

基于自注意力的双流结构 门控循环单元的 自动调制识别

张依宁, 张志超

南京信息工程大学数学与统计学院, 江苏 南京

收稿日期: 2023年4月22日; 录用日期: 2023年6月18日; 发布日期: 2023年6月21日

摘要

自动调制识别是各种无线通信场景中的一项重要任务。受通信环境日益复杂的影响, 如何有效、高效地提取无线电信号的时间特征是当前亟待解决的问题。为了解决这个问题, 本文提出了一种双流结构的基于自注意力的门控循环单元(Gated Recurrent Unit, GRU)模型。本文的方法首先使用卷积层提取浅层空间特征, 然后将特征分为两部分, 在两个并行流中使用堆叠的GRU来充分提取时间特征。自注意力的加入使网络能够关注输入序列的不同部分, 从而提高特征提取的能力。本文提出的模型在基准数据集上对信噪比(Signal-to-Noise Ratio, SNR)大于等于-6 dB的信号实现了最高的识别准确率。特别地, 当SNR为-6 dB时, 识别准确率达到58.9%, 比其他的模型提高了5%以上。此外, 本文提出的模型具有很强的适应性, 即使在较小的数据集上训练也能达到最高的准确率。本文还研究了初始学习率对模型性能的影响, 给出了模型达到最高准确率和最高效率的初始学习率。

关键词

自动调制识别, 双流架构, 门控循环单元, 自注意力

Automatic Modulation Recognition Using Dual-Stream Gated Recurrent Unit Based on Self-Attention

Yining Zhang, Zhichao Zhang

School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing
Jiangsu

Received: Apr. 22nd, 2023; accepted: Jun. 18th, 2023; published: Jun. 21st, 2023

文章引用: 张依宁, 张志超. 基于自注意力的双流结构门控循环单元的自动调制识别[J]. 建模与仿真, 2023, 12(4): 3286-3298. DOI: 10.12677/mos.2023.124302

Abstract

Automatic modulation recognition is an important task in various wireless communication scenarios. Due to the increasingly complex communication environment, how to extract the temporal characteristics of radio signals effectively and efficiently is an urgent problem to be solved. To solve this problem, this paper proposes a dual-stream gated recurrent units (GRUs) model based on self-attention. The approach first uses the convolutional layers to extract the shallow spatial features, then splits the features into two parts and uses stacked GRUs in two parallel streams to fully extract the temporal features. The addition of self-attention enables the network to focus on different parts of the input sequences, thus improving the ability of feature extraction. Our model achieves the highest recognition accuracy for signals with a signal-to-noise ratio (SNR) greater than or equal to -6 dB on the baseline dataset. In particular, when the SNR is -6 dB, the recognition accuracy reaches 58.9%, which is more than 5% higher than other models. In addition, our model has strong adaptability and can achieve the highest accuracy even when trained on a small dataset. The paper also studies the impact of the initial learning rate on the model performance and gives the initial learning rate for the model to achieve the highest accuracy and efficiency respectively.

Keywords

Automatic Modulation Recognition, Dual-Stream Architecture, Gated Recurrent Unit, Self-Attention

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 简介

自动调制识别是接收机自动识别接收到的无线电信号的调制方案的技术。调制识别是信号解调和信息提取前的关键步骤，它在频谱传感、干扰识别和信息战等领域有着重要的应用。在无线电技术飞速发展的今天，无线通信环境越来越复杂，迫切需要一种即使在恶劣环境下也能达到较高准确率的自动调制识别方法。

传统的自动调制识别方法可分为两类：基于似然的方法[1] [2] [3]和基于特征的方法[4] [5] [6]。基于似然的方法在贝叶斯估计的意义上是最优的，但严重依赖于先验知识，且计算量过大，很难进行实时预测。基于特征的方法包括两个部分：特征提取和分类器设计。无线电信号常用的特征有小波特征、高阶累积量特征、星座图特征等。机器学习中经典的分类算法，如人工神经网络[7] [8] [9]、决策树[10]和支持向量机[11] [12] [13]是基于特征的方法中经常使用的分类器。基于特征的方法的优点是不需要先验知识，计算复杂度较低。然而，基于特征的方法需要手动提取信号特征，费时且需要专业知识。此外，基于特征的方法的识别准确率严重依赖于所选择的信号特征的有效性，因而难以适应不同的环境。

近年来，深度学习在各个领域都取得了突破。研究人员将深度学习用于自动调制识别任务，取得了比传统基于特征的方法更好的性能。在[14]中，引入卷积神经网络(Convolutional Neural Network, CNN)对无线电信号的调制方案进行分类，模型在低信噪比下表现出更好的性能。很快更多的基于卷积神经网络或循环神经网络(Recurrent Neural Network, RNN)的深度学习模型被提出。使用CNN的方法有的使用一

维 CNN 直接处理序列数据, 有的使用某些信号处理的方法将信号转化为二维图片输入 CNN。在[15]中, 对振幅序列和相位序列分别基于一维 CNN 采用并行融合的方法进行训练。[16]使用平滑伪 Wigner-Ville 和 Born-Jordan 分布来获取 CNN 的图像。[17]使用短时傅里叶变换(Short-Time Fourier Transform, STFT)创建了不同信号谱图的图像数据集。长短期记忆(Long Short-Term Memory, LSTM)作为一种优秀的 RNN 模型有强大的提取序列数据时间特征的能力, [18]提出了一种基于 LSTM 的自动调制识别模型, 该模型从调制方案的幅度和相位信息中学习时间特征。还有一些研究者发现, 使用混合模型, 比如结合 CNN 和 LSTM, 可以提高模型的准确性和鲁棒性。在[19]中提出的模型将 CNN 与 LSTM 结合, 在高信噪比下获得了更高的精度。在[20]中, 提出了一个多通道的深度学习框架, 结合一维、二维卷积和 LSTM, 从时间和空间角度提取信号特征。[21]提出了一种高效的深度学习模型, 基于相位参数估计和变换, 结合 CNN 和 GRU, 该模型在保持较高精度的同时减少了需要计算的参数量。注意力机制于 2014 年提出[22], 最早应用于机器翻译, 而后在多个领域最先进的模型中被采用。为了提高自动调制识别的准确性, [23]提出了一种包含四个自适应注意力机制模块的自动调制识别框架。[24]提出了一种基于 CNN 的时频注意力机制, 取得了较高的识别精度。

