

doi:10.3969/j.issn.1003-3106.2025.11.004

引用格式:薄丹,王凯,刘云升,等.基于时频特征融合的自动调制识别方法[J].无线电工程,2025,55(11):2163-2173.

[BO Dan, WANG Kai, LIU Yunsheng, et al. Automatic Modulation Recognition Method Based on Time-Frequency Feature Fusion[J]. Radio Engineering, 2025, 55(11): 2163-2173.]

## 基于时频特征融合的自动调制识别方法

薄丹<sup>1,2</sup>,王凯<sup>1,2</sup>,刘云升<sup>3</sup>,王树彬<sup>1,2\*</sup>

- (1. 内蒙古大学电子信息工程学院, 内蒙古呼和浩特 010021;
2. 内蒙古智能通信与传感信号处理重点实验室, 内蒙古呼和浩特 010021;
3. 内蒙古自治区大数据中心, 内蒙古呼和浩特 010021)

**摘要:**针对自动调制识别(Automatic Modulation Recognition, AMR)在实际应用中受限于小样本数据、时频多模态信息融合不充分,进而导致识别准确率较低的问题,对AMR领域现有技术的局限性进行了简要分析,提出了一种融合扩散模型与对比学习机制的跨模态自监督学习框架。该框架通过引入扩散模型,利用其生成能力实现通信信号高质量数据合成与增强,有效缓解小样本数据对模型训练的约束;同时结合跨模态对比学习机制,构建模态间关联学习模块,充分挖掘和利用时频不同模态表示之间的内在关联与互补信息,解决多模态信息融合不充分的痛点,最终基于上述设计构建了“扩散-对比混合网络(Diffusion-Contrastive Hybrid Network, DCHN)”模型。实验结果显示,该模型在RML2016.10a数据集上的识别准确率较其他网络模型有较大提升,具备较好的识别能力。

**关键词:**自动调制识别;特征融合;扩散模型;对比学习

中图分类号:TN911.3

文献标志码:A

开放科学(资源服务)标识码(OSID):



文章编号:1003-3106(2025)11-2163-11

## Automatic Modulation Recognition Method Based on Time-Frequency Feature Fusion

BO Dan<sup>1,2</sup>, WANG Kai<sup>1,2</sup>, LIU Yunsheng<sup>3</sup>, WANG Shubin<sup>1,2\*</sup>

- (1. School of Electronic Information Engineering, Inner Mongolia University, Hohhot 010021, China;
2. Inner Mongolia Key Laboratory of Intelligent Communication and Sensing Signal Processing, Hohhot 010021, China;
3. Inner Mongolia Autonomous Region Big Data Center, Hohhot 010021, China)

**Abstract:** To solve the problem that Automatic Modulation Recognition (AMR) is limited by small-sample data and insufficient fusion of time-frequency multimodal information in practical applications, which in turn leads to low recognition accuracy, the limitations of existing technologies in the AMR field are analyzed and a cross-modal self-supervised learning framework integrating a diffusion model and a contrastive learning mechanism is proposed. By introducing the diffusion model, the framework leverages its generative capability to achieve high-quality data synthesis and augmentation of communication signals, effectively alleviating the constraints of small-sample data on model training. Meanwhile, combined with the cross-modal contrastive learning mechanism, it constructs an inter-modal association learning module to fully explore and utilize the inherent correlations and complementary information between different time-frequency modal representations, thus solving the problem of insufficient multimodal information fusion. Finally, based on the above design, a Diffusion-Contrastive Hybrid Network (DCHN) model is established. Experimental results show that the recognition accuracy of this model on the RML2016.10a dataset is significantly higher than that of other network models, indicating that it possesses excellent recognition capability.

**Keywords:** AMR; feature fusion; diffusion model; contrastive learning

收稿日期:2025-07-23

基金项目:国家自然科学基金(62361048);内蒙古自治区重点研发与成果转化计划项目(2023YFHH0069)

Foundation Item: National Natural Science Foundation of China (62361048); Key Research and Development and Achievement Transformation Plan Project of Inner Mongolia Autonomous Region (2023YFHH0069)

## 0 引言

AMR<sup>[1]</sup>是指在接收端对接收到的无线电信号的调制方式进行盲估计和判决的技术,即在未知发送端采用何种调制方案的情况下,仅依据接收信号的特征来准确判定其调制类型。在非合作通信系统中,AMR是信号检测与后续解调、译码之间的关键桥梁,其性能直接影响整个信息获取链路的有效性,扮演着举足轻重的角色<sup>[2]</sup>。在日益拥挤和动态变化的频谱环境中,认知无线电(Cognitive Radio, CR)系统依赖AMR技术来实现智能化的频谱感知、频谱共享和动态频谱接入。通过准确识别现有信号的调制类型,CR系统能够有效避免或减轻对授权用户的干扰,优化频谱资源的分配策略,从而保障整体通信质量和频谱利用效率。在电子情报及其他领域AMR技术是实现信号监听、识别、干扰和破译的核心能力之一。通过截获并分析敌方通信信号,快速准确地识别其调制方式,可以为后续的信号参数估计、信息解密以及制定有效的对抗策略提供关键情报支持<sup>[3]</sup>。由此可见,无论是在民用通信、频谱管理,还是在国防安全等领域,AMR都发挥着重要作用,具有显著的研究意义和实际应用价值。

AMR主要可划分为2类:传统AMR与基于深度学习的AMR。传统AMR技术中,核心方法包括基于似然理论(Likelihood-Based, LB)<sup>[4]</sup>与基于特征(Feature-Based, FB)<sup>[5]</sup>的2类方案。LB-AMR方法虽在信道特性已知时可达到理论最优性能,但对信道先验知识依赖性极强且计算复杂度呈指数增长,难以适配实际动态场景。FB-AMR方法通过人工设计瞬时幅度<sup>[6]</sup>、相位<sup>[7]</sup>、频率的统计特性<sup>[8]</sup>、高阶累积量(Higher-Order Cumulants, HOCs)<sup>[9]</sup>、循环谱特征(Cyclic Spectrum Features)<sup>[10]</sup>、小波变换(Wavelet Transform, WT)特征<sup>[11]</sup>等特征结合分类器完成识别,虽降低了对先验知识的依赖,但人工特征普适性差,易受信道损伤影响,且特征工程耗时费力,难以适应调制技术的快速演进。

