



系统工程与电子技术  
Systems Engineering and Electronics  
ISSN 1001-506X, CN 11-2422/TN

## 《系统工程与电子技术》网络首发论文

题目：基于集成学习与特征降维的小样本调制识别方法  
作者：史蕴豪，许华，郑万泽，刘英辉  
收稿日期：2020-06-08  
网络首发日期：2020-10-15  
引用格式：史蕴豪，许华，郑万泽，刘英辉. 基于集成学习与特征降维的小样本调制识别方法[J/OL]. 系统工程与电子技术.  
<https://kns.cnki.net/kcms/detail/11.2422.TN.20201014.1326.028.html>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 基于集成学习与特征降维的小样本调制识别方法

史蕴豪, 许 华, 郑万泽, 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, 通过集成人工特征与深度学习自动提取特征构成特征集合, 再针对性设计特征选择算法对特征合集进行优选生成高效特征子集, 最后利用可快速收敛的高性能分类器对信号进行区分, 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, 算法对八种数字信号进行调制识别, 在信噪比为 20dB 时, 本文算法可将信号最高识别率提升至 96%, 同时本算法设计简单, 具有较大应用价值。

**关键词:** 调制识别、小样本、集成学习、特征选择

**中图分类号:** TN 975 **文献标志码:** A

## Few-shot modulation recognition method based on ensemble learning and feature reduction

SHI Yunhao, XU Hua, ZHENG Wanze, LIU Yinghui

(Information and Navigation College, Air Force Engineering University, Xi'an 710077, China)

**Abstract:** Aiming at the problem of modulation recognition with few labeled samples, forming the feature set by integrating handcrafted features and deep learning features, and then the feature selection algorithm is designed to optimize the feature set to generate an optimal feature subset. Finally, the signals are distinguished by a fast convergent high-performance classifier, which can 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 when the SNR is 20 dB, the algorithm can improve the signal recognition rate to 96%. At the same time, the algorithm is simple and has great application value.

**Keywords:** modulation recognition; few-shot; ensemble learning; feature selection

## 0 引 言

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, 无论是传统方法还是近些年兴起的深度学习方法均已取得杰出的研究成果。传统方法方面, 文献[1-3]利用信号高阶累积量实现调制信号分类, 文献[4-5]通过提取信号循环谱特征区分信号不同调制方式, 文献[6-7]利用信息熵特征实现信号分类。深度学习方法方面, Timothy 等人<sup>[8-9]</sup>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构建端到端的学习模型, 成

功对 11 种数字或模拟调制方式进行了识别。Jeong 等学者<sup>[10]</sup>在论文中提出了一种算法利用短时傅里叶变换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 7 种调制方式的识别,其在 -4dB 的信噪比下仍有 90% 以上的识别正确率。Meng 等学者<sup>[11]</sup>提出了一种联合噪声估计的调制识别算法,该算法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接近理论识别率的上限。Zhang 等人<sup>[12]</sup>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 8 种调制方式进行识别,在 -4dB 时仍有 92.5% 的识别准确率。

虽然近些年基于深度学习的调制识别方法在识别性能方面逐渐超越传统人工特征方法,但是深度学习类方法需要大量带标签信号样本作为支撑,一旦带标签样本量不足,模型的识别性能就会发生急剧下降。随着信号采集方式的多样化以及存储技术的高速发展,现阶段获取大量无类别标记信号样本变得相当简单,但是若想获得同样量级的有类别标记信号则相当困难,这是由于数据的类别标记工作一般需要耗费大量的人力、物力和时间。在实际战场环境中,电磁频谱中会出现大量信号,将这些信号采集下来后逐一打上标签是不现实的,无法适应瞬息万变的战场态势,因此研究有标签信号样本不足的小样本调制识别问题就显得尤为重要。

通过总结大量专家研究成果发现,小样本问题产生的根本原因是传统方法通过固定模式处理信号序列提取信号特征,从而可利用低维度特征表示信号,而深度学习类方法由于网络参数多、

拟合难,需要大量的样本才能完成低维特征提取;另一方面,分类器训练特征时所需样本量也是随着特征维度呈指数增长。因此如何有效减小特征数量,并尽可能利用低维度特征表示高维信号序列是解决小样本问题的核心所在,若能利用低维度特征表征原始时序信号,则可极大减小分类器所需训练样本。针对上述思路,本文提出一种基于集成学习与特征降维的小样本调制方式分类模型,该模型通过集成人工设计特征与机器学习自动提取特征构成融合特征集,再针对性设计特征选择算法对融合特征集进行优选生成高效特征子集,最后设计高性能分类器对少量有标签信号进行训练,从而解决有标签训练样本不足导致的小样本问题。

## 1 识别模型

基于集成学习与特征降维的小样本模型本质思想是利用低纬度特征表征原始信号,降维对高维小样本问题行之有效,信号降维主要包括特征提取和特征选择两方面。本文拟提取信号最具表征性、区分度的特征并降维,从而达到减少有标签训练样本数的目的。如图 1 所示,本文算法主要分为 3 个阶段,第 1 阶段为特征提取阶段,利用传统方法提取对信号区分能力强的人工特征,与此同时利用机器学习类方法中的自编码器神经网络对样本进行无监督训练,自动提取低维信号特征,而后将两类特征进行融合重组。第 2 阶段为特征选择阶段,运用自主设计定的特征选择算法综合选出一定数量的最具区分能力的特征,生成最优特征子集。第 3 阶段为分类阶段,利用少量有标签样本对高性能分类器进行有监督训练,从而完成信号小样本条件下的分类识别。

