

DOI: 10.20271/j.cnki.1673-9833.2026.3010

一种基于时频注意力机制 Swin Transformer 的 调制识别方法

周俊, 陈青辉, 文鸿, 高子雄, 彭聪

(湖南工业大学 空天技术学院, 湖南 株洲 412007)

摘要: 自动调制识别是无线通信和非合作通信软件无线电系统中的关键技术之一。为了增强信号特征识别的鲁棒性, 提高接收端对未知信号调制识别的准确性, 提出了一种基于时频注意力 (time-frequency attention, TFA) 机制 Swin Transformer 的自动调制识别方法。该方法结合了时频注意力模块和 Swin Transformer, 有效提高了信号调制识别的准确率。考虑到信号频率随时间变化是区分不同调制类型无线电信号的重要特征, 该方法先将一维无线电信号转换为二维时频图像, 作为 Swin Transformer 模型的输入。在此基础上, 引入时频注意力模块增强模型对信号特征的识别能力。实验结果表明, 与传统算法相比, 所提出的模型在识别性能上具有显著优势, 同时与深度神经网络相比, 其训练成本更低。

关键词: 调制识别; 注意力机制; Swin Transformer; 时频图像

中图分类号: TP393.094

文献标志码: A

文章编号: 1673-9833(2026)03-0070-07

引文格式: 周俊, 陈青辉, 文鸿, 等. 一种基于时频注意力机制 Swin Transformer 的调制识别方法 [J]. 湖南工业大学学报, 2026, 40(3): 70-76.

A Modulation Recognition Method Based on Time-Frequency Attention Mechanism Swin Transformer

ZHOU Jun, CHEN Qinghui, WEN Hong, GAO Zixiong, PENG Cong

(School of Aerospace, Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: Due to the fact that modulation recognition is applied as a key technology in wireless and non-cooperative communication software-defined radio systems, an automatic modulation recognition method has thus been proposed based on the time-frequency attention module (TFA) Swin Transformer to enhance the robustness of signal feature recognition as well as improve the accuracy of unknown signal modulation recognition at the receiving end. With the time-frequency attention module combined with Swin Transformer, the proposed method effectively improves the accuracy of signal modulation recognition. Considering that the variation of signal frequency over time is an important characteristic for distinguishing different modulation types of radio signals, the one-dimensional radio signal is firstly converted into a two-dimensional time-frequency image, which serves as the input for the Swin Transformer model. On this basis, a time-frequency attention module is introduced for an enhancement of the model's ability to recognize signal features. The experimental results show that, compared to traditional algorithms, the proposed model is characterized with a significant advantage in recognition performance, with a lower training cost compared to deep

收稿日期: 2024-10-04

基金项目: 湖南省自然科学基金资助项目 (2023JJ50197); 湖南省教育厅科研基金资助重点项目 (22A0418, 23A0444)

作者简介: 周俊, 男, 湖南工业大学硕士生, 主要研究方向为深度学习, 调制格式识别, E-mail: arts0720@foxmail.com

通信作者: 文鸿, 男, 湖南工业大学教授, 博士, 硕士生导师, 主要研究方向为光通信, 数字信号处理与物理层通信安全, E-mail: wenhhut@163.com

neural networks.

Keywords: modulation recognition; attention mechanism; Swin Transformer; time-frequency image

1 研究背景

在现代无线通信领域中,信号调制方式的多样性和复杂性不断增加,这给自动调制识别(automatic modulation recognition, AMR)带来了前所未有的挑战^[1]。AMR技术的核心任务是实时、准确地识别信号的调制类型,这对于频谱资源的有效利用、信号监测、干扰检测以及军事和民用通信系统的性能优化等至关重要^[2]。因此,AMR技术引起了国内外学者们的广泛关注和深入研究。

