

基于 SVD 与混合神经网络模型的自动调制识别^{*}

周顺勇^{1,2} 陆欢^{1,2} 胡琴^{1,2} 彭梓洋^{1,2} 张航领^{1,2}

(1. 四川轻化工大学自动化与信息工程学院 宜宾 644000; 2. 四川轻化工大学人工智能四川省重点实验室 宜宾 644000)

摘要: 随着现代无线通信环境中调制类型复杂性和多样性的显著增加,对自动调制识别技术的性能提出了更高要求。文章提出一种由卷积神经网络、挤压与激励模块、长短期记忆网络、门控循环单元和全连接层网络组成的混合神经网络模型,提升 AMR 技术的效率和准确性。首先,针对低信噪比环境下调制信号识别精度受限的问题,引入奇异值分解算法对接收的 I/Q 信号进行去噪,在提高信号质量的基础上提高低信噪比下调制信号的识别精度。然后,利用卷积神经网络对去噪后的信号进行多通道空间特征提取,随后加入挤压与激励模块提升特征提取的针对性,将门控循环单元和长短期记忆网络相结合,捕获信号的时间序列特征,最后,通过全连接层网络将提取的特征映射到调制方式的分类空间进行分类识别。实验结果表明,提出的网络模型在低信噪比环境下显著提高了调制识别精度,在 RadioML2016.10b 数据集上的平均识别准确率达到 64.63%,同时增强和提高了对 QAM16 与 QAM64 的区分与识别精度。

关键词: 自动调制识别; 奇异值分解; 卷积神经网络; 门控循环单元; 长短期记忆网络

中图分类号: TN92 文献标识码: A 国家标准学科分类代码: 510.5

Automatic modulation recognition based on SVD and hybrid neural network model

Zhou Shunyong^{1,2} Lu Huan^{1,2} Hu Qin^{1,2} Peng Ziyang^{1,2} Zhang Hangling^{1,2}

(1. School of Automation and Information Engineering, Sichuan University of Science & Engineering, Yibin 644000, China;

2. Artificial Intelligence Key Laboratory of Sichuan Province, Sichuan University of Science & Engineering, Yibin 644000, China)

Abstract: With the significant increase in the complexity and diversity of modulation types in modern wireless communication environments, higher requirements are placed on the performance of automatic modulation recognition technology. This paper proposes a hybrid neural network model consisting of a convolutional neural network, a squeeze and excitation module, a long short-term memory network, a gated recurrent unit, and a fully connected layer network to improve the efficiency and accuracy of AMR technology. First, to address the problem of limited modulation signal recognition accuracy in low signal-to-noise ratio environments, a singular value decomposition algorithm is introduced to denoise the received I/Q signal, thereby improving the recognition accuracy of modulation signals under low signal-to-noise ratios while improving signal quality. Then, a convolutional neural network is used to extract multi-channel spatial features from the denoised signal. Then, a squeeze and excitation module is added to improve the pertinence of feature extraction. The gated recurrent unit and the long short-term memory network are combined to capture the time series characteristics of the signal. Finally, the extracted features are mapped to the classification space of the modulation mode through a fully connected layer network for classification and recognition. Experimental results show that the proposed network model significantly improves the modulation recognition accuracy in a low signal-to-noise ratio environment. The average recognition accuracy on the RadioML2016.10b dataset reaches 64.63%. At the same time, it enhances and improves the distinction and recognition accuracy of QAM16 and QAM64.

Keywords: automatic modulation recognition; singular value decomposition; convolutional neural networks; gated recurrent units; long short-term memory network

0 引言

从未知的接收信号中分类并识别其调制技术的过程被

叫做自动调制识别^[1] (automatic modulation recognition, AMR)。随着无线通信的快速发展,通信环境日益复杂,信号的调制方式也变得越来越复杂和多样。调制识别成为连

接信号检测和信号解调之间的关键步骤,在噪声环境下,且部分或者是全部先验信息缺失的情况下,AMR 通过精心设计一系列算法,提取接收信号的特征,进而识别其调制方式,并有效估计相关调制参数,为后续的解调及其他信号处理步骤提供至关重要的支持,确保了无线通信的准确性和可靠性^[2]。

AMR 在军事和民用的通信领域均扮演者举足轻重的角色。在军事通信领域,AMR 技术不仅助力拦截敌方无线通信信号,对其实施针对性的干扰和严密监控^[3],更成为争夺和控制电磁能量的关键方法;在民用通信中,主要应用于频谱监测、动态频谱访问、发射机监控和认知无线电等多个方面,如政府及相关部门利用调制识别技术验证和规范无线信号,以提高频谱利用率,同时也防止非法使用无线频谱。

在 5G 等大规模无线通信系统中,AMR 对无线电信号的调制类型分类,被视为多类决策问题^[1]。在噪声和多径衰落的复杂条件下,确保对调制类型的高精度识别成为一项关键任务。面对这一挑战,将人工智能 (artificial intelligence, AI) 算法^[4]运用在 AMR 方面,被认为是未来的主流解决方案。

近年来,深度学习 (deep learning, DL)^[5] 广泛应用于计算机视觉和自然语言处理等领域,并获得了巨大成功。O’Shea 等^[6]首次将 DL 算法运用在 AMR 当中,将卷积神经网络 (convolutional neural networks, CNN) 用于调制识别,显示了 CNN 对复值时间信号域的适应及改进性能,同时也证明了 CNN 在低信噪比下对大型且密集编码的时间序列进行盲时间学习的方案是可行的。陶志勇等^[7]提出了一种基于双注意力机制与 Ghost 模块的轻量级 CNN 模型 AG-CNN 的调制识别方法,将调制信号映射至复空间,得到高阶特征星座图,输入到模型进行学习训练,大量节省了学习参数量的同时提高了模型性能。Li 等^[8]提出了一种混合深度学习模型 LAGNet,结合长短期记忆网络 (long short term memory, LSTM) 和图卷积网络 (graph convolutional network, GCN),使用注意力机制将 LSTM 输出映射到图中,然后 GCN 提取信号的空间特征,减少了模型的可学习参数。Elsagheer 等^[9]结合了残差神经网络 (residual neural network, ResNet) 和 LSTM,利用 ResNet 网络最小化梯度消失和 LSTM 的时间敏感性,在 18 SNR 下实现的峰值分类准确率为 92%,提升了高信噪比的分类准确率。Zhang 等^[10]提出一种将 CNN 与 Transformer 相结合的新型自动调制识别网络 CTRNet,利用 Transformer 充分捕获全局序列之间长距离依赖关系的能力及 CNN 提取信号局部特征的能力,在参数效率上表现出明显优势。

在无线电信号调制识别的任务中,DL 因其卓越的特征学习能力,展现出显著提升识别精度的潜力。在高信噪比的环境下,调制信号识别准确率通常较高。但在低信噪比条件时,由于噪声干扰和信道环境的复杂多变,调制信号

