

基于轻量级深度神经网络的电磁信号调制识别技术

张思成¹, 林云¹, 涂涯¹, Shiwen Mao²

(1. 哈尔滨工程大学信息与通信工程学院, 黑龙江 哈尔滨 150001; 2. 奥本大学电子和计算机工程学院, 奥本 36849)

摘要: 针对 6G 时代将会是移动通信与人工智能紧密结合的时代, 产生数量庞大的边缘智能信号处理节点的趋势, 提出了一种可部署于资源受限的边缘设备上的高效智能电磁信号识别模型。首先, 通过绘制电磁信号的星座图将电磁信号具象为二维图像, 并根据归一化点密度对星座图上色以实现特征增强; 然后, 使用二值化深度神经网络对其进行识别, 在保证识别准确率的同时明显降低了模型存储开销以及计算开销。采用电磁信号调制识别问题进行验证, 实验选取常用的 8 种数字调制信号, 选择加性高斯白噪声为信道环境。实验结果表明, 所提方案可以在信噪比为 $-6\sim 6$ dB 的噪声条件下获得 96.1% 的综合识别率, 网络模型大小仅为 166 KB, 部署于树莓派 4B 的执行时间为 290 ms, 相比于同规模的全精度网络, 准确率提升了 0.6%, 模型缩减到 $\frac{1}{26.16}$, 运行时间缩减到 $\frac{1}{2.37}$ 。

关键词: 6G; 边缘智能; 电磁信号调制识别; 图像具象化; 二值化深度神经网络

中图分类号: TN971

文献标识码: A

doi: 10.11959/j.issn.1000-436x.2020237

Electromagnetic signal modulation recognition technology based on lightweight deep neural network

ZHANG Sicheng¹, LIN Yun¹, TU Ya¹, Shiwen Mao²

1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

2. Department of Electrical and Computer Engineering, Auburn University, Auburn 36849, USA

Abstract: In response to the trend that in the 6th generation wireless (6G) era, mobile communications and artificial intelligence will be closely integrated, and a huge number of edge intelligent signal processing nodes will be deployed, an efficient and intelligent electromagnetic signal recognition model was proposed, which could be deployed on resource-constrained edge devices. The constellation diagram of electromagnetic signal was firstly drawn to visualize electromagnetic signal as a two-dimensional image, and color the constellation diagram according to the normalized point density to achieve feature enhancement. Then, a binary deep neural network was adopted to recognize the colored constellation diagrams. It was shown that the approach can guarantee a high recognition accuracy, which significantly reduced the model storage and calculation costs. For verification, the proposed approach was applied to the problem of electromagnetic signal modulation recognition. The experiment uses eight commonly used digital modulation signals and selects additive white Gaussian noise as the channel environment. The experimental results show that the scheme can achieve a comprehensive recognition rate of 96.1% under the noise condition of $-6\sim 6$ dB, while the size of the network model is only 166 KB. Further, the execution time, when executed on a Raspberry Pi 4B, is only 290 ms. Compared to a full-precision network of the same scale, the accuracy is increased by 0.6%, the model is reduced to $\frac{1}{26.16}$, and the running time is reduced to $\frac{1}{2.37}$.

Key words: 6G, edge intelligence, electromagnetic signal modulation recognition, image visualization, binary deep neural network

收稿日期: 2020-08-21; 修回日期: 2020-11-06

基金项目: 国家自然科学基金资助项目 (No.61771154); 中央高校基本科研业务费基金资助项目 (No.3072020CF0813)

Foundation Items: The National Natural Science Foundation of China (No.61771154), The Fundamental Research Funds for the Central Universities (No.3072020CF0813)

1 引言

移动通信系统由 1G、2G 到 3G、4G，现已进入 5G 的商业发展阶段。5G 的商业部署工作已经展开，将提供大规模连接、高可靠性和低时延的服务。本着“商用一代、规划一代”的发展理念，部分国家已经展开了对 6G 的研究与探索。6G 将依赖新的使能技术实现全球覆盖范围，更高的频谱、能源、成本效率，更加安全，更高的数据速率，更低的时延，更高的连接密度，更高的全自动智能性，以获得更好的服务质量、体验质量，实现万物互联^[1]。同时，6G 时代将会是移动通信技术与人工智能技术紧密结合、深度融合的时代，放眼于智能、通信与人类未来三者之间的关系，深入挖掘用户的需求，实现人-机-物-灵的“万物互联”^[2]。

伴随 6G 时代的到来，将会产生大量分布式节点，这些分布式节点的智能连接是实现 6G 时代万物互联的关键，而智能连接的基础是物理层智能信号处理。其中，物理层信号识别将促进自适应调制编码、联合信道估计和符号检测、自适应传输功率控制以及更加广泛的物理层技术的发展^[3]。分布式节点实现智能、自动的电磁信号识别将极大地推动 6G 万物互联的实现。人工智能技术凭借其强大的拟合复杂模型的能力，近几年得到了迅猛的发展，在各项研究领域中都取得了突破性的进展，其在信号识别与处理问题中已经有丰富的成果。Li 等^[4]提出一种基于贝叶斯随机推断和最大后验准则的联合估计模型，该模型深层表征了调制方案和衰落信道之间的耦合关系，实现识别信号调制方式的同时顺序获取时间相关的衰落状态。Wang 等^[5]使用卷积神经网络 (CNN, convolutional neural network) 对信号的眼图进行识别，可以在宽信噪比范围内对 4 种调制信号实现 100% 的识别率。查雄等^[6]通过将电磁信号的波形域数据转化为眼图与矢量图形式，获得信号的浅层次表达，进一步利用多端卷积神经网络对其进行识别，当信噪比为 5 dB 时，识别准确率可达 95%。Peng 等^[7]将 8 种调制信号转换为星座图并处理得到灰度图像和增强灰度图像，使用 CNN 对其进行训练和测试，仿真结果表明该方法可以有效地对 8 种调制信号进行分类。吴佩军等^[8]设计了 33 层卷积神经网络 ReSENet，采集实测信号作为实验样本，使用短时傅里叶变换将信号转换为时频图，实验结果表明该方法可对多进制相位调制信号实