传统的深度学习方法大多侧重于从时域或者频域单独提取信号特征,这在处理具有丰富时频特性的信号时,可能会忽略一些关键信息^[3-4]。为了克服这一局限性,研究者们开始探索时频域分析技术,以便更全面地捕捉信号的本质特征。时频图像作为一种将信号的时域和频域信息融合在二维平面上的视觉表示,不仅保留了信号的时变特性,还揭示了频率随时间的变化规律,是区分不同调制方式的重要依据。文献[5]提出了利用卷积神经网络(convolutional neural network, CNN)对时频图像进行调制识别,展示了CNN在处理时频图像时的有效性。但CNN对时频图像的特征提取主要依赖于卷积层和池化层的组合,容易忽略全局特征,导致在处理长时间依赖关系时存在不足。文献[6]提出了一种结合CNN和长短期记忆网络(long-short memory neural network, LSTM)的混合模型,这为调制识别任务提供了强大的特征提取能力,从而提高了识别准确率,但是由于LSTM层的引入,导致模型在处理长序列数据时需要更多的计算资源。

随着深度学习的快速发展,作为一种新型深度学习模型的Swin Transformer被提出,其基于Transformer架构,通过自注意力机制捕捉数据中的

全局依赖关系,使其在处理具有复杂时频特性的信号时表现出色。与传统的深度学习模型相比,Swin Transformer通过其多尺度特征表示和层次化的注意力结构,能够更有效地处理长序列数据并提取深层次特征。已有研究者尝试将Swin Transformer应用于自动调制识别领域^[7]。如文献[8]探索了结合时频图像和Swin Transformer的调制识别方法,通过多尺度特征提取和全局上下文建模,有效提高了识别准确率。文献[9]提出在CNN模型中集成Swin Transformer的新型双分量雷达信号调制识别方法,增强了特征表示能力的同时降低了计算成本。此外,文献[10]提出了一种ResSwinT-SwinT调制识别框架,该框架利用Swin Transformer网络进行特征提取,显著提升了Swin Transformer处理调制识别能力。

在此背景下,本文提出一种基于时频注意力机制的Swin Transformer调制识别方法。该方法通过短时傅里叶变换(short-time Fourier transform, STFT)将一维无线信号转换为时频图像,并进一步利用Swin Transformer作为骨干架构进行特征提取和调制识别。为增强模型对时频图像特征的识别能力,引入时频注意力模块(time-frequency attention module, TFA),该模块能够突出图像中的关键区域和特征通道。最后,针对RadioML2016.10b数据集进行了性能分析和对比实验,结果表明本文所提模型能有效完成调制信号的分类任务,并提高调制信号识别准确率。

2 模型构建

2.1 基于时频注意力模块的Swin Transformer模型

本文提出的基于Swin Transformer模型并嵌入时频注意力模块的网络模型TF-Swin Transformer结构,如图1所示。

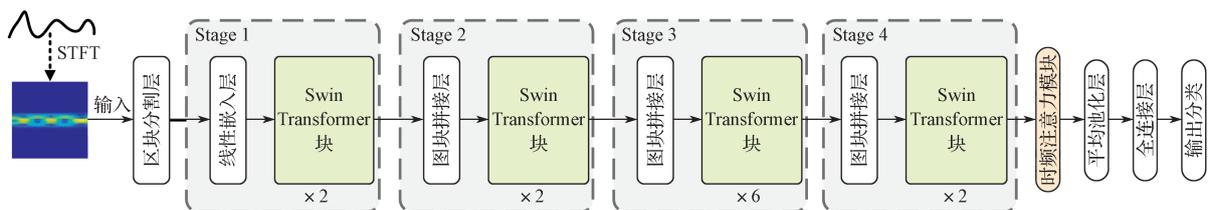


图1 TF-Swin Transformer 模型架构

Fig. 1 TF-Swin Transformer model architecture

时频图像首先经区块分割层划分为若干不重叠的图像块,再通过线性嵌入层对每个图像块进行线性变换。在每个 Stage 中, Swin Transformer 块通过自注意力机制处理图像特征,并增强图像特征之间的交互性。随后,利用下采样图块拼接层操作,缩减特征图的宽度与高度,并且扩展通道维度。为了进一步捕捉信号的时频分布特征,在 Stage 4 后引入了时频注意力模块,使网络在关注不同区域图像特征的同时,能够捕捉到高频的图像特征,从而提高调制识别的能力。模型最后通过全连接层对特征进行处理,输出分类结果。

相比于传统的调制识别方法,基于时频注意力模块的 TF-Swin Transformer 网络模型,主要有以下改进:

