

# 基于改进 Transformer 的自动调制识别方法

战权海, 张雄伟, 宋磊, 孙蒙, 周振吉, 李涛

(陆军工程大学指挥控制工程学院, 南京 210007)

**摘要:** 调制识别技术在认知无线电、电子侦察对抗领域已得到广泛应用。近年来, 得益于深度神经网络强大的特征提取能力, 基于深度学习的自动调制识别研究已经取得了很大的进展。在实际的调制识别场景中, 调制信号通常传输没有语义的比特序列, 并且每个调制符号以均匀的概率出现在波形中, 因此其特征信息均匀地分布在信号数据当中。但现有的基于深度学习的自动调制识别方法通常采用卷积神经网络或循环神经网络结构, 难以适应上述场景中的数据分布特点, 未能充分利用长序列数据中的全局特性信息, 调制识别率有待进一步提升。本文提出了一种基于改进 Transformer 的自动调制识别方法——AMR-former。首先对信号数据进行预处理, 强化信号数据中的时序特性; 随后, 结合多头注意力机制与长短期记忆网络设计实现了用于特征提取的 AMR-Encoder 结构, 有效提高了全局时序特征的提取能力, 为后续识别分类提供了更为丰富的数据表示。通过在开源数据集 RadioML 2016.10a 上的实验表明, AMR-former 方法在信噪比为 0~18 dB 条件下的平均识别准确度达到 91.90%, 相比于典型的 GRU、PET-CGDNN、LSTM 和 MCLDNN 等网络结构分别提高了 6.38%、2.15%、1.99% 和 1.75%。

**关键词:** 深度学习; 调制识别; Transformer; 时序特征; 注意力机制

**中图分类号:** TP391 **文献标志码:** A

## Automatic Modulation Recognition Method Based on Improved Transformer

ZHAN Quanhai, ZHANG Xiongwei, SONG Lei, SUN Meng, ZHOU Zhenji, LI Tao

(College of Command & Control Engineering, Army Engineering University of PLA, Nanjing 210007, China)

**Abstract:** Modulation recognition technology has been widely used in cognitive radio and electronic reconnaissance countermeasures. In recent years, thanks to the powerful feature extraction ability of deep neural networks, the research of automatic modulation recognition based on deep learning has made great progress. In practical modulation recognition scenarios, modulation signals usually transmit bit sequences without semantic information, and each modulation symbol appears in waveforms with uniform probability, so its feature information is uniformly distributed in signal. However, existing automatic modulation recognition methods based on deep learning usually use structures of convolutional neural network (CNN) or recurrent neural network (RNN). They are difficult to be adapted to the data distribution in the scenarios above and thus fail to make full use of the global characteristics of long sequential data. Therefore, the accuracy of modulation recognition can be further improved by exploiting the sequential information. In this paper, an automatic modulation recognition method based on improved Transformer, AMR-former, is proposed. Firstly, the input signal is preprocessed to strengthen the

temporal characteristics. Then, the AMR-Encoder structure for feature extraction is designed and implemented by combining the multi-head attention mechanism and long short-term memory (LSTM) network, which effectively improves the ability of global temporal feature extraction and provides richer representations for the subsequent recognition and classification. Experiments on the RadioML 2016.10a dataset show that the average recognition accuracy of the AMR-former method reaches 91.90% with the signal-to-noise ratio (SNR) from 0 dB to 18 dB, which is 6.38%, 2.15%, 1.99% and 1.75% higher than the typical networks of GRU, PET-CGDNN, LSTM and MCLDNN, respectively.

**Key words:** deep learning; modulation recognition; Transformer; temporal feature; attention mechanism

## 引 言

随着无线通信技术的飞速进步,自动调制识别(Automatic modulation recognition, AMR)技术已经在认知无线电、电子侦察对抗等领域得到了广泛的应用<sup>[1-3]</sup>。AMR技术旨在未知的信道条件下从接收到的信号中识别调制样式,但现代通信系统中出现了越来越多的调制类型和信号变体,对自动调制识别系统提出了更高的要求<sup>[4]</sup>。传统的AMR方法可以分为基于似然理论的方法和基于特征分类的方法。基于似然理论的方法<sup>[5]</sup>通过计算未知信号调制样式可能性的似然度来估计分类,此类方法的计算成本较高,且需要较为理想的信道环境;基于特征分类的方法<sup>[6]</sup>需要通过手动构建特征表示,此类方法可以凭借较低的复杂度而实现次优的分类效果,但往往需要较为丰富的人工特征提取经验。而基于深度学习的自动调制识别方法以数据为驱动,通过神经网络自动提取信号数据中的特征信息,避免了对人工经验的高度依赖,在低成本下实现了较好的识别效果。因此,基于深度学习的自动调制识别方法被广泛研究,大量研究者也提出了基于深度学习的AMR(Deep learning AMR, DL-AMR)模型,这些方法凭借强大的特征自动提取能力取得了较好的性能与效果<sup>[7-8]</sup>。

