C]//Proceedings of the 2019 International Conference on Computing, Networking and Communications (ICNC). Piscataway: IEEE Press, 2019: 1000-1004.
- [6] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks[C]//Proceedings of the JAYNE C, ILIADIS L, eds. Communications in Computer and Information Science. Cham: Springer International Publishing, 2016: 213-226.
- [7] WEST N E, O'SHEA T. Deep architectures for modulation recognition[C]//Proceedings of the 2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN). Piscataway: IEEE Press, 2017: 1-6.
- [8] CHEN Y F, SHAO W, LIU J, et al. Automatic modulation classification scheme based on LSTM with random erasing and attention mechanism[J]. IEEE Access, 2020, 8: 154290-154300.
- [9] PENG S L, JIANG H Y, WANG H X, et al. Modulation classification based on signal constellation diagrams and deep learning [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(3): 718-727.



- [10] LIU K, XIANG X, LIANG Y, et al. Automatic modulation recognition through wireless sensor networks in aeronautical wireless channel[J]. IEEE Sensors Journal, 2021, 21(20): 23125-23132.
- [11] WU Y L, LI X J, FANG J. A deep learning approach for modulation recognition via exploiting temporal correlations[C]//Proceedings of the 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Piscataway: IEEE Press, 2018: 1-5.
- [12] PERENDA E, RAJENDRAN S, POLLIN S. Automatic modulation classification using parallel fusion of convolutional neural networks[J]. BalkanCom'19, 2019.
- [13] ZHANG Z F, LUO H, WANG C, et al. Automatic modulation classification using CNN-LSTM based dual-stream structure [J]. IEEE Transactions on Vehicular Technology, 2020, 69(11): 13521-13531.
- [14] HUYNH-THE T, HUA C H, PHAM Q V, et al. MCNet: an efficient CNN architecture for robust automatic modulation classification[J]. IEEE Communications Letters, 2020, 24(4): 811-815.
- [15] 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.
- [16] XU J L, LUO C B, PARR G, et al. A spatiotemporal multi-channel learning framework for automatic modulation recognition[J]. IEEE Wireless Communications Letters, 2020, 9(10): 1629-1632.
- [17] WANG X, LIU D J, ZHANG Y H, et al. A spatiotemporal multi-stream learning framework based on attention mechanism for automatic modulation recognition[J]. Digital Signal Processing, 2022(130): 103703.
- [18] ZHANG F X, LUO C B, XU J L, 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.
- [19] LI L, ZHU Y X, ZHU Z G. Automatic modulation classification using ResNet-GRU with deep feature fusion[J]. IEEE Transactions on Instrumentation and Measurement, 2023(72): 2519710.
- [20] ELSAGHEER M M, RAMZY S M. A hybrid model for automatic modulation classification based on residual neural networks and long short term memory[J]. Alexandria Engineering Journal, 2023(67): 117-128.
- [21] LIN W S, HOU D B, HUANG J S, et al. Transfer learning for automatic modulation recognition using a few modulated signal samples[J]. IEEE Transactions on Vehicular Technology, 2023, 72(9): 12391-12395.
- [22] PAN Z S, WANG S F, LI Y J. Residual attention-aided U-net GAN and multi-instance multilabel classifier for automatic waveform recognition of overlapping LPI radar signals[J]. IEEE Transactions on Aerospace and Electronic Systems, 2022, 58(5): 4377-4395.
- [23] NANDI A K, AZZOUZ E E. Algorithms for automatic modulation recognition of communication signals[J]. IEEE Transactions on Communications, 1998, 46(4): 431-436.
- [24] HONG D H, ZHANG Z L, XU X D. Automatic modulation classification using recurrent neural networks[C]//Proceedings of the 2017 3rd IEEE International Conference on Computer and Communications (ICCC). Piscataway: IEEE Press, 2017: 695-700.
- [25] 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.
- [26] LIU X Y, YANG D Y, EL GAMAL A. Deep neural network architectures for modulation classification[C]//Proceedings of the 2017 51st Asilomar Conference on Signals, Systems, and Computers. Piscataway: IEEE Press, 2017: 915-919.
- [27] HERMAWAN A P, GINANJAR R R, KIM D S, et al. CNN-based automatic modulation classification for beyond 5G communications[J]. IEEE Communications Letters, 2020, 24(5): 1038-1041.

[作者简介]



万进华 (2001-), 男, 杭州电子科技大学硕士生, 主要研究方向为通信信号调制识别。

尚俊娜 (1979-), 女, 博士, 杭州电子科技大学教授、硕士生导师, 主要研究方向为卫星通信、导航与定位方面。

张华娣 (1978-), 女, 中国电子科技集团公司第36研究所正高级工程师、硕士生导师, 主要研究方向为通信信号分析识别。