基于轻量化分布式学习的自动调制分类方法

杨洁，董标，付雪，王禹，桂冠
（南京邮电大学通信与信息工程学院，江苏 南京 210003）

摘 要：为了解决集中式学习存在的问题，提出了一种基于轻量化网络的分布式学习方法。分布式学习利用边缘设备进行本地训练和模型权重共享的方法训练同一个全局模型，既充分利用了各边缘设备的训练数据，又避免了边缘设备数据泄露。轻量化网络是一种由多个轻量化神经网络块堆叠而成的深度学习模型，相较于传统的深度学习模型，轻量化网络以较低的空间复杂度和时间复杂度实现较高的调制分类性能，有效地解决了分布式学习在实际部署中存在的边缘设备算力不足、存储空间有限及通信开销较高的问题。实验结果表明，基于分布式学习的自动调制信号分类技术在 RadioML.2016.10A 数据集的分类准确率为 62.41%，相比于集中式学习，分类准确率仅降低了 0.68%，训练效率提高了近 5 倍。实验结果也证明了在分布式学习下，部署轻量化网络可以有效降低通信开销。

关键词：自动调制分类；分布式学习；轻量化网络；深度学习
中图分类号：TP391
文献标志码：A
DOI: 10.11959/j.issn.1000−436x.2022145

Lightweight decentralized learning-based automatic modulation classification method

YANG Jie, DONG Biao, FU Xue, WANG Yu, GUI Guan
College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Abstract: In order to solve the problems in centralized learning, a lightweight decentralized learning-based AMC method was proposed. By the proposed decentralized learning, a global model was trained through local training and model weight sharing, which made full use of the dataset of each communication nodes and avoided the user data leakage. The proposed lightweight network was stacked by a number of different lightweight neural network blocks with a relatively low space complexity and time complexity, and achieved a higher recognition accuracy compared with traditional DL models, which could effectively solve the problems of computing power and storage space limitation of edge devices and high communication overhead in decentralized learning based AMC method. The experimental results show that the classification accuracy of the proposed method is 62.41% based on RadioML.2016.10 A. Compared with centralized learning, the training efficiency is nearly 5 times higher with a slight classification accuracy loss (0.68%). In addition, the experimental results also prove that the deployment of lightweight models can effectively reduce communication overhead in decentralized learning.

Keywords: automatic modulation classification, decentralized learning, lightweight network, deep learning

0 引言

物联网体系中众多边缘设备广泛连接，实现了设备之间的无缝通信。然而，随着终端数量的增加，大量的中间数据存储在边缘设备中，当边缘设备受到外部恶意攻击时，设备中的数据有丢失或外泄的风险。因此，有效识别并拦截外部恶意攻击是物联网设备安全部署的基础[1-7]。自动调制分类（AMC, automatic

收稿日期：2022−04−18；**修回日期：**2022−06−20
通信作者：桂冠，guiguan@njupt.edu.cn
基金项目：科技创新 2030“新一代人工智能”重大基金资助项目（No.2021ZD0113003）
Foundation Item: The National Key Research and Development Program of China (No.2021ZD0113003)

modulation classification）部署在通信系统的接收机端，可实现自动识别不同种类调制信号的功能[8-10]。因此，AMC 是一种识别物理层恶意攻击的重要方式。

传统的 AMC 方法包括基于似然比的方法和基于特征的方法。基于似然比的方法计算复杂度高，并且需要通信系统的信道状态信息。基于特征的方法虽然计算复杂度相对较低，但是需要依靠专业知识构建复杂的特征工程。近年来，深度学习技术在图像领域和自然语言处理领域取得了重大突破[11]，因此许多研究者尝试将深度学习技术应用于 AMC。然而，现有的基于深度学习的 AMC 研究普遍采用集中式学习进行优化[12-14]，即每台边缘设备将本地数据集上传到一台中心设备训练，然后边缘设备从中心设备下载训练好的全局模型权重。显然，集中式学习有如下缺点：1) 本地设备上传数据集到中心设备，本地设备的数据隐私安全无法得到保证；2) 海量的数据汇聚到中心设备，给中心设备带来巨大的存储压力；3) 中心设备对海量数据进行训练，计算压力较大且训练效率低。

针对集中式学习出现的问题，研究者尝试采用分布式学习策略[15-20]。分布式学习中，本地边缘设备与远程中心设备之间进行模型权重信息交互而非数据共享，避免了本地设备的数据外泄，也减轻了中心设备的存储压力。通过多台本地边缘设备协同训练的方式可以有效减轻中心设备的计算压力，从而缩短训练时间。Wang 等[20]提出了一种基于分布式学习框架训练卷积神经网络（CNN, convolutional neural network）的自动调制信号分类方法，实验结果证明分布式训练框架可以提高训练效率。但是在分布式训练框架中部署 CNN 存在如下缺点：1) CNN 较大的模型权重导致分布式训练中因模型权重交互带来的通信开销剧增；2) CNN 较高的模型复杂度给边缘设备的有限算力提出挑战。为了保证分类性能，同时进一步降低通信开销和模型复杂度，有必要根据信号特征设计网络。Xu 等[12]将信号分解为多信道输入，充分提取信号同相分量和正交分量中的特征，以较小的模型权重和较低的模型复杂度取得了较好的分类性能。Zhang 等[13]将信号预处理为幅度相位形式，充分提取信号中的时间特征和空间特征，提高了分类性能，模型权重和复杂度与传统的 CNN 相比有所降低。为了进一步降低模型权重和复杂度，本文尝试在分布式学习框架上部署更轻量的网络。

本文的主要贡献总结如下。1) 提出了一种基于分布式学习的轻量化网络，本文将其命名为 MCMBNN，旨在解决集中式学习下数据隐私外泄、中心设备存储压力大和训练效率低等问题。2) 在分布式学习框架下部署不同的网络进行对比实验，结果证明了 MCMBNN 可以有效降低通信开销，同时具备较好的分类性能。

1　问题描述

1.1　信号模型

本文使用的基于单载波的信号模型如式(1)所示。

$$u(k)=\lambda \mathrm{e}^{\mathrm{j}\left(\frac{2\pi f_0 k}{K}+\theta\right)}\sum_{l=0}^{L-1}h[l]q[k-l-\tau](k) \tag{1}$$

