

文章编号:1671-6833(2026)03-0076-07

基于BLR并行结构的多模态调制识别方法

江桦,肖科杰,胡坡,巩克现,赵振禹

(郑州大学 电气与信息工程学院,河南 郑州 450001)

摘要:针对现有基于卷积神经网络(CNN)的调制识别方法对单一模态数据(如IQ序列)依赖性强、难以充分提取信号多维特征等问题,提出了一种基于双向长短时记忆网络(BiLSTM)和残差网络(ResNet)的多模态并行结构调制识别方法(BLR网络)。首先,通过上支路的BiLSTM提取IQ数据的时序特征,通过下支路的ResNet-18提取星座图的空间特征;其次,在决策融合模块采用串行特征融合,更好地挖掘多模态数据的互补性;最后,借助模型的特征提取能力对信号调制样式进行识别,并在公开数据集RML2018.01a上进行了实验验证。实验结果表明:BLR网络在6~30 dB信噪比区间内的整体识别准确率稳定在96.48%,相较于单一模态的ResNet和BiLSTM模型分别提升了2.61%和3.91%,相较于并联结构的CNN-LSTM模型提高了1.25%,验证了所提模型在调制识别问题上的有效性。

关键词:自动调制识别;卷积神经网络;多模态;特征融合;并联结构

中图分类号:TN911.7 **文献标志码:**A **doi:**10.13705/j.issn.1671-6833.2025.03.019

自动调制识别(automatic modulation recognition, AMR)是通信侦察和电子对抗中的关键技术,旨在未知信号样式和缺乏先验信息的条件下,自动识别目标信号的调制方式。传统AMR方法主要分为两类:基于假设检验的最大似然方法^[1]和基于特征提取的模式识别方法^[2]。前者依赖大量先验信息,计算复杂度高,难以适应复杂电磁环境;后者虽理论成熟,但在低信噪比(signal-to-noise ratio, SNR)条件下识别精度有限。尽管这些方法在某些场景下表现良好,但其对先验知识和人工特征的依赖限制了其在实际应用中的鲁棒性和泛化能力。

近年来,深度学习^[3](deep learning, DL)在图像分类^[4-5]、自然语言处理^[6]等领域的成功为AMR提供了新的解决方案。基于深度学习的AMR能够自动提取信号特征,显著降低对先验信息的依赖,同时在复杂电磁环境中展现出优越的识别性能。2016年O'Shea等^[7]率先提出基于CNN(convolutional neural network)的模型,利用RML2016.10a数据集中信号的I/Q特性进行调制识别,打破了传统调制识别方法的固有技术路线。随后,Wu等^[8]利用长短时记忆网络(long short-term memory network, LSTM)提取信号的时间序列特性,在CNN的基础上

引入了LSTM,并取得了更好的性能。为了在复杂的调制方式下进一步提高精度,Wang等^[9]设计了一种基于星座图的模型,在低信噪比下也能实现较好的识别效果。王晓明等^[10]采用ResNet在公开数据集上取得约95%的分类效果。章昕亮等^[11]引入双向长短期记忆网络(BiLSTM)更好地提取样本的时序特征,从而提高了多径衰落信道下信号的识别率。查雄等^[12]利用多端卷积神经网络将电磁信号转换成眼图和矢量图进行学习,在低信噪比条件下,取得了较好的识别性能。任彦洁等^[13]提出了一种基于时间卷积网络的调制识别方法,解决了端到端调制识别方法下识别率较低和神经网络参数数量大的问题。West等^[14]提出CLDNN混合架构,通过CNN、LSTM和DNN的级联设计实现了时空特征的联合提取。翁建新等^[15]提出了一种CNN-LSTM并联的调制样式识别算法,该算法将I路和Q路数据作为模型输入,通过上下两条支路分别提取信号的空间特征和时间特征,从而提高模型的特征表达能力。

尽管上述研究在提高识别精度方面取得了成果,但大多数方法仍主要依赖单一的IQ^[16]数据,并未充分利用信号中的多模态信息。随着多模态数据融合的发展,研究者们逐渐认识到不同模态间的互

收稿日期:2025-10-16;修订日期:2025-11-24

基金项目:国家自然科学基金资助项目(61901417)

作者简介:江桦(1956—),男,江苏南通人,郑州大学教授,博士,博士生导师,主要从事通信信号处理、信号检测及电子对抗等研究,E-mail:jh653@sina.com。

引用本文:江桦,肖科杰,胡坡,等.基于BLR并行结构的多模态调制识别方法[J].郑州大学学报(工学版),2026,47(3):76-82,116.(JIANG H, XIAO K J, HU P, et al. Multimodal modulation recognition method based on BLR parallel structure[J]. Journal of Zhengzhou University (Engineering Science), 2026, 47(3): 76-82, 116.)

