

基于镜像数据增强与多尺度特征卷积融合的 Transformer调制识别技术

李国军^{1,2}, 陈世奥^{1,2*}, 王 杰¹, 张哲铭¹, 郑建忠²

(1. 重庆邮电大学通信与信息工程学院, 重庆 400065; 2. 重庆邮电大学超视距可信信息传输研究所, 重庆 400065)

摘要: 调制识别是通信对抗的关键技术, 现有的基于深度学习的调制识别研究大多在仿真数据集或开源数据集上进行, 导致训练所得的模型在实际应用时面临着不能适应具体场景的巨大挑战。本文首先提出了一种针对调制信号的镜像数据增强方法, 以矢量信号源发射且用接收机接收得到的信号作为原始数据, 通过滤波、不同的速率抽取、相移、频移、加噪操作, 实现信号增强。这样增强生成的数据集能够适应实际场景中不同的符号速率、多普勒频移、接收机载波偏移、信噪比 (Signal-to-Noise Ratios, SNRs)、接收机特征等各种要素的影响, 与实际信号相近。接着设计一个信号模态变换模块, 对采集的同向正交信号样本进行模态变换, 为后续的多模态处理提供数据基础。随后设计了多模态多尺度卷积融合加压缩激励去噪机制的Transformer调制识别模型。之后, 用RadioML2018.10a的数据对提出设计模型进行训练, 信噪比 12 dB 以上时测试集准确率达到 98.3%, 但此时训练的模型用于实际场景测试时的结果仅 10.4%。最后使用增强的数据集对提出的模型进行训练, 信噪比 16 dB 以上平均准确率为 90.1%; 使用训练所得的模型进行在线落地测试, 信噪比 12 dB 时识别率为 91.9%。

关键词: 深度学习; 调制识别; 数据增强; 特征融合; 模态变换; 多头自注意力

基金项目: 国家自然科学基金 (No.U22A2006); 重庆市自然科学基金 (No.CSTB2023NSCQ-LZX0082); 科技委基础加强项目 (No.2023-ZQ-080); 科技委基础加强重点项目 (No.2022-173-SD-1)

中图分类号: TN911.3; TP18

文献标识码: A

文章编号: 0372-2112(2026)01-0329-11

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20250566

Transformer-Based Modulation Recognition with Mirroring Data Augmentation and Multi-Scale Convolutional Feature Fusion

LI Guojun^{1,2}, CHEN Shi'ao^{1,2*}, WANG Jie¹, ZHANG Zheming¹, ZHENG Jianzhong²

(1. School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

2. Laboratory of Beyond LOS Reliable Information Transmission, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Modulation recognition is a key technology in communication countermeasures. Most of the existing modulation recognition studies based on deep learning are conducted on simulated datasets or open-source datasets. As a result, the models obtained through training face huge challenges of being unable to adapt to specific scenarios in practical applications. First, this paper proposes a mirror data augmentation method for modulated signals. The signals transmitted by a vector signal source and received by a receiver are used as the original data. Signal augmentation is achieved through operations such as filtering, different rate sampling, phase shift, frequency shift, and noise addition. The augmented dataset generated in this way can adapt to the influence of various factors in real-world scenarios, such as different symbol rates, Doppler frequency shifts, receiver carrier offsets, signal-to-noise ratios (SNRs), and receiver characteristics, and is similar to real signals. Next, a signal modality transformation module is designed to perform modality transformation on IQ sampling data samples, providing a data basis for subsequent multi-modality processing. Then, a Transformer-based modulation recognition model with multi-modality multi-scale convolution fusion and SE denoising mechanism is designed. After that, the proposed model is trained using the RadioML2018.10a dataset. When the SNR is above 12 dB, the accuracy of the test set reaches 98.3%. However, when the trained model is used for testing in real-world scenarios, the result is only 10.4%. Finally, the proposed model is trained using the augmented dataset. When the SNR is above 16 dB, the average accuracy is 90.1%. The trained model is used for online practical testing, and the recognition rate reaches 91.9% when the SNR is 12 dB.

Keywords: deep learning; modulation recognition; data augment; feature fusion; modal transformation; multi-head self-attention

Foundation Item(s): National Natural Science Foundation of China (No.U22A2006); Natural Science Foundation of Chongqing, China (No.CSTB2023NSCQ-LZX0082); Basic Strengthening Project of the Science and Technology Committee (No.2023-ZQ-080); Key Basic Research Project of the Science and Technology Committee (No.2022-173-SD-1)

0 引言

自动调制识别 (Automatic Modulation Recognition, AMR) 常用于非协作通信与电子侦察环境中对接收信号进行调制类型的识别。作为信号检测和解调之间的关键步骤^[1], 其在军事领域和民用领域中都扮演着重要角色。随着近年来无线通信技术的迅速发展, 信号调制方式越发复杂和多样化, 无线电环境愈加恶劣, 这是对信号调制识别的巨大挑战。

传统的调制识别方式基本上可划分为基于似然函数的调制识别算法以及基于特征的调制识别算法两类。基于似然函数的方法^[2]要求准确建立未知信号模型, 将调制识别问题视为多假设问题, 并通过比较似然比和预设阈值来作出最终判断。然而, 这种方法对先验信息依赖性大, 对似然函数模型的精确性要求较高、计算复杂^[3]。相比之下, 基于特征的方法^[4]则是从信号中提取特征, 再利用分类器进行调制识别。此方法相对计算复杂度低、实用性强^[5]。然而, 在提取的特征与分类器不匹配时, 识别结果会受到较大影响^[6]。

基于深度学习的调制识别方法较传统方法优势明显。深度模型中包含多个隐藏层, 因此不需要传统方法中复杂的特征提取步骤或设定阈值即可自动判别特征, 其灵活性使得其相较于传统方法具有更低的计算复杂度与更高的识别精度。数据设计是深度学习解决问题过程中的关键, 现有用于调制识别的数据集为软件定义无线电 (Software Defined Radio, SDR) 产生的全仿真数据集, 存在信号特征覆盖不全及参数分布与实际应用场景存在偏差的问题。

针对目前数据集存在的问题, 本文提出一种基于真实数据的镜像数据增强技术, 原始信号由矢量信号源 SMCV100B 生成, 使用接收机接收, 使其包含发射过程与接收过程应包含的影响, 后对其进行数据增强处理, 使其更加符合目前调制信号的实际情况, 并且包含接收到的信号可能存在的一定范围内的各类因素, 如不同的符号速率、多普勒频移、采样率、信噪比、相移、载波偏移等。针对所制作数据集, 构建特征提取能力更强的模型, 使用该模型在现有数据集上进行训练, 对比现有先进模型以验证模型设计有效性。最后, 使用所构建数据集进行模型训练, 并进行离线和在线测试。

本文主要贡献如下:

(1) 提出了一种针对现实条件下调制识别任务的半仿真数据镜像增强设计方法, 较目前所用仿真数据集更贴合现实条件下的信号特征, 同时避免真实数据所带来的特征分布不均匀等数据设计问题。该方法相对全部以实际的采集信号制作数据集而言, 大大节省了工作量, 提高了工作效率, 同时体现实际场景数据集样本的各态历经性。

(2) 提出了多模态多尺度卷积融合加去噪机制的 Transformer 调制识别模型。首先设计一个数据模态变换模块, 为后续的多模态处理提供数据基础。接着设计了多尺度卷积模块, 实现不同尺度特征的提取。然后设计一个卷积融合模块, 实现特征融合与数据压缩, 从而降低后续模型的复杂度。之后设计压缩激励模块 (Squeeze and Excitation, SE) 用于噪声压缩, 突出重要特征。最后连结 Transformer 的编码器进一步进行特征提取, 所得特征用于调制类型识别。

(3) 针对所设计的模型, 使用开源的 RadioML2018.01a 数据集与所设计的增强数据集, 在模型训练、测试与在线落地测试三个方面进行了比较研究, 证实所构建模型与数据集的有效性, 这对基于深度学习调制识别技术落地具有重要的意义。

1 相关工作

随着深度学习技术的发展, 越来越多研究者致力于构建神经网络模型来解决调制识别这一分类问题, 常用的深度学习模型包括卷积神经网络 (Convolutional Neural Networks, CNN)、循环神经网络 (Recurrent Neural Network, RNN)、Transformer 架构。CNN 能够学习数据的空间特征, 但缺乏时间建模能力, RNN 可以利用时间序列对输入信号进行分析, 但无法处理含噪声的无序序列。Transformer 架构包含多头自注意力机制和位置编码, 拥有更好的全局处理能力。

李钦等人^[7]通过时频分析将雷达信号转化为时频图像, 提升了低信噪比下雷达信号调制识别精度。Chen 等人^[8]将连续样本信号分为片段并拼接为二维特征矩阵, 摆脱了传统模型依赖局部特征的缺陷, 并平衡了模型推导速度与精度。Pan 等人^[9]提出了基于双输入卷积神经网络的调制识别方法, 将信号转化为高功率谱灰度图与时频灰度图作为双输入, 改善了

相移键控信号(Phase Shift Keying, PSK)和正交幅度调制信号(Quadrature Amplitude Modulation, QAM)信号的识别效果。上述方法将信号信息转化为矩阵信息,使用CNN进行识别,一定程度上忽略了信号的时序信息,且模型复杂度较高。

Gao 等人^[10]使用门控网络(Gated Recurrent Unit, GRU)学习信号样本的信噪比特征,并使用不同的网络对不同信噪比的样本进行识别,在低信噪比取得了较高识别率。Zhu 等人^[11]提出一种多特征轻量化自动调制识别网络,采用深度可分离卷积和GRU降低模型参数,多路输入以同步提取空间、时间与频率特征,并使用Dropout算法抑制了过拟合。Liu 等人^[12]将基于对比学习的多模态融合模型用于自动调制识别,将信号与关键特征进行融合作为输入。Xing 等人^[13]提出了一种使用深度可分离卷积的轻量化调制识别网络,通过信噪比(Signal-to-Noise Ratios, SNRs)特征校准模块动态调整特征以提升抗噪鲁棒性。Lu 等人^[14]提出了动态自适应加权特征融合神经网络,将正交和幅度(In-phase and Quadrature, IQ)信号数据与多个特征参数构成的循环特征矩阵作为双输入,提升了模型识别性能。上述方法在输入IQ信号的同时输入了信号特征,使用深度卷积层降低了模型复杂度,为模型落地测试作了准备。

Vaswani 等人^[15]提出了一种包含自注意力模块和位置编码的Transformer架构,与传统的卷积模型依赖局部感受野的概念不同,其通过自注意力机制有效捕捉输入中各个部分之间的全局关系。自注意力机制使得每个输入元素都能与其他元素互动,从而更好地捕捉到长程依赖关系。Dao 等人^[16]将Transformer架构与视觉变换器(Vision Transformers, ViT)^[17]结合,在数据输入使用了多个卷积核并行提取特征,牺牲了模型复杂度但提高了模型特征提取能力。李广丽等人^[18]以残差网络提取局部特征,设计多视角Transformer编码模块捕获全局上下文,实现双网络双向知识传递与优势互补。Zeng 等人^[19]以信号瞬时幅度、相位和频率为输入,引入轻量化分组卷积残差模块进行特征预提取,使用Transformer作为核心架构,取得较高精度的同时计算复杂度增加有限。然而,上述Transformer架构相关模型受限于单一数据集,缺乏真实场景验证,未能充分检验其泛化能力及抗噪能力之外的其他鲁棒性指标。

随着许多种不同设计理念的模型相继问世,基于特定数据集的识别准确率也在不断提高,常用数据集有O'Shea 等人^[20]相继公开的RadioML2016和RadioML2018系列数据集,以及Tekbiryk 等人^[21]设计的HisarMod2019.1数据集,其中RadioML2018.01a数据集

使用最为广泛,其包含24种调制方式以及26种信噪比。此类数据集通过参数化建模生成包含多调制类型、动态信噪比及基础信道损伤的标准化样本,但未包含相关物理层损伤,如信号上变频发射与接收机滤波与下变频过程中所产生的影响,并且相关参数范围不可控,生成流程复杂,在面对需要大量数据的场景时无法便捷生成数据集。

2 镜像数据增强技术

本文提出的镜像增强技术,类比物理场景中单一真实场景经多镜反射生成多个镜像的现象,该技术以单一真实场景采集的数据为基础,通过符号速率、多普勒、初始相位、载波偏置等维度的镜像变换生成数据,而非额外采集真实场景数据。镜像数据增强的目的是实现场景的多样性,进而实现数据的多样性,提升模型的泛化能力与落地效果。

为使数据充分包含发射过程与接收过程产生的影响,在数据采集时进行真实的信号发射与接收步骤。具体为使用矢量信号源将信号加以载波发射,接收机接收时进行滤波和下变频操作。同时采用半仿真的数据设计,采用仿真设计模拟真实信道条件,避免完全数据采集构建数据集带来的参数分布不均的问题,减少数据采集的重复工作。如图1所示,接收机以高采样率采集原始信号序列,经过平滑滤波、下变频和下采样处理后,分别存储为IQ两路信号。之后进行包括符号速率增强、载波偏置增强、初始相位增强、多普勒增强、噪声增强在内的镜像数据增强。