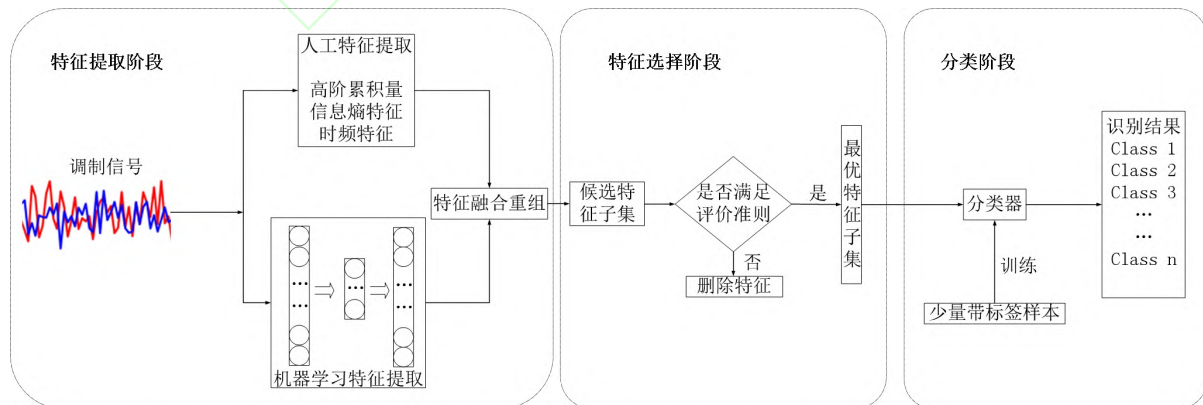


图 1 算法总体框架

Fig. 1 Algorithm block diagram

## 2 信号特征提取

### 2.1 人工特征提取

在调制识别领域中, 众多研究人员已进行大量特征工程研究, 设计了很多卓有成效的特征用于信号分类, 并取得了优异的结果。因此, 利用前人的研究成果提升信号分类效率是重要的研究切入点。本文拟提取多类表征性强、区分度高的特征作为信号人工特征, 包括信号高阶累积量、信息熵特征以及时频特征。

#### 2.1.1 信号高阶累积量特征

在调制技术领域, 高阶累积量是应用非常广泛的特征之一, 由于其具有较强的周期分量, 可用于准确识别不同的调制信号。为提取高阶累积量, 首先要计算信号的高阶矩, 序列信号  $x(n)$  的高阶矩计算公式为

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

通过信号各阶高阶矩便可计算得到各类高阶累积量, 本文选择下列高阶累积量, 这些累积量均已被证明在调制信号分类识别中有较好鉴别能力<sup>[13-14]</sup>, 即

$$C_{40} = M_{40} - 3M_{20}^2 \quad (2)$$

$$C_{42} = M_{42} - M_{20}^2 - 2M_{21}^2 \quad (3)$$

$$C_{60} = M_{60} - 15M_{40}M_{20} + 30M_{20}^3 \quad (4)$$

$$C_{61} = M_{61} - 5M_{40}M_{21} - 10M_{20}M_{41} + 30M_{21}^2M_{20} \quad (5)$$

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

#### 2.1.2 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, 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。本文拟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<sup>[15-16]</sup>。

##### 2.1.2.1 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

$$y(k) = \sum_{n=0}^{N-1} x(n) e^{-i \frac{2\pi}{N} nk} = \sum_{n=0}^{N-1} x(n) W_N^{nk} \quad (7)$$

式中,  $W = e^{-i \frac{2\pi}{N}}$ ,  $k$  表示傅里叶变换的第  $k$  个频

谱,  $N$  表示变换点数, 一般要求  $N=2$  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, 即

$$S(k) = \frac{1}{N} |y(k)|^2, k = 0, 1, 2, \dots, N-1 \quad (8)$$

记

$$p_i = \frac{S(k)}{\sum_{k=0}^{N-1} S(k)} \quad (9)$$

将式(9)代入香农熵计算公式, 即可得到功率谱香农熵。香农提出的熵的计算公式如式(20)所示。

##### 2.1.2.2 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。假设一段离散时间序列为

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

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

$$M = \begin{bmatrix} x_1 & x_{1+n} & \dots & x_{1+(m-1)n} \\ x_2 & x_{2+n} & \dots & x_{2+(m-1)n} \\ \dots & \dots & \dots & \dots \\ x_{N-(m-1)n} & x_{N-(m-1)n+n} & \dots & x_N \end{bmatrix}_{K \times J} \quad (10)$$

对式(10)进行奇异值分解, 可得

$$M_{K \times J} = U_{K \times K} \Sigma_{K \times J} V_{J \times J}^T \quad (11)$$

式中,  $U$  和  $V$  均为正交矩阵,  $U$  为左奇异矩阵,  $V$  为右奇异矩阵,  $\Sigma$  矩阵可化为对角阵, 即

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

式中,  $\sigma_k$  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

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

记  $p_i$  表示非零奇异值  $\sigma_i$  占有所有非零奇异值之和的比值:

$$p_i = \frac{\sigma_i}{\sum_{i=1}^r \sigma_i} \quad (12)$$



将式(12)代入香农熵计算公式,即可得到奇异值香农熵。

将式(12)代入指数熵计算公式,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。

### 2.1.2.3 能量谱熵

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

$$S(\omega) = \frac{1}{N} |X(\omega)|^2 \quad (13)$$

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

$$p_i = \frac{S(i)}{\sum_{i=1}^N S(i)} \quad (14)$$

将式(14)代入指数熵计算公式,即可得到信号能量谱指数熵。

### 2.1.3 归一化中心瞬时振幅的功率密度最大值

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<sup>[17]</sup>,其定义为

$$\gamma_{\max} = \frac{\max |FFT|a_{cn}(i)|^2}{N_s} \quad (15)$$

$$a_{cn}(i) = a_n(i) - 1 \quad (16)$$

$$a_n(i) = a(i) / m_a \quad (17)$$

$$m_a = \frac{1}{N_s} \sum_{i=1}^{N_s} a(i) \quad (18)$$

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。

## 2.2 自编码器特征提取

自编码器作为一种无监督机器学习算法可以在不借助标签信息的前提下学习数据的稀疏表示,近些年已有许多学者将其应用至信号调制方式识别中,例如文献[18]使用两个并行自编码器完成信号调制方式识别,文献[19]利用卷积自编码器直接作用于信号序列以提取信号特征。

自编码器结构包括编码器和解码器两部分,输入数据首先通过编码器进行降维操作,再通过解码器进行数据重构,通过约束中间层的维度并最小化重构数据与输入数据间的误差,从而达到特征映射的目的,在整个过程中虽然并未涉及输入信号的标签信息但却完成了信号特征提取的功能。本文设计图2所示的自编码器结构用作信

号特征自动提取,自编码器网络总体上由卷积层和全连接层构成。在信号输入自编码器网络之前,首先对信号进行加噪处理,由于通过自编码器后需重构原始无噪数据,因此自编码器提取的特征具有抗噪能力,也更具鲁棒性,本文网络选用高斯噪声干扰信号,噪声系数为0.1。信号输入后首先通过两层卷积层对其进行特征提取,而后通过三层全连接层对特征进行降维、再升维处理,最后再通过两层卷积层对信号进行重构,使得网络输出信号格式与输入格式保持一致。

