

文章编号: 2095-4980(2020)02-0228-07

## 具有可解释性的 OFDM 雷达信号识别方法

葛 鹏<sup>1</sup>, 张文强<sup>1</sup>, 金炜东<sup>1</sup>, 郭 建<sup>1</sup>, 何贤坤<sup>2</sup>

(1.西南交通大学 电气工程学院, 四川 成都 610031; 2.中国电子科技集团 第二十九研究所, 四川 成都 610036)

**摘 要:** 针对目前正交频分复用(OFDM)雷达信号识别方法存在的问题, 提出了一种具有可解释性的 OFDM 雷达信号识别方法。该方法是通过基于树结构的流程优化(TPOT)和与模型无关的局部可理解的解释性(LIME)相结合对 OFDM 雷达信号进行识别。针对 OFDM 雷达信号特性提取了复杂度特征和基于时频图矩阵的奇异值熵, 组成特征向量; 通过 TPOT, 得到表现最佳的机器学习流程; 通过“解释器”解释预测结果, 对识别结果做出是否识别正确的风险评估, 同时可根据 OFDM 雷达信号的解释性, 得到哪些信号不易区分。实验表明, 该方法对信噪比为 0 dB 时的 OFDM 雷达信号的识别率达 91%, 通过 LIME 给出的解释性可以判断数据集中不易区分的雷达信号类型。

**关键词:** OFDM 雷达信号; 机器学习; 奇异值熵; 流程优化; 局部可理解的解释性

**中图分类号:** TN911

**文献标志码:** A

**doi:** 10.11805/TKYDA2018191

## An interpretable method for recognition of OFDM radar signals

GE Peng<sup>1</sup>, ZHANG Wenqiang<sup>1</sup>, JIN Weidong<sup>1</sup>, GUO Jian<sup>1</sup>, HE Xiankun<sup>2</sup>

(1.College of Electrical Engineering, Southwest Jiaotong University, Chengdu Sichuan 610031, China; 2.The 29th Research Institute of China Electronics Technology Group Corporation, Chengdu Sichuan 610036, China)

**