

领域知识驱动结合深度学习的调制识别方法

刘高辉,张宇

(西安理工大学 自动化与信息工程学院,西安 710048)

E-mail:liugh68@xaut.edu.cn

摘要:针对数据驱动深度学习调制识别中的数据冗余及低信噪比下训练性能欠佳的问题,提出了一种基于领域知识驱动的双流网络模型,并引入了自适应注意力机制.首先,利用通信信号领域知识和主分量分析算法提取信号关键动态特性,并初始化卷积神经网络,提高训练效率;其次,将信号的I/Q数据送入初始化后的卷积神经网络提取信号时域特征;同时,对具有噪声抑制特性的双谱数据进行对角切片处理,送入结合软阈值降噪算法的深度残差收缩网络提取频域特征,将时、频域特征同领域知识组成联合特征向量;在分类阶段,自适应注意力机制通过动态调整注意力头数量与权重筛选冗余特征,最后经全连接层完成分类.仿真结果表明:提出的模型在SNR=0dB时识别率达到87.9%,最高识别率达到95.8%,训练时间减少6.40%~39.68%.相比其他深度学习模型,本方法在较低参数量下表现出更好的性能.

关键词:调制识别;知识驱动;深度学习;网络初始化;注意力机制

中图分类号:TN911

文献标识码:A

文章编号:1000-1220(2026)02-0343-08

Domain Knowledge-driven Modulation Recognition Method Combined with Deep Learning

LIU Gaohui, ZHANG Yu

(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

Abstract: Aiming at the problem of data redundancy and poor training performance under low SNR in data-driven deep learning modulation recognition, a dual-stream network model based on domain knowledge-driven approach was proposed, and an adaptive attention mechanism was introduced. The domain knowledge of communication signals and principal component analysis algorithm were used to extract key dynamic characteristics, which were then employed to initialize the convolutional neural network and improve initial training performance. The I/Q data of the signal were input into the initialized CNN to extract time domain features. Meanwhile, the bispectrum data with noise suppression characteristics were processed by diagonal slices and sent to a deep residual shrinkage network combined with a soft threshold denoising algorithm to extract frequency domain features. The time and frequency domain features, along with domain knowledge, were combined to form a joint feature vector. In the classification stage, the adaptive attention mechanism filtered redundant features by dynamically adjusting the number and weight of attention heads. Finally, classification was completed through a fully connected layer. Simulation results show that the recognition rate of the proposed model reaches 87.9% when SNR = 0dB, with a maximum recognition rate of 95.8%. The training time was reduced by 6.40% to 39.68%. Compared to other deep learning models, the proposed method demonstrates better performance with fewer parameters.

Keywords: modulation recognition; knowledge driven; deep learning; network initialization; attention mechanism

0 引言

通信调制方式自动识别(Automatic Modulation Classification, AMC)作为信号处理和通信系统的核心部分在军用通信侦察、民用频谱监测和认知定义无线电领域都有着重要的研究意义.在复杂电磁环境中,信号通常受到多径干扰和强噪声等因素的影响,给调制识别任务带来了严峻挑战.尤其是在低信噪比场景下,传统基于特征提取或深度学习的方法识别率往往显著下降.因此,如何提升低信噪比条件下的识别性能成为当前研究的重要课题.

传统的AMC算法大多是基于领域知识来设计的,这些

领域知识可能来自与通信和信号处理领域的理论相关的推定统计模型或确定性模型^[1].常用特征包括信号的瞬时特征^[1-3]、统计量特征^[4,5]和变换域特征^[6,7],如S. Zheng等人^[1]对信号的瞬时特征、统计特征和光谱特征进行集成,构建了两个包含传统特征和原始I/Q数据的调制分类数据集,在小样本分类场景下通过结合深度学习框架完成信号的识别分类;李正权等人^[4]将信号的高阶累积量作为特征输入,结合判别式受限波尔兹曼机,在动态信噪比条件下显著提高了识别性能.基于知识的方法通常只需估计少量参数即可实现调制分类,因此不需要大量的训练样本来学习输入与期望输出之间的关系.然而,这类方法在复杂、动态信道环境中表现出适应

性不足,并且对不同调制类型的特征选择依赖性较强,限制了其实际应用。

近年来,数据驱动的深度学习技术广泛应用在 AMC 领域.数据驱动深度学习的方法主要是将大量数据作为训练输入,通过含有隐藏节点的多层分布式网络进行处理,最后通过输出分类层输出预测分类结果^[8].如 Liu. B 等人^[9]将信号的频谱作为领域知识,与信号的 I/Q、A/P 数据联合输入网络进行特征提取,利用 3 种模态的互补性增强分类性能;Chen. Tao 等人^[10]通过替换离散小波变换分解的细节系数,利用不同的小波基重构生成扩展样本,改进训练效果.然而,纯粹的数据驱动方法强调更大的输入数据量来训练网络,这增加了计算复杂度,并可能导致分类性能的下降.此外,许多研究通过对网络整体结构进行设计优化或引入权重“惩罚项”约束提高系统的识别性能,如 Gao. Jiaxin 等人^[11]提出一种基于混合专家(MoE)模型的框架,将 Transformer 模块用于低 SNR 信号处理,ResNet 模块用于高 SNR 信号处理,实现了在不同 SNR 条件下的有效识别;D. Zhang 等人^[12]引入高阶卷积和多光谱注意力机制,将高阶统计特征和频谱信息结合到神经网络中,提升分类性能及网络可解释程度;Thorpe. Adam 等人^[13]提出将先验知识结合到使用内核嵌入的数据驱动控制算法中,将先验知识作为内核学习的偏置项,把有偏差的学习问题表述为具有正则项的最小二乘问题进行有效计算的封闭解.尽管这些研究在提高噪声环境下的识别性能方面取得了一定成果,但它们往往侧重于数据样本的不充分和无效的属性而忽略了领域知识提供的信号关键动态特性,而这些特性对于信号的本质表征至关重要。