现 99.9% 的识别准确率。Ji 等^[9]针对多径衰落信道提出盲均衡辅助的深度学习网络，实验表明该方法相比不均衡的深度学习方法对正交振幅调制 (QAM, quadrature amplitude modulation) 类信号的识别精度有显著提升。Xu 等^[10]针对运用 CNN 进行调制信号识别的课题展开研究，发现较深的网络模型不适用于调制信号识别，并使用迁移学习对训练过程进行了优化，同时使用实测信号进行实验，识别性能得到了提升。

然而，在追求模型性能提升的同时，模型复杂度与参数量也成倍增长，随之而来的是过高的计算复杂度与存储开销，这对其应用于 6G 智能边缘设备产生了极大的阻碍。桂冠等^[11]对人工智能技术应用于物理层无线通信问题中的研究进行了总结与分析，说明了人工智能技术具有诸多优点，模型压缩与加速技术将是未来无线通信中影响系统带宽、时延以及安全问题的重要因素。研究者开展了大量工作，研究如何将基于人工智能技术实现的智能信号处理技术部署于资源受限的设备中。Tu 等^[12]在文献[6]的基础上提出基于点密度的星球图数据处理方法，使用 AlexNet 对其进行识别，准确率有明显提升，此外，他还根据文献[13]的方法将模型进行剪枝，在保证准确率的同时，将参数量压缩至原模型的 1.5%~5%，部署在 NVIDIA Jetson TX2 上的运行 1 000 次的时间为 6~8 ms。文献[14-15]进一步提出激活最大化的评估指标对模型进行评估与剪枝，实验结果显示，该方法在不同压缩率下针对 RML2016.a 数据集获得了同等或更高的识别准确率。陶冠宏等^[16]将信号的 I/Q 两路分别提取出来，构建 2 行 N 列的二维图像，输入 CNN 中进行训练，测试结果表明该方法可以对 8 种调制信号实现 85% 的识别准确率，实际部署后的单信号识别时间为 0.1 ms。周鑫等^[17]将信号的 I/Q 两路数据经过离散傅里叶变换映射到频域，经过图像化处理将多次变换的结果组合为二维频谱瀑布图输入 CNN 中进行训练，实验结果表明，该方法可以对 12 种信号实现 86.04% 的识别准确率，部署于中等配置的台式机上的运行时间为 33 ms。

以上的研究成果都很难取得准确率与速度上的双重提升，是因为以往的全精度神经网络 (FNN, full-precision neural network) 仅从网络模型的结构入手进行修剪或稀疏处理，有些方法对模型的容量以及稳健性会有损失；有些方法对计算机运算过程并不友好，在减小模型参数存储的同时增加了模型

的运算时间。考虑计算机本身的数据存储、访问、运算以及网络分类性能，研究人员发现神经网络在逐层计算过程中，参数的量化空间不需要非常完备也可以对问题模型进行拟合，从而完成正确的分类或检测，精简后的网络参数运算可以用简单的位操作完成，而不需要复杂的浮点乘法运算，这样兼顾了保持精度与加速的效果。基于此想法，Dettmers 等^[18]开发测试了 8 bit 浮点的网络模型，再利用 GPU 集群测试其性能，在保证模型与预测性能的前提下，速度提升到原来的 2 倍。此后出现了各种模型量化的研究，包括 16 bit^[19]、三值量化^[20]与二值量化^[21]。目前，主流的模式量化方法分为朴素量化方法和基于优化的量化方法，后者又分为最小量化误差量化法、改善网络损失函数量化法和减小梯度误差量化法^[22]。

本文针对 6G 中智能边缘设备的信号识别问题，在已有的电磁信号识别算法研究的基础上，将电磁信号的 I/Q 数据绘制为星座图，以将电磁信号从波形域具象化为二维图像，并根据归一化点密度对其添加颜色信息以实现特征增强；分析现有的信号识别模型的资源开销，设计并构建了二值化神经网络 (BNN, binary neural network) 模型对经过特征增强的图像域信号进行识别。经过初步实验证明，所提模型能够在保证对电磁信号识别准确率的前提下，兼具内存开销小、数据访问量低和计算速度高的特点，适合于 6G 中资源受限的智能边缘设备。

2 电磁信号模型

2.1 电磁信号波形域分析

本文实验选取 8 种常用的数字调制信号，分别

为 4ASK、BPSK、QPSK、OQPSK、8PSK、16QAM、32QAM 和 64QAM。信号的基带波形域为

$$s(t) = \sum_n a_n g(t - nT) + v(t) \quad (1)$$

其中， $v(t)$ 表示信道环境噪声，本实验选择加性高斯白噪声； $g(t)$ 表示包括成型滤波器、信道滤波器以及匹配滤波器的等效滤波器； a_n 表示发送端符号序列。

对于 ASK 类信号，有

$$a_n = I_n + jQ_n \quad (2)$$

其中， $Q_n = 0$ ， $I_n \in \{1, 2, \dots, M-1\}$ ， M 为调制进制数。

对于 PSK 类信号，有

$$a_n = e^{j\frac{(2i+1)\pi}{M}}, i = 0, 1, \dots, M-1 \quad (3)$$

对于 QAM 类信号，有

$$a_n = I_n + jQ_n \quad (4)$$

其中， $I_n, Q_n \in \left\{ 2i - \frac{M}{4} + 1, i = 0, 1, \dots, \frac{M}{4} - 1 \right\}$ ， M 为调制进制数。

2.2 电磁信号的图像域具象方法

实验首先通过正交解调获得电磁信号的 I/Q 两路原始数据，进一步利用 I/Q 数据绘制电磁信号的星座图以将电磁信号具象为二维图像，并根据归一化点密度对星座图上色以实现特征增强。以 QPSK 为例说明电磁信号图像域具象方法，如图 1 所示。将 I/Q 两路信号绘制在以同向分量为横轴，正交分量为纵轴的坐标系中，得到 QPSK 信号的星座图，如图 1 左半部分所示。对 QPSK 每一个信号点计算其归一化点密度，并根据其归一化点密度的高低对

