

一种基于伪标签半监督学习的小样本调制识别算法

史蕴豪, 许华, 刘英辉

(空军工程大学 信息与导航学院, 陕西 西安 710077)

摘要:针对有标签样本较少条件下的通信信号调制识别问题,提出了一种基于伪标签半监督学习技术的小样本调制方式分类算法,通过优选人工特征集、设计高性能分类器以及基于输出概率的伪标签数据选择方法,构建高效的伪标签标注系统,然后通过该伪标签标注系统与基于深度学习的信号分类方法配合,实现在少量有标签样本和大量无标签样本条件下的调制方式分类。仿真结果表明,对6种数字信号进行调制识别,在信噪比大于5 dB时,伪标签算法可将模型识别性能提高5%~10%,该算法设计简单,具有较大应用价值。

关键词:调制识别; 伪标签; 半监督学习

中图分类号: TN975

文献标志码: A

文章编号: 1000-2758(2020)05-1074-10

通信信号调制方式识别是通信侦察、认知电子战领域的关键技术之一,在实际战场环境中,由于敌我双方的非协作特性,使得在还原敌方通信信号、获取敌方情报信息之前必须进行调制方式识别,此外,通信信号调制方式识别也为后续的信号解调、比特流分析、协议识别、信号解密、灵巧干扰等提供了重要支撑。

调制识别技术发展至今,已逐渐形成两大类基于似然比的调制识别方法和基于特征的调制识别方法。基于似然比的调制识别方法主要是通过计算似然概率模型估计不同调制方式的概率,然后将各个信号代入检验,最后判定似然概率最大的为识别结果。基于似然比的调制识别算法主要分为3类^[1-3],但该类方法普适性差的缺点较为明显,由于概率密度函数都是针对特殊环境提出来的,因此参数微小的偏差和模型失配都会导致识别率大幅降低。基于特征的调制识别方法通过提取信号不同特征搭配分类算法从而做到对信号的识别,如文献[4]使用高阶累积量对8种数字信号进行识别,文献[5]混合高阶累积量与循环谱特征对信号进行分类,文献[6-7]使用信号熵特征对不同调制信号进行分类。但是基于人工提取特征的调制识别算法中,对特征区分度要求较高,一旦特征对于不同信号辨别性能差,

识别模型效果就会显著下降,因此有必要对自动提取特征的方法进行研究。

近年来,深度学习在模式识别、计算机视觉等领域取得了显著突破,利用深度学习的方法自动提取信号特征并对信号调制方式进行识别也已经取得了不错的效果,O'Shea等^[8-9]最早于2016年利用有监督深度学习技术实现调制方式识别,该论文直接使用卷积神经网络(convolutional neural network, CNN)构建端到端的学习模型,成功对包括WBFM、DSB、BPSK、16QAM在内的11种数字或模拟调制方式进行了识别。Jeong等^[10]学者在论文中提出了一种算法,利用短时傅里叶变换(short-time Fourier transform, STFT)将信号从时域转换为时频域,并通过深度卷积神经网络提取时频域特征,最终完成了对2-FSK、4-FSK、8-FSK等7种调制识别方式的识别,其在-4 dB的信噪比下仍有90%以上的识别正确率。Meng等^[11]学者提出了一种联合噪声估计的调制识别算法,该研究提出了一种巧妙的网络结构,同时将原始信号数据和信噪比作为神经网络的输入,仿真结果显示这种算法在不同信噪比下、不同频偏下的识别成功率已经非常接近理论识别率的上限。文献[12]提出利用时频图的纹理信息进行调制识别,在大样本条件下,可取得良好的分类效果。

收稿日期: 2019-12-31

基金项目: 国家自然科学基金(61601500)资助

作者简介: 史蕴豪(1996—),空军工程大学硕士研究生,主要从事智能通信对抗、信号处理研究。

Zhang 等^[13]利用卷积神经网络提取信号 SPWVD 时频图特征和 BJD 时频图特征并与大量手工特征融合对 BPSK、QPSK、OFDM 等 8 种调制方式进行识别,在 -4 dB 时仍有 92.5% 的识别准确率。

虽然基于深度学习的调制识别方法已经取得了非常显著的成果,但是国内外专家的研究重点都偏向于有监督深度学习方法,而有监督深度学习方法需要大量有标签信号样本作为支撑。但在实际应用中,一些敏感的军用通信信号由于受敌方保密要求的限制以及地理环境条件的制约,在平时只能获取十分有限的信号样本,而这些信号真正到了战场又会大量出现。因此,如何利用少量有标签信号结合大量无标签信号的小样本调制方式识别逐渐成了大家关注的重点方向之一。现阶段,国内外针对小样本调制识别的研究尚处于起步阶段,但在其他领域已有一些半监督学习算法出现。半监督学习解决的就是有标签样本数量较少情况下如何提升模型性能的问题,其核心思想为结合无标签数据进行优化。尽管无标签数据没有标签信息但是它们和有标签数据一样都是从相同的数据源独立同分布采样得到的,因此它们包含的关于数据分布的信息对优化模型大有裨益^[14]。半监督学习方法主要包括自训练学习方法^[15]、生成式学习方法^[16]以及半监督支持向量机等^[17]。

为了充分利用无标签信号解决小样本调制识别问题,本文提出了一种基于伪标签半监督学习的小样本调制识别模型。该模型通过优选人工特征集结合高性能分类器为无标签信号进行预测打上伪标签,然后利用深度学习类方法联合训练带标签样本与伪标签样本从而实现小样本调制识别。