互补性对提升识别性能的重要性。例如,张富强等^[17]利用多模态数据融合的方法,解决了动态手势识别过程中的信息缺乏导致的识别准确率难以提高的问题。Qi 等^[18]提出了基于 ResNet 的波形-频谱多模态融合方法,相较于传统的基于 CNN 的单模态方法取得了显著的性能提升,并在 256QAM 和 1024QAM 等高阶数字调制类型下取得了较好的效果。

信号表示大致可以分为 3 类^[19]:基于专家特征表示(如高阶统计量特征、频谱特征)、序列表示(如 IQ 序列、振幅相位序列)和图像表示(如星座图、时频图、循环谱图)。其中,IQ 数据包含信号的相位和幅度信息,而星座图能够直观反映信号的噪声情况和调制方式分布差异。因此,本文选取 IQ 数据和星座图作为多模态输入,以便充分挖掘不同模态的互补性。

基于以上背景,本文提出了一种基于双向长短时记忆网络和残差网络并联结构的多模态调制识别方法(BiLSTM-ResNet,简称为 BLR 网络)。该方法将信号的同向分量(I 路)、正交分量(Q 路)和星座图数据并行输入 BiLSTM 和 ResNet 网络,BiLSTM 通过双向结构捕捉 IQ 数据的时序信息,并利用遗忘门优化梯度消失问题;ResNet 通过卷积层提取星座图中的几何特征和分布模式。BiLSTM 从时域动态变化中提取时序特征,ResNet 从静态的几何分布中获取空间特征,加强了一维和二维数据的不同层次特征的互补性。最后,将两支路输出的特征向量经过串行融合,并送入全连接层,实现了对常见调制信号的自动调制识别。

1 系统模型与问题分析

基于 DL 的 AMR 系统架构^[18]如图 1 所示。发射端的信号 $x_{sig}(t)$ 经过调制后得到 $s_{sig}(t)$, $s_{sig}(t)$ 经过信道时受到衰落和噪声干扰的影响,接收端接收到的信号 $r_{sig}(t)$ 可以表示为

$$r_{sig}(t) = s_{sig}(t) * h_{sig}(t) + n_{sig}(t) \quad (1)$$

式中: $h_{sig}(t)$ 代表信道冲激响应; $n_{sig}(t)$ 代表加性噪声;* 表示卷积操作。

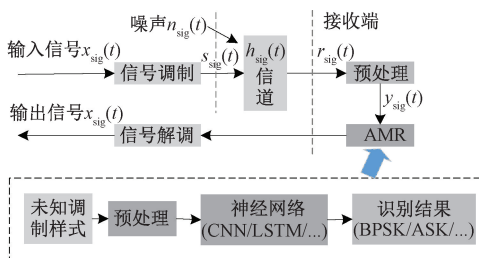


图 1 基于 DL 的 AMR 系统架构

Figure 1 AMR system architecture based on DL

在预处理阶段,经过参数估计和下变频处理后,可以获得基带信号 $y_{sig}(t)$, 可以表示为

$$y_{sig}(t) = y_{sigI}(t) + jy_{sigQ}(t) \quad (2)$$

式中: $y_{sigI}(t)$ 表示基带信号的 I 路信号; $y_{sigQ}(t)$ 表示基带信号的 Q 路信号; $y_{sig}(t)$ 表示基带信号的复数形式。

I、Q 两路既可以转换为复数表达形式,也可以捕获和表征信号的幅度和相位信息,可以表示为

$$A_{sig}(t) = \sqrt{y_{sigI}^2(t) + y_{sigQ}^2(t)} \quad (3)$$

$$P_{sig}(t) = \arctan \frac{y_{sigQ}(t)}{y_{sigI}(t)} \quad (4)$$

式中: $A_{sig}(t)$ 表示信号瞬时幅度; $P_{sig}(t)$ 表示信号瞬时相位。

2 多模态下的 BLR 并联网络结构

本文所提 BLR 网络模型图如图 2 所示,主要分为样本预处理、网络框架和决策融合 3 个部分。

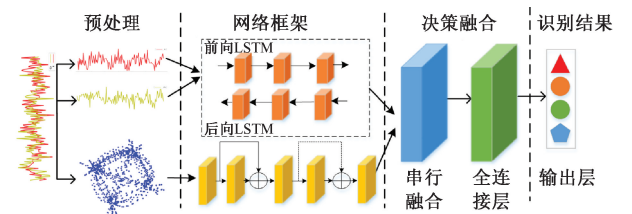


图 2 BLR 网络模型图

Figure 2 BLR network model diagram

预处理模块主要用于完成数据输入网络前的准备工作。每个样本为 2×1024 大小的矩阵,其中 2 代表 I、Q 两路;1024 代表每个样本在 I 路和 Q 路上的采样点数目。将输入的原始数据直接输入上支路,经过星座映射转换为星座图后输入下支路。