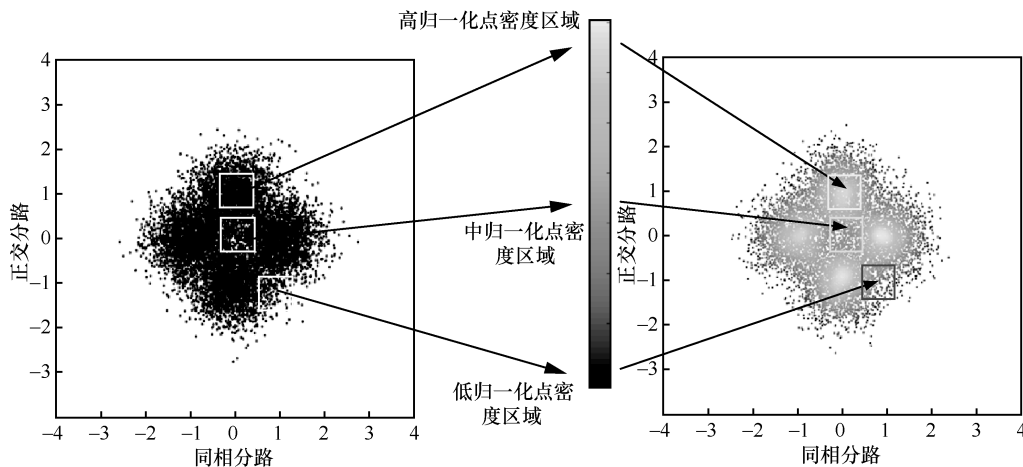


图 1 电磁信号图像域具象方法示意

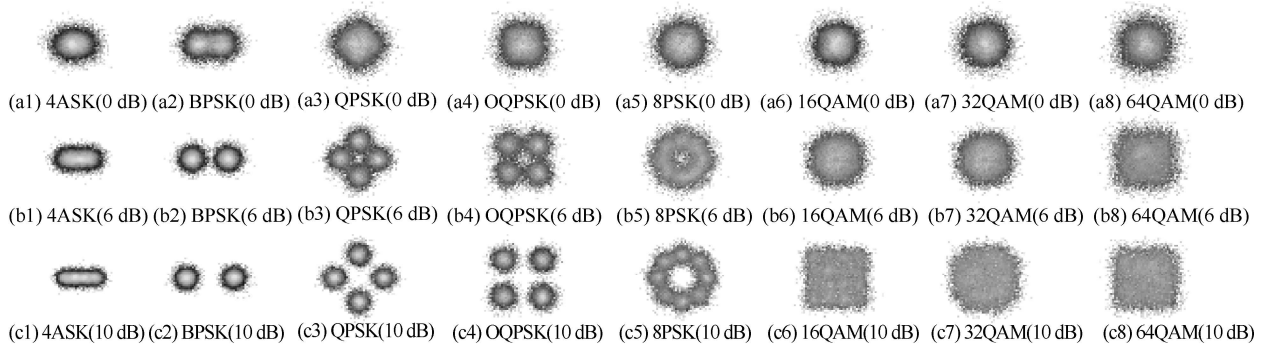


图 2 不同信噪比噪声下的数据处理效果

星座图进行上色以实现特征增强，上色的方法与效果如图 1 右半部分所示。从图 1 中可以看出，经过预处理后的星座图中每个点不再是独立以及等信息量的。这种处理方法给星座图中增加了时间累积的特征，数据特征维度更高，凝聚了更多的调制信号先验知识，增强其可分性，实现了特征增强的目的。

计算某一点的归一化点密度时，以该点为中心绘制的边长为 $2r$ 的正方形内，其他信号点的个数与总共接收到信号点的个数的比值为该点的归一化点密度。第 i 点的归一化点密度 $d(i)$ 为

$$d(i) = \frac{\sum_{j=1}^N f[|h(i) - h(j)| < r \& \& |v(i) - v(j)| < r]}{N} \quad (5)$$

其中， $h(i)$ 表示获取第 i 个点的横坐标值， $v(i)$ 表示获取第 i 个点的纵坐标值， N 表示点的总数， $f(x)$ 函数为

$$f(x) = \begin{cases} 1, & x = \text{true} \\ 0, & x = \text{false} \end{cases} \quad (6)$$

由此可得，当信噪比为 0 dB、6 dB 以及 10 dB 的高斯白噪声时，8 种信号的数据处理效果如图 2 所示。

3 二值化深度神经网络

3.1 二值化深度神经网络原理

本文研究利用 BNN 完成图像域的电磁信号调制识别问题。BNN 来源于神经网络低比特量化的思想的极端情况，其思想为神经网络在逐层计算过程中，参数的量化空间不需要非常完备也可以对问题模型进行拟合，从而完成正确的分类或检测。而在模型规模方面，原有的 32 bit 浮点数据被 1 bit 数据代替，可实现参数总占用减小近 32 倍的压缩效果。

另外，位宽为 1 的参数之间的计算可以使用简单的位操作来完成，可省去 32 bit 浮点型数据乘法的烦琐计算，实现加速效果。

本实验采用的神经网络二值化方法为神经网络朴素二值化分类中的确定性二值化方法^[22]。在确定性二值化方法中，前向推理时，网络中二值化层的权重与输入被量化为 ± 1 。量化规则为，大于或等于 0 的权重或输入被量化为 +1，而小于 0 的被量化为 -1，如式(7)所示。

$$w_b = \begin{cases} +1, & w \geq 0 \\ -1, & w < 0 \end{cases} \quad (7)$$

其中， w 为网络中原始的全精度参数， w_b 为经过二值化后得到的二值化参数。代码实现时可以用取符号函数 $\text{sign}(\cdot)$ 对参数和输入取得符号，从而实现二值化。 $\text{sign}(\cdot)$ 函数曲线如图 3 所示。