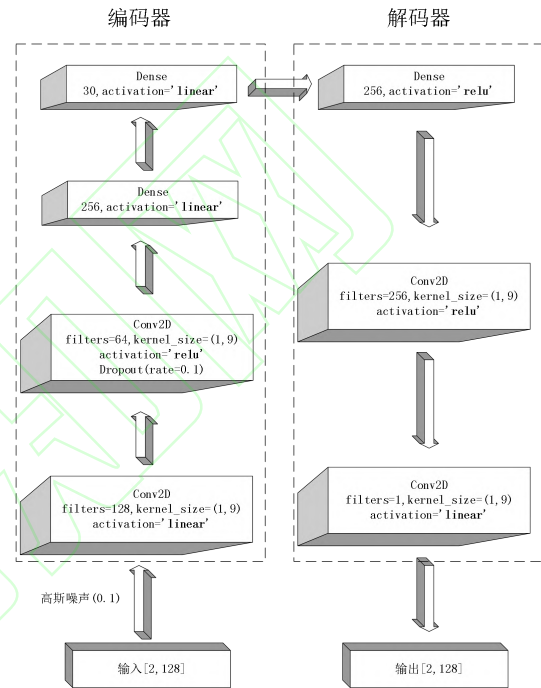


图2 自编码器结构

Fig.2 Auto-coder structure

假设自编码器网络输入信号为  $x$ , 经过加噪声处理后为  $\hat{x}$ , 自编码器输出需要重构原始无噪信号  $x$ , 则自编码器损失函数可化简表示为

$$J(x, \hat{x}) = \sum \left\| h(f(\hat{x})) - x \right\|_2^2 \quad (19)$$

式中,  $f(x)$  表示自编码器中编码器函数部分,  $h(x)$  表示自编码器中解码器函数部分。通过最小化损失函数,约束输入输出之间的重构误差,便可以低维中间层特征表示高维信号,从而达到特征提取的目的。

## 3 特征选择算法

与前两节所述的特征提取方法不同,特征选择方法指的是从全部原始特征集中选择最相关

的特征子集以此降低特征向量维度,相较于特征提取方法,特征选择方法更加侧重于揭示特征与特征间、特征与类别间的因果关系。在以往的调制方式识别算法中,选用传统手工特征还是利用机器学习方法自动提取信号特征一直存在争议,两类方法也各自具有优势,因此本文先对两类特征进行融合,再使用特征选择算法进行综合筛选,选取其中最具区分度的特征。一个好的特征选择算法不仅可以帮助我们降低训练分类器所需样本数量,剔除冗余或不相关特征,还可以提升模型运行速度,加快算法收敛速度和降低硬件要求。

由于本文研究的是小样本条件下的信号识别问题,因此在特征选择时既有少量有标签的信号,同时还存在大量无标签的信号,即需要使用半监督特征选择算法进行特征选择,近些年,国内外学者也提出了各式各样的特征选择算法,例如文献[20-26],其中文献[25]通过信息熵理论设计了随机变量的对称不确定性,用于度量特征与特征之间的冗余性、特征与类别之间的相关性,在多个数据集上均有良好的表现。文献[26]则最大程度的利用了无标签样本,确保了所选特征子集优于仅用有类别标记数据的特征选择算法。本文综合文献[25]的快速过滤特征选择算法(fast correlation-based filter solution, FCBF)和文献[26]的半监督代表特征选择算法(semi-supervised representatives feature selection, SRFS)设计了新的半监督特征选择算法。

该方法的流程图如图3所示,整个特征选择过程主要分为两个阶段,具体内容如下。

(1)删除不相关特征。有标签信号可利用标签信息直接求出特征与标签之间的互信息量作为衡量特征重要性的标准。互信息是两个随机变量共同信息量的度量,假设特征为随机变量 $X$ ,信号标签为 $C$ , $p(x)$ 、 $p(c)$ 和 $p(x,c)$ 分别表示 $X$ 、 $C$ 、 $(X,C)$ 的概率密度函数,则随机变量 $X$ 与 $C$ 的互信息量 $I(X,C)$ 定义为

$$I(X,C)=\sum_{i=1}^m \sum_{j=1}^n p(x_i, c_j) \log_2 \frac{p(x_i, c_j)}{p(x_i)p(c_j)} \quad (20)$$

无标签信号虽然没有标签信号可以利用,但是其本身包含的自信息量也可在一定程度上指导特征选择。在信息论中,熵可用来度量特征自身包含的信息量,对于随机变量 $X$ ,熵的计算公式为

$$H(X)=\sum_{i=1}^m p(x_i) \log \frac{1}{p(x_i)} \quad (21)$$

当计算出各个特征与标签的互信息量与自信息量后,通过 $F$ -相关性判别特征 $F_i$ 是否为不相关特征, $F$ -的计算公式为

$$F\_Rel(F_i, C)=\beta I(F_i, C)+(1-\beta)H(F_i) \quad (22)$$

式中, $\beta$ 是一个大于0但小于1的参量,可用于控制互信息量与熵之间的比例。本文设定 $\beta=\sqrt{\frac{N}{N+M}}$ ,其中 $N$ 表示带标签信号数量, $M$ 表示无标签信号数量。

若 $F$ -大于设定的阈值表示该特征与类别相关且携带的信息量较大,则将其加入相关特征子集,若 $F$ -小于设定阈值则将其视为不相关特征删除。本文设计相关度阈值 $\alpha=\left\lceil \frac{D}{2\log_2 D} \right\rceil$ ,其中 $D$ 表示特征个数。

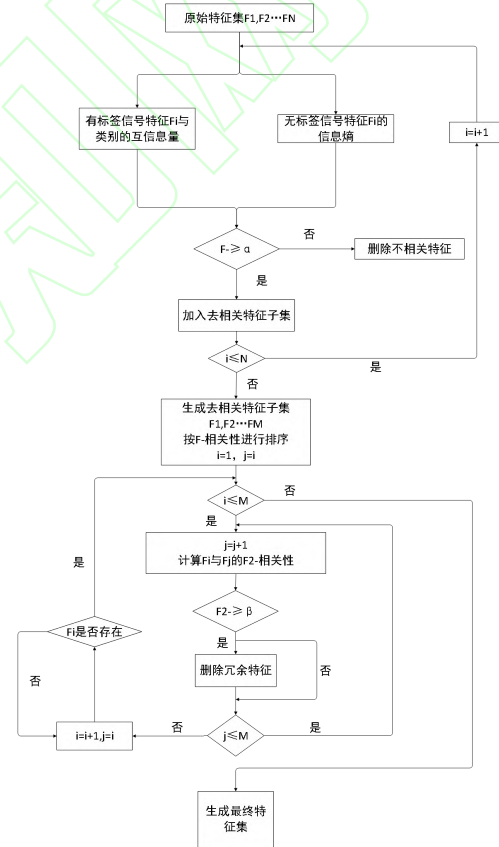


图3 特征选择过程  
Fig.3 Feature selection algorithm

(2)删除冗余特征。第一步删除不相关特征后得到去相关特征子集 $\{F_1, F_2, \dots, F_M\}$ ,将待选的特征按照 $F$ -相关性进行降序排列, $F$ -数值越大则排名越靠前。而后从前到后依次计算特征与类别间的 $F1$ -相关性,以及特征与特征之间的 $F2$ -相关性, $F1$ -的计算公式为