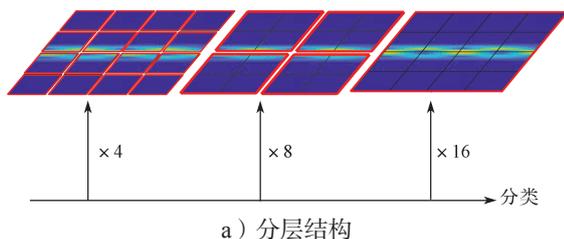
1) 对数据的预处理和降噪。数据预处理使用基于帧的 STFT 得到时频图像矩阵,经过归一化处理和高斯滤波器后,转换为二维时频图像。

2) 引入了时频注意力模块。更深入地挖掘图像的时频分布特征,通过对信号的不同部分分配不同的权重,增强了模型对信号特征的认识能力。

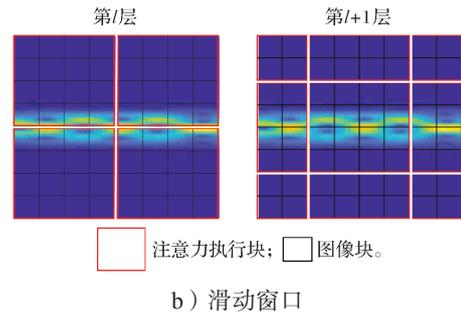
3) 将预训练的 Swin Transformer 模型的权重参数迁移到调制识别任务中,提高了模型训练速度的同时,还提升了调制识别的准确率。

2.2 Swin Transformer 网络模型

Swin Transformer 是一种专为计算机视觉任务设计的 Transformer 架构,相较于早期的 Vision Transformer (ViT),其有两个关键的改进之处。首先, Swin Transformer 采用分层结构使得特征图具有层次性,从而实现多尺度特征表示。如图 2a 所示,随着特征层加深,特征图的高度和宽度逐渐减小,分别对应 4 倍、8 倍和 16 倍的下采样。其次, Swin Transformer 引入了滑动窗口机制来实现窗口之间的信息交互。如图 2b 所示,输入一个 8×8 的特征图被划分为 HW/M^2 个窗口,其中 H 、 W 均为 8, M 为 4。随后,窗口向右下方平移 $(M/2, M/2)$ 。通过调整窗口位置,窗口能够捕捉图像中不同位置像素的信息,使模型能够利用图像中不同位置像素的特征信息,增强自相邻的非重叠窗口之间的信息交换。



a) 分层结构



b) 滑动窗口

图 2 层次化和滑动窗口操作

Fig. 2 Hierarchical and sliding window operation

Swin Transformer 块如图 3 所示,为由窗口多头自注意力 (window multi-head self-attention, W-MSA)、移动窗口多头自注意力 (shifted window multi-head self-attention, SW-MSA)、归一化 (layer norm, LN) 层、残差连接和具有高斯误差线性单元 (Gaussian error linear unit, GELU) 激活函数组成的多层感知器 (multilayer perceptron, MLP)。采用移动窗口划分方法后,两个连续的 Swin Transformer 块中的计算过程如下。

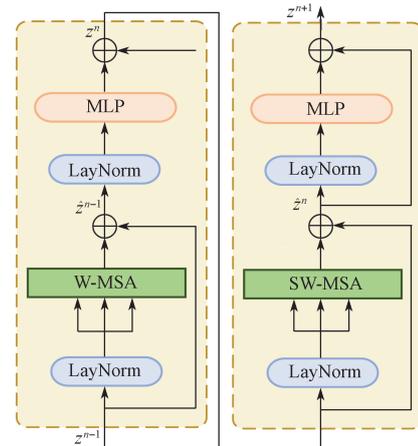


图 3 Swin Transformer 块

Fig. 3 Swin Transformer blocks

在首个 Swin Transformer 块中,输入特征 z^{n-1} 首先经过 W-MSA 模块,以实现自注意力计算;接着,通过 MLP 的 F_{MLP} ,将自注意力计算后得到的特征 \hat{z}^{n-1} 进行非线性变换,以进一步提取特征的深层次信息;随后,再次通过 LN 层的 F_{LN} 对该特征进行归一化处理。该步骤的数学表达式如下:

$$\hat{z}^{n-1} = F_{W-MSA} \left(F_{LN} \left(z^{n-1} \right) \right) + z^{n-1}, \quad (1)$$

$$z^n = F_{MLP} \left(F_{LN} \left(\hat{z}^{n-1} \right) \right) + \hat{z}^{n-1}. \quad (2)$$

