

基于注意力机制的门控密集卷积网络调制识别算法

杨驰 龚晓峰 雒瑞森

(四川大学电气工程学院 四川 成都 610065)

摘要 自动调制识别(AMR)是非合作通信系统中的重要组成部分,也是一个通信领域的研究难点。针对该难点,利用深度学习,将密集卷积网络(DenseNet)、门控循环单元(GRU)和注意力机制(Attention)三者结合,提出一种基于注意力机制的门控密集卷积网络(AGDCN)的调制识别算法。该算法提取了信号的空间特征和时序特征,将两者相结合解决了信号识别率低的问题。同时,在网络中加入注意力机制,对GRU训练过程进行权重的自适应调整,有效地加强关键特征的学习。通过实验验证了AGDCN模型性能优于其他神经网络算法,在信噪比超过2 dB时,对11种调制类型的识别率可以达到90%。

关键词 自动调制识别 深度学习 密集卷积网络 门控循环单元 注意力机制

中图分类号 TP391.4

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.10.019

A MODULATION RECOGNITION ALGORITHM BASED ON ATTENTION MECHANISM AND GATED DENSE CONVOLUTIONAL NETWORKS

Yang Chi Gong Xiaofeng Luo Ruisen

(College of Electrical Engineering, Sichuan University, Sichuan 610065, Chengdu, China)

Abstract Automatic modulation recognition (AMR) is an important part of non cooperative communication system, and it is also a research difficulty in the field of communication. In order to solve the above problems, this paper proposes an attention based gated dense convolutional network (AGDCN) modulation recognition algorithm by combining dense convolutional network (DenseNet), gated recurrent unit (GRU) and attention mechanism. The algorithm extracted the spatial and temporal features of the signal, and combined them to solve the problem of low recognition rate. Attention mechanism was added to the network to adaptively adjust the weight of GRU training process and effectively strengthen the learning of key features. Experiments show that the proposed AGDCN model outperforms other mainstream neural network algorithms. Specifically, when SNR exceeds 2 dB, the recognition rate of 11 modulation types can reach 90%.

Keywords Automatic modulation recognition Deep learning Dense convolutional network Gated recurrent unit Attention mechanism

0 引言

无线电信号自动调制识别(Automatic Modulation Recognition, AMR)作为非合作通信系统的一项重要技术,在认知无线电、干扰识别和信号监测等民用和军用应用中发挥着重要作用^[1],它可以在没有先验知识的情况下,识别被截获信号的调制类型,为后续的解调提

供信息。传统的AMR算法可以分为两大类:基于似然比方法和基于特征识别方法^[2]。相比于似然方法,特征识别的方法是目前研究的主要方向,主要特征包括瞬时特征、高阶累计量^[3]、循环谱^[4]和小波特性和小波特性^[5]。使用的分类器也是基于机器学习算法的,其中包括决策树、支持向量机(Support Vector Machine, SVM)^[6]和人工神经网络(Artificial Neural Network, ANN)^[7]。这些传统的以人工提取特征为核心的AMR算法识别精

度不够理想,局限性太大,因此,近年来深度学习(Deep Learning, DL)逐渐成为 AMR 领域的研究热点。

深度学习作为一种强大的机器学习方法,在图像处理、情感分析、语音识别等多个领域都获得了巨大的成功^[8]。相比于传统的数据分析和处理方法,DL 能够自动地、更好地表征复杂数据,而无须设计人工特征。因此,许多学者也将这一方法引入了信号调制识别中来。文献[9]利用多层的卷积神经网络(Convolutional Neural Network, CNN),结合信号的时频图像进行分类识别。文献[10]利用深度学习中的 Alex 网络对信号的星座图像进行分类识别。文献[11]提出了一种结合谱相关函数特征和深度信念网络模型(Deep Believe Network, DBN)生成的 AMC 方案。上述基于深度学习的方法相较于传统机器学习方法,不需要人为提取特征,识别精度也有所提升。但是,将原始信号序列转换为二维图像需要大量的计算,并且对于信号的特征能力也不足。2016 年, O'shea 等^[12]使用开源无线电软件(GNU Radio)生成了一组调制信号的公共数据集用于调制识别,这些信号具有同相和正交(I/Q)信息,命名为 RML2016. 10a,同时他们提出了一种双层 CNN 模型^[13],对该数据集中的 11 种调制信号进行分类。文献[14]提出了一种深度残差网络,利用 IQ 信号作为输入,在 5 dB 信噪比以上时分类效果良好。虽然 CNN 和其改进结构可以有效提取空间局部特征,但随着网络深度加深,梯度消失现象明显,同时忽略了时间序列的上下文信息。循环神经网络(Recurrent Neural Network, RNN)可以很好解决上下文依赖的关系,比如常见的长短期记忆(Long Short-Term Memory, LSTM)网络和门控循环单元(Gated Recurrent Unit, GRU)网络。研究表明^[15]直接使用 LSTM 或 GRU 作为网络模型结构进行调制方式的识别也可以达到不错的分类精度,不过与 LSTM 相比,GRU 使用的参数更少,训练速度和执行速度更快。但是这些网络结构在低信噪比时性能不佳,且本身存在着训练难度较大的问题。