通过研究和分析这些论文可以发现, 将 CNN 和 RNN 结合起来, 从原始信号中提取特征进行调制识别是一个优秀而有潜力的想法。然而, 如何有效地从无线电信号中提取时间特征, 这仍然是模型设计过程中的一个问题。过去的模型往往采用单通道进行时间特征的提取[19] [20] [21], 而无线电信号本身有 I 和 Q 两个分量, 在采用单通道进行时间特征提取的过程中可能会有部分特征信息被丢失。因此, 本文将卷积处理后的特征图分为两部分, 分别进行时间特征提取, 以减少特征丢失, 增强模型对信号时间特征的提取能力, 提高调制识别的准确率。所提出的模型被称为基于自注意力的双流 GRU 神经网络(Dual-Stream GRU Based on Self-Attention Neural Network, DSGANN)。具体来说, 首先使用卷积层来提取信号的浅层特征, 对最后一个卷积层采用 same 填充不改变输入输出形状, 然后建立双流结构的时间特征提取模块, 在每个时间特征提取流中采用堆叠的 GRU, 并创新性地每个流的两个 GRU 层之间增加自注意力, 更好地捕获输入特征序列之间的依赖关系, 使网络在进行预测时能够有效地权衡输入序列中不同部分的重要性。DSGANN 在一个开放数据集上的表现优于一些现有模型, 对信噪比大于等于-6dB 的信号具有更高的识别精度。并且在减少数据集样本量训练时, DSGANN 仍然取得了最高的精度, 这说明 DSGANN 设计良好, 不依赖于大量的信号样本来获得良好的性能。

本文的其余部分组织如下:第二部分详细描述了提出的模型的结构, 第三部分展示实验结果和讨论, 第四部分给出结论。

2. 方法

接收到的信号表示为:

$$y[l] = A[l]e^{j(\omega l + \varphi)}x[l] + n[l], l = 1, 2, \dots, L$$

式中, $y[l]$ 为接收信号, $A[l]$ 为信道增益, ω 为频率偏移, φ 为相位偏移, $x[l]$ 为发射信号, $n[l]$ 为噪声, L 为信号的采样点数, 接收到的信号可以以同相/正交(I/Q)的形式存储。

噪声会严重影响自动调制识别的精度。图 1 显示了不同调制方案的信号在不同信噪比下的波形。从图中可以看出, 高信噪比下清晰的信号波形在低信噪比下变得杂乱而难以区分。

DSGANN 的整体架构如图 2 所示。它可以分为三部分:卷积模块, 双流结构的自注意力 GRU 和全连接层。首先, 用两个通道的一维卷积层分别提取 I、Q 分量的特征, 再将两个通道融合后送到二维卷积层提取浅层特征。第二部分是带有自注意力的双流 GRU。为了充分提取信号的时间特征, 设计了双流架

构的 GRU，对每个时间特征提取流采用堆叠 GRU，并在两个 GRU 层之间增加自注意力层，以提高模型的性能。第三部分是全连接层。DSGANN 集成了 CNN、GRU 和注意力机制的优点，分离通道和融合通道的设计大大提高了网络的特征提取能力。

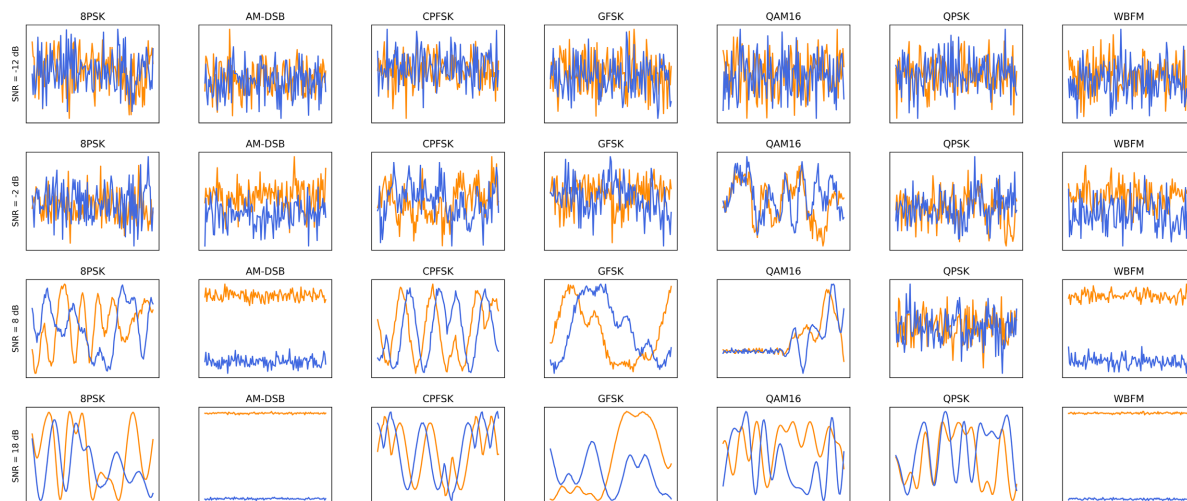


Figure 1. Radio signals with different modulation schemes at -12 , -2 , 8 , 18 dB SNR

图 1. 信噪比分别为 -12 , -2 , 8 , 18 dB 的不同调制方案的无线电信号

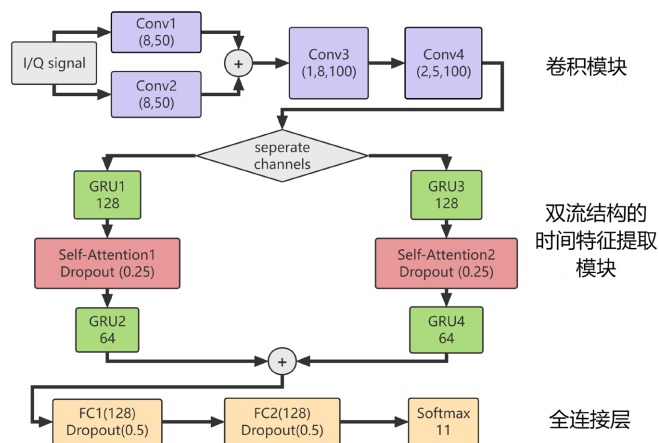


Figure 2. The architecture of DSGANN

图 2. DSGANN 的结构

2.1. 卷积模块

CNN 具有很强的二维图像特征提取能力，这使得它在物体识别和图像分类任务中表现优异。而对于一维信号，CNN 可以有效地提取信号序列的浅层特征。先前的工作表明，CNN 和 RNN 的结合在多数情况下比单独使用 CNN 或 RNN 表现更好。