的识别准确率往往不尽如人意,这使得低信噪比下的调制分类识别成为一项极具挑战性的任务。

本文的核心工作是创新地将奇异值分解方法应用于调制信号的降噪处理,针对信号噪声在调制识别中的影响,有效降低了噪声环境对识别精度的干扰,显著提升低信噪比下的识别精度;此外,还提出了一种由 CNN、挤压与激励 (squeeze and excitation, SE)^[11] 模块、GRU-LSTM 及全连接层网络组成的全新网络模型。该模型利用 CNN 与 SE 模块结合提取多通道 I/Q 信号的空间特征信息,丰富了其空间特征的表达。采用 GRU 与 LSTM 组合使得到的时间序列信息更加充分完整。与现有的一些模型相比,此模型在提升性能的同时其复杂度有一定程度的减小,并显著提高了低信噪比下的识别准确率及整体平均准确率,特别是增强并提高了对 QAM16 和 QAM64 的区分与识别准确率,为无线电信号调制识别领域带来了新的突破,也为未来的研究提供了有价值的参考。

1 调制信号模型

在通信系统中,包含 I/Q 信号分量的等效基带信号^[12]可以表示为:

$$y(l) = Ae^{j(2\pi f_0 Tl + \theta_l)} \cdot \sum_{n=-\infty}^{\infty} x(n)h(lT - nT + \epsilon_r T) + g(l) \quad (1)$$

其中, $x(n)$ 是复基带符号序列, A 表示未知幅度的因子, $h(\cdot)$ 表示残余通道效应, T 是符号间距, ϵ_r 指时间误差,且 $0 \leq \epsilon_r \leq 1$, f_0 表示频率偏移或残余载波频率, θ_l 是相位抖动, $g(l)$ 表示复加性高斯白噪声 (additive white gaussian noise, AWGN) 序列。

接收的调制信号以离散 I/Q 分量存储,信号采样频率为 $f_s = 1/T$, 样本长度设为 L 的离散时间序列 \mathbf{y} 为:

$$\mathbf{y} = [y[0], y[1], \dots, y[L-1]]^T \in \mathbb{C}^L \quad (2)$$

$$y(k) = y_I[k] + jy_Q[k], k = 0, 1, \dots, L-1 \quad (3)$$

式中: $y(k)$ 表示同相分量 $y_I(k)$ 和正交分量 $y_Q(k)$ 的组合,即 I/Q 通道的实部和虚部。

2 奇异值分解去噪原理

奇异值分解^[13] (singular value decomposition, SVD) 作为一种矩阵分解方法,其被广泛应用于信号处理领域,特别是在信号的降噪处理中,能够有效地利用原始信号和噪声信号在能量上的可分离性^[14],通过将信号看作数据矩阵, SVD 根据原始信号和噪声信号间的奇异值对应关系,对信号进行准确的特征分析。尽管在高信噪比下,无线电信号的自动调制识别技术已经能实现较高的识别精度,但现代电子环境的噪声环境逐渐变得复杂,导致低信噪比下识别精度并不理想。为了减小噪声环境对调制识别的影响,从而提高低信噪比下信号识别的精确度,在时间序列信号预处理过程中引入奇异值分解,其通常分为矩阵构造、奇异值

分解和信号重构 3 个部分。通过对含噪信号构成的矩阵进行分解即降维和去噪处理,保留信号特征奇异值,对去噪后的信号矩阵进行重构以此得到去噪后的信号。

对带有噪声干扰的离散信号 $x(i)(i=1,2,\dots,N)$ 进行奇异值分解, $x(i)$ 为观测的时间序列, N 为信号采样点数,需要将其映射到 $m \times n$ 维的相空间,构造矩阵 \mathbf{X} :

$$\mathbf{X} = \begin{bmatrix} s(1) & s(2) & \cdots & s(n) \\ s(2) & s(3) & \cdots & s(n+1) \\ \vdots & \vdots & \vdots & \vdots \\ s(m) & s(m+1) & \cdots & s(N) \end{bmatrix} \quad (4)$$

此矩阵称为 Hankel 矩阵,是一种特殊的矩阵类型,其反对角线上的元素相等,其中 $N = m + n - 1$, m 为信号样本数, n 表示信号的特征维度且 $1 < n < N$ 。

奇异值分解实现了具有任意维数的矩阵数据集的精确分解,对于实矩阵 \mathbf{X} ,存在正交矩阵 $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_m] \in \mathbf{R}^{m \times m}$ 和 $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n] \in \mathbf{R}^{n \times n}$,使得式(5)成立。

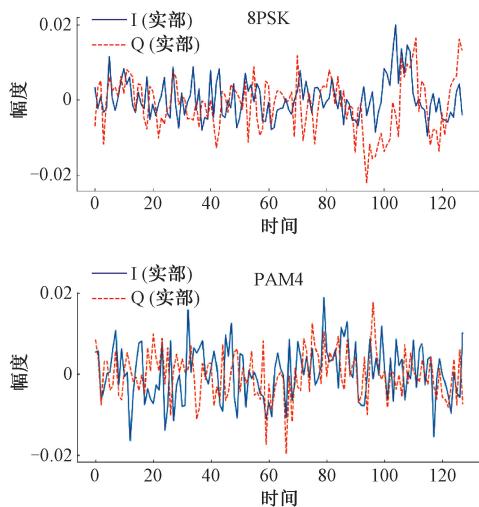
$$\mathbf{X} = \mathbf{U} \mathbf{S} \mathbf{V}^T \quad (5)$$

其中, \mathbf{U} 是 $m \times m$ 的正交矩阵,为矩阵 \mathbf{X} 的左奇异矩阵,矩阵 \mathbf{X} 的列向量 $\mathbf{u}_j(j=1,2,\dots,m)$ 为矩阵 \mathbf{X} 的左奇异向量; \mathbf{V} 是 $n \times n$ 的正交矩阵,为矩阵 \mathbf{X} 的右奇异矩阵,矩阵 \mathbf{X} 的列向量 $\mathbf{v}_l(l=1,2,\dots,n)$ 为矩阵 \mathbf{X} 的右奇异向量; \mathbf{S} 表示奇异值组成的 $m \times n$ 对角矩阵, $\mathbf{S} \in \mathbf{R}^{m \times n}$, $\mathbf{R}^{m \times n}$ 表示实数矩阵。

对角矩阵 \mathbf{S} 表示为:

$$\mathbf{S} = \begin{cases} (\text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r), 0), & m < n \\ (\text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)), & m = n \\ (\text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r), 0)^T, & m > n \end{cases} \quad (6)$$

式中: $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq \dots \geq \sigma_r \geq 0$ ($p = 1, 2, \dots, r$),为矩阵 \mathbf{X} 的全部非零奇异值,若矩阵 \mathbf{S} 为行满秩或列满秩矩阵,则 $r = \min\{m, n\}$,若矩阵 \mathbf{S} 不满秩,则 $r = \text{rank}(\mathbf{X})$ 。