针对以上问题,本文提出一种基于注意力机制的门控密集卷积网络(Attention Based Gated Dense Convolutional Network, AGDCN),该模型无须对原始信号序列进行复杂的转换,利用密集卷积网络(Dense Convolutional Networks, DenseNet)进行空间特征提取^[16],通过特征的重复利用缓解了梯度消失现象,加强了特征传播,再经过 GRU 网络提取更多的时间性特征,充分利用了 GRU 网络处理长时间依赖序列的优势,同时结合 Attention 机制^[17],能让模型在训练过程中更加关注那些比较重要的特征,加快模型训练效率,提升识别准确率。

1 基本算法

1.1 DenseNet 网络结构

DenseNet 网络是一种密集连接方式^[16],其核心思想在于特征重用,与残差网络(Residual Network, ResNet)相比,其结构设计更加激进,每一层的输出都会与前面所有层在通道上相连,这样最终的分层器就可以根据整个网络的所有特征做出决策,如图 1 所示。其优点在于拥有很好的抗过拟合能力,同时泛化性能也很优越。

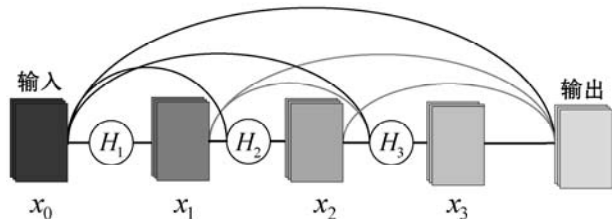


图 1 DenseNet 内部结构

若网络层数为 N ,则网络中共有 $N(N+1)/2$ 个连接,其中第 l 层的输出是前面所有层的映射,如式(1)所示。

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (1)$$

式中: x_0 指的是输入信号, x_1, x_2, \dots, x_{l-1} 表示前面所有层的输出的密集连接; $H(\cdot)$ 是非线性转换函数,包含了批处理归一化(Batch Normalization, BN)、ReLU 激活函数、填充(Padding)、和卷积(Convolution)等复合函数。

由于特征重用导致后面卷积层的输入维度很大,因此每一层的输入都会先经过一个 1×1 的卷积,减少输入的特征映射数量,既能减少计算量,还能融合各个通道的特征,如图 2 所示。其中 C 表示输入特征的通道数, g 代表每个 H 层输出的特征通道数。

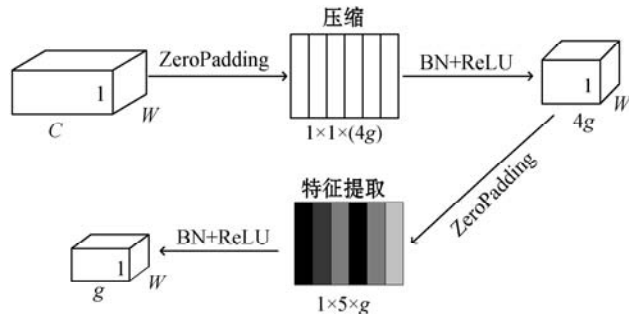


图 2 H 层内部结构

1.2 GRU 模型和 Attention 机制

本文使用 GRU 网络来进行时序特征的提取,其内部结构如图 3 所示。GRU 作为一种特殊的 RNN 网络,能够检测到长期序列模型中潜在的依赖关系,同时与 LSTM 比较起来,内部结构和参数更加简单,训练效率