近年来,深度学习(Deep Learning, DL)技术迅猛发展,促使学术界和工业界开始将DL技术引入AMR领域,以克服传统方法的瓶颈。在深度学习模型应用于AMR的研究中,多种经典及前沿的网络架构得到了广泛探索。

卷积神经网络(Convolutional Neural Network, CNN)<sup>[12]</sup>:利用其局部感知和权值共享的特性,有效提取信号在时间或频率域(如t同相正交(In-phase

Quadrature, IQ)序列、频谱图、星座图)的局部模式和结构特征。

密集连接网络(Densely Connected Networks, DenseNet)<sup>[13]</sup>:通过建立网络内所有层之间的密集连接,促进特征的复用和信息流动,有助于缓解梯度消失问题,并能以较少的参数量达到高性能。

残差网络(Residual Networks, ResNet)<sup>[14]</sup>:引入残差学习块,使得网络可以学习恒等映射,有效解决了深度网络训练中的退化问题,能够构建更深的网络结构以学习更复杂的特征。

长短时记忆网络(Long Short Term Memory, LSTM)<sup>[15]</sup>:作为循环神经网络(Recurrent Neural Network, RNN)的一种变体,特别擅长处理序列数据,能够捕捉信号在时间维度上的长距离依赖关系,对于识别具有时序相关性的调制信号具有优势。

门控循环单元网络(Gated Recurrent Unit, GRU)<sup>[16]</sup>:LSTM的一种简化变体,具有更少的门控单元和参数,计算效率更高,同时在许多任务中表现出与LSTM相当的性能。

然而单一深度学习模型存在固有局限:不同模型对特征的捕捉存在偏倚(如CNN擅长空间特征、RNN擅长时序特征),难以全面提取多维度判别信息;在低信噪比(Signal to Noise Ratio, SNR)环境下信号特征被噪声淹没,导致识别精度急剧下降;面对信道变化或新型调制类型时泛化能力薄弱。

为突破单一模型局限,多模型融合成为研究热点。部分研究者构建了CNN与LSTM相结合的特征融合模型<sup>[17]</sup>;融合CNN与GRU网络的创新架构<sup>[18]</sup>,文献[19]在此基础上增加了自适应降噪模块,通过整合不同模型的特征提取优势,增强对信号特征的捕捉能力。文献[20]提出一种基于多注意力残差网络和GRU的AMR模型,通过结合信号的时空特征,提高调制识别精度。此外,文献[21]提出卷积长短时密集网络(Convolutional Long Short Term Deep Neural Network, CLDNN),该网络专门针对无线通信信号调制识别任务设计,能够从时域和频域2个维度深度挖掘输入信号的特征信息。尽管这些基于特征融合模型在AMR领域取得了一定进展,但仍存在以下局限性:计算资源与时间成本高、特征冗余与冲突、过拟合风险加剧、调参难度大、低SNR适应性仍不足。

针对上述低SNR性能瓶颈、数据稀缺性、多模态信息利用效率等问题,提出了一种创新的解决方法,通过引入扩散模型生成高质量信号样本扩充训

练数据,尤其补充低 SNR 与复杂信道条件下的稀缺样本,缓解数据不足问题;设计自适应时频分析模块精细化捕捉信号时变频变特性;采用跨模态对比学习机制融合 IQ 域、幅度相位 (Amplitude Phase, AP) 域信息,强制模型学习模态不变的、更具判别力的特征表示。该方法旨在实现高精度、高鲁棒性与强泛化能力的 AMR 技术突破,其理论贡献在于:验证了扩散模型在通信信号合成中的有效性,为数据稀缺问题提供新范式;深化了多域特征融合的理论基础,揭示模态不变性特征对提升鲁棒性的核心作用。

## 1 信号模型

本文所涉及的算法模型主要聚焦于单输入单输出通信系统。完整的通信系统示意如图 1 所示。

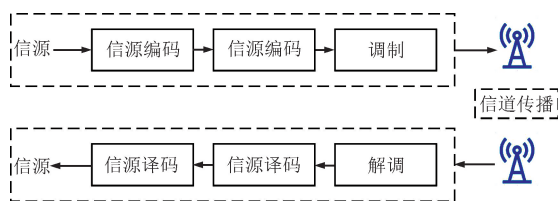


图 1 完整的通信系统示意

Fig. 1 Schematic diagram of a complete communication system

在此系统中,接收机接收到的信号模型可表示为:

$$r(t) = \left[ \int_0^{\tau_0} s(\Delta_{\text{co}}(t - \tau)) \times h(t - \tau) d\tau \right] \cdot e^{j\Delta_{\text{fo}}(t)} + n(t), \quad (1)$$

式中: $s(t)$ 表示信源所发送的调制信号, $\tau_0$ 表示最大延迟拓展, $h(t)$ 表示信号通过无线信道时的信道函数, $\Delta_{\text{fo}}(t)$ 表示载波频率偏移, $n(t)$ 表示加性高斯白噪声, $\Delta_{\text{co}}(t)$ 表示采样频率偏移, $r(t)$ 表示接收机接收到的调制信号,在数字信号处理中,连续时间信号 $r(t)$ 需要经过采样得到离散时间序列 $r[n]$ 。通常采用正交两路 IQ 格式进行采样。此信号由同相分量 $r_1$ 和正交分量 $r_Q$ 构成,具体表示为:

$$r[n] = r_1[n] + j \cdot r_Q[n]. \quad (2)$$

尽管原始的 IQ 数据序列 $r[n]$ 是信号的完整表示,但直接将其作为深度学习模型的输入,有时难以让模型高效地学习到区分不同调制类型的关键特征,尤其是在低 SNR 或存在复杂信道损伤时。为了增强特征的代表能力和鲁棒性,有必要考虑信号的其他等价或派生表示形式,利用信号的多模态特性。