$$F1\_Rel(F_i, C) = \beta UI(F_i) / H(F_i) + (1 - \beta) SU(F_i, C) \quad (23)$$

式中,  $UI(F_i)$  表示特征  $F_i$  与其它所有特征的互信息量的均值, 其计算表达式为

$$UI(F_i) = \frac{1}{n} (H(F_i) + \sum_{j=1, n, j \neq i} I(F_i; F_j)) = \frac{1}{n} [H(F_i) + \sum_{j=1, n, j \neq i} H(F_i, F_j) - \sum_{j=1, n, j \neq i} H(F_j)] \quad (24)$$

式中,  $SU(F_i, F_j)$  为特征  $F_i$  与  $F_j$  的对称不确定性, 其计算公式为

$$SU(F_i, F_j) = 2 \left[ \frac{I(F_i; F_j)}{H(F_i) + H(F_j)} \right] \quad (25)$$

$F2$ -的计算公式为

$$F2\_Rel(F_i, F_j) = \beta SU(F_i, F_j) + (1 - \beta) USU(F_i, F_j) \quad (26)$$

式中,  $USU(F_i, F_j)$  表示特征  $F_i$  与  $F_j$  的无监督对称不确定性, 具体表述为

$$USU(F_i, F_j) = 2 \left[ \frac{UI(F_i; F_j)}{H(F_i) + H(F_j)} \right] \quad (27)$$

$$UI(F_i; F_j) = UI(F_i) - UI(F_i | F_j) = UI(F_i) - \frac{UI(F_i)}{H(F_i, F_j) + UI(F_i)} \quad (28)$$

本文设定式(29)为冗余判别条件, 若满足该条件则将  $F_j$  视为冗余特征删除, 在删除过程中优先保留  $F$ -排序靠前的特征, 迭代结束后最终剩余的即为最终选出的特征子集。

$$F1\_Rel(F_j, C) \geq F1\_Rel(F_i, C) \cap F2\_Rel(F_i, F_j) \geq F1\_Rel(F_i, C) \quad (29)$$

## 4 仿真验证

本文选用的调制信号集为 {BPSK、4PSK、8PSK、8QAM、16QAM、64QAM、4PAM、8PAM}, 共计 8 种调制信号, 各个信号序列长度  $L=128$ ,

包括 I、Q 两路数据, 信号数据格式为 [2, 128], 训练集每类信号生成 5000 个信号样本, 信噪比随机, 共计 40000 个信号样本, 这其中包含 800 个带标签信号样本, 每类 100 个, 其余均为无标签样本; 测试集每个信噪比点生成 100 个样本, 信噪比从 -10dB 至 20dB, 间隔为 2dB, 共计 16 个信噪比点, 12800 个信号, 所有信号均由 MATLAB R2016a 仿真生成。

网络训练均基于 python 下的 tensorflow、keras 深度学习框架实现, 硬件平台中 CPU 使用 Intel(R)Core(TM)i7-8700, GPU 使用 NVIDIA GeForce 1060Ti。

### 4.1 人工特征提取分析

本节将对第 2 节遴选出的人工特征性能进行仿真验证, 实验对 8 类调制信号在 -10dB 至 20dB 间每个信噪比点取 100 个样本, 提取信号 10 种手工特征并取平均值, 取其中具有代表性的特征并绘图, 得到图 4 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  $L=128$ , 且基带序列为随机生成无周期特性, 因此设置分段长度  $m=43$ ; 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  $N=128$ 。由图 4 可以看出, 随着信噪比的不断上升, 不同类别信号的特征值差距逐渐增大并逐渐趋于平稳, 这有助于对信号进行分类; 但同时可以看出每个特征均存在不好区分的调制类别, 因此本文选择将所有特征融合起来再由特征选择算法进行自动选择。通过本节可视化的特征实验可以看出, 本文所选人工特征具有较好的调制信号区分能力。

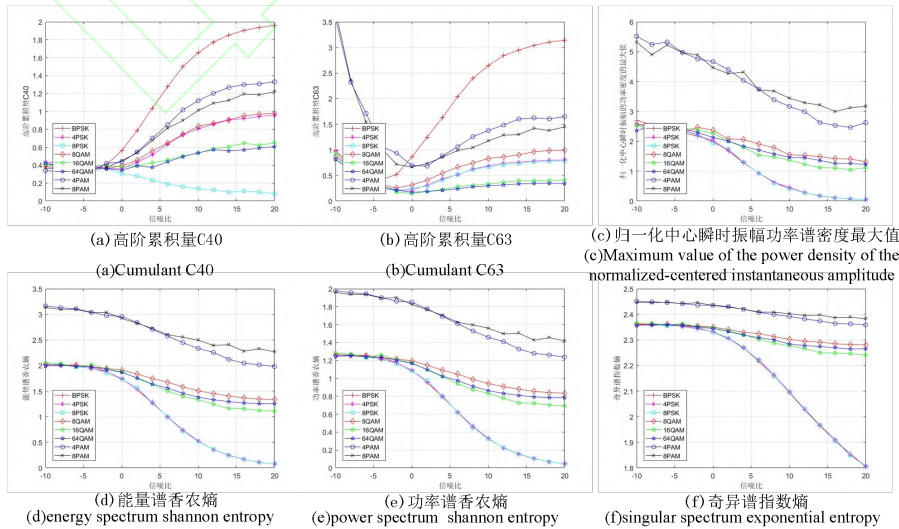


图 4 信号特征曲线

Fig.4 Signal feature curve



## 4.2 自编码器特征提取分析

本节将对自编码器特征提取效果进行实验验证, 使用每类 5000 个无标签信号样本, 共计 40000 个无标签训练样本对自编码器进行无监督训练, 自编码器结构如 2.2 节所示, 网络训练过程中 batch\_size 设置为 500, 共迭代 100 个 epochs。输入和输出间使用 MSE 衡量损失, 使用 ADAM 优化器进行优化, 网络输入、输出信