现有的DL-AMR模型已使用多种神经网络实现了自动调制识别方法<sup>[9-10]</sup>。文献[11-12]考虑了多种信道条件下的无线信号,使用卷积神经网络(Convolutional neural network, CNN)对信号中包含的不同信息进行特征提取,证明了其在无线信号调制识别方面的突出性能。后续的研究发现,循环神经网络(Recurrent neural network, RNN)具有保留信号动态时域信息的特性,因此基于RNN的调制识别模型<sup>[7]</sup>利用递归结构对输入信号的时序性特征进行提取,同样取得了较好的识别性能。此外,文献[13-15]将CNN与RNN进行了组合,充分利用两者对空间与时间特性的提取能力,识别性能得到了进一步提升。

在实际的调制识别任务中,调制信号通常传输没有语义信息的比特序列,每个调制符号以均匀的概率出现在波形中<sup>[16]</sup>。因此,对于AMR任务更有价值的特征信息应该均匀分布于调制信号中,前后采样点之间较强的时序相关性应作为支撑识别分类的重要依据。然而,基于CNN的自动调制识别方法往往聚焦于信号的局部信息,未能充分利用信号数据中的上下文信息进行特征提取,因此在调制识别性能上可进一步获得提升。基于RNN的自动调制识别方法采用线性结构提取调制信号中的时序特性,过程中易受到梯度消失和梯度爆炸问题的制约,在处理较长序列的信号数据时,识别性能难以得到保证。Transformer模型<sup>[17]</sup>最早被应用于自然语言处理领域,其利用基于注意力机制的编码器-解码器结构处理长序列数据和捕获全局特性信息。得益于自注意力机制的并行处理能力,Transformer结构更适合处理较长的序列输入,同时在计算数据之间相关性时所需的操作次数不会随着距离增长,相比CNN与RNN能够更好地捕获序列数据的全局依赖关系,从而提供更优异的特征提取能力。

本文受Transformer模型启发,将编码器结构应用于调制识别任务中,提出了一种名为

AMR-former的自动调制识别方法。AMR-former方法首先对原始信号数据进行预处理,强化前后数据的时序相关性;随后,利用融合多头注意力机制的长短时记忆(Long short-term memory, LSTM)网络设计AMR-Encoder编码器结构,保留了信号数据中的全局特性,可有效实现对调制信号长期时序特性的提取;最后,由多层感知机(Multi-layer perception, MLP)网络根据提取到的特征完成调制信号的识别分类。在公开数据集RML 2016.10a上的实验结果表明,本文所提方法能够更充分地利用信号数据中的全局信息进行特征提取,在0~18 dB的信噪比(Signal-to-noise ratio, SNR)条件下,相较于典型的GRU、PET-CGDNN、LSTM和MCLDNN等网络,平均识别准确度分别提高了6.38%、2.15%、1.99%和1.75%。

## 1 AMR-former自动调制识别方法

为了适应调制信号特征均匀分布的特点,本文提出了一种基于改进Transformer的自动调制识别方法——AMR-former。该方法旨在对输入信号数据的全局特征信息进行提取分类,输出相应的调制样式类别,模型结构如图1所示。AMR-former方法包括数据预处理、特征提取和分类器3个模块,各个模块串联相接。数据预处理模块负责接收原始信号数据,首先通过固定步长的分割操作实现对编码器输入长度的适配,形成多个等长的信号序列块,随后由多个Dense层完成线性映射,获取信号数据的高维表示,最后嵌入位置编码信息,用以增强信号序列块的时序特性。特征提取模块名为AMR-Encoder,采用Transformer网络中的编码器结构,以融合多头注意力机制的LSTM网络为核心,并行接收多个信号序列块,实现时序信号的全局特征提取。在获得信号数据的特征矩阵后,由MLP作为分类器完成识别分类工作。下面分别对方法中的各个模块进行详细阐述。

### 1.1 信号模型

一般地,典型通信信号模型可被表示为

$$r(t) = F(s(t), m) * h(t) + n(t) \quad (1)$$

式中: $s(t)$ 表示发射器的基带信号, $h(t)$ 表示无线信道脉冲响应, $n(t)$ 表示传输过程中的加性高斯白噪声。调制的过程首先选定调制样式 $m$ ,随后利用调制器 $F$ 将基带信号 $s(t)$ 映射为调制信号,最后转化为适合在信道中传输的信号。AMR任务的目的是在缺少先验信息的情况下从接收信号 $r(t)$ 中提取信号特征,估计调制样式 $m$ 。

### 1.2 信号数据表示

本文主要对11种常见的无线电通信信号的调制样式进行识别,其中包含8种数字调制样式(8PSK、BPSK、CPFSK、GFSK、PAM4、16QAM、64QAM和QPSK)和3种模拟调制样式(AM-DSB、