主要考虑以下几种表示。

(1) 瞬时幅度 (Instantaneous Amplitude) 与瞬时相位 (Instantaneous Phase)

信号的瞬时幅度和相位能够反映信号包络和相位的变化规律,这对于区分幅度调制 (Amplitude Modulation, AM)、脉冲幅度调制 (Pulse Amplitude Modulation, PAM)、相移键控 (Phase Shift Keying, PSK) 和频移键控 (Frequency Shift Keying, FSK) 等类型的信号至关重要。信号的瞬时幅度和瞬时相位数据的计算过程分别由式 (3) 和式 (4) 给出:

$$r_P[n] = \arctan \left( \frac{r_I[n]}{r_Q[n]} \right), \quad (3)$$

$$r_A[n] = \sqrt{(r_I[n])^2 + (r_Q[n])^2}, \quad (4)$$

式中: $r_P[n]$ 表示信号的瞬时相位, $r_A[n]$ 表示信号的瞬时幅度,为表述简便,将其简称为信号的 AP 格式数据。AP 数据提供了信号在极坐标系下的表示,有助于模型关注信号的包络稳定性和相位跳变特性。

(2) 离散傅里叶变换 (Discrete Fourier Transform, DFT)

信号的 DFT 可以将时域信号 $r[n]$ 变换到频域,得到信号的频谱表示 $X[k]$ 。频谱特征对于识别 FSK、正交频分复用 (Orthogonal Frequency Division Multiplexing, OFDM) 以及区分不同带宽的信号非常重要。其计算过程由式 (5) 和式 (6) 表示。

$$X_I[k] = DFT[x_I[n]] = \sum_{n=0}^{N-1} x_I[n] \cdot w_N^{nk}, \quad k = 0, 1, 2, \dots, N-1, \quad (5)$$

$$X_Q[k] = DFT[x_Q[n]] = \sum_{n=0}^{N-1} x_Q[n] \cdot w_N^{nk}, \quad k = 0, 1, 2, \dots, N-1, \quad (6)$$

式中: $X_I[k]$ 表示 I 路信号的 DFT 数据序列, $X_Q[k]$ 表示 Q 路信号的 DFT 数据序列, $w_N^{nk}$ 表示 DFT 系数,其计算公式如式 (7) 所示。

$$w_N^{nk} = e^{-j\frac{2\pi}{N}nk}. \quad (7)$$

通常更关心信号的幅度谱 $|X[k]| = \sqrt{|X_I[k]|^2 + |X_Q[k]|^2}$ 或功率谱密度 (Power Spectral Density, PSD)  $|X[k]|^2/N$ 。频域能够揭示信号的能量在不同频率上的分布情况。

通过利用 IQ、AP、DFT 等多种信号表示形式,可以为后续的特征提取和融合提供更丰富的信息源,有助于模型从不同角度捕捉信号的判别性特征。

## 2 网络结构

### 2.1 基于扩散模型的数据增强

在 AMR 任务中,尤其是在低 SNR 或小样本 (Small Sample Size) 场景下,获取大量高质量、多样性的标注信号数据往往是困难且昂贵的。原始数据集可能存在样本不均衡、覆盖信道条件有限等问题,导致训练出的模型泛化能力不足。

传统的数据增强技术(如旋转、加噪、插值、仿射变换等),虽然能在一定程度上扩充数据量,但它们往往难以真实模拟无线信道中复杂的物理效应,例如频率选择性衰落、时间选择性衰落、相位噪声、非线性失真等。生成的样本可能与真实信号的统计特性存在偏差,增强效果有限。

为了克服这些局限性,创新性地引入了扩散模型(Diffusion Models)<sup>[22-23]</sup>进行通信信号的数据增强。扩散模型是一类强大的深度生成模型,其灵感来源于非平衡热力学中的扩散过程。它们通过 2 个核心过程来学习数据的分布并生成新样本:前向扩散过程(Forward Diffusion Process)和反向扩散过程(Reverse Diffusion Process)。

#### 2.1.1 前向扩散过程

这个过程逐步地向原始数据样本  $x_0$  (例如一段 IQ 信号序列) 中注入预定强度的高斯噪声,经过  $T$  个时间步,最终将其转化为一个接近纯高斯噪声的隐变量  $x_T$ 。该过程是一个固定的、无需学习的马尔可夫链:

$$q(x_{1:T} | x_0) = \prod_{t=1}^T q(x_t | x_{t-1}), \quad (8)$$

式中:每一步的转移核  $q(x_t | x_{t-1})$  定义为向  $x_{t-1}$  添加少量高斯噪声。

$$q(x_t | x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t \mathbf{I}), \quad (9)$$

式中: $\mathbf{I}$  为单位矩阵,  $\beta_t \in (0, 1)$  为预先设定的噪声方差调度(Noise Schedule),它控制了在第  $t$  步加入噪声的强度。通常  $\beta_t$  随  $t$  增大而增大或保持不变。一个重要的性质是,任意时刻  $t$  的含噪样本  $x_t$  可以直接从原始样本  $x_0$  通过一个闭式解得到:

$$q(x_t | x_0) = \mathcal{N}(x_t; \sqrt{\alpha_t}x_0, (1-\alpha_t)\mathbf{I}), \quad (10)$$

式中: $\alpha_t = 1 - \beta_t$  且  $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$ 。式(10)可以高效地在训练过程中采样任意噪声水平的  $x_t$ 。当前向过程步数  $T$  足够大且噪声调度  $\beta_t$  合理设置时,  $x_T$  的分布将非常接近标准高斯分布  $\mathcal{N}(0, \mathbf{I})$ 。