1 小样本识别模型

本文采用基于伪标签的半监督学习方法,在仅有少量有标签信号样本的条件下进行调制方式识别。基于伪标签的半监督学习是一种增量算法,算法流程图如图 1 所示,在分类器训练部分,首先通过少量有标签样本提取优选人工特征送入高性能分类器中进行训练,将信号识别准确率提升到一个较高水准,此时若环境中存在大量无标签样本,则进入伪标签生成部分,利用训练好的高性能分类器对无标签样本进行预测,通过预测的概率对无标签样本进行排序,给可靠的无标签样本打上伪标签并加入到

训练集中,但该过程并非选择所有的无标签样本均加入训练集,因为全部无标签样本内包含的错误标签样本会严重干扰模型收敛。当环境中无标签样本不足时,也可利用分类器训练部分模块直接进行信号分类。

伪标签算法不断迭代增加有标签样本数量,当有标签样本数量满足迭代要求时,进入深度学习训练部分,将所有带标签的样本数据组合新的训练集,使用分类能力更强的深度学习类方法对所有标签样本进行联合训练,最后利用训练好的模型对测试信号进行预测。本文选取 BP 神经网络作为生成伪标签的高性能分类器,选取 CNN 卷积神经网络作为深度学习类方法训练真实标签样本与伪标签样本的合集。

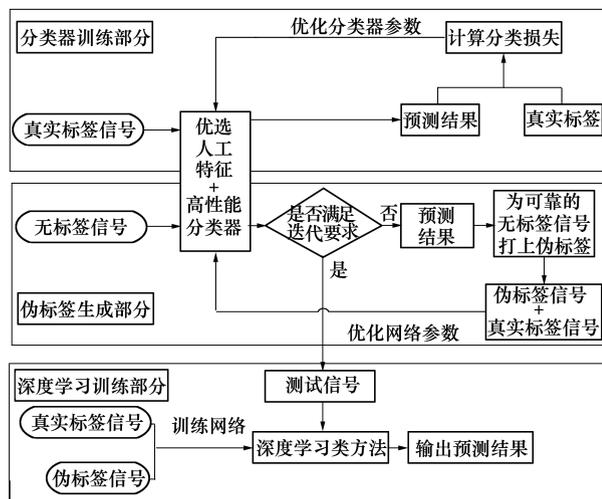


图 1 算法流程图

2 信号人工特征

本文选取多类具有良好区分性能的人工特征作为 BP 神经网络输入,这些特征均已被证明具有良好的区分能力,其中包括高阶累积量特征、熵特征以及时频特征。

2.1 信号高阶累积量特征

在调制识别技术领域,高阶累积量是应用非常广泛的特征之一,其具有较强的周期分量,可用于准确识别不同的数字调制信号。为了提取高阶累积量,首先要计算信号的高阶矩,信号高阶矩由(1)式计算:

$$M_{pq} = E[x(n)^{p-q}(x^*(n))^q] \quad (1)$$

通过信号各高阶矩便可计算得到许多高阶累积量,本文选择下列高阶累积量,这些累积量均已被证明在调制样式分类中有较好鉴别性能^[18-19]

$$C_{40} = M_{40} - 3M_{20}^2 \tag{2}$$

$$C_{41} = M_{41} - 3M_{20}M_{21} \tag{3}$$

$$C_{42} = M_{42} - M_{20}^2 - 2M_{21}^2 \tag{4}$$

$$C_{60} = M_{60} - 15M_{40}M_{20} + 30M_{20}^3 \tag{5}$$

$$C_{61} = M_{61} - 5M_{40}M_{21} - 10M_{20}M_{41} + 30M_{21}M_{20}^2 \tag{6}$$

$$C_{62} = M_{62} - 6M_{20}M_{42} - 8M_{21}M_{41} - M_{22}M_{40} + 6M_{20}^2M_{22} + 24M_{21}^2M_{20} \tag{7}$$

$$C_{63} = M_{63} - 6M_{41}M_{20} - 9M_{21}M_{42} + 18M_{21}M_{20}^2 + 12M_{21}^3 \tag{8}$$

2.2 信息熵特征

熵是用于评价信号或系统状态平均不确定性的指标。在信息论领域,熵用于衡量信息的信息量大小,信息的不确定程度越大,则其熵值越大,因此信息熵理论为我们提供了一个很好的信号特征描述方法。本文提取信号的功率谱熵、奇异谱熵和能量谱熵,以此作为信号的特征^[20-21]。

2.2.1 功率谱熵

假设时间序列 X 长为 L , 对其进行离散傅里叶变换, 变换结果为

$$y(k) = \sum_{n=0}^{N-1} x(n)e^{-i\frac{2\pi}{N}nk} = \sum_{n=0}^{N-1} x(n)W_N^{nk} \tag{9}$$

式中: $W = e^{-i\frac{2\pi}{N}}$, k 表示傅里叶变换后的第 k 个频谱; N 表示变换点数, 一般要求 N 为 2 的整数幂且接近序列 X 的长度。计算频谱序列 Y 在 $y(k)$ 处的功率谱

$$S(k) = \frac{1}{N} |y(k)|^2, k = 0, 1, 2, \dots, N-1 \tag{10}$$

记

$$p_i = \frac{S(k)}{\sum_{k=0}^{N-1} S(k)} \tag{11}$$

将(11)式代入香农熵计算公式, 可得到功率谱香农熵。香农熵计算公式为: $H = E[-\log p_i] =$

$\sum_{i=1}^n p_i \log p_i$ 。式中, H 表示熵值; p_i 表示信号概率分布。

2.2.2 奇异谱熵

奇异谱分析是近年来非常流行的一种研究非线性时间序列数据的强大方法, 它结合相空间重构和奇异值分解对时间序列维数进行估计。若一段离散

时间序列为

$$\mathbf{X} = [x_1, x_2, x_3, \dots, x_N]$$

首先将信号分段, 假设分段长度为 m , 在奇异谱分析过程中, m 最好为信号周期的整数倍且不宜超过信号序列长度的 $1/3$, 重构后的序列轨迹矩阵为

$$\mathbf{M} = \begin{bmatrix} x_1 & x_2 & \dots & x_{N-m+1} \\ x_2 & x_3 & \dots & x_{N-m+2} \\ \dots & \dots & \dots & \dots \\ x_m & x_{m+1} & \dots & x_N \end{bmatrix}_{K \times J} \tag{12}$$