网络框架由双向长短期记忆网络(BiLSTM)和残差网络(ResNet)并联组成。上支路的 BiLSTM 网络实现了对原始 IQ 数据在时间序列上的特征提取。BiLSTM 网络结构图如图 3 所示。

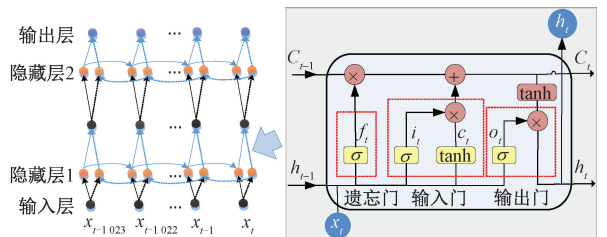


图 3 BiLSTM 网络结构图

Figure 3 BiLSTM network structure diagram

BiLSTM 是一种带有正反向连接的 LSTM 网络,研究表明,针对长序列建模任务,循环神经网络在捕捉双向时序依赖方面仍具有独特优势^[20]。正向 LSTM 按时间步从左到右处理输入序列,LSTM 的每一个隐藏层都包括遗忘门、输入门、输出层 3 个门结

构^[21]。遗忘门可以控制历史信息的输入,滤掉没用的信息。输入门可以通过对细胞状态进行添加或移除信息,从而有选择性地更新细胞状态。输出门可以决定当前时间下输出的隐藏状态 h_t 。LSTM 的具体计算过程为

$$f_t = \sigma(\mathbf{W}_f \cdot [h_{t-1}, x_t] + \mathbf{b}_f); \quad (5)$$

$$i_t = \sigma(\mathbf{W}_i \cdot [h_{t-1}, x_t] + \mathbf{b}_i); \quad (6)$$

$$\tilde{c}_t = \tanh(\mathbf{W}_c \cdot [h_{t-1}, x_t] + \mathbf{b}_c); \quad (7)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t; \quad (8)$$

$$o_t = \sigma(\mathbf{W}_o \cdot [h_{t-1}, x_t] + \mathbf{b}_o); \quad (9)$$

$$h_t = o_t \odot \tanh c_t. \quad (10)$$

式中: f_t, i_t, o_t 分别表示遗忘门、输入门、输出门的输出; σ 表示 Sigmoid 激励函数; \mathbf{W}_f 表示用于连接上一时刻隐藏状态 h_{t-1} 和当前输入 x_t 的权重矩阵; $[h_{t-1}, x_t]$ 表示将上一时刻隐藏状态和当前输入拼成向量; $\mathbf{W}_i, \mathbf{W}_c, \mathbf{W}_o$ 分别表示输入门、备选状态、输出门的权重矩阵; h_t 表示当前时刻的隐藏状态; $\mathbf{b}_f, \mathbf{b}_i, \mathbf{b}_o, \mathbf{b}_c$ 分别表示遗忘门、输入门、输出门、备选状态的偏置向量; \tanh 为双曲正切激活函数; \tilde{c}_t 表示当前时刻用来备选的更新内容; c_t 和 c_{t-1} 分别表示当前时刻的单元状态和上一时刻的单元状态; \odot 表示逐元素乘法。