#### 2.1.2 反向扩散过程

生成新样本的核心在于学习反向过程  $p_\theta(x_{0:T}) = p(x_T) \prod_{t=1}^T p_\theta(x_{t-1} | x_t)$ , 即从纯噪声  $x_T \sim \mathcal{N}(0, \mathbf{I})$  开始,逐步去除噪声,最终恢复出符合原始数据分布  $q(x_0)$  的样本  $x_0$ 。这个反向过程也是一个马尔可夫链,其每一步转移核  $p_\theta(x_{t-1} | x_t)$  由一个深度神经网络(通常是 U-Net 架构)来参数化,用  $\theta$  表示网络参数。该网络的目标是预测  $x_{t-1}$  的分布(或等价地,预测加入的噪声)。理论上,如果已知  $x_0$ , 反向转移  $q(x_{t-1} | x_t, x_0)$  也是高斯分布:

$$q(x_{t-1} | x_t, x_0) = \mathcal{N}(x_{t-1}; \tilde{\mu}_t(x_t, x_0), \tilde{\beta}_t \mathbf{I}), \quad (11)$$

式中: $\tilde{\mu}_t(x_t, x_0)$  和  $\tilde{\beta}_t$  表示通过  $x_t, x_0, \beta_t$  计算得到的均值和方差。扩散模型的关键思想是训练神经网络  $p_\theta(x_{t-1} | x_t)$  来逼近这个真实的后验分布。常用的做法是让网络  $\epsilon_\theta(x_t, t)$  预测在第  $t$  步添加到  $x_0$  上的噪声  $\epsilon$ , 然后利用这个预测来估计  $x_{t-1}$  的均值。

#### 2.1.3 训练目标

模型的训练通常通过优化变分下界(Variational Lower Bound, VLB) 或一个简化的目标函数来进行。一个广泛使用的简化目标是最小化预测噪声  $\epsilon_\theta(x_t, t)$  与真实添加的噪声  $\epsilon$  之间的均方误差(Mean Squared Error, MSE):

$$L_{\text{simple}}(\theta) = \mathbb{E}t \sim u(1, T), x_0 \sim q(x_0), \\ \epsilon \sim \mathcal{N}(0, \mathbf{I}) \left| \epsilon - \epsilon_\theta(\sqrt{\alpha_t}x_0 + \sqrt{1-\alpha_t}\epsilon, t) \right|^2, \quad (12)$$

式中: $\mathbb{E}$ 表示期望,通过最小化这个损失函数,神经网络  $\epsilon_\theta$  学会了如何在给定含噪样本  $x_t$  和时间步  $t$  的情况下,有效地估计出对应的噪声成分。

#### 2.1.4 数据增强应用

在 AMR 任务中,将扩散模型应用于生成特定调制类型、特定 SNR 和模拟特定信道条件(如衰落、频偏等)下的信号样本,可以通过条件扩散模型(Conditional Diffusion Models)来实现。在训练时,除了输入含噪信号  $x_t$  和时间步  $t$ ,还需要将条件信息  $c$  (如调制类型标签、目标 SNR 值、信道参数等)输入到噪声预测网络  $\epsilon_\theta(x_t, t, c)$  中。训练好的条件扩散模型可以根据给定的条件  $c$  从噪声  $x_T$  开始采样,生成符合该条件的信号样本  $x_0$ 。扩散模型的信号数据增强流程如图 2 所示。



③ 特征提取:经过注意力增强的时频特征图被送入后续的特征提取网络 ResNet 中,进行更深层次的特征学习和分类。

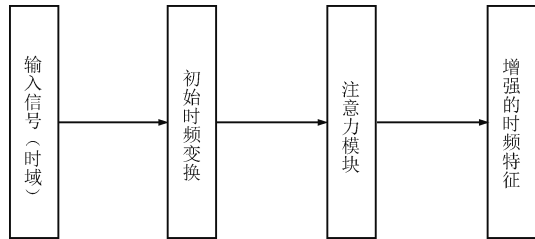


图3 自适应时频分析网络模块示意

Fig.3 Schematic diagram of adaptive time-frequency analysis network module

通过这种自适应时频分析网络,模型能够根据数据自动学习最优的时频表示方式或关注点,减少对固定变换参数的依赖,更有效地捕捉复杂通信信号中蕴含的、对调制识别任务至关重要的时变频变特征,从而提升识别精度和鲁棒性。

### 2.3 跨模态对比学习

无线通信信号天然具有多模态(Multi-modal)特性,同一信号可以从不同角度进行描述和表示,例如IQ序列、AP序列、频域表示(DFT谱)、星座图等。这些不同的模态各自蕴含了信号的部分信息,且互为补充。例如,IQ序列保留了最原始的相位和幅度信息,AP序列强调了包络和相位的动态变化,频谱揭示了频率成分。然而,许多现有的AMR方法仅利用单一模态的输入,未能充分挖掘和利用这些不同模态表示之间的内在关联和互补信息。跨模态对比学习(Cross-modal Contrastive Learning)提供了一种强大的自监督学习范式,旨在学习不同模态数据之间的共享语义信息或不变性特征,从而提升模型的表示能力和泛化性能。

在AMR的背景下,跨模态对比学习的目标是:对于同一个底层调制信号,使其不同模态的表示在特征空间中相互靠近;而对于不同底层调制信号,使其任意模态的表示在特征空间中相互远离。通过这种方式,模型被强制学习那些在不同表示模态下保持一致的、能够区分不同信号来源的本质特征。

#### 2.3.1 数据对构建

① 正样本对(Positive Pairs):对于一个原始信号样本 $x$ ,生成其不同模态的表示 $x_{IQ}$ (IQ序列)和 $x_{AP}$ (AP序列)。 $(x_{IQ}, x_{AP})$ 构成一个正样本对,因为它们源自同一个信号实例。