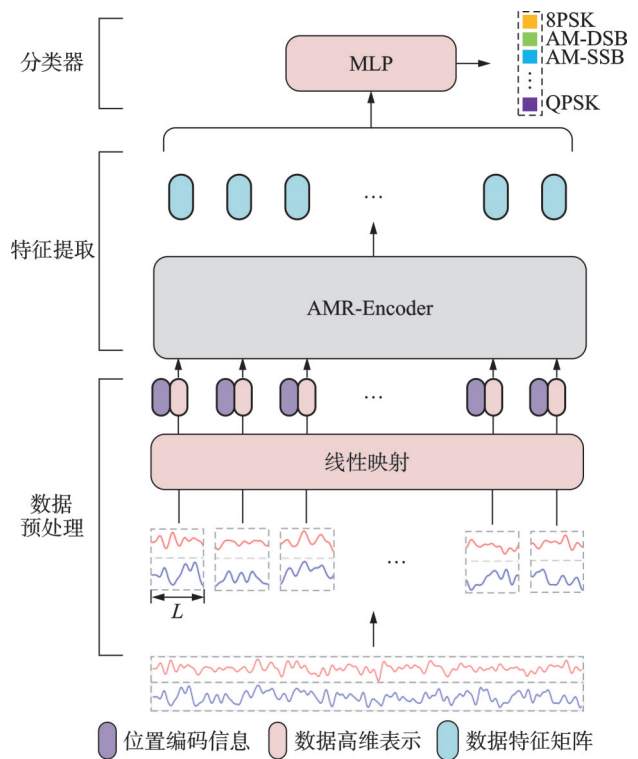


图1 AMR-former模型结构

Fig.1 Structure of AMR-former model

AM-SSB 和 WBFM)。接收信号使用同相正交(In-phase and quadrature, IQ)的数据形式进行表示,其中 I 分量代表信号的实部, Q 分量代表信号的虚部,其中内含了不同信号调制样式特有的幅度与相位信息。上述 11 种信号的 IQ 时域表示如图 2 所示。本文采用的调制信号数据是以固定采样率  $f_s = 200$  kHz 对调制信号的 IQ 数据进行采样得到的时域离散复信号  $r(t)$ ,其数学表达式为

$$r(n) = r_I(n) + ir_Q(n) \quad n = 0, 1, \dots, M - 1 \quad (2)$$

式中:  $r_I(n)$  和  $r_Q(n)$  表示离散信号  $r(n)$  的第  $n$  个同向正交分量;  $M$  为对应序列的长度;  $i = \sqrt{-1}$ 。

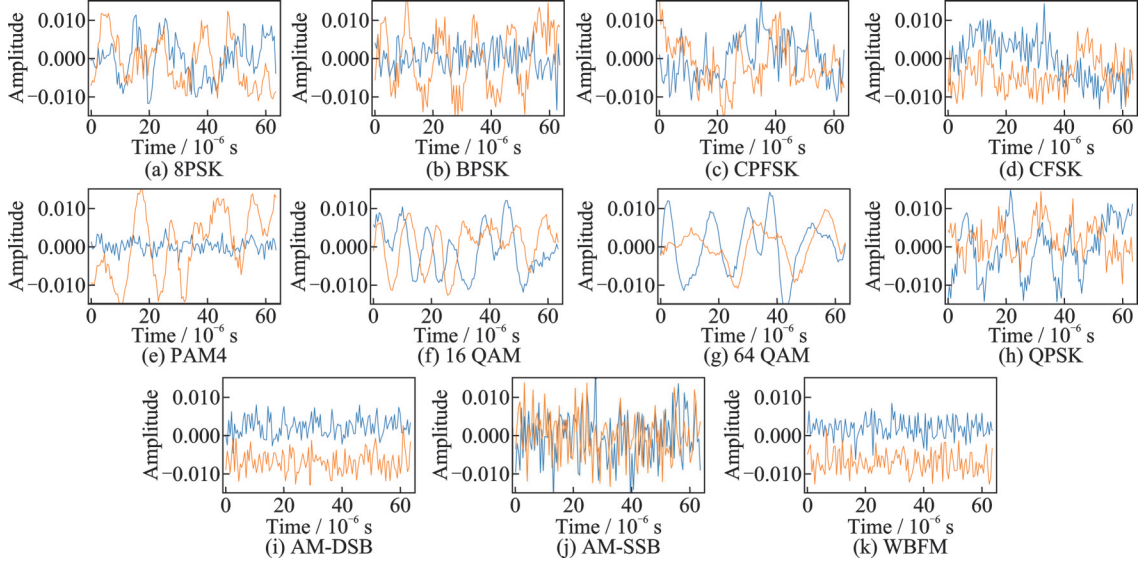


图2 调制信号 I/Q 时域图

Fig.2 Time-domain I/Q waveforms of modulated signal

### 1.3 数据预处理

现有的 DL-AMR 方法通常利用 CNN 或 RNN 直接对原始信号数据进行特征嵌入与提取,过程中未能充分考虑特征均匀分布的特点,从而忽略了全局特性对于识别性能的提升。为解决这一问题,AMR-former 的数据预处理模块采用序列结构,通过数据分割、线性映射与位置编码信息嵌入等手段,将更多的信号序列关联起来,增强前后数据的时序特性。首先,为保证原始调制信号的信息完整性并结合 Transformer 网络在特征提取上的优势,对输入的信号数据进行标准归一化与分割处理,如图 1 所示。输入信号  $x \in \mathbf{R}^{H \times W \times C}$  将以  $L$  为步长被分割为一组信号序列块  $x_p \in \mathbf{R}^{N \times (H \times L \times C)}$ ,其中  $(H, W)$  为原始输入信号数据的行列数,  $C$  为通道数,  $(H, L)$  为每个序列块的分辨率,  $N = W/L$  为得到的序列块个数,这也是多头注意力处理模块的有效序列处理长度。此时,输入数据  $X$  即可表示为:  $X = [x_1, x_2, \dots, x_N]$ ,其中  $x_t$  可表示为

$$x_t = \begin{bmatrix} r_I(tL), r_I(tL + 1), \dots, r_I(tL + L - 1) \\ r_Q(tL), r_Q(tL + 1), \dots, r_Q(tL + L - 1) \end{bmatrix} \quad t = 1, 2, \dots, N \quad (3)$$