(a) 原始信号示例 (SNR=0 dB)
(a) Original signal example (SNR=0 dB)

矩阵 \mathbf{S} 中前几个值较大,包含了矩阵 \mathbf{X} 的全部信息。

通过保留奇异值矩阵 \mathbf{S} 的前 k 个最大的奇异值,将信号矩阵 \mathbf{X} 降维到更低维度的信号矩阵 \mathbf{X}_k :

$$\mathbf{X}_k = \mathbf{U}_k \mathbf{S}_k \mathbf{V}_k^T \quad (7)$$

式中: \mathbf{U}_k 是左奇异矩阵 \mathbf{U} 中前 k 列构成的矩阵, \mathbf{S}_k 是对角矩阵 \mathbf{S} 中前 k 个最大的奇异值构成的对角矩阵, \mathbf{V}_k 是右奇异矩阵 \mathbf{V} 中前 k 行构成的矩阵。

在降维后的信号矩阵 \mathbf{X}_k 中,奇异值按大小递减排列,较小的奇异值对应的成分是噪声信息,将较小的奇异值置零或只保留前面部分奇异值。在本文中,选择保留的奇异值数量为 $k = 20$,保留的 k 个奇异值包含了信号能量,后续的奇异值对应信号的噪声分量, k 的大小决定了控制去噪的程度,以得到去除后的信号矩阵 $\mathbf{X}_{k,\text{denoised}}$,实现对信号的去噪效果。之后使用去噪后的信号矩阵 $\mathbf{X}_{k,\text{denoised}}$ 和保留的矩阵 \mathbf{U}_k 和 \mathbf{V}_k 重新构建信号,得到最终的去噪信号矩阵 $\mathbf{X}_{\text{denoised}}$:

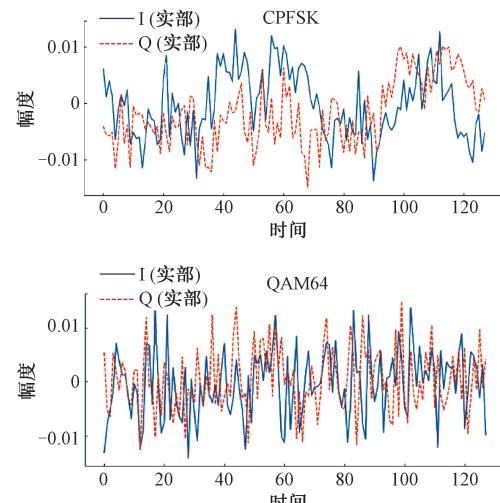
$$\mathbf{X}_{\text{denoised}} = \mathbf{U}_k \mathbf{S}_{k,\text{denoised}} \mathbf{V}_k^T \quad (8)$$

其中, $\mathbf{S}_{k,\text{denoised}}$ 是 \mathbf{S}_k 中保留了较大奇异值对应的对角矩阵。

图 1 展示了不同调制信号在信噪比为 0 dB 条件下原始信号与经过奇异值分解去噪之后的效果图。

3 调制模型结构

在 SVD 对信号去噪处理基础上,构建了一种基于深度学习的自动调制识别模型,通过融合不同网络结构的优势,实现对无线通信信号调制方式的精确识别。该模型由卷积神经网络模块、SE 模块、循环神经网络 (recurrent neural network, RNN) 及全连接深度网络模块构成,其网络结构如图 2 所示。无线电信号样本的 I/Q 分量可以看作时空特征的表示,所以通过构建 CNN 结构提取信号 I/Q 分量的空间特征,加入 SE 模块进一步提取通道信息特征。将



(a) 原始信号示例 (SNR=0 dB)
(a) Original signal example (SNR=0 dB)