号格式均为[2, 128]。训练结束后保存网络参数, 再将所有信号输入网络进行计算, 利用编码层对信号样本进行压缩, 观察编码器与解码器的输出。信号压缩重构后的结果如图 5 所示, 其中图 5(a)~图 5(d)分别表示 BPSK、8PSK、16QAM、64QAM 信号输入 I、Q 序列、输出 I、Q 序列、中间层编码器输出特征向量。

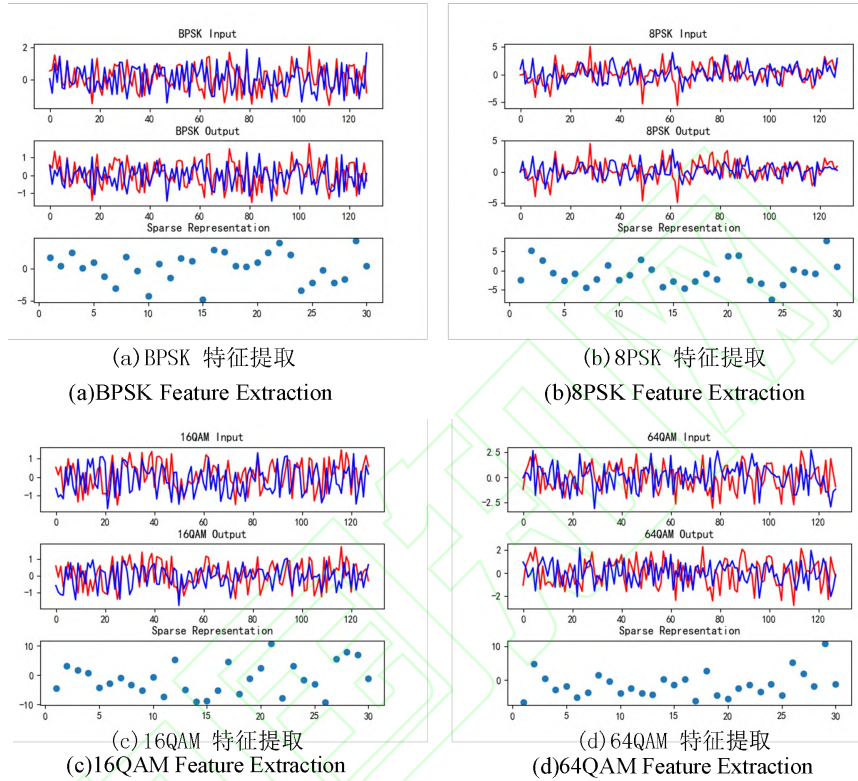


图 5 自编码器输入输出

Fig.5 Input and output of auto-coder

通过图 5 可以看出, 经过自编码器重构后信号波形并未发生明显变化, 这表示信号经过网络重构后输入端与输出端误差较小, 意味着中间层的低维 1\*30 特征在一定程度上可代表原始信号 2\*128 的 IQ 数据。

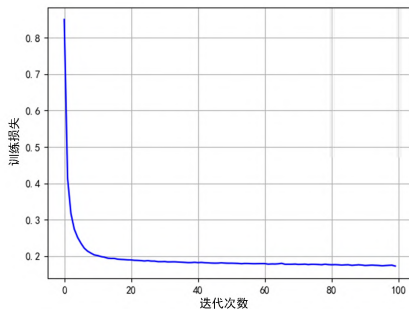


图 6 自编码器训练损失

Fig.6 Loss of Auto-coder

自编码器训练过程中的损失值随迭代次数的变化情况如图 6 所示, 可以看出, 随着网络不断迭代, 输入与输出间的重构误差即训练损失不断减小, 输入信号与重构信号不断逼近。

## 4.3 调制信号识别性能分析

本节将对调制信号识别性能进行综合探究, 首先利用第 4.2 节训练好的自编码器提取所有信号样本低维度特征, 而后将低维度特征与第 4.1 节中的手工特征融合, 共计 40 维送入特征选择器内, 通过选择后生成优选特征子集, 再利用少量带标签样本对分类器进行训练。

在本节实验中, 送入特征选择器的特征集为  $\{f_1, f_2, f_3, \dots, f_{39}, f_{40}\}$ , 其中前十维特征为手工特征  $\{f_1, \dots, f_{10}\}$ , 分别代表{奇异谱香农熵、功率谱香农熵、 $C_{40}$ 、 $C_{42}$ 、 $C_{60}$ 、 $C_{61}$ 、 $C_{63}$ 、能量谱香农熵、奇异



谱指数熵、归一化中心瞬时振幅功率密度最大值},后三十维特征 $\{f_{11}, \dots, f_{40}\}$ 为自编码器网络自动提取的编码器输出层特征。

特征选择过程中,本文设定参数 $\beta=0.38$ , $\alpha=4$ 。经过特征选择算法选择后,选出的最优特征子集为 $\{f_3, f_7, f_8, f_9, f_{10}, f_{12}, f_{16}, f_{19}, f_{21}, f_{27}, f_{32}, f_{36}, f_{37}\}$ 共计13维特征,所得特征比率为32.5%。

将特征选择后的13维特征子集送入分类器中,利用800个带标签信号对其进行监督训练。本文最终选择的分类器为浅层BP神经网络,之所以选择浅层BP网络作为特征分类器是因为浅层网络参数量级小,仅需少量训练样本就可快速收敛拟合,因此非常适用于本文设定的小样本场

景;此外,BP网络相较于传统分类算法具有可学习性的优势,可自动学习各个特征对分类结果影响的权重,这在一定程度上也起到了抑制弱相关特征的目的。BP网络结构如图7所示,网络输入层神经元个数与特征维度相同,特征输入后首先对其进行批量归一化,而后接入全连接层,全连接层神经元个数分别为16、32、16,再将全连接层的输出送入SOFTMAX分类器。各神经元均采用ReLU激活函数。为提升网络的泛化能力,在第2、第3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,本文Dropout比例设置为0.1。