当前,与调制分类中神经网络分类器性能的大量研究相比,较少有研究文献关注通过对领域知识引导的特征提取与数据驱动的深度学习神经网络分类器相结合来提升低信噪比下系统的识别性能.考虑到知识驱动和数据驱动方法的优势互补,本文提出了一种基于领域知识驱动的时频域双流网络模型,并结合自适应注意力机制用于通信信号调制识别.领域知识特征的引入能够有效增强信号的表征能力,采用主分量分析算法(Principal Components Analysis, PCA)对其进行主分量特征分析提取并初始化网络权重,提升网络训练效率和提取特征的有效性,从而引导网络更好地捕捉信号有效特征.在此基础上,本文进一步设计了一种时频域双流网络结构,对信号的 I/Q 和双谱数据分别通过经 PCA 初始化后的卷积神经网络和深度残差收缩网络(Dccp Residual Shrinkage Networks, DRSN)进行特征提取,全面提升模型对多模态信号特征的表征能力.此外,为了筛选出高辨识度的特征并筛选冗余信息,本文设计了基于门控自注意力机制的自适应特征筛选策略,以提高识别的准确性和鲁棒性。

1 通信符号波形信号模型

在通信系统中,信号调制是指将原始基带信号通过频谱搬移变换成高频信号以适应信道传输的过程.原始的数据序列首先经过星座映射将数据比特转换为复数符号,每个符号包含同相(In-phase, I)和正交(Quadrature, Q)两个分量,随后将其分别与正交载波相乘,两者相和得到带通调制信号。

不同的调制方式决定了信号的具体表示形式以及如何在接收端进行解析.当通过改变载波的幅度、频率或相位来承载并传输信息,即信号采用移幅键控(ASK)、移频键控(FSK)或移相键控(PSK)调制方式时,对接收到的通信信号按符号周期分割为连续的符号波形信号,其数学模型可以表示为:

$$s(t) = [A \sum_{n=-\infty}^{+\infty} a_n g(t-nT)] \cos(2\pi(f_c + f_n)t + \phi_0 + \phi_n) \quad (1)$$

式中, A 为载波幅度; a_n 为发送的符号序列; $g(t)$ 为低通脉冲波形; T 为符号周期; f_c 为载波频率; f_n 为调制频率; ϕ_0 为载波初始相位; ϕ_n 为调制相位。

而正交振幅调制(QAM)属于高阶数字调制,是一种同时变化信号幅度和相位的调制方式.它基于正交信号的叠加,通过将数据分成多个比特组,并将每个比特组映射到一个特定的信号点,对接收到的 QAM 调制信号按符号周期分割成连续的符号波形信号,每个符号波形信号都包含了完整的 I、Q 信息,其数学模型可表示为:

$$s(t) = [A \sum_{n=-\infty}^{+\infty} a_I g(t-nT)] \cos(2\pi f_c t + \phi_n) + [A \sum_{n=-\infty}^{+\infty} a_Q g(t-nT)] \sin(2\pi f_c t + \phi_n) \quad (2)$$

式中, a_I, a_Q 分别为通信信号的同相、正交符号序列。

2 领域知识驱动结合深度学习的调制识别算法

2.1 时域特征提取

2.1.1 PCA 卷积网络初始化

在研究符号波形信号的识别与分类时,瞬时特征参数如瞬时频率、瞬时幅度和瞬时相位,是描述信号动态特性的关键.传统方法根据这些特征的二阶特性在时频域的差异来区分不同调制信号.本文为了使网络更好地捕捉单个符号周期内调制信号的动态特性,选取了信号的 7 个瞬时统计二阶特征统计量作为领域知识特征引导网络的初始化,这些特征参数包括:

- 零中心归一化瞬时幅度谱密度最大值 $\gamma_{\max}^{[1]}$: 衡量瞬时幅度频谱的集中程度;
- 零中心归一化瞬时幅度绝对值的标准偏差 σ_{aa} 和非弱信号段瞬时频率绝对值的标准偏差 $\sigma_{af}^{[3]}$: 分别反映瞬时幅度变化和瞬时频率变化的离散程度;
- 零中心归一化瞬时幅度的紧致性 μ_{a2}^a 和瞬时频率的紧致性 μ_{f2}^f : 分别评估瞬时幅度分布和瞬时频率分布的紧密性;
- 零中心非弱信号段瞬时相位的非线性分量的标准偏差 σ_{ap} 及绝对值标准差 $\sigma_{ap}^{[3]}$: 分别衡量相位非线性变化的离散程度和绝对偏差,排除弱信号段以减少噪声干扰。

为充分利用选取的领域知识,本文采用 PCA 算法来评估输入特征的有效信息,通过计算主成分特征组合来筛选出对调制方式识别最具贡献的关键动态特性. PCA 通过线性映射学习到的主成分特征,与自编码器通过非线性映射得到的隐藏层特征在结构上具有一定的相似性.基于这一特性,本文将 PCA 计算得到的特征向量用于自编码神经网络中编码层的初始化,替代传统的随机初始化方式.这种初始化方法不仅能

够有效提供训练良好的单层前馈神经网络使网络能够在较少的迭代次数内收敛至更好的模型参数,还能在网络训练初期生成更高质量的特征表示,从而有助于捕捉信号的关键动态特性.

将接收到的通信信号按符号周期分段,其中各信号样本 $s_d (d=1,2,\dots,N)$ 包含 q 维特征,即每个信号样本的特征向量为 $F = (s_d^{(1)}, s_d^{(2)}, \dots, s_d^{(q)})$,由于本文选取的领域知识为 7 个特征,因此本文中 $q=7$.