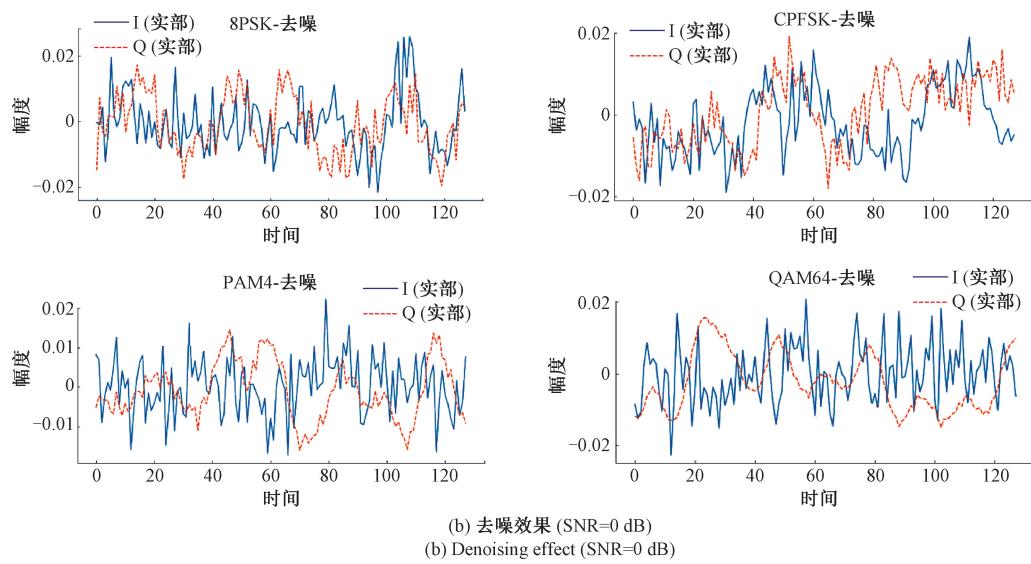


图 1 调制信号去噪前后对比

Fig. 1 Comparison of modulation signal before and after denoising

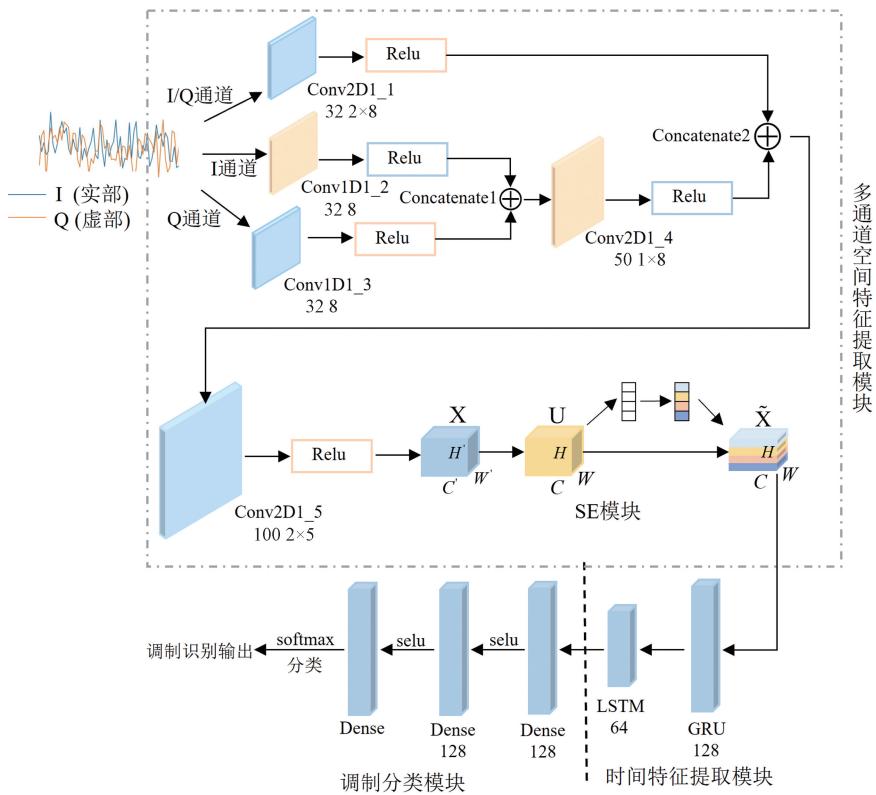


图 2 调制识别网络结构

Fig. 2 Modulation recognition network structure

RNN 中的门控循环单元 (gated recurrent units, GRU) 和 LSTM 结合用以提取信号的时间序列特征, 利用 CNN、GRU 和 LSTM 的互补性及其相互的协同作用尽可能提取信号的时空特征, 以提高对调制信号的识别准确率, 最后的分类识别由全连接层网络完成。

3.1 多通道特征提取

接收的调制信号为同相和正交分量, 尽管他们彼此正交, 但因为信号的幅度和相位特征影响其正交关系, 使得两个通道出现差异, 因此考虑将输入的基带 I/Q 无线电信号分为单独的 I/Q 通道、I 通道和 Q 通道来充分提取信号

的数据特征。CNN 的优势在于局部特征信息的提取,通过设定多层 CNN 对输入数据进行卷积运算,随着 CNN 层的卷积核逐步增加,实现了对数据深层特征的提取。在构造的 CNN 中采用二维卷积层网络 Conv1D1_1 提取 I/Q 通道的空间特征,其卷积核数量设定为 32,大小为 2×8 ,并采用 Relu 激活函数来提高非线性映射能力。Conv1D1_2 和 Conv1D1_3 分别表示提取 I 通道和 Q 通道的一维卷积层,卷积核数量为 32,长度为 8,在利用一维卷积层提取空间特征的同时加深信号之间的特征融合^[15],并尽可能地提取 I/Q 信号各分量的隐藏瞬时特征,同样采用 Relu 激活函数。然后将 Conv1D1_2 及 Conv1D1_3 两个卷积层进行拼接形成新的特征图作为下一个卷积层网络的输入。对拼接后的特征图进行二维卷积即卷积层 Conv2D1_4,卷积核数量为 50,大小为 1×8 ,通过进一步卷积来提取特征。其卷积输出结果与第一层卷积层网络输出进行合并以丰富特征,最后,利用卷积层 Conv2D1_5 对拼接后的特征进行二维卷积,卷积层 Conv2D1_5 的卷积核数量为 100,大小为 2×5 。通过对输入数据利用不同大小的卷积核对多通道提取不同尺度的特征信息,有效利用 I/Q 通道、I 通道及 Q 通道的信息相关性及互补性。

为了加强 CNN 对特征信息的关注程度,以提高网络性能。在构造的 CNN 网络中加入了通道注意力机制即 SE 模块,如图 1 的 SE 模块,这种注意力机制包括挤压与激励两个部分。挤压部分也是指全局信息的嵌入,因为输入特征中信号的学习过滤器使用局部感受野来操作,输出 \mathbf{U} 的每个单元无法利用该区域之外的上下文信息^[11]。将全局空间信息压缩并通过全局平均池化生成通道统计数据,统计量 $z \in \mathbb{R}^c$ 通过空间维度 $H \times W$ 缩小 \mathbf{U} 生成, z 的第 c 个元素表示为:

$$z_c = \mathbf{F}_{sq}(\mathbf{u}_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (9)$$

z 的每个元素对应变换输出 \mathbf{U} 的全局感受野。

激励部分通过两个全连接层的维度变换来学习通道之间的非线性交互,利用学习获得的权重调整特征图,来提高网络的表示能力。

3.2 时间特征提取和分类

RNN 在时序数据处理能力方面有很大的优势,LSTM^[16]和 GRU^[17]都属于改进的 RNN。当 RNN 展开为深度前馈神经网络时,梯度可能会消失,从而导致 RNN 无法捕获输入时间序列的长期依赖性。GRU 和 LSTM 作为 RNN 特定结构的变体,本质上为克服这一缺点产生,在处理时间序列问题方面表现出较强的能力。基于信号样本的时间特性,LSTM 对其具有长期依赖性,在不同时间节点的动态变化特征可以避免梯度消失。GRU 相较于 LSTM,计算参数和训练时间有所减少且收敛速度加快,整体性能有所提升。因此考虑将通过 CNN 结构捕获的空间特征输入到 GRU-LSTM 组合结构中提取其中的时序特

征,在减少模型计算量的同时能较为完整地提取时间序列特征信息。

基本的 LSTM 模型由输入门、遗忘门、任意数量的隐藏层和输出门构成,如图 3 所示。

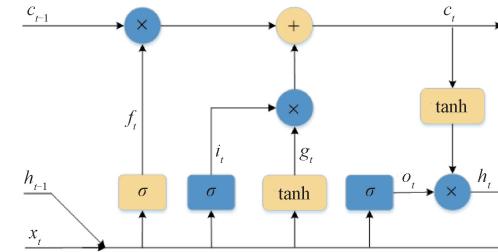


图 3 LSTM 基本结构

Fig. 3 LSTM basic structure

其中,输入门决定更新的值,其神经元的数量由输入数据的特征数量决定,遗忘门决定是否是部分还是完全忘记之前学习的信息,输出门决定最终输出内容。执行步骤表示为:

$$f_t = \sigma(\mathbf{W}_f \cdot [h_{t-1}, x_t] + \mathbf{b}_f) \quad (10)$$

式中: \mathbf{W}_f 表示权重矩阵, h_{t-1} 是前一个时间段的输出, x_t 表示当前输入, \mathbf{b}_f 表示遗忘门偏差向量。输入通过 sigmoid 激活函数 σ 进行处理并投影到 $[0, 1]$ 范围内, 选择数据以保留先前的单元状态或删除数据。

