

基于双模态混合神经网络的自动调制识别方法^①



郭业才^{1,3}, 王孟杰^{1,2}, 毛湘南¹, 胡晓伟¹

¹(南京信息工程大学 电子与信息工程学院, 南京 210044)

²(南京信息工程大学天长研究院, 滁州 239399)

³(无锡学院 电子信息工程学院, 无锡 214105)

通信作者: 郭业才, E-mail: guo-yecai@163.com

摘要: 为提升调制识别的准确性和鲁棒性, 本文提出了一种全新的双模态混合调制识别模型. 模型同时考虑原始时域同相正交 (in-phase and quadrature, I/Q) 和幅度相位 (amplitude and phase, A/P) 双模态数据以探索信号的时空相关性. 采用双路对称结构对 A/P 模态数据进一步处理, 更有效地学习数据间的重复特征, 避免信息冗余. 模型中引入双向长短时记忆网络 (bidirectional long short-term memory network, BiLSTM), 利用其双向时序特征提取能力, 增强模型对复杂时序信息的理解. 实验结果表明, 所提模型在数据集 RadioML2016.10A 上表现良好. 当 SNR 低于 -8 dB 时, 平均识别精度比主流模型提升 6%, 而 SNR 在 0-18 dB 时, 平均识别精度比主流模型提高 2%-10%, 且在 SNR 为 16 dB 时, 识别精度高达 94.32%. 另外, 将模型迁移到数据集 RadioML2016.10B 所得结果同样最优, 且当 SNR 为 18 dB 时识别精度高达 93.91%.

关键词: 自动调制识别; 深度学习; 双路对称; 双向提取特征

引用格式: 郭业才, 王孟杰, 毛湘南, 胡晓伟. 基于双模态混合神经网络的自动调制识别方法. 计算机系统应用, 2025, 34(8): 169-178. <http://www.c-s-a.org.cn/1003-3254/9943.html>

Automatic Modulation Recognition Method Based on Bimodal Hybrid Neural Network

GUO Ye-Cai^{1,3}, WANG Meng-Jie^{1,2}, MAO Xiang-Nan¹, HU Xiao-Wei¹

¹(School of Electronics & Information Engineer, Nanjing University of Information Science & Technology, Nanjing 210044, China)

²(Tianchang Research Institute, Nanjing University of Information Science & Technology, Chuzhou 239399, China)

³(School of Electronics and Information Engineer, Wuxi University, Wuxi 214105, China)

Abstract: To improve the accuracy and robustness of modulation recognition, this study proposes an improved bimodal hybrid modulation recognition model. The model incorporates both the original time-domain in-phase and quadrature (I/Q) data, as well as amplitude and phase (A/P) format data to explore the spatiotemporal correlations within the signal. A two-branch symmetric structure is applied to further process the A/P data, enabling more effective learning of repetitive patterns while mitigating information redundancy. A bidirectional long short-term memory (BiLSTM) network is introduced to enhance the model's capacity for complex temporal feature extraction. Experimental results demonstrate that the proposed model performs well on the RadioML2016.10A dataset. When the signal-to-noise ratio (SNR) is below -8 dB, the average recognition accuracy surpasses mainstream models by 6%. Within the SNR range of 0 to 18 dB, the average recognition accuracy improves by 2% to 10%, reaching 94.32% at 16 dB. In addition, when applied to the RadioML2016.10B dataset, the model continues to achieve superior performance, attaining a recognition accuracy of 93.91% at 18 dB.

Key words: automatic modulation recognition; deep learning; two-way symmetric; bidirectional extract feature

① 收稿时间: 2024-12-18; 修改时间: 2025-02-12; 采用时间: 2025-03-06; csa 在线出版时间: 2025-06-20
CNKI 网络首发时间: 2025-06-23

自动调制识别 (automatic modulation recognition, AMR)^[1,2]作为非协作通信中的一种关键技术,在先验知识不充分或信号数据有限的情况下,通过对接收信号的处理和分析,最终判断出信号的调制类型,从而为后续的信号处理提供更有价值的信息. AMR 是信号检测与解调之间的关键步骤,传统的 AMR 方法通常分为两大类:基于似然 (likelihood-based, LB)^[3]的方法和基于特征 (feature-based, FB)^[4]的方法. LB 方法依赖于对未知信号的正确建模,并通过多假设检验来进行调制识别. 这种方法在充分了解通道条件的情况下最大化了正确分类的概率,在贝叶斯意义上是最优的. 然而, LB 方法严重依赖于先验知识和参数估计,且计算复杂度. 相比之下, FB 方法的成功取决于信号特征工程的经验和知识,同时在对抗多径衰落的负面影响方面非常可靠. 但 FB 方法存在一个缺点,即生成的特征不够具有代表性,这使得其在处理具有挑战性的调制类型时没有足够的判别能力.

近年来,深度学习的发展为科学技术各个领域的进步做出了巨大贡献,尤其是在自然语言处理^[5]和计算机视觉^[6]领域表现优异,这也很大程度上促进了信号调制分类的研究. 文献[7]提出了一种用于调制信号特征提取和分类的卷积神经网络 (convolutional neural network, CNN),通过将原始 I/Q 模态数据输入网络训练,效果显著优于传统方法. 但是对 8PSK 识别准确率较低,同时模型的计算复杂程度较大. 文献[8]采用了一种基于长短期记忆 (long short-term memory, LSTM) 网络的模型,输入 A/P 模态数据能够有效提升在高 SNR 范围内的识别准确率. 然而,在低 SNR 范围内的识别率较低. 文献[9]研究了基于 CNN-LSTM 的双流输入网络结构,实现对信号的时域 I/Q 模态数据和 A/P 模态数据的并行输入,探索信号的时空相关性. 但是产生了新的问题,16-QAM 和 64-QAM 之间的识别容易混淆. 文献[10]开发了一种双向循环神经网络 (bidirectional recurrent neural network, BiRNN),尽管能够捕获更丰富的上下文信息,但仍然存在与传统 RNN 类似的梯度消失和梯度爆炸问题. 文献[11]利用了具有不对称的卷积核学习在多尺度特征图上保留更多初始残差信息,有效防止了梯度消失的问题. 然而,随着卷积块的增加会导致模型需要更多的可训练参数和更长的推理时间. 文献[12]采用了时空并行编码器 (ConvLSTMAE) 提取时空特征,以较低成本实现更高的分类精度. 值得

注意的是,该方法在一定程度上损失了模型稳定性.