对(12)式矩阵进行奇异值分解, 可得

$$\mathbf{M}_{K \times J} = \mathbf{U}_{K \times K} \mathbf{\Sigma}_{K \times J} \mathbf{V}_{J \times J}^T \tag{13}$$

式中: \mathbf{U} 和 \mathbf{V} 均为正交矩阵; \mathbf{U} 为左奇异矩阵; \mathbf{V} 为右奇异矩阵; $\mathbf{\Sigma}$ 矩阵可化为对角阵

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & & \dots & & \\ & \sigma_2 & & \dots & \\ & & \dots & \sigma_k & \dots \\ & & & & \dots \end{bmatrix}_{K \times J}$$

式中, σ_k 表示矩阵 M 的奇异值且除对角线上元素以外其余值均为零, 对角线上的非零元素便构成了序列的奇异值谱, 即

$$\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_i, \dots, \sigma_j \mid j < K\}$$

记 p_i 表示非零奇异值 σ_i 占有所有非零奇异值之和的比值

$$p_i = \frac{\sigma_i}{\sum_{i=1}^r \sigma_i} \tag{14}$$

将(14)式代入香农熵计算公式, 即可得到奇异值香农熵。

将(14)式带入指数熵计算公式, 即可得到奇异谱指数熵。指数熵的计算公式为: $H = E[e^{-p_i}] = \sum p_i e^{-p_i}$ 。式中, H 表示熵值; p_i 表示信号概率分布。

2.2.3 能量谱熵

对于序列信号 $\mathbf{X} = \{x_1, x_2, \dots, x_N\}$, 其能量谱定义为

$$S(\omega) = \frac{1}{N} |X(\omega)|^2 \tag{15}$$

式中, $X(\omega)$ 表示序列 X 的离散傅里叶变换。记 p_i 为

$$p_i = \frac{S(i)}{\sum_{i=1}^N S(i)} \tag{16}$$

将(16)式代入指数熵计算公式,即可得到信号能量谱指数熵。

2.3 时频特征

2.3.1 归一化中心瞬时振幅功率密度最大值

归一化中心瞬时振幅的功率密度最大值可在一定程度上反应不同信号的谱特征,其定义如下^[22]:

$$\gamma_{\max} = \frac{\max | \text{FFT}[a_{cn}(i)] |^2}{N_s} \quad (17)$$

$$a_{cn}(i) = a_n(i) - 1 \quad (18)$$

$$a_n(i) = a(i)/m_a \quad (19)$$

$$m_a = \frac{1}{N_s} \sum_{i=1}^{N_s} a(i) \quad (20)$$

式中: N_s 表示信号序列长度; m_a 表示信号瞬时幅度的均值。

2.3.2 归一化中心瞬时振幅绝对值的标准差

序列信号归一化中心瞬时振幅绝对值的标准差可由(18)式求得,其定义如下

$$\sigma = \sqrt{\frac{1}{N_s} \left(\sum_{i=1}^{N_s} a_{cn}^2(i) \right) - \left(\frac{1}{N_s} \sum_{i=1}^{N_s} [a_{cn}(i)] \right)^2} \quad (21)$$

3 网络结构

本节将对选择的 BP 网络与 CNN 卷积神经网络作详细介绍。选用 BP 网络主要原因是 BP 网络具有学习能力,可以自动学习各个特征对分类结果影响的权重,即送入网络的特征中若有部分在分类方面作用不明显那么其所属权重就会减小。因此对于 BP 网络而言,其对输入的特征会自动进行“特征筛选”,这在一定程度上可以减轻对先验信息的依赖,此外,由于深度神经网络等模型参数庞大、权重太多,在小样本条件时模型过拟合现象非常严重,因此浅层 BP 网络用作小样本阶段训练具有明显优势。

BP 网络的训练过程如图 2 所示,在开始训练时,由于仅有少量的有标签信号样本,训练出的模型泛化能力弱,而人工特征不依赖于训练数据,可以弥补训练数据量较少时特征表达能力不足的问题,因此本文设计的 BP 网络模型输入为已具备分类能力的优选特征集。首先对信号进行归一化处理,然后计算信号熵特征、高阶累积量特征和时频特征,将所有特征组合后一起送入 BP 网络训练,最后利用 softmax 分类器输出信号预测概率。