在特征空间 F 中,对信号样本 s_d 计算对应的协方差矩阵 C^ϕ 为:

$$C^\phi = \frac{1}{N} \sum_{d=1}^N \phi(s_d) \phi(s_d)^T \quad (3)$$

在特征空间 F 中分析提取 C^ϕ 的主分量可看作求解 C^ϕ 特征值 λ 及对应特征向量 η 的问题:

$$\lambda^\phi \eta = C^\phi \eta \quad (4)$$

将计算得到的特征向量根据其对应的特征值大小进行降序依次排列,并将这些特征向量按行组成矩阵.特征值的大小直接反映了其对应特征向量在数据总体方差中的贡献率,即特征值 λ 越大,其对应的特征向量对数据方差的贡献率越高.求解前 l 个贡献率大于 90% 的特征值 λ 所对应特征向量 η 作为主分量特征,阈值 ψ 求解表示为:

$$\psi = \frac{\sum_{d=1}^l \eta_d}{\sum_{d=1}^q \eta_d} \quad (5)$$

用于初始网络的主分量特征表示为:

$$\text{PCA}(s_d^{(1)}, s_d^{(2)}, \dots, s_d^{(q)}) = (\eta_1', \eta_2', \dots, \eta_l') \quad (6)$$

设对 L 个卷积核进行初始化,利用 PCA 提取的主分量特征矩阵近似代替编码层的学习结果作为卷积核的初始权值,寻找标准正交矩阵最小化重构误差对卷积神经网络的卷积核进行初始化:

$$\min \|Y - V_l V_l^T Y\|_F^2, \text{ s. t. } V_l^T V_l = E \quad (7)$$

式中, Y 代表信号样本集中的领域知识数据; V_l 为提取的前 l 个特征向量组成的矩阵; E 为单位矩阵.

2.1.2 卷积神经网络结构

对于卷积神经网络 (Convolutional neural network, CNN) 主体结构的搭建,利用二维卷积核对抗符号波形信号数据中的 I、Q 两路分量进行处理可以对抗多径效应引起的符号间频率选择性衰落^[14].因此本文特征提取网络的具体结构如图 1 所示,其中第 1 层是基于符号波形信号领域知识预先构建的

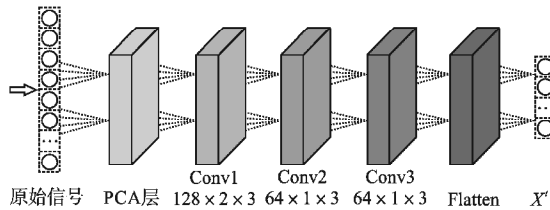


图 1 经 PCA 初始化的卷积神经网络结构

Fig. 1 CNN structure initialized with PCA

PCA 层,其他层分别选用尺寸为 $2 \times 3, 1 \times 3$ 的卷积核,通过堆叠三层数量分别为 128、64、64 的卷积层来完成卷积神经网络主体架构的搭建.

第 1 层为领域知识卷积层,它通过量化输入信号的领域知识特征与对应信号特征之间的相似性,将不同调制方式的领域知识信息纳入特征提取过程.因此,在构建物理层时不会引入额外的超参数.这一特性使得在不增加训练工作量和过拟合风险的情况下尽可能地提高网络的特征提取能力.

2.2 频域特征提取

2.2.1 符号波形信号频域高阶谱特征

双谱 (Bi-spectrum) 指通信信号高阶谱中的三阶累积量谱,对高斯白噪声有很好的抑制作用.与功率谱相比,它包含了信号的幅度和相位特征信息,能够在低信噪比下对识别效果不佳的信号更准确地识别.

为了确保符号波形信号高阶累积量的 Fourier 变换和双谱的存在,要求离散符号序列 $s(i)$ 的 3 阶累积量 $c_{3s}(k_1, k_2)$ 绝对可求和,即:

$$\sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} |c_{3s}(k_1, k_2)| < +\infty \quad (8)$$

则信号的三阶累积量 $c_{3s}(k_1, k_2)$ 可表示为:

$$c_{3s}(k_1, k_2) = E\{s^*(i) s(i+k_1) s(i+k_2)\} \quad (9)$$

式中, $E\{\cdot\}$ 为均值, k_1, k_2 为时延, $*$ 表示共轭计算.

对信号的随机序列 $s(i)$ 三阶累积量进行二维傅里叶变换得到信号的双谱表示形式,由于高斯白噪声的三阶累积量为 0,则信号的双谱 $B_s(\omega_1, \omega_2)$ 可表示为:

$$B_s(\omega_1, \omega_2) = \sum_{k_1=-\infty}^{+\infty} \sum_{k_2=-\infty}^{+\infty} E\{s^*(i) s(i+k_1) s(i+k_2)\} e^{-j(\omega_1 k_1 + \omega_2 k_2)} \quad (10)$$

式中, $s(i) = x(i) + n(i)$; $x(i)$ 为信号零均值平稳随机序列; $n(i)$ 为第 i 个序列的高斯白噪声; ω_1, ω_2 表示频率.

根据双谱分布的对称特性,信号的双谱表示在平面上有 12 个扇形区,即 6 个对称关系.于是,只要根据双谱部分区域内的双谱特征即可完全描述信号的频域信息.同时,考虑到计算成本,对信号的双谱数据进行对角切片 (Diagonal Slice, DS) 处理并作为 DRSN 的输入特征序列.令 $\omega_1 = \omega_2 = \omega$ 得到双谱对角切片 $B_{DS}(\omega_1, \omega_2)$ 的表达式为:

$$B_{DS}(\omega_1, \omega_2) = B_s(\omega_1, \omega_2) |_{\omega_1 = \omega_2 = \omega} \quad (11)$$

2.2.2 软阈值降噪网络结构

为了实现特征序列的软阈值降噪处理,DRSN 通过引入“软阈值化 (Soft Thresholding, ST)”处理将输入的特征数据中有效信息的绝对值放大,而噪声信息尽可能缩小到 0,从而达到降噪的目的,同时软阈值函数导数非 0 即 1 的性质可以防止网络训练中出现“梯度消失”和“梯度爆炸”的问题.软阈值化处理可表示为:

$$f(B_D) = \begin{cases} B_D - \tau, & B_D > \tau \\ 0, & -\tau \leq B_D \leq \tau \\ B_D + \tau, & B_D < -\tau \end{cases} \quad (12)$$

式中, τ 是软阈值并且是正数, B_D 为软阈值化前的双谱切片特征, $f(B_D)$ 为软阈值化后的双谱切片特征.