基于上述分析,本文提出一种基于双模态混合神经网络 (dual-modal hybrid neural network, DMHNet) 模型. 具有以下优点:对 I/Q 和 A/P 双模态数据进行并行提取特征,以探索信号的时空相关性;减少卷积层的数目以降低计算参数量,同时更好地学习数据之间的相互关系;引入 BiLSTM 与 LSTM 级联,有效增强对序列信息中长期依赖关系的捕获,进行更全面的时间建模,提升了网络结构之间的信号特征传递和交互性,为调制识别任务提供更强的特征提取和建模能力,进一步提高模型的调制识别率.

1 信号模型和预处理

在单输入单输出 (SISO) 系统模型中,考虑完全频率偏移和时间偏移恢复,并且假设信道是频率非选择性的加性高斯白噪声 (AWGN) 信道时,接收到的信号 $r(t)$ 通常表示为:

$$r(t) = \sum_{i=1}^N h_i(t) \cdot s(t - \tau_i) + n(t) \quad (1)$$

其中, $s(t - \tau_i)$ 是来自发射机的调制信号在第 i 条路径上经过时延 τ_i 后的信号, $h_i(t)$ 是第 i 条路径的信道脉冲响应, $n(t)$ 通常表示加性白色高斯噪声.

无线信号的传输涉及信号的强度和信号的相位变化. 信号数据样本通常会转化为 I/Q 模态作为输入,是为了同时捕获到这两方面的信息. 故接收信号 $r(t)$ 由模数转换器以 $f_s = 1/T_s$ 的采样速率采样 n 次,生成离散时间观测信号 $r[n]$. 它是由同相分量 $r^I \in \mathbb{R}$ 和正交分量 $r^Q \in \mathbb{R}$ 组成,因此离散信号 $r[n]$ 可以表示为:

$$r[n] = r^I[n] + jr^Q[n] \quad (2)$$

对 I/Q 模态数据样本进行归一化处理,并填充零值以确保每一个样本具有相同的长度. 为提升模型对多调制信号的识别能力,同时确保处理流程轻量化,对 I/Q 模态数据进行简单的预处理. 首先,通过极坐标变换将复数域信号转化为 A/P 模态数据,增强不同调制模式的固有可分性,转换公式为:

$$r_A[n] = \sqrt{(r^I[n])^2 + (r^Q[n])^2} \quad (3)$$

$$r^\phi[n] = \arctan(r^Q[n]/r^I[n]) \quad (4)$$

其中, $r_A[n]$ 、 $r^\phi[n]$ 分别为第 n 个采样点的幅度分量和

相位分量。

对转化后的 A/P 模态数据沿时间维度分割为等长片段 (每段 128 采样点), 提取幅度分量作为单通道表征。原始三维张量 $r_{A/P}[n] \in \mathbb{R}^{B \times 128 \times 2}$ (B 为批次大小, 128 为采样点, 2 对应 A/P 双通道) 经切片操作后降维为 $r_A[n] \in \mathbb{R}^{B \times 128}$ 。通过添加通道维度, 将 $r_A[n] \in \mathbb{R}^{B \times 128}$ 重构为 $r'_A[n] \in \mathbb{R}^{B \times 128 \times 1}$ 。并采用与 I/Q 数据一致的归一化处理。对相位分量重复相同操作。将 I/Q 模态数据和预处理后得到的 A/P 单通道模态数据同时作为所提模型的输入, 一方面缓解利用传统单一特征训练网络模型时, 可能会出现信息不足和泛化能力差等问题, 另一方面可以使所提模型学习到更为鲁棒的特征表示, 提高模型的识别性能。

2 DMHNet 网络模型

基于双模态混合神经网络 (dual-modal hybrid neural network, DMHNet) 模型结构, 如图 1 所示。调制的

本质是将信息加载到信号波形的幅度或相位中, I/Q 通道可能存在由于幅度和相位的不平衡导致模型产生分类误差的问题, 故所提模型考虑利用两种模态数据构建双模态数据并行提取的结构, 更全面地捕获信号的时空相关性。并且先前的研究已经表明^[9], I/Q 模态数据和 A/P 模态数据都是映射信号特征的有效方法。所提模型中 ICL-Block 模块利用 I/Q 模态数据作为输入, PCAE-Block 模块利用双路对称结构提取 A/P 模态数据中的特征信息, 并作为 ICL-Block 模块的特征补充, 能够提供与原始复杂信号的幅度和相位相关的有效时间和空间特征, 提高模型对各调制方式的可行性。随后, 通过 FBFC-Block 模块利用双向处理多流输入提取全局瞬时特征, 增强模型对复杂序列信息的特征理解能力和时间建模能力, 从而提高模型的调制识别性能。最后, 输出层采用 11 个神经元组成的 Softmax 激活函数作为分类器输出识别信号的调制方式。接下来将对各模块结构进行阐述。