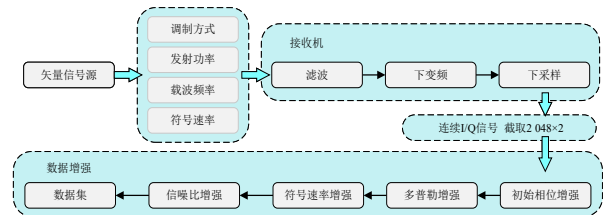


图1 镜像数据增强技术流程图

Figure 1 Flowchart of mirroring data augmentation

2.1 符号速率增强

在实际通信过程中,信号不同的符号速率 R_s 或接收端采样率 f_s 会映射为一个符号的采样点数 N_s ,其数学关系可表示为

$$N_s = \frac{f_s}{R_s} \quad (1)$$

通过调整 R_s 和 f_s ,可模拟实际通信场景中因收发设备差异导致的 N_s 动态变化,增强数据集对符号速率与采样率联合变化的表征能力。具体的,通过矢量信号源控制发射信号的 R_s ,联合接收机的高采样率 f_s

生成原始高分辨率信号。随后对采集数据统一进行不同下采样因子($M=f_s/f'_s$,其中, f'_s 为目标采样率)的抽取操作,模拟实际采样接收效果。不同的符号速率,在相同采样速率下,意味着每个符号采样点数不一样;同样的符号速率,在不同的采样速率下,每个符号的采样点数也不一样;数据符号速率上的增强,本质上反映了每个符号采样点数在一定范围内变化。

2.2 多普勒增强

多普勒频移主要是由于多普勒效应引起的通信载波频率的偏移,为模拟现实情况下广泛存在的该现象,对接收机初步数据采集后的数据进行一定范围内的随机多普勒频移处理,保证每个信号数据的多普勒频移量不同,增强数据对多普勒频移的适应性。数据集中多普勒频移范围大小结合发射信号频率进行设置。

设 W_0 为多普勒频移,分别产生相互正交的两路频率为 W_0 的信号,与原始信号或经延迟的多径信号相乘即可。即: $I = \cos(W_0 t)$, $Q = \sin(W_0 t)$,且

$$S_0(t)^*(I + jQ) = \exp(j(\omega_0 + W_0)t) \quad (2)$$

从上式可见,原始信号频率 ω_0 叠加上了多普勒频移 W_0 。

2.3 信噪比增强

由于接收机的设计也会对信号信噪比产生影响,这种情况能更真实地模拟实际情况下信号的传输。此外,为了有效评估使用此数据集进行训练的模型对噪声的适应能力,在在线测试过程中,本方法通过改变矢量信号源发射功率以改变信噪比大小,对采集后的数据进行噪声标定估计,得出某功率下发射与接收信号的信噪比范围。在镜像数据增强时,计算出信号样本功率后,按信噪比大小的不同添加高斯白噪声,计算公式如式(3):

$$P_n = \frac{P_s}{10^{\frac{\text{SNR}}{10}}} \quad (3)$$

其中, P_n 表示需要添加的噪声功率; P_s 为原信号样本功率; SNR 为所需信噪比。

对于信噪比标定,同种调制方式信号频谱具有一定相似性,绘制其频谱后可以去除信号本身来估计噪声功率,并估计出该条件下某种调制信号的信噪比。改变发射功率即可使信号信噪比等比例下降,同时估计低功率下发射信号的信噪比,以此即可在不同信号源与接收机条件下,估计得到所得数据信号的信噪比。经测算,实验所用信噪比范围对应的信号源发射功率如图2所示,在-8 dB到30 dB的范围内,信号源发射功率与信号信噪比线性拟合方差为0.39,呈现出良好线性关系,证实了此范围内信号信噪比标定的有

效性,为后续在线测试提供了信噪比相关测试的参考。

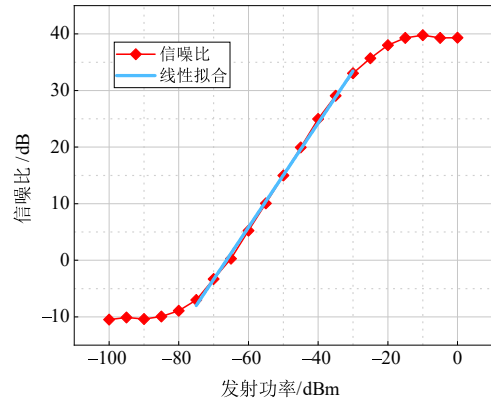


图2 实际发射功率与信噪比标定关系曲线以及二者线性拟合关系
Figure 2 Calibration curve of actual transmission power versus signal-to-noise ratio and their linear fitting relationship

2.4 平滑滤波

信号下采样处理前的滤波操作是接收端的关键环节,此过程由接收机的带通滤波完成,以限制信号的带宽,去除带外噪声,从而提高信号的质量和后续处理的精确性。在带通滤波之前,选择采用均值滤波或者其他滤波方式来进行平滑滤波以提高信号质量。本文实验使用均值滤波进行平滑滤波处理。设接收信号为长度为 N 的序列 $x[n]$,采用滑动窗口长度为奇数 M ,则有输出信号 $y[n]$ 如下式:

$$y[n] = \frac{1}{M} \sum_{k=-(M-1)/2}^{(M-1)/2} x[n+k], \quad n=0, 1, \dots, N-1 \quad (4)$$

为避免滤波引入的时域失真,采用合适的边缘处理方法。采用取更长的信号序列进入滤波器,滤波后去除两端数据。在原始信号的两端分别扩展 L 个样本点,得到长度为 $N+2L$ 的扩展信号序列 $\tilde{x}[n]$,对扩展信号序列进行滤波操作,得到滤波后的序列 $\tilde{y}[n]$,去除两端的 L 个样本点,得到最终的滤波序列 $y[n]$,其长度仍为 N 。

2.5 初始相位增强

2.5.1 样本偏置增强

实际接收机由于符号定时偏差,可能从任意位置开始采样,致使信号截取的初始相位随机化,因此设计以采样时间间隔为单位的样本偏置增强。设时域偏移量 $\Delta t \in [0, T_s)$,其中 T_s 为符号周期。则对应的离散信号偏移点数为

$$n_0 = \frac{\Delta t}{T_{\text{sample}}} \quad (5)$$

其中, T_{sample} 为采样周期,是采样率的倒数。在进行数据增强时,每条信号的偏移点数 $n_0 \sim U(0, N \cdot N_s)$,其中

N 为符号数, N_s 为每个符号的采样点数。随机偏移后的信号可表示为

$$s_{\text{shift}}[n] = s[(n_0 + n) \bmod (N \cdot N_s)], n = 0, 1, \dots, L - 1 \quad (6)$$