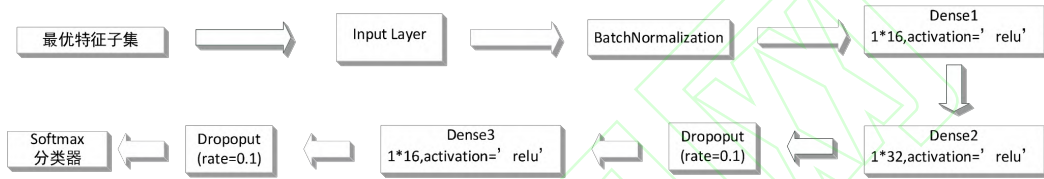
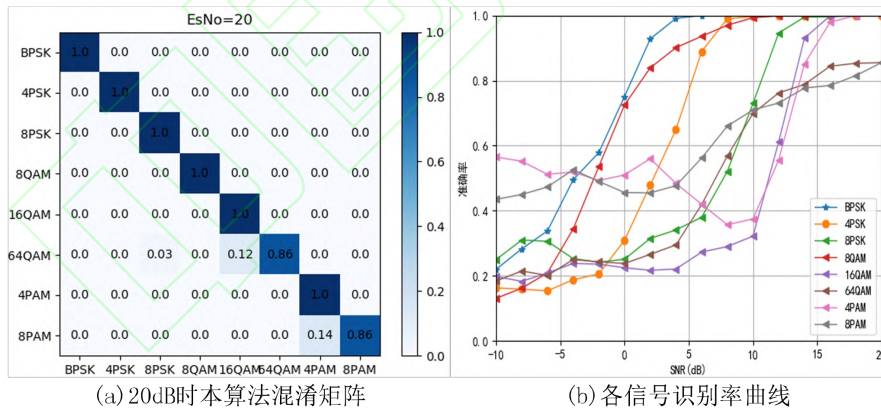


图7 BP网络结构

Fig.7 BP network structure

经过训练后分类器的识别性能如图8所示,其中图8(a)表示信噪比为20dB时的信号识别混淆矩阵,不同信噪比下各个信号的识别率见图8(b)所示,可以看出分类器在高信噪比条件下除

对64QAM、8PAM的识别有一定错误率外,其余信号均可以做到准确识别,这在小样本条件下已算较好性能。



(a) confusion matrix of this paper at 20dB

(b) Recognition Rate of Each Signal

图8 本文算法在20dB时识别性能

Fig.8 Classification performance at 20dB

当不使用特征选择算法,分别直接利用10维手工特征、30维自编码器特征、40维融合特征对BP网络分类器进行训练,得到图(9)所示在20dB时的混淆矩阵,其中图9(a)表示10维手工

特征训练所得的混淆矩阵,图9(b)表示30维自编码器特征训练所得的混淆矩阵,图9(c)表示联合10维手工特征与30维自编码器不进行特征选择直接送入BP网络训练所得的混淆矩阵。

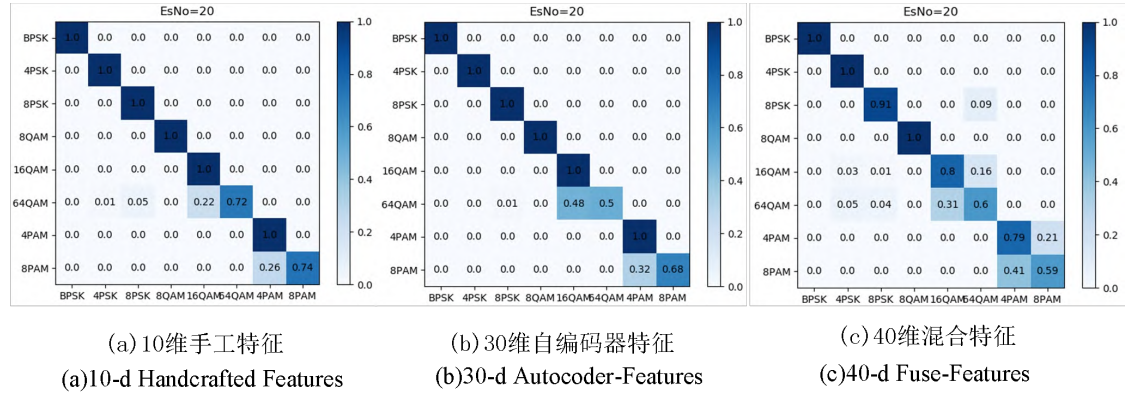


图 9 对比效果  
Fig.9 Comparison effect

通过图 9 可以看出,不论是分别采用某种单独方法或是联合特征后不进行特征优选,信号的识别性能相较本文算法较差。当利用 10 维手工特征时,信号最高识别率在 20dB 为 93.1%,利用 30 维自编码器特征时,信号最高识别率在 20dB 为 89.8%。当使用 10 维手工特征与 30 维自编码器不进行特征优选时,信号最高识别率在 20dB 仅有 83.7%。本文算法的识别率在信噪比大于 14dB 时可达 90%以上,最高识别率在 20dB 时可达 96%,可以看出通过特征融合优选后信号识别率得到了一定程度的提升。此外,可以看出当特征维度较高时信号识别性能有所下降,比如直接对 40 维特征直接进行训练,信号的最高识别率只有 83.7%,分析这是因为在小样本条件下分类器无法对高维特征较好拟合所导致的性能下降。上述 4 类方法在各个信噪比点的识别率对比曲线如图 10 所示。此外,本节选择 3 类特征选择算法作为比较对象以验证本文特征选择算法的效果,分别为 mRMR<sup>[27]</sup>、FCBF、SRFS,这其中 mRMR、FCBF 都是利用数据标签信息的有监督特征选择算法,SRFS 是基于图论的半监督特征选择算法。考虑到上述 3 个算法中均需设置相关度阈值以剔除不相关特征,因此本节设定相关度阈值  $\alpha=4$  与本文算法相同,在 SRFS 算法

中设置控制变量  $\beta=0.38$  也与本文相同。由于对比的 3 种特征选择方法都与具体分类器算法是无关的,但需通过分类算法来评估特征选择方法所选特征子集的优劣,因此本节将各算法选择出的特征集送入 BP 网络,利用 800 个带标签样本进行监督训练,对比其分类识别准确率。