其中，$k\in\{0,1,\cdots,K-1\}$，$u(k)$ 表示接收机收到的调制信号，λ 表示信道增益，f_0 表示载波偏移，θ 表示相位偏移，$h[l]$ 表示瑞利衰弱信道脉冲响应，L 表示信道脉冲响应的长度，K 表示信号采样点数，$\sigma(k)$ 表示加性白高斯噪声。接收信号被分解为同相分量 $\boldsymbol{I}$ 和正交分量 $\boldsymbol{Q}$ 并输入神经网络中，如式(2)所示。

$$\begin{aligned}\boldsymbol{I}&=\{\mathrm{Re}[u(k)]\}_{k=0}^{K-1}\\ \boldsymbol{Q}&=\{\mathrm{Im}[u(k)]\}_{k=0}^{K-1}\end{aligned} \tag{2}$$

1.2　基于深度学习的自动调制分类系统框架

基于深度学习的自动调制分类系统框架如图 1 所示，深度学习网络在系统中充当一个分类器。设收到的调制信号的类别为 $D=\{d_j, j=0,1,\cdots, J-1\}$，其中，$J$ 表示信号的类别数目。基于深度学习的自动调制分类的决策式为

$$d_j=F_{d_j\in D}\left([\boldsymbol{I};\boldsymbol{Q}],\boldsymbol{P}\right) \tag{3}$$

其中，F 表示深度学习作为分类器的函数，$\boldsymbol{P}$ 表示模型参数。此外，选用交叉熵函数作为损失函数，如

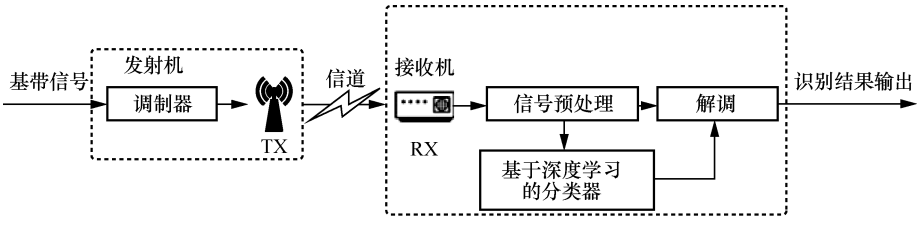


图 1　基于深度学习的自动调制分类系统框架

式(4)所示，利用 L2 正则化作为平衡项的惩罚因子，防止模型出现过拟合。

$$E=-\frac{1}{R}\sum_{r=1}^{R-1}y_r\log\left[F\left([\boldsymbol{I};\boldsymbol{Q}],\boldsymbol{P}\right)\right]+\mu K\left(F\left([\boldsymbol{I};\boldsymbol{Q}],\boldsymbol{P}\right)\right) \tag{4}$$

其中，R 表示训练样本的规模，y_r 表示样本的实际标签，$K(\cdot)$ 表示惩罚函数，μ 表示惩罚函数的平衡系数。

1.3 传统的集中式学习

集中式学习框架如图 2 所示。每个边缘设备都有一个本地数据集 $M_n=\{(X_1,Y_1),(X_2,Y_2),\cdots,(X_{S_n},Y_{S_n})\}$，$n$ 表示第 n 个边缘设备，S_n 表示第 n 个边缘设备中数据集的规模，因此全局数据集可以表示为 $M_{\mathrm{g}}=M_1\cup M_2\cup\cdots\cup M_N$，全局数据集的规模可以表示为 $S_{\mathrm{g}}=\sum_{n=1}^{N}S_n$，另外 $M_i\cap M_j=0$，$i\neq j$。

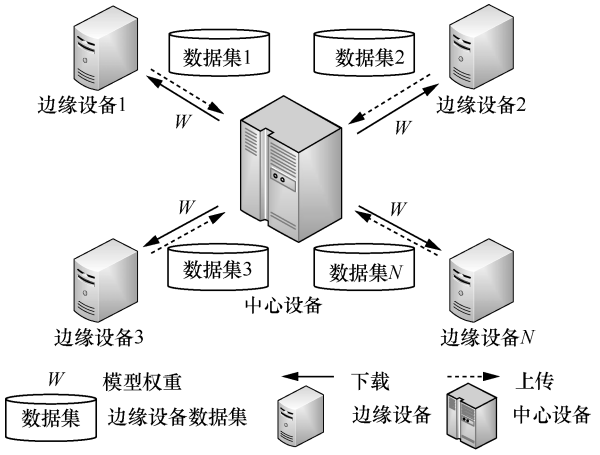


图 2 集中式学习框架

在集中式学习中，各个边缘设备将本地数据集上传到远程的中心设备（在上传过程中边缘设备数据对外暴露），中心设备收到数据集后开始训练本地模型，这要求中心设备具备足够的存储空间和较强的算力。训练结束后将模型权重 W 共享给每个边缘设备。集中式学习采用最小化经验损失函数准则训练模型，如式(5)所示；同时采用 Adma 优化器来更新权重信息，如式(6)所示。

$$E'=\frac{1}{S_n}\sum_{n=1}^{N}S_nE \tag{5}$$

$$w_t=w_{t-1}-\eta_t\frac{\hat{m}_t}{\sqrt{\hat{v}_t}+\varepsilon} \tag{6}$$

其中，w_t 表示第 t 个训练周期后更新后的模型权重，$\hat{m}_t$ 表示更新后的有偏一阶矩估计，$\hat{v}_t$ 表示更新后的有偏二阶矩估计。

2 本文方法原理

2.1 分布式学习

分布式学习框架如图 3 所示。与集中式学习相比，分布式学习中每个边缘设备利用本地数据集训练模型，将模型权重共享给中心设备。具体地，分布式学习包括以下 4 个步骤。