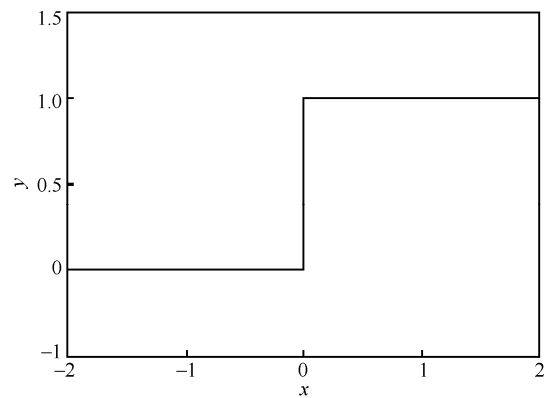


图 3 $\text{sign}(\cdot)$ 函数曲线

在计算机实际存储中，置 0 bit 代表参数值为 -1，置 1 bit 代表参数值为 +1，其运算的权值真实数值和计算机变量的乘法运算真值如表 1 和表 2 所示。从表 1 中可以看出，经过以上量化规则得到的 ± 1 权重乘法运算与计算机 1 bit 变量的同或

运算等效,这就将 FNN 中的 32 bit 浮点乘法运算简化为 1 bit 的同或运算,实现对网络卷积运算中的乘法进行加速的效果。

表 1 权值实际运算真实数值

W_a	W_b	$W_a W_b$
1	1	1
1	-1	-1
-1	1	-1
-1	-1	1

表 2 计算机运算真值

V_a	V_b	$V_a \odot V_b$
1	1	1
1	0	0
0	1	0
0	0	1

在反向传播求解梯度进行优化过程中,例如随机梯度下降等优化算法的实现都是通过每次对网络进行小步微调,逐渐逼近整个网络的最优解。而 BNN 中,二值参数的值只有+1 和-1,这意味着量化值的每一次更新都会使网络状态产生很大的改变,增加了对 BNN 中参数更新标准的设定。因此,每次训练时依然需要对原始全精度参数进行优化更新,而在前向推导时进行量化参数操作。

sign(·) 函数的梯度在大部分情况下都为 0,这将导致二值化参数与原始参数之间的梯度传播被阻断。为了避免梯度阻断问题,将二值化权重的梯度跳过 sign(·) 函数直接传播给原始的全精度参数,使二值化参数的梯度作为原始参数的梯度近似,如图 4 所示。Identity 函数的计算式为 $y = x$, 其正向传播相当于不具任何处理的直连函数,而其函数梯度恒等于 1,起到了将二值化参数的梯度直接回传的作用。将 Identity 函数与随机梯度下降等优化

算法结合,即可实现原始全精度参数的逐步更新。

3.2 用于电磁信号调制识别的二值化深度神经网络

本文设计并构建了用于电磁信号调制识别的二值化深度神经网络模型,在模型整体结构上综合参考经典的 AlexNet 网络结构^[23],先采用卷积层提取特征,然后采用全连接层对特征分类。

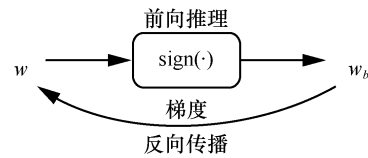


图 4 BNN 训练极端的前/反向传播示意

本文分析数据特征设计网络中每层的类型。从电磁信号的图像域表现形式可以看出,图像中每个像素点的量化空间为 8 bit 整型,并且其中具有很多明显的细粒度的信号点的分布特征,但是不存在计算机视觉分类问题中的目标被遮盖、平移、旋转等问题,所以在网络中仅设计第一层为全精度网络层来提取电磁信号图像域中的细粒度浅层特征,其参数量化空间为 32 bit 浮点型;利用连续的三层二值化网络层综合提取不同深度的粗粒度特征。受到本文所选的二值化推理框架 DABNN (daqexian's accelerated binary neural network) 未兼容全连接网络的限制,实验使用卷积核尺寸为 1×1 的卷积层替代全连接层^[24]。在 1×1 的卷积层与特征提取部分之间插入过渡层以完成数据尺度的衔接。网络输出层得到的结果经过 Softmax 层,得到对不同分类的判别概率分布。图 5 为实验构建的用于电磁信号识别的 BNN 结构。

如图 6 所示,除输出层外,网络中的层均由三部分构成。Float/Binary Layer 用于特征提取或分类; BatchNormal Layer 用于将数据分布归一化以加快训练; MaxPooling Layer 用于实现特征聚合,增强网络的抗噪声性能。

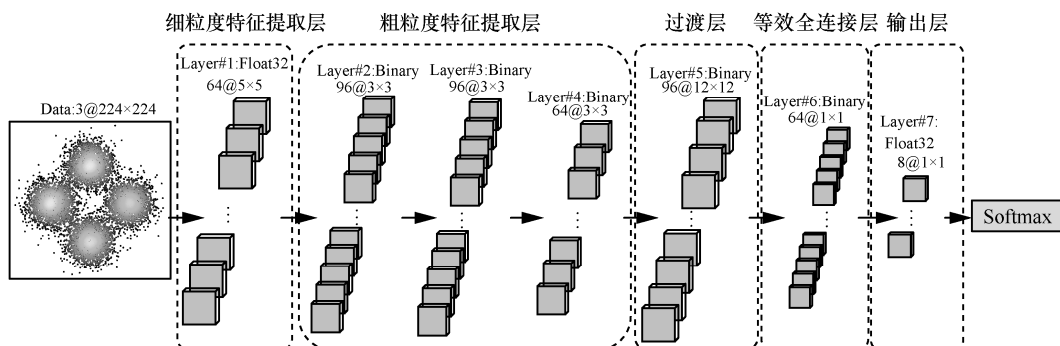


图 5 BNN 结构

需要说明的是，在卷积运算的累加运算过程中，因为 1 bit 数据无法表达求和运算的结果，所以在累加求和过程中的中间变量仍然需要用更高位数的数据类型。本实验所设计的 BNN 中用到的 BatchNormal layer 和 MaxPooling Layer 中涉及更精确的运算同样是二值化参数所无法表达的。综合考虑各层对中间变量的数据类型的要求以及编程复杂度，统一将中间变量的数据类型定义为 32 bit 浮点型数据。本实验构建二值化网络的前向传播与反向传播算法如算法 1 所示。其中， (a_0, a^*) 为输入批数据与标签， W 为原始全精度参数， θ 为批归一化层的参数集， η 为学习率， C 为损失函数， L 为二值化网络的层数， \otimes 为二值/全精度卷积运算， $\text{Binarize}(\cdot)$ 为参数的二值化处理， $\text{BatchNorm}(\cdot)$ 为批归一化处理， $\text{BackBatchNorm}(\cdot)$ 为批归一化处理的反向传播， $\text{Update}(\cdot)$ 为相应优化器的参数更新过程以及此优化器规则或人为设定规则的学习率更新过程。