式 (1) (2) 中: \hat{z}^{n-1} 为第 $n-1$ 个块中 W-MSA 模块的输出特征; z^n 为第 n 个块中 MLP 模块的输出特征。

在第一个模块的基础上,特征 z^n 通过 SW-MSA

模块进行自注意力计算; 同样, 将 z^n 应用到 MLP 的 F_{MLP} , 并经过 LN 层 F_{LN} 归一化, 该步骤的数学表达式如下:

$$z^n = F_{SW-MSA} \left(F_{LN} \left(z^n \right) \right) + z^n, \quad (3)$$

$$z^{n+1} = F_{MLP} \left(F_{LN} \left(z^n \right) \right) + z^n. \quad (4)$$

式 (3) (4) 中: z^n 为第 n 个块中 SW-MSA 模块的输出特征; z^{n+1} 为第 $n+1$ 个块中 MLP 模块的输出特征。

计算自注意力时, Swin Transformer 为每个头引入相对的位置偏置, 以优化相似性计算, 其表达式为

$$Attention(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax} \left(\mathbf{Q}\mathbf{K}^T / \sqrt{d} + \mathbf{B} \right) \mathbf{V}. \quad (5)$$

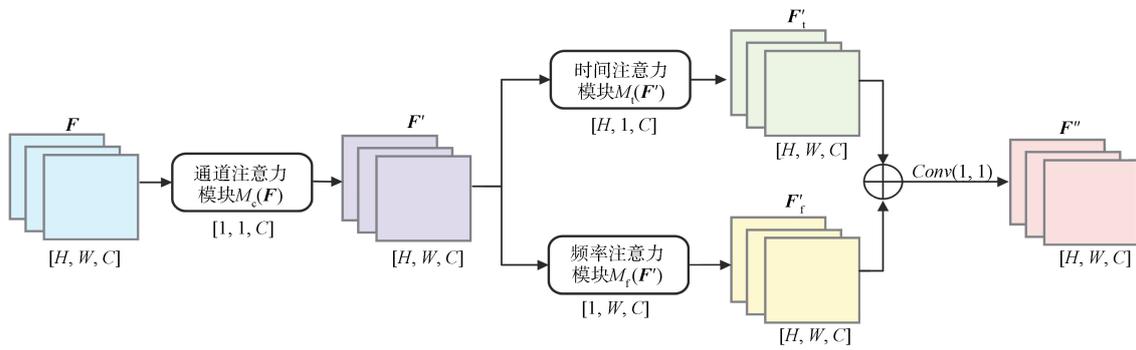


图4 时频注意力模块

Fig. 4 Time-frequency attention modules

图4中, F 为特征图像, 向量 F' 经过通道注意力模块, 向量 F'_f 和向量 F'_t 经过频率注意力模块和时间注意力模块, 在通道轴上连接向量 F'_f 和 F'_t 经过卷积操作得到最终的输出特征向量 F'' , 故特征图 F 经过时频注意力模块的表达式如下:

$$F' = M_c(F) \otimes F, \quad (6)$$

$$F'_f = M_f(F') \otimes F', \quad (7)$$

$$F'_t = M_t(F') \otimes F', \quad (8)$$

$$F'' = f_{1 \times 1} \left(\left[F'_f; F'_t \right] \right). \quad (9)$$

式 (6) ~ (9) 中: \otimes 为逐元素相乘; $f_{1 \times 1}$ 为括号内的特征向量与大小为 1×1 的卷积核进行卷积操作。

具体来说, 在通道注意力模块 CAM 中, 输入特征 F 经过全局平均池化操作 F_{Avg} 和最大池化操作 F_{Max} , 然后通过具有共享权重 MLP 的 F_{MLP} , 生成具有上下文描述的特征。接着对两个输出特征逐元素相加, 并通过 Sigmoid 激活函数计算最终的通道注意力权重 $M_c(F)$, 其表达式为

$$M_c(F) = \sigma \left(F_{MLP} \left(F_{Avg}(F) \right) + F_{MLP} \left(F_{Max}(F) \right) \right). \quad (10)$$