信号的 I 分量和 Q 分量彼此正交，它们提供了关于调制信号的不同信息。DSGANN 分别取 I 分量和 Q 分量作为模型的输入来获取信号信息。对于每个通道，选择卷积核大小为 8 的一维卷积层，采用因果填充，得到 50 个特征图。使用因果填充是为了保持输入信号的时序性。

对这两个通道的输出张量，都添加一个维度，并沿着新的维度将它们连接起来。将张量重塑成合适

的形状后, 将其送入两个二维卷积层(图 2 中的 Conv3 和 Conv4)进行进一步的特征提取。这两个卷积层的作用是融合单独的 I 分量和 Q 分量的特征, 捕获信号的整体表示特征。在二维卷积层中使用 same 填充来保持输出形状与输入形状相同。所有卷积层都采用“relu”激活函数。

2.2. 双流结构的 GRU

考虑到信号本身是作为时间序列数据输入模型, 对信号的时间特征的提取是至关重要的。受到信号本身由 I 分量和 Q 分量组成的启发, 产生了设计双流架构分别提取时间特征的想法。双流结构的输入是通过使用 Lambda 层拆分最后一个卷积层(图 2 中的 Conv4)的输出得到的。需要注意的是, 双流通道的输入是经过二维卷积计算后融合特征的不同部分, 并不是分别代表 I 分量和 Q 分量的特征, 但在两个通道中表示信号特征的想法是一样的。

GRU 和 LSTM 都是处理序列数据的优秀模型。这两个模型通过设置门来选择性地保留或遗忘前一时间段的信息, 从而解决了传统 RNN 中梯度消失的问题。GRU 有两个门: 一个更新门和一个重置门。这两个门是用来控制通过网络的信息流的, 决定保留多少先前的信息和有多少新输入与先前的信息混合。本文选择 GRU 而不是 LSTM, 因为 GRU 需要的参数比 LSTM 少, 计算速度更快, 并且在某些任务上可能比 GRU 表现出更好的性能。LSTM 对于需要处理复杂的长期依赖关系的任务可能更有效。在下一个实验章节中, 将通过实验比较 DSGANN 模型使用 GRU 和使用 LSTM 的性能差异。

堆叠 GRU 将多个 GRU 层组合在一起, 对序列数据中的复杂关系进行建模。每一层 GRU 学习不同深度层次的特征, 允许模型捕获数据中更复杂的模式。DSGANN 在时间特征提取模块的每个流堆叠了两个 GRU 层, 中间加入自注意力, 第一层 GRU 输出 128 个单元, 且返回输入序列中每个时间步的隐藏状态的完整序列。第二层 GRU 输出 64 个单元, 只返回最后一个时间步的输出。

2.3. 自注意力与全连接层

自注意力是神经网络在处理输入序列中的每个元素时关注输入序列的不同部分的一种机制。自注意力允许模型选择性地关注输入序列的不同部分, 有效地学习哪些部分对手头的任务更重要。此外, 自注意力层可以帮助解决在训练深度神经网络时可能出现的梯度消失问题。

DSGANN 模型在每个流的两个 GRU 层之间添加了自注意力, 以提高模型捕获时间信息的能力, 自注意力层的 num_heads 设为 6, key_dim 设为 32, 并随机舍弃 25% 的参数以防止过拟合。自注意力机制将输入序列分成多个头部, 每个头部关注输入序列的不同部分, 每个头部的注意权值在训练过程中学习, 从而允许模型捕捉输入特征之间的依赖关系。在每个流中的 GRU 层之间添加自注意力层, 可以使模型适应不同的输入序列和任务, 更好地捕获输入序列中的长期依赖关系。这是因为注意力机制使模型能够选择性地关注输入序列的不同部分, 同时考虑每个部分的相对重要性, 提高模型学习复杂模式的能力。

在完成时间特征的提取后, 两个并行 GRU 流的输出被连接, 并送到全连接层。对前两个全连接层使用“selu”激活函数, 通过 dropout 防止过拟合, 最后一个全连接层使用“softmax”激活函数, 给出最终预测的调制方案。

3. 实验

为了验证所提方法的有效性, 设计了三组实验。第一个实验比较了 DSGANN 与四个现有的先进模型的性能; 并且减少数据集的样本数, 在一个较小的数据集上比较 DSGANN 和其他模型的性能; 第二个实验是消融实验, 在原数据集和小数据集上比较了 DSGANN 和不同的模型变体的性能; 第三个实验研究了初始学习率对 DSGANN 性能的影响。

本文以测试集的认识精度作为评价指标。用于模型训练和评价的批次大小设置为 400。训练集、验证集和测试集按 3:1:1 的比例进行划分。所有实验均在 python 3.7.12 和 tensorflow 2.11.0 环境下进行, 使用 Nvidia Tesla P100 GPU 和 Intel(R) Xeon(R) CPU @ 2.20GHz。

3.1. 数据集简介

RadioML2016.10a 是 O'shea 等人使用 GNU Radio 生成的开放的无线电信号数据集[25]。该数据集一共有 22 万个无线电信号样本, 有包含 8PSK、BPSK、CPFSK、GFSK、PAM4、16QAM、64QAM、QPSK、AM-DSB、AM-SSB、WBFM 这 11 种调制方案的无线电信号, 信噪比范围为 -20 dB~ 18 dB。信号以 I/Q (同相/正交) 的形式存储, 信号长度为 128。由于该数据集在仿真信号时充分考虑了现实世界中可能遇到的恶劣环境因素, 包括中心频偏、采样率偏移、加性高斯白噪声、多径和衰落等对信号的影响, 该数据集里的信号与真实通信环境中的信号接近。

3.2. 在 RadioML2016.10a 上比较 DSGANN 和一些现有模型的性能

这个实验环节使用 RadioML2016.10a 数据集训练和评估了 DSGANN, 并复现了四个现有的优秀模型 (1DCNNPF [15]、LSTM [18]、MCLDNN [20]和 PET [21]) 进行比较。将所有模型的初始学习率设置为 0.001, 如果验证损失在 5 个训练周期没有减小, 则学习率减小为原来的 0.8 倍。如果在 30 个周期验证损失没有减小, 则终止训练过程。对其他模型的复现参考了综述论文[26]的工作。