由于 Transformer 网络通常不能直接学习训练序列中的位置关系,为了使 AMR-former 能够更好地提取复杂信号数据中的有效信息,数据预处理模块还采用了线性映射与位置编码信息嵌入手段。线性映射将分割后获得的信号序列通过 Dense 层执行全连接映射,以此得到数据的高维度表示  $\tilde{X}$ 。随后,位置编码算法使用正弦和余弦函数生成可学习的位置编码矩阵  $X_{\text{pos}}$ ,其位置信息代表了序列块在整个信号样本中的时序信息,使得后续自注意力机制能够有效地考虑不同序列块的时序特性。最后,将两者进行融合作为特征提取模块的输入,上述过程可具体表示为



$$\tilde{X} = Wx_t + b \quad (4)$$

$$X_{(\text{pos}, n)} = \begin{cases} \sin(\text{pos}/10\,000^{2i/d_{\text{model}}}) & n = 2i \\ \cos(\text{pos}/10\,000^{2i/d_{\text{model}}}) & n = 2i + 1 \end{cases} \quad (5)$$

$$X_{\text{Input}} = \tilde{X} + X_{\text{pos}} \quad (6)$$

式中:  $W$  表示权重矩阵,  $b$  表示偏置矩阵,  $\text{pos}$  表示序列所处位置,  $n$  表示对应特征维度值,  $d_{\text{model}}$  表示模型嵌入向量的维度。

#### 1.4 AMR-Encoder 特征提取

经过数据预处理模块的处理,接收信号被分割为若干个包含位置信息的序列块。为了有效提取序列块中的全局特征,提出了由 12 层融合多头注意力机制的 LSTM 网络所构成的特征提取模块——AMR-Encoder,其内部由层归一化、多头自注意力机制、残差连接、层归一化、LSTM 网络和残差连接顺序组成,结构如图 3 所示。其中,多头注意力机制由  $h = 16$  个缩放点乘注意力模块组合而成,其内部结构如图 4 所示。

首先将数据预处理模块的输出变换成维度为  $d_k$  的查询矩阵( $Q$ )、键矩阵( $K$ )和值矩阵( $V$ ),随后根据式(7)计算相应的注意力特征,即

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (7)$$

式中  $d_k$  为  $\tilde{X}$  所规定的特征维度,通过标准归一化至  $[0, 1]$  区间。输入序列块被划分为  $h$  个子空间,分别经由每个独立的注意力变换矩阵进行计算,得到加权后的特征矩阵( $\text{head}_1 \sim \text{head}_h$ ),可表示为

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (8)$$

式中  $W_i^Q$ 、 $W_i^K$ 、 $W_i^V$  分别表示  $Q$ 、 $K$ 、 $V$  的权重矩阵。

最后,将每个 head 进行拼接,通过线性变换的整合得到更丰富、更全面的序列特征表示,最终的输出结果可由 MultiHead( $Q, K, V$ ) 表示为

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W^O \quad (9)$$

式中:Concat 表示拼接操作,  $\text{head}_i$  表示注意力机制的第  $i$  个特征矩阵,  $W^O$  表示线性变换的权重。

在多头注意力机制所计算出的全局特征矩阵中,包含了特征序列中的大量位置相关性信息。由于信号序列的位置编码信息是根据时序序列生成而来,因此可进一步利用 LSTM 来提取和建模全局特征矩阵中的时序特性。LSTM 是一种被广泛用于解决长序列依赖问题的循环神经网络,可以有效地表达和传递较长时间序列中的信息,其模型结构如图 5 所示,包含记忆单元状态的遗忘门、输入门和输出门。遗忘门决定了需要保存或遗忘的某一时刻的状态数据,输入门确定了来自输入的某个特定时段相应信息的重要性,而输出门控制了输出的状态数据流量,其计算过程分别如下

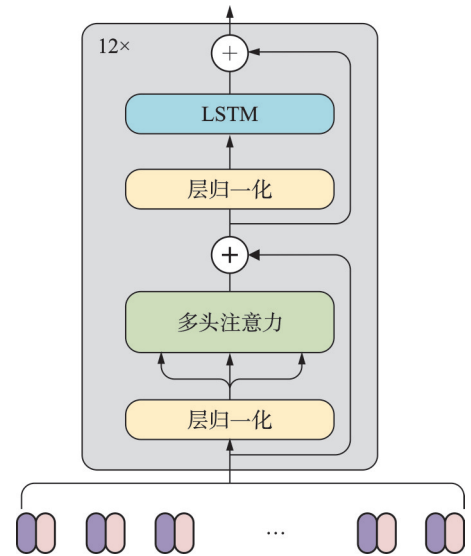


图 3 AMR-Encoder 结构

Fig.3 Structure of AMR-Encoder

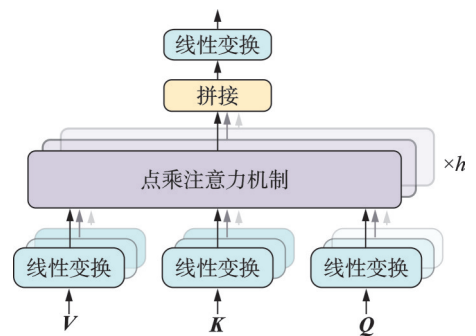


图 4 多头注意力机制

Fig.4 Multi-head attention mechanism

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

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

$$\tilde{C}_t = \tanh(\mathbf{W}_c[h_{t-1}, x_t] + \mathbf{b}_c) \quad (12)$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t \quad (13)$$