反向 LSTM 按时间步从右到左处理输入序列,每个时间步的隐藏状态 h'_t 和单元状态 c'_t 可以由类似的公式计算,而 BiLSTM 的输出由 2 个方向的隐藏状态拼接而成,如式(11)所示。

$$y_t = [h_t, h'_t]. \quad (11)$$

下支路为多个残差块拼接而成的 ResNet-18 网络,用于获取信号的空间特征。如图 4 所示。

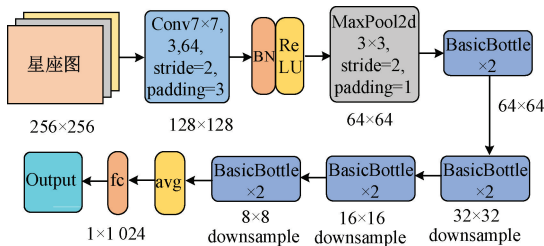


图 4 ResNet-18 模型结构图

Figure 4 ResNet-18 model structure

ResNet 通过在不同层之间添加捷径连接 (shortcut) 形成残差块 (residual block) 实现了模型的训练速度的提升,并解决了网络层数加深时出现的网络退化问题^[22]。加入批量规范化层 (batch normalization, BN) 在一定程度上解决了梯度消失和梯度爆炸的问题。ResNet 结构如图 5 所示。

ResNet-18 网络由多个残差块堆叠而成,每个残差块后有一个激活函数,从而使得网络学习到恒等

映射,并避免了网络过深而造成的梯度消失的问题。残差块的详细结构如图 5 所示。

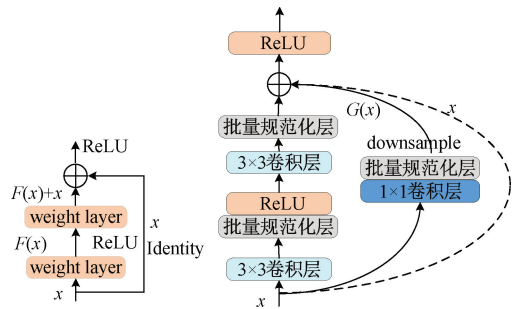


图 5 ResNet 结构

Figure 5 ResNet structure

当不需要改变前后输入输出维度的层时,ResNet-18 的残差块和残差映射可以分别表示为

$$H(x) = F(x) + x; \quad (12)$$

$$F(x) = \text{ReLU}(\text{BN}(\text{Conv}(x))). \quad (13)$$

式中: $H(x)$ 表示网络映射关系; $F(x)$ 表示残差映射; x 表示输入; Conv 表示卷积操作; BN 表示批量规范化处理; ReLU 表示激活函数。

当上一层的输出通道不符合当前层所需的输入通道时,则应对残差块做出如下调整。

$$H(x) = \text{ReLU}(\text{BN}(\text{Conv}(x))) + \text{Conv}_{1 \times 1}(x). \quad (14)$$

式中: $\text{Conv}_{1 \times 1}$ 表示对输入信号进行 1×1 卷积操作。

当前的研究中,模态融合主要有 3 个阶段:数据级融合、特征级融合和决策级融合^[23]。本文采用特征级融合和决策级融合来提高信号识别准确率。将原始 IQ 两路信号以 $(N_{\text{all}}, 1, 1024, 2)$ 的维度送入 BiLSTM 网络,将星座图数据以 $(N_{\text{all}}, 3, 256, 256)$ 的维度送入 ResNet 网络 (N_{all} 表示样本数量),决策融合就是将 BiLSTM 和 ResNet 网络输出的特征矩阵 (1×1024) 按特征维度进行拼接,从而形成 (1×2048) 的特征矩阵。串行融合策略借鉴了时空特征联合的思想^[24],通过不同层次特征的互补增强模型对调制信号的判别能力。最后,经过全连接层通过权重参数学习到特征之间的非线性关系,实现信号多种调制方式的分类。

在串行融合前,由于 BiLSTM 和 ResNet 网络的输出结果可能具有不同的尺度和单位,使用了归一化确保不同尺度和单位的特征对模型的影响是均衡的。串行融合后的特征维度的增加,导致模型参数量增大,本文使用正则化方法来防止过拟合。

3 算法仿真及实验分析

3.1 数据集介绍

本文采用了由 DeepSig 公司于 2018 年发布的

RML2018.01a 数据集^[25]。该数据集是经过数字化处理的无线通信信号数据集,模拟了真实通信场景下的调制信号,包含多种调制方式(如 FM, QAM, BPSK, AM 等),并以 I/Q 数据格式存储,适用于调制方式分类的研究。RML2018.01a 数据集相比于 RML2016 数据集涵盖了更多的调制信号类型和更广泛的信噪比范围,以及实际硬件采集的信号数据。