图 3 显示了 DSGANN 和其他模型在 RadioML2016.10a 上的性能。可以看出, 当信噪比大于等于 -6 dB 时, DSGANN 具有更高的识别精度, 这充分验证了由于双通道时间特征提取流和自注意力等的设计, DSGANN 模型从原始信号中提取特征的能力比其他模型更强大, 因而具有更出色的调制识别性能。当信噪比为 -6 dB 时, DSGANN 的识别准确率达到 58.9%, 相比其他模型提高了 5% 以上。图 4 为信噪比为 -6 dB 时不同模型的混淆矩阵。如果把识别精度按照不同调制方案划分, 不同模型的每种调制方案的识别精度如图 5 所示。从图 5 可以看出, 当信噪比大于等于 2 dB 时, 只有 DSGANN 对除了 WBFM 以外的所有调制方案, 识别准确率都超过了 90%。

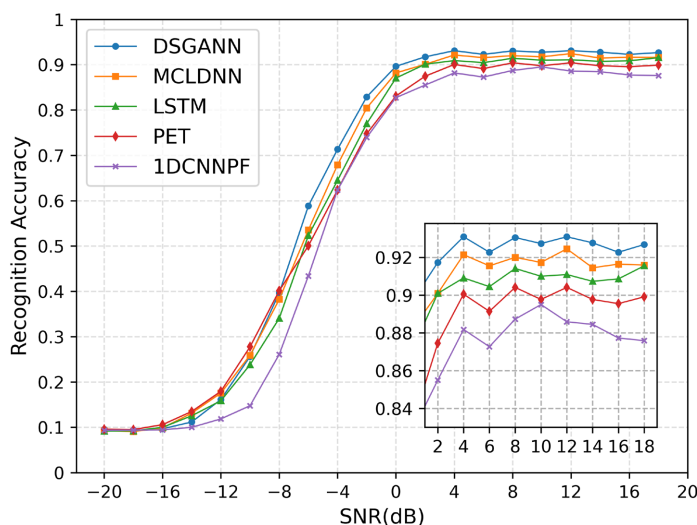


Figure 3. Recognition accuracy of DSGANN and other models on RadioML2016.10a

图 3. 在 RadioML2016.10a 数据集上 DSGANN 与其他模型的识别精度

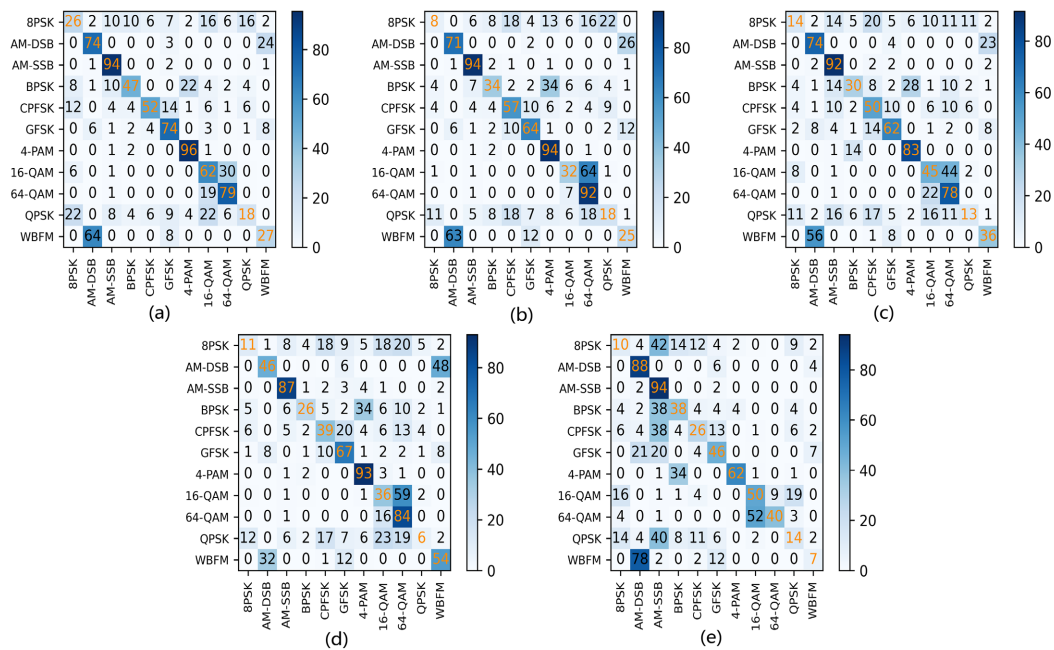
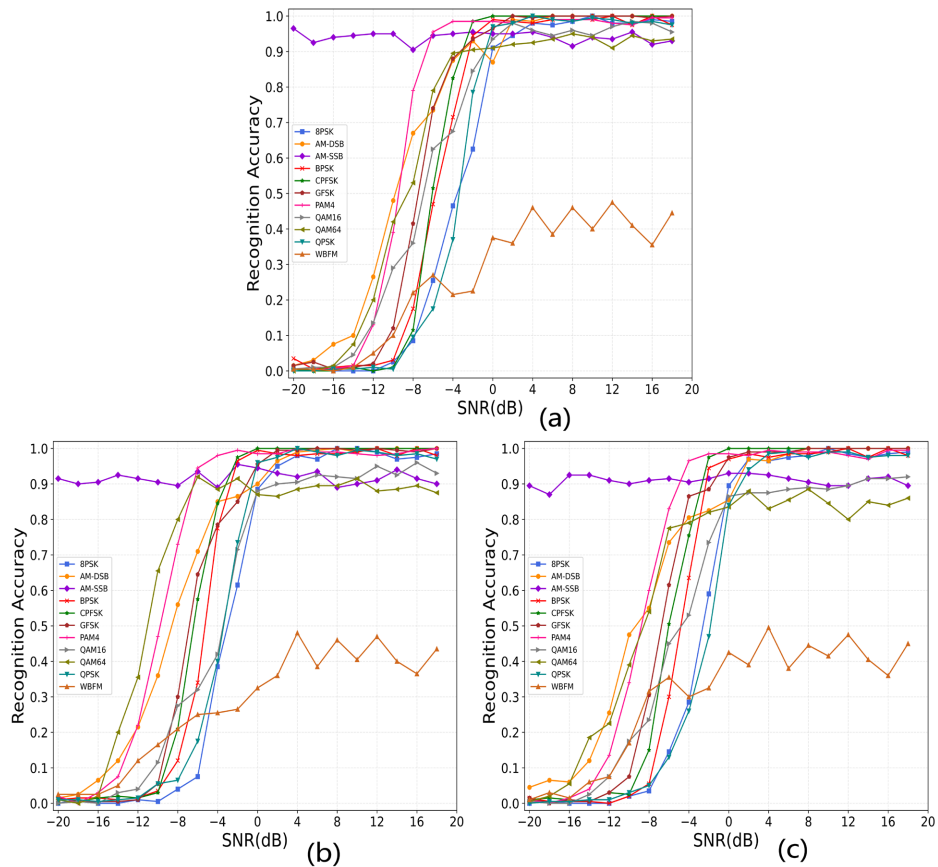