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

$$h_t = O_t \otimes \tanh C_t \quad (15)$$

式中:  $x_t, h_t$  分别为输入序列和输出值;  $f_t, i_t$  和  $O_t$  分别代表遗忘门、输入门和输出门;  $\tilde{C}_t$  为  $t$  时刻的候选状态;  $C_{t-1}$  和  $C_t$  分别表示  $t-1$  与  $t$  时刻的细胞状态;  $\mathbf{W}$  与  $\mathbf{b}$  分别代表相应的权重矩阵与偏置项;  $\sigma, \tanh$  和  $\otimes$  分别对应 sigmoid 激活函数、双曲正切函数与 Hadamard 积。

### 1.5 MLP 分类器

在经过 AMR-Encoder 的处理后, 每个序列块都被计算出相应的特征矩阵, 此时由 MLP 分类器完成最后的识别分类任务。MLP 是一种基于前馈结构的神经网络模型, 由输入层、隐藏层和输出层全连接构成, 其结构如图 6 所示。输入层接收来自特征提取模块的特征向量; 隐藏层由 2 层全连接网络构成, 采用 ReLU 作为激活函数来获取特征矩阵中所蕴涵的特征信息; 输出层通过 Softmax 函数计算各个类别的概率分布, 并选择最高概率的类别作为识别分类的结果。

## 2 实验与结果分析

本节首先介绍用于仿真实验的基准数据集和实验设置, 随后研究不同分割长度对 AMR-former 的性能影响, 并将优化后的 AMR-former 与其他 4 种调制识别方法进行性能对比, 最后对 AMR-former 性能方面进行详细分析。

### 2.1 数据集与实验设置

实验采用调制识别领域中 RadioML 2016.10a 公开数据集作为信号输入<sup>[18]</sup>。该数据集由 GNU Radio 对动态信道模型进行仿真生成, 其中包含了 11 种常见调制样式的 220 000 个调制信号, 综合考虑了加性高斯白噪声、选择性衰落(莱斯+瑞利)、中心频率偏移和采样率偏移等信道环境因素, 样本信噪比以 2 dB 为间隔均匀分布在  $-20 \sim 18$  dB 之间, 每个样本由 IQ 两路信号的 128 个采样点组成。

在上述数据集上, 按 6:2:2 的比例对不同调制样式进行分层抽样, 将数据集划分为训练集、验证集与测试集。训练过程中, Batch size 设置为 400, 初始学习率为 0.01, 若连续 10 个阶段验证损失没有减少, 则学习率将下调 50%, 若连续 50 个阶段验证损失没有减少, 则训练提前终止。具体优化参数如表 1 所示。

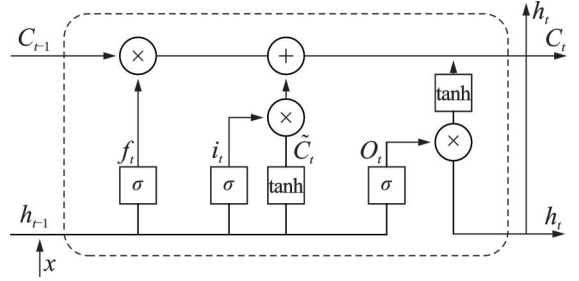


图 5 LSTM 模型结构

Fig.5 Structure of LSTM model

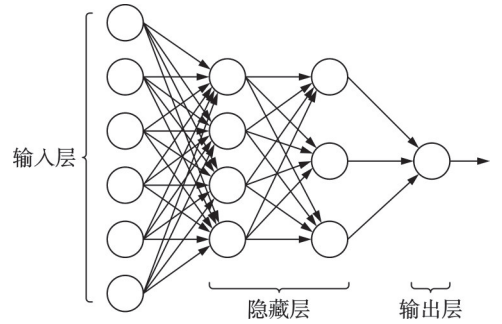


图 6 MLP 分类器结构

Fig.6 Structure of MLP classifier

表 1 数据集和模型训练参数

Table 1 Overview of datasets and model training parameters

训练参数	参数值
调制类别	8PSK, BPSK, CPFSK, GFSK, PAM4, 16QAM, 64QAM, QPSK, AM-DSB, AM-SSB, WBFM
SNR 范围	$-20 \text{ dB} : 2 \text{ dB} : 18 \text{ dB}$
样本长度	128
数据集划分	6:2:2
优化器	Adam
Batch size	400
最大训练轮数	500
初始学习率	0.01