$$i_t = \sigma(\mathbf{W}_i \cdot [h_{t-1}, x_t] + \mathbf{b}_i) \quad (11)$$

式中: i_t 表示输入门, 同样地, \mathbf{W}_i 表示权重矩阵, \mathbf{b}_i 表示偏差向量。

$$g_t = \tanh(\mathbf{W}_g \cdot [h_{t-1}, x_t] + \mathbf{b}_g) \quad (12)$$

$$c_t = f_t * c_{t-1} + i_t * g_t \quad (13)$$

式中: g_t 表示候选单元, 将输入数据通过 \tanh 函数投影到 $[-1, 1]$ 范围, 计算输入门的候选值即 g_t , \mathbf{W}_g 表示权重矩阵, \mathbf{b}_g 表示偏差向量, c_t 是新的单元状态, c_{t-1} 表示前一时间段学习的信息, $*$ 表示哈达玛积, 即逐元素相乘。

$$o_t = \sigma(\mathbf{W}_o \cdot [h_{t-1}, x_t] + \mathbf{b}_o) \quad (14)$$

$$h_t = o_t * \tanh(c_t) \quad (15)$$

式中: o_t 是输出门, 决定模型的输出, h_t 表示隐藏状态, 其输出由 sigmoid 和 \tanh 激活函数决定。 \mathbf{W}_o 表示权重矩阵, \mathbf{b}_o 表示偏差向量。

与 LSTM 相比,GRU 的优势是网络复杂性减小,缩短了网络训练时间,减少了所需的计算资源。GRU 利用网络神经元内的门,由更新门、重置门和输出门构成,如图 4 所示。

GRU 将 LSTM 的遗忘门和输入门合并为更新门,并且合并了数据单元状态和隐藏状态,因此整体结构比 LSTM 更简单。其执行步骤表示如下:

$$r_t = \sigma(\mathbf{W}_r \cdot [h_{t-1}, x_t]) \quad (16)$$

$$z_t = \sigma(\mathbf{W}_z \cdot [h_{t-1}, x_t]) \quad (17)$$

$$\tilde{h}_t = \tanh(\mathbf{W} \cdot [r_t * h_{t-1}, x_t]) \quad (18)$$

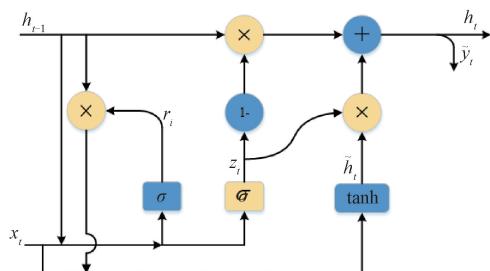


图 4 GRU 基本结构

Fig. 4 GRU basic structure

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (19)$$

式中: r_t 是重置门, 将新的输入信息与之前记忆相结合, z_t 是更新门, 控制先前输入保留的数据量, 从而允许网络更长时间地保留信息, 同时也可以更新隐藏层的状态即 \tilde{h}_t , 并减少了网络参数。 h_{t-1} 表示先前的输出, x_t 是当前输入, h_t 是输出门, σ 表示 sigmoid 激活函数, 取 0~1 之间的值, 利于数据的遗忘和保留, $*$ 表示哈达玛积, 即逐元素相乘。

在本文构造的时间特征提取结构中, 将 128 个单元的 GRU 和 64 个单元的 LSTM 结合构成 GRU-LSTM 混合模型, 利用 LSTM 的长时间记忆能力和 GRU 处理短时间依赖关系的优势, 更好地提取输入信号的时间序列特征。GRU 因为结构相对简单, 且计算效率高, 与 LSTM 复杂的门机制组成相比, 平衡了计算效率, 也增强了其记忆能力的互补性。此外, GRU-LSTM 的混合结构可以减少过拟合, 使模型具备更好的适应能力, 同时, 在面对输入的复杂时间序列数据时, GRU-LSTM 模型能够更灵活地选择合适的门机制, 以适应不同的需求。

采用两个具有 128 个神经元的全连接层, 将 LSTM 层的输出映射到新的特征分类空间, 来加深网络, 并使用缩放指数线性单元 (scaled exponential linear units, selu) 激活^[18], 如式(20)所示, selu 激活函数可以进行自我归一化, 降低了梯度消失的概率, 提高了训练效率和稳定性。由于 selu 激活函数能保持激活值的均值和方差恒定, 使得梯度流动更加稳定, 因此可以加速模型的收敛。最后的 Dense 层网络对调制信号进行分类, 利用 softmax 激活函数将其转换为概率分布, 使其适用于多分类分布的目标预测, 输出分类结果。

$$selu(x) = \begin{cases} x, & x > 0 \\ \alpha e^x - \alpha, & x \leq 0 \end{cases} \quad (20)$$

4 实验数据集与仿真环境

此实验采用的数据集是使用 GNU 无线电生成的调制数据集 RadioML2016.10a 与 RadioML2016.10b。数据集 RadioML2016.10a 的详细参数如表 1 所示。RadioML2016.10b 包含 1 200 000 个信号样本和除模拟调制方式 AM-SSB 的

10 种调制信号, 是数据集 RadioML2016.10a 的扩展版, 两者都是在类似于实际环境的情况下生成的。将实验的数据集按照 6:2:2 的比例划分为训练集、验证集和测试集。模型训练批量大小 (batch size) 为 1 024, 优化器是 Adam, 损失函数是分类交叉熵。当实验验证在 50 个 epoch 内没有减少时, 则停止训练, 训练好模型将以最小的验证损失保存。实验的仿真环境内容如表 2 所示。

表 1 RadioML2016.10 数据集

Table 1 RadioML2016.10 dataset

数据集	GNU 无线电生成
调制格式	数字调制格式: 8PSK、BPSK、CPFSK、GFSK、PAM4、16QAM、64QAM、QPSK; 模拟调制格式: AM-DSB、AM-SSB、WBFM
数据格式/维度	同相正交分量(IQ)/2×128
样本长度	128
样本总量	220 000
采样率	128 μs
符号调制率	8
信噪比范围	-20:2:18 dB
信道环境	多径衰落、中心频率偏移、加性高斯白噪声、采样率偏移等

表 2 实验仿真环境

Table 2 Experimental simulation environment

配置内容	配置参数
硬件配置	Intel(R) Core(TM) i5-12490F; NVIDIA GeForce RTX 3060
内存	32 GB
操作系统	Windows 10 专业版
深度学习框架	Keras2.9.0

5 实验结果讨论

RadioML2016.10a 不同信噪比下各类调制信号的调制分类识别精度如图 5 所示。随着信噪比的不断提高, 除 AM-DSB 和 WBFM 外, 各类调制信号的识别准确率在不断上升直至平稳, 平稳之后的波动范围较小。调制信号 CPFSK、GFSK、PAM4 都逐渐达到了 100% 的识别准确率, QPSK、BPSK 和 8PSK 的识别精度接近 100%。因为 AM-DSB 和 WBFM 在静默期时信号几乎为零, 在此期间, WBFM 作为语音信号本身是没有内容的, 网络是对无效信号进行分类的, 所以识别精度无法达到理想效果。此外, 本文改进的网络模型也提高了 16QAM 和 64QAM 的识别精度, 使得其识别精度在 18 dB 下都达到了 90% 以上, 两者的时间序列数据很类似, 因此在分类过程中很容易受到噪声影响。