更高^[18]。

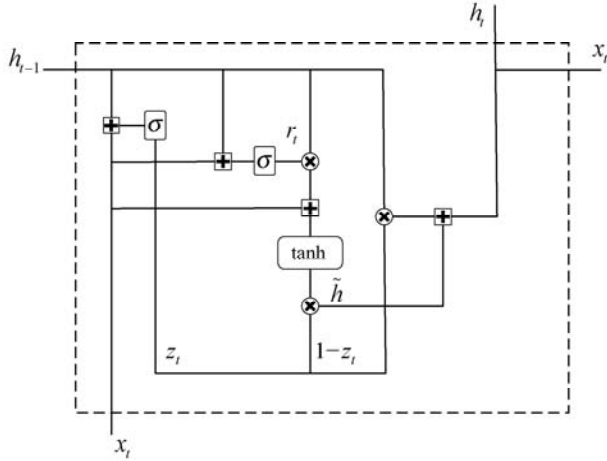


图3 GRU 内部结构

GRU 模型由更新门 z_t 和重置门 r_t 组成,在时刻 t , 其输出更新状态表达式如下:

$$r_t = \sigma(\mathbf{w}_r x_t + \mathbf{U}_r h_{t-1} + \mathbf{b}_r) \quad (2)$$

$$z_t = \sigma(\mathbf{W}_z x_t + \mathbf{U}_z h_{t-1} + \mathbf{b}_z) \quad (3)$$

$$\tilde{\mathbf{h}} = \tanh(\mathbf{W}_h x_t + r_t(\mathbf{U}_h \times h_{t-1} + \mathbf{b}_h)) \quad (4)$$

$$\mathbf{h}_t = (1 - z_t) \times \mathbf{h}_{t-1} + z_t \times \tilde{\mathbf{h}} \quad (5)$$

式中: h_{t-1} 表示上一时刻的输出状态, z_t 用来控制上一时刻的信息被带入到当前时刻状态的程度, $\tilde{\mathbf{h}}$ 表示当前时刻的候选状态, r_t 控制了前一刻状态对当前候选状态的影响, \mathbf{W} 、 \mathbf{U} 、 \mathbf{b} 分别为权重和偏置, $\sigma(\cdot)$ 为激活函数。

虽然利用 GRU 可以提取到长期序列的特征,但是不同位置的特征对最终分类结果的影响是不一样的,特别需要加强关键特征的学习。基于上述考虑,我们将 GRU 与 Attention 机制相结合,模型结构如图 4 所示。

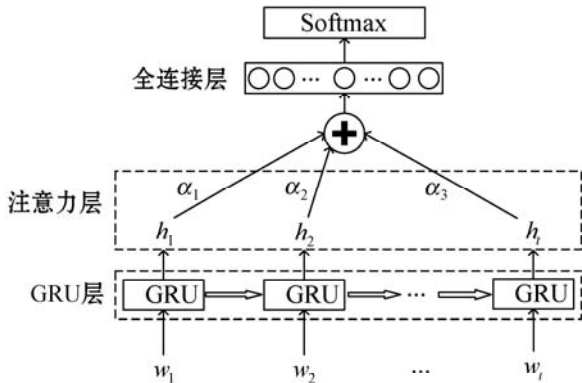


图4 基于 Attention 机制的 GRU 结构

式中: h_t 为 GRU 网络输出的第 t 个特征向量,将其输入 Attention 层得到初始状态向量 s_t ,再与权重系数 α_t 相乘并累加求和,得到最终输出的状态向量 \mathbf{Y} 。计算公式如下^[17]:

$$s_t = \tanh(\mathbf{w}_t h_t + \mathbf{b}_t) \quad (6)$$

$$\alpha_t = \frac{\exp(s_t)}{\sum_{i=1}^t s_i} \quad (7)$$

$$\mathbf{Y} = \sum_{i=1}^n \alpha_i s_i \quad (8)$$

式中: w_t 和 b_t 分别表示第 t 个特征向量的权重和偏置,根据式(7)和式(8)将输入的初始状态转换为最终输出的状态向量 \mathbf{Y} ,最后将全连接层与之相连,再经过 Softmax 函数对最后的输出进行计算,得到最终的分类结果,即:

$$\tilde{y} = \operatorname{argmax}(\operatorname{softmax}(\mathbf{W}_y \mathbf{Y} + \mathbf{b}_y)) \quad (9)$$