引入 ST 后的 DRSN 网络提取特征流程如图 2 所示,其中 BRC 由批量归一化 (Batch Normalization, BN) 层、ReLU 激活函数及卷积层组成.首先,通过批量归一化 (Batch Normalization, BN) 层对特征进行标准化处理,使得特征服从均值 μ 为 0、方差 σ 为 1 的标准正态分布,随后利用训练过程中学习

到的尺度参数 γ 和平移参数 β 对标准化后的分布进行自适应调整. 对于第 N 层的批量归一化过程可表示为:

$$\mu = \frac{1}{N_{\text{Batch}}} \sum_{i=1}^{N_{\text{Batch}}} B_D^i \quad (13)$$

$$\sigma = \frac{1}{N_{\text{Batch}}} \sum_{i=1}^{N_{\text{Batch}}} (B_D^{i+1} - \mu)^2 \quad (14)$$

$$\hat{B}_D^{i+1} = \frac{B_D^{i+1} - \mu}{\sqrt{\sigma^2 + \xi}} \quad (15)$$

$$B_D^{i+1} = \gamma \hat{B}_D^{i+1} + \beta \quad (16)$$

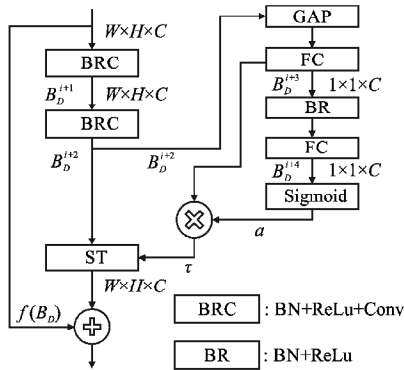


图2 双谱特征的软阈值处理

Fig. 2 Soft thresholding processing of Bi-spectrum features

随后,应用 ReLU 激活函数增加网络非线性;对 ReLU 激活后的特征进行卷积操作提取频域特征. 将经过 BRC 处理后的双谱特征同时输入至残差结构中,通过全局均值池化 (GAP) 得到特征矩阵 A , 计算其绝对值的平均值并组成一组

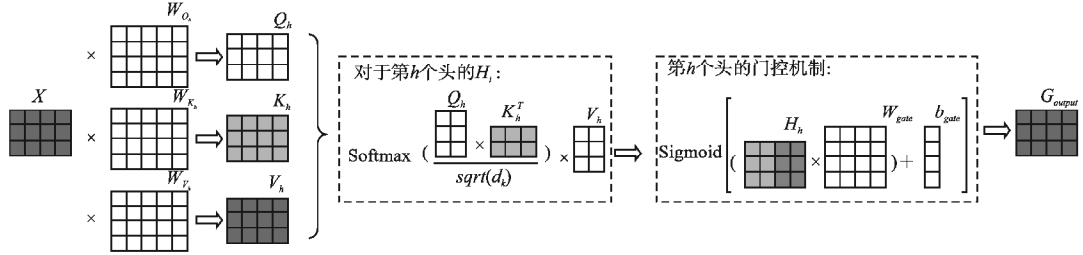


图3 第 h 个头的自适应注意力计算

Fig. 3 Adaptive attention calculation of the h -th head

对融合后的多特征序列 X 进行线性变换,通过自学习获得各注意力头的重要值,并根据该重要值为获得的特征赋予权重分别得到查询矩阵 Q , 键矩阵 K 和值矩阵 V , 公式如下:

$$Q = XW_Q \quad (18)$$

$$K = XW_K \quad (19)$$

$$V = XW_V \quad (20)$$

式中, W_Q, W_K, W_V 表示可学习的权重矩阵,通过学习过程中自动调整的参数以优化模型在训练数据上的性能.

将得到的查询、键、值矩阵分别拆分成多个头. 对于第 h 个头进行以下计算:

$$H_h = A_h(QW_{Qh}, KW_{Kh}, VW_{Vh}) \quad (21)$$

式中, H_h 表示第 h 个头的输出; A_h 表示第 h 个头的注意力函数计算; W_{Qh}, W_{Kh}, W_{Vh} 分别表示第 h 个头的权重矩阵. 接着通过点积的计算方式获得各特征之间的相关性,其中 A_h 的

一维向量,由两层全连接层 (FC) 获得一个权重系数 α ,接着使用 Sigmoid 函数将权重系数 α 转换为 $0 \sim 1$ 之间, α 与特征矩阵 A 的每个通道的特征分别相乘得到最终的阈值 τ ,整个网络由全连接层、批量归一化、ReLU 和 Sigmoid 激活函数完成自动学习阈值,有效消除信号中的噪声,改善低信噪比下部分信号识别效果不佳的问题.

在 DRSN 网络主体结构选取中,与 ResNet50 相比, ResNet34 在关键性能指标几乎相当,但训练所需参数量较低,且网络整体架构较简单^[15],因此本文选择 ResNet34 为频域特征提取网络主体结构.

2.3 自适应注意力机制

通过嵌入领域知识的特征提取网络对符号波形信号提取特征得到信号的时域特征序列 X^t , 同时将符号波形信号的双谱数据送入 DRSN 网络学习得到信号的频域特征序列 X^f , 最后联合领域知识特征序列 X^c 进行拼接融合得到多特征序列 X , 表示为:

$$X = X^t + X^f + X^c \quad (17)$$

为了解决传统多头注意力机制 (Multi-Head Attention, MHA) 数据量繁杂的问题,本文设计自适应注意力机制 (Adaptive Attention Mechanism, AAM) 通过在 MHA 的基础上增加门控机制以控制信息流.