② 负样本对(Negative Pairs):对于一个信号样本 $x_i$ 的某种模态表示(如 $x_{i,IQ}$ ),来自其他不同信号

样本 $x_j(j \neq i)$ 的任意模态表示(如 $x_{j,IQ}$ 或 $x_{j,AP}$ )都构成负样本对。

#### 2.3.2 特征编码器(Feature Encoders)

特征编码器的核心在于,为对比学习中涉及的每个模态分别设计专属的特征编码器网络。对于IQ序列,可采用CNN构建编码器 $f_{IQ}(\cdot)$ 进行处理;而针对AP序列,则可借助Transformer设计编码器 $f_{AP}(\cdot)$ 来完成特征提取。这些编码器将各自模态的输入映射到一个共享的或模态特定的嵌入空间(Embedding Space),得到特征向量 $\mathbf{v}_{IQ} = f_{IQ}(x_{IQ})$ 和 $\mathbf{v}_{AP} = f_{AP}(x_{AP})$ 。通常在特征向量之后还会接一个投影头(Projection Head) $g(\cdot)$ ,将特征映射到用于计算对比损失的空间,例如 $\mathbf{z}_{IQ} = g_{IQ}(\mathbf{v}_{IQ})$ 和 $\mathbf{z}_{AP} = g_{AP}(\mathbf{v}_{AP})$ 。

#### 2.3.3 对比损失函数(Contrastive Loss Function)

核心是设计一个损失函数,使得正样本对的特征向量相似度最大化,负样本对的相似度最小化。InfoNCE(Noise Contrastive Estimation)损失<sup>[24]</sup>是目前最常用的一种对比损失函数。假设在一个批次(Batch)中有 $N$ 个原始信号样本,为每个样本生成了2种模态的表示(例如IQ和AP)。对于第 $i$ 个样本的表示对 $(\mathbf{z}_{i,IQ}, \mathbf{z}_{i,AP})$ ,InfoNCE损失(以IQ到AP方向为例)计算如下:

$$L_{IQ \rightarrow AP}^{(i)} = - \ln \frac{\exp(\sin(\mathbf{z}_{i,IQ}, \mathbf{z}_{i,AP})/\tau)}{\sum_{j=1}^N \exp(\sin(\mathbf{z}_{i,IQ}, \mathbf{z}_{j,AP})/\tau)}, \quad (15)$$

式中: $\sin(\mathbf{u}, \mathbf{v})$ 是衡量2个向量 $\mathbf{u}, \mathbf{v}$ 相似度的函数,常用余弦相似度 $\sin(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u}^T \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$ ;  $\tau > 0$ 是温度系数(Temperature Parameter),用于调节相似度得分的分布。较小的 $\tau$ 会放大相似度差异,使得模型更难区分负样本,有助于学习更精细的特征;较大的 $\tau$ 则使分布更平滑。

分子是正样本对 $(\mathbf{z}_{i,IQ}, \mathbf{z}_{i,AP})$ 的相似度得分的指数。分母是 $\mathbf{z}_{i,IQ}$ 与批次内所有 $N$ 个AP表示(包括一个正样本 $\mathbf{z}_{j,AP}$ 和 $N-1$ 个负样本 $\mathbf{z}_{j,AP}, j \neq i$ )的相似度得分的指数之和。这个损失项的目标是让正样本对的相似度(分子)在所有样本对的相似度总和(分母)中占据主导地位。

对称地,还需要计算从AP到IQ方向的损失:

$$L_{AP \rightarrow IQ}^{(i)} = - \ln \frac{\exp(\sin(\mathbf{z}_{i,AP}, \mathbf{z}_{i,IQ})/\tau)}{\sum_{i=1}^N \exp(\sin(\mathbf{z}_{i,AP}, \mathbf{z}_{j,IQ})/\tau)}. \quad (16)$$

最终的批次对比损失 $L_{contrastive}$ 通常是这2个方向损失的平均值:

$$L_{\text{contrastive}} = \frac{1}{2N} \sum_{i=1}^N (L_{\text{IQ} \rightarrow \text{AP}}^{(i)} + L_{\text{AP} \rightarrow \text{IQ}}^{(i)}) \quad (17)$$

通过最小化  $L_{\text{contrastive}}$ , 编码器  $f_{\text{IQ}}, f_{\text{AP}}$  和投影头  $g_{\text{IQ}}, g_{\text{AP}}$  被驱动去学习能够将来自同一信号的不同模态表示映射到嵌入空间中邻近位置的特征, 同时将来自不同信号的特征推开。

在 DCHN 模型中, 跨模态对比学习框架示意如图 4 所示, 被用作一种自监督预训练策略或辅助损

失。利用 IQ 数据、AP 数据以及可能由自适应时频分析网络生成的时频特征图作为不同的模态输入。通过最小化它们之间的对比损失, 模型可以在没有调制标签的情况下学习到信号的内在结构和模态不变特征。这些学习到的特征表示随后可以用于初始化下游的调制识别分类器, 或者对比损失可以作为总训练损失的一部分, 与监督分类损失 (如交叉熵损失) 联合优化。

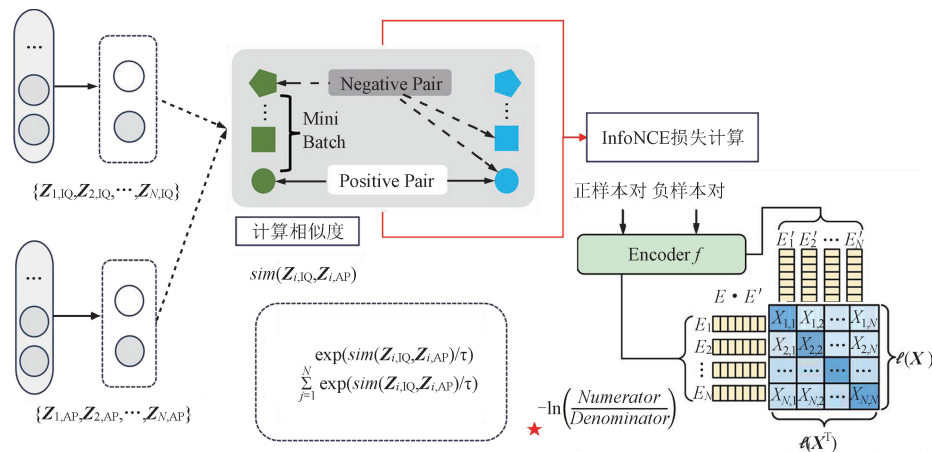


图 4 跨模态对比学习框架示意

Fig. 4 Schematic diagram of cross-modal contrastive learning framework

跨模态对比学习的引入, 旨在充分利用信号的多维度信息, 克服单一模态输入的限制性, 学习到更具鲁棒、更具判别力的特征表示, 尤其有望提升模型在低 SNR 和复杂信道条件下的性能, 并增强其对未见数据的泛化能力。