该数据集中的信号是基于 GNU Radio 软件生成,考虑到了信号在无线信道中传输的多种干扰因素,生成的信号共有 24 种不同的调制方式,其中包含了 5 种模拟调制信号 AM-SSB-WC, AM-SSB-SC, AM-DSB-WC, AM-DSB-SC, FM, 19 种数字调制信号 OOK, 4ASK, 8ASK, BPSK, QPSK, 8PSK, 16PSK, 32PSK, 16APSK, 32APSK, 64APSK, 128APSK, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM, GMSK, OQPSK。该数据集包括 2 555 904 个调制信号,具有 26 种不同的信噪比(SNR),范围从 -20 dB 到 30 dB,步长为 2 dB,每种调制方式中每信噪比有 4 096 个信号。数据集每个信号都是由 1 024×2 的 IQ 信号组成。

图 6 是 6 dB 下 24 种调制方式的星座图。其中,横轴表示同相分量 x_1 ,纵轴表示正交分量 y_0 。研究表明,当信噪比低于 6 dB 时,RML2018.01a 的 24 类调制样式在星座图上难以区分,无法进行准确识别。

因此,本文采用了信噪比从 6 dB 到 30 dB,步进为 2 dB,包含 24 种调制方式的新数据集 RML2018.01_

New 进行实验。为了方便训练和验证模型,按照 8 : 2 的比例将新数据集划分为训练集和测试集。训练批量的大小受限于 GPU 显存,采用渐进式增加批量的方式,最终每个批次选取 256 个样本,使用自适应学习率方案,初始学习率设置为 0.001。采用 Adam 优化器,通过动量和梯度平方的均值动态调整学习率,加快模型的收敛速度并达到最优解。采用交叉熵损失函数,星座图尺寸为 256×256 像素。为了防止模型在训练过程中出现过拟合问题,采用 Dropout 正则化方法,在每次训练迭代时随机舍弃部分神经元,减少模型对某些特征的过度依赖。同时,训练过程中保存每个 Epoch 训练的模型。

所有的训练、测试实验都是基于两块 Nvidia GeForce RTX3090 GPU 实现,借助轻量级的 VS Code 平台进行实验仿真和调试,BLR 算法流程如图 7 所示。

本文算法步骤如下。

步骤 1 加载读取 RML2018.01a 数据集,通过预处理模块生成新的数据集 \mathcal{D}_{new} 。

步骤 2 将 \mathcal{D}_{new} 进行批量归一化处理,并将其划分为训练集 \mathcal{D}_{train} 和测试集 \mathcal{D}_{test} 。