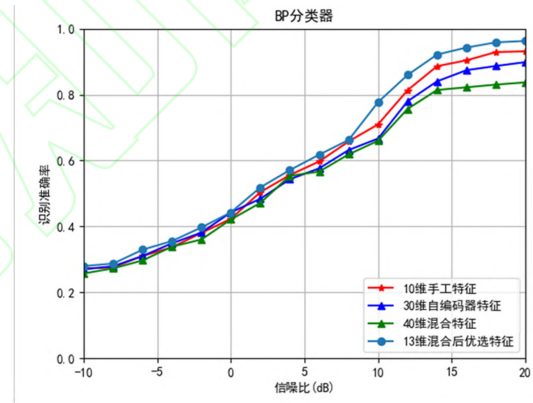


图 10 不同方法识别率对比曲线  
Fig.10 Classification rate of different methods

各特征选择算法所选特征数目、特征子集与选择时间如表 1 所示,可以看出不利用无标签样本的特征选择算法 mRMR 与 FCBF 的选择时间非常短,但其所选特征子集不够精简,而 SRFS 算法由于其需要构建有向无环图且仅从各子图中选取一个代表特征,因此其子集精简但比较耗时。

表 1 各选择算法所选特征子集  
Table1 Comparison of each selection algorithms

特征选择算法	所选特征数	特征子集	时间/s
mRMR	26	$f_2, f_3, f_5, f_6, f_7, f_8, f_9, f_{10}, f_{11},$ $f_{12}, f_{14}, f_{15}, f_{16}, f_{19}, f_{21}, f_{22}, f_{27},$ $f_{28}, f_{30}, f_{31}, f_{32}, f_{33}, f_{34}, f_{35}, f_{36}, f_{37}$	0.7

FCBF	19	$f_3, f_6, f_7, f_8, f_9, f_{10}, f_{12}, f_{14}, f_{15}, f_{16}, f_{19}, f_{21}, f_{27}, f_{28}, f_{30}, f_{32}, f_{34}, f_{36}, f_{37}$	1.3
SRFS	7	$f_3, f_9, f_{12}, f_{16}, f_{26}, f_{27}, f_{36}$	37.4
本文算法	13	$f_3, f_7, f_8, f_9, f_{10}, f_{12}, f_{16}, f_{19}, f_{21}, f_{27}, f_{32}, f_{36}, f_{37}$	9.8

各个算法所得特征比率如图 11 所示, 其中 mRMR 所选特征比率为 65%, FCBF 所选特征比率为 47.5%, SRFS 所选特征比率为 17.5%, 本文算法所得特征比率为 37.5%。

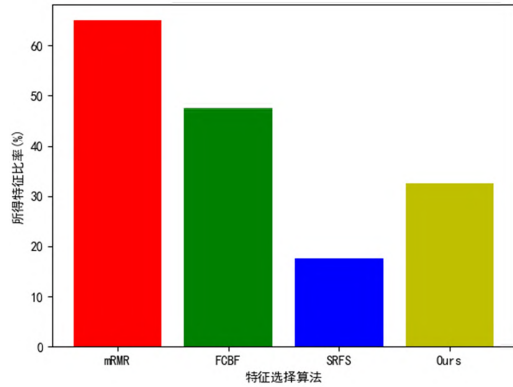


图 11 特征比率直方图

Fig.11 Feature ratio histogram

将各特征选择算法选出的特征子集送入 BP 网络, 利用少量带标签样本进行有监督训练, 所得不同信噪比下的识别率曲线如图 12 所示, 可以看出本文算法所选出的特征子集相较于其他 3 种算法在浅层 BP 网络中的识别性能有一定优势。当信噪比为 20dB 时, 利用 mRMR 算法所提特征训练出的分类器最高识别率为 90.5%, 利用 FCBF 算法所提特征训练出的分类器最高识别率为 92.1%, 利用 SRFS 算法所提特征训练出的分类器最高识别率为 93.9%, 但均低于本文算法的 96%。

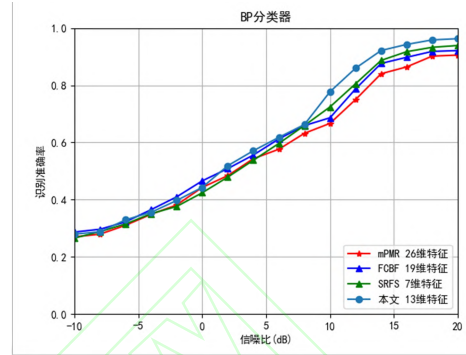


图 12 不同特征选择算法识别性能

Fig.12 Classification rate of different feature selection algorithms

为验证本文设计的浅层 BP 网络参数的最优性以及相较其他分类算法是否具有优势, 本文首先对比不同激活函数、优化器、隐藏层数条件下神经网络的识别性能, 得到了图 13 所示的仿真结果。通过实验结果可以看出, 选用 Relu 激活函数的网络识别率相较于 Sigmoid 与 Tanh 激活函数有一定优势; 基于 Adam 优化器的网络识别性能相较于 SGD 与 Adagrad 优化器有一定提升。在基于不同隐藏层数的对比实验中, 本文堆叠神经元个数为 32 的中间隐藏层, 分别对比中间隐藏层数为 1 层、3 层、5 层时不同信噪比下的识别准确率, 可以看出随着中间隐藏层数目的增加, 网络识别准确率下降幅度很大, 这是由于随着隐藏层数目的增加网络参数规模呈指数式上涨, 仅有少量的有标签训练样本根本无法优化如此量级的网络, 从而造成了严重的过拟合问题, 最终导致识别性能的急剧下降。

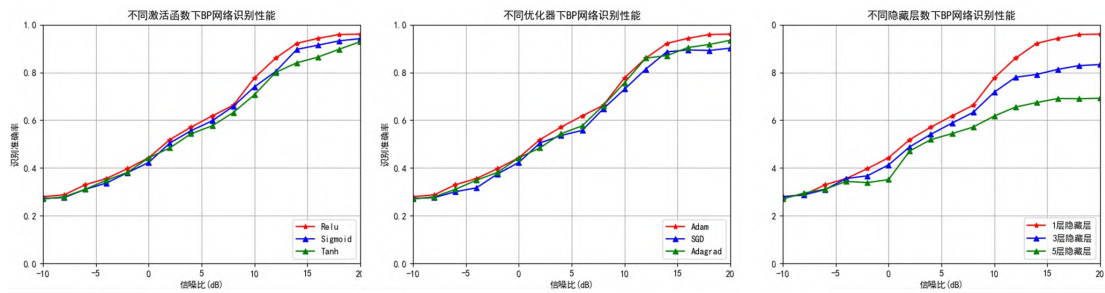


图 13 BP 网络参数性能分析

Fig.13 Performance analysis of different network parameters

最后, 本文对比 BP 网络, XGBOOST 分类器, KNN 分类器<sup>[28]</sup>的识别性能, 分别利用 800 个有标签信号的 13 维优选特征对上述 3 个分类器进行训练, 得到各个分类器的收敛速度和最高识别率, 如