Figure 4. Confusion matrix of (a) DSGANN; (b) MCLDNN; (c) LSTM; (d) PET; (e) 1DCNNPF on RadiomL2016.10a at -6 dB SNR

图 4. 在 RadiomL2016.10a 上 SNR = -6 dB 时 (a) DSGANN; (b) MCLDNN; (c) LSTM; (d) PET; (e) 1DCNNPF 的混淆矩阵



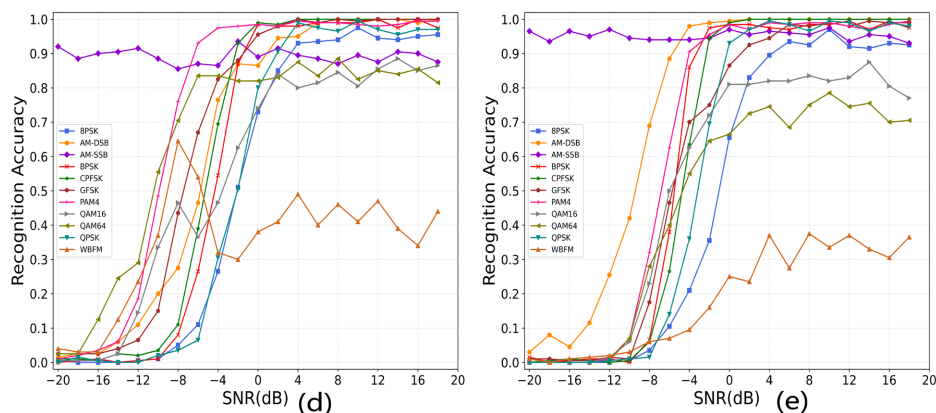


Figure 5. Recognition accuracy of 11 modulation schemes using (a) DSGANN; (b) MCLDNN; (c) LSTM; (d) PET; (e) 1DCNNPF on RadioML2016.10a

图 5. 在 RadioML2016.10a 数据集上使用 (a) DSGANN; (b) MCLDNN; (c) LSTM; (d) PET; (e) 1DCNNPF 的 11 种调制方案的识别精度

3.3. 在较小的数据集上的模型性能

考虑到在实际应用场景下可能没有足够多的信号样本用于模型的训练, 使用一个较小的数据集训练 DSGANN 和其他模型。这个较小的数据集基于 RadioML2016.10a, 但在每种调制方案和信噪比下只有 60% 的信号样本被随机采样用于训练。实验结果如图 6 所示。可以很容易地发现, 样本量的减少并没有太多地影响 DSGANN 的优异性能。DSGANN 的平均识别准确率下降仅不到 1%, 而用于训练的样本量则大大减少。与其他模型相比, DSGANN 在信噪比大于等于 -8 dB 时有更高的识别精度, 在更低的信噪比下则和其他模型性能接近。DSGANN 在 0 dB 时的识别准确率达到到了 88.8%, 远远高于其他模型。图 7 显示了在较小的数据集上, 当信噪比为 0 dB 时, DSGANN 与其他模型的混淆矩阵。

3.4. 消融实验

本小节设计了消融实验来研究模型中不同部分对模型性能的影响, 基于原模型 DSGANN 建立了五个变体后的模型。它们被称为 DSGANN-A (输入不是分开的 I/Q 通道, 卷积层特征图的参数保持不变),