1) 模型初始化和参数广播

中心设备构建一个模型。首先，初始化模型权重、学习速率、训练周期数和训练批次等参数；然后，中心设备将模型和初始化的参数广播给边缘设备。

2) 边缘设备权重更新和上传

边缘设备先从中心设备下载模型和初始化的

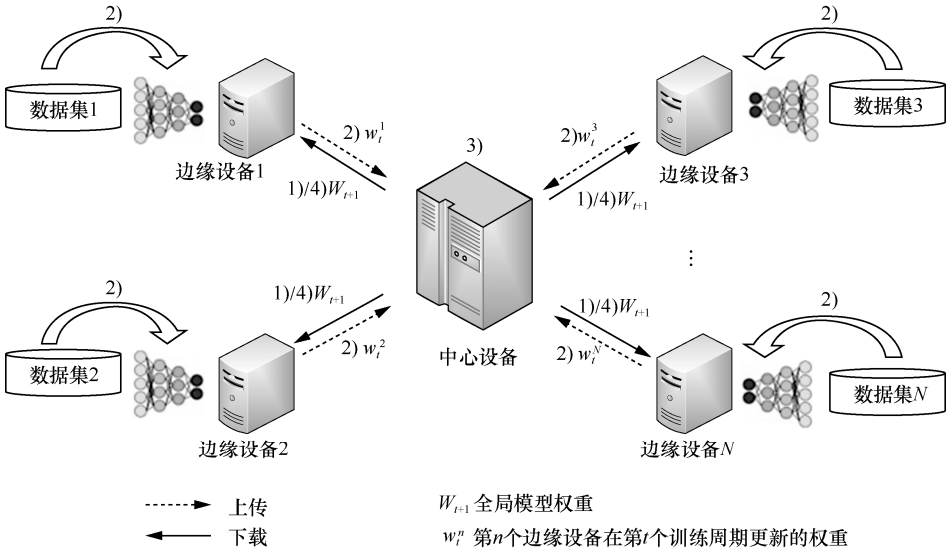


图 3 分布式学习框架

参数，利用本地数据集训练模型；然后根据 Adma 优化更新模型权重。与集中式学习中所有数据被上传到中心设备训练相比，分布式学习采用边缘设备在本地训练的方式，可以缓解中心设备的存储压力和计算压力。

3) 模型聚合

一个训练周期完成后，边缘设备将本地模型权重共享给中心设备，由中心设备对模型权重作加权平均得到全局模型权重，如式(7)所示。

$$W_{t+1}=\frac{\sum_{n=1}^{N}S_n\tilde{w}_t^n}{S_{\mathrm{g}}} \tag{7}$$

其中，W_{t+1}是第 t 个训练周期聚合的全局权重，$\tilde{w}_t^n$是步骤 2)中更新的模型权重。与集中式学习相比，分布式学习通过共享更新的模型权重而非共享数据集的方式，达到保证数据隐私安全的目的。

4) 全局权重更新

边缘设备从中心设备下载全局模型权重代替原来的模型权重，重复步骤 2)～步骤 4)，直到模型收敛。

2.2 轻量化网络

本文在分布式学习的步骤 1) 中部署了轻量化网络 MCMBNN，其整体结构如图 4 所示，网络由相位参数评估模块、空间特征提取模块和时间特征提取模块 GRU(128)组成。

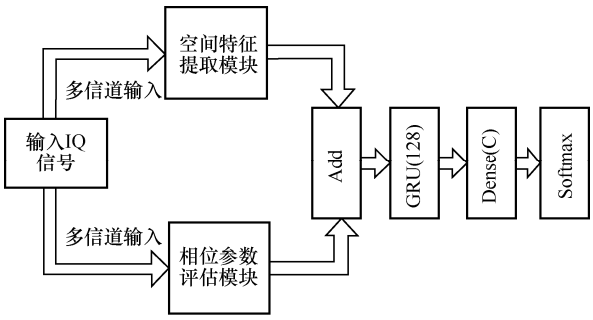


图 4　轻量化网络 MCMBNN 整体结构

1) 多信道输入

因为 I 信道和 Q 信道的信号之间存在极大差异，所以输入信号通过 3 路信道输入轻量化网络，即 IQ 信道、I 信道和 Q 信道，这样可以提取到不同信道间的互补特征，进一步提高分类性能。

2) 相位参数评估模块

信号通过信道时受到噪声的影响，所以输入 IQ 信号通常携带相位偏移信息。

相位参数评估模块的作用是提取相位偏移信息[14]。如图 5 所示，相位参数评估模块由一个 Flatten 层和一个包含单个神经元的 Dense 层组成，原始 IQ 信号通过 Flatten 层实现维度变化以满足 Dense 层的输入要求，经过 Dense 层输出的张量提取到丰富的相位特征，最后用线性激活函数获取相位评估参数ψ。参数评估之后是相位转置，相位转置的计算式为

$$\hat{\boldsymbol{u}}(\boldsymbol{k})=\boldsymbol{u}(\boldsymbol{k})\mathrm{e}^{-\mathrm{j}\psi}=\begin{bmatrix}\boldsymbol{I}\cos\psi+\boldsymbol{Q}\sin\psi\\ \boldsymbol{I}\cos\psi-\boldsymbol{Q}\sin\psi\end{bmatrix} \tag{8}$$

其中，$\hat{\boldsymbol{u}}(\boldsymbol{k})$是相位转置模块的输出。

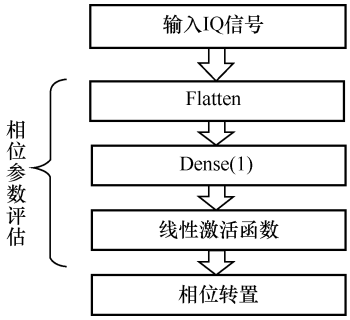


图 5　相位参数评估模块

3) 空间特征提取模块

空间特征提取模块如图 6 所示，3 路信道并行输入不同的卷积块。第一路 IQ 混合信道输入 MC2D-Block1，其结构如图 7(a)所示。MC2D-Block1 采用 2 个非对称卷积核(2,8)、(8,2)和一个(1,1)的卷积核来代替(8,8)对称卷积核，每一个卷积层 Conv2D 后利用 ReLU 层增强网络的非线性，同时避免梯度消失，最后在信道维度进行拼接。