其中, L 代表截取长度, 且 $L \leq N - n_0$, 以避免循环填充。

2.5.2 相位旋转增强

与上述时域随机偏移不同, 相位旋转增强引入了非线性因素和频域扰动, 是 $0 \sim 2\pi$ 范围内均匀分布的相位偏移增强, 更针对对于相位敏感的调制方式, 如 QAM、PSK 等。定义接收信号复基带表示为 $s(n) = A(n)e^{j\varphi(n)}$, 其中 $A(n)$ 为幅度, $\varphi(n)$ 为相位, 则施加随机相位偏移后得到的信号表示为

$$s_{\text{rot}}(n) = s(n)e^{j\theta} \quad (7)$$

其中, $\theta \sim U(0, 2\pi)$ 。

3 模型结构

面对多径衰落、干扰等各种因素对于通信性能的影响, 接收机部分需要较低复杂度的算法来进行自适应调制识别以提供更高标准的通信服务。本节中提出了一种使用压缩激励模块的 Transformer 结构自动调制识别模型, 模型框架如图 3 所示。模型包括信号多模态变换、多尺度卷积特征融合、位置编码、Transformer 编码器和分类层。经计算, 模型参数量共 168 000, 浮点运算次数 (Floating Point operations, FLOPs) 为 12 M, 其中 Transformer 结构 FLOPs 共约 9.5 M, 其余卷积层、池化层和全连接层的 FLOPs 共约 2.5 M。

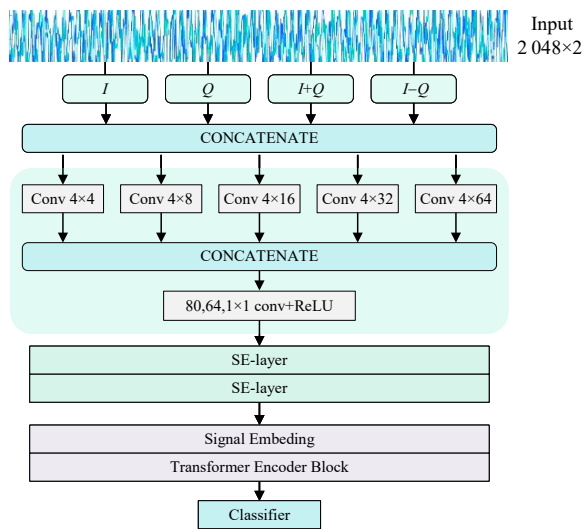


图 3 模型结构

Figure 3 Model structure

3.1 信号模态变换

信号输入为 IQ 复数信号, 其中 I 和 Q 分别代指同

向量和正交分量, 它们分别是复数信号的实部和虚部。将两路信号进行线性运算后重构以增强信号的代表能力, 具体方法如下式所示:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} I \\ Q \end{bmatrix} = \begin{bmatrix} I \\ Q \\ I+Q \\ I-Q \end{bmatrix} \quad (8)$$

其中, Y_1 和 Y_2 分别代表原 I 路和 Q 路的信号样本; Y_3 和 Y_4 代表线性处理后的信号样本, 将 Y_1 、 Y_2 、 Y_3 和 Y_4 拼接以实现多模态特征提取, 所得信号覆盖了更多信号特征, 如原始 IQ 信号不能表征的两路分量叠加和差分后的联合幅度分布信息。

设 $\begin{bmatrix} I' \\ Q' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} I \\ Q \end{bmatrix} = \begin{bmatrix} I+Q \\ I-Q \end{bmatrix}$, 以 32 QAM、

128 QAM、32 PSK、32 APSK (32 Amplitude and PSK) 为例, 原始信号的 I、Q 星座图和预处理信号的星座图分别如图 4(a)~(d) 所示。变换后的星座图与原始星座图相比, 其总体覆盖的范围更为广泛, 且星座图中点与点之间的间隔更大。这种间隔的增加有助于提高信号的可区分性, 使得模型能够更加清晰地识别不同的信号状态。

3.2 多尺度卷积特征提取与融合

将线性处理后的信号样本并行通过五个不同卷积核的卷积层, 卷积核大小分别为 (4, 4)、(4, 8)、(4, 16)、(4, 32) 和 (4, 64), 以实现不同尺度层次的特征提取, 多尺寸卷积核的并行架构利用不同尺寸的卷积核分别捕捉不同尺度的细粒度特征, 通过差异化的感受野设计实现特征提取的协同优化。五个并行卷积操作以及用于特征融合的点卷积层具体参数如表 1 所示。在多尺度卷积特征提取后, 在通道维度连接所有并行层的输出张量, 并使用一个跨度为 (1, 1) 的点卷积层和 ReLU 激活函数进行特征融合, 同时实现数据压缩, 以减少后续模块处理的规模, 所得输出为 $X_0 \in \mathbb{R}^{C_{\text{out}} \times H_{\text{out}} \times 1}$ 。

3.3 SE 去噪模块

SE-Net 是一种通过动态调整通道权重来增强特征表达能力的注意力机制模块, 其核心结构包括压缩和激励两部分。压缩通过全局平均池化将输入特征图的空间维度压缩为的通道描述符, 获取全局感受野并表征各通道的全局分布。SE 模块结构如图 5 所示, 给定输入 $x \in \mathbb{R}^{c \times h \times w}$, 经过给定的映射条件 F_{tr} 后得到特征图大小为 $c_1 \times h \times w$, 之后该特征图分别经过压缩操作 F_{sc} 和激励操作 F_{ex} 生成原特征图的每个通道的权重, 大小为 $c_1 \times 1 \times 1$, 将此通道权重与特征图逐个相乘, 形成新的特征图 \tilde{x} 。