表 2 所示。对比实验中, 所选的 XGBOOST 分类器为 KERAS 库中所属。可以看出浅层 BP 网络在训练时间基本相当的情况下, 识别准确率较 XGBOOST、KNN 算法也有一些优势。

表 2 不同分类器性能对比  
Table 2 Performance of different classifier

分类器	最高识别准确率	训练时间
BP 网络	96.0%	7.04s
XGBOOST	94.1%	6.26s
KNN	91.2%	8.31s

## 5 结 论

本文针对实际战场中信号调制识别领域可能出现的小样本情况进行探究, 提出了一个全新的结构, 该结构将传统方法与机器学习方法巧妙结合, 融合了分类能力强的人工特征与自编码器自动提取出的特征, 再利用特征选择算法对融合后特征进行自动选择, 从而利用尽可能低维的特征表征原始信号, 最后使用分类器训练低维度特征, 形成了小样本条件下通信信号分类新的解决方案, 通过实验也验证了本文设计方案的可行性。但本文仍存在许多不足, 比如本文算法对低信噪比条件下信号的识别率并无明显提升, 这也将作为下一步研究的重点。

## 参考文献

- [1] KHARBECH S, DAYOUB I, SIMON E, et al. On classifiers for blind feature-based automatic modulation classification over MI-MO Channels[J]. IET Communications, 2016, 10(7): 790-795.
- [2] ZHE X, YONG G. Method to reduce the signal-to-noise ratio required for modulation recognition based on logarithmic properties[J]. IET Communications, 2018, 12(11): 1360-1366.
- [3] GHASEMI S, GANGAL A. An effective algorithm for automatic modulation recognition[C]// Proc. of the IEEE 22nd Conference on Signal Processing and Communications Applications, 2014: 903-906.
- [4] SHERME A E. A novel method for auto-matic modulation recognition[J]. Applied Soft Computing, 2012, 12(1): 453-461.
- [5] WANG H, GUO L L. A new method of automatic modulation recognition based on dimension reduction[C]// Proc. of the IEEE Forum on Cooperative Positioning and Service, 2017: 316-320.
- [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, et al. Convolutional radio modulation recognition networks[C]// Proc. of the International Conference on Engineering Applications of Neural Networks, 2016: 213-226.
- [9] O'SHEA T J, ROY T, CLANCY T C, et al. Over-the-air deep learning based radio signal classification[J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(1): 168-179.
- [10] JEONG S, LEE U, KIM S C. Spectrogram-based Automatic Modulation Recognition Using Convolution Neural Network[C]// Proc. of the IEEE 10th



- International Conference on Ubiquitous and Future Networks, 2018: 843-845.
- [11] FAN M, PENG C, LENAN W, et al. Automatic modulation classification: a deep learning enabled approach[J]. IEEE Trans. on Vehicular Technology, 2018, 67(11): 10760-10772.
- [12] 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, 16(4): 1-15.
- [13] 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 Y, GUO X J, ZHAO X. Study on modulation recognition based on higher-order cumulants[J]. Journal of Southwest University of Science and Technology, 2018, 33(3): 64-68.
- [14] 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, 40(1): 171-177.
- TAN X H, ZHU G X, ZHANG X J, et al. Modulation recognition algorithm based on highorder-cumulants and wavelet transform[J]. Systems Engineering and Electronics, 2018, 40(1): 171-177.
- [15] 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 C, YANG J A, LIU H. Modulation recognition algorithm based on information entropy and GA-ELM[J]. Systems Engineering and Electronics, 2020, 42(1): 223-229.
- [16] 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): 198-206.
- [17] ALI A, FAN Y 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.
- [18] YA T, LIN Y, WANG H. Modulation recognition of digital signal based on deep auto-ancoder network[C]// Proc. of the IEEE International Conference on Software Quality, Reliability and Security Companion, 2017: 256-260.
- [19] O'SHEA T J, CORGAN J, CLANCY T C. Unsupervised representation learning of structured radio communication signals[C]// Proc. of the Learning for Intelligent Machines, 2016.
- [20] ASLAM M W, ZHU Z, NANDI A K. Automatic modulation classification using combination of genetic programming and KNN[J]. IEEE Trans. on Wireless Communication, 2012, 11(8): 2742-2750.
- [21] ANG J C, HARON H, NUZLY H, et al. Semi-supervised SVM-based feature selection for cancer classification using microarray gene expression data[C]// Proc. of the International Conference on Industrial, 2015: 468-477.
- [22] HAN Y H, YANG Y, YAN Y et al. Semi-supervised feature selection via spline regression for video semantic recognition[J]. IEEE Trans. on Neural Networks Learning System, 2015, 26(6): 252-264.
- [23] BENABDESLEM K, HINDAWI M. Efficient semi-supervised feature selection: constraint, relevance, and redundancy[J]. IEEE Trans. on Knowledge and Data Engineering, 2014, 26(5): 1131-1143.
- [24] SHEIKHPOUR R, SARRAM M A, GHARAGHANI S, et al. A survey on semi-supervised feature selection methods[J]. Pattern Recognition, 2016, 31(3): 35-45.
- [25] YU L, LIU H. Efficient feature selection via analysis of relevance and redundancy[J]. Journal of Machine Learning Research, 2004, 5(12): 1205-1224.
- [26] WANG Y T, WANG J D, LIAO H. An efficient semi-supervised representatives

feature selection algorithm based on information theory[J]. Pattern Recognition, 2017, 61(5): 11-52.

- [27] PENG H, LONG F, DING C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005, 27(4): 1226-1238.

- [28] ZHANG S C, LI X L. Learning K for KNN classification[J]. ACM Trans. on Intelligent System and Technology, 2017, 8(8): 1-19.

## 作者简介

**史蕴豪**(1996-), 男, 硕士研究生, 主要研究方向为智能通信对抗, 通信信号识处理。

E-mail: shiyunhaoai@163.com

**许 华**(1976-), 男, 教授, 博士研究生导师, 博士, 主要研究方向为通信信号处理, 盲信号处理, 通信对抗。

E-mail: 13720720010@139.com

**郑万泽**(1986-), 男, 讲师, 硕士, 主要研究方向为通信工程, 训练评估。

E-mail: 107011650@qq.com

**刘英辉**(1996-), 男, 硕士研究生, 主要研究方向为通信信号处理, 机器学习。

E-mail: YingHui\_Liu@163.com