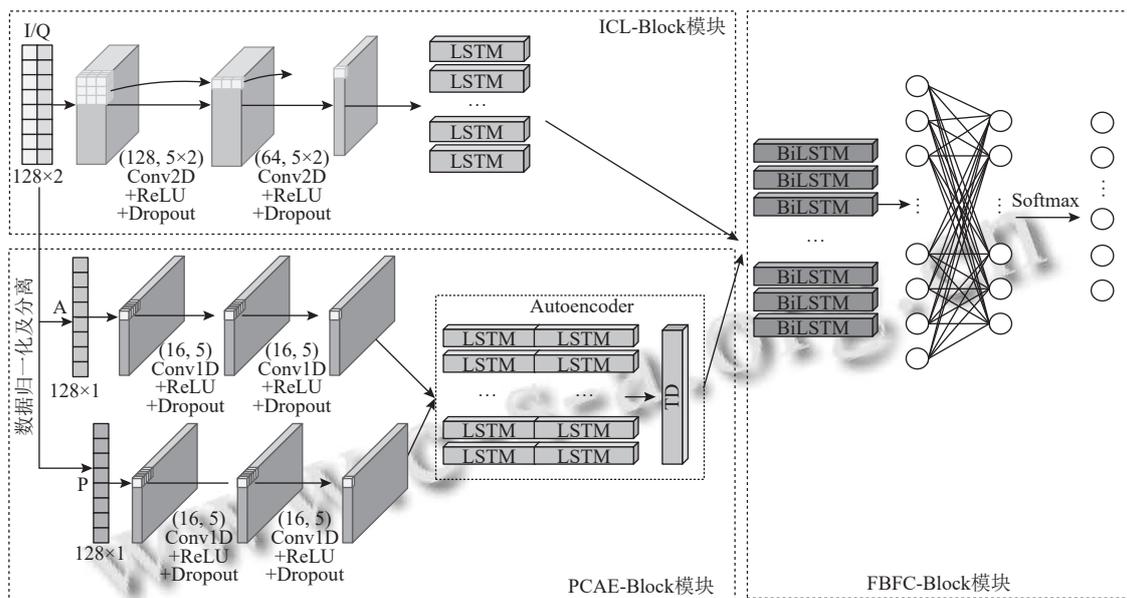


图 1 DMHNet 模型结构

1) ICL-Block 模块

ICL-Block 模块采用两个二维卷积层从原始 I/Q 模态输入信号中学习空间结构和局部相关特征。其中二维卷积的尺寸为 5×2 , 膨胀率为 1, 通道数量分别为 128 和 64。通过级联组合方式, 在不显著增加模型参数的情况下扩大网络感受野, 有助于提取信号的高级空间特征。ICL-Block 模块采用逐层降低通道数量的策

略, 减少计算成本, 使模型在训练过程中不能过于依赖某些特定的神经元, 从而增强模型的泛化能力。与传统卷积^[13]处理相同, 卷积层采用整流线性单元 (ReLU) 作为激活函数。为防止训练过程中模型过度拟合训练数据, 每个二维卷积层后添加一个丢弃率为 0.2 的 Dropout 层, 增强模型在测试数据上的特征表现。将提取的特征按顺序聚合形成与输入特征具有相同大小的输出特征,

并作为 LSTM 层 (50 个隐藏单元) 的输入特征, 是为了捕获空间特征在时间上的依赖性. 同时, 从底层的原始序列信息学习到更复杂的特征表示, 这对于区分高度相似的调制信号尤为重要.

2) PCAE-Block 模块

PCAE-Block 模块由双尺度一维卷积 (dual scale 1D convolution, DSC) 和自编码器 (autoencoder, AE) 按顺序连接而成, 如图 2 所示. 该模块利用 A/P 模态放大不同调制类型之间的差异, 专注于对 A/P 单通道信息进行特征提取, 显著提高信号的可分性. DSC 采用双路对称的 Conv1D+Conv1D 组合形式, 由上下分支独立处理输入数据, 以学习不同的特征表示, 其中 Conv1D 尺寸为 5. DSC 利用层次学习结构充分地提取输入数据中的局部模式和特征, 并捕获浅层特征中更具有代表性的特征信息, 有助于模型对输入数据进行更高级的抽象处理来提取优越的空间特征. 双路特征通过 Concatenate 层融合输出作为 AE 的输入特征. AE 利用 LSTM 层有助于捕捉输入特征中的长短期依赖关系, 增强对时间序列数据的建模能力. 在 AE 中引入时间分布 (temporal distribution, TD) 层, 通过对时间分布特征的学习, 使得模型能够更好地提取出时间维度上的深层次特征, 从而提高表达能力和对时序数据的理解与处理能力. 此外, 该模块具有滤除局部噪声和恢复序列的全局结构作用, 有助于充分交互特征中的局部与全局信息. 值得注意的是, 在每一个 IDCNN 层和 LSTM 层后添加一个丢弃率为 0.2 的 Dropout 层以防止模块过拟合. 该模块作为 ICL-Block 的特征补充模块, 独立处理和提取 A/P 模态数据特征更直观反应调制关键参数, 并通过幅度特征稳定分类, 也缓解因 ICL-Block 模块提取特征信息不充分的问题.

3) FBFC-Block 模块

FBFC-Block 模块由一层 BiLSTM 和两层 64 个神经元的全连接层组成. 通过 Concatenate 层拼接融合 ICL-Block 提取的空间增强特征和 PCAE-Block 捕获的时间动态特征形成时空联合输入表征, 其操作流程如下:

$$F_{FBFC}^{in} = Concat(F_{ICL}^{out}, F_{PCAE}^{out}) \quad (5)$$

其中, Concat 表示融合操作, F^{in} 和 F^{out} 分别表示相应模块的输入和输出特征.

BiLSTM 层由正向和逆向两个独立的 LSTM 组成,

包括门控结构 (遗忘门、输入门、输出门) 和细胞状态, 其结构如图 3 所示.

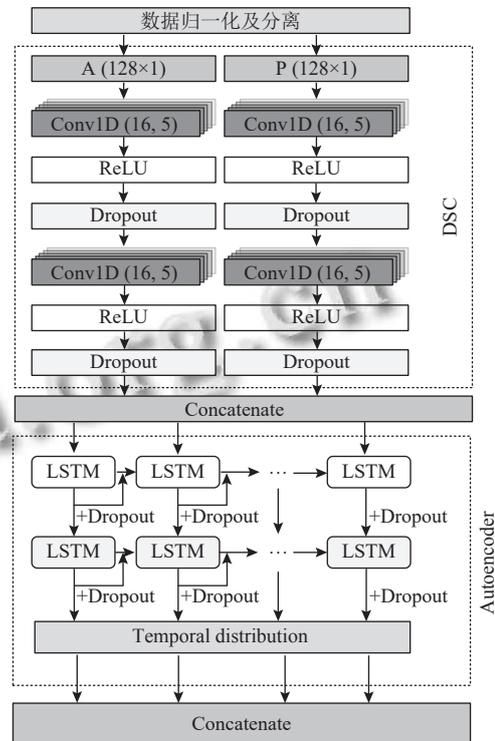


图 2 PCAE-Block 模块结构图

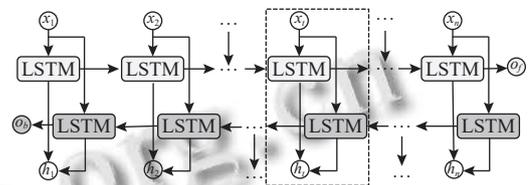


图 3 BiLSTM 结构图

与 LSTM 不同的是, BiLSTM 层同时考虑过去和未来的状态, 通过内部的双向门控结构对输入特征进行滤波, 保留部分特征信息, 进一步捕获信号对时间序列的依赖性.