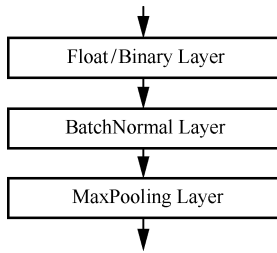


图 6 网络块组成结构

算法 1 BNN 训练算法

前向传播

- 1) for $k = 1$ to L do
- 2) if $k = 1$ or $k = L$ then
- 3) $s_k \leftarrow a_{k-1} \otimes W_k$
- 4) $a_k \leftarrow \text{BatchNorm}(s_k, \theta_k)$
- 5) else then
- 6) $W_k^b \leftarrow \text{Binarize}(W_k)$
- 7) $a_{k-1}^b \leftarrow \text{Binarize}(a_{k-1})$
- 8) $s_k \leftarrow a_{k-1}^b \otimes W_k^b$
- 9) $a_k = \text{BatchNorm}(s_k, \theta_k)$
- 10) end if
- 11) end for

反向传播

- 1) Compute $g_{a_L} = \frac{\partial C}{\partial a_L}$ with (a_0, a^*) and C
- 2) for $k = L$ to 1 do

- 3) $(g_{s_k}, g_{\theta_k}) \leftarrow \text{BackBatchNorm}(g_{a_k}, s_k, \theta_k)$
- 4) if $k = 1$ then
- 5) $g_{a_{k-1}} \leftarrow g_{s_k} W_k^b$
- 6) $g_{W_k} \leftarrow g_{s_k} a_{k-1}^b$
- 7) else if $1 < k < L$ then
- 8) $g_{a_{k-1}} \leftarrow g_{s_k} W_k$
- 9) $g_{W_k} \leftarrow g_{s_k} a_{k-1}$
- 10) end if
- 11) $\theta_k \leftarrow \text{Update}(\theta_k, \eta, g_{\theta})$
- 12) $W_k \leftarrow \text{Update}(W_k, \eta, g_{W_k})$
- 13) $\eta \leftarrow \text{Update}(\eta)$
- 14) end for

4 实验及分析

本文设计并构建 BNN 完成电磁信号调制识别任务，实验选取 8 种常用数字调制信号对本文方法进行验证，分别为 4ASK、BPSK、QPSK、OQPSK、8PSK、16QAM、32QAM 和 64QAM，噪声环境为加性高斯白噪声，信号的 I/Q 波形域为通用调制基带信号，波形图可参考文献[25]。实验构建的数据集中，每个信号样本生成 10 000 个采样点，并利用 2.2 节所述方法进行预处理，得到特征增强的星座图。

数据集规模比例为训练集:交叉验证集:测试集=3:2:2。在训练集中，每个信噪比点下每种调制信号生成 3 000 个信号样本，在测试集与交叉验证集中每种调制信号生成 1 000 个信号样本。实验在 Windows 10 环境下选择 Pytorch 深度学习框架完成 3.2 节中设计的二值化网络的构建。同时，设计具有相同规模的 FNN 作为对照组。实验的 GPU 环境为 GTX2080。

4.1 二值化网络训练分析

本节设计基于 BNN 的电磁信号调制识别实验，考虑到在二值化网络反向传播过程中二值化参数的梯度直接传播给原始参数带来的梯度误差，实验寻找适用于 BNN 的训练优化器。

实验选择 SGD (stochastic gradient descent)、RMSprop (root mean square prop) 与 Adam 这 3 种常用的优化器，在试验性数据集上分别对 FNN 与 BNN 进行训练，评估其训练效果。试验性数据集的调制信号类型、噪声环境与处理方法均与实验数据集相同，仅仅将规模调整为训练集、交叉验证集与测试集均为每个信噪比下每种调制信号 100 个样