## 2.4 网络结构

本文设计的网络模型结构示意如图 5 所示。该框架是一个端到端的深度学习模型, 系统地整合了数据增强、多模态特征提取与融合以及自监督与监督联合学习。

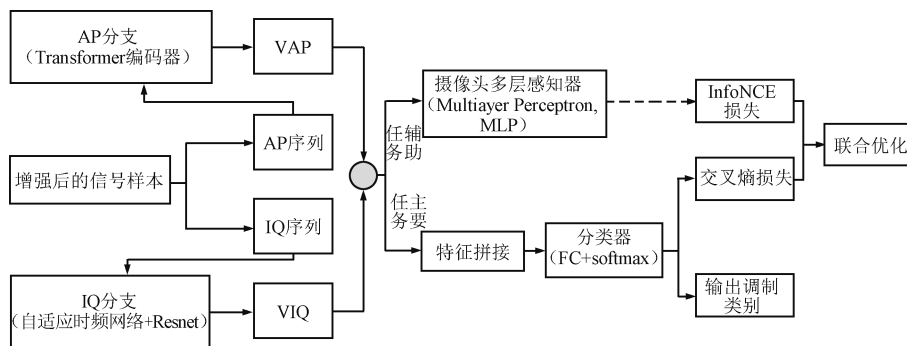


图 5 模型结构示意图

Fig. 5 Schematic diagram of model structure

## 3 实验

本节旨在通过一系列精心设计的实验来全面评估所提出的 DCHN 模型的有效性和性能优势。实

验内容包括数据集介绍、参数设置、消融研究 (Ablation Study) 以验证各核心模块的贡献, 以及与当前主流的基于深度学习的 AMR 模型进行的对比实验。

### 3.1 实验数据与参数设置

本研究采用公开的、广泛使用的基准数据集 RML2016. 10a<sup>[25]</sup> 开展实验,其详细参数配置如表 1 所示。选择该数据集是因为其包含了多种常见的数字和模拟调制类型,覆盖了较宽的 SNR 范围,并且是当前 AMR 领域进行模型性能比较的标准平台之一。

表 1 RML2016. 10a 数据集参数  
Tab. 1 RML2016. 10a dataset parameters

类别	内容
数据样式	L/Q 两路
数据维度	2×128
调制样式	WBFM、AM-DSB、AM-SSB、 BPSK、CPFSK、GFSK、4-PAM、 16-QAM、64-QAM、QPSK、8PSK
SNR	[-20,18]dB,间隔 2 dB,共 20 个
单 SNR 样本数	1 000
总样本数	220 000

在实验过程中,将数据集按照 6 : 2 : 2 的比例随机划分为训练集、验证集和测试集。

训练时,为了保证实验的可复现性和公平比较,所有模型的训练均采用统一或相似的参数设置:批量大小设置为 400;选用 Adam 优化器,Adam 是一种广泛使用的自适应学习率优化算法,能够为不同参数计算独立的自适应学习率,通常收敛速度较快且性能稳健;以交叉熵损失函数作为模型训练的优化目标;为提升模型训练效率与稳定性,采用动态学习率调整策略:若验证损失在连续 10 个训练周期内未出现下降,则将学习率减半,若验证损失在 50 个周期内持续无改善,将自动终止训练。

所有实验均在统一的硬件和软件环境下进行,以确保结果的可比性。具体的实验工作环境配置如表 2 所示。

表 2 实验工作环境  
Tab. 2 Experimental working environment

实验环境	硬件信息、软件版本
系统环境	Windows 10
Python	3.9
TensorFlow	2.8
CUDA	11.6
CUDNN	8.4
GPU	NVIDIA GeForce RTX 3050 Laptop GPU

### 3.2 消融实验

为系统验证模型各模块的有效性,将数据增强、自适应时频分析、跨模态对比学习模块分别去除,建立了 4 个模型,开展了消融实验,具体结果如表 3 所示。

① 模型 1:将数据增强模块去除,在公共数据集上进行训练、验证和测试。

② 模型 2:将自适应时频分析模块去除,在公共数据集上进行训练、验证和测试。

③ 模型 3:将跨模态对比学习模块去除,在公共数据集上进行训练、验证和测试。

④ 模型 4:保留所有模块。

数据表明,完整模型的识别准确率表现最优,相较于其他消融版本模型,平均准确率分别提升了 1.22、3.12、1.34 个百分点。这一结果充分证实了数据增强、自适应时频分析、跨模态对比学习模块在增强模型特征提取能力方面的显著作用。

表 3 消融实验准确率

Tab. 3 Accuracy of ablation experiment

单位:%

模型	平均准确率	最高准确率
模型 1	63.66	92.95
模型 2	61.76	91.23
模型 3	63.54	92.77
模型 4	64.88	93.75

消融实验结果如图 6 所示。

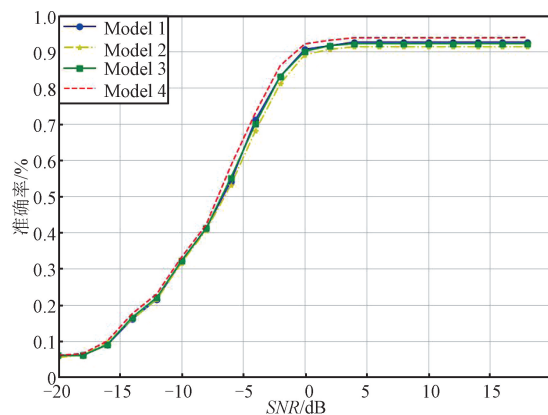


图 6 消融实验

Fig. 6 Ablation experiment

由图 6 可以看出,数据增强通过调节条件参数,能快速模拟不同 SNR、不同衰落环境下的信号,显著缓解小样本问题;而自适应时频分析模块则凭借对信号空间特征的高效提取,推动模型实现更高的