针对第 h 个头的自适应注意力计算如图 3 所示,首先通过多头自学习获得各注意力头的重要值,并根据该重要值为获得的特征赋予权重; AAM 通过学习门控单元权重以及偏置向量将注意力头的数量 h 变为一个可学习参数,通过门控机制 G_h 来控制每个注意力头的激活状态,以达到模型根据任务的动态调整注意力头的数量.

计算公式表示为:

$$A_h = \text{Softmax} \left(\frac{QW_{Qh} \cdot (KW_{Kh})^T}{\sqrt{d_k}} \right) \cdot V_h \quad (22)$$

式中, d_k 表示头的维度; Softmax 函数用于归一化权重和偏差,确保总和为 1.

对每个注意力头进行学习门控单元权重以及偏置向量,门控机制的动态学习过程可表示为:

$$G_h = \sigma(W_{gate} H_h + b_{gate}) \quad (23)$$

式中, W_{gate} 和 b_{gate} 为可学习参数, $\sigma(\cdot)$ 为 sigmoid 激活函数,随后每个添加门控单元的注意力头 H_h' 可表示为:

$$H_h' = H_h \odot G_h \quad (24)$$

式中, \odot 表示哈达玛积, H_h' 表示学习门控单元权重后的注意力头.

最后将自适应调整后多个头的输出进行加权平均,通过

合并来自注意力头的特征信息得到最终的多头注意力输出,表示为:

$$\text{AAM}_{\text{output}} = \sum_{h=1}^H \varepsilon_h \cdot H'_h \quad (25)$$

式中, $\text{AAM}_{\text{output}}$ 为动态学习后各个头的加权平均输出; H 为头的数量; ε_h 表示第 h 个头的权重,可用公式表示为:

$$\varepsilon_h = \frac{\exp(\text{score}_h)}{\sum_{h=1}^H \exp(\text{score}_h)} \quad (26)$$

式中, $\text{score}_h = \frac{Q_h K_h^T}{\sqrt{d_k}}$.

2.4 算法流程

整个算法架构包括数据预处理、多特征提取网络模型构建和识别分类 3 个主要部分,完整的识别流程如图 4 所

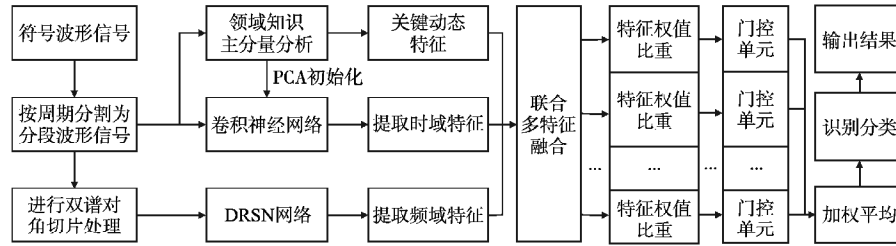


图 4 领域知识驱动多特征融合分类算法流程

Fig. 4 Process of domain knowledge-driven multi-feature fusion classification algorithm

3 实验与结果讨论

为了验证基于领域知识驱动特征提取和自适应注意力机制的有效性,从不同初始化方法,不同输入特征识别、不同模型识别、不同信噪比下识别进行仿真实验,在通信信号数据集上测试网络模型的性能。

3.1 实验条件设置

3.1.1 数据集介绍

本文使用 GNU Radio + Python 形式生成的 Radioml2016.10a 公开数据集作为本文模型训练和测试的数据样本来源,数据集中原始信号 I、Q 路两分量数据以 2×128 格式存放,其中 128 代表每个样本长度。该数据集包含了频率选择性衰落、采样率偏移、高斯加性白噪声传播条件下的 11 类常用调制信号(包含 8 种数字调制:8PSK, BPSK, CPFSK, GFSK, PAM4, 16QAM, 64QAM, QPSK; 3 种模拟调制: AM-DSB, AM-SSB, WBFM)共 22 万个样本,信噪比从 -20dB 以 2dB 为间隔取至 18dB 。具体参数见表 1。

在实验中,数据集按照 8:1:1 随机分为 3 组,其中 80% 的样本被分配到训练集,10% 的样本被分配到验证集,剩下的 10% 被分配到测试集。同时为消除信道增益的影响,对获取的符号波形信号数据进行归一化处理。本文采用线性归一化方法,即 min-max 标准化,规定原始数据通过线性变换得到的结果映射到 $[0, 1]$ 之间,转换函数为:

$$\hat{s} = \frac{s_i - s_{\min}}{s_{\max} - s_{\min}} \quad (27)$$

其中, s_{\max} 为样本对应数据值的最大值, s_{\min} 为样本对应数据值的最小值, \hat{s} 为归一化处理过后的数据。

示。首先,考虑到多模态信号间的协同效应有助于低信噪比下的识别性能提升,数据预处理阶段对信号进行双谱对角切片处理以获取更多的有效信息;接着,鉴于信号的关键动态特性对调制方式识别至关重要,本文设计了一种基于领域知识驱动的时频双流网络模型。该模型通过领域知识引导的 PCA 初始化卷积神经网络提取信号的时域特征,并采用 DRSN 对信号的双谱数据提取频域特征;在分类阶段,为了进一步筛选冗余特征并聚焦于有效信息,本文创新性地引入自适应注意力机制,根据特征的重要性自适应地为每个注意力头分配权重,从而提高模型对关键特征的关注度。该方法通过充分挖掘信号在时、频域上的丰富特征信息,并结合单个符号周期内信号关键动态特性,提升网络模型在低信噪比条件下的识别性能。

表 1 RML2016. a 数据集参数

Table 1 RML2016. a dataset parameters

参 数	数值
采样频率	200 kHz
采样率偏移标准偏差	0.01Hz
最大采样率偏移	50 Hz
载波频率偏移标准偏差	0.01 Hz
最大载波频率偏移	500 Hz
频率选择性衰落中使用的正弦波数	8
部分信号样本延迟参数	[0.0,0.9,1.7]
每个延迟参数对应的幅度	[1,0.8,0.3]