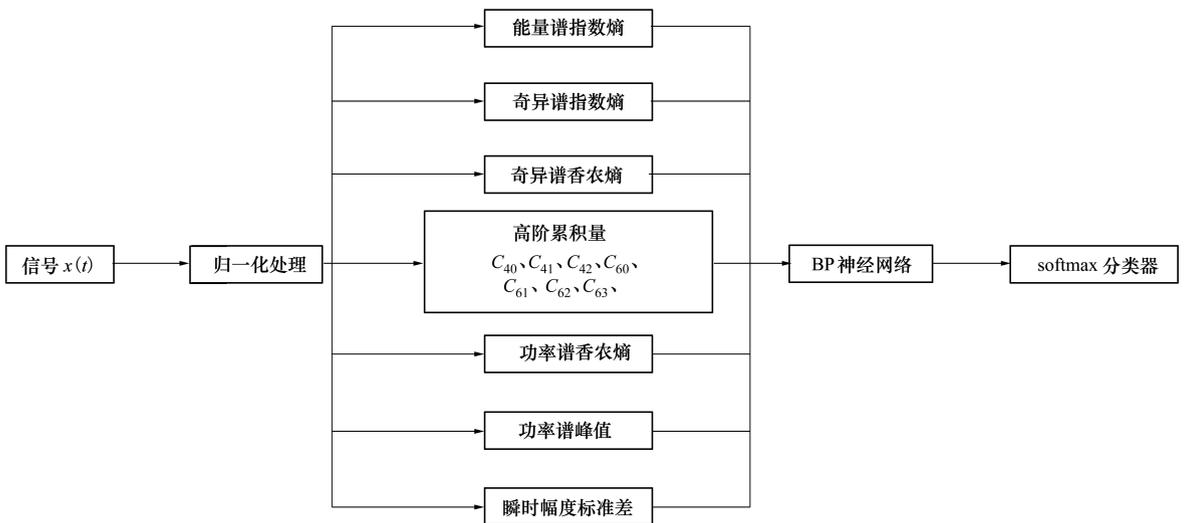


图 2 网络训练过程

BP 网络结构如图 3 所示,本文设计的 BP 网络由 1 层输入层,4 层全连接层以及 softmax 分类器构成,输入层大小为 $1 * 13$,分别代表 13 种人工提取的特征,各全连接层神经元个数分别为 16, 32, 64, 32,各神经元均采用 ReLU 激活函数。为提升网络

的泛化能力,在第三、四层全连接层后使用 dropout 技术干扰训练以防止网络过拟合,以提高网络在测试样本上的泛化能力,但干扰只发生在网络训练阶段,在网络模型生成伪标签时则停止干扰,Dropout 比率设置为 0.5。BP 网络训练产生的损失函数如下

所示:

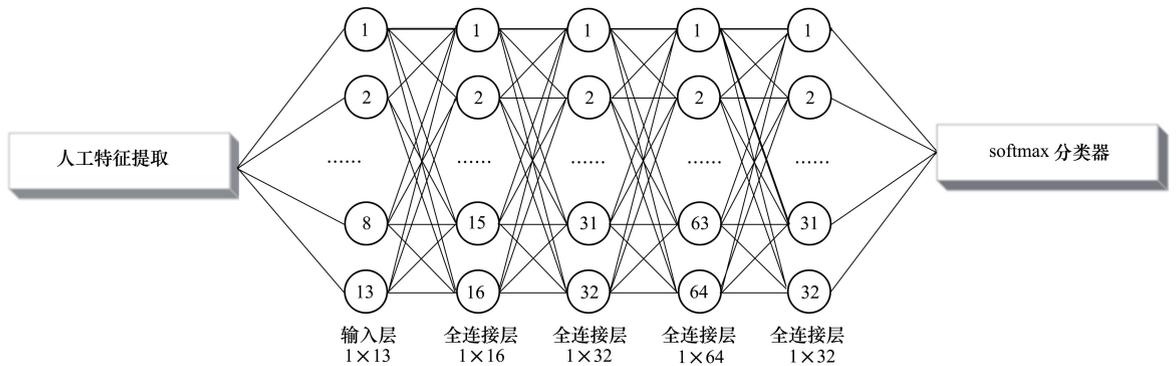


图 3 BP 神经网络结构

$$L_{total} = \frac{1}{N} \sum_{m=1}^N \sum_{i=1}^C L_{real_label}(y_i^m, f_i^m) + \mu \frac{1}{N'} \sum_{m=1}^{N'} \sum_{i=1}^C L_{pseudo_label}(y_i^m, f_i^m) \quad (22)$$

式中: $L_{real_label}(y_i^m, f_i^m)$ 表示有标签信号产生的分类损失; $L_{pseudo_label}(y_i^m, f_i^m)$ 表示伪标签信号产生的分类损失; N 表示有标签信号数量; N' 伪标签信号数量, μ 用于控制 2 类损失的比重。

在利用伪标签算法迭代训练 BP 网络过程中, 有时难免给无标签信号样本打上错误的伪标签, 因此伪标签的准确性对网络最终的识别率有决定性影响。由于 softmax 分类器输出的是各个类别的预测概率, 因此本文提出可靠条件为

$$p_2 + p_3 \leq p_1 \quad (23)$$

式中: p_1 表示 softmax 分类器输出的最大概率; p_2, p_3 依次表示 softmax 分类器输出的第二大、第三大 概率。即只有 softmax 输出的最大概率大于第二大 概率与第三大 概率之和时, 才会给该无标签信号打上

伪标签, 通过此基于输出概率的样本选择算法便可在一定程度上保证伪标签的可靠性。

无标签样本标注伪标签结束后, 将真实标签样本与伪标签样本联合起来送入 CNN 卷积神经网络训练, CNN 结构如图 4 所示。为充分发挥卷积神经网络的特征提取能力, 本文直接将 IQ 信号输入卷积神经网络, 信号输入网络前, 将其转换为二维数据, 即若信号长为 L , 则转换后的数据格式为 $[2, L]$, 两路分别代表 I 路数据与 Q 路数据。网络共包含 3 层卷积层和 3 层全连接层, 其中第一个卷积层卷积核个数为 64, 核尺寸为 $2 * 4$, 第二个卷积层卷积核个数为 32, 核尺寸为 $1 * 4$, 第三个卷积层卷积核个数为 16, 核尺寸为 $1 * 4$ 。卷积层完成对序列数据的特征提取, 将提取后的特征转换为一维序列送入全连接层, 3 层全连接层的维度分别为 64, 32, 16, 将第三个全连接层输出的特征送入 softmax 分类器进行分类, 输出预测结果, 并通过反向传播不断优化网络参数。