步骤 3 将训练集中 IQ 序列 $X_i = [I_1, Q_1, \dots, I_{1024}, Q_{1024}]$ 映射为星座图 M_i 。

步骤 4 构建 BiLSTM-ResNet 网络模型。

步骤 5 按照批次将从训练集中提取的 I/Q 数据、星座图和标签送入 BiLSTM-ResNet 网络行训练,

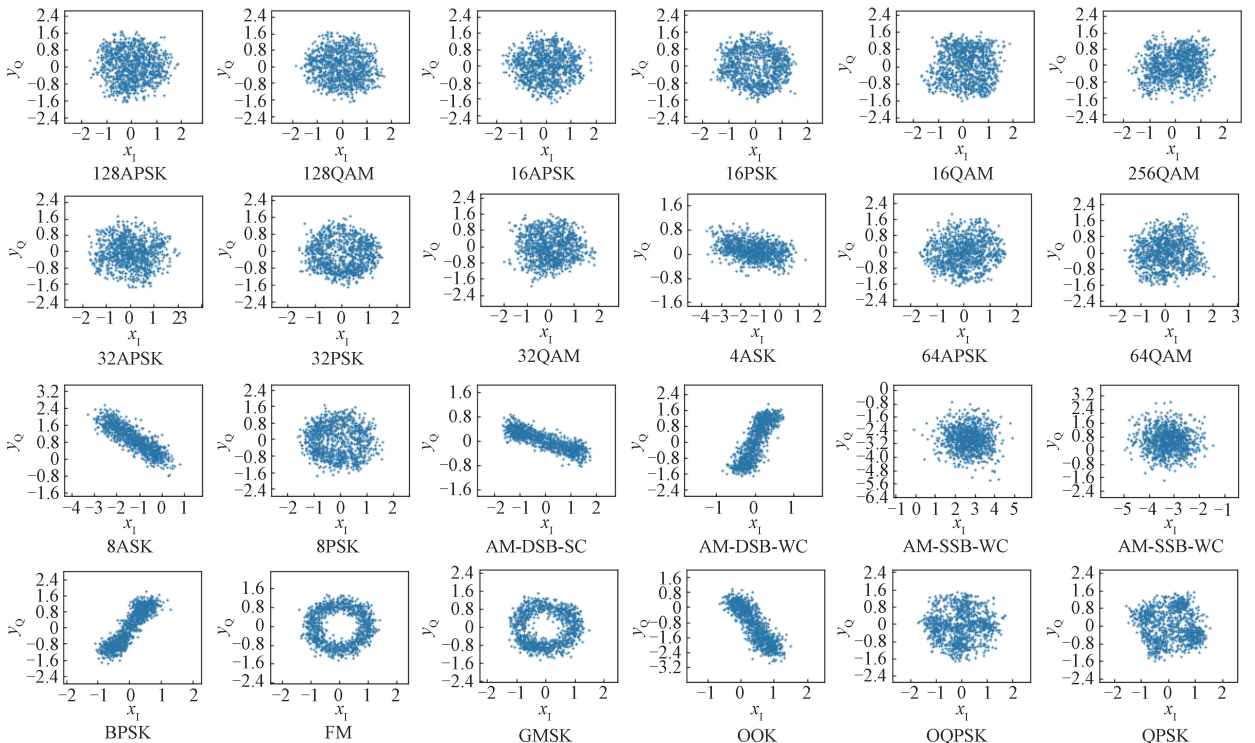


图 6 6 dB 下 24 种调制方式的星座图

Figure 6 Constellation diagram of 24 modulation modes at 6 dB

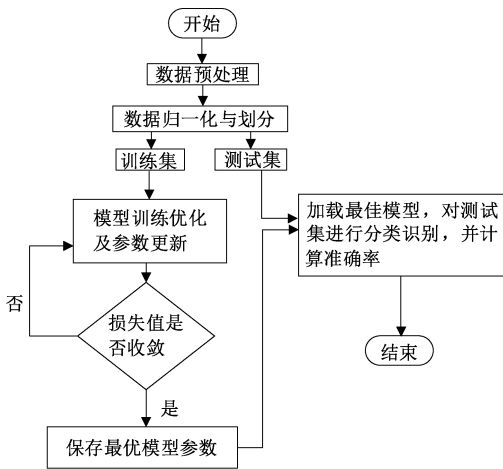


图7 BLR 算法流程

Figure 7 Flowchart of BLR algorithm

计算损失并更新模型参数,当损失值收敛时,网络达到稳态时,保存最优模型参数 W^* 。

步骤6 利用保存的最佳模型对测试数据集进行分类识别,得到信号的预测结果并计算准确率。

3.2 实验结果与分析

将数据集 RML2016.10a、RML2018.01a 进行预处理后的样本作为输入,通过对 CNN 网络进行训练,保存最佳的训练参数和模型,然后加载测试训练集,自动对信号的调制方式作出预测,从而得到不同数据集下 CNN 的平均识别率,如图 8 所示。

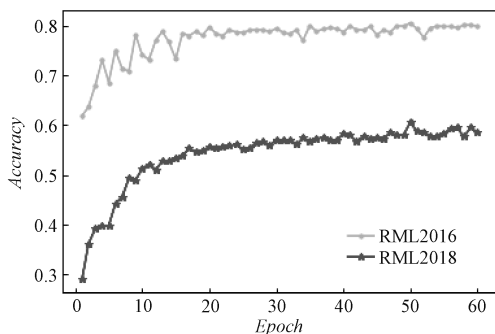


图8 不同数据集下 CNN 的平均识别率

Figure 8 Recognition rate of CNN under different datasets

由图 8 可以看出,对调制样式较少的 RML2016.10a 数据集 CNN 识别率较高,但对调制样式更为丰富的 RML2018.01a 数据集, CNN 识别效果出现了显著下降。结合实验结果分析,当训练迭代次数 (Epoch) 达到 50 轮时,信号的平均识别正确率最高,随后识别率趋于稳定,表明网络训练到一定程度后,识别性能并不会随着 Epoch 的增加不断提升,甚至会产生过拟合导致性能下降。

为了探究 BiLSTM 不同层数对模型的性能影响,本文分别选取单层、双层和三层 BiLSTM,对 24 种信号的平均识别率进行实验。实验结果表明,当上支路采用两层 BiLSTM 网络时,平均识别率最高,

达 92.57%,表明网络能够充分提取信号的时序信息。当层数增多或减少时,平均识别率 (Accuracy) 均出现不同程度的下降。因此,本研究选择采用双层 BiLSTM 结构进行实验。