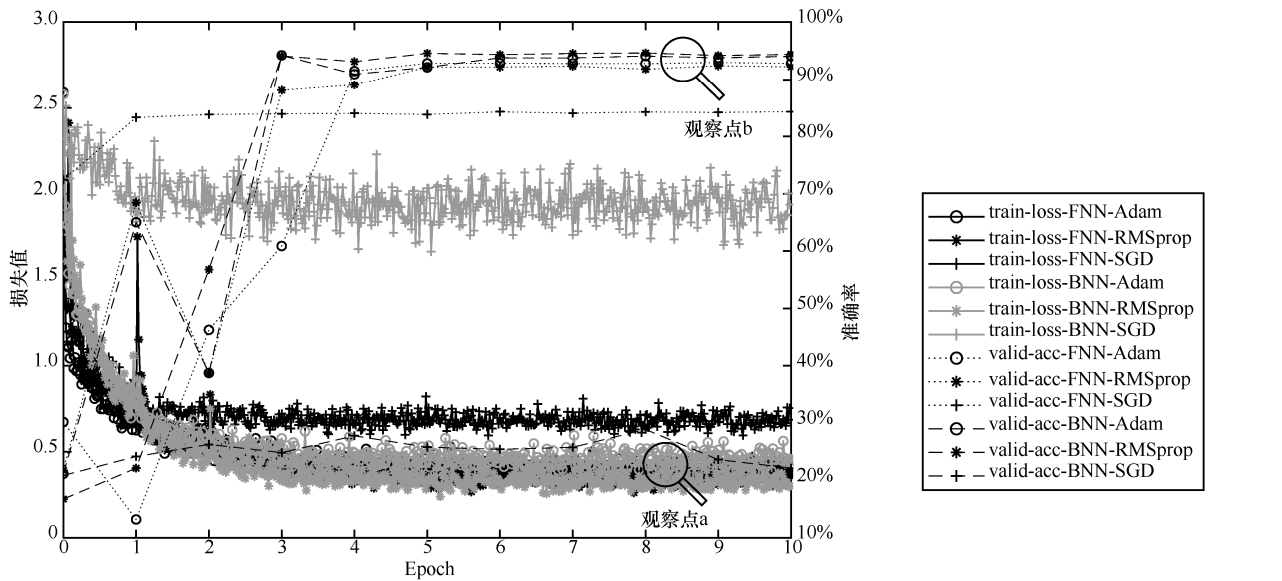
本。训练结果如图 7 所示。

图 7(a)中，实线为损失值变化曲线，对应 y 轴左侧；虚线为准确率变化曲线，对应 y 轴右侧。从图 7(a)中可以看出，在 3 种优化器下，2 种网络在验证集的准确率均随着批数据 (Batch) 的损失值的下降而波动上升，在达到 2~3 个训练周期 (Epoch) 后，训练结果都趋于稳定，稳定阶段的损失值越低，验证集下的准确率越高。结合图 7(b)可以发现，FNN 在 3 种优化器下均可以实现准确率达 80% 以上，训练稳定后，损失值由大到小的顺序为 SGD、RMSprop 和 Adam，而准确率正好相反，说明 Adam 更适合作为 FNN 的训练优化器。结合图 7(c)可以发现，BNN 在 SGD 优化器下基本不能完成训练，而在 Adam 与 RMSprop 优化器下准确率均可以达到 93% 以上，其中 RMSprop 优化器下的结果要高于 Adam

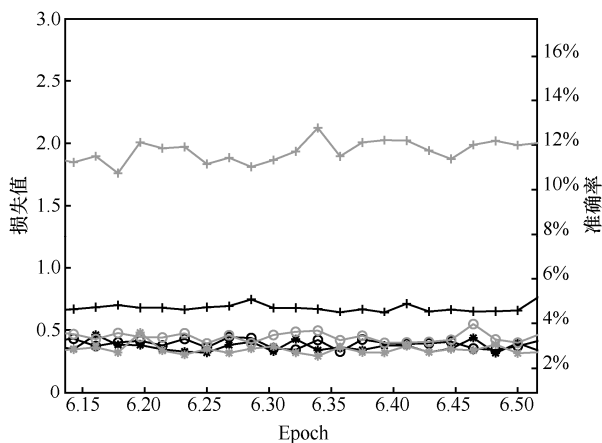
优化器，说明 RMSprop 更适合作为 BNN 的训练优化器。对比 2 种网络可以发现，BNN 除不能在 SGD 优化器下完成训练外，在另外 2 种优化器下训练稳定后，BNN 的损失值均小于 FNN，而准确率均高于 FNN。由此可以得出结论，BNN 相较于 FNN 的训练难度更高。通过选取更合适的优化器训练，可以得到泛化性能更强的模型。

4.2 模型的识别效果分析

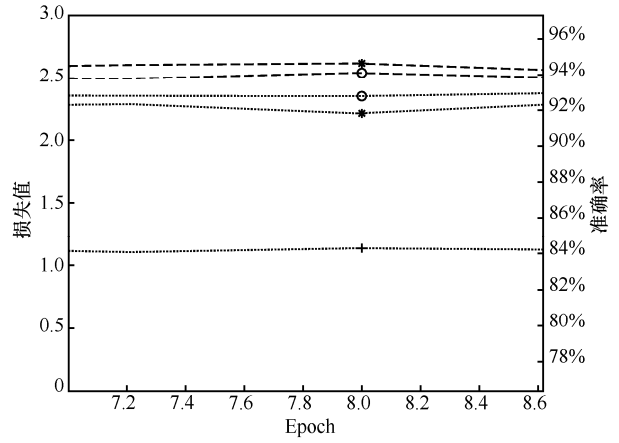
根据 4.1 节的测试结果，本实验分别为 BNN 与 FNN 选取合适的优化器 (即 BNN 使用 RMSprop 优化器，FNN 使用 Adam 优化器) 在实验数据集上进行训练。分析其不同信噪比下对 8 种信号的识别准确率，并将其结果与文献 [6,12] 进行对比。文献 [6,12] 与本文选取的电磁信号的调制类型相同，对比结果如图 8 所示。



(a) 训练整体情况



(b) 观察点a的放大演示



(c) 观察点b的放大演示

图 7 BNN/FNN 在 3 种优化器下的训练过程

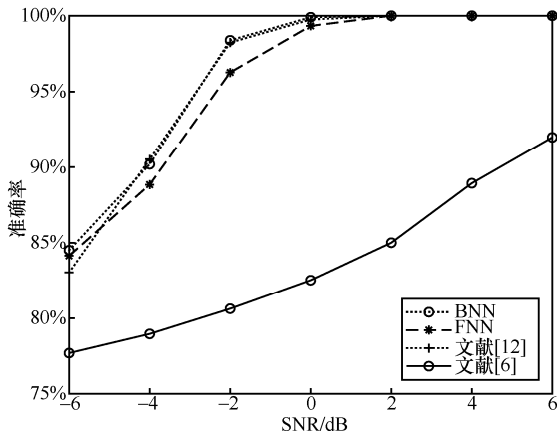


图 8 各信噪比噪声下的结果对比

通过对比分析可以发现，随着信噪比的提升，4 种模型的准确率均呈现上升趋势。当信噪比大于 0 dB 时，本文设计的 BNN 和 FNN 与文献[12]的识别准确率均达到 100%，整体分类性能明显高于文献[6]。对比各个网络的结果可以发现，本实验的 BNN 模型仅在 -4 dB 下的准确率低于文献[12]的结果，在其余信噪比的噪声下均为最高准确率，由此证明了本文基于 BNN 对调制信号识别的有效性；在各个信噪比的噪声下，BNN 的分类准确率均高于 FNN，可以发现在相同规模的网络结构下，BNN 具有更强的泛化性能。