2 AGDCN 模型

2.1 AGDCN 结构和参数

本文的网络结构如图 5 所示,其配置信息如表 1 所示,该网络由普通二维卷积层、DenseNet 层、GRU 层、Attention 层和全连接层组成。网络输入信号的 I/Q 数据,经过 3 个 DenseBlock 层,提取信号的高维特征,再利用 GRU 层提取特征长序列之间潜在的依赖特征,配合上 Attention 层对 GRU 训练过程的优化,将更多的注意力集中到关键特征,使得整个网络的训练效率、识别速度和分类精度都达到了较高的水准。

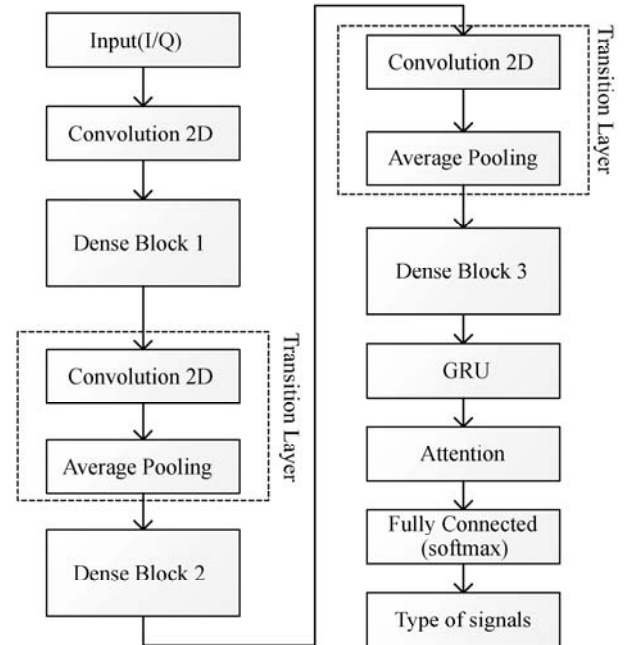


图5 AGDCN 模型

由于每一个 DenseBlock 模块输出的维度是很大的,因此为了进一步压缩参数,在两个 DenseBlock 之间加入了一个由维度为的卷积层和平均池化层(Average Pooling)组成的过渡层(Transition Layer),如图 5 中虚线框所示。

另外,该网络结构中所有的卷积层都包括了填充、卷积、批处理化、ReLU 激活函数和 dropout 五个步骤,其中 dropout 等于 0.5。并且,设置两个 DenseBlock 模块中 H 函数的数量为 4,每个 H 函数的增长率为 32,过渡层中卷积核的数量为输入特征通道数的 0.5。设置 GRU 单元数量为 50,最后输出 11 种调制类型的分类结果。具体参数见表 1。

表 1 AGDCN 网络结构参数

网络层	输出维度	配置参数
Convolution	$(1 \times 128, 64)$	2×5 Conv, 64 kernels
DenseBlock 1	$(1 \times 128, 192)$	$[1 \times 5$ Conv, 32] $\times 4$
Transition Layer 1	$(1 \times 128, 96)$	1×1 Conv, 96 kernels
	$(1 \times 64, 96)$	1×2 average pooling, stride 2
DenseBlock 2	$(1 \times 64, 224)$	$[1 \times 5$ Conv, 32] $\times 4$
Transition Layer 2	$(1 \times 64, 112)$	1×1 Conv, 112 kernels
	$(1 \times 32, 96)$	1×2 average pooling, stride 2
DenseBlock 3	$(1 \times 64, 240)$	$[1 \times 5$ Conv, 32] $\times 4$
GRU + Attention	$(1 \times 240, 50)$	50 units
Classification	11	fully connection, softmax

2.2 训练流程

为了省去信号处理和转换的时间,我们直接将信号的 IQ 序列作为模型的输入:

步骤 1 标签处理。对不同类别的标签进行独立热编码。

步骤 2 建立数据集。将数据和标签相互匹配,在将它们打乱后,从样本中以一定的比例构成训练集和测试集。

步骤 3 添加早停(EarlyStopping)机制,训练过程中,验证集的误差会表现为先减小后增大,表明出现了过拟合现象。设置 patience 值为 10,当验证集误差经过 10 轮的训练都不再减小时,就提前中止训练。

步骤 4 设置日期时间,以获取网络训练过程的用时。

步骤 5 设置模型的训练过程。其中损失函数采用 CategoricalCrossentropy,优化器采用 Adam,初始学习率等于,批次大小设置为 256,训练轮数为 100,最后开启训练。