BiLSTM 层的输出隐藏状态为两个隐藏状态的连接, 即:

$$h_t = [\vec{h}_t, \overleftarrow{h}_t] \quad (6)$$

其中, \rightarrow 和 \leftarrow 分别表示正向和逆向过程, h_t 表示当前时刻的隐藏状态.

为了将特征映射到更可分离的空间, 堆叠两个具有 64 个神经元的全连接层来深化网络模型. 第 1 层全连接层可以提取输入数据的高层次特征, 而第 2 层全连接层可以将这些高层次特征进一步组合和变换, 从

而生成更为抽象和有用的特征表示. 这种逐层抽象的过程有助于模型更好地理解和处理输入数据的局部特征和全局特征. 作为全局时序建模模块, 通过融合多模态特征获取更全面的表征信息, 捕获信号的动态时序依赖关系, 使得信息可以在网络中更深层次地流动和共享, 帮助模型学习到更多的特征交互和关联关系.

3 实验和结果分析

3.1 实验设置

3.1.1 数据集

为验证所提模型性能, 所有实验均是在软件无线平台 GNU Radio 信道模型生成的 RadioML2016.10A

和 RadioML2016.10B 两个知名开源基准数据集^[14]上进行训练和测试. RadioML2016.10A 数据集包括 11 种常用的调制信号, 如 8PSK、WBFM、AM-DSB、AM-SSB 等. 数据集中的信号包括多种实际无线环境中的典型损伤, 如功率延迟分布、本振偏移和频率选择性衰落等. 目的是提高模型在实际场景中的处理性能. 数据集覆盖[-20, 18] dB 的 SNR, 增量为 2 dB. 每种调制方式下每个 SNR 有 1000 个样本, 共计 220000 个调制信号. RadioML2016.10B 数据集是 RadioML2016.10A 数据集的扩展版本, 但不包括 AM-SSB 调制信号, 可以有效避免因数据集不足而导致的过拟合问题. 数据集的具体参数信息见表 1.

表 1 数据集参数

数据集名	识别种类	识别类型	样本维度	样本数量	SNR范围 (dB)	信道环境
RadioML2016.10A	11	8PSK, BPSK, CPFSK, GFSK, PAM4, AM-DSB, AM-SSB, 16-QAM, 64-QAM, QPSK, WBFM	2×128	220000	[-20, 18]	加性高斯白噪声、功率延迟分布、本振偏移、
RadioML2016.10B	10	8PSK, BPSK, CPFSK, GFSK, PAM4, AM-DSB, 16-QAM, 64-QAM, QPSK, WBFM	2×128	1200000	[-20, 18]	频率选择性衰落、多径传播等

3.1.2 参数设置

所有实验都是在 NVIDIA GeForce GTX 1080Ti 上进行, 搭建神经网络的深度学习库是 Keras, 后端选择 TensorFlow. 训练过程中经逐步实验, 批量大小为 400 时最优; 使用自适应矩阵估计 (Adam) 优化器用于更新参数. 初始学习速率为 10^{-3} , 模型使用均方差 (mean squared error) 作为损失函数进行训练, 当验证损失在 10 个 epoch 内未降低, 学习速率调整为原来速率的 0.5 倍, 以帮助模型更快的逼近最优解. 此外, 当验证损失在 30 个 epoch 内没有发生降低时, 利用早停机制避免模型出现过拟合现象. 同时, 使用具有最小验证损失的训练模型来预测测试信号样本的调制类型, 有助于提高模型的泛化能力和鲁棒性, 减少过拟合风险, 优化模型性能.

3.1.3 评估指标

为了对所提模型进行客观有效的评价, 与不同模型进行公平对比, 采用识别率和混淆矩阵两个指标进行评估.

1) 识别率

识别率是通过计算正确分类的信号数量与总信号数量的比率来衡量, 采用百分比来表示. 较高的识别率意味着系统能够更准确地识别和区分不同类型的调制

信号. 识别率越高, 说明模型的分类效果越好. 通常使用以 SNR 为横轴、识别率为纵轴的曲线图可视化不同的 SNR 下的识别率.

$$acc = \frac{tp + tn}{tp + tn + fp + fn}$$

其中, tp 表示标记为正确的样本和正确的模型预测, tn 表示标记为错误的样本但错误的模型预测, fp 表示标记为错误的样本和正确的模型预测, fn 表示标记为正确但错误的模型预测.

2) 混淆矩阵

混淆矩阵以矩阵形式提供模型在数据集上分类结果的详细信息, 展示模型在各个类别上的正确分类和错误分类情况. 矩阵横轴表示预测样本调制类型, 纵轴表示真实样本调制类型. 在混淆矩阵中, 对角线元素表示正确分类的样本数, 非对角线元素则表示错误分类的样本数. 从而更好地理解神经网络模型, 指导模型的进一步改进.

3.2 实验结果分析

3.2.1 对比实验

为了验证所提模型 DMHNet 对于调制信号识别的准确性, 将其与 Double-CNN-LSTM^[15]、IC-AMCNet^[16]、BMCCLDNN^[17]、CNN-LSTM^[9]、CNN2^[18]这 5 种网络结

构进行对比验证. 图4展示了6种模型在数据集 RadioML2016.10A 和 RadioML2016.10B 下不同 SNR 的识别率.