为验证不同深度学习模型在提升 RML2018.01a 数据集中调制样式识别准确率的有效性,进行了基于 CNN、ResNet、BiLSTM 和 CNN-LSTM 方法的训练。实验选取 24 种调制方式,每种调制方式下选取 13 种信噪比(从 6 dB 到 30 dB,步长为 2 dB),每种信噪比下包含 4 096 个样本。表 1 展示了不同算法的时间复杂度和平均识别率。

表1 对比实验

Table 1 Comparison experiment

方法	时间复杂度	平均识别率/%
CNN	$O(N \times K_h \times K_w \times D)$	60.59
ResNet ^[10]	$O(L \times N \times K_h \times K_w \times D)$	93.87
BiLSTM ^[11]	$O(2 \times N \times H^2)$	92.57
CNN-LSTM ^[14]	$O(N \times K_h \times K_w \times D) + O(N \times H^2)$	95.23
BiLSTM-ResNet	$O(2 \times N \times H^2) + O(L \times N \times K_h \times K_w \times D)$	96.48

由表 1 可以看出, BiLSTM 和 ResNet 在识别准确率上相较于传统的 CNN 提升了 31.98% ~ 33.28%,说明了 ResNet, BiLSTM 在 RML2018.01a 数据集有更高效率的识别效果。结合实验过程,对上述网络结构的时间复杂度进行分析,如表 1 第 2 列所示。其中 $N = 2 \times 1024$ 表示输入数据有 2 个特征维度和 1 024 个时间步长; K_h, K_w 分别表示对应网络模型的卷积核的高度和宽度; $K_h \times K_w$ 为计算卷积操作时间复杂度的重要组成; D 表示输出通道数; L 表示卷积层的总数; $H = 256$ 表示隐藏层维度。

为了验证本文所提网络的有效性,将表 1 中的 BiLSTM 和 ResNet 进行并联,形成 BLR 网络。对 BLR 网络的多模态特征融合实验结果如图 9 所示。数据与上述实验规格相同,一种调制方式下的单一信噪比随机挑选 400 个样本。