3 实验

3.1 数据集和实验环境

实验数据采用 O'shea 等提出的 RadioML2016.10a

开放数据集^[12],该数据集涵盖了 11 种被广泛使用的调制方式的 IQ 信号,为了能真实地模拟各种情况下的信号,其中还加入了加性高斯白噪声、频率偏移等信道误差,信道模型也包括了采样率和载波率偏移量,具体参数见表 2。

表 2 RadioML2016.10a 数据集参数

数据集	参数配置
调制方式	WE-FM, AM-SSB, AM-DSB, BPSK, QPSK, 8PSK, 16QAM, 64QAM, BFSK, CPFSK, PAM4
数据维度	2×128
样本总数	220000
信噪比/dB	-20:2:18
采样频率	1 MHz

随机选取 20% 的样本作为测试集,80% 的样本作为训练集,实验运行在 Linux 服务器上,CPU 为 Intel i5-7500,GPU 为 NVIDIA GeForce GTX 2070,并且通过 Python 中的 TensorFlow 框架完成网络的搭建、训练和测试。

3.2 实验结果及分析

实验 1 为了探究 LSTM 和 GRU 各自对分类性能的影响,以及注意力机制给网络模型带来的性能提升,分别在有无注意力机制的情况下进行实验,以测试集在 -20 dB ~ 18 dB 范围内的平均分类准确率、训练集训练轮数和训练时间作为评价标准,实验结果如表 3 所示。

表 3 探究 LSTM、GRU 和 Attention 对网络分类能力的影响

序号	网络	有无 Attention	训练轮数	训练时间/s	准确率/%
1	LSTM	有	32	806	62.79
2	LSTM	无	63	2016	61.13
3	GRU	有	28	644	62.83
4	GRU	无	56	1680	61.16

从表 3 可以看出,采用 GRU 的分类效果只比 LSTM 高出 0.04%,但训练轮数和训练时间却比 LSTM 少,原因是 GRU 的模型相对简单,结构参数和训练难度都比 LSTM 低,采用 GRU 作为网络结构更为合适。另外,从表中还可以看出,GRU 与注意力机制结合后,测试集平均分类精度达到了 62.83%,分类性能提升了 1.67%,并且训练轮数和训练时间都有所减少。实验结果证明,注意力机制能有效提升网络对调制信号的分类精度和识别效率。

实验 2 由于网络深度对于网络的能力有着很大的影响,为了探究不同的网络结构对于训练结果的影

响,图 6 给出了 AGDCN 模型中,不同的 DenseBlock 层数和 GRU 层数的性能表现。

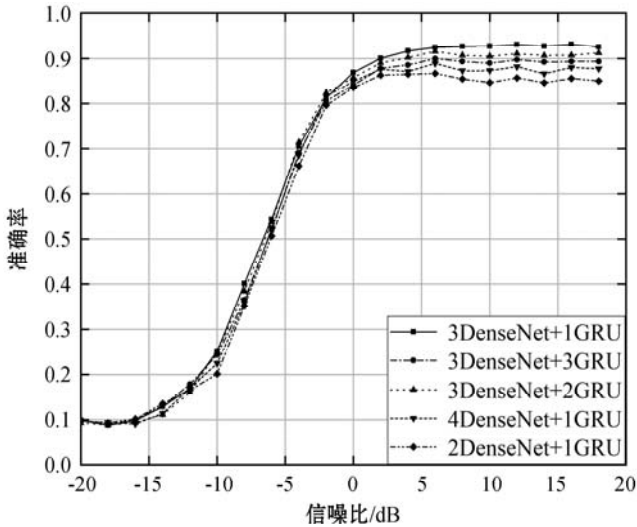


图 6 不同信噪比下网络结构的识别率

从图 6 可以看出 3 个 DenseBlock 加 1 个 GRU 的识别准确率最高,网络性能最好。通过实验结果证明了,当 DenseBlock 层数增加和减少时,网络的识别率都有所下降,说明基于特征重用的卷积层不能太多,也不能太少,太多会导致网络结构更加复杂,太少特征提取能力不足。实验结果还说明了,增加 GRU 层的数量并不能有效提升识别率,当层数为 1 时刚刚好,既有效改善了网络的性能,同时也不会导致模型变复杂。