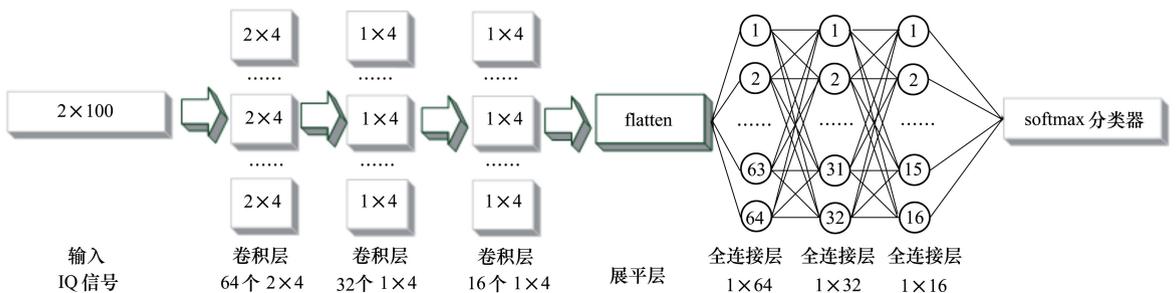


图 4 卷积神经网络结构图

4 仿真结果及分析

本文选用的调制信号集为 {BPSK, 4PAM, 4PSK, 8PSK, 16QAM, 64QAM}, 共计 6 种数字调制信号, 信号序列长度 $L=100$, 信噪比从 -10 至 20 dB, 间隔为 2 dB。训练集每类信号生成 4 000 个信号样本, 信噪比随机, 共计 24 000 个样本。测试集每类信号每个信噪比点生成 500 个信号, 共计 48 000 个信号, 所有信号均由 MATLABR2016a 仿真生成。

网络训练均基于 Python 下的 Keras 深度学习框架实现, 硬件平台为 Intel (R) Core (TM) i7-

8700CPU, GPU 为 NVIDIA GeForce 1060Ti。

4.1 信号特征提取分析

本节对 6 类数字信号在 -10 至 20 dB 间每个信噪比点生成 500 个样本, 提取信号 13 种特征并取平均值, 取其中具有代表性的特征并绘图, 得到图 5 所示的特征曲线图。在进行信号奇异谱分析时, 由于信号序列长度 $L=100$, 且基带序列为随机生成无周期特性, 因此设置分段长度 $m=33$; 在计算序列离散傅里叶变换时, 由于要求傅里叶变换点数靠近序列长度且为 2 的整数幂, 因此设置傅里叶变换点数 $N=128$ 。

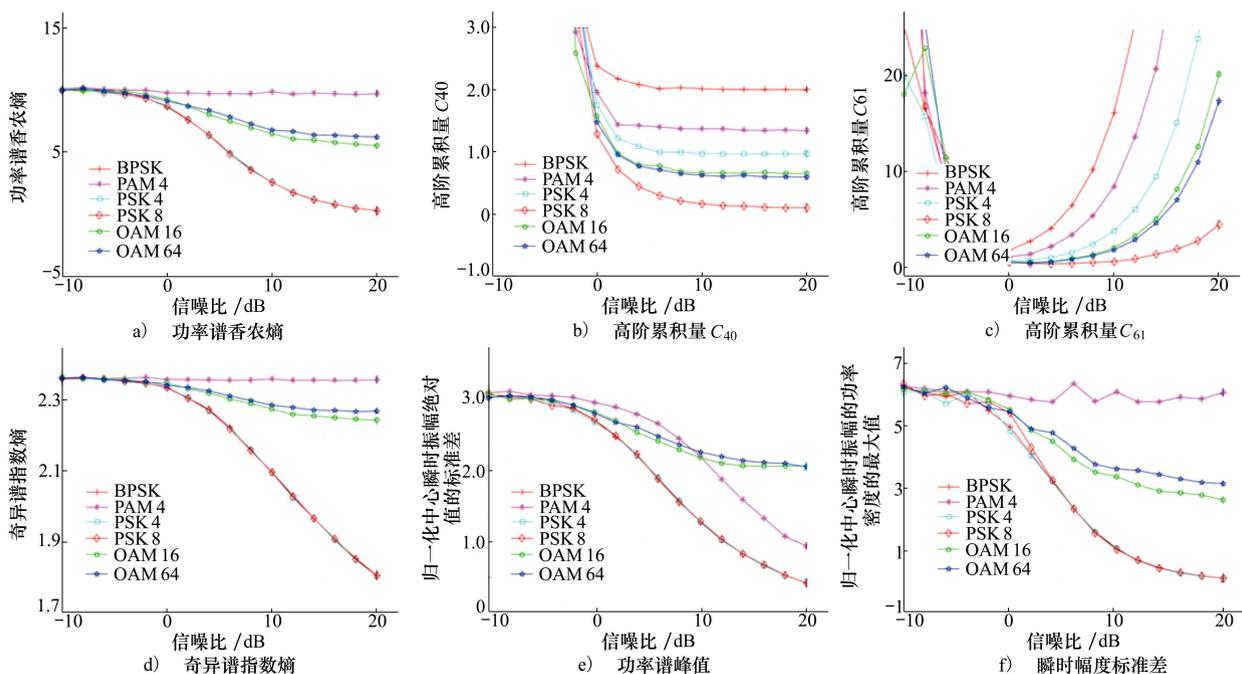


图5 信号特征曲线图

图 5a) 至 5e) 分别代表信号功率谱香农熵、高阶累积量 C_{40} 、高阶累积量 C_{61} 、奇异谱指数熵、功率谱峰值和瞬时幅度标准差随信噪比变化的曲线。可以看出, 随着信噪比的不断上升, 各类信号特征间的差异越来越大, 代表着特征的区分能力越来越强, 但每个特征均有无法区分的调制信号, 因此本文选择综合 13 类特征一起作为分类特征。