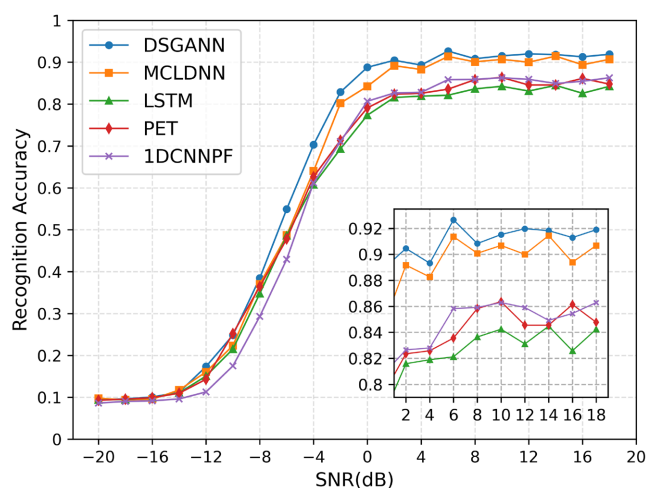


Figure 6. Recognition accuracy of DSGANN and other models on the smaller dataset

图 6. 在小数据集上 DSGANN 与其他模型的识别精度

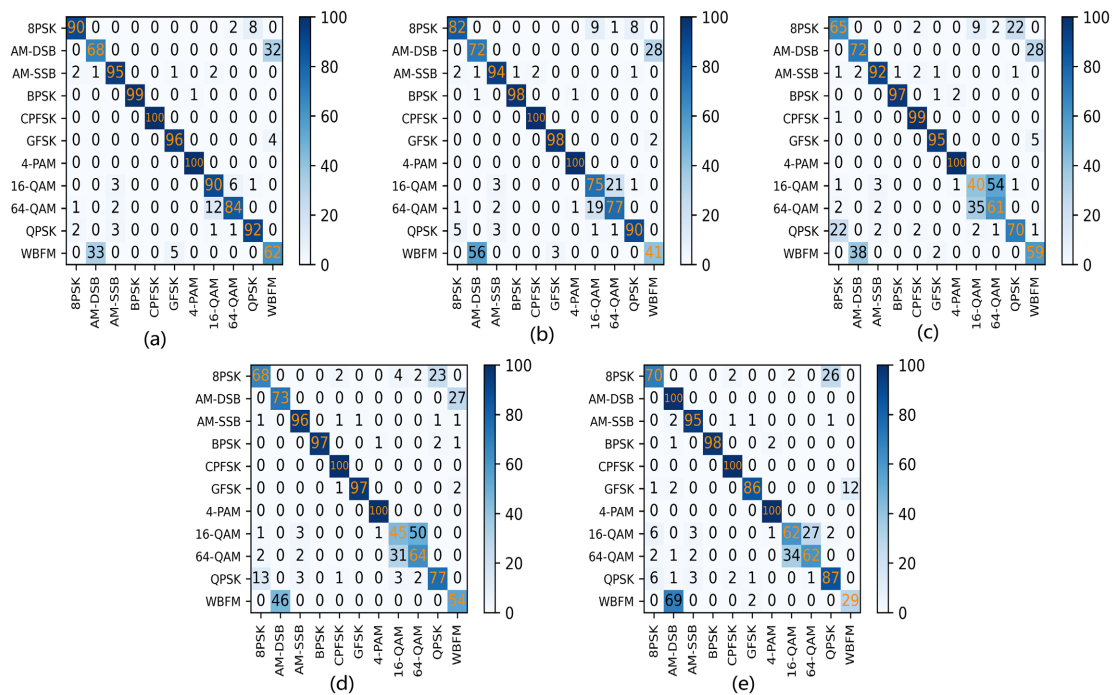


Figure 7. Confusion matrix of (a) DSGANN; (b) MCLDNN; (c) LSTM; (d) PET; (e) 1DCNNPF on the smaller dataset at 0 dB SNR

图 7. 在小数据集上 SNR = 0 dB 时(a) DSGANN; (b) MCLDNN; (c) LSTM; (d) PET; (e) 1DCNNPF 的混淆矩阵

DSGANN-B (最后一个卷积层使用 valid 填充, 采用单通道的时间特征提取而不是双流结构的时间特征提取), DSGANN-C (去除了自注意力)、DSGANN-D (用 LSTM 代替 GRU)、DSGANN-E (去除一个全连接层)。分别在 RadioML2016.10a 数据集和 0.6 倍采样后的小样本数据集进行实验, 结果如图 8 和表 1 所示。从图 8 和表 1 中可以看出, DSGANN 比它的所有变体后的模型都表现得更好, 特别是在更高的信噪比下, 这表明 DSGANN 的每个部分的选取都是合理有效的。如果用 LSTM 代替原网络中的 GRU, 在两个不同大小的数据集上的识别精度都会略有下降。此外, LSTM 需要更大的计算量和更大的内存空间, 因此 GRU 是提取时间特征的更好选择。当去除两个 GRU 之间的自注意力时, 性能下降最大。这说明了添加自注意力的重要性, 它可以显著提高网络提取特征的能力。在去除了自注意力之后, 模型在小样本数据集的平均准确率甚至下降到 57.29%, 而原模型在小样本数据集的平均准确率为 61.98%, 这充分说明自注意力通过关注输入序列不同部分之间的依赖关系优化特征提取能力, 可以较好地弥补样本量不足的缺陷。DSGANN-B 的识别精度相较原模型也有较大的下降, 尤其是在小样本数据集上, 说明 DSGANN 在时间特征提取模块的双流结构的设计是合理并且有效的, 通过采用双流结构的 GRU 充分提取信号的时间特征, 在一定程度上可以弥补信号样本不足的缺陷。正是这些模块结构的组合产生了一个优秀的模型, 在较小的数据集上也表现良好。

3.5. 初始学习率对模型性能的影响

初始学习率是影响模型性能的一个重要参数, 过低或过高的初始学习率都不利于模型的表现。为了研究初始学习率对 DSGANN 性能的影响, 设计了不同初始学习率的对比实验。实验结果如图 9 所示。

首先将初始学习率设置为 0.001, 在测试集的平均准确率为 62.88%。为了研究初始学习率对模型性能的影响, 将初始学习率分别设置为 0.0015、0.002 和 0.0025。在这些初始学习率下, 训练所需的时间减

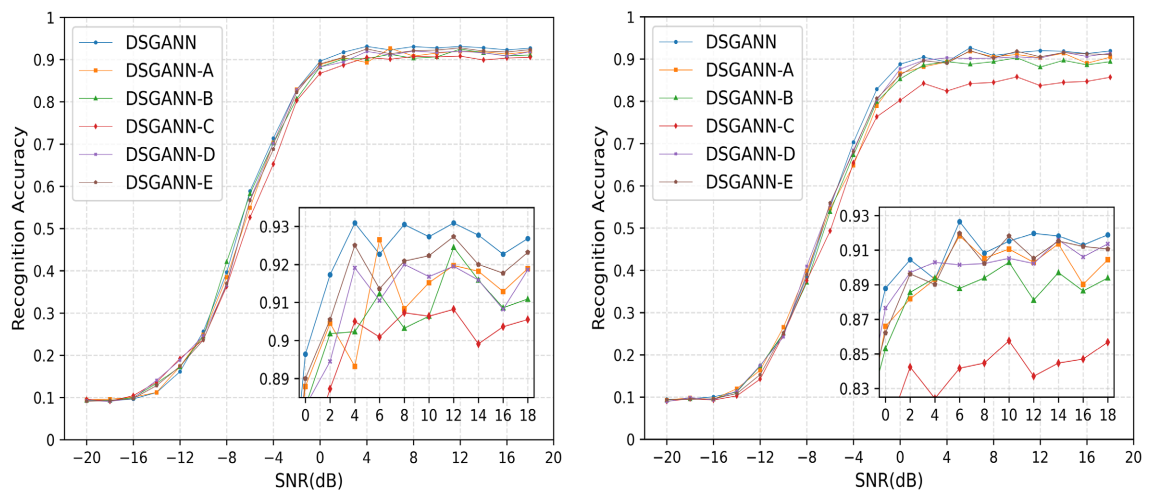


Figure 8. Recognition accuracy of DSGANN and its variants on (a) RadioML2016.10a and (b) the smaller dataset
图 8. DSGANN 及其变体在数据集(a)RadioML2016.10a 和(b)小数据集上的识别精度

