

DOI:10.20079/j.issn.1001-893x.240625003

基于相位参数估计和空间重建的自动调制识别*

张子胤^a, 李大鹏^a, 单国强^b

(南京邮电大学 a. 通信与信息工程学院; b. 波特兰学院, 南京 210003)

摘要:针对现有深度学习调制识别算法在复杂信号环境下的鲁棒性和泛化能力不足的问题,提出了一种基于相位参数估计和空间重建的多通道网络(Phase Estimation and Spatial Reconstruction-based Attention Mechanism Multi-channel Network, PET-SAMCL)。首先,将输入的同相正交信号(In-phase Quadrature, IQ)通过相位参数估计转换,分成 3 个模块分别提取 IQ 的幅度-相位特征、IQ 合路以及分路特征。在特征提取模块中加入空间重建单元(Spatial Reconstruction Unit, SRU),减少冗余特征的影响。利用全局平均池化和软注意力操作对空间特征进行提炼与融合,通过门控循环单元(Gated Recurrent Unit, GRU)及双向门控循环单元(Bidirectional Gated Recurrent Unit, BiGRU)提取时间和空间特征。通过消融实验确定了最优模型结构。该模型在 RML2016.10a 数据集上表现优异,在 14 dB 时达到了 93.9% 的最高识别准确率,平均识别率相较其他模型最大提高了 7.7%。

关键词:自动调制识别;深度学习;相位参数估计;空间重建单元;注意力机制

开放科学(资源服务)标识码(OSID):



中图分类号:TN911.3;TP183 文献标志码:A 文章编号:1001-893X(2025)11-1859-10

Automatic Modulation Recognition Based on Phase Parameter Estimation and Spatial Reconstruction

ZHANG Ziyin^a, LI Dapeng^a, SHAN Guoqiang^b

(a. College of Telecommunications & Information Engineering;

b. Portland Institute, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: For the problem that the existing deep learning modulation recognition algorithms are not robust enough and have insufficient generalization ability in complex signal environments, a multi-channel network based on phase parameter estimation and spatial reconstruction (PET-SAMCL) is proposed. First, the input in-phase quadrature (IQ) signal is converted by phase parameter estimation and divided into three modules to extract the amplitude-phase feature, IQ combination and branching features of IQ respectively. A spatial reorganization unit (SRU) is added to the feature extraction module to reduce the influence of redundant features. The spatial features are refined and fused by global average pooling and soft attention operations, and the temporal and spatial features are extracted by gated recurrent units (GRU) and bidirectional gated recurrent units (BiGRU). Ablation study determines the optimal model structure. The model performs well on the RML2016.10a dataset, achieving a maximum recognition accuracy of 93.9% at 14 dB, and the average recognition rate is increased by 7.7% compared with that of other models.

Key words: automatic modulation recognition; deep learning; phase parameter estimation; spatial reconstruction unit; attention mechanism

* 收稿日期:2024-06-25;修回日期:2024-09-26

基金项目:国家自然科学基金资助项目(62371245)

通信作者:李大鹏 Email:dapengli@njupt.edu.cn

0 引言

自动调制识别 (Automatic Modulation Recognition, AMR) 可以为接入的无线电信号尤其是非合作无线电信号提供基本的调制信息, 在频谱感知、干扰识别等重要场景发挥了重要的作用^[1], 在近年来引起了人们广泛的研究兴趣^[2]。传统的调制识别研究可以分为两种: 第一种是基于最大似然决策的方法, 但其过于依赖先验知识与参数估计^[3], 计算复杂度较高; 第二种是基于特征提取的方法^[4], 需要对特征进行手动提取, 使用场景受限。得益于人工智能技术的普及, 深度学习在调制识别方面的应用愈发成熟^[5]。相较于传统调制识别方法而言, 神经网络可以自动从大量样本中学习特征, 无需人工提取。

O' Shea 等人^[2]首次应用卷积神经网络 (Convolutional Neural Network, CNN) 处理原始 IQ 数据进行调制识别, 并公开了 RML2016. 10a 数据集, 吸引了大量研究者的关注。然而, CNN 在时间序列数据处理中存在局限, 无法有效捕捉信号的时间相关性, 导致某些情况下识别精度不足。为此, Hermawan 等人^[6]在 CNN 结构基础上提出了改进的卷积神经网络自动调制分类网络 (Improved Convolutional Neural Network-based Automatic Modulation Classification Network, IC-AMCNet), 进一步提升了识别精度, 但复杂的网络结构增加了计算成本和训练时间, 且在不同信噪比下的适应性仍需改进。同时, 递归神经网络 (Recurrent Neural Network, RNN) 也被应用于调制识别领域^[7], 其在时间序列特征提取上具有优势, 取得了较好的效果。然而, RNN 存在梯度消失和梯度爆炸问题, 影响模型的稳定性和识别精度。West 等人^[8]提出将 CNN 和长短时记忆网络 (Long Short-Term Memory, LSTM) 结合的卷积长短期记忆深度神经网络 (Convolutional Long Short-Term Memory Deep Neural Network, CLDNN), 通过更好地提取 I/Q 样本的时间和空间特征, 提高了识别精度。然而, CLDNN 结构更加复杂, 训练时间和计算资源需求增加。Xu 等人^[9]提出了多通道卷积长短期记忆深度神经网络 (Multi-channel Convolutional Long Short-Term Memory Deep Neural Network, MCLDNN), 通过多种数据流从空间和时间角度提取特征, 实现了较高的识别精度,

并且该模型的输入端具有较高的扩展性, 可引入星座、幅度和相位等特征, 实现更精确的分类。然而, MCLDNN 的多通道结构增加了实现和部署难度, 计算成本较高。为了解决梯度消失问题, 文献^[10]中引入了 GaussianDropout 层, 通过乘性高斯噪声随机腐蚀特征, 提升了模型的鲁棒性和泛化能力, 但在某些情况下, 可能导致模型过度依赖噪声, 影响稳定性。Zhang 等人^[11]提出了基于参数估计与转换的卷积神经网络-门控循环单元深度神经网络 (Parameter Estimation and Transformation-based Convolutional Gated Recurrent Deep Neural Network, PET-CGDNN), 其识别准确率超过 90%, 且参数数量不到 1/8。尽管如此, PET-CGDNN 在复杂信号环境下的鲁棒性仍需进一步验证。文献^[12]在轻量化模型基础上加入了注意力机制, 进一步提升了性能。文献^[13]提出了一种新颖的空间和通道重建模块, 通过减少卷积中的空间和通道冗余, 降低计算成本和模型存储, 同时提升了 CNN 模型的性能。