3.1.2 实验设置

在模型训练过程中,以标准交叉熵函数作为损失函数,由结合了 L2 正则化思想的 Adam 优化器训练,初始学习率设为 0.001,批处理大小设为 256,若验证损失在连续 10 epochs 内没有减少触发提前停止机制,停止模型训练。

实验系统环境为 Windows 11 操作系统,在 NVIDIA GeForce RTX 4060 GPU 搭载 i9-13900HX 处理器的计算设备上完成,深度学习神经网络训练采用 Tensorflow2.1.0 + Python3.7 利用 GPU 加速训练。

3.2 PCA 初始化仿真实验

为验证 PCA 算法初始化权重对网络整体训练的有效性,将本文提出的 PCA 领域知识驱动权重初始化算法与随机初始化算法、Xavier 初始化算法^[16]和 He 初始化算法^[17]应用至 2.1 节中提出的 CNN 结构中进行了对比实验。在使用随机初始化算法时,设定从均值为 0,标准差为 0.01 的高斯分布中随机采样获得权重;Xavier 和 He 初始化算法原理类似,都是

根据前一层输入与后一层输出神经元的数量确定权重初始化的范围,但两者不同的是,He 初始化将权重初始化均匀分布或高斯分布中的较大随机值。

实验以训练准确率和训练成本作为判断不同算法初始化卷积权重的性能标准,为避免随机因素对实验的影响,每组均设置迭代次数最高为 40,设置学习率为 0.001,重复进行 10 次实验后以各项的平均值作为评估标准。

表 2 不同初始化方法性能对比

Table 2 Performance comparison of different initialization methods

方法	最高准确率/%	整体准确率/%	Epoch/轮	训练时间/s
Xavier	74.1	50.7	40	640
随机初始化	73.6	49.3	34	510
He 初始化	80.7	56.3	24	408
PCA 初始化	82.5	60.2	31	393

由表 2 可知,在训练准确率指标上,PCA 初始化算法较其他算法均有不同幅度的提升,最高准确率达到 82.5%,整体准确率达到 60.2%。此外,在训练成本指标上,PCA 初始化所需的时间(393s)与 He 初始化(408s)持平,这是因为 He 初始化算法使用不同的系数来适应 ReLU 激活函数的性质,减少了训练次数,但 PCA 算法用时明显少于随机初始化(510s)和 Xavier 方法(640s)。最后,在 Epoch 次数上,PCA 初始化平稳需要进行 31 次,比 He 初始化(24 次)多,但比随机初始化(34 次)和 Xavier 方法(40 次)要少。

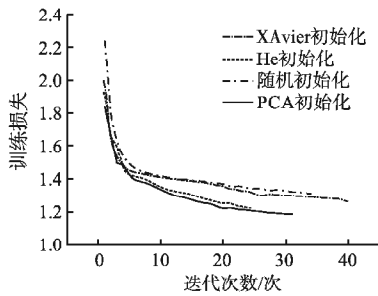


图 5 不同初始化方法对比实验

Fig. 5 Comparison experiment of different initialization methods

如图 5 所示,PCA 初始化算法在初期阶段展现出了较快的下降趋势,尤其是在前 10 次迭代内,它的损失函数值下降得最快。相比之下,Xavier 算法和随机初始化算法虽然也在下降,但是下降的速度相对较慢。当迭代次数继续增加到 20 次左右时,He 初始化算法开始展现出与 PCA 初始化算法相似的快速下降趋势,然而,Xavier 算法和随机初始化算法的下降速度依然较慢。通过对比不同初始化方法的损失函数收敛性能,可以发现 PCA 初始化算法在初期阶段有着明显的收敛优势,能够在短时间内快速降低损失函数值,相比 He 初始化训练整体准确率提升了 3.9%。同时,He 初始化算法在中期阶段表现出色,但 PCA 初始化算法在整个训练过程中始终保持了较好的收敛性能。

3.3 模型仿真及性能分析

为分析评估本文模型的识别分类性能,本文从网络模型角度分别选取经典 CNN 模型^[18]、CLDNN 模型^[19]、Transformer 模型^[20]、ResNet 模型^[21]和 CNN-LSTM^[22]模型进行对比试验,使用本文选取的测试集分别进行识别验证,各模型在不同信噪比下的识别准确率如图 6 所示。

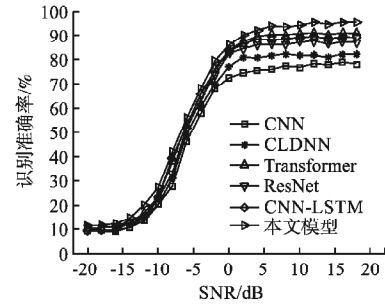


图 6 不同模型识别准确率

Fig. 6 Recognition accuracy of different models

由实验可得,各模型的识别性能均随着 SNR 的增加而提升,当信噪比小于 0dB 时,噪声信息带来的干扰导致信号有效信息湮没,因此各模型的识别效果表现不佳;当信噪比大于 0dB 时,由于数据中有效信息被提取,各模型识别效果逐渐趋于稳定。本文模型相较其他模型识别率在全信噪比条件下均有不同幅度的提升。由于本文模型利用多特征输入和特征融合注意力机制获取特征之间的互补信息,所以在不同信噪比条件下较其他模型识别率均有不同幅度的提高,当 SNR = 0dB 或更低时,本文模型较其他模型相比识别优势明显。本文模型在 SNR = 0dB 时识别率达到 87.9%,SNR = 18dB 时识别率达到 95.8%,这表明本文模型在处理噪声环境下具有更好的鲁棒性。