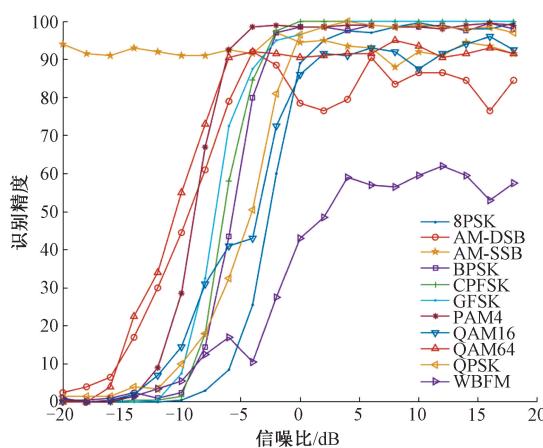


图 5 不同 SNR 下每类调制信号的分类精度

Fig. 5 Classification accuracy of each type of modulation signal under different SNR

图 6 展示了训练损失和验证误差的过程, 图中给出了 100 个 epoch 内的损失和误差。从图中可以看出, 整个训练过程的损失逐渐下降直至平稳。模型在进行训练期间, 会设置提早停止训练, 这是因为神经网络训练时性能可能会逐渐慢慢下降, 提前停止训练就是为了让网络模型的有效性提高, 也是为了保证神经网络映射的能力不会过度拟合。在验证数据集的泛化误差开始增大时停止训练, 可以有效减少过拟合, 并提高神经网络的泛化能力。

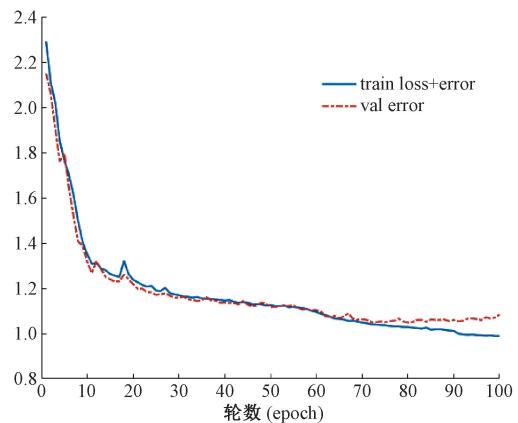


图 6 训练和验证损失过程

Fig. 6 Training and validation loss process

5.1 对比实验结果

图 7 展示了数据集 RadioML2016.10a 和 RadioML2016.10b 与其他网络模型调制识别精度的对比结果。对比模型选取了以下几种模型:

CGDNet^[19]: 从输入的原始信号中提取特征, 由 CNN、GRU 和 DNN 组合的网络结构, CNN 由三层滤波器个数为 50 的卷积层串联形成, GRU 学习信号的时间特征, DNN 进行分类;

CNN1^[6]: CNN 结构由两个卷积层和两个密集全连接层组成, 密集层大小为 128, 由 n 类神经元组成, 各层均使

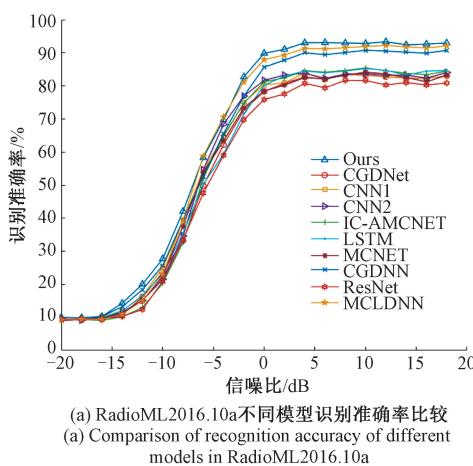
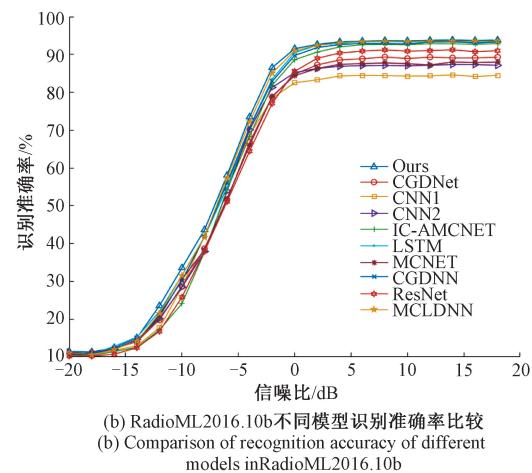
(a) RadioML2016.10a 不同模型识别准确率比较
(a) Comparison of recognition accuracy of different models in RadioML2016.10a(b) RadioML2016.10b 不同模型识别准确率比较
(b) Comparison of recognition accuracy of different models in RadioML2016.10b

图 7 不同模型识别准确率比较

Fig. 7 Comparison of recognition accuracy of different models

用 Relu 激活函数;

CNN2^[20]: 由 4 个卷积层和 2 个密集全连接层组成, 模型复杂度较低且训练时间较短;

IC-AMCNet^[21]: 由 4 层卷积层、最大池化层和全连接层组成, 在前 2 个卷积层和第 4 个卷积层后加入最大池化层, 以减少学习参数量, 还加入 Flatten 层将二维矩阵转换为向量;

LSTM^[22]: 将时域调制信号的幅度和相位作为输入, 由两层 128 个单元的 LSTM 组成, softmax 层进行分类;

MCNET^[1]: 利用最大池化层进行下采样, 由 6 个非对称核的卷积块执行, 同时减少参数, 部署跳越连接以防止梯度消失问题;

CGDNN^[23]: 由参数估计器、参数变换器、CNN 和 GRU 组合形成的模型, 将经过参数估计和变换后的数据

输入 CNN 中, CNN 由 75 个卷积核和 25 个卷积核的卷积层组成, GRU 有 128 个单元, 为了压缩模型, 采用网络修剪方法以减少冗余;

ResNet^[24]: 两层 256 个滤波器和两层 80 个滤波器的卷积层组成, 第一层的输出转发到后两层以此缓解梯度消失问题;

MCLDNN^[25]: 利用 CNN 形成多通道特征提取, 采用两个 LSTM 网络层提取时间信息, FC 层进行分类识别。

从图中的对比结果中可以明显看出本文实验设计的自动调制识别网络模型的整体平均分类识别准确率要高于其他网络模型, 从 -14 dB 开始, 本文的模型对 RadioML2016. 10a 和 RadioML2016. 10b 两者的识别精度值相对于其他模型在逐渐增大, 随着信噪比的增加逐渐趋于