尽管现有的网络模型在识别精度上取得了一定的进展, 但它们仍然存在一些不足: 首先, 部分模型在处理时间序列数据时未能充分捕捉信号的时空特征, 限制了识别性能; 其次, 复杂的多通道结构导致计算成本较高, 模型的实现与部署难度增加; 最后, 在复杂的信噪比条件下, 现有模型的鲁棒性和泛化能力有待进一步提升。基于上述问题, 本文设计了一种基于相位参数估计和空间重建的多通道网络, 主要贡献如下: ①在空间特征提取模块中加入空间重建单元, 对输入空间特征进行重建, 增强网络的特征提取能力; ②结合通道重建单元 (Channel Reconstruction Unit, CRU), 通过全局平均池化和软注意力操作, 进一步丰富了提取的空间特征; ③在多通道输入框架中增加一个通道用于提取 I/Q 数据的幅度和相位特征。与 MCLDNN 模型相比, 本文所提模型具备更高的信号识别率、更少的参数, 且在复杂信号环境下表现更优异。

1 信号模型

调制信号经过传输信道并被采样之后, 其等效基带信号可以表示为

$$y[n] = A[n] e^{j(\omega n + \varphi)} x[n] + v[n], n = 1, 2, \dots, N \quad (1)$$

式中: $y[n]$ 代表接收信号; $x[n]$ 为发射机发射的调制信号; $A[n]$ 代表信道增益; $v[n]$ 是加性高斯白噪声; ω 和 φ 则分别代表信号的角频率偏移与相位偏移。

在自动调制识别中, 通常将上述信号处理成 I/Q 分量的数据格式, 即

$$y = \begin{bmatrix} i[1] & i[2] & i[3] & \cdots & i[n] \\ q[1] & q[2] & q[3] & \cdots & q[n] \end{bmatrix} \quad (2)$$

$$i[n] = \text{Re}\{y[n]\}, n = 1, 2, \dots, N \quad (3)$$

$$q[n] = \text{Im}\{y[n]\}, n = 1, 2, \dots, N \quad (4)$$

式中: N 为信号的采样长度; $\text{Re}\{y[n]\}$ 表示信号实部; $\text{Im}\{y[n]\}$ 表示信号虚部。

2 基于相位参数估计和空间重建的多通道网络模型

本节将介绍一种基于相位参数估计和空间重建的多通道网络模型, 其整体结构如图 1 所示, 网络由空间特征提取模块 (part-A)、空间特征融合模块 (part-B) 和时间特征提取模块 (part-C) 组成。

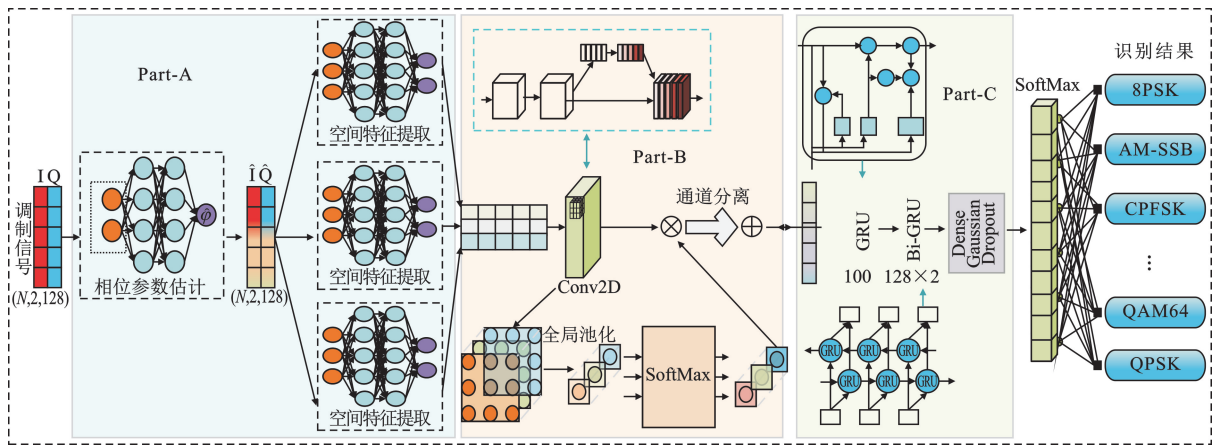


图 1 PET-SAMCL 模型网络结构

2.1 空间特征提取模块

2.1.1 相位参数估计模块

在实际情况中, 传输信号由于受到信道噪声的干扰, 输入网络中的 IQ 数据通常带有相位偏移信息, 因此在空间特征提取模块中添加了相位参数估计模块, 可以去除相位偏移降低模型识别准确率的不利影响。

如图 2 所示, 该模块使用 Flatten 层、Dense 层以及线性激活函数通过与后续模型的协同训练^[11] 来获取估计的相位参数, 然后通过相位转置公式计算来实现相位偏移的纠正:

$$\hat{y}[n] = y[n] e^{-j\hat{\varphi}} = \begin{bmatrix} \text{Re}\{y[n]\} \cos \hat{\varphi} + \text{Im}\{y[n]\} \sin \hat{\varphi} \\ \text{Im}\{y[n]\} \cos \hat{\varphi} - \text{Re}\{y[n]\} \sin \hat{\varphi} \end{bmatrix} \quad (5)$$

式中: $\hat{\varphi}$ 是相位参数; $\text{Re}\{y[n]\}$ 、 $\text{Im}\{y[n]\}$ 代表信号的实部与虚部, 即 I、Q 数据。