为了进一步分析 BNN 对调制信号的分类性能，本文绘制了模型分类的混淆矩阵，并计算了分类整体的准确率与每个类别下的召回率和精准度，如图 9 所示。图 9 中真实标签与预测标签交叉对应的区域为分类的混淆矩阵，其中，数字为每个分类下的样本数量与样本占测试集的百分比；右下角位置为整体的分类准确率，其代表模型对数据集的综合性能；最下边一排的其余位置为每个分类的召回率，其代表模型对这个类数据的分类性能；最右边一列的其余位置为每种分类的精准度，其代表模型预测一个数据为该分类的可信度。从图 9 中可以看出，模型在选定的信噪比范围内对数据集的综合识别率为 96.1%，对 4ASK、BPSK、16QAM 这 3 种调制信号实现了 100% 的识别，而分别在 QPSK 与 OQPSK 和 32QAM 与 64QAM 之间产生轻微混淆。另外，模型对于 8PSK 信号的分类性能较差，其置信度与精准度都明显低于其他调制类型，容易产生混淆的类别为 QPSK 与 OQPSK。从整体情况可以看出，BNN 模型可以准确地对低阶调制类型进行分类，对不同调制类型的调制信号不易产生混淆，而对相同调制类型下不同阶调制信号容易产生混淆。

4.3 模型规模及运行时间分析

建立在 BNN 对电磁信号调制识别的可行性基础上，本文进一步从模型规模以及运行速度两方面对模型的高效性进行评估。实验选择的测试环境为京东 AI 的 DABNN，其为针对 ARM 指令集高度优化的二值化网络推理框架^[24]；选取的平台为单一 ARM 内核、主频 1.5 GHz 的树莓派 4B。

4ASK	7 000 12.5%								100% 0.0%
BPSK		7 000 12.5%							100% 0.0%
QPSK			6 478 11.6%	58 0.1%	534 1.0%				91.6% 8.4%
OQPSK				52 0.1%	6475 11.6%	562 1.0%			91.3% 8.7%
8PSK			470 0.8%	467 0.8%	5 904 10.5%				86.3% 13.7%
16QAM						7 000 12.5%			100% 0.0%
32QAM							6 988 12.5%	6 0.0%	99.9% 0.1%
64QAM							12 0.0%	6 994 12.5%	99.8% 0.2%
	100% 0.0%	100% 0.0%	92.5% 7.5%	92.5% 7.5%	84.3% 15.7%	100% 0.0%	99.8% 0.2%	99.9% 0.1%	96.1% 3.9%
	4ASK	BPSK	QPSK	OQPSK	8PSK	16QAM	32QAM	64QAM	

图 9 信噪比为 -6~6 dB 噪声下的 BNN 分类的混淆矩阵

实验将模型文件的大小作为模型规模的评价指标，将模型实际部署后的运行时间作为运行速度的评价指标。为了避免偶然因素影响运行时间的测量，本实验设计在开机稳定运行 30 min 后连续测试 100 次，求运行时间的平均数作为测试结果。BNN 与 FNN 的测试结果如图 10 所示。

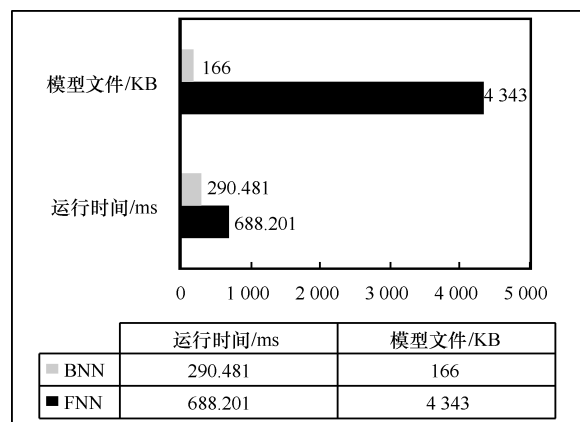


图 10 相同规模 BNN 与 FNN 的模型大小与运行时间比较

通过实验结果可以发现，FNN 模型文件大小是 BNN 的 26.16 倍，分析网络规模可以发现，在二值

化层中每一个参数由原来的 32 bit 浮点型数据缩减到 1 bit, 而 BNN 中的第一层与输出层仍然保持全精度层不变, 综合得到 26.16 倍的压缩率是合理的。另一方面, FNN 的运行时间是 BNN 的 2.37 倍, 本实验设计的网络为 7 层网络, 其中 5 层为二值化层。模型运行时间的压缩倍数会随着二值化层的占比增加而增加。综合以上结果, 相较于 FNN, BNN 可以作为更高效可行的电磁信号调制识别方案。

5 结束语

边缘智能信号识别是 6G “万物互联” 时代的至关重要的一项技术。本文针对该问题, 总结了已有的研究成果, 提出了一种基于 BNN 的信号图像域智能识别方案。所提方案通过绘制信号的星座图将信号从波形域转化为图像域, 进而用 BNN 对图像域信号进行识别。本文以电磁信号调制识别问题为例对该方案进行验证, 在信噪比为 $-6 \sim 6$ dB 的加性高斯白噪声环境下获得 96.1% 的准确率, 且在不同信噪比下的识别准确率均高于同规模 FNN, 整体分类性能优于相同问题的已有算法, 证明了方案的可行性。其次, 实验将所提方案部署于具有单一 ARM 内核、主频 1.5 GHz 的树莓派 4B 平台上, 模型文件仅有 166 KB, 平均运行时间仅为 290.481 ms, 证明了所提方案的高效性。

综上所述, 本文提出了一种兼具可行性与高效性的边缘电磁信号调制识别方案, 但是该方案仍具有很多值得提升与研究之处。例如, 本文的信号预处理部分利用信号处理的先验知识, 将电磁信号的高维抽象的波形数据具象化为图像域, 实现了对调制信号的特征增强, 也增强了数据的可解释性与可分性, 但这种方法会引入大量的参数, 如何实现电磁信号的波形数据进行端到端的智能处理是下一步的研究内容。本文相信 BNN 仍然存在网络参数以及连接的冗余, 进一步对网络结构进行精简或将得到更为小巧高效的网络模型; 其次, 卷积神经网络自动提取数据中有助于分类的高维特征, 但电磁信号的卷积层特征是抽象的, 其可理解性难以与计算机视觉问题相比, 如何对电磁信号的特征进行可视化或对其机理进行解释也是未来的研究方向。另外, 本文所选用的调制信号类型以及噪声范围具有一定局限性, 应进一步扩大研究范围以发现问题, 对该方案提出改善。