## 2.2 信号截取长度的选择

在数据预处理过程中,信号数据需要通过分割为固定的长度 $L$ ,才能输入到深度网络做进一步的处理。然而,信号截取长度 $L$ 的选择直接影响到特征提取时的有效输入长度,长度越小,特征提取的有效输入长度越长;反之,有效输入长度越短。为了验证分割长度对识别分类精度的影响,分别以分割长度 $L=2,4,8,16,32$ 开展实验测试,其结果如图7所示。

观察实验结果发现,分割步长在低信噪比条件下( $\text{SNR} < -2 \text{ dB}$ )对识别准确度的影响较小。此外,在实际调制识别任务中,通常考虑信噪比 $0 \text{ dB}$ 以上的信号作为输入。因此,选取 $\text{SNR} \geq 0 \text{ dB}$ 的识别准确度数据进行具体分析,结果如表2所示。可以明显看出,当分割步长每缩小1倍时,识别准确提升 $1\% \sim 4\%$ 。特别当步长由16变为8时,平均准确度提升 $4.15\%$ 。由此可见,不同的分割步长所引起的有效输入长度的变化对识别准确度有较大影响。当分割步长 $L=2$ 时,实验获得了最佳的识别效果,推测是由于更多的位置信息被用于强化信号数据的时序特性,保留了更多的全局特征信息,为此选取 $L=2$ 来完成模型对比实验。

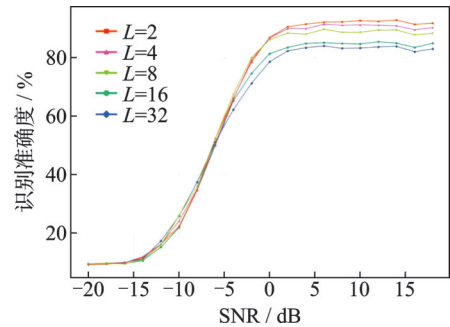


图7 不同分割步长下的识别准确度

Fig.7 Recognition accuracy at different segmentation step sizes

表2  $\text{SNR} \geq 0 \text{ dB}$  时不同分割步长下的识别准确度

Table 2 Recognition accuracy at  $\text{SNR} \geq 0 \text{ dB}$  for different segmentation step sizes

SNR/dB	分割步长				
	$L=2$	$L=4$	$L=8$	$L=16$	$L=32$
0	0.869 5	0.868 2	0.862 3	0.814 1	0.786 4
2	0.906 4	0.899 5	0.884 5	0.835 5	0.823 6
4	0.915 0	0.900 0	0.880 5	0.849 1	0.834 5
6	0.922 3	0.914 5	0.897 7	0.851 4	0.840 5
8	0.922 7	0.911 4	0.887 3	0.848 6	0.832 3
10	0.927 3	0.912 7	0.887 3	0.847 3	0.833 6
12	0.925 0	0.911 4	0.894 5	0.855 0	0.837 3
14	0.929 1	0.909 5	0.895 5	0.850 0	0.839 1
16	0.914 5	0.895 9	0.878 2	0.835 5	0.820 9
18	0.918 2	0.903 2	0.883 6	0.849 5	0.830 5
平均值	0.915 0	0.902 6	0.885 1	0.843 6	0.827 9

## 2.3 模型对比实验

为进一步验证所提模型的性能,选取现有较为先进的调制识别方法 MCLDNN<sup>[15]</sup>、LSTM<sup>[19]</sup>、PET-CGDNN<sup>[20]</sup>和GRU<sup>[21]</sup>4种模型进行对比实验。

(1)MCLDNN网络采用互补式处理方法,将IQ数据拆解成3路输入网络,使用卷积模块实现特征的提取与融合,再通过2层LSTM和全连接网络进行特征提取,过程中的卷积核分别设置为(8,50)、(2,8,50)、(1,8,50)和(2,5,100),LSTM与全连接的单元数分别为250与128。

(2)LSTM网络使用幅度/相位特征,经由2层LSTM网络,最后使用1层全连接输出分类结果。

(3)PET-CGDNN和GRU模型使用卷积层与门控循环单元实现了IQ数据的识别工作,过程中卷积核分别设置为(2,8,75)和(1,5,25),门控循环单元数设置为128。其他参数设置与本文所提模型一致。

图8展示了上述模型与 AMR-former 在 RadioML 2016.10a 上的对比实验结果。GRU 模型总体识别准确率为 58.42%，而在 0 dB 以上的识别准确率则为 85.52%，该数值明显低于其他模型。PET-CGDNN 模型与 LSTM 模型的整体识别准确率相同，均为 60.47%，但 0 dB 以上的 PET-CGDNN 与 LSTM 模型分别实现了 89.75% 和 89.81% 的识别准确率，相比之下 LSTM 表现更加出色；MCLDNN 模型的识别准确率达到 61.53%，且在 0 dB 及以上的识别率超过 90.15%；本文提出的 AMR-former 模型总体准确率为 61.43%，其中在 0 dB 以上环境下的识别准确率高达 91.90%，总体识别精度与最优的 MCLDNN 近似，但在高信噪比条件下表现出更为优秀的分类识别效果。

2.4 性能分析

在模式识别领域，通常使用混淆矩阵来比较预测值与真实值之间的差异。AMR-former 模型在 SNR 分别为 2、6、10 和 14 dB 条件下的混淆矩阵结果如图 9 所示。可以发现，随着信噪比的提高，模型