式中 σ 为 Sigmoid 激活函数。

式中: \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} 分别为查询矩阵、键矩阵和值矩阵; d 为输入通道; \mathbf{B} 为偏置矩阵。

2.3 时频注意力模块

在 Swin Transformer 模型中引入时频注意力模块, 其结构如图4所示, 它结合了通道注意力模块 (channel attention module, CAM)、时间注意力模块 (time attention module, TAM) 和频率注意力模块 (frequency attention module, FAM), 可以从通道、时间和频率域3个维度提取有用的特征, 突出时频图像中的关键区域和特征通道, 因此非常适用于调制识别任务。

TAM 和 FAM 的结构类似。TAM 沿着 x 轴对特征向量 F' 进行平均池化 F_{Avg} , 专注于输入时频图像的时间轴, 然后, 经过平均池化 F_{Avg} 和最大池化操作 F_{Max} , 进一步提取时间特征, 之后将特征图输入3个串联的 3×3 卷积层和 Sigmoid 激活函数中, 生成时间注意力权重 $M_t(F')$ 。FAM 沿着 y 轴对特征向量 F' 进行平均池化, 专注于输入时频图像的频率轴, 然后并行执行与 TAM 相同的操作, 得到频率注意力权重 $M_f(F')$, 其表达式为

$$M_t(F') = \sigma \left(f_{3 \times 3}^3 \left[F_{Avg}(F', x); F_{Max}(F', x) \right] \right), \quad (11)$$

$$M_f(F') = \sigma \left(f_{3 \times 3}^3 \left[F_{Avg}(F', y); F_{Max}(F', y) \right] \right). \quad (12)$$

式 (11) (12) 中: $f_{3 \times 3}^3$ 为括号内的特征向量与大小为 3×3 的卷积核进行卷积操作; (F', x) 为沿 x 轴对 F' 进行平均池化; (F', y) 为沿 y 轴进行平均池化。

3 实验设置与结果分析

3.1 数据集与预处理

本研究中用于模型训练和测试的数据集是 RadioML2016.10b^[11-12]。它由名为 GNU Radio 的无线电软件定义无线电工具包生成。该数据集中考虑了采

样率偏移、中心频率偏移、相位差异、多径衰落和加性白高斯噪声,其中包含8种数字调制方式:QPSK (quadrature phase shift keying)、16QAM (quadrature amplitude modulation)、64QAM、BPSK (binary phase shift keying)、8PSK (8 phase shift keying)、GFSK (Gauss frequency shift keying)、PAM4 (4-level pulse amplitude modulation)、CPFSK (continuous phase frequency shift keying),以及2种模拟调制方式:WBFM (wide band frequency modulation)、AM-DSB,每个无线电信号样本由128个连续波调制的正交相位(I/Q)信号组成,共2个通道,每个样本对应的标签包括调制类型和信噪比。数据集包含20种不同的信噪比,范围为-20~18 dB,步长为2 dB。RadioML2016.10b数据集的详细参数如表1所示。

表1 数据集参数
Table 1 Dataset parameters

参数名称	取值或类别
数据维度	2×128
采样频率 / MHz	1
采样率偏移 / Hz	50
中心频率 / kHz	200
最大载波频率偏移 / Hz	500
噪声	加性高斯白噪声

通信信号的I/Q数据经过STFT预处理,使用长度为40的汉明窗作为窗函数,对功率进行归一化处理,并通过高斯滤波器对信号的能量谱进行降噪处理,生成时频谱图,并将其保存为RGB图像,通过对时频图像随机水平翻转进行数据增强,扩展模型的泛化能力,防止过拟合并提高分类性能。

3.2 模型训练

为确保实验条件的一致性,本文提出的算法与对比算法将在相同的学习策略和训练环境下进行比较。使用Python语言开发,环境为Pycharm,通过Pytorch深度学习框架构建网络。训练和测试的服务器配置为Intel(R) Core(TM) i5-12490F, GPU为NVIDIA GeForce GTX 2060 Super。使用自适应学习率Adam优化器进行优化,学习率设置为0.0001,学习率衰减为0.0005,批次大小为64,采用交叉熵损失函数计算模型预测值与真实值的差距,具体识别步骤如下。