平稳。其中, RadioML2016. 10a 的最大识别准确率达到了 93.36%, RadioML2016. 10b 的最大识别准确率达到 93.79%, 在 -8 dB 时, 两者的识别精度都达到了 40% 以上。不管是整体识别准确率还是低信噪比下的识别准确率, 都表明了本文构造的网络模型在两个数据集都表现良好。

表 3 展示了两类数据集关于模型总参数量、训练时间和整体的平均调制识别准确率的对比。本文的算法结构较其他大部分算法的训练周期要长一点, 但整体识别精度得到了大幅度的提升, 模型参数量与 CNN1、CNN2、IC-AMCNet、ResNet 及 MCLDNN 相比也相对减少了许多。因此, 训练时间的差异是可以被接受的, 以延长训练时间的代价获得了更好的识别精度, 从整体性能方面证明了本文构造深度学习网络模型的有效性。

表 3 模型性能比较

Table 3 Model performance comparison

框架类型	RadioML2016. 10a			RadioML2016. 10b		
	总参数量	训练时间 (s/周期)	平均识别 准确率/%	总参数量	训练时间 (s/周期)	平均识别 准确率/%
Ours	261 410	10	62.28	261 281	53	64.63
CGDNet	124 933	5	50.01	124 676	27	60.56
CNN1	1 592 383	2	55.50	1 592 126	13	58.77
CNN2	858 123	2	55.54	857 994	67	60.43
IC-AMCNet	1 264 011	3	56.80	1 263 882	19	62.40
LSTM	201 099	7	57.78	200 970	38	63.51
MCNET	121 611	20	56.56	121 226	110	59.56
CGDNN	71 871	3	59.90	71 742	33	63.85
ResNet	3 098 283	20	54.46	3 098 154	106	60.94
MCLDNN	406 199	13	61.70	406 070	68	63.76

5.2 去噪效果分析

-8、-6、-4、-2 和 0 dB 下不同模型的分类识别结果。

从表 4 内数据可以看出, 本文的模型与其他模型相比,

表 4 低信噪比下 RadioML2016. 10a 不同模型的调制识别结果

Table 4 Modulation recognition results of different models of RadioML2016. 10a under low signal-to-noise ratio

框架类型	识别准确率/%						
	-12 dB	-10 dB	-8 dB	-6 dB	-4 dB	-2 dB	0 dB
Ours	20.05	27.72	42.09	58.32	69.77	82.73	89.86
CGDNet	16.09	23.68	33.27	49.82	62.23	72.68	78.32
CNN1	16.81	25.14	38.50	55.04	64.63	75.05	80.32
CNN2	15.14	21.55	34.64	54.68	68.41	76.95	81.68
IC-AMCNet	13.00	20.32	32.68	52.09	65.00	74.77	81.27
LSTM	16.64	23.18	35.50	50.27	59.41	71.50	80.18
MCNET	15.00	22.41	37.64	53.95	63.41	73.36	78.45
CGDNN	18.41	26.18	39.45	52.64	65.69	76.91	85.68
ResNet	12.41	20.95	33.50	47.59	59.05	69.77	75.86
MCLDNN	15.00	22.91	39.36	58.77	70.59	81.22	87.95

识别精度在低信噪比下有不同程度的提升,识别精度分别最大提高了 7.64%、7.4%、9.41%、12.96%、14%。尽管在-6 dB 和-4 dB 时分别低于 MC LDNN 模型 0.45% 和 0.82%,但整体提升效果是比较良好的,也证明了奇异值分解对调制信号降噪的有效。

如图 8 所示,显示了 RadioML2016.10a 在信噪比为 0 dB 下各种网络模型的混淆矩阵。混淆矩阵是指不同调制信号的错误分类情况,如果预测精度的结果越高,那么对角线结果就越高,对应的蓝色深度也越深,表示了预测标签与实际标签的一致程度,对角线的结果对应为调制类型的识

别准确率。在 RadioML2016.10a 中,AM-DSB 与 WBFM 两类调制信号因为调制信息易丢失所以很容易导致错误分类,从图中可以看出,本文的模型与 ResNet 相比,对两者的预测精度是更高的。较其他模型而言,对 8PSK 的预测精度也得到了不同程度的提升,最大与 CGDNet 模型相比提升了 24%。此外,QAM16 和 QAM64 由于其具有重叠的星座映射,在低信噪比下的噪声环境干扰也很容易造成错误识别。本文构造的网络模型大大减少了调制信号 QAM16 和 QAM64 之间的混淆,两类调制类型的识别准确率都达到了 90% 以上,这明显优于其他模型。