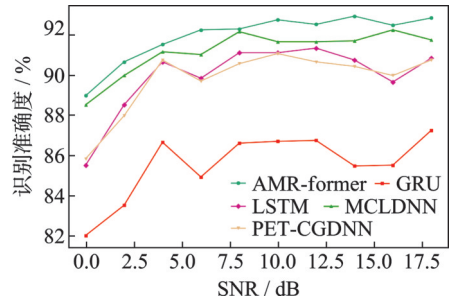
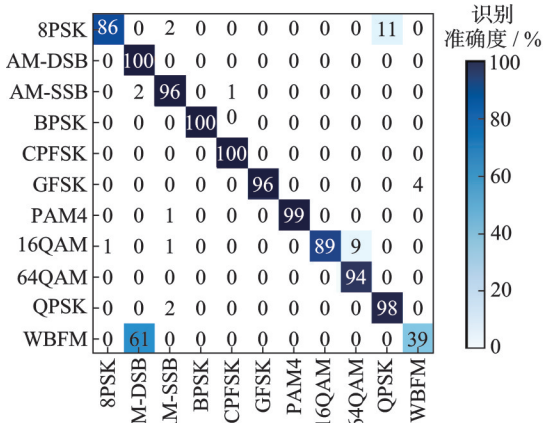
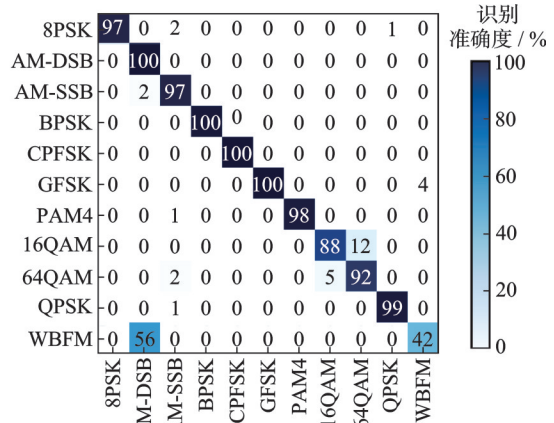


图8 模型对比实验结果

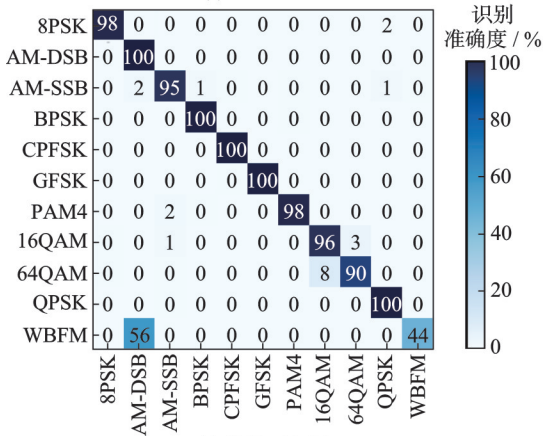
Fig.8 Comparison of experimental results across models



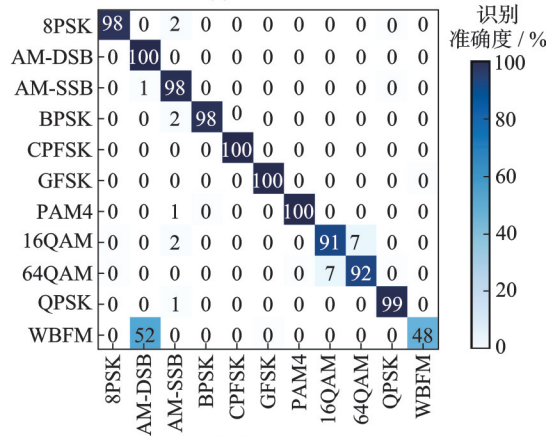
(a) SNR=2 dB



(b) SNR=6 dB



(c) SNR=10 dB



(d) SNR=14 dB

图9 AMR-former在信噪比分别为2、6、10和14 dB时的混淆矩阵

Fig.9 Confusion matrices of AMR-former at SNRs of 2, 6, 10 and 14 dB



的调制识别性能越来越强,对 AM-DSB、AM-SSB、BPSK、CPFSK、GFSK、PAM4、QPSK 7类调制样式的识别精度均超过 96%。

此外,实验发现,即便在高信噪比条件下,模型对于 WBFM 与 AM-DSB、16QAM 与 64QAM 的识别上仍存在一定的混淆。对于 WBFM 与 AM-DSB 两种调制方式,两者均属于模拟连续调制方式,同时在生成信号数据时,两类信号的采样过程存在较长的静默期,因此在高维特征的表现上差异较小,加大了对两种信号区分的难度。而对于 16QAM 和 64QAM 两种调制方式,两者都基于正交振幅调制,调制符号存在着大量相似分布,尽管符号数量有所不同,但符号间距相对较小。因此,两者调制符号之间的特征差异更加微弱,分类的难度相对较高。