步骤一 使用STFT获得原始信号的时频图像矩阵,并通过高斯滤波器和归一化等预处理操作,生成 224×224 的图像数据集;

步骤二 将数据集按7:3的比例分为训练集和测试集;

步骤三 搭建基于时频注意力模块TF-Swin Transformer模型,利用迁移学习的方法,加载在ImagNet上预训练的Swin Transformer模型参数;

步骤四 将训练集输入TF-Swin Transformer网络进行训练至收敛,并保存模型;

步骤五 利用训练好的TF-Swin Transformer模型对测试集进行识别。

3.3 不同Transformer架构模型对比

为了验证上文所提出框架的有效性,对Vision Transformer (ViT)网络、Swin Transformer (Swin-T)以及本文中的框架进行模型的训练和测试。本文的TF-Swin Transformer框架采用基于预训练的Swin-T模型的迁移学习方法进行训练,以确保模型能够快速收敛;同时,使用时频注意力模块,以提高模型的特征提取能力。本文方法与传统Transformer的准确率性能比较曲线见图5。

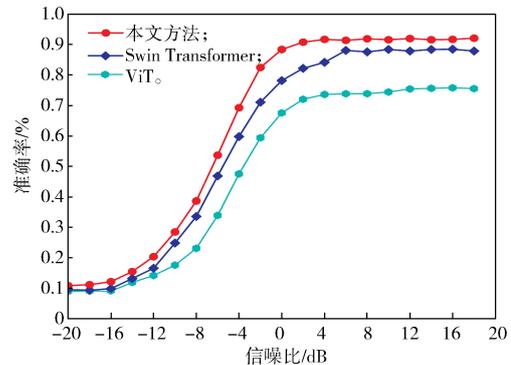


图5 准确率性能比较曲线

Fig. 5 Comparison curves of accuracy performance

如图5所示,随着信噪比增加,信号特征变得更加明显,调制信号的识别率也随之提高,本文提出的TF-Swin Transformer网络在调制识别方面的性能优于没有注意力机制模块的网络性能,总体的平均识别率提高了2.14%,尤其是在信噪比范围为[-6, 18] dB时,能够更有效地提取区分性特征,这表明时频注意力模块能够显著提高提取特征的效果,从而提高调制识别的准确性。同时,所提出框架的识别精度明显高于ViT的识别精度,并且在信噪比为16 dB时,本文所提方法的识别精度比ViT网络的识别精度高出了16.6%,这表明Swin-T网络性能优于ViT网络性能。这是由于Swin-T网络通过引入局部注意力机制、分层结构和移位窗口机制,提高了计算效率和特征提取能力。