在本模型中, SE-layer 由池化层、两个全连接层和

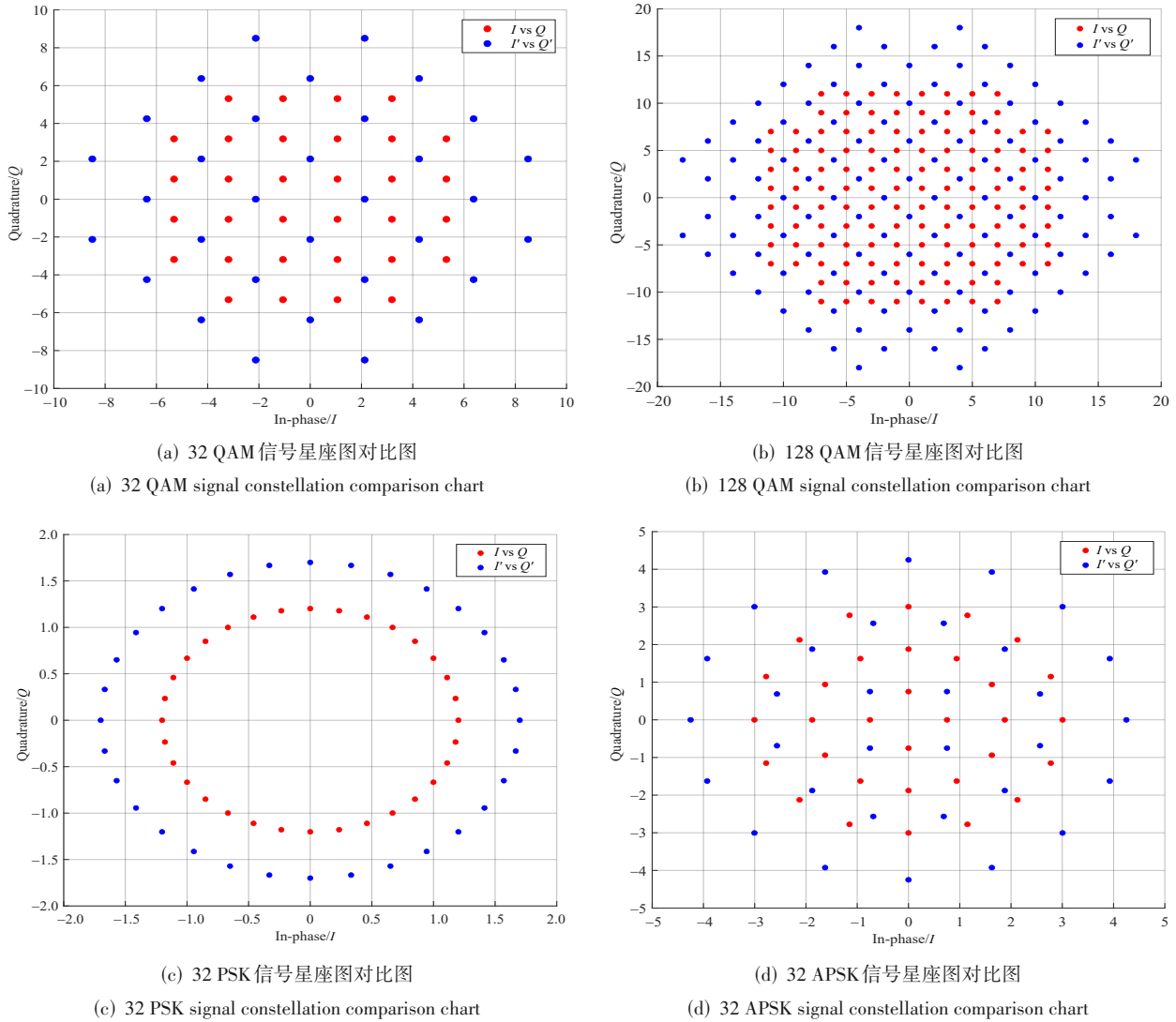


图4 四种调制信号的星座图对比

Figure 4 Comparison of constellation diagrams of four modulated signals

表1 并行卷积层参数

Table 1 Parallel convolutional layer parameters

卷积层	卷积核 (kernel size)	步长(stride)	填充(padding)	输入/输出 通道数
CONV1	(4,4)	(4,8)	(0,0)	1/16
CONV2	(4,8)	(4,8)	(0,2)	1/16
CONV3	(4,16)	(4,8)	(0,6)	1/16
CONV4	(4,32)	(4,8)	(0,12)	1/16
CONV5	(4,64)	(4,8)	(0,30)	1/16
CONV6	(1,1)	(1,1)	(0,0)	80/64

非线性激活函数 ReLU (Rectified Linear Unit) 组成, 将 X_0 作为输入, 使用两个全连接层与非线性激活函数 ReLU 生成通道权重, 学习通道间的非线性依赖关系, 并通过 Sigmoid 函数归一化, 最终将权重与原特征图

相乘, 实现特征重标定, 输出尺寸不变。其独特的通道注意力机制使其拥有降噪能力, 压缩的过程提取了通道级信息, 并在后续特征图与通道权重的逐个相乘实现噪声抑制和信号增强, 这一过程通过动态调整特征通道权重来增强有效信号并抑制噪声。

3.4 位置编码

位置编码操作如图 6 所示, 对压缩激励模块输出

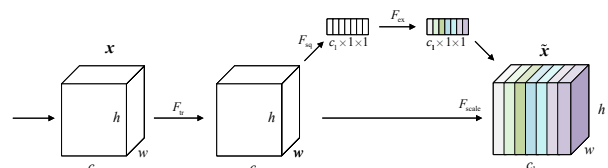


图5 SE 结构图

Figure 5 SE structure

X_0 转置, 并拼接加入类别标记 $X_{cls} \in \mathbb{R}^{1 \times H_{out}}$, 得到 $X_1 \in \mathbb{R}^{(H_{out}+1) \times C_{out}}$ 。设计待结合的正余弦交替的位置编码表示如下:

$$E_t^i = \begin{cases} \sin(\omega_k t), & i = 2k \\ \cos(\omega_k t), & i = 2k + 1 \end{cases} \quad (9)$$

其中, t 表示在类别标记序列中的实际位置; $i = 0, 1, 2, 3, \dots, \frac{C_{out}}{2} - 1$, 表示位置编码向量中的具体维度; $\omega_k = \frac{1}{10\,000^{2k/C_{out}}}$, 表示正余弦函数采用的频率。此位置编码过程可完整表达为 $X_2 = \text{Concat}(x_{cls}, X_0^T) + E_t^i$ 。

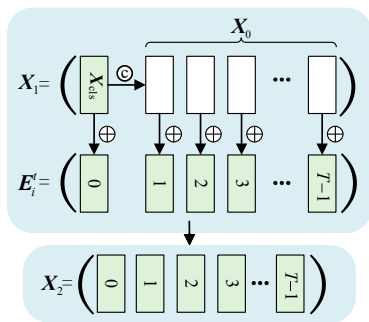


图6 位置编码

Figure 6 Position encoding

3.5 Transformer 编码器

如图 7, 将输入 X_2 通过三个独立的权重矩阵 W_q, W_k, W_v 获得查询向量 Q 、键向量 K 和值向量 V , 然后通过计算每一对元素的查询向量和键向量的点积, 计算它们之间的注意力权重, 接着通过键向量维度的平方根进行归一化, 并通过 Softmax 函数得到最终的权重。具体公式如下:

$$\begin{aligned} Q &= X_1 W_Q \\ K &= X_1 W_K \\ V &= X_1 W_V \end{aligned} \quad (10)$$

$$S(X_2) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) \cdot V$$

其中, $X_2 \in \mathbb{R}^{(H_{out}+1) \times C_{out}}$ 为位置编码模块输出; Q, K, V 分别代表线性变换得到的查询矩阵、键矩阵、值矩阵; W_Q, W_K, W_V 是可训练的权重矩阵; d_k 为缩放因子, 等同于键向量维度大小。

与此同时, 在上述过程中, 并行使用 N_h 个自注意力操作来增强得到的 $S(X_2)$, 具体操作为在通道维度拼接注意力头生成的不同的 $S(X_2)$ 。之后将拼接后的特征输入并联结构的前馈神经网络 (Feed-Forward Neural network, FFN) 进行特征整合, 引入非线性激活函数 GELU (Gaussian Error Linear Units), 增强模型表