参考文献:

- [1] YOU X H, WANG C X, HUANG J, et al. Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts[J]. SCIENCE CHINA Information Sciences, 2021,64(1): 1-84.
- [2] 张平, 牛凯, 田辉, 等. 6G 移动通信技术展望[J]. 通信学报, 2019,40(1):141-148.
ZHANG P, NIU K, TIAN H, et al. Technology prospect of 6G mobile communications[J]. Journal on Communications, 2019, 40(1): 141-148.
- [3] JAGANNATH A, JAGANNATH J, MELODIA T. Redefining wireless communication for 6G: signal processing meets deep learning[J]. arXiv Preprint, arXiv:2004.10715, 2020.
- [4] LI B, LI S, HOU J, et al. A Bayesian approach for adaptively modulated signals recognition in next-generation communications[J]. IEEE Transactions on Signal Processing, 2015, 63(16):4359-4372.
- [5] WANG D S, ZHANG M, LI Z, et al. Modulation format recognition and OSNR estimation using CNN-based deep learning[J]. IEEE Photonics Technology Letters, 2017, 29(19): 1667-1670.
- [6] 查雄, 彭华, 秦鑫, 等. 基于多端卷积神经网络的调制识别方法[J]. 通信学报, 2019, 40(11): 30-37.
ZHA X, PENG H, QIN X, et al. Modulation recognition method based on multi-inputs convolution neural network[J]. Journal on Communications, 2019, 40(11): 30-37.
- [7] 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.
- [8] 吴佩军, 侯进, 吕志良, 等. 基于卷积神经网络的多进制相位调制信号识别算法[J]. 计算机应用与软件, 2019, 36(11): 202-209.
WU P J, HOU J, LYU Z L, et al. Multi-phase modulation signal recognition algorithm based on convolutional neural network[J]. Computer Applications and Software, 2019, 36(11): 202-209.
- [9] JI X F, WANG J, LI Y, et al. Modulation recognition in maritime multipath channels: a blind equalization-aided deep learning approach[J]. China Communication, 2020,17(3):12-25.
- [10] XU Y, LI D Z, WANG Z Y, et al. A deep learning method based on convolutional neural network for automatic modulation classification of wireless signals[J]. Wireless Networks, 2019,25(7): 3735-3746, .
- [11] 桂冠, 王禹, 黄浩. 基于深度学习的物理层无线通信技术: 机遇与挑战[J]. 通信学报, 2019, 40(2): 19-23.
GUI G, WANG Y, HUANG H. Deep learning based physical layer wireless communication techniques: opportunities and challenges[J]. Journal on Communications, 2019, 40(2): 19-23.
- [12] TU Y, LIN Y. Deep neural network compression technique towards efficient digital signal modulation recognition in edge device[J]. IEEE Access, 2019, 7: 58113-58119.
- [13] HU H, PENG R, TAI Y W, et al. Network trimming: a data-driven neuron pruning approach towards efficient deep architectures[J]. arXiv Preprint, arXiv:1607.03250, 2016.
- [14] LIN Y, TU Y, DOU Z. An improved neural network pruning technology for automatic modulation classification in edge devices[J]. IEEE Transactions on Vehicular Technology, 2020, 69(5): 5703-5706.

- [15] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks[C]//International Conference on Engineering Applications of Neural Networks. Piscataway: IEEE Press, 2016: 213-222.
- [16] 陶冠宏, 廖开升, 周林. 一种基于深度学习的电磁信号建模与调制识别新方法[J]. 电子信息对抗技术, 2019, 34(5): 10-15.
TAO G H, LIAO K S, ZHOU L. A novel signal modeling and modulation classification method with deep learning[J] Electronic Information Countermeasure Technology, 2019, 34(5): 10-15.
- [17] 周鑫, 何晓新, 郑昌文. 基于图像深度学习的无线电信号识别[J]. 通信学报, 2019, 40(7): 114-125.
ZHOU X, HE X X, ZHENG C W. Radio signal recognition based on image deep learning[J]. Journal on Communications, 2019, 40(7): 114-125.
- [18] DETTMERS T. 8 bit approximations for parallelism in deep learning[J]. arXiv Preprint, arXiv: 1511.04561, 2015.
- [19] GUPTA S, AGRAWAL A, GOPALAKRISHNAN K, et al. Deep learning with limited numerical precision[C]//International Conference on Machine Learning. New York: ACM Press, 2015: 1737-1746.
- [20] ZHU C, HAN S, MAO H, et al. Trained ternary quantization[J]. arXiv Preprint, arXiv:1612.01064, 2016.
- [21] COURBARIAUX M, BENGIO Y, DAVID J P. Binaryconnect: training deep neural networks with binary weights during propagations[C]// Advances in Neural Information Processing Systems. New York: ACM Press, 2015: 3123-3131.
- [22] QIN H, GONG R, LIU X, et al. Binary neural networks: a survey[J]. Pattern Recognition, 2020, 105: 1737-1746.
- [23] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. Massachusetts: MIT Press, 2012: 1097-1105.
- [24] ZHANG J, PAN Y, YAO T, et al. DABNN: a super fast inference framework for binary neural networks on arm devices[C]//Proceedings of the 27th ACM International Conference on Multimedia. New York: ACM Press, 2019: 2272-2275.
- [25] 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 Sig-

nal Processing, 2018, 12(1): 168-179.

[作者简介]



张思成（1996-），男，山东临沂人，哈尔滨工程大学博士生，主要研究方向为基于深度学习的智能电磁信号处理。



林云（1980-），男，黑龙江哈尔滨人，哈尔滨工程大学在站博士后，哈尔滨工程大学教授、博士生导师，主要研究方向为人工智能、深度学习、信号识别和智能信息对抗等。



涂涯（1994-），男，湖北十堰人，哈尔滨工程大学博士生，主要研究方向为信号处理、深度学习等。



Shiwen Mao（1971-），男，湖南岳阳人，博士，奥本大学教授，主要研究方向为无线网络、多媒体通信和智能电网。