实验 3 为了进一步研究本文模型的性能,将其与目前先进的方法做对比,我们选择文献[13]提出的 CNN2 模型、文献[14]提出的 Resnet 模型和 Inception 模型、文献[15]提出的双层门控回归单元模型(GRU2)和文献[19]提出的 CNN-LSTM 模型作为基准,这 6 种识别算法的性能如图 7 所示,它们的训练用时和网络收敛训练的轮数如表 4 所示。

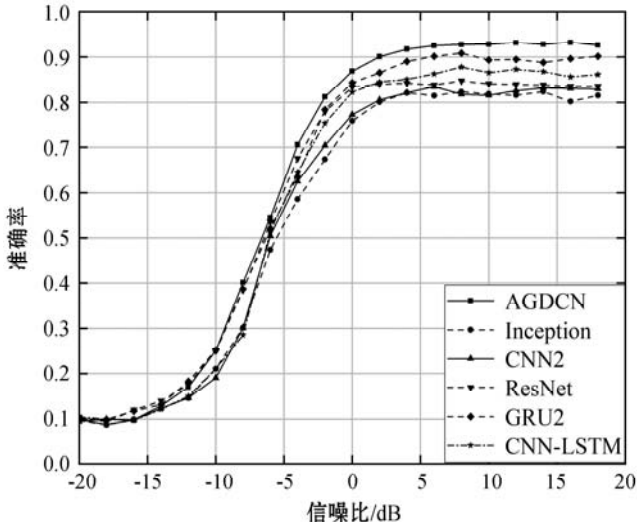


图 7 6 种模型在不同信噪比下的识别准确率

表 4 6 种模型的训练情况

模型	训练轮数	训练用时/s
AGDCN	28	644
CNN2	24	432
ResNet	62	2 046
Inception	52	1 144
GRU2	43	1 247
CNN-LSTM	84	2 432

结合图 7 和表 4 可以看出,AGDCN 模型比其他经典的深度学习模型识别准确率要高,提升效果明显,同时训练用时也较短,说明了其训练过程收敛速度较快,训练相对容易。CNN2 的训练用时最短,但是其识别精度很低,这是由于 CNN2 结构简单,但是其无法提取长期信息,所以精度不高。ResNet 和 Inception 都属于 CNN 的改进模型,虽然加强了对空间特征的提取,但是依然无法有效利用时序特征,导致识别准确率上限较低,并且网络的加深导致训练难度增大,训练用时增多。GRU2 模型虽然其识别精度比 CNN2、ResNet 和 Inception 高,最高达到 90.8%,但是未能有效利用空间特征,且本身训练难度较大,最终识别准确率低于本文所提出的模型。CNN-LSTM 结构与本文模型相似,但是其模型结构简单,泛化能力不强,收敛速度慢,所以准确度和训练速度都比较低,相比之下,本文提出的模型,利用特征复用抑制了过拟合,加强了泛化能力,通过 Transition 层降低了数据的维度,还降低了 GRU 层所需要处理的序列长度,同时 Attention 机制还优化了 GRU 模块,使其每次只关注重要的特征,降低了训练难度,加快了收敛速度,也提升了识别准确率。在信噪比为 -2 dB 时,准确率依然在 80% 以上,在 2 dB 时准确率已经达到了 90%,在 16 dB 时,识别准确率达到最高的 93.16%,此时 11 类信号的混淆矩阵如图 8 所示。

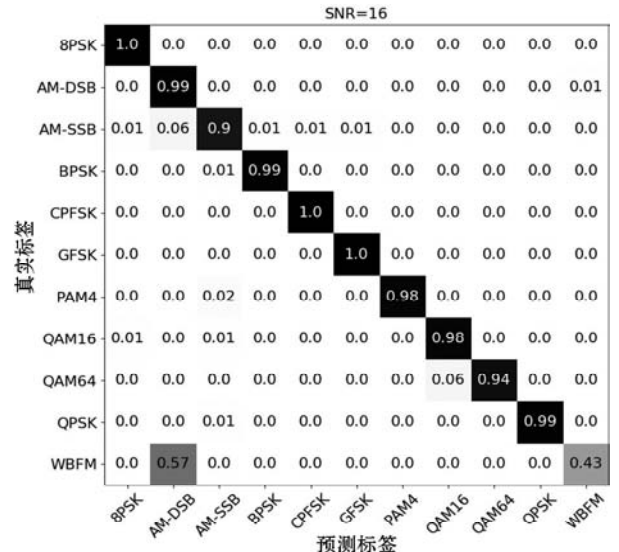


图 8 SNR = 16 dB 时 11 类信号的混淆矩阵

从图 8 中可以看出,在模型识别准确度最高时,WBFM 信号容易和 AM-DSB 信号混淆,其识别准确率最低,只有 43%,另外,AM-SSB 信号也有 10% 的误判率,QAM64 信号有 6% 的概率误判为 QAM16 信号,其他的信号基本没有误判。