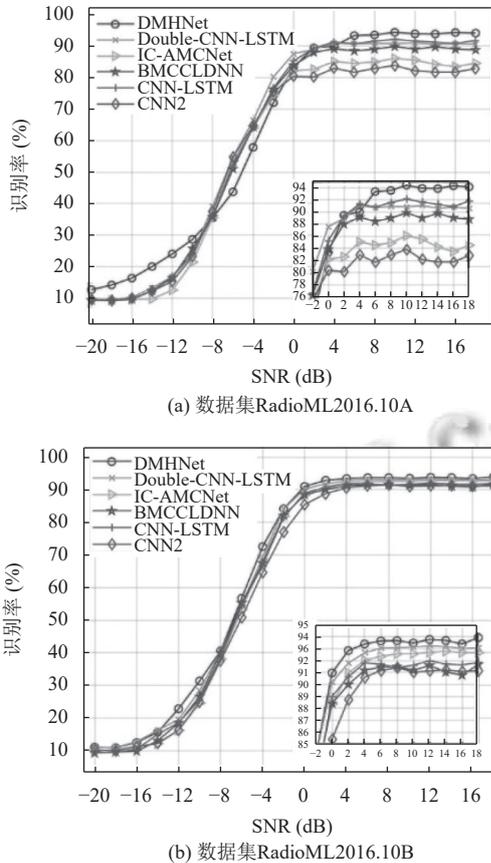


图4 DMHNet模型与主流模型在RadioML2016.10A和RadioML2016.10B数据集上的识别率

图4(a)中结果表明,以单模态数据作为输入的CNN2和IC-AMCNet网络结构的识别率最高仅为86.09%,识别效果明显逊色.而以双模态数据输入的其余模型识别效果优异.这表明,仅依赖一种数据模态容易导致部分原始特征信息的丢失,从而影响识别性能.相比于Double-CNN-LSTM、BMCCLDNN和CNN-LSTM这3种采用I/Q和A/P两种模态数据进行时空特征提取并融合处理的主流网络结构,DMHNet不仅采用了两种模态数据的输入形式,还通过DSC的双路对称结构充分提取了A/P模态数据中的特征信息,同时利用双向多层结构更为有效地提取了双向时序特征,从而进一步提升了模型的识别性能.在SNR为16 dB情况下,本模型的识别率高达94.32%.

表2对比所提模型与主流模型在数据集RadioML2016.10A上不同SNR下的识别率.结果显示,在SNR

为[-20, -12] dB的范围内,本模型比主流模型最大增幅分别为3.32%、6.77%和11.50%.在SNR为[8, 18] dB的范围内,最大增幅分别为10.64%、11.72%、12.54%和11.27%.值得注意的是,在SNR为[-8, 4] dB范围内时,本模型识别率低于主流模型.这是由于样本不足导致模型对特定模式发生过拟合现象.因此,在RadioML-2016.10B数据集上重复上述实验,如图4(b)所示.结果表明,在SNR为[-8, 4] dB范围内模型的识别能力优于主流模型,说明模型对训练样本丰富度有一定要求.同时,在SNR为[0, 18] dB的范围内,识别率超过90.97%,平均识别率超过93.26%.

表2 DMHNet模型与主流模型在不同SNR下识别率(%)

SNR (dB)	DMHNet	Double-CNN-LSTM	IC-AMCNet	BMCC-LDNN	CNN-LSTM	CNN2
-20	12.63	9.72	9.50	9.36	9.31	9.36
-16	16.31	10.22	9.54	9.68	9.54	9.81
-12	24.00	16.68	12.50	16.36	16.90	15.27
-8	35.50	39.13	35.54	35.81	36.13	37.54
-4	57.86	66.45	64.00	64.31	64.13	64.72
0	83.36	87.59	82.22	83.86	85.31	80.40
4	90.00	91.09	85.09	89.13	91.31	82.95
8	93.54	90.90	84.95	89.00	91.59	82.90
12	93.90	91.04	85.54	89.00	91.63	82.18
16	94.32	90.77	83.50	89.04	90.86	81.77
18	94.13	90.63	84.54	88.81	91.81	82.86

从表3可看出, BMCCLDNN模型通过简化网络结构,减少了计算开销,但在本实验中识别率低于所提出的DMHNet模型.

表3 DMHNet模型与主流模型的计算复杂度对比

模型	参数数量	收敛迭代次数	每轮训练时间(s)
DMHNet	167667	140	28
Double-CNN-LSTM	1118743	146	54
IC-AMCNet	1264011	262	6
BMCCLDNN	160569	108	14
CNN-LSTM	1273443	165	58
CNN2	858123	106	5

结合图4(a)分析知CNN2模型在各个指标上均表现出良好的性能,特别是在较低SNR范围内具有较好的识别精度,但是高SNR范围内识别能力不及主流模型.相比之下,Double-CNN-LSTM和CNN-LSTM在SNR较高的情况下表现优异,但由于模型复杂度较高,导致计算开销过大,每轮训练时间较长,效率最低.另一方面,IC-AMCNet模型的收敛迭代次数达到了262轮次,说明模型收敛性能有待提高.通过对比参数量、收敛次数和每轮训练时间3个方面来分析模型的计算

复杂度, DMHNet 模型计算开销较小且收敛迭代次数适中, 整体性能表现平衡, 各个指标都在可接受范围内, 具有较好的实际应用价值. 这也验证了一个观点, 提高模型性能不仅依赖于增加参数量, 还需要有效利用数据复杂性和增强模型的特征提取能力.

3.2.2 消融实验

为了验证所提 DMHNet 模型结构的有效性, 分别移除 ICL-Block 模块、PCAE-Block 模块、FBFC-Block 模块进行消融实验, 分别命名为 DMHNet-A、DMHNet-B 和 DMHNet-C, 如表 4 所示, 其中√表示各模块的使用情况.