### 3 结束语

通信信号的调制识别技术是无线通信的重要组成部分,在自动频谱监测等领域已得到广泛的应用。为了更加充分地利用调制信号中的全局特征并提高识别精度,本文提出了一种基于改进 Transformer 的自动调制识别方法——AMR-former。首先通过对接收信号的预处理,增强了信号数据的时序相关性。随后采用编码器结构,结合多头注意力机制与 LSTM 设计了全局时序特征提取模块——AMR-Encoder,更加有效地实现了对不同调制样式的特征提取,为 MLP 分类器提供了识别分类的关键特征。为验证本文所提方法的识别效果,在公开数据集 RML 2016.10a 上进行了实验。结果表明,在信噪比为 0~18 dB 的条件下,AMR-former 的平均识别准确度达到 91.90%,相比于较为先进的 GRU、PET-CGDNN、LSTM 和 MCLDNN 等网络结构分别提高了 6.38%、2.15%、1.99% 和 1.75%。在未来研究中,可考虑进一步提升模型的计算效率,通过降低复杂度和优化算法流程等方式来满足调制识别任务的实时性需求。

#### 参考文献:

- [1] BHATTI F A, KHAN M J, SELIM A, et al. Shared spectrum monitoring using deep learning[J]. *IEEE Transactions on Cognitive Communications and Networking*, 2021, 7(4): 1171-1185.
- [2] ZHANG W, FENG M, KRUNZ M, et al. Signal detection and classification in shared spectrum: A deep learning approach [C]//*Proceedings of IEEE Conference on Computer Communications*. [S.l.]: IEEE, 2021: 1-10.
- [3] 江汉, 胡林, 李文, 等. 信号调制识别的对抗样本攻防技术研究进展[J]. *数据采集与处理*, 2023, 38(6): 1235-1256.  
JIANG Han, HU Lin, LI Wen, et al. Research progress of adversarial attack and defense for signal modulation recognition[J]. *Journal of Data Acquisition and Processing*, 2023, 38(6): 1235-1256.
- [4] 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.
- [5] DULEK B. Online hybrid likelihood based modulation classification using multiple sensors[J]. *IEEE Transactions on Wireless Communications*, 2017, 16(8): 4984-5000.
- [6] HAZZA A, SHOAB M, ALSHEBEILI S A, et al. An overview of feature-based methods for digital modulation classification [C]//*Proceedings of the 1st International Conference on Communications, Signal Processing, and Their Applications (ICCSPA)*. Sharjah: IEEE, 2013: 1-6.
- [7] KE Z, VIKALO H. Real-time radio technology and modulation classification via an LSTM auto-encoder[J]. *IEEE Transactions on Wireless Communications*, 2021, 21(1): 370-382.
- [8] TEKBIYIK K, EKTİ A R, GÖRÇİN A, et al. Robust and fast automatic modulation classification with CNN under multipath fading channels[C]//*Proceedings of 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*. [S.l.]: IEEE, 2020: 1-6.
- [9] ZHANG F, LUO C, XU J, et al. Deep learning based automatic modulation recognition: Models, datasets, and challenges[J]. *Digital Signal Processing*, 2022, 129: 103650.

- [10] XIAO W, LUO Z, HU Q. A review of research on signal modulation recognition based on deep learning[J]. Electronics, 2022, 11(17): 2764.
- [11] PERENDA E, RAJENDRAN S, POLLIN S. Automatic modulation classification using parallel fusion of convolutional neural networks[C]//Proceedings of the 3rd International Balkan Conference on Communications and Networking. Skopje, North Macedonia: [s.n.], 2019.
- [12] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks[C]//Proceedings of the 17th International Conference on Engineering Applications of Neural Networks. Aberdeen, UK: [s.n.], 2016: 213-226.
- [13] ZHANG Z, 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.
- [14] 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.
- [15] XU J, LUO C, PARR G, et al. A spatiotemporal multi-channel learning framework for automatic modulation recognition[J]. IEEE Wireless Communications Letters, 2020, 9(10): 1629-1632.
- [16] CHEN Y, DONG B, LIU C, et al. Abandon locality: Frame-wise embedding aided transformer for automatic modulation recognition[J]. IEEE Communications Letters, 2022, 27(1): 327-331.
- [17] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.
- [18] O'SHEA T J, WEST N. Radio machine learning dataset generation with GNU radio[C]//Proceedings of the 6th GNU Radio Conference. [S.l.]: [s.n.], 2016.
- [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.
- [20] ZHANG F, LUO C, XU J, 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.
- [21] HONG D, ZHANG Z, XU X. Automatic modulation classification using recurrent neural networks[C]//Proceedings of the 3rd IEEE International Conference on Computer and Communications (ICCC). [S.l.]: IEEE, 2017: 695-700.

#### 作者简介:



战权海(1997-),男,硕士研究生,研究方向:调制识别、通信对抗、无线安全, E-mail: zhanquanhai@163.com。



张雄伟(1965-),通信作者,男,教授,博士生导师,研究方向:智能语音处理与信息内容安全, E-mail: xw-zhang9898@163.com。



宋磊(1976-),男,副教授,硕士生导师,研究方向:无线网络安全、智能网络对抗。



孙蒙(1984-),男,副教授,研究方向:智能语音处理与机器学习。



周振吉(1985-),男,副教授,研究方向:漏洞挖掘、网络对抗。



李涛(1995-),男,硕士研究生,研究方向:网络对抗、系统安全。

(编辑:王静)