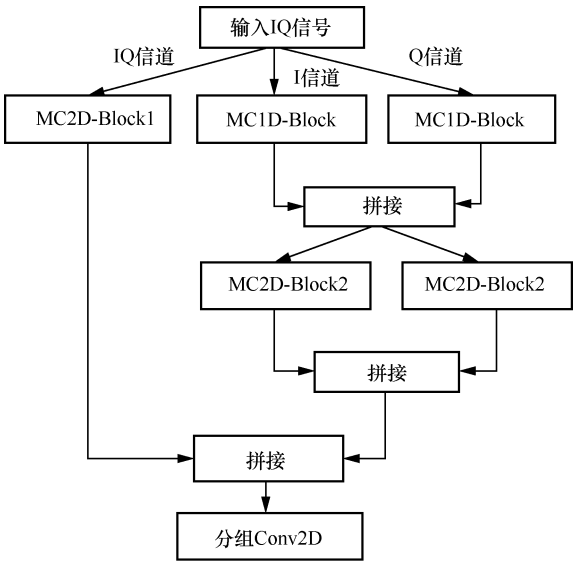


图 6　空间特征提取模块

第二路 I 信道和第三路 Q 信道分别输入 MC1D-Block，如图 7(b)所示。MC1D-Block 的整体

结构和 MC2D-Block1 相似，区别是 MC1D-Block 采用了一维卷积，原因是 I 信道和 Q 信道的输入信号是 2 个一维序列。此外，3 个卷积层使用的卷积核的大小分别是 2、4 和 8，这样设计可以用较少的参数提取相对丰富的空间特征。2 个 MC1D-Block 的输出张量在深度和维度上进行拼接形成新张量，新张量融合了 I 信道和 Q 信道的空间特征。为了提取更加高级的特征，新张量被并行输入 MC2D-Block2。与 MC2D-Block1 相比，MC2D-Block2 采用(1,8)和(8,1)的非对称卷积，MC2D-Block2 的输出张量和 MC2D-Block1 的输出张量在通道维度上拼接，进一步实现特征融合。融合的张量输入分组卷积层，分组卷积层采用 2 个分组的 3×3 卷积核代替标准卷积，实现模型复杂度的降低。

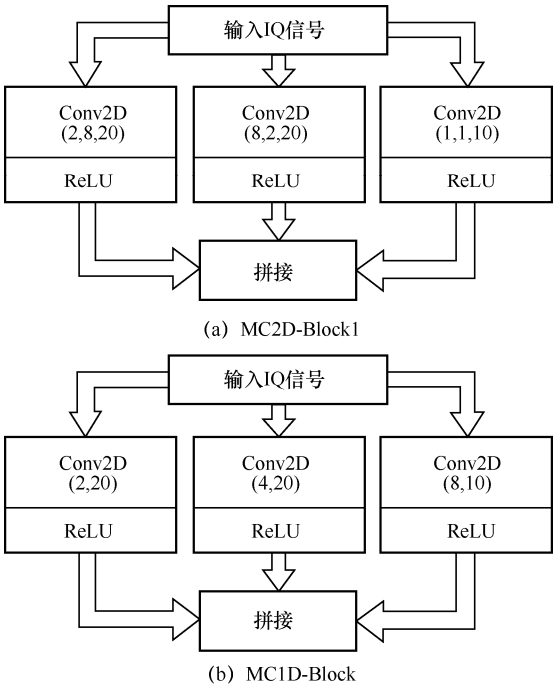


图 7 MC2D-Block1 和 MC1D-Block

4) 时间特征提取模块

MCMBNN 利用 Add 层实现信号相位特征和空间特征的融合，采用包含 128 个神经元的 GRU 层实现信号时间特征的提取。

2.3 轻量化网络的复杂度分析

一般地，模型复杂度包括时间复杂度和空间复杂度[20]，本文中时间复杂度以每秒浮点运算次数（FLOPS, floating-point operations per second）呈现，空间复杂度以参数量的形式呈现。

1) 相位参数评估模块

相位参数评估模块的复杂度集中在 Dense 层，Dense 层的复杂度分别如式(9)和式(10)所示。

$$P_{\text{Dense}} \sim O\left(C_{\text{out}}\left(C_{\text{in}}+1\right)\right) \tag{9}$$

$$F_{\text{Dense}} \sim O\left(2C_{\text{out}}C_{\text{in}}\right) \tag{10}$$

其中，P_{Dense} 和 F_{Dense} 分别是 Dense 层的参数量和 FLOPS，C_{out} 和 C_{in} 分别是输出和输入信道数。相位参数评估模块的 Dense 层仅有一个神经元，因此 $C_{\text{out}}=1$。显然，相位参数评估模块的复杂度和 C_{in} 处于一个量级，可忽略不计。

2) 空间特征提取模块

为了降低模型复杂度，空间特征提取模块部署了非对称卷积核而非对称卷积核。标准卷积的参数量和 FLOPS 分别如式(11)和式(12)所示

$$P_{\text{sta}} \sim O\left(K_{\text{h}}K_{\text{w}}C_{\text{out}}C_{\text{in}}\right) \tag{11}$$

$$F_{\text{sta}} \sim O\left(M_1K_{\text{h}}K_{\text{w}}C_{\text{out}}C_{\text{in}}\right) \tag{12}$$

其中，M_1 是输出特征图的尺寸，K_{h} 和 K_{w} 分别是卷积核的 2 个维度。显然，部署 $K_{\text{h}}\times K_{\text{w}}=8\times2$ 和 $K_{\text{h}}\times K_{\text{w}}=2\times8$ 非对称卷积核的复杂度小于部署 2 个 $K_{\text{h}}\times K_{\text{w}}=8\times8$ 对称卷积核的复杂度。

此外，与文献[3,12,20]不同，MCMBNN 利用分组卷积取代标准卷积进一步提取空间特征，分组卷积的复杂度分别如式(13)和式(14)所示。

$$P_{\text{group}} \sim O\left(K_{\text{h}}K_{\text{w}}C_{\text{out}}C_{\text{in}}\frac{1}{G}\right) \tag{13}$$

$$F_{\text{group}} \sim O\left(M_1K_{\text{h}}K_{\text{w}}C_{\text{out}}C_{\text{in}}\frac{1}{G}\right) \tag{14}$$