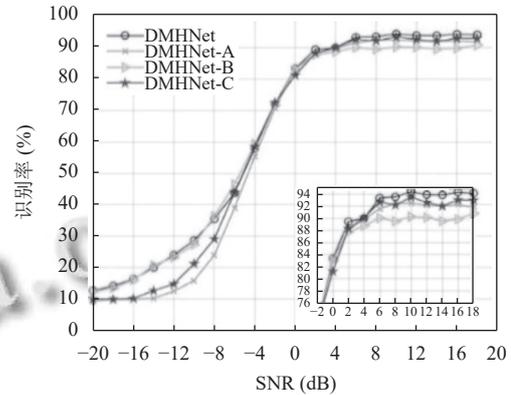
表 4 DMHNet 模型与变体模型

变体	ICL-Block 模块	PCAE- Block模块	FBFC- Block模块	平均精度 (%)
DMHNet-A	—	√	√	58.12
DMHNet-B	√	—	√	60.83
DMHNet-C	√	√	—	59.55
DMHNet	√	√	√	62.26

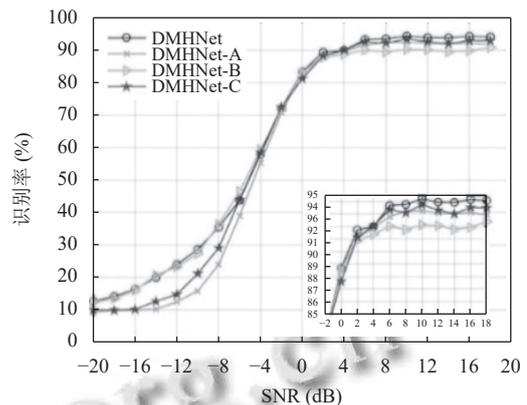
通过表 4 对变体模型的统计结果和图 5(a) 绘制出的 DMHNet 模型与变体模型在数据集 RadioML2016.10A 上不同 SNR 下的识别率显示, 变体 A 与所提模型相比, 得出结论 ICL-Block 模块使模型的平均精度提升了 4.14%, 主要得益于其从原始复杂信息中提取有效的时空特征, 使分类识别性能提升. 这进一步说明了 ICL-Block 模块在捕获和处理时空特征方面对模型性能提升的重要性. 与变体 B 相比, 所提模型中的 PCAE-Block 模块通过提供额外的幅度和相位信息有效地增强了特征表达的多样性和复杂性, 高信噪比范围的识别率明显提升. 因此, 证实了 PCAE-Block 模块作为特征补充的重要性. 另外, 通过对变体 A、变体 B 和变体 C 的分析可观察到, 当移除 ICL-Block 模块或 PCAE-Block 模块时, FBFC-Block 模块仅对单一模态特征进行建模提取双向时序特征, 模型的识别精度显著下降, 平均精度最高仅为 60.83%. 这说明了时空特征联合优化的必要性, 即单一模态难以充分捕捉空间局部性与时间连续性的交互关系. 其次, 变体 C 的平均精度为 59.55%, 这表明简单的特征拼接无法代替 FBFC-Block 模块的深度时空融合能力. 通过 BiLSTM 与全连接层发级联优化, 所提模型能够有效建模跨模态依赖, 提升关键特征的判别性.

图 5(b) 显示了 DMHNet 模型与其他变体模型在

RadioML2016.10B 数据集上的泛化性实验, 所提模型在不同 SNR 下均表现出色, 尤其是在高 SNR 范围内 (8 dB 及以上), 其识别率超过了其他变体模型. 这表明所提模型在设计上更好地捕捉了信号中的特征, 具有更强的特征提取能力和泛化能力.