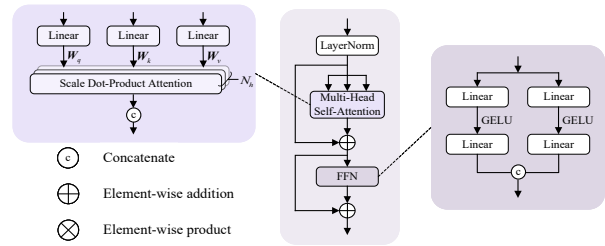


图7 Transformer 编码器

Figure 7 Transformer encoder

达能力。

4 实验结果

本节通过离线和在线测试证明所构建模型与所设计数据集的有效性。首先使用 RadioML2018.01a 对所提出的模型进行训练、测试, 并用所得模型进行在线测试; 接着使用增强数据集进行对所提出的模型进行训练、测试, 并进行在线落地测试。

4.1 模型性能评估

将数据集 RadioML2018.01a 按 6:2:2 分为训练集、测试集和验证集, 选取自适应矩估计优化器 (Adaptive moment estimation, Adam), 初始学习率设置为 0.000 2, 每 20 轮次降为先前轮次的 0.8 倍, 训练 300 轮次后损失函数已收敛。选择类别标签的交叉熵作为损失函数, 并设置批次大小为 512。

图 8 展示了各模型在 RadioML2018.01a 中正信噪比样本上的识别性能。在 6 dB 及以上时, 本文所设计模型在所有对比模型中识别率均为最高。低信噪比时, 模型也展现了极大性能优势, 6 dB 时识别率已经达到了 90% 以上, 较其他模型高出 6.4%。信噪比在 6 dB 以下时, SCSNet (Spatial and Channel-wise Shrinkage neural Network)^[22] 识别率最高, 这是因为其是专注于降噪的模型, 且复杂度较高。在 26 dB 时, 本文模型较其他模型也取得了平均 1.3% 的优势。

表 2 展示了各模型的复杂度与识别率比较。使用可训练参数量和每秒浮点数运算次数进行模型复杂度对比, 由于所构建模型使用了 SE-layer 和并行的 FFN 等低复杂度设计, 且多头自注意力模块本身复杂度较低, 其可训练参数数量仅高于 SCGNet^[23] 这一稀疏连接的 CNN 网络, 较 SCSNet 等模型有较大提升。同时, 网络在模型中拥有最低的 FLOPs。对比各模型准确度, 所构建模型较低复杂度的 SCGNet 的最高识别率提高了 4%, 但可训练参数量仅增加了 61 000, 相比其他复杂度更高的模型也有平均 0.56% 的优势。这表明所构建模型较好地平衡了参数量与准确率, 且更强的特征提取能力使其在面对更复杂的调制信号时也有更好效果。

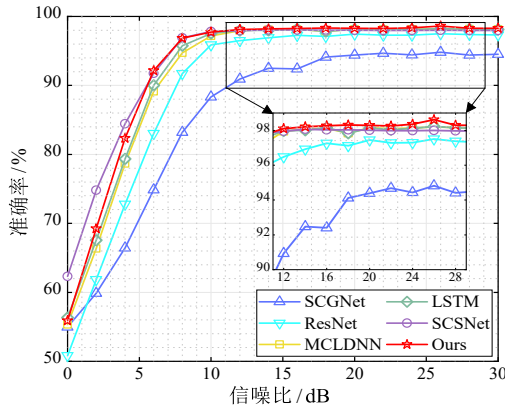


图 8 各模型在 RadioML2018.01a 上的分类效果

Figure 8 Classification performance of each model on RadioML2018.01a

表 2 模型参数量与性能比较

Table 2 Comparison of model parameters and performance

网络名称	参数量(params.)	FLOPs/M	最高/平均准确率/%
SCGNet ^[23]	107 000	36	94.39/58.82
ResNet ^[20]	239 000	23.45	97.46/60.08
MCLDNN ^[24]	407 000	49.15	98.20/62.51
LSTM ^[25]	203 000	25.82	98.22/62.51
SCSNet ^[22]	580 000	144.05	98.36/65.00
Ours	168 000	12.02	98.62/63.89

4.2 数据集泛化能力评估

经实验,使用现有仿真数据集训练的模型无法完成真实条件下的调制识别任务,如使用 RadioML2018.01a 训练的本文模型无法实现对本文设计数据集的识别,识别结果的混淆矩阵如图 9 所示,平均准确率仅 10.4%,数字调制方式大多识别为二进制振幅键控(Binary Amplitude Shift Keying, 2ASK)和二进制相移键控(Binary PSK, BPSK),模拟调制方式均识别为四进制频移键控(4-ary Frequency Shift Keying, 4FSK)。这证明针对现存仿真数据集训练的模型泛化性较差,无法适应于真实数据,相关实验模型在用于实际条件下的调制信号识别时仍然充满较大不确

定性。使用增强数据集训练的模型识别结果的混淆矩阵如图 10 所示,对比图 9 可明显看出明显的改善效果。

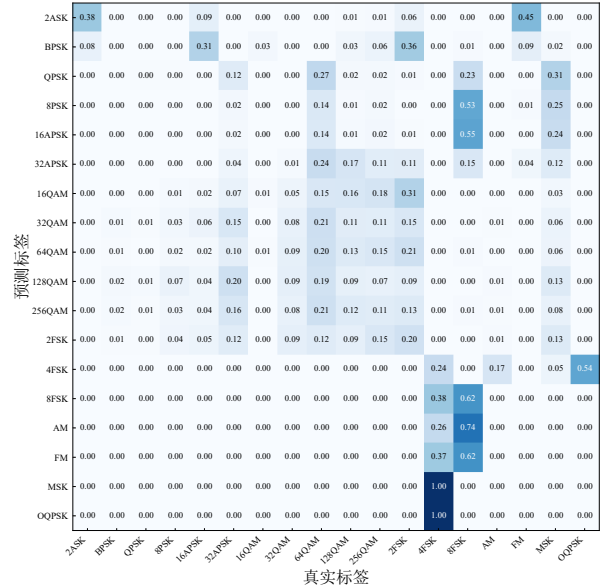


图 9 仿真数据集训练模型在真实数据下的识别效果

Figure 9 Recognition performance of the model trained on simulated datasets when applied to real data

4.3 增强数据集上的训练与测试

4.3.1 数据集信息

新设计的增强数据集预计包含多种信号特征,涵盖了实际通信环境中的多样性。其包含不同的多普勒频移、初始相位、信噪比、符号速率、滤波窗口大小和采样率等变化,从而确保训练过程中能够涉及多种可能的通信环境。具体信息如表 3 所示。

为保证测试的真实性与有效性,此节数据中离线测试结果为增强数据集中的测试集结果,在线落地测试为不断注入数据的不间断测试,训练参数设置与上一节中相同。通过修改发射机信号参数与接收机参数,得到在不同信号速率、发射功率、多普勒频移、载

表 3 数据集信息

Table 3 Dataset information

信息类别	内容
调制方式数量	18
模拟调制类别	AM, FM
数字调制类别	ASK, MSK, BPSK, QPSK, OQPSK, 8 PSK, 16APSK, 32APSK, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM, 2 FSK, 4 FSK, 8 FSK
单类别调制方式信号数量	4 096
数据格式	2×2 048
信噪比	-10:2:30
单信噪比样本数量	65 536