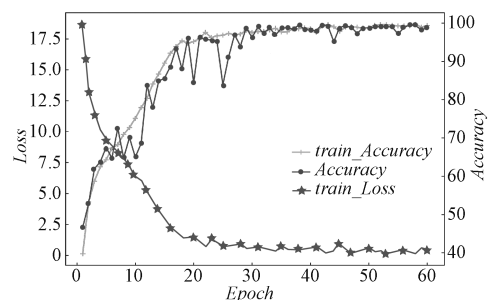


图9 BiLSTM-ResNet 下的损失与准确率曲线

Figure 9 Loss and accuracy curve of BiLSTM-ResNet

- radio modulation recognition networks [M]//Engineering Applications of Neural Networks. Cham: Springer International Publishing, 2016: 213-226.
- [8] WU Y L, LI X J, FANG J. A deep learning approach for modulation recognition via exploiting temporal correlations [C]//2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Piscataway: IEEE, 2018: 1-5.
- [9] 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.
- [10] 王晓明, 许梦竹. 基于深度学习的调制识别方法研究[J]. 中国无线电, 2023(5): 47-51.
WANG X M, XU M Z. Radio signal modulation recognition method based on deep learning [J]. China Radio, 2023(5): 47-51.
- [11] 章昕亮, 李天昀, 龚佩, 等. 联合深度学习和专家先验特征的信号调制识别[J]. 信息工程大学学报, 2023, 24(2): 129-134.
ZHANG X L, LI T Y, GONG P, et al. Signal modulation recognition based on joint deep learning and expert prior features[J]. Journal of Information Engineering University, 2023, 24(2): 129-134.
- [12] 查雄, 彭华, 秦鑫, 等. 基于多端卷积神经网络的调制识别方法[J]. 通信学报, 2019, 40(11): 30-37.
ZHA X, PENG H, QIN X, et al. Modulation recognition method based on multi-inputs convolution neural network [J]. Journal on Communications, 2019, 40(11): 30-37.
- [13] 任彦洁, 唐晓刚, 张斌权, 等. 基于时间卷积网络的通信信号调制识别算法[J]. 无线电工程, 2023, 53(4): 807-814.
REN Y J, TANG X G, ZHANG B Q, et al. Communication signal modulation recognition algorithm based on a temporal convolutional network [J]. Radio Engineering, 2023, 53(4): 807-814.
- [14] WEST N E, O'SHEA T. Deep architectures for modulation recognition[C]//2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN). Piscataway: IEEE, 2017: 1-6.
- [15] 翁建新, 赵知劲, 占锦敏. 利用并联 CNN-LSTM 的调制样式识别算法[J]. 信号处理, 2019, 35(5): 870-876.
WENG J X, ZHAO Z J, ZHAN J M. Modulation recognition algorithm by using parallel CNN-LSTM[J]. Journal of Signal Processing, 2019, 35(5): 870-876.
- [16] SNOAP J A, POPESCU D C, SPOONER C M. On deep learning classification of digitally modulated signals using raw I/Q data [C]//2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC). Piscataway: IEEE, 2022: 441-444.
- [17] 张富强, 曾夏, 白筠妍, 等. 多模态数据融合的加工作业动态手势识别方法[J]. 郑州大学学报(工学版), 2024, 45(5): 30-36.
ZHANG F Q, ZENG X, BAI J Y, et al. Dynamic gesture recognition method for machining operations based on multi-modal data fusion [J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(5): 30-36.
- [18] QI P H, ZHOU X Y, ZHENG S L, et al. Automatic modulation classification based on deep residual networks with multimodal information[J]. IEEE Transactions on Cognitive Communications and Networking, 2021, 7(1): 21-33.
- [19] 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[J]. 无线电通信技术, 2022, 48(4): 697-710.
ZHANG X X, WANG Y, LIN Y, et al. A comprehensive survey of deep learning-based automatic modulation recognition methods [J]. Radio Communications Technology, 2022, 48(4): 697-710.
- [20] BAI S J, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[EB/OL]. (2018-04-19) [2025-08-27]. <https://arxiv.org/abs/1803.01271>.
- [21] YU L, CHEN J, DING G R, et al. Spectrum prediction based on taguchi method in deep learning with long short-term memory[J]. IEEE Access, 2018, 6: 45923-45933.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [23] 黄杰, 张顺生, 陈爽. 基于深度学习网络融合的自动调制分类方法[J]. 信号处理, 2023, 39(1): 42-50.
HUANG J, ZHANG S S, CHEN S. Automatic modulation classification method based on deep learning network fusion [J]. Journal of Signal Processing, 2023, 39(1): 42-50.
- [24] 邓远征, 白立云. 基于时空联合与注意力机制的调制识别研究与实现[J]. 无线电工程, 2024, 54(12): 2913-2922.
DENG Y Z, BAI L Y. Research and implementation of modulation recognition based on spatiotemporal joint and attention mechanism [J]. Radio Engineering, 2024, 54(12): 2913-2922.
- [25] 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.

olutional entity alignment method that integrates multiple information was proposed. Firstly, the improved PageRank algorithm was used to filter triplets and alleviate the impact of differences in knowledge graph structure. Next, we learn the embedding representations of entities and attributes were learnt through graph convolutional networks, and the relationships between entities were iteratively updated by using these representations. Finally, based on the PBAB method, text description information was integrated and weighted with graph structure information to enhance the effectiveness of entity alignment. The experimental results showed that the proposed method improved the Hits@1 metric by approximately 3% compared to the optimal baseline, with corresponding improvements observed in other evaluation metrics as well.

Keywords: knowledge graph; graph convolutional network; entity alignment; graph attention network; pre-trained languages model

(上接第 82 页)

Multimodal Modulation Recognition Method Based on BLR Parallel Structure

JIANG Hua, XIAO Kejie, HU Po, GONG Kexian, ZHAO Zhenyu

(School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: Aiming at the problem that existing convolutional neural network (CNN)-based modulation recognition methods are highly dependent on single modal data (e.g., IQ sequences) and difficult to adequately extract multi-dimensional features of signals, in this study a multimodal parallel structural modulation recognition method was proposed based on bidirectional long short-term memory network (BiLSTM) and residual network (ResNet), termed the BiLSTM-ResNet (BLR network). Firstly, the temporal features of IQ data were extracted by BiLSTM in the upper branch, and the spatial features of constellation maps were extracted by ResNet-18 in the lower branch. Secondly, serial feature fusion was used in the decision fusion module to better exploit the complementary nature of the multimodal data. Lastly, the signal modulation styles were recognised with the help of the model's feature extraction capability. In this study, experimental validation was carried out on the publicly available dataset RML2018.01a. The experimental results showed that the overall recognition accuracy of BLR network in the 6–30 dB SNR interval was stable at 96.48%, 2.61% and 3.91% higher than that of the single-modal ResNet and BiLSTM models, respectively, and 1.25% higher than that of the CNN-LSTM model with concatenated structure, which verified that the model proposed in this paper had the modulation recognition problem Effectiveness.

Keywords: automatic modulation recognition; convolutional neural network; multimodal; feature fusion; parallel structure