3.4 TF-Swin Transformer性能分析

为了更直观地分析本文模型对不同调制方式识别的准确性,图6展示了在信噪比为0 dB和16 dB条件下的混淆矩阵。

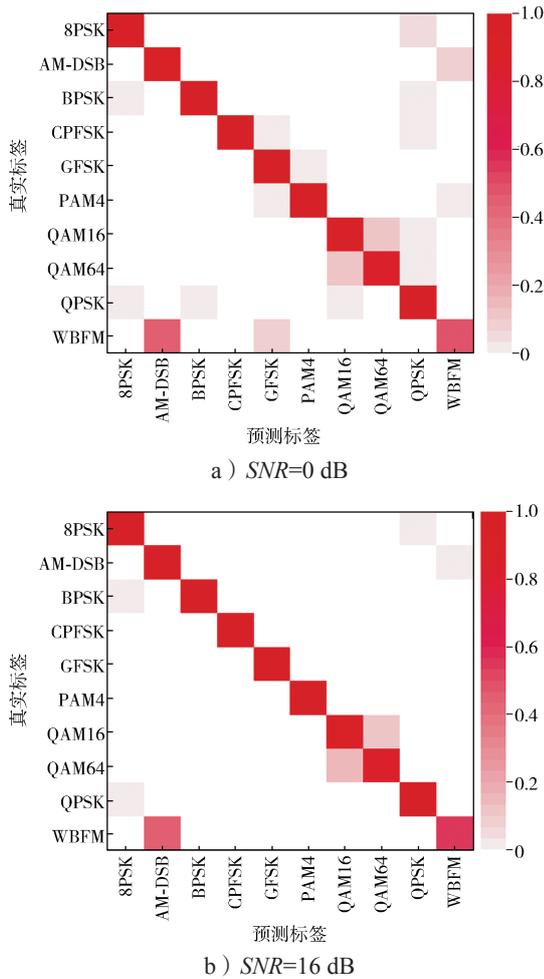


图6 不同信噪比下的混淆矩阵

Fig. 6 Confusion matrices under different SNR

由混淆矩阵可以看出, 在信噪比为 0 dB 时, 模型对 8PSK、BPSK、CPFSK、GFSK、PAM4、QPSK 共 6 种信号的识别准确率均超过 95%, 对 WBFM 的识别准确率最低, 为 48%, 而对 AM-DSB 信号的识别准确率为 92%, 同时对于 QAM16 和 QAM64 调制格式的数据容易混淆, 主要原因是 QAM 信号的星座点有重叠且具有相似的时频谱图。

在信噪比为 16 dB 时, 错误主要集中在模拟调制模式, 原因是 AM-DSB 和 WBFM 信号都属于连续波调制, 它们间的特征区别较小。此外, 由于数据集中信号的观测窗口较小, 导致调制信息过度丢失, 这一现象在文献 [13] 中也有提到, 但本文所提出模型的整体识别能力表现优异, 且具有较强的泛化能力。

3.5 本文方法与其他方法对比

为了充分验证本文所提出来的网络框架在调制识别领域的可行性, 将其与其他 4 种现有网络模型架构, 包括 CNN4 (convolutional neural network 4-layer)、CLDNN (convolutional long short-term memory deep neural network)、ResNet (residual

network) 及 SCNN2 (spectrum convolutional neural Network 2), 进行了对比分析^[14-17]。图 7 展示了在 RadioML2016.10b 数据集上, 本文所提框架与其他方法的准确性比较结果。

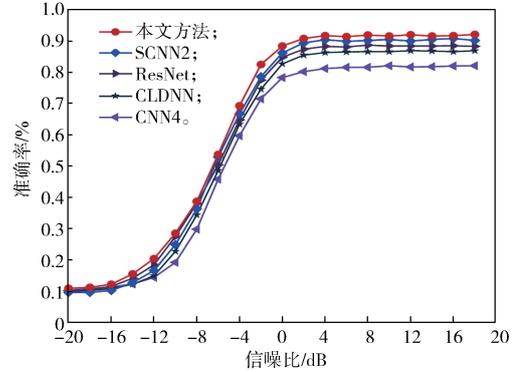


图7 不同方法的识别准确率比较曲线

Fig. 7 Comparison curves of recognition accuracy for different methods

图 7 所示结果表明, 当信噪比低于 -8 dB 时, 5 种网络模型的分类效果相近, 当信噪比高于 -12 dB 时, CNN 模型的性能明显低于其他模型的, 识别准确率最高为 82%。CLDNN 网络结合了 CNN、LSTM 和 DNN, 能够提取部分空间和时序特征, 但无法提取双向时序特征, 其最高准确率为 86.81%。本文所提方法的识别性能在 -8 dB 和 -4 dB 的信噪比条件下, 与 SCNN2 和 ResNet 相似, 而在信噪比范围为 0~16 dB 时, 本文所提方法的平均识别准确率分别比 SCNN2、ResNet、CLDNN、CNN4 高了 1.8%、3.6%、5.5%、10.8%。

这些结果表明, 具有时频注意力模块的 Swin Transformer 框架在调制识别领域能有效地进行调制识别, 同时由于输入数据的规模较大, 因此提出的方法在推理时间上也高于现有方法的。

4 结语

本文提出了一种 TF-Swin Transformer 调制识别模型, 通过将调制信号转换为时频图像, 并利用迁移学习预训练模型进行调制识别和分类。实验结果表明, 时频注意力模块充分利用了信号的时间和频率信息, 使二者相互补充, 从而扩大了不同调制类型之间的差异, 显著提高了识别准确率。与传统 CNN 模型相比, 在高信噪比下, 总体识别准确率提高约 10.8%, 验证了所提出的方法可以有效地提高调制识别的识别准确率, 且具备较好的泛化性能。但目前的研究仍具有一定的局限性: 仅使用时频图作为特征, 导致特征较为单一。未来的工作中, 将考虑整合更多