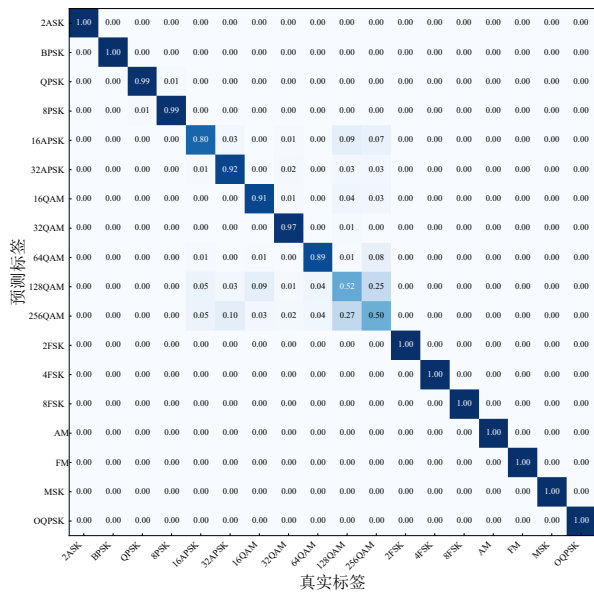


图 10 增强数据集训练模型在真实数据下的识别效果

Figure 10 Recognition performance of the model trained on the augmented dataset under real data

波频率下模型识别率。

4.3.2 测试结果

所构建模型在不同信噪比下离线测试和在线测试的识别准确率如图 11 所示,其中 16 dB 以上测试集平均准确率为 90.1%,在线落地测试平均准确率为 90.6%。由于所制作数据集中信号样本的各参数范围更广,因此在低信噪比下在线落地测试准确率更高。

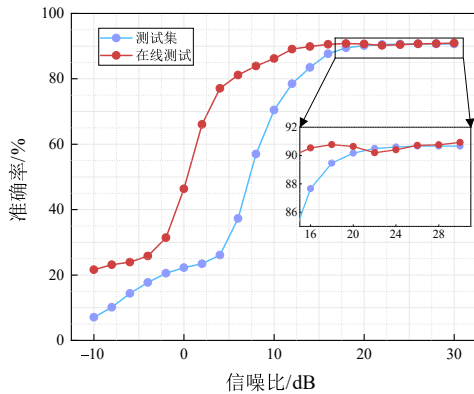


图 11 不同信噪比下模型在测试集和在线落地测试时的识别准确率

Figure 11 Recognition accuracy of the model on the test set and in online deployment tests under different signal-to-noise ratios

图 12 和图 13 显示了所设计网络在线识别过程中不同信噪比下各类调制方式的识别精度结果。网络能以高准确度识别出各类低阶数字及模拟调制、PSK 以及 FSK 调制;在信噪比大于 6 dB 时,网络对 16 QAM 和 32 APSK 识别精度均高于 85%,对 32 QAM、64 QAM 和 16 APSK 识别精度均高于 93%;但对于 128 QAM 以

及 256 QAM 网络的识别精度较低阶 QAM 调制下降了约 35%,这是由高低阶 QAM 调制的层次关系引起的,且高阶 QAM 调制更易受到不利传输条件的影响,数据设计中各类因素的模拟加剧了这一特点,导致高阶 QAM 准确率较低。

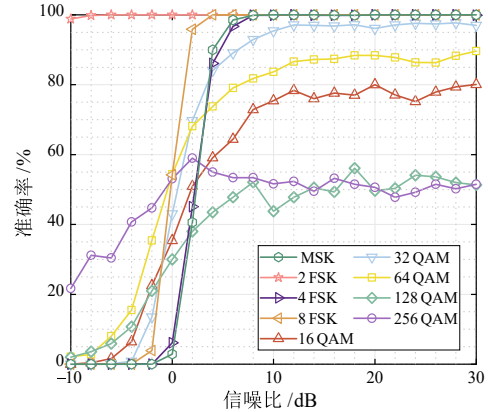


图 12 FSK 和 QAM 信号在不同信噪比下的识别准确率

Figure 12 Recognition accuracy of FSK and QAM signals at different signal-to-noise ratios

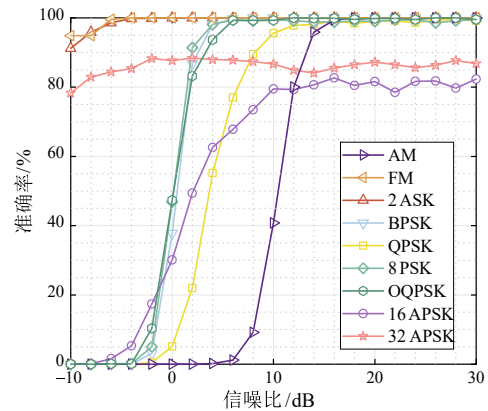


图 13 模拟调制、低阶调制和 APSK 信号在不同信噪比下识别准确率

Figure 13 Recognition accuracy of simulated modulation, low-order modulation, and APSK signals at different signal-to-noise ratios

4.4 数据模态消融测试

通过对比实验,研究样本尺寸和输入模态数量对模型准确率的影响机制,确定合适的数据尺寸与输入模态。如表 4 所示,IQ 信号长度从 1 024 增加到 2 048 时,模型准确率显著提升 8.32%,这是因为模型能在更长的时序信号中捕获更多高频特征与相位信息。在此基础上,进一步对模型输入的通道数量进行调整。模型输入原本分为 I 路和 Q 路两路,在增加 I+Q 和 I-Q 两路后准确率提高了 8.22%,达到 91.9%。这证明了数据模态设计对模型性能的关键作用。

表 4 样本尺寸和模态对模型准确率的影响

Table 4 Effect of sample size and modality on model accuracy

样本尺寸	输入模态数	准确率/%
2×1 024	2	75.36
2×2 048	2	83.68
2×2 048	4	91.90

5 结论

针对调制识别开源数据集训练的模型在落地面临的困境,本文提出基于少量实际采样信号进行镜像增强的方法,设计了增强数据集用于模型训练。同时设计基于多模态多尺度卷积融合加去噪机制的编码器调制识别模型。通过开源数据集进行模型训练,在 12 dB 以上信噪比的测试数据上,获得了 98.3% 的总体识别精度。但以开源训练的模型,用于实际场景进行在线测试的时候,识别精度仅约 10.4%。使用增强数据进行模型训练,所得的模型参数用于在线测试,获得了极大的识别精度改善,高信噪比下在线调制识别精度达到 91.9%。在线测试表明,本文提出的数据增强技术与深度调制识别网络是有效的,这对调制识别、信号识别的落地具有重要意义。