分类精度;跨模态对比学习模块利用多模态特征相互补充,大幅提高了低 SNR 场景下的识别准确率。消融实验结果从多维度验证了模型各组成模块的必要性与有效性,为其在 AMR 任务中的优势提供了有力支撑。

### 3.3 对比实验

为充分验证模型的性能优势,开展了与主流深度学习调制识别模型的对比实验。实验选取了 CGDNet<sup>[26]</sup>、CLDNN、ResNet 以及 DenseNet 等当前表现优异的网络作为对照基准。所有模型均采用相同的数据集划分策略,并统一配置 Adam 优化器与交叉熵损失函数,确保实验条件的一致性与结果的可比性。

从图 7 所示的实验结果来看,在全 SNR 范围内,模型的识别准确率均显著优于各对比模型。特别在-2~18 dB 的关键 SNR 区间,模型的识别准确率较其他模型提升幅度达 3%~10%,平均识别准确率更是达到 92.59%,展现出强大的稳健性与高效性。进一步结合表 4 数据可知,相较于 CLDNN 和 ResNet,本文模型在降低参数数量的同时大幅提升了识别精度,实现了性能与效率的双重优化;即便与参数量最少的 CGDNet 相比,本文模型在调制分类准确率上依然具有显著优势,充分证明了其在调制识别任务中的优越性。

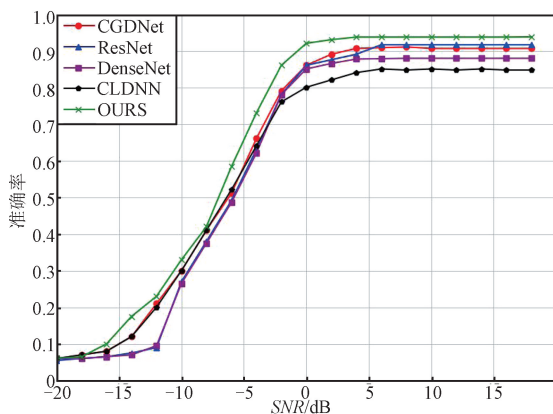


图 7 对比实验

Fig. 7 Comparative experiment

表 4 不同模型的调制识别

Tab. 4 Modulation recognition of different models

单位:%

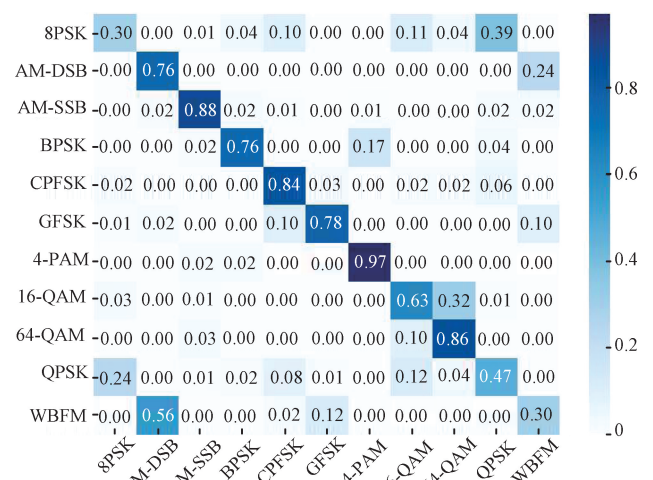
模型	平均准确率	最高准确率
OURS	64.88	93.75
CGDNet	62.03	90.58

续表

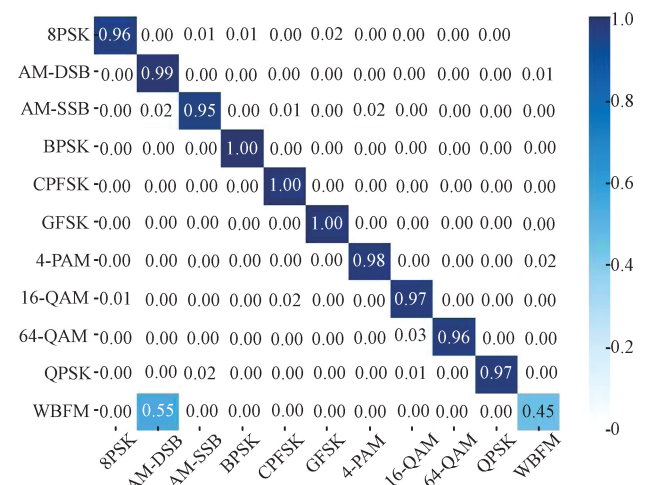
单位:%

模型	平均准确率	最高准确率
ResNet	60.78	91.60
CLDNN	58.30	84.70
DenseNet	59.94	87.98

图 8(a) 呈现了本文模型在-6 dB SNR 时的混淆矩阵,其中每一行对应真实的调制类型,每一列对应预测的调制类型。图 8(b) 展示了本文模型在 12 dB SNR 时的混淆矩阵,此时,除 WBFM 外,其余 9 种信号的识别准确率皆在 95% 以上。