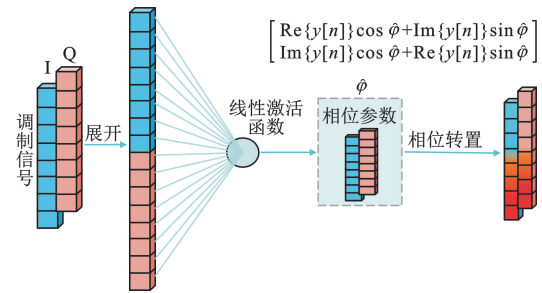


图 2 相位参数估计模块

2.1.2 多通道提取空间特征

输入 IQ 数据通过相位矫正之后, 将数据分别输入到 3 个模块 (IQ 幅度-相位特征提取模块、IQ 合路特征提取模块、IQ 分路特征提取模块) 进行不同特征的提取, 这样可以提取到不同信道间 (即 IQ 信道、I 信道和 Q 信道) 的互补特征, 进一步提高模型的性能。总体的公式可以表示为

$$F_{\text{concat}} = \text{Concat}(f_1(X; \hat{\varphi}), f_2(X; \hat{\varphi}), f_3(X; \hat{\varphi})) \quad (6)$$

式中: Concat 表示拼接操作; f_1 、 f_2 、 f_3 分别表示 IQ 数据的幅度-相位特征模块、合路特征提取模块、分路

特征提取模块,对应图 3 中的(a)(b)(c); X 为输入的 IQ 数据; $\hat{\phi}$ 表示相位参数估计模块输出的值。

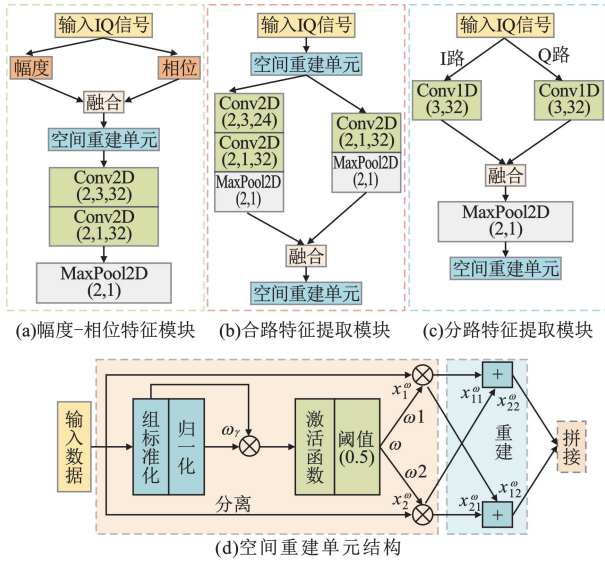


图 3 空间特征提取模块及空间重建单元结构

在图 3(a)中,由 IQ 数据计算出幅度与相位特征,通过两个卷积层来提取特征,卷积核数量均为 32,大小分别为(2,3)和(2,1)。最后通过步长为(2,1)的最大池化层,对输入的数据在时域方向进行降维,减少后续的计算成本。在图 3(b)中,将输入的 IQ 数据分别通过两个分支进行特征的提取;第一个分支先使用数量为 24 的(2,3)卷积核来提取 IQ 数据的通用特征,再通过(2,1)卷积核来提取深层 IQ 数据的相关特征,最后用最大池化层来缩减尺寸;第二个分支则直接使用(2,1)卷积核来提取 IQ 的相关特征,通过(2,1)池化层来缩减尺寸。然后,将两个分支的张量数据进行拼接。在图 3(c)中,分别使用(3,32)的一维卷积对传入的 I 路数据与 Q 路数据进行时域特征提取,将其拼接后通过池化进行尺寸缩减。

2.1.3 空间重建单元

在图 3(a)~(c)中都添加了空间重建单元。该模块分为分离与重建两个部分。分离部分是为了将数据中信息丰富和相对较少的空间特征分开。首先通过组标准化层^[14]中的缩放因子来评估不同数据特征的信息内容,接着使用归一化来表示不同特征的重要性,将元素相乘得到相关权重,将权重通过 sigmoid 函数映射到(0,1)范围并设定 Gate 函数的阈值为 0.5(在权重映射到[0,1]范围内时,0.5 作为中间值,能够均衡地分离高于和低于中间值的特征,从而保证模型的稳定性和鲁棒性),高于阈值的

部分为 ω_1 ,代表有信息丰富的权重;低于阈值的部分为 ω_2 ,表示匮乏的权重。将输入特征 Input 分别与权重相乘,得到信息丰富的空间特征数据与信息匮乏的冗余数据。分离部分的公式如下:

$$\omega = \text{Gate}_{0.5}(\text{Sigmoid}(\omega_y \otimes \text{GN}(X))) = \begin{cases} \omega_1, & \text{映射值} \geq 0.5 \\ \omega_2, & \text{映射值} < 0.5 \end{cases} \quad (7)$$

$$\text{GN}(X) = \gamma \frac{X - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta \quad (8)$$

$$\omega_y = \{\omega_i\} = \frac{\gamma_i}{\sum_{j=1}^C \gamma_j}, i, j = 1, 2, \dots, C \quad (9)$$

式中:GN 表示组标准化操作; μ, σ 分别代表输入数据 X 的均值和标准差; ϵ 为保证除法运算稳定的常数; γ, β 代表网络中的可训练参数;Gate 是门限函数,阈值为 0.5; \otimes 表示元素相乘操作; $\omega_y \in \mathbb{R}^C$ 是归一化权重; C 表示输入数据的通道数。

在重建部分中,通过将上述两块数据进行交叉重建来充分组合加权的两个不同信息特征并加强它们之间的信息流,从而抑制原数据中的冗余特征,进一步增强模型的特征提取能力。具体公式如下:

$$\begin{cases} x_{1,2}^{\omega} = \omega_{1,2} \otimes X \\ x_i^{\omega} = [x_{i1}^{\omega}; x_{i2}^{\omega}], i = 1, 2 \\ \text{Sru}(X) = [x_{11}^{\omega} \oplus x_{22}^{\omega}; x_{12}^{\omega} \oplus x_{21}^{\omega}] \end{cases} \quad (10)$$