为进一步验证本文模型在不同信噪比下的性能和对各类调制信号的识别有效性,实验考虑了 4 种信噪比场景:(a) SNR = -6dB, (b) SNR = 0dB, (c) SNR = 6dB, (d) SNR = 18dB 的识别结果并以混淆矩阵的形式呈现。如图 7 所示,由于双谱特征经过切片和网络软阈值降噪处理后对噪声有较好的抑制作用,同时信号双谱切片特征结合信号同相正交分量特征和信号领域知识增加特征之间的互补性,在 SNR = 0dB 下,调制信号易混淆现象得到缓解,在 SNR = 18dB 时,整体识别效果达到最佳。

从图 7 中可以看出本文模型有效改善了 QAM 类内信号的识别性能,同时 WBFM 信号与 AM-DSB 信号识别也得到改善,但仍然存在小概率的误判问题,其原因是 WBFM 信号与 AM-DSB 信号在数据采样过程中存在静默期的影响;16QAM 和 64QAM 信号的内在特征差异在噪声环境中可能会被模糊,即使在高 SNR 条件下,长度为 128 的信号序列也不能提供足够有效的特征信息,导致神经网络对该类信号识别易混淆。总体来说,本文模型可以更有效的提取和分析数据中的特征信息,并以较高的精准度识别出调制信号种类。

为了对比所提出的方法整体性能及网络参数的优化效果,表 3 给出了本文模型与其他模型在参数量、识别率和训练所需时间 3 个指标的对比结果。由表 3 可知,文中模型与传统 CNN 模型相比,以较少的参数量和显著提升的识别率取得了

较佳的识别性能,识别率相比提高约 6% ~ 8%;相比 ResNet 和其他复杂模型,识别率及训练时间均表现出优异的性能,且

参数量处在较低水平. 综合来看,本文模型在性能与资源需求之间实现了良好的平衡.

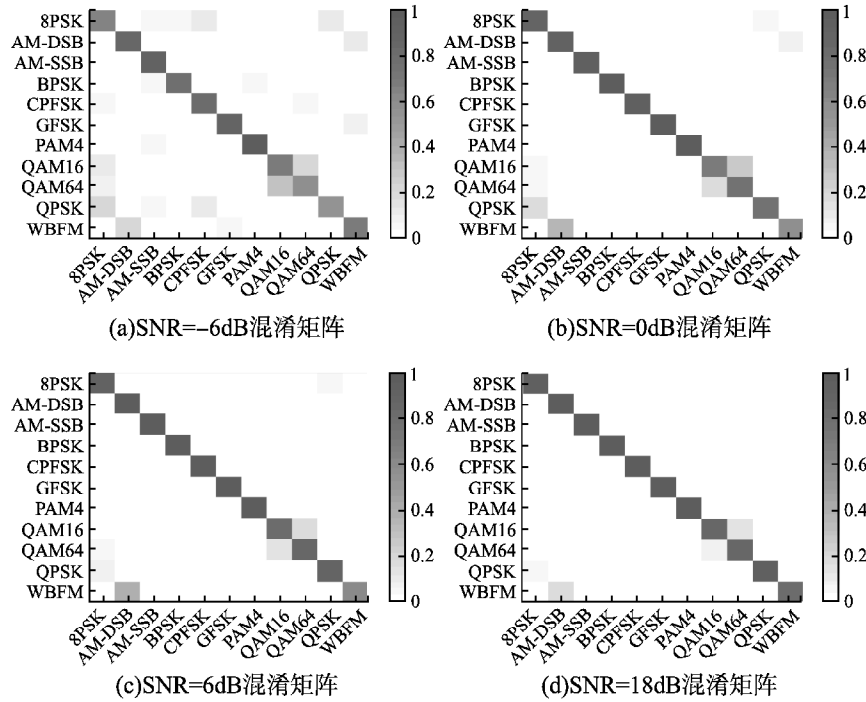


图 7 不同信噪比下模型混淆矩阵

Fig. 7 Confusion matrix of the model under different SNR

表 3 不同模型训练性能及资源需求
Table 3 Training performance and resource requirements of different models

模型	参数量/K	整体准确率/%	最高识别率/%	训练时间/s
CNN	4526.21	56.08	79.1	510
CLDNN	642.27	61.87	82.5	572
Transformer	5187.95	62.83	91.3	478
ResNet	4398.67	60.5	87.7	630
CNN-LSTM	10592.81	62.43	89.6	630
本文模型	4174.69	64.12	95.8	380

3.4 消融对比实验

为测试特征输入不同和添加注意力机制对模型识别效果的影响,针对单特征输入、多特征输入和多特征加注意力机制 3 种方案进行实验,实验结果如图 8 所示,随着迭代次数的增加,5 种不同的训练策略的识别准确率都随信噪比的升高而逐渐提升. 当采用信号多特征结合自适应注意力机制时,通过对信号时域特征和频域双谱切片特征的有效提取,并通过自适应注意力机制进行筛选冗余信息,相较于其他模型识别效果为最佳,最高可达到 95.8%.

实验表明,使用 I/Q 数据和双谱信息的组合在大多数信噪比下比仅使用单一数据的模型有更好的识别准确率,这证明了采用信号时频多特征对于改善模型性能的重要性;在信噪比大于 0dB 范围时,相对于只使用单一特征信息的模型,多特征识别有更好的识别效果,同时在 SNR = 0dB 时,添加 PCA 算法的模型相较于未添加的识别率提升 2.2%,原因是

网络模型经领域知识 PCA 初始化后能够更有效地捕捉符号周期内信号的动态特性;在添加自适应注意力机制后,在 SNR = 0dB 时识别率达到 87.9%,相较于未添加注意力机制提

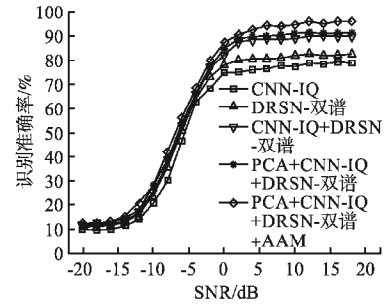


图 8 不同输入特征识别准确率

Fig. 8 Recognition accuracy with different input features

升 3.4%,表明了自适应注意力机制能够有效增强模型对重要特征的关注度,验证了自适应注意力机制的有效性和可行性. 综上,本文模型能够有效提高模型在噪声环境下识别性能.