参考文献

- [1] Dobro O A, Abdi A, Bar-Ness Y, et al. Survey of automatic modulation classification techniques: Classical approaches and new trends[J]. IET Communications, 2007, 1(2): 137-156.
- [2] Zhu Zhechen, Nandi A K. Blind digital modulation classification using minimum distance centroid estimator and non-parametric likelihood function[J]. IEEE Transactions on Wireless Communications, 2014, 13(8): 4483-4494.
- [3] Zheng Jianping, Lv Yanfang. Likelihood-based automatic modulation classification in OFDM with index modulation[J]. IEEE Transactions on Vehicular Technology, 2018, 67(9): 8192-8204.
- [4] Chang D C, Shih P K. Cumulants-based modulation classification technique in multipath fading channels[J]. IET Communications, 2015, 9(6): 828-835.
- [5] Wang Sese, Sun Zhuo, Liu Siyuan, et al. Modulation classification of linear digital signals based on compressive sensing using high-order moments[C]//2014 European Modeling Symposium. Piscataway: IEEE, 2014: 145-150.
- [6] Qu Zhiyu, Hou Chenfan, Hou Changbo, et al. Radar signal intra-pulse modulation recognition based on convolutional neural network and deep Q-learning network[J]. IEEE Access, 2020, 8: 49125-49136.
- [7] 李钦, 刘伟, 牛朝阳, 等. 低信噪比下基于分裂 EfficientNet 网络的雷达信号调制方式识别[J]. 电子学报, 2023, 51(3): 675-686.
- [8] Li Qin, Liu Wei, Niu Chaoyang, et al. Radar signal modulation recognition based on split EfficientNet under low signal-to-noise ratio[J]. Acta Electronica Sinica, 2023, 51(3): 675-686. (in Chinese)
- [9] Chen Yantao, Dong Binhong, Liu Cuiting, et al. Abandon locality: Frame-wise embedding aided transformer for automatic modulation recognition[J]. IEEE Communications Letters, 2023, 27(1): 327-331.
- [10] Pan Yiyu, Gao Feng. Signal modulation recognition of satellite Internet based on dual input convolutional neural network[C]//2024 9th International Conference on Signal and Image Processing. Piscataway: IEEE, 2024: 13-16.
- [11] Gao Jiabin, Cao Qinglong, Chen Yutian. MoE-AMC: Enhancing automatic modulation classification performance using mixture-of-experts[PP/OL]. V1.arXiv (2023-12-04)[2025-06-01]. <https://doi.org/10.48550/arXiv.2312.02298>.
- [12] Zhu Zhengyu, Zhou Ning, Wang Zixuan, et al. MSGNet: A multi-feature lightweight learning network for automatic modulation recognition[J]. IEEE Communications Letters, 2024, 28(11): 2553-2557.
- [13] Liu Fugang, Pan Jingyi, Zhou Ruolin. Contrastive learning-based multimodal fusion model for automatic modulation recognition[J]. IEEE Communications Letters, 2024, 28(1): 78-82.
- [14] Xing Jingyi, Xu Shengfa, Xu Chunpeng, et al. A lightweight modulation recognition network with depthwise separable convolution and SNR enhancement[C]//2025 7th International Conference on Electronics and Communication, Network and Computer Technology. Piscataway: IEEE, 2025: 530-534.
- [15] Lu Yitong, Hou Shujuan, Zhang Qin, et al. DAWFNN: An automatic modulation recognition method based on multi feature fusion[C]//2025 10th International Conference on Computer and Communication System. Piscataway: IEEE, 2025: 480-485.
- [16] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2017: 6000-6010.
- [17] Dao T T, Noh D I, Pham Q V, et al. VT-MCNet: High-accuracy automatic modulation classification model based on vision transformer[J]. IEEE Communications Letters, 2024, 28(1): 98-102.
- [18] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is

worth 16×16 words: Transformers for image recognition at scale[PP/OL]. V2. arXiv (2021-06-03) [2025-06-01]. <https://doi.org/10.48550/arXiv.2010.11929>.

- [18] 李广丽, 叶艺源, 吴光庭, 等. 联合多视角Transformer编码与在线融合互学习的乳腺癌病理图像分类模型[J]. 电子学报, 2024, 52(7): 2369-2381.
Li Guangli, Ye Yiyuan, Wu Guangting, et al. Breast cancer pathological image classification model via combining multi-view transformer coding and online fusion mutual learning[J]. Acta Electronica Sinica, 2024, 52(7): 2369-2381. (in Chinese)
- [19] Zeng Rui, Lu Zhilin, Zhang Xudong, et al. Convolutional neural network assisted transformer for automatic modulation recognition under large CFOs and SROs[J]. IEEE Signal Processing Letters, 2024, 31: 741-745.
- [20] 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.
- [21] Tekbıyık K, Ekti A R, Görçin A, et al. Robust and fast au-

tomatic modulation classification with CNN under multipath fading channels[C]//2020 IEEE 91st Vehicular Technology Conference. Piscataway: IEEE, 2020: 1-6.

- [22] Wei Tao, Li Zan, Bi Dexin, et al. Adaptive multi-dimensional shrinkage block for automatic modulation recognition[J]. IEEE Communications Letters, 2023, 27(11): 2968-2972.
- [23] Tunze G B, Huynh-The T, Lee J M, et al. Sparsely connected CNN for efficient automatic modulation recognition[J]. IEEE Transactions on Vehicular Technology, 2020, 69(12): 15557-15568.
- [24] Xu Jialang, Luo Chunbo, Parr G, et al. A spatiotemporal multi-channel learning framework for automatic modulation recognition[J]. IEEE Wireless Communications Letters, 2020, 9(10): 1629-1632.
- [25] Rajendran S, Meert W, Giustiniano D, et al. Deep learning models for wireless signal classification with distributed low-cost spectrum sensors[J]. IEEE Transactions on Cognitive Communications and Networking, 2018, 4(3): 433-445.

作者简介



李国军 男, 1978年10月出生于四川省资阳市。博士, 教授, 博导, 重庆邮电大学超视距可信信息传输研究所所长。主要研究方向为极端环境超视距无线通信与网络。
E-mail: ligj@cqupt.edu.cn



陈世奥 男, 2001年7月出生于河南省许昌市。重庆邮电大学硕士研究生。主要研究方向为调制识别。
E-mail: chensa01@126.com



王杰 男, 1994年12月出生于重庆市。重庆邮电大学博士研究生。主要研究方向为超视距无线通信与网络。
E-mail: d230101027@stu.cqupt.edu.cn



张哲铭 男, 1999年3月出生于辽宁省铁岭市。重庆邮电大学博士研究生。主要研究方向为短波通信、信号识别。
E-mail: 1193660497@qq.com



郑建忠 男, 1970年出生于四川省南充市。博士, 现任重庆邮电大学视距外可靠信息传输实验室高级工程师。主要研究方向为电子对抗。
E-mail: 1286510096@qq.com