特征, 探讨不同特征对各种调制格式识别的影响。

参考文献:

- [1] WANG Y, LIU M, YANG J, et al. Data-Driven Deep Learning for Automatic Modulation Recognition in Cognitive Radios[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(4): 4074–4077.
- [2] 赵 泽, 尚鹏飞, 陈海明, 等. 无线传感器网络干扰分类识别机制的研究 [J]. *通信学报*, 2013, 34(10): 28–36.
ZHAO Ze, SHANG Pengfei, CHEN Haiming, et al. Interference Identification and Classification Mechanism for Wireless Sensor Network[J]. *Journal on Communications*, 2013, 34(10): 28–36.
- [3] AFAN A L, FAN Y Y. Automatic Modulation Classification Using Deep Learning Based on Sparse Autoencoders with Nonnegativity Constraints[J]. *IEEE Signal Processing Letters*, 2017, 24(11): 1626–1630.
- [4] NAMDEV D, BANSAL A. Frequency Domain Analysis for Audio Data Forgery Detection[C]//2015 Fifth International Conference on Communication Systems and Network Technologies. Gwalior: IEEE, 2015: 702–705.
- [5] ZENG Y, ZHANG M, HAN F, et al. Spectrum Analysis and Convolutional Neural Network for Automatic Modulation Recognition[J]. *IEEE Wireless Communications Letters*, 2019, 8(3): 929–932.
- [6] HASSAN A N, MOHASSEL FEGHHI M, ESMAEILI V. A Fast Automatic Modulation Classification Based on STFT Using Hybrid Deep Neural Network[J]. *Journal of Communication Engineering*, 2021, 10(2): 328–343.
- [7] LIU X Y. Automatic Modulation Classification Based on Improved R-Transformer[C]//2021 International Wireless Communications and Mobile Computing (IWCMC). Harbin: IEEE, 2021: 1–8.
- [8] LI Z X, SUN X H, ZHAO M Y, et al. Swin Transformer-Based Time-Frequency Image Classification of Frequency Hopping Modulated Signals[C]//2023 3rd International Symposium on Computer Technology and Information Science (ISCTIS). Chengdu: IEEE, 2023: 858–862.
- [9] WAN C X, ZHANG Q H. A Novel Dual-Component Radar-Signal Modulation Recognition Method Based on CNN-ST[J]. *Applied Sciences*, 2024, 14(13): 5499.
- [10] REN B, TEH K C, AN H Y, et al. Automatic Modulation Recognition of Dual-Component Radar Signals Using ResSwinT-SwinT Network[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2023, 59(5): 6405–6418.
- [11] HUANG J H, FANG Y Y, WU Y Z, et al. Swin Transformer for Fast MRI[J]. *Neurocomputing*, 2022, 493: 281–304.
- [12] O’ SHEA T J, CORGAN J, CLANCY T C. Unsupervised Representation Learning of Structured Radio Communication Signals[C]//2016 First International Workshop on Sensing, Processing and Learning for Intelligent Machines (SPLINE). Aalborg: IEEE, 2016: 1–5.
- [13] O’ SHEA T J, CORGAN J, CLANCY T C. Convolutional Radio Modulation Recognition Networks[C]//Engineering Applications of Neural Networks: 17th International Conference. Aberdeen: Springer International Publishing, 2016: 213–226.
- [14] RAMJEE S, JU S, YANG D, et al. Fast Deep Learning for Automatic Modulation Classification[J/OL]. *IEEE Transactions on Cognitive Communications and Networking*. (2019–01–16). <https://arxiv.org/abs/1901.05850>.
- [15] WEST N E, O’ SHEA T. Deep Architectures for Modulation Recognition[C]//2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN). Piscataway: IEEE, 2017: 1–6.
- [16] 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.
- [17] GU S, SU J, LIU C, et al. Time-Frequency Fusion Deep Learning Algorithm for Modulation Recognition[C]//2025 10th International Conference on Computer and Communication System (ICCCS). [S. l.]: IEEE, 2025: 470–474.

(责任编辑: 廖友媛)