(a) SNR 为-6 dB 时不同模型的混淆矩阵



(b) SNR 为 12 dB 时不同模型的混淆矩阵

图 8 SNR 为-6 dB 和 12 dB 时不同模型的混淆矩阵

Fig. 8 Confusion matrix of different models under SNR of -6 dB and 12 dB

## 4 结束语

针对 AMR 中低 SNR、数据稀缺、多模态信息利

用不足及新调制类型适应困难等挑战,提出了DCHN框架。通过将扩散模型引入数据合成环节,结合物理层信道先验生成高保真的低SNR信号,有效缓解了数据稀缺问题;设计自适应时频分析网络,以可微分变换层和多尺度特征融合策略,突破传统时频分析方法的局限性;构建跨模态对比学习机制,通过多模态特征协同与自监督训练,显著提升模型对复杂信道环境的鲁棒性。实验结果表明,DCHN在全SNR范围内的识别准确率优于现有主流模型,尤其在低SNR场景下优势显著,同时实现了参数效率与识别精度的平衡。本文工作为AMR技术在动态频谱接入、认知无线电等领域的实际应用提供了新的解决方案与技术路径,未来将进一步探索其在更多复杂通信场景中的泛化能力与优化空间。



#### 参考文献

- [1] 何继爱,张文启. 通信信号调制识别技术及其发展[J]. 高技术通讯,2016,26(2):157-165.
- [2] TU Y, WANG M Y, WANG S. A Data-driven Deep Neural Network Pruning Approach Towards Efficient Digital Signal Modulation Recognition [C] // 2019 IEEE Globecom Workshops (GCWkshps). Waikoloa; IEEE, 2019: 1-5.
- [3] JDID B, HASSAN K, DAYOUB I, et al. Machine Learning Based Automatic Modulation Recognition for Wireless Communications: A Comprehensive Survey [J]. IEEE Access, 2021, 9: 57851-57873.
- [4] XU J L, SU W, ZHOU M C. Likelihood-ratio Approaches to Automatic Modulation Classification [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2011, 41(4): 455-469.
- [5] HAZZA A, SHOAI B, ALSHEBEILI S A, et al. An Overview of Feature-based Methods for Digital Modulation Classification [C] // 2013 1st International Conference on Communications, Signal Processing, and Their Applications (ICCSPA). Sharjah; IEEE, 2013: 1-6.
- [6] 胡延平,李广森,李纲,等. 利用参数统计方法自动识别数字调制信号[J]. 通信学报,2002,23(2):58-65.
- [7] AZZOUZ E E, NANDI A K. Automatic Identification of Digital Modulations [J]. Signal Processing, 1995, 47(1): 55-69.
- [8] NANDI A K, AZZOUZ E E. Algorithms for Automatic Modulation Recognition of Communication Signals [J]. IEEE Transactions on Communications, 1998, 46(4): 431-436.
- [9] WU H C, SAQUIB M, YUN Z F. Novel Automatic Modulation Classification Using Cumulant Features for Communications via Multipath Channels [J]. IEEE Transactions on Wireless Communications, 2008, 7(8): 3098-3105.
- [10] YAN X, FENG G Y, WU H C, et al. Innovative Robust Modulation Classification Using Graph-based Cyclic-spectrum Analysis [J]. IEEE Communications Letters, 2017, 21(1): 16-19.
- [11] PARK C S, CHOI J H, NAH S P, et al. Automatic Modulation Recognition of Digital Signals Using Wavelet Features and SVM [C] // 2008 10th International Conference on Advanced Communication Technology. Gangwon; IEEE, 2008: 387-390.
- [12] 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.
- [13] 王安义,李立. 基于密集连接机制的无线信号调制识别[J]. 科学技术与工程,2023,23(3):1104-1110.
- [14] LIU X Y, YANG D Y, GAMAL A E. Deep Neural Network Architectures for Modulation Classification [C] // 2017 51st Asilomar Conference on Signals, Systems, and Computers. Pacific Grove; IEEE, 2017: 915-919.
- [15] 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.
- [16] SUN S Q, WANG Y Y. A Novel Deep Learning Automatic Modulation Classifier with Fusion of Multichannel Information Using GRU [J]. EURASIP Journal on Wireless Communications and Networking, 2023, 2023(1): 66.
- [17] WANG N, LIU Y X, MA L, et al. Multidimensional CNN-LSTM Network for Automatic Modulation Classification [J]. Electronics, 2021, 10(14): 1649.
- [18] LIU F G, ZHANG Z W, ZHOU R L. Automatic Modulation Recognition Based on CNN and GRU [J]. Tsinghua Science and Technology, 2021, 27(2): 422-431.
- [19] 陈昊,郭文普,康凯,等. 基于卷积自适应降噪网络的自动调制识别方法[J]. 无线电工程,2025,55(2):291-297.
- [20] 李鸣皓,解志斌,颜培玉,等. 基于多注意力残差网络和GRU的自动调制识别算法[J]. 无线电工程,2025, 55(1): 36-44.
- [21] WESTN E, O' SHEA T. Deep Architectures for Modulation Recognition [C] // 2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN). Baltimore; IEEE, 2017: 1-6.
- [22] HO J, JAIN A, ABBEEL P. Denoising Diffusion Probabilistic Models [C] // 34th Conference on Neural Information Processing Systems. Vancouver; Curran Associates Inc., 2020: 6840-6851.

- [23] ROMBACH R, BLATTMANN A, LORENZ D, et al. High-resolution Image Synthesis with Latent Diffusion Models [C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans: IEEE, 2022: 10674–10685.
- [24] VAN DEN OORD A, LI Y Z, VINYALS O. Representation Learning with Contrastive Predictive Coding [EB/OL]. (2018-07-10) [2025-05-22]. <https://arxiv.org/abs/1807.03748>.
- [25] O'SHEA T, CORGAN J, CLANCY T C. Convolutional Radio Modulation Recognition Networks [C] // 17th International Conference on Engineering Applications of Neural Networks. Aberdeen: Springer, 2016: 213–226.
- [26] 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.

---

#### 作者简介

薄丹 女, (1996—), 硕士研究生。主要研究方向: 自动调制识别。

王凯 男, (1998—), 硕士研究生。主要研究方向: 频谱感知。

刘云升 男, (1981—)。主要研究方向: 信息化与大数据。

(\*通信作者) 王树彬 男, (1971—), 博士, 教授。主要研究方向: 智能无线通信及其应用、物联网、网络人工智能。