Table 1. Average recognition accuracy of DSGANN and its variants on (a) RadioML2016.10a and (b) the smaller dataset
表 1. DSGANN 及其变体在数据集(a) RadioML2016.10a 和(b) 小数据集上的平均识别精度(%)

数据集	DSGANN	DSGANN-A	DSGANN-B	DSGANN-C	DSGANN-D
a	62.88	62.01	62.10	60.97	62.16
b	61.98	61.01	60.39	57.29	61.33

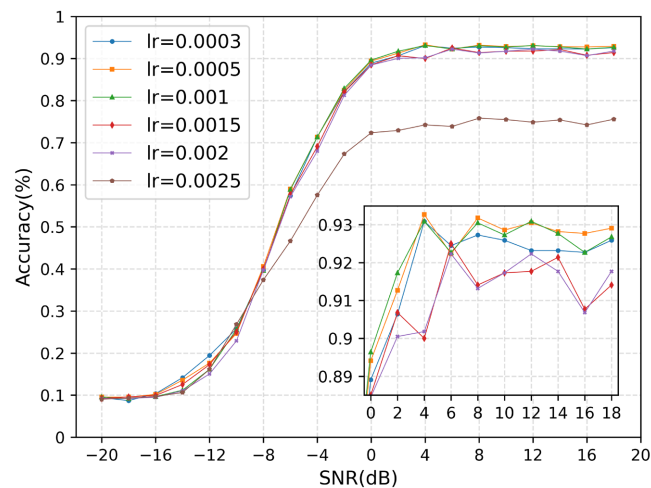


Figure 9. Recognition accuracy of DSGANN under different initial learning rates

图 9. DSGANN 在不同初始学习率下的识别准确率

少, 但识别准确率下降了。因此, 推测初始学习率更低的话, 识别准确率可能更高。于是将初始学习率设置为 0.0005, 尽管训练时间增加了, 平均识别准确率提高到了 63.12%。这是由于在较大的初始学习率下, 模型参数的更新可能过大, 导致优化算法超过最优值, 从而降低了准确性, 而适当减小初始学习率解决了这一问题。然而, 当进一步将学习率降低到 0.0003 时, 训练时间增加了, 但准确率并没有提升。这表明初始学习率太小, 训练模型的更新太小, 无法在识别准确率上取得进一步提升, 因此 0.0003 并不

是更适合 DSGANN 的初始学习率。这表明, 从平均识别准确率的角度来看, 0.0005 可能是 DSGANN 最合适的初始学习率。

虽然初始学习率为 0.0005 时的整体的准确率要高于初始学习率为 0.001 时的准确率, 但是训练模型所需的时间增加了大约 24 分钟, 这在实际应用中可能并不合适。为了综合衡量模型性能的改进和训练模型所需要的时间, 定义了一个新的度量标准, 称为“IE”(Improving Efficiency)。它是平均识别准确率(%)乘 100 减去基准识别准确率(设置为 60%)乘 100 再除以训练时间(以小时为单位)。

通过计算发现, 当初始学习率为 0.001 时, IE 是最高的。因此, 可以得出结论, 如果只追求更高的准确率, 应该将初始学习率设置为 0.0005。但是, 如果在要求达到较高精度的同时需要最小化训练时间, 那么 0.001 是最佳的选择。表 2 给出了 DSGANN 在不同初始学习率下的实验结果, 包括测试集的损失、测试集的平均准确率、训练周期、训练时间和 IE 指标。

Table 2. The recognition accuracy and computational complexity of DSGANN under different initial learning rates
表 2. DSGANN 在不同初始学习率下的识别精度和计算复杂度

初始学习率	测试集损失	平均准确率(%)	训练周期	训练时间(h)	IE
0.0003	1.035	62.92	141	2.4214	1.206
0.0005	1.029	63.12	109	1.8728	1.666
0.001	1.039	62.88	85	1.4708	1.958
0.0015	1.043	62.17	71	1.2092	1.795
0.002	1.052	61.68	73	1.2544	1.339
0.0025	1.234	51.77	34	0.5964	-13.799

4. 结论

本文提出的模型融合了一维二维 CNN、GRU 和自注意力机制的优势。具体来说, 使用一维和二维卷积层提取信号的浅层空间特征, 将卷积处理后的输出分成两个流进行对时间特征的提取。GRU 可以有效地处理序列数据, 并且与 LSTM 相比降低了计算复杂度。双流结构的设计使得模型对时间特征的提取更加充分。在 GRU 之间加入自注意力机制, 使模型关注到输入序列不同部分之间的相关性。在基准数据集上的实验结果验证了 DSGANN 在信噪比大于等于 -6 dB 时具有更高的识别精度。消融实验证明了模型各部分的有效性。在实践中, 用于训练调制识别网络的样本数量不足是一个常见的问题。当减小数据集的规模时, DSGANN 仍然实现了最高的识别精度, 这表明 DSGANN 在更小的数据集上的适用性。本文还研究了初始学习率对模型性能的影响, 给出了最高准确率的初始学习率和相对最高效的初始学习率。本文的方法为改进自动调制识别提供了一种创新的解决方案, 可以应用于各种无线通信场景。

参考文献

- [1] Dulek, B. (2017) Online Hybrid Likelihood Based Modulation Classification Using Multiple Sensors. *IEEE Transactions on Wireless Communications*, **16**, 4984-5000. <https://doi.org/10.1109/TWC.2017.2704124>
- [2] Wei, W. and Mendel, J.M. (2000) Maximum-Likelihood Classification for Digital Amplitude Phase Modulations. *IEEE Transactions on Communications*, **48**, 189-193. <https://doi.org/10.1109/26.823550>
- [3] Xu, J.L., Su, W. and Zhou, M. (2010) Likelihood-Ratio Approaches to Automatic Modulation Classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, **41**, 455-469. <https://doi.org/10.1109/TSMCC.2010.2076347>
- [4] Hazza, A., Shoaib, M., Alshebeili, S.A. and Fahad, A. (2013) An Overview of Feature-Based Methods for Digital Modulation Classification. *2013 1st International Conference on Communications, Signal Processing, and Their Applications*, Sharjah, 12-14 February 2013, 1-6. <https://doi.org/10.1109/ICCSPA.2013.6487244>