4.2 伪标签算法可行性分析

本节对比 3 类识别算法在不同训练样本总量下的识别性能, 分别为人工提取特征结合 BP 神经网络、时序 IQ 序列结合 CNN 卷积神经网络、时序 IQ 序列结合 LSTM 循环神经网络, LSTM 网络为

KERAS 框架下的 CuDNNLSTM。3 类算法均迭代 200 次, 使用 Adam 函数优化。当不同样本量条件下进行调制方式识别时, 各算法在测试集上的性能如图 6 所示, 从仿真结果可以看出, 当训练样本数量有限时, 例如图 6a) 所示总样本量为 600 个, 通过人工提取特征结合 BP 神经网络的识别率要远高于 CNN 卷积神经网络的识别率以及 LSTM 循环神经网络的识别率。但随着样本数量的不断增加, 人工提取特征结合 BP 神经网络的识别率变化不大, 而 CNN 卷积神经网络的识别率和 LSTM 网络的识别率则上升的非常快, 当总样本量达到 6 000 时, CNN 的最高识别性能已经优于人工提取特征结合 BP 网络的性

能,其最高识别率可以达到 85%以上,当总样本量达到 24 000 时,网络的最高识别率可以达到 95%以上。因此,利用人工特征结合 BP 神经网络在样本

量较少的情况下采用伪标签增量算法增加训练数据总样本量,而后采用 CNN 卷积神经网络进行训练是可行的。

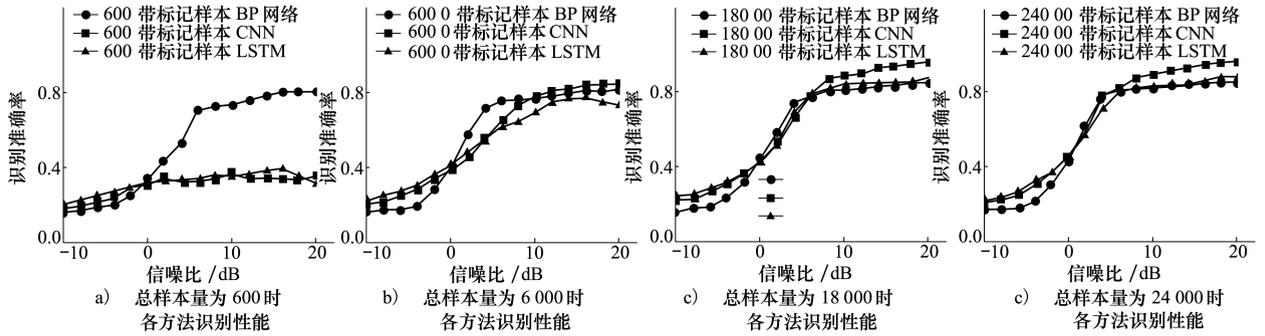


图 6 不同样本总量下各算法识别率曲线

4.3 伪标签算法识别率分析

通过 4.2 节中的仿真结果可分析出,在样本量充足的条件下本文提出的 CNN 深度卷积神经网络模型性能优于 CuDNNLSTM 网络,因此本文选用 CNN 作为训练模型对打伪标签后的样本集进行训练。

假设初始条件为有每类调制信号有 100, 500, 1 000 个带标签样本,共 600, 3 000, 6 000 个带标签样本以及 48 000 个无标签样本,首先通过人工提取特征结合 BP 神经网络对 600, 3 000, 6 000 个带标签样本进行训练,而后通过伪标签增量算法对无标签样本打伪标签,最后利用 CNN 对有标签样本、伪标签样本的 IQ 序列进行联合训练。在 CNN 网络训练过程中,由于真实标签样本与测试集样本分布相同且标签准确,因此本文选用真实标签样本作为验证集,以此提升网络在测试集上的识别性能。

当使用 600 个真实标签样本对大量无标签样本预测伪标签时,样本数随迭代次数的变化如表 1 所示。

表 1 600 个真实样本

迭代次数	真实标签样本数	伪标签样本数	总标签样本数
1	600	25 843	26 443
2	600	26 451	27 051
3	600	26 953	27 533
4	600	28 426	29 026
5	600	28 920	29 520

从表 1 的结果可以看出,随着迭代次数的不断增加,伪标签数量不断提升,但由于设置了可靠条件,所以伪标签样本总量并未达到 48 000。将所有真实标签样本和伪标签样本一并送入 CNN 卷积神经网络进行训练,共计 29 520 个样本,测试集的识别率如图 7 所示。

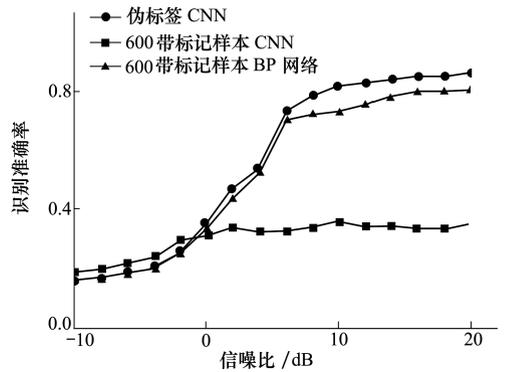


图 7 伪标签算法识别率对比

当训练样本总量为 29 520 时,算法各部分运行时间如表 2 所示,当使用 3 000 个真实标签样本对大量无标签样本预测伪标签时,样本数随迭代次数的变化如表 3 所示,从表 3 的结果可以看出,随着迭代次数的不断增加,伪标签数量不断提升,经过 5 次迭代后训练样本数总数可达 36 566 个,测试集的识别率如图 8 所示,当训练样本总量为 36 566 个时,算法各部分运行时间如表 4 所示。