4 结 语

本文提出了一种基于深度学习的调制识别算法,实现了对 11 种不同类型的调制信号的识别。该算法融合了 DenseNet 网络、GRU 网络和 Attention 机制的优点,无须进行信号转换,直接将 IQ 信号序列作为输入,提取数据集的空间局部特征和时序特征,具有较高的识别速度和识别准确率。通过开源数据集 RML2016.10a 对方案进行验证,最高可以达到 93.16% 的高识别率,在低信噪比 -2 dB 时,仍有 80% 以上的识别率。实验结果证明,与其他先进的网络模型对比,在识别准确度上有显著提高。

参 考 文 献

- [1] 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.
- [2] Dobre O A. Signal identification for emerging intelligent radios: classical problems and new challenges[J]. IEEE Instrumentation & Measurement Magazine, 2015, 18(2): 11-18.
- [3] Huang S, Yao Y, Wei Z Q, et al. Automatic modulation classification of overlapped sources using multiple cumulants[J]. IEEE Transactions on Vehicular Technology,2016,66(7):6089-6101.
- [4] Yan X, Liu G, Wu H C, et al. New automatic modulation classifier using cyclic-spectrum graphs with optimal training features[J]. IEEE Communications Letters,2018,22(6):1204-1207.
- [5] Avci E, Avci D. A novel approach for digital radio signal classification: Wavelet packet energy-multiclass support vector machine (WPE-MSVM)[J]. Expert Systems with Applications,2008,34(3):2140-2147.
- [6] Park C S, Choi J H, Nah S P, et al. Automatic modulation recognition of digital signals using wavelet features and SVM[C]//10th International Conference on Advanced Communication Technology,2008:387-390.
- [7] Popoola J, Olst R. Effect of training algorithms on performance of a developed automatic modulation classification using artificial neural network[C]//Africon,2013:1-6.
- [8] Dargan S, Kumar M, Ayyagari M R, et al. A survey of deep learning and its applications: A new paradigm to machine learning[J]. Archives of Computational Methods in Engineering,2019,27(4):1071-1092.
- [9] Zeng Y, Zhang M, Han F, et al. Spectrum analysis and convolutional neural network for automatic modulation recognition[J]. Wireless Communications Letters IEEE,2019,8(3):929-932.
- [10] Peng S L, Jiang H Y, Wang H X, et al. Modulation classification based on signal constellation diagrams and deep learning[J]. IEEE Transactions on Neural Networks and Learning Systems,2018,30(3):718-727.
- [11] Mendis G J, Wei J, Madanayake A. Deep learning-based automated modulation classification for cognitive radio[C]//IEEE International Conference on Communication Systems, 2016:1-6.
- [12] O'Shea T J, West N. Radio machine learning dataset generation with gnu radio[C]//6th GNU Radio Conference,2016:2151-2162.
- [13] O'Shea T J, Corgan J, Clancy T C. Convolutional radio modulation recognition networks[C]//International Conference on Engineering Applications of Neural Networks,2016:213-226.
- [14] West N E, O'Shea T. Deep architectures for modulation recognition[C]//IEEE International Symposium on Dynamic Spectrum Access Networks,2017:1-6.
- [15] Hong D, Zhang Z L, Xu X D. Automatic modulation classification using recurrent neural networks[C]//3rd IEEE International Conference on Computer and Communications, 2017:695-700.
- [16] Huang G, Liu S C, Maaten L V, et al. CondenseNet: An efficient DenseNet using learned group convolutions[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition,2018:2752-2761.
- [17] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[EB]. arXiv:1409.0473,2014.
- [18] Pudikov A, Brovko A. Comparison of LSTM and GRU recurrent neural network architectures[C]//International Scientific and Practical Conference in Control Engineering and Decision Making,2020:114-124.
- [19] Wu Y L, Li X J, Fang J. A deep learning approach for modulation recognition via exploiting temporal correlations[C]//19th International Workshop on Signal Processing Advances in Wireless Communications,2018:1-5.