4 结论

本文提出了一种将信号的领域知识整合并驱动网络初始化,同时联合信号多特征完成对通信调制信号识别分类的方法. 该模型利用 PCA 算法对信号领域知识进行重建主分量分析,以提取的 PCA 特征向量对神经网络模型输入层进行初始化,使网络模型更好地捕捉信号关键动态特性;同时,通过提

取通信信号多特征增强了低信噪比下特征的表征能力,并设计了自适应注意力机制筛选冗余特征信息.通过仿真实验表明,所提出的模型识别性能较其他深度学习方法表现出更好的训练性能,且低信噪比条件下识别率较其他算法有明显的提升,全信噪比下识别率最高达 95.8%.下一步工作将在现有模型基础上考虑针对实际通信中的高阶信号和多载波信号的领域知识选取和扩展,从而进一步促进领域知识驱动识别.

References:

- [1] Zheng S L, Zhou X Y, Zhang L X, et al. Toward next-generation signal intelligence; a hybrid knowledge and data-driven deep learning framework for radio signal classification[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2023, 9(3): 564-579.
- [2] Nandi A K, Azzouz E E. Algorithms for automatic modulation recognition of communication signals[J]. *IEEE Transactions on Communications*, 1998, 46(4): 431-436.
- [3] Moser E, Moran M K, Hillen E, et al. Automatic modulation classification via instantaneous features[C] // *Proceedings of 2015 National Aerospace and Electronics Conference*, 2015: 218-223.
- [4] LI Z Q, LIN Y, LI M Y, et al. Digital modulation recognition based on discriminative restricted Boltzmann machine [J]. *Journal on Communications*, 2021, 42(2): 81-91.
- [5] Yang W C, Ren K, Du Y, et al. Modulation recognition method of mixed signals based on cyclic spectrum projection [J]. *Scientific Reports*, 2023, 13(1): 21459-21459.
- [6] YANG W C, DU Y, WEN W, et al. Research on satellite signal modulation recognition based on multifractal spectrum intelligent analysis [J]. *Acta Electronica Sinica*, 2022, 50(6): 1336-1343.
- [7] Wang T, Hou Y H, Zhang H Y, et al. Deep learning based modulation recognition with multi-cue fusion[J]. *IEEE Wireless Communications Letters*, 2021, 10(8): 1757-1760.
- [8] Zhou L M, Sun Z, Wang W B. Learning to short-time fourier transform in spectrum sensing [J]. *Physical Communication*, 2017, 25(2): 420-425.
- [9] Liu B H, Ge R X, Zhu Y X, et al. IDAF: iterative dual-scale attentional fusion network for automatic modulation recognition [J]. *Sensors*, 2023, 23(19): 8134, doi:10.3390/s23198134.
- [10] Chen T, Zheng S L, Qiu K F, et al. Augmenting radio signals with wavelet transform for deep learning-based modulation recognition [J]. *IEEE Transactions on Cognitive Communications and Networking*, 2024, 10(6): 2029-2044.
- [11] Gao J X, Cao Q L, Chen Y T. MoE-AMC: enhancing automatic modulation classification performance using mixture-of-experts [J]. *arXiv:2312.02298*, 2023, 23(12): 0298.
- [12] Zhang D N, Lu Y Y, Ding W R, et al. Knowledge embedding networks based on deep learning for automatic modulation classification in cognitive radio[J]. *IEEE Transactions on Communications*, 2024, 7(2): 7814-7825.
- [13] Thorpe A J, Neary C, Djeumou F, et al. Physics-informed kernel embeddings; integrating prior system knowledge with data-driven control [J]. *American Control Conference*, 2024, 7(10): 3130-3137.
- [14] LIU Q S, LIU J J, WEI G C, et al. Research on end-to-end wireless communication system based on two-dimensional convolution [J]. *Acta Electronica Sinica*, 2023, 51(7): 1725-1733.
- [15] LIU G H, SONG B W. Individual emitter identification method of OFDM based on DRSN and integrated fusion [J]. *Signal Processing*, 2024, 40(6): 1062-1073.
- [16] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks [C] // *Proceedings of International Conference on Artificial Intelligence and Statistics*, 2010: 249-256.
- [17] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers; surpassing human-level performance on image net classification [J]. *IEEE International Conference on Computer Vision*, 2015, 12(7): 1026-1034.
- [18] O'Shea T J, Corgan J, Clancy T C. Convolutional radio modulation recognition networks [J]. *International Conference on Engineering Applications of Neural Networks*, 2016, 16(2): 41-55.
- [19] Liu X Y, Yang D Y, Gamal A E. Deep neural network architectures for modulation classification [C] // *Proceedings of 51st Asilomar Conference on Signals, Systems, and Computers*, 2017: 915-919.
- [20] Hamidi Rad S, Jain S. MCformer: a transformer based deep neural network for automatic modulation classification [J]. *IEEE Global Communications Conference*, 2021, 7(11): 1-6.
- [21] 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.
- [22] 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.

附中文参考文献:

- [4] 李正权, 林媛, 李梦雅, 等. 基于判别式受限玻尔兹曼机的数字调制识别[J]. *通信学报*, 2021, 42(2): 81-91.
- [6] 杨伟超, 杜宇, 文伟, 等. 基于多重分形谱智能分析的卫星信号调制识别研究[J]. *电子学报*, 2022, 50(6): 1336-1343.
- [14] 刘乔寿, 刘俊杰, 蔚淦丞, 等. 基于二维卷积的端到端无线通信系统研究[J]. *电子学报*, 2023, 51(7): 1725-1733.
- [15] 刘高辉, 宋博武. DRSN 与集成融合的 OFDM 辐射源个体识别方法[J]. *信号处理*, 2024, 40(6): 1062-1073.