其中，G 是分组卷积的分组数。分组卷积和标准卷积的空间复杂度和时间复杂度之比分别如式(15)和式(16)所示。与标准卷积相比，分组卷积的参数量和 FLOPS 降低为原来的 $\frac{1}{G}$。

$$\frac{P_{\text{group}}}{P_{\text{sta}}}=\frac{K_{\text{h}}K_{\text{w}}C_{\text{out}}C_{\text{in}}\frac{1}{G}}{K_{\text{h}}K_{\text{w}}C_{\text{out}}C_{\text{in}}}=\frac{1}{G} \tag{15}$$

$$\frac{F_{\text{group}}}{F_{\text{sta}}}=\frac{M_1K_{\text{h}}K_{\text{w}}C_{\text{out}}C_{\text{in}}\frac{1}{G}}{M_1K_{\text{h}}K_{\text{w}}C_{\text{out}}C_{\text{in}}}=\frac{1}{G} \tag{16}$$

3) 时间特征提取模块

MCMBNN 采用单层的 GRU 模块而非 LSTM 模块作为时间特征提取模块。GRU 模块的参数量和 FLOPS 分别如式(17)和式(18)所示，文献[3,12]中使用的 LSTM 模块参数量和 FLOPS 分别如式(19)和式(20)所示。

$$P_{\mathrm{GRU}}\sim O\left(3\left(E_{\mathrm{s}}(E_{\mathrm{s}}+H_{\mathrm{s}})+E_{\mathrm{s}}\right)\right) \tag{17}$$

$$F_{\mathrm{GRU}}\sim O\left(3\times 2H_{\mathrm{s}}(E_{\mathrm{s}}+H_{\mathrm{s}})\right) \tag{18}$$

$$P_{\mathrm{LSTM}}\sim O\left(4\left(E_{\mathrm{s}}(E_{\mathrm{s}}+H_{\mathrm{s}})+E_{\mathrm{s}}\right)\right) \tag{19}$$

$$F_{\mathrm{LSTM}}\sim O\left(\left(4\times 2H_{\mathrm{s}}(E_{\mathrm{s}}+H_{\mathrm{s}})\right)\right) \tag{20}$$

其中，E_{s} 和 H_{s} 分别是词向量维度和隐藏层节点数。GRU 模块和 LSTM 模块的空间复杂度和时间复杂度之比分别如式(21)和式(22)所示。可以看出，GRU 模块的复杂度是 LSTM 模块的 $\frac{3}{4}$。

$$\frac{P_{\mathrm{GRU}}}{P_{\mathrm{LSTM}}}=\frac{3\left(E_{\mathrm{s}}(E_{\mathrm{s}}+H_{\mathrm{s}})+E_{\mathrm{s}}\right)}{4\left(E_{\mathrm{s}}(E_{\mathrm{s}}+H_{\mathrm{s}})+E_{\mathrm{s}}\right)}=\frac{3}{4} \tag{21}$$

$$\frac{F_{\mathrm{GRU}}}{F_{\mathrm{LSTM}}}=\frac{(3\times 2H_{\mathrm{s}}(E_{\mathrm{s}}+H_{\mathrm{s}})}{(4\times 2H_{\mathrm{s}}(E_{\mathrm{s}}+H_{\mathrm{s}})}=\frac{3}{4} \tag{22}$$

3　实验与评估

3.1　实验设置

实验在 Geforce GTX 2080ti GPU 计算设备上实施。仿真平台是 Tensorflow 1.10 + Keras 2.2.4 深度学习框架，环境采用 Python 3.6。采用的数据集是公开数据集 RadioML.2016.10A，RadioML.2016.10A 共有 220 000 个样本，其调制类别包括{BPSK, 8PSK, CPFSK, GFSK, PAM4, 16QAM, 64QAM, QPSK, AM-DSB, AM-SSB, WBFM}，信噪比为−20～18 dB。本节将所提轻量化网络与目前较先进的几个深度学习网络进行性能对比，包括 MCLDNN[12]、CNN[20]、CNN-LSTM[3]。同时进行了分布式学习和集中式学习的对比实验，观察分类性能的差异。超参数设置如表 1 所示。

表 1　超参数设置

超参数	值
训练周期/Epoch	500
学习率	0.001
优化器	Adma
批处理规模	400
边缘设备数/个	10

另外，集中式学习数据集样本数为 220 000 个，均在中心设备数据集中可用。本文在分布式学习中引入 10 个边缘设备和一个中心设备。对应地，RadioML.2016.10A 数据集被等分成 10 份并存储在 10 个边缘设备中（每个边缘设备数据集样本数为 22 000 个）。

3.2　实验结果

3.2.1　关键超参数设置

分组卷积的卷积核尺寸 K_S^G 和时间特征提取模块 GRU 中神经元的数量 U 影响着模型分类性能和复杂度。分组卷积核尺寸设置如表 2 所示，随着 K_S^G 的增加，模型分类性能得到改善，但是模型复杂度也随之增加。综合考虑模型分类性能和复杂度，本文设置 K_S^G 为 3×3。时间特征提取模块神经元设置如表 3 所示，本文设 U=128，与 U=32 相比，虽然模型复杂度增加，但是模型分类性能提升显著。换言之，MCMBNN 牺牲部分模型复杂度换取较大的分类性能提升。

表 2　分组卷积核尺寸设置

K_S^G	FLOPS	参数量/个	模型权重/KB	分类性能
1×1	15.476×10^3	91 046	485	57.46%
3×3	20.596×10^3	111 046	565	62.84%
5×5	30.836×10^3	151 046	725	62.27%
7×7	46.196×10^3	211 046	965	62.77%

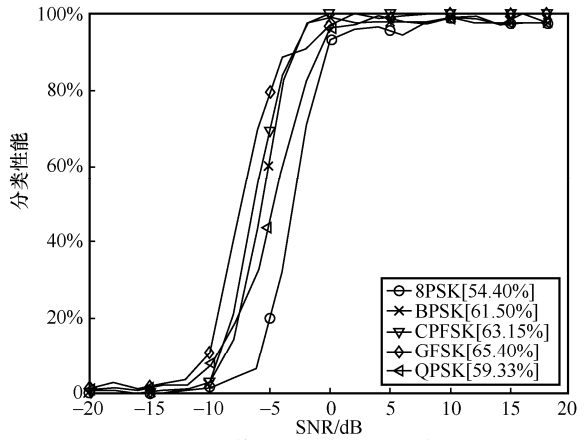
表 3　时间特征提取模块神经元设置

U	FLOPS	参数量/个	模型权重/KB	分类性能
32	20.474×10^3	48 934	316	57.08%
64	20.502×10^3	63 494	375	55.82%
128	20.596×10^3	111 046	565	62.84%
256	20.931×10^3	279 878	1228.8	62.66%