-
- [5] Chang, D. and Shih, P. (2015) Cumulants-Based Modulation Classification Technique in Multipath Fading Channels. *IET Communications*, **9**, 828-835. <https://doi.org/10.1049/iet-com.2014.0773>
- [6] Huang, S., Yao, Y., Wei, Z., Feng, Z. and Zhang, P. (2017) Automatic Modulation Classification of Overlapped Sources Using Multiple Cumulants. *IEEE Transactions on Vehicular Technology*, **66**, 6089-6101. <https://doi.org/10.1109/TVT.2016.2636324>
- [7] Xu, Y., Ge, L. and Wang, B. (2008) Digital Modulation Recognition Method Based on Self-Organizing Map Neural Networks. 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing, Dalian, 12-14 October 2008, 1-4. <https://doi.org/10.1109/WiCom.2008.418>
- [8] Kharbech, S., Dayoub, I., Zwingelstein-Colin, M., Simon, E.P. and Hassan, K. (2014) Blind Digital Modulation Identification for Time Selective MIMO Channels. *IEEE Wireless Communications Letters*, **3**, 373-376. <https://doi.org/10.1109/LWC.2014.2318041>
- [9] Jagannath, J., Polosky, N., O'Connor, D., Theagarajan, L.N., Sheaffer, B., Foulke, S. and Varshney, P.K. (2018) Artificial Neural Network Based Automatic Modulation Classification over a Software Defined Radio Testbed. 2018 IEEE International Conference on Communications (ICC), Kansas City, 20-24 May 2018, 1-6. <https://doi.org/10.1109/ICC.2018.8422346>
- [10] Jiang, Y., Zhang, Z. and Qiu, P. (2004) Modulation Classification of Communication Signals. Proc. IEEE MILCOM 2004. Military Communications Conference, Monterey, 31 October-3 November 2004, 1470-1476.
- [11] Park, C.S., Choi, J.H., Nah, S.P., Jang, W. and Kim, D.Y. (2008) Automatic Modulation Recognition of Digital Signals Using Wavelet Features and SVM. 2008 10th International Conference on Advanced Communication Technology, Gangwon, 17-20 February 2008, 387-390. <https://doi.org/10.1109/ICACT.2008.4493784>
- [12] Almohamad, T.A., Salleh, M.F.M., Mahmud, M.N. and Sa'D, A.H.Y. (2018) Simultaneous Determination of Modulation Types and Signal-to-Noise Ratios Using Feature-Based Approach. *IEEE Access*, **6**, 9262-9271. <https://doi.org/10.1109/ACCESS.2018.2809448>
- [13] Yang, F., Yang, L., Wang, D., Qi, P. and Wang, H. (2018) Method of Modulation Recognition Based on Combination Algorithm of K-Means Clustering and Grading Training SVM. *China Communications*, **15**, 55-63.
- [14] O'Shea, T.J., Corgan, J. and Clancy, T.C. (2016) Convolutional Radio Modulation Recognition Networks. 17th International Conference, EANN 2016, Aberdeen, 2-5 September 2016, 213-226. https://doi.org/10.1007/978-3-319-44188-7_16
- [15] Perenda, E., Rajendran, S. and Pollin, S. (2019) Automatic Modulation Classification Using Parallel Fusion of Convolutional Neural Networks. *Proceedings of Third International Balkan Conference on Communications and Networking*, Skopje, North Macedonia, 10-12 June 2019.
- [16] Zhang, Z., Wang, C., Gan, C., Sun, S. and Wang, M. (2019) Automatic Modulation Classification Using Convolutional Neural Network with Features Fusion of SPWVD and BJD. *IEEE Transactions on Signal and Information Processing over Networks*, **5**, 469-478. <https://doi.org/10.1109/TSIPN.2019.2900201>
- [17] Zhang, Q., Xu, Z. and Zhang, P. (2019) Modulation Scheme Recognition Using Convolutional Neural Network. *The Journal of Engineering*, **2019**, 9075-9078. <https://doi.org/10.1049/joe.2018.9188>
- [18] Rajendran, S., Meert, W., Giustiniano, D., Lenders, V. and Pollin, S. (2018) Deep Learning Models for Wireless Signal Classification with Distributed Low-Cost Spectrum Sensors. *IEEE Transactions on Cognitive Communications and Networking*, **4**, 433-445. <https://doi.org/10.1109/TCCN.2018.2835460>
- [19] Liu, X., Yang, D. and Gamal, A.E. (2017) Deep Neural Network Architectures for Modulation Classification. 2017 51st Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, 29 October-1 November 2017, 915-919. <https://doi.org/10.1109/ACSSC.2017.8335483>
- [20] Xu, J., Luo, C., Parr, G. and Luo, Y. (2020) A Spatiotemporal Multi-Channel Learning Framework for Automatic Modulation Recognition. *IEEE Wireless Communications Letters*, **9**, 1629-1632. <https://doi.org/10.1109/LWC.2020.2999453>
- [21] Zhang, F., Luo, C., Xu, J. and Luo, Y. (2021) An Efficient Deep Learning Model for Automatic Modulation Recognition Based on Parameter Estimation and Transformation. *IEEE Communications Letters*, **25**, 3287-3290. <https://doi.org/10.1109/LCOMM.2021.3102656>
- [22] Bahdanau, D., Cho, K. and Bengio, Y. (2014) Neural Machine Translation by Jointly Learning to Align and Translate.
- [23] Liang, Z., Tao, M., Wang, L., Su, J. and Yang, X. (2021) Automatic Modulation Recognition Based on Adaptive Attention Mechanism and ResNeXt WSL Model. *IEEE Communications Letters*, **25**, 2953-2957. <https://doi.org/10.1109/LCOMM.2021.3093485>
- [24] Lin, S., Zeng, Y. and Gong, Y. (2022) Learning of Time-Frequency Attention Mechanism for Automatic Modulation Recognition. *IEEE Wireless Communications Letters*, **11**, 707-711. <https://doi.org/10.1109/LWC.2022.3140828>

- [25] O'shea, T.J. and West, N. (2016) Radio Machine Learning Dataset Generation with GNU Radio. *Proceedings of the GNU Radio Conference*, **1**. <https://pubs.gnuradio.org/index.php/grcon/article/view/11>
- [26] Zhang, F., Luo, C., Xu, J., Luo, Y. and Zheng, F. (2022) Deep Learning Based Automatic Modulation Recognition: Models, Datasets, and Challenges. *Digital Signal Processing*, **129**, Article ID: 103650. <https://doi.org/10.1016/j.dsp.2022.103650>