式中: $x_{1,2}^{\omega}$ 代表不同 ω 与输入数据 X 进行元素相乘;Sru 代表输出重建数据; \oplus 表示元素相加操作。

2.2 空间特征融合模块

在图 4 中,将拼接好的空间特征数据通过数量为 64 的(5,5)卷积核进一步提取特征,然后引入通道注意力机制模块(挤压激励模块)^[15]来提高信息利用率与网络性能。与传统方法(仅使用全局平均池化层或线性变换)相比,挤压激励模块具有显著优势。全局平均池化层虽然能减小特征图尺寸,但无法动态调整各通道的重要性,容易丢失重要信息。而简单的线性变换虽然计算量小,但对特征选择的效果有限,无法自适应地调整各特征通道的重要性。相比之下,挤压激励模块能够自适应地分配不同通道的重要性权重,从而更有效地选择关键特征和抑制冗余信息,增强模型对细节的处理能力,在不显著增加模型复杂度的基础上,提升模型的性能。将数据通过全局平均池化层、SoftMax 层来生成特征的重要性参数,将其与提取后的特征数据相乘得到新的特征数据,最后将各通道空间数据进行分离并相加,得到最终的空间特征数据。

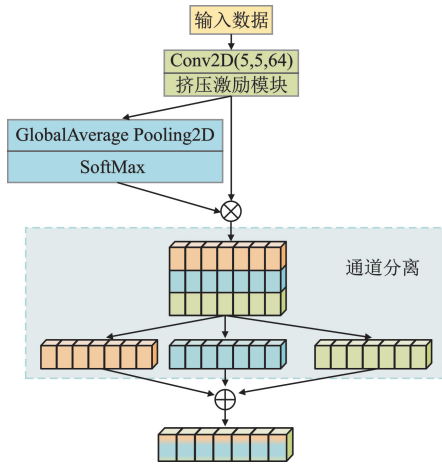


图 4 空间特征融合模块

2.3 时间特征提取模块

在 MCLDNN 模型中,通过加入双层的 LSTM 层可以有效处理序列数据以提取时间相关性。受到该模型的启发,在本文中将提取好的空间数据通过 GRU(100 cells)、Bi-GRU(128×2 cells)层来提取数据的时域特征。GRU 相较于 LSTM 具有更少的训练参数,在实际应用中具有更好的性能。如图 5 所示,Bi-GRU 网络在 GRU 的基础上增加了从未来到过去的数据流,并且过去使用的隐藏层和未来使用的隐藏层之间没有联系。因此, Bi-GRU 相较于 GRU 可以更好地探索数据的时间特征^[16]。Bi-GRU 网络的公式如下:

$$\vec{h}_t = \text{GRU}_f(x_t, \vec{h}_{t-1}) \quad (11)$$

$$\overleftarrow{h}_t = \text{GRU}_b(x_t, \overleftarrow{h}_{t-1}) \quad (12)$$

$$y_t = [\vec{h}_t; \overleftarrow{h}_t] \quad (13)$$

式中: GRU_f 表示正向 GRU 输出; GRU_b 表示反向 GRU 输出; $[\ ;]$ 表示拼接操作。

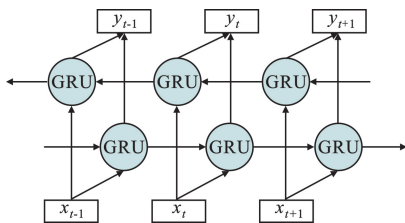


图 5 Bi-GRU 结构

最后,使用一个具有 256 个神经元的全接层并使用 SELU 作为激活函数。为了防止过拟合,引入了 GaussianDropout 层。该层为输入的数据添加均值为 1、标准差为 0.5 的乘性高斯噪声,起到了正则化的作用,提高模型的鲁棒性和识别准确性。

2.4 模型结构对比

表 1 显示,混合模型结合了 CNN 对局部特征的

提取和 RNN 对时序信息的处理能力,相较于单一 CNN 或 RNN,能同时捕捉信号的空间和时间特征,提高准确率和鲁棒性。本文提出的 PET-SAMCL 模型在多通道输入的基础上扩展了 IQ 数据的幅度/相位特征提取,并引入空间重建单元,有效去除冗余特征,提升提取精度。时间特征模块使用 GRU + BiGRU 代替单一 GRU/LSTM,增强了对信号时序前后文的理解能力,提高模型对不同信号的识别准确率。综合上述改进,PET-SAMCL 的模型结构更加合理,对比其他模型,能够更好地适应复杂的信号环境。

表 1 模型结构对照

模型类型	模型名称	输入数据	主要结构
卷积神经网络	CNN1 ^[2]	I/Q	2 卷积层+全连接层
	CNN2 ^[18]	I/Q	4 卷积层+全连接层
递归神经网络	ICAMCNET ^[6]	I/Q	CNN+Gaussian noise
	GRU2 ^[7]	I/Q	2 GRU layers
混合模型	LSTM2 ^[19]	Amplitude/Phase	2LSTM layers
	CLDNN ^[8]	I/Q	CNN+LSTM+Skip connections
混合模型	CGDNet ^[10]	I/Q	CNN+GRU+DNN
	PET-CGDNN ^[11]	I/Q	CNN+GRU+DNN
	MCLDNN ^[9]	I/Q, I and Q	多通道(I/Q,I,Q)+CNN+2 LSTM layers
	PET-SAMCL	I/Q	多通道(IQ 分路、合路、幅度/相位)+CNN(添加空间重建单元)+GRU+BiGRU

3 实验与结果分析

3.1 数据集介绍

为了方便验证和对比 PET-SAMCL 的性能,本文将使用开源数据集 RadioML2016. 10a^[17] 和 RadioML2016. 10b 进行实验。RadioML2016. 10a 包含 11 种调制类型:WBFM、AM-DSB、AM-SSB、BPSK、CPFSK、GFSK、4-PAM、16-QAM、64-QAM、QPSK 和 8PSK。RadioML2016. 10b 数据集是 RadioML2016. 10a 的扩展版本,包含 1 200 000 个信号和除 AM-SSB 之外的 10 种调制信号。数据集中的信号有 128 个复杂的浮点时间 I/Q 样本,在恶劣的传播环境中生成,并受到加性高斯白噪声(Additive White Gaussian Noise, AWGN)、多径衰落、采样率偏移和中心频率偏移的破坏,以此来模拟真实环境。按 6 : 2 : 2 的比例划分数据集,随机选择 600 个信号作为训练数据,200