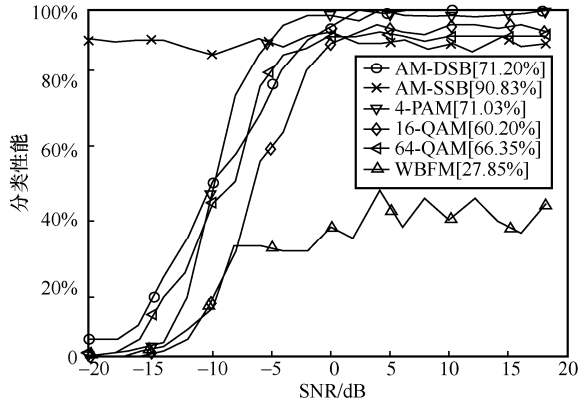
3.2.2　基于集中式学习的轻量化网络

集中式学习下的分类性能比较如图 8 所示，其中中括号内的数值代表相应的平均分类性能。从图 8(a)可以看出，所提轻量化网络 MCMBNN 在不同信噪比情况下对 PSK 和 FSK 都有较好的分类性能，且当 SNR = 4 dB 时，CPFSK 和 GFSK 分类性能达到 100%。图 8(b)给出了 MCMBNN 在不同信噪比情况下对 QAM、PAM 和模拟调制识别的性能。从图 8(b)可以看出，除 WBFM 外，其他调制的分类性能均超过 60%。WBFM 分类性能差是因为 WBFM 和 AM-DSM、AM-SSB 都属于模拟信号，它们之间的幅度相位频率等特征差异较小，因此容易造成混淆。

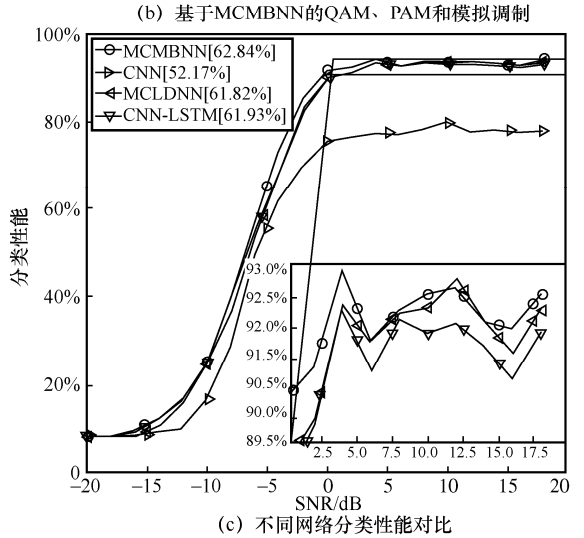
不同网络分类性能对比如图 8(c)所示。与 MCLDNN 和 CNN-LSTM 相比，当−6 dB≤SNR≤−2 dB 时，MCMBNN 分类性能比 MCLDNN 和 CNN-LSTM 高 3%。与 CNN 相比，MCMBNN 在所有信噪比下的分类性能均大于 CNN，尤其当 SNR≥−4 dB 时，MCMBNN 的分类性能比 CNN 高 10%。



(a) 基于MCMBNN的PSK和FSK



(b) 基于MCMBNN的QAM、PAM和模拟调制



(c) 不同网络分类性能对比

图 8 集中式学习下的分类性能比较

不同网络的复杂度比较如表 4 所示。相比于其他网络，MCMBNN 的时间复杂度和空间复杂度都较低。为了进一步分析 MCMBNN 的时间复杂度和空间复杂度，本文对 MCMBNN 实施了消融实验，消融实验是对模型中某一模块控制变量，观察这一模块对模型指标的影响。如表 5 所示，V_1 中只去除相位参数评估模块，V_2 中用对称卷积核代替非对称卷积核，V_3 中用标准卷积代替分组卷积，V_4 中用 LSTM 模块代替 GRU 模块。根据 V_1 和 MCMBNN 的对比可以发现，相位参数评估模块复杂度较低并且可以有效提高分类性能。根据 V_2、V_3 和 MCMBNN 的对比可以发现，非对称卷积核和分组卷积的使用可以在保证分类性能的同时有效降低模型复杂度。根据 V_4 和 MCMBNN 的对比可以发现，相比于 LSTM 模块，GRU 模块的使用可以有效降低模型复杂度。

表 4 不同网络的复杂度比较

网络	FLOPS	参数量/个	模型权重/MB	分类性能
CNN	42.664×10^3	2 190 283	8.793	52.17%
MCLDNN	36.322×10^3	406 199	1.669	61.82%
CNN-LSTM	38.966×10^3	340 939	1.396	61.93%
MCMBNN	20.596×10^3	111 046	0.552	62.84%

表 5 基于 MCMBNN 的消融实验

网络	FLOPS	参数量/个	模型权重/KB	分类性能
V_1	20.596×10^3	110 789	527	62.12%
V_2	21.087×10^3	112 966	571	62.83%
V_3	32.129×10^3	133 596	647	62.54%
V_4	20.642×10^3	134 086	657	62.33%
MCMBNN	20.596×10^3	111 046	565	62.84%

3) 基于分布式学习的轻量化网络

分布式学习下的分类性能比较如图 9 所示。在分布式学习中，MCMBNN 依然保持较好的分类性能。对比集中式学习，分布式学习有 0.68%的分类性能损失，原因是在分布式学习中，边缘设备不再共享本地数据集给中心设备，而是共享本地训练好模型的权重信息。

下面，对分布式学习中通信开销和训练效率 2 个指标展开分析，通信开销定义为

$$C_{\text{Decent}} = 2NW_{\text{m}}T \tag{9}$$

其中，N 表示边缘设备的数量，W_{m} 表示模型权重的大小，T 表示模型收敛时的训练周期数。当训练周期数一定时，通信开销与模型权重大小呈正相关。因此在分布式学习中部署轻量化模型可以有效减少通信开销，不同网络的通信开销如图 10 所示。相比于其他网络，轻量化网络 MCMBNN 有较低的通信开销。