(a) 数据集RadioML2016.10A

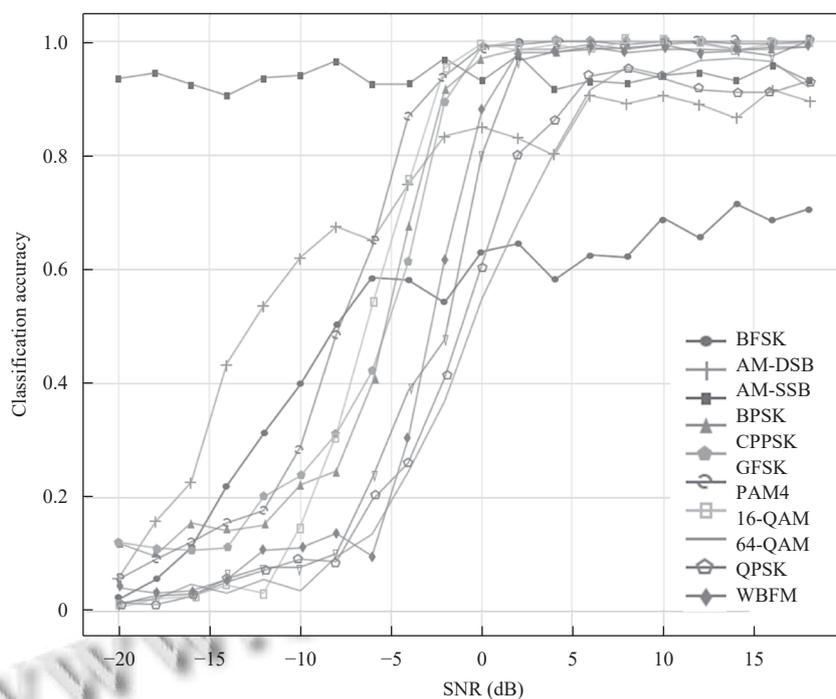


(b) 数据集RadioML2016.10B

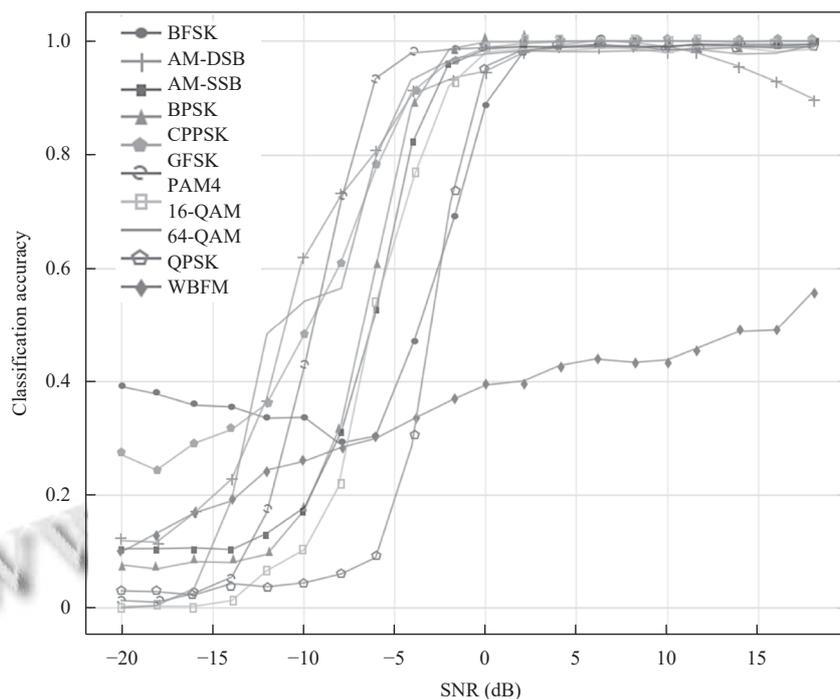
图 5 DMHNet 模型与变体模型在 RadioML2016.10A 和 RadioML2016.10B 数据集上的识别率

3.2.3 不同信噪比条件下各种调制信号的识别性能分析

DMHNet 网络模型在 RadioML2016.10A 数据集上对不同调制信号的识别结果如图 6(a) 所示, 所提模型对 AM-SSB 调制信号识别效果最好, 在全信噪比下有 90% 以上的识别精度; 8PSK、BPSK、CPFSK、GFSK、PAM4 调制信号在 SNR 为[0, 20] dB 范围内识别率接近 100%; AM-DSB、QPSK、16-QAM、64-QAM 调制信号在 SNR 为[0, 20] dB 范围内识别效果良好, 识别率在 80% 左右; 所提模型对 WBFM 调制信号识别能力有待提升, 在 SNR 为[0, 20] dB 范围内识别率最高仅为 70.5%. 同样, 在 RadioML2016.10B 数据集上, 所提模型在 SNR 为[0, 20] dB 范围内, 11 种调制信号中有 9 种达到了 90% 以上, 7 种接近 100%. 如图 6(b) 所示.



(a) 数据集RadioML2016.10A



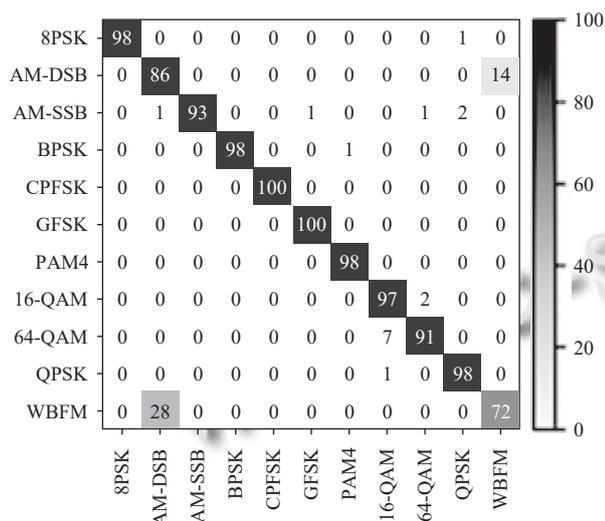
(b) 数据集RadioML2016.10B

图6 模型在不同信噪比下对不同调制信号的识别率曲线

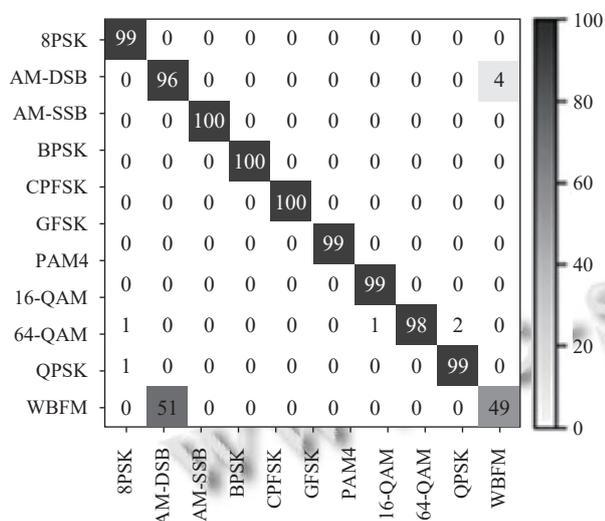
11种调制信号在SNR为14 dB时的混淆矩阵如图7所示。结果表明,除WBFM和AM-DSB外,其他调制方式都能在本模型上达到90%以上的识别精度。尤其值得注意的是,16-QAM和64-QAM调制之间的

区分具有挑战性,因为前者是后者的子集^[19]。这种关系使得在识别时容易出现混淆,难以准确地区分这两种调制方式。所提模型能够有效捕捉到这两种调制方式的细微差异,从而对这两种高阶调制信号能保持高精

度的识别. 另外, 模型将大量宽带频率调制 (WBFM) 信号识别为双边带调幅 (AM-DSB) 信号, 这是因为两者是通过采样模拟音频信生成的. 音频信号存在静默期, 期间信号中仅有载波音^[20]. 且 WBFM 和 AM-DSB 在频谱特征上有部分重叠, 特别是在信号能量较低或噪声较大的情况下, 区分两者变得更加困难.



(a) 数据集RadioML2016.10A



(b) 数据集RadioML2016.10B

图7 DMHNet 模型在 SNR 为 14 dB 时的混淆矩阵

4 结束语

本文提出了一种基于双模态混合神经网络的 AMR 方法. 该方法同时采用 I/Q 和 A/P 双模态数据, 并通过信号预处理简化了信号特征的复杂性. 具体而言, ICL-Block 模块利用 I/Q 模态数据对原始复杂信号进行特

征提取; 而 PCAE-Block 模块采用双路对称结构处理 A/P 模态数据并展示了强大的特征提取能力, 增强了对信号幅度和相位信息的捕捉. 后者为前者的特征补充, 强化了从两种表示中提取的不同特征之间的相互作用, 这种特征互补的设计显著提升了识别性能. 此外, FBFC-Block 模块通过双向时序特征提取能力, 有效捕获了信号的时序动态特征, 提升了模型对复杂时序信息的理解能力, 进一步改善了模型在低 SNR 环境下的识别效果.