个信号作为验证数据,200 个信号作为测试数据。

3.2 网络参数设置

实验的所有框架使用分类交叉熵损失函数和 Adam 优化器。初始学习率从 0.001 开始,如果验证损失在 10 个 epoch 内没有下降,则乘以 0.8 以提高训练效率。梯度更新的批量大小为 128。当验证损失在 60 个 epoch 内没有改善时,则停止训练过程,并使用具有最小验证损失的训练模型来预测每个测试信号的调制类型。实验环境为 Windows 操作系统,搭载 NVIDIA RTX 4090 GPU,使用 Python 接口的 tensorflow2.9.0 深度学习框架完成仿真实验。

3.3 实验仿真

3.3.1 消融实验

在构建 PET-SAMCL 网络模型的过程中,为了确

定现有框架中的特征提取模块能否有效提高网络的性能,建立了 4 组对照模型,具体含义如表 2 所示。

对照模型	网络框架
PET-SAMCL_A	PartA(无 SRU 模块)+PartB+PartC
PET-SAMCL_B	PartA+PartB(无 Conv(5,5,64))+PartC
PET-SAMCL_C	PartA+PartB+PartC(GRU 代替 Bi-GRU)
PET-SAMCL_D	PartA+PartB+PartC(增加全连接层)
PET-SAMCL	PartA+PartB+PartC

实验结果如图 6 所示,其中 (a)(b) 折线图末端部分放大,可以清晰地看到 0~18 dB 信噪比的对比曲线,以便直观分析。5 个网络获得的识别精度通过直方图直观地表示,(c)(d) 的 Macc(最大精度)代表了 SNR 范围内网络的最大识别精度,Aacc(平均精度)表示信噪比大于 0 dB 时网络的平均识别精度。

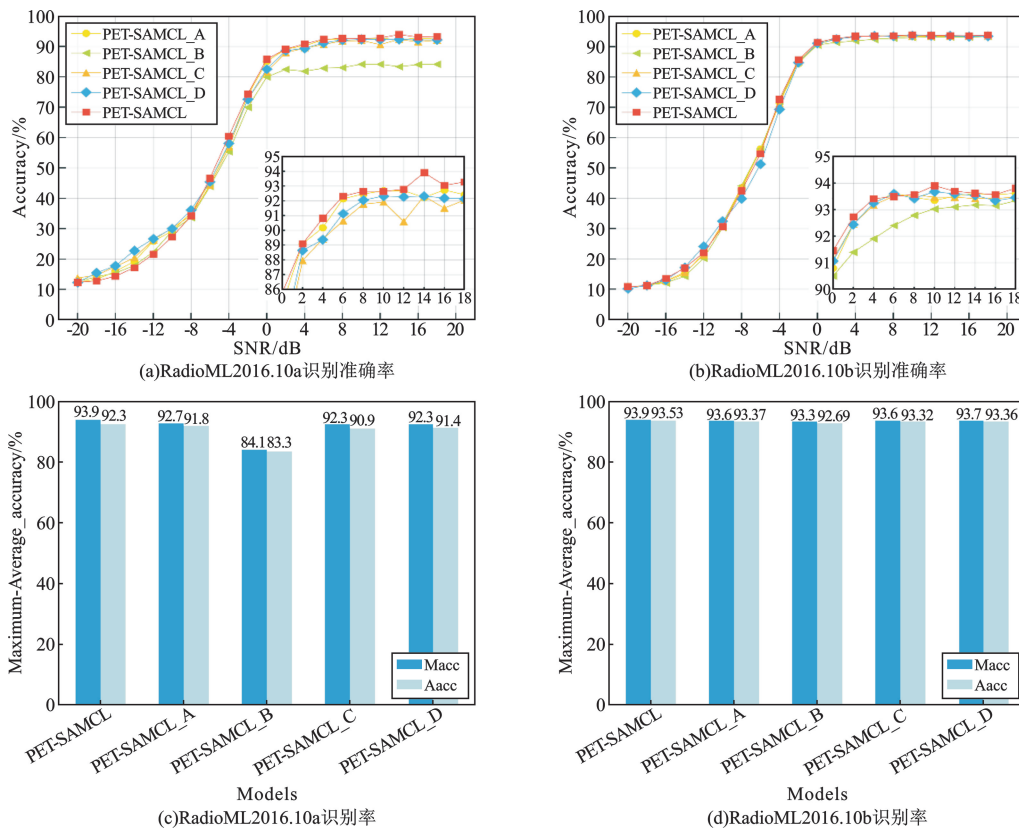


图 6 消融实验结果

对比 PET-SAMCL_A 和 PET-SAMCL,去除空间重建单元后,由于冗余空间特征的影响,模型性能下降,(c)中 Macc 和 Aacc 分别下降 1.2% 和 0.5%,(d)中分别下降 0.3% 和 0.16%。PET-SAMCL_B 的实验结果表明,PartB 中对于拼接好的空间特征进行卷积操作是必不可少的步骤,对识别精度至关重要。PET-SAMCL_C 中使用 Bi-GRU 替代 GRU 模块

后,(c)中 Macc 和 Aacc 分别提升 1.6% 和 1.4%,(d)中分别下降 0.3% 和 0.21%,表明 Bi-GRU 有助于更好地探索时间特征。PET-SAMCL_D 增加了一层 FC 层,但并未提升识别精度,反而在 (c) 中 Macc 和 Aacc 分别下降 1.6% 和 0.9%,(d)中分别下降 0.2% 和 0.17%,说明过多的 FC 层导致过拟合,降低泛化能力。

综上, PET-SAMCL 模型各部分互为补充, 其组合使得模型性能更优, 因此最终选择 PET-SAMCL 网络结构。

3.3.2 不同网络模型的对比试验

上节通过对比实验确定了 PET-SAMCL 模型的最终结构, 本节将该网络模型与 CNN1^[2]、CNN2^[18]、CLDNN^[8]、CGDNet^[10]、GRU2^[7]、ICAMCNET^[6]、LSTM2^[19]、PET-CGDNN^[11]、MCLDNN^[9] 等 9 种现有的先进调制信号识别网络进行性能比较。