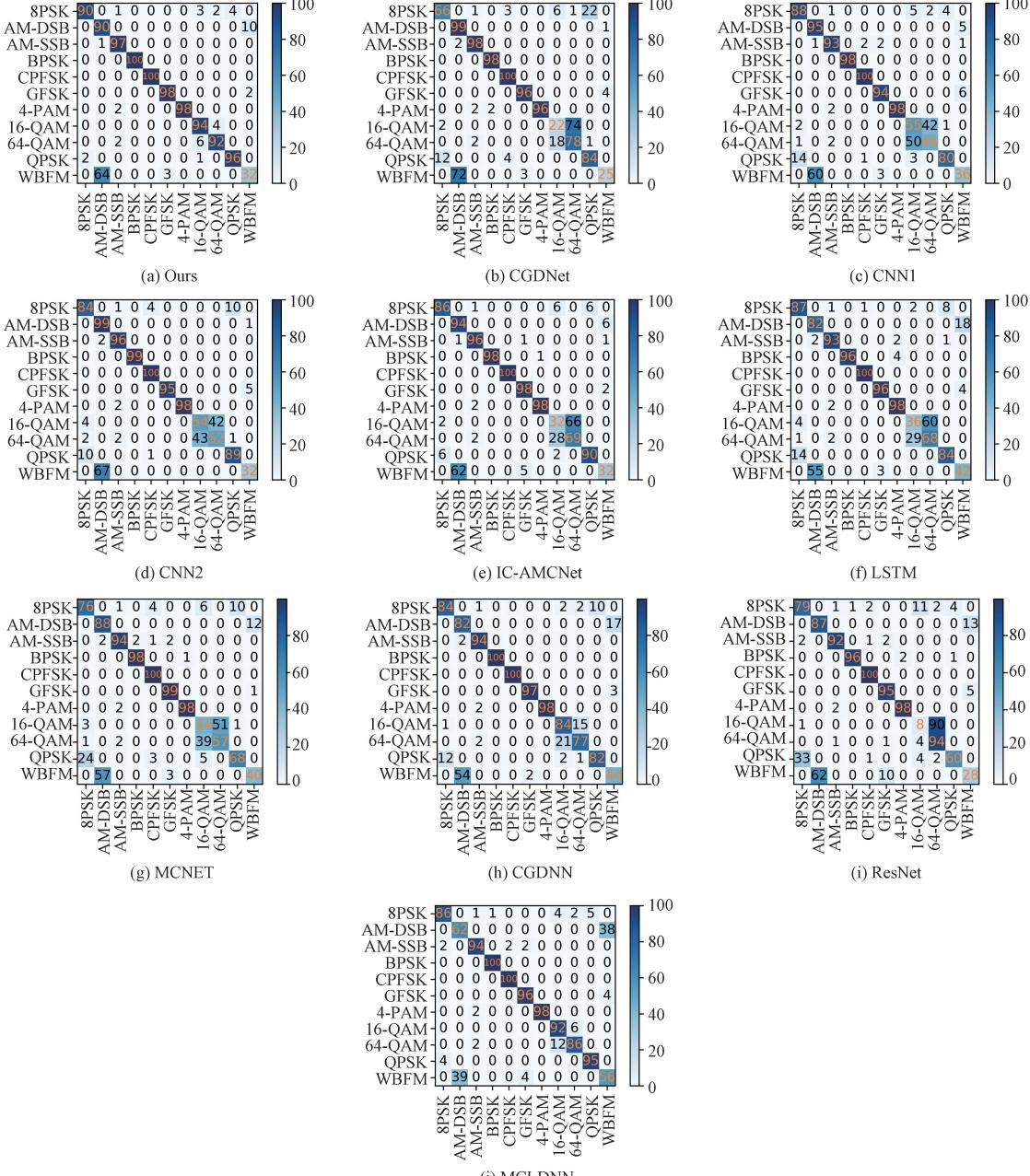


图 8 不同框架在 0 dB 的混淆矩阵对比

Fig. 8 Confusion matrix comparison of different frameworks at 0 dB

6 结 论

针对低信噪比下的噪声环境干扰导致调制信号分类识别精度无法提高的问题,本文提出了一种基于奇异值分解去噪的神经网络模型。该模型首先利用 SVD 对时间序列信号进行降维和去噪处理,保留信号的特征奇异值,通过重构获得清晰的信号。随后,将经去噪后的调制信号作为网络模型的输入,利用新构造的神经网络模型对其进行特征提取和分类识别。本文提出的网络模型由 CNN、SE 模块和 GRU-LSTM 组成。模型将输入的 I/Q 序列信号分为多通道输入,利用 CNN 网络的卷积层分别提取 3 个通道的空间特征,增强信息互补性^[26],然后,用 SE 模块加强 CNN 对特征信息提取的关注。接着,使用 GRU-LSTM 结构捕捉信号的时间序列特征,最后通过全连接层网络实现分类识别。实验结果表明,本文算法提高了低信噪比下对调制信号的识别精度,并在减少模型参数量的同时提升了整体的识别准确率。尤其是,该模型提高了对调制信号 QAM16 和 QAM64 的区分,对两者的识别准确率均超过 90%。然而,该模型还无法有效区分 AM-DSB 信号和 WBFM 信号,这将成为未来工作的关注点,同时,也计划探索融合多种特征的方法提升模型在低信噪比下对调制信号的识别性能。

参考文献

- [1] 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.
- [2] ZHANG W N, SUN Y Q, XUE K L, et al. Research on modulation recognition algorithm based on channel and spatial self-attention mechanism [J]. IEEE Access, 2023, 11: 68617-68631.
- [3] WU X W, WEI SH Y, ZHOU Y. Deep multi-scale representation learning with attention for automatic modulation classification[C]. IEEE 2022 International Joint Conference on Neural Networks(IJCNN), 2022: 1-8.
- [4] ASLAM M W, ZHU ZH CH, NANDI A K. Automatic modulation classification using combination of genetic programming and KNN [J]. IEEE Transactions on Wireless Communications, 2012, 11(8): 2742-2750.
- [5] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [6] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks[C]. Engineering Applications of Neural Networks: 17th International Conference, 2016: 213-226.
- [7] 陶志勇, 闫明豪, 刘影, 等. 基于 AG-CNN 的轻量级调制识别方法[J]. 电子测量与仪器学报, 2022, 36(4): 241-249.
- [8] TAO ZH Y, YAN M H, LIU Y, et al. Light weight modulation recognition method based on AG-CNN[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(4): 241-249.
- [9] LI ZH, LU G Y, LI Y X, et al. LAGNet: A hybrid deep learning model for automatic modulation recognition[C]. 2024 IEEE Wireless Communications and Networking Conference, 2024: 1-6.
- [10] 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.
- [11] ZHANG W N, XUE K L, YAO AI Q, et al. CTRNet: An automatic modulation recognition based on transformer-CNN neural network[J]. Electronics, 2024, 13(17): 3408.
- [12] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [13] SWAMI A, SADLER B M. Hierarchical digital modulation classification using cumulants[J]. IEEE Transactions on Communications, 2000, 48 (3): 416-429.
- [14] LI L N, WEI X L. Suppression method of partial discharge interferences based on singular value decomposition and improved empirical mode decomposition[J]. Energies, 2021, 14(24): 8579.
- [15] BEARD J K. Singular value decomposition of a matrix representation of the Costas condition for Costas array selection[J]. IEEE Transactions on Aerospace and Electronic Systems, 2020, 57(2): 1139-1161.
- [16] ZHANG H, NIE R H, LIN M H, et al. A deep learning based algorithm with multi-level feature extraction for automatic modulation recognition[J]. Wireless Networks, 2021, 27(7): 4665-4676.
- [17] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [18] CHO K, MERRIENBOER B V, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. ArXiv preprint arXiv:1406.1078, 2014.
- [19] KLAMBAUER G, UNTERTHINER T, MAYR A, et al. Self-normalizing neural networks[J]. Advances in Neural Information Processing Systems, 2017, 30,

DOI:10.48550/arXiv.1706.02515.

- [19] 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.
- [20] TEKBİYİL K, EKTI 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 (VTC2020-Spring), 2020: 1-6.
- [21] 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.
- [22] 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.
- [23] ZHANG F X, LUO CH 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.
- [24] 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: 103650.
- [25] XU J L, LUO CH 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.
- [26] 钱磊,吴昊,乔晓强,等.基于特征融合的调制识别增强与迁移演化 [J]. 电子测量技术, 2022, 45 (18): 153-160.
- QIAN L, WU H, QIAO X Q, et al. Modulation recognition enhancement and migration evolution based on feature fusion [J]. Electronic Measurement Technology, 2022, 45(18):153-160.

作者简介

周顺勇,副教授,主要研究方向为图像处理、信息与信号处理。

E-mail: zsy22219750517@163.com

陆欢(通信作者),硕士研究生,主要研究方向为信号处理。

E-mail: 15883522610@163.com

胡琴,硕士研究生,主要研究方向为信号处理。

E-mail: 2094758200@qq.com

彭梓洋,硕士研究生,主要研究方向为图像处理、目标检测。

E-mail: 1281082359@qq.com

张航领,硕士研究生,主要研究方向为信号处理。

E-mail: 1294755782@qq.com