实验结果显示, 所提模型不仅显著降低了网络复杂度, 而且增强了对时空特征的提取能力, 有效提升了模型的识别性能. 尤其是在低信噪比条件下, 模型的识别准确率明显优于其他主流识别算法, 展示了其在实际应用中的潜力和优势.

参考文献

- Alnajjar KA, Ghunaim S, Ansari S. Automatic modulation classification in deep learning. Proceedings of the 5th International Conference on Communications, Signal Processing, and their Applications (ICCSPA). Cairo: IEEE, 2022. 1-5.
- 张茜茜, 王禹, 林云, 等. 基于深度学习的自动调制识别方法综述. 无线电通信技术, 2022, 48(4): 697-710. [doi: 10.3969/j.issn.1003-3114.2022.04.017]
- Zheng JP, Lv YF. Likelihood-based automatic modulation classification in OFDM with index modulation. IEEE Transactions on Vehicular Technology, 2018, 67(9): 8192-8204. [doi: 10.1109/TVT.2018.2839735]
- Tunze GB, Huynh-the T, Lee JM, et al. Sparsely connected CNN for efficient automatic modulation recognition. IEEE Transactions on Vehicular Technology, 2020, 69(12): 15557-15568. [doi: 10.1109/TVT.2020.3042638]
- Otter DW, Medina JR, Kalita JK. A survey of the usages of deep learning for natural language processing. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(2): 604-624. [doi: 10.1109/TNNLS.2020.2979670]
- Jeyaraj PR, Asokan SP, Kathiresan AC. Views of deep learning algorithm applied to computer vision knowledge discovery. National Academy Science Letters, 2022, 45(6): 561-566. [doi: 10.1007/s40009-022-01157-z]
- O'Shea TJ, Corgan J, Clancy TC. Convolutional radio modulation recognition networks. Proceedings of the 17th International Conference on Engineering Applications of Neural Networks. Aberdeen: Springer, 2016. 213-226.

- 8 Rajendran S, Meert W, Giustiniano D, *et al.* Deep learning models for wireless signal classification with distributed low-cost spectrum sensors. *IEEE Transactions on Cognitive Communications and Networking*, 2018, 4(3): 433–445. [doi: [10.1109/TCCN.2018.2835460](https://doi.org/10.1109/TCCN.2018.2835460)]
- 9 Zhang ZF, Luo H, Wang C, *et al.* Automatic modulation classification using CNN-LSTM based dual-stream structure. *IEEE Transactions on Vehicular Technology*, 2020, 69(11): 13521–13531. [doi: [10.1109/TVT.2020.3030018](https://doi.org/10.1109/TVT.2020.3030018)]
- 10 Schuster M, Paliwal KK. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 1997, 45(11): 2673–2681. [doi: [10.1109/78.650093](https://doi.org/10.1109/78.650093)]
- 11 Huynh-The T, Hua CH, Pham QV, *et al.* MCNet: An efficient CNN architecture for robust automatic modulation classification. *IEEE Communications Letters*, 2020, 24(4): 811–815. [doi: [10.1109/LCOMM.2020.2968030](https://doi.org/10.1109/LCOMM.2020.2968030)]
- 12 Shi YH, Xu H, Jiang L, *et al.* ConvLSTMAE: A spatiotemporal parallel autoencoders for automatic modulation classification. *IEEE Communications Letters*, 2022, 26(8): 1804–1808. [doi: [10.1109/LCOMM.2022.3179003](https://doi.org/10.1109/LCOMM.2022.3179003)]
- 13 Xu JL, Luo CB, Parr G, *et al.* A spatiotemporal multi-channel learning framework for automatic modulation recognition. *IEEE Wireless Communications Letters*, 2020, 9(10): 1629–1632. [doi: [10.1109/LWC.2020.2999453](https://doi.org/10.1109/LWC.2020.2999453)]
- 14 O'Shea TJ, West N. Radio machine learning dataset generation with GNU radio. *Proceedings of the 6th GNU Radio Conference*. Boulder: GNURadio, 2016. 213–226.
- 15 皮爽. 基于数据增强与卷积神经网络的信号调制识别算法研究 [硕士学位论文]. 天津: 天津理工大学, 2023.
- 16 Hermawan AP, Ginanjar RR, Kim DS, *et al.* CNN-based automatic modulation classification for beyond 5G communications. *IEEE Communications Letters*, 2020, 24(5): 1038–1041. [doi: [10.1109/LCOMM.2020.2970922](https://doi.org/10.1109/LCOMM.2020.2970922)]
- 17 Pathak NK, Bajaj V. Automatic modulation classification using bimodal parallel multichannel deep learning framework for spatial multiplexing MIMO system. *Physical Communication*, 2023, 59: 102071. [doi: [10.1016/j.phycom.2023.102071](https://doi.org/10.1016/j.phycom.2023.102071)]
- 18 Tekbıyık K, Ekti AR, Görçin A, *et al.* Robust and fast automatic modulation classification with CNN under multipath fading channels. *Proceedings of the 91st IEEE Vehicular Technology Conference*. Antwerp: IEEE, 2020. 1–6.
- 19 Yang HG, Zhao LZ, Yue GX, *et al.* IRLNet: a short-time and robust architecture for automatic modulation recognition. *IEEE Access*, 2021, 9: 143661–143676. [doi: [10.1109/ACCESS.2021.3121762](https://doi.org/10.1109/ACCESS.2021.3121762)]
- 20 Ke ZQ, Vikalo H. Real-time radio technology and modulation classification via an LSTM auto-encoder. *IEEE Transactions on Wireless Communications*, 2022, 21(1): 370–382. [doi: [10.1109/TWC.2021.3095855](https://doi.org/10.1109/TWC.2021.3095855)]

(校对责编: 张重毅)