图 7(a) 显示了在 RML2016. 10a 数据集上各模型的识别准确率, 可见随着信噪比的增加, 准确率逐渐提升并趋于稳定。在信噪比位于 -20~-4 dB 之间时, 受噪声影响, 各模型的识别精度较低。然而, PET-SAMCL 网络结构的识别准确率高干大多数其他网络, 验证了其通过提取不同空间特征进行互补,

能在低信噪比下表现出良好的分类性能。尤其在信噪比高于 0 dB 时, PET-SAMCL 的识别效果优于 MCLDNN。图 7(b) 展示了在 RML2016. 10b 数据集上的各模型表现, 信噪比较低时, PET-SAMCL 相较于其他模型识别效果稍好一些, 在信噪比大于 -4 dB 时, 其表现更佳。

图 7(c)~(d) 分别展示了各模型在两个公共数据集上的最高信号识别精度和在信噪比大于 0 dB 时的平均识别精度, 通过折线图可以直观对比模型表现。尽管 MCLDNN 在其他对比模型中表现良好, 本文提出的 PET-SAMCL 在 RML2016. 10a 和 RML2016. 10b 数据集上的表现更优。在图 7(c) 中, PET-SAMCL 的最高和平均识别精度分别为 93.9% 和 92.3%, 比 MCLDNN 提高了 1.3% 和 0.8%。在图 7(d) 中, PET-SAMCL 则分别高出 0.2% 和 0.2%。

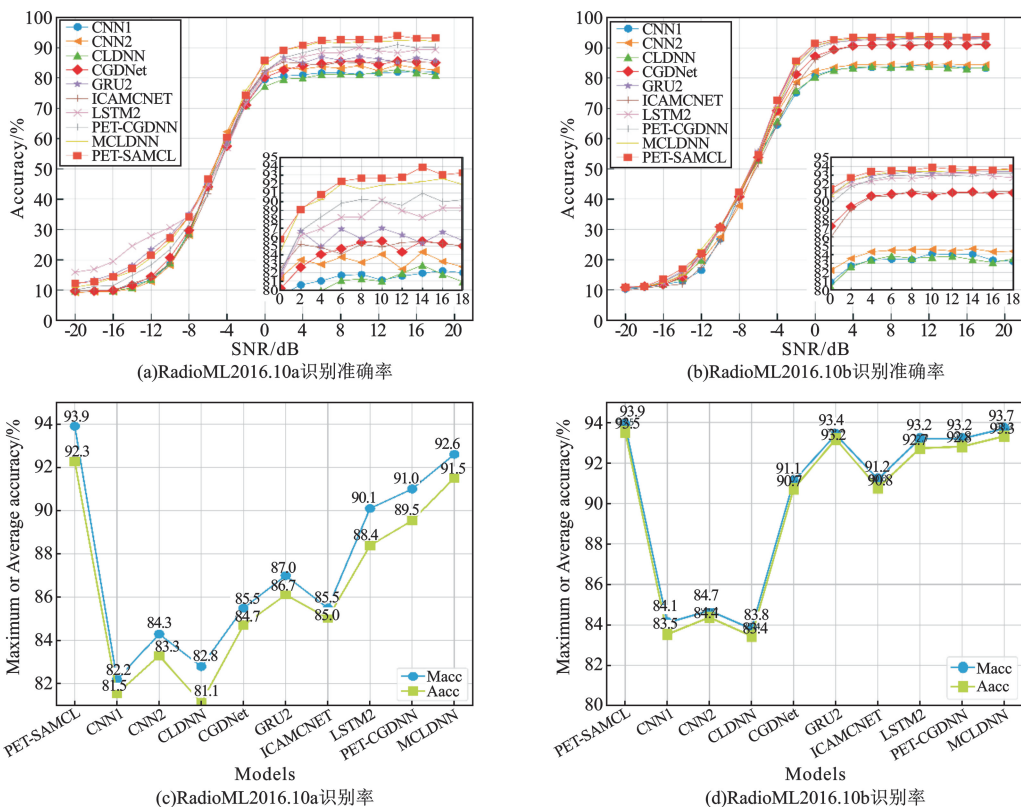


图 7 多模型对比实验结果

表 3 给出了各模型在数据集 RML2016. 10a 和 RML2016. 10b 上的模型学习参数、训练的最小验证损失、整体信噪比上的平均识别精确度以及训练时间。对比最优的两个模型 PET-SAMCL 与 MCLDNN, PET-SAMCL 有着更少的学习参数, 比 MCLDNN 减少了 13.58% 和 13.62%。对比其他模型, PET-SAMCL 在两个数据集上有着最小的验

证损失与最高的平均识别精度。然而, 通过对比各模型的训练时间可以看出, PET-SAMCL 模型的计算复杂度仍然较高, 特别是由于引入了空间重建单元和注意力机制, 进一步增加了计算成本。与 MCLDNN 模型相比, PET-SAMCL 模型在牺牲计算时间的情况下换取了更高的识别准确率。

表 3 模型性能比较

模型	学习参数		最小验证损失		平均识别精确度		训练时间(秒/周期)	
	a	b	a	b	a	b	a	b
PET-SAMCL	351 015	350 758	1.062 6	0.884 3	0.618 6	0.644 7	21	90
MCLDNN	406 199	406 070	1.084 8	0.889 2	0.612 1	0.643 2	16	82
PET-CGDNN	71 871	71 742	1.147 4	0.900 3	0.590 3	0.637 7	9	47
CNN1	1 592 383	1 592 126	1.215 2	0.997 8	0.545 3	0.578 5	3	14
CNN2	858 123	857 994	1.193 0	0.986 5	0.554 8	0.587 4	5	25
CLDNN	164 433	164 176	1.236 0	1.012 8	0.541 6	0.582 7	7	32
CGDNet	124 933	124 676	1.193 2	0.928 2	0.560 5	0.624 7	7	33
GRU2	151 179	151 050	1.124 3	0.889 0	0.589 7	0.643 9	14	67
ICAMCNET	1 264 011	1 263 882	1.181 4	0.965 2	0.561 7	0.611 7	4	20
LSTM2	201 099	200 970	1.063 7	0.889 7	0.609 7	0.640 0	14	63