表 2 算法运行时间 单位:s

步骤	标签预测	网络训练	测试
时间	8.6	197.8	3.6

表 4 算法运行时间 单位:s

步骤	标签预测	网络训练	测试
时间	11.5	245.6	3.7

表 3 3 000 个真实样本

迭代次数	真实标签样本数	伪标签样本数	总标签样本数
1	3 000	28 154	31 154
2	3 000	29 514	32 514
3	3 000	31 102	34 102
4	3 000	32 158	35 158
5	3 000	33 566	36 566

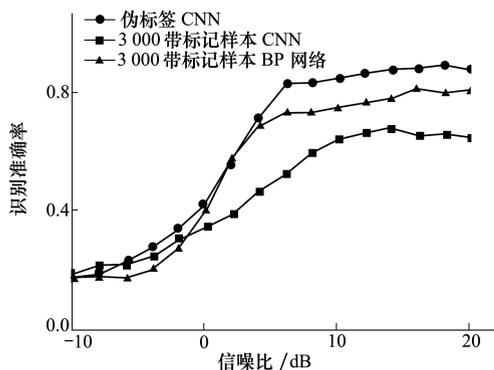


图 8 伪标签算法识别率对比

当使用 6 000 个真实标签样本对大量无标签样本预测伪标签时,样本数随迭代次数的变化如表 5 所示。从表 5 的结果可以看出,随着迭代次数的不断增加,伪标签样本数量不断提升,经过 5 次迭代后训练样本总数可达 41 063 个,测试集的识别率如图 9 所示。

表 5 6 000 个真实样本

迭代次数	真实标签样本数	伪标签样本数	总标签样本数
1	6 000	31 009	37 009
2	6 000	32 151	38 151
3	6 000	33 548	39 548
4	6 000	34 158	40 158
5	6 000	35 063	41 063

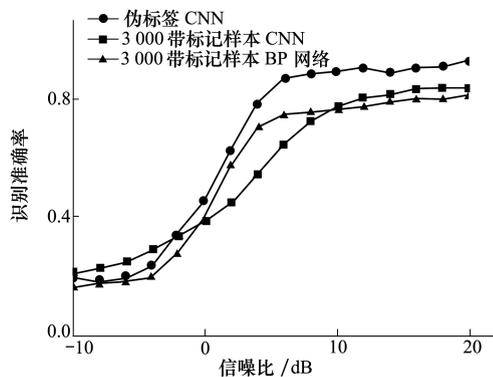


图 9 伪标签算法识别率对比

当训练样本总量为 41 063 时,算法各部分运行时间如表 6 所示。

表 6 算法运行时间 单位:s

步骤	标签预测	网络训练	测试
时间	13.4	273.6	3.7

通过仿真结果可以看出,在小样本条件下,利用伪标签增量算法对传统特征结合 BP 神经网络的识别性能有一定的提升,当信噪比大于 5 dB 时,本文提出的伪标签 CNN 的网络识别率相较于 BP 神经网络结合人工特征的识别率可提高约 5%~10%。

实际上本文提出的伪标签半监督模型,其性能与生成伪标签所用的人工特征集以及分类器密切相关,通过大量实验发现,如果伪标签标注分类器不能很好区分信号类型,那么能够获得的有效伪标签样本就会较少,总体识别准确率就会偏低,但如果能进一步提升伪标签分类器的识别率,将其控制到 90% 以上,那么整体算法的识别率甚至可以接近 100%。在实际应用过程中,一方面可以实时监控伪标签样本数量分布,从而掌握深度学习类方法是否有足够的样本量支撑,另一方面还可以不断改进特征集和分类器,继续研究更具区分能力的特征并选择更复杂的模型作为伪标签生成器,这样深度学习类方法的性能也会不断提升。

5 结 论

本文针对战场信号调制识别领域可能出现的小样本情况进行研究,设计了基于人工优选特征集与 BP 神经网络的信号伪标签标注方法,并结合基于

CNN 的通信信号分类模型,形成了小样本条件下通信信号分类新的解决方案,通过大量的实验验证了方案的可行性,在单个信号类型的带标签样本量为

100 以上时,本文模型就可有效工作,且在不同标签样本量条件下,其总体性能均比常规方法有明显提升。

参考文献:

- [1] PANAGIOTOU P, ANASTASOPOULOS A, POLYDOROS A. Likelihood Ratio Tests for Modulation Classification[C]//MILCOM 2000 Proceedings of 21st Century Military Communications, Architectures and Technologies for Information Superiority, 2000: 670-674
- [2] POLYDOROS A, KIM K. On the Detection and Classification of Quadrature Digital Modulations in Broad-Band Noise[J]. IEEE Trans on Communications, 1990, 38(8): 1199-1211
- [3] ABDI A, DOBRE O A, CHOUDHRY R, et al. Modulation Classification in Fading Channels Using Antenna Arrays[C]//IEEE Military Communications Conference, 2004: 211-217
- [4] 赵雄文,郭春霞,李景春. 基于高阶累积量和循环谱的信号调制方式混合识别算法[J]. 电子与信息学报,2016,38(3): 674-680
ZHAO Xiongwen, GUO Chunxia, LI Jingchun, et al. Mixed Recognition Algorithm for Signal Modulation Schemes by High-Order Cumulants and Cyclic Spectrum[J]. Journal of Electronics & Information Technology, 2016,38(3): 674-680 (in Chinese)
- [5] 刘亚冲,唐智灵. 基于 Softmax 回归的通信辐射源特征分类识别方法[J]. 计算机工程,2018,44(2): 98-102
LIU Yachong, TANG Zhiling. Classification and Identification Method of Communication Radiation Source Feature Based on Softmax Regression[J]. Computer Engineering, 2018, 44(2): 98-102 (in Chinese)
- [6] ZHANG Z, LI Y B, JIN S S, et al. Modulation Signal Recognition Based on Information Entropy and Ensemble Learning[J]. Entropy, 2018, 20(3): 198-198
- [7] WANG H, GUO L L, DOU Z, et al. A New Method of Cognitive Signal Recognition Based on Hybrid Information Entropy and D-S Evidence Theory[J]. Mobile Networks and Applications, 2018, 23(4): 677-685
- [8] 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
- [9] 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 Process, 2018, 12(1): 168-179
- [10] JEONG S, LEE U, KIM S C. Spectrogram-Based Automatic Modulation Recognition Using Convolution Neural Network[C]//2018 Tenth International Conference on Ubiquitous and Future Networks, 2018: 843-845
- [11] MENG Fan, CHEN Peng, WU Lenan. Automatic Modulation Classification: A Deep Learning Enabled Approach[J]. IEEE Trans on Vehicular Technology, 2018, 67(11): 10760-10772
- [12] 白芑远,许华,孙莉. 基于卷积神经网络与时频图纹理信息的信号调制方式分类方法[J]. 西北工业大学学报,2019,37(4): 816-823
BAI Pengyuan, XU Hua, SUN Li. A Recognition Algorithm for Modulation Schemes by Convolution Neural Network and Spectrogram Texture[J]. Journal of Northwestern Polytechnical University, 2019, 37(4): 816-823 (in Chinese)
- [13] ZHANG Z, WANG C, GAN C, et al. Automatic Modulation Classification Using Convolutional Neural Network with Features Fusion of SPVWD and BJD[J].IEEE Trans on Signal and Information Processing over Networks, 2019, 5(3): 469-478
- [14] 龙明盛. 迁移学习问题与方法研究[D]. 北京:清华大学,2014
LONG Mingsheng. Transfer Learning: Problems and Methods[D]. Beijing: Tsinghua University, 2014 (in Chinese)
- [15] MAULIK U, CHAKRABORTY D. A Self-Trained Ensemble with Semi-Supervised SVM: an Application to Pixel Classification of Remote Sensing Imagery[J]. Pattern Recognition,2011,44(3): 615-623
- [16] PAWAR S, BHATTACHARYYA P, PALSHIKAR G K. Semi-Supervised Relation Extraction Using EM Algorithm[C]//10th Int Conf on Natural Language Processing, 2013
- [17] 陈波. 半监督迁移学习算法研究[D]. 西安:西北大学,2015

- CHEN Bo. Research on Semi-Supevised Transfer Learning[D]. Xi'an: Northwest University, 2015 (in Chinese)
- [18] 李跃,郭兴吉,赵欣. 基于高阶累积量的调制方式识别研究[J]. 西南科技大学学报,2018,33(3): 64-68
- LI Yue, GUO Xingji, ZHAO Xin. Study on Modulation Recognition Based on Higher-Order Cumulants[J]. Journal of Southwest University of Science and Technology, 2018, 33(3): 64-68 (in Chinese)
- [19] 谭晓衡,褚国星,张雪静,等. 基于高阶累积量和小波变换的调制识别算法[J]. 系统工程与电子技术, 2018(1): 171-177
- TAN Xiaoheng, CHU Gouxing, ZHANG Xuejing, et al. Modulation Recognition Algorithm Based on High-Ordercumulants and Wavelet Transform[J]. Engineering and Electronic, 2018(1): 171-177 (in Chinese)
- [20] 李晨,杨俊安,刘辉. 基于信息熵和 GA-ELM 的调制识别算法[J]. 系统工程与电子技术, 2020, 42(1): 223-229
- LI Chen, YANG Junan, LIU Hui. Modulation Recognition Algorithm Based on Information Entropy and GA-ELM[J]. Engineering and Electronic, 2020, 42(1): 223-229 (in Chinese)
- [21] ZHEN Z, YIBING L, SHANSHAN J, et al. Modulation Signal Recognition Based on Information Entropy and Ensemble Learning [J]. Entropy, 2018, 20(3): 198
- [22] ALI A, FAN Y. Automatic Modulation Classification Using Deep Learning Based on Sparse Autoencoders with Non-Negativity Constraints[J]. IEEE Signal Processing Letter, 2017, 24(11): 1626-1630

A Few-Shot Modulation Recognition Method Based on Pseudo-Label Semi-Supervised Learning

SHI Yunhao, XU Hua, LIU Yinghui

(Institute of Information & Navigation, Air Force Engineering University, Xi'an 710077, China)

Abstract: In order to solve the problem of insufficient labeled samples in modulation recognition, this paper proposes a few-shot modulation recognition algorithm based on pseudo-label semi-supervised learning (pseudo-label algorithm). First of all, high quality artificial feature, excellent classifier and data-labeling method are used to build efficient pseudo label system, and then the pseudo label system is combined with signal classification method based on the deep learning to realize the modulation classification under the condition of a small number of labeled samples and a large number of unlabeled samples. The simulation results show that the pseudo-label algorithm can improve the model recognition performance by 5%-10% when the six kinds of digital signals are classified and identified and its SNR is greater than 5 dB. At the same time, the algorithm has a simple network design and is of great application value.

Keywords: modulation recognition; pseudo-label algorithm; semi-supervised learning; simulation