训练效率是衡量模型训练速度的指标，定义为

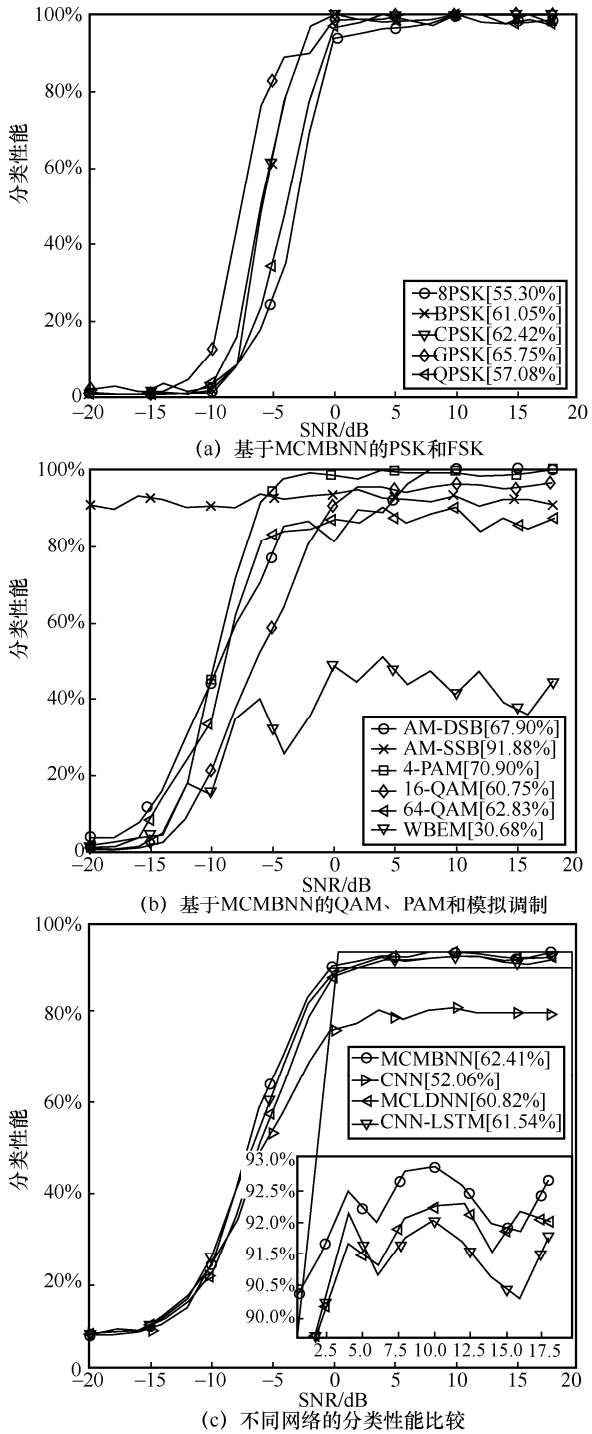


图 9　分布式学习下的分类性能比较

$$E_{\mathrm{f}}=\frac{1}{T_{\mathrm{train}}} \tag{10}$$

其中，T_{train} 是模型在一个训练周期的时间。不同模型在分布式学习和集中式学习下的训练效率对比如图 11 所示。从图 11 可以看出，分布式学习训练效率约为集中式学习的 5 倍，这得益于分布式学习中有多个边缘设备协同训练。

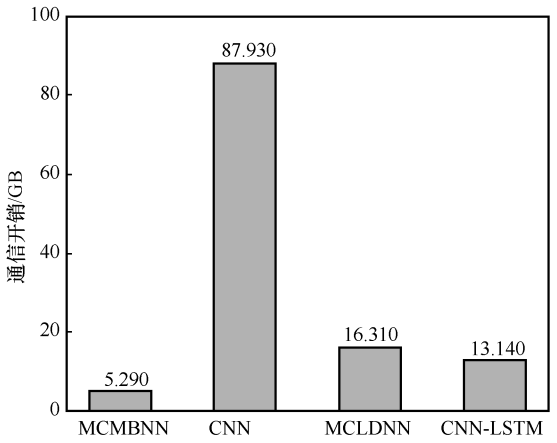


图 10　不同网络的通信开销

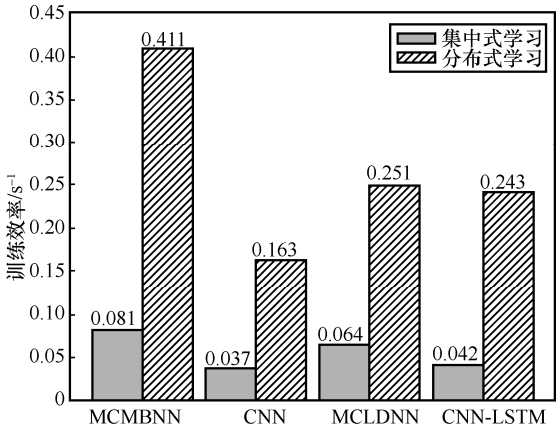


图 11　不同模型在分布式学习和集中式学习下的训练效率对比

4　结束语

本文提出了一种基于轻量化网络的分布式自动调制信号分类方法，该方法采用分布式学习优化和轻量化网络 MCMBNN。在设计轻量化网络的过程中，综合考虑模型分类性能和模型复杂度，应用不同的轻量化设计思想，充分提取调制信号的特征（相位、时间和空间特征）。不同于传统的集中式学习，分布式学习利用多个边缘设备训练一个全局模型，并且共享模型权重，因此分布式学习既充分利用了边缘设备上的数据，又避免了数据隐私外泄的风险，同时减轻了中心设备的计算压力和存储压力，提高了训练效率。此外，利用分布式学习训练轻量化网络，可以在保证分类性能的基础上，降低分布式学习由于权重信息的反复传输带来的通信开销。

参考文献：

[1]　化存卿. 物联网安全检测与防护机制综述[J]. 上海交通大学学报, 2018, 52(10): 1307-1313.