图 8~9 显示了信噪比为 0 dB 条件下不同模型的混淆矩阵,选取的模型为平均识别精确度较高的前 6 位。在混淆矩阵中,调制信号的识别精度可以通过颜色的深浅来判断,识别正确率越高,对应对角线上的蓝色越深。图右侧的色标表示识别正确率的百分比(%),纵坐标代表调制信号的真实类别,横坐标代表对应的预测类别。图 8 展示了在 RML2016. 10a 数据集下的混淆矩阵,可以看出,影响低信噪比条件下识别精度的因素有两点:首先是 AM-DSB 和 WBFM 信号存在混淆问题,这是由于 AM-DSB 和 WBFM 信号都有静默期(信号传输时没

有有效载荷数据的时间段),在静默期间信号几乎为零,因此网络模型无法对其进行有效的识别分类;然后是 QAM16 与 QAM64 信号也存在分类错误,这是由于数据集为 IQ 数据,在星座图上可以看出存在重叠的星座点,模型难以进行识别。相较于其他模型,PET-SAMCL 在这两点上有着一定的改善。图 9 展示了在 RML2016. 10b 数据集下的混淆矩阵,对比 (d) 中出现 AM-DSB 与 WBFM 之间的混淆问题和 (f) 中的 QAM 信号混淆问题,PET-SAMCL 都有了很大的改善;对比 (b) 中的 MCLDNN 模型,PET-SAMCL 的整体识别效果更加出色。

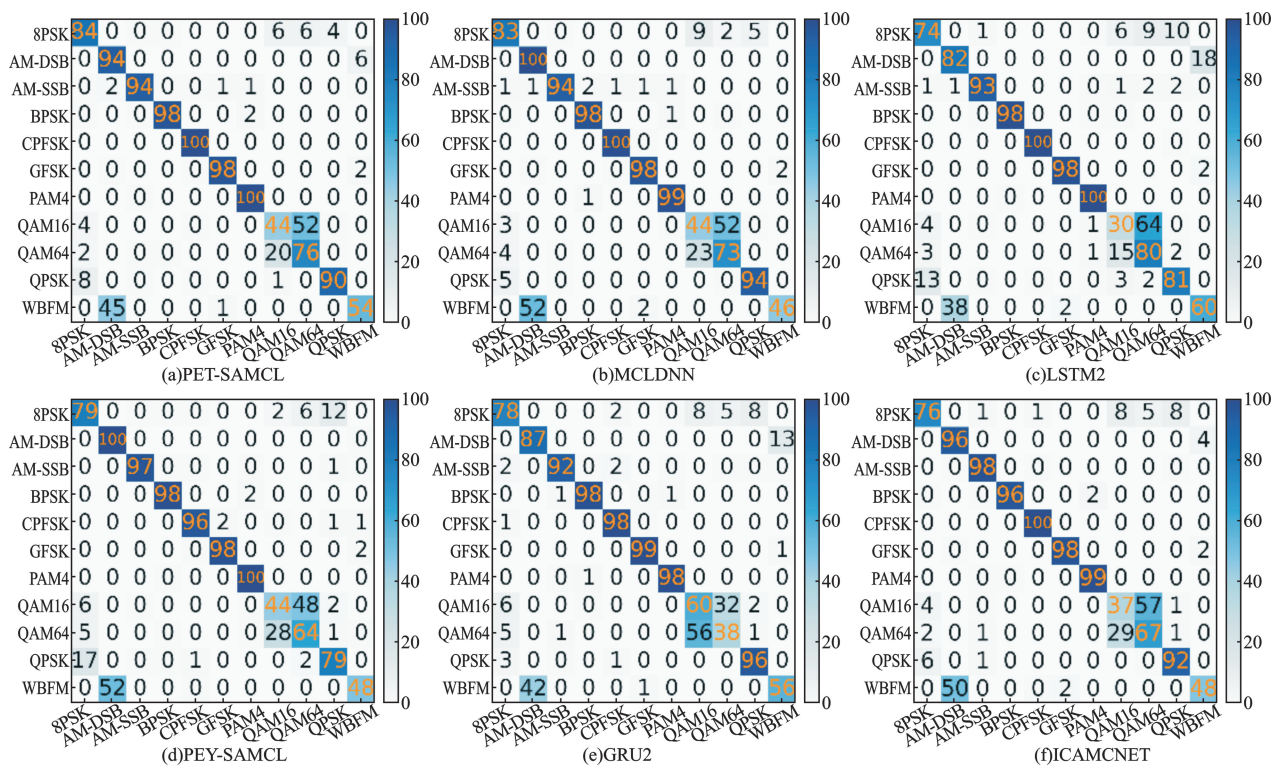


图 8 数据集 a 混淆矩阵

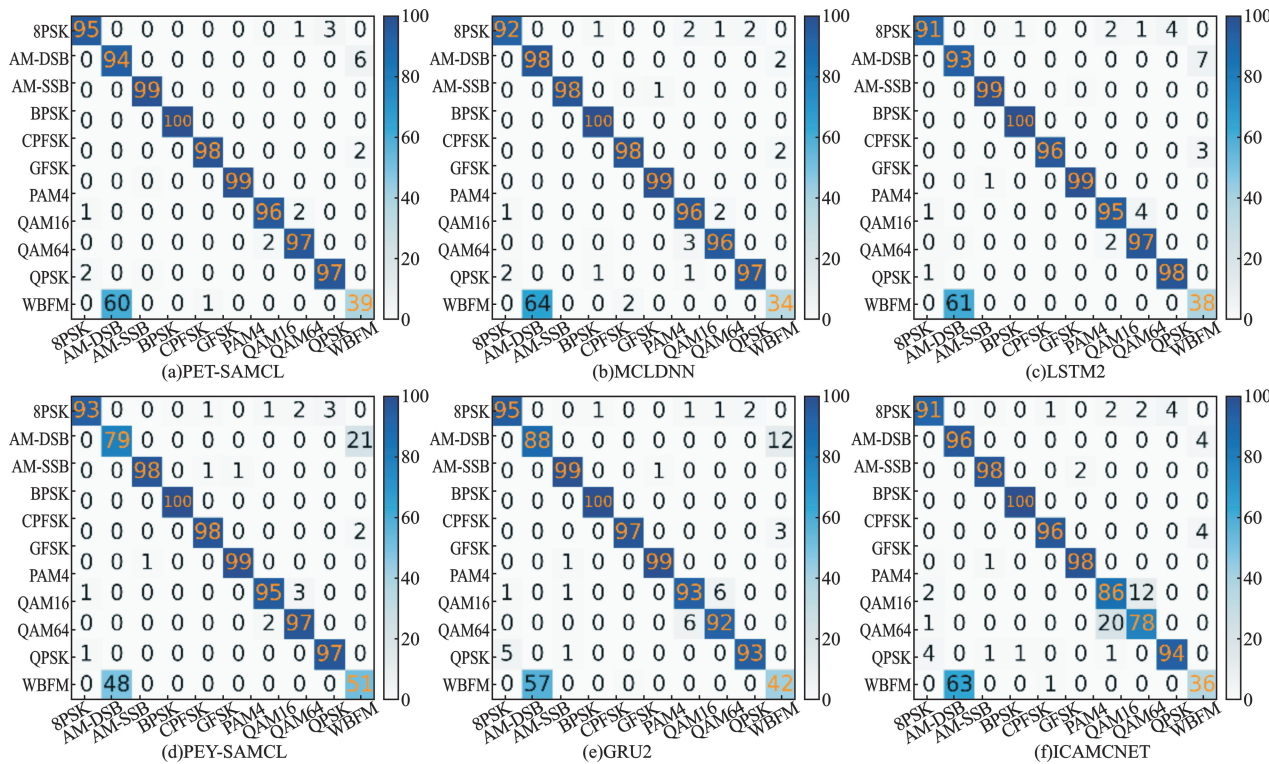


图 9 数据集 b 混淆矩阵

4 结束语

本文针对现有深度学习调制识别算法在复杂信号环境下的鲁棒性和泛化能力不足的问题,提出了一种基于相位参数估计和空间重建的多通道网络 PET-SAMCL。先通过相位参数估计模块提取 IQ 信号的相位偏移信息来实现相位变换,在现有多通道输入框架的基础上进行扩展,分别提取 IQ 的幅度-相位特征、合路以及分路特征;通过引入空间重建单元来优化对于空间特征的提取过程,减少冗余特征的影响,加入了注意力机制以改善模型的性能;最后结合 GRU 和 Bi-GRU 来高效地提取时间特征,从而增强模型对复杂信号的识别能力。实验结果表明,该模型在 RML2016. 10a 数据集上表现优异,在 14 dB 时达到了 93.9% 的最高识别准确率,平均识别率相较于其他模型最大提高了 7.7%,验证了其在复杂信号环境下的卓越性能。

未来计划继续在现有模型基础上融合更多特征,并对算法进行轻量化处理,以进一步加速训练过程。

参考文献:

[1] ZHANG F X, LUO C B, XU J L, et al. Deep learning based automatic modulation recognition: models, datasets, and challenges [J]. Digital Signal Processing,

2022, 129: 103-115.
 [2] O' SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks [C] // Engineering Applications of Neural Networks. Cham: Springer, 2016: 213-226.
 [3] DULEK B. Online hybrid likelihood based modulation classification using multiple sensors [J]. IEEE Transactions on Wireless Communications, 2017, 16 (8): 4984-5000.
 [4] CHANG D C, SHIH P K. Cumulants-based modulation classification technique in multipath fading channels [J]. IET Communications, 2015, 9(6): 828-835.
 [5] XIAO W S, LUO Z Q, HU Q. A review of research on signal modulation recognition based on deep learning [J]. Electronics, 2022, 11(17): 150-170.
 [6] 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.
 [7] HONG D H, ZHANG Z L, XU X D. Automatic modulation classification using recurrent neural networks [C] // 2017 3rd IEEE International Conference on Computer and Communications. Chengdu: IEEE, 2018: 695-700.
 [8] WEST N E, O' SHEA T. Deep architectures for modulation recognition [C] // 2017 IEEE International Symposium on Dynamic Spectrum Access Networks. Baltimore: IEEE, 2017: 1-6.

- [9] 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.
- [10] NJOKU J N, MOROCHO-CAYAMCELA M E, LIM W. CGDNet: efficient hybrid deep learning model for robust automatic modulation recognition [J]. IEEE Networking Letters, 2021, 3(2): 47–51.
- [11] 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.
- [12] JIA F, YANG Y Y, ZHANG J Y, et al. A hybrid attention mechanism for blind automatic modulation classification [J]. Transactions on Emerging Telecommunications Technologies, 2022, 33(7): 640–655.
- [13] LI J F, WEN Y, HE L H. SCConv: spatial and channel reconstruction convolution for feature redundancy [C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 6153–6162.
- [14] WU Y X, HE K M. Group normalization [M]// 2018 European Conference on Computer Vision. Cham: Springer International Publishing, 2018: 3–19.
- [15] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132–7141.
- [16] ZOU B H, ZENG X D, WANG F Q. Research on modulation signal recognition based on CLDNN network [J]. Electronics, 2022, 11(9): 1379–1404.
- [17] O'SHEA T J, WEST N. Radio machine learning dataset generation with gnu radio [C]// The GNU Radio Conference. Boulder: ACM, 2016: 69–74.
- [18] TEKBIYIK K, EKI 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. Antwerp: IEEE, 2020: 1–6.
- [19] 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.

作者简介:

张子胤 男, 2000 年生于江苏兴化, 2022 年获学士学位, 现为硕士研究生, 主要研究方向为调制识别算法研究、深度学习。

李大鹏 男, 1982 年生于山东临清, 2011 年于上海交通大学获博士学位, 现为教授, 主要研究方向为新一代通信、人工智能、无人系统等新技术。

单国强 男, 1996 年生于江苏盐城, 2019 年获学士学位, 现为硕士研究生, 主要研究方向为多智能体强化学习。