HUA C Q. A survey of security detection and protection for Internet of things[J]. Journal of Shanghai Jiao Tong University, 2018, 52(10): 1307-1313.

[2] POPOOLA S I, ANDE R, ADEBISI B, et al. Federated deep learning for zero-day botnet attack detection in IoT-edge devices[J]. IEEE Internet of Things Journal, 2022, 9(5): 3930-3944.

[3] 霍添财. 物联网终端设备恶意软件检测研究与设计[D]. 西安: 西安电子科技大学, 2021.
HUO T C. Research and design of malware detection of terminal devices in IoT networks[D]. Xi'an: Xidian University, 2021.

[4] 彭安妮, 周威, 贾岩, 等. 物联网操作系统安全研究综述[J]. 通信学报, 2018, 39(3): 22-34.
PENG A N, ZHOU W, JIA Y, et al. Survey of the Internet of things operating system security[J]. Journal on Communications, 2018, 39(3): 22-34.

[5] STOYANOVA M, NIKOLOUDAKIS Y, PANAGIOTAKIS S, et al. A survey on the Internet of things (IoT) forensics: challenges, approaches, and open issues[J]. IEEE Communications Surveys & Tutorials, 2020, 22(2): 1191-1221.

[6] 张玉清, 周威, 彭安妮. 物联网安全综述[J]. 计算机研究与发展, 2017, 54(10): 2130-2143.
ZHANG Y Q, ZHOU W, PENG A N. Survey of Internet of things security[J]. Journal of Computer Research and Development, 2017, 54(10): 2130-2143.

[7] 梁浩然, 伍军, 赵程程, 等. 基于博弈优化边缘学习的物联网入侵检测研究[J]. 物联网学报, 2021, 5(2): 37-47.
LIANG H R, WU J, ZHAO C C, et al. Leveraging edge learning and game theory for intrusion detection in Internet of things[J]. Chinese Journal on Internet of Things, 2021, 5(2): 37-47.

[8] 林冲, 闫文君, 张立民, 等. 通信信号调制识别综述[J]. 中国电子科学研究院学报, 2021, 16(11): 1074-1085.
LIN C, YAN W J, ZHANG L M, et al. An overview of communication signals modulation recognition[J]. Journal of China Academy of Electronics and Information Technology, 2021, 16(11): 1074-1085.

[9] 代翱, 张海剑, 孙洪. 联合时域和时频域特征的数字调制信号自动分类[J]. 信号处理, 2016, 32(11): 1283-1292.
DAI A, ZHANG H J, SUN H. Digital modulations automatic classification using the combination of several features extracted from time and time-frequence domain[J]. Journal of Signal Processing, 2016, 32(11): 1283-1292.

[10] 向建, 高勇. 基于 GRU-CNN 并联神经网络的自动调制识别[J]. 电讯技术, 2021, 61(11): 1339-1343.
XIANG J, GAO Y. Automatic modulation recognition based on GRU-CNN parallel neural network[J]. Telecommunication Engineering, 2021, 61(11): 1339-1343.

[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] XU J L, LUO C B, PARR G, et al. A spatiotemporal multi-channel learning framework for automatic modulation recognition[J]. IEEE Wireless Communications Letters, 2020, 9(10): 1629-1632.

[13] ZHANG Z F, LUO H, WANG C, et al. Automatic modulation classification using CNN-LSTM based dual-stream structure[J]. IEEE Transactions on Vehicular Technology, 2020, 69(11): 13521-13531.

[14] ZHANG F X, LUO C B, XU J L, et al. An efficient deep learning model for automatic modulation recognition based on parameter estimation and transformation[J]. IEEE Communications Letters, 2021, 25(10): 3287-3290.

[15] HUYNH-THE T, HUA C H, PHAM Q V, et al. MCNet: an efficient CNN architecture for robust automatic modulation classification[J]. IEEE Communications Letters, 2020, 24(4): 811-815.

[16] 张立志, 冉浙江, 赖志权, 等. 分布式深度学习通信架构的性能分析[J]. 计算机工程与科学, 2021, 43(3): 416-425.
ZHANG L Z, RAN Z J, LAI Z Q, et al. Performance analysis of distributed deep learning communication architecture[J]. Computer Engineering & Science, 2021, 43(3): 416-425.

[17] 刘艺璇, 陈红, 刘宇涵, 等. 联邦学习中的隐私保护技术[J]. 软件学报, 2022, 33(3): 1057-1092.
LIU Y X, CHEN H, LIU Y H, et al. Privacy-preserving techniques in federated learning[J]. Journal of Software, 2022, 33(3): 1057-1092.

[18] 陈世达, 刘强, 韩亮. 降低分布式训练通信的梯度稀疏压缩方法[J]. 浙江大学学报(工学版), 2021, 55(2): 386-394.
CHEN S D, LIU Q, HAN L. Gradient sparsification compression approach to reducing communication in distributed training[J]. Journal of Zhejiang University (Engineering Science), 2021, 55(2): 386-394.

[19] 赵羽, 杨洁, 刘淼, 等. 面向视频监控基于联邦学习的智能边缘计算技术[J]. 通信学报, 2020, 41(10): 109-115.
ZHAO Y, YANG J, LIU M, et al. Federated learning based intelligent edge computing technique for video surveillance[J]. Journal on Communications, 2020, 41(10): 109-115.

[20] WANG Y, GUO L, ZHAO Y, et al. Distributed learning for automatic modulation classification in edge devices[J]. IEEE Wireless Communications Letters, 2020, 9(12): 2177-2181.

[作者简介]

杨洁（1980− ），女，江苏南京人，博士，南京邮电大学讲师，主要研究方向为分布式学习、边缘计算和智能无线通信等。

董标（1998− ），男，江苏淮安人，南京邮电大学硕士生，主要研究方向为基于分布式学习的自动调制信号分类技术。

付雪（1997− ），女，贵州遵义人，南京邮电大学博士生，主要研究方向为基于分布式学习的自动调制信号分类技术。

王禹（1996− ），男，江苏盐城人，南京邮电大学博士生，主要研究方向为基于分布式学习的自动调制信号分类技术。

桂冠（1982− ），男，安徽枞阳人，博士，南京邮电大学教授，主要研究方向为人工智能、深度学习、智能通信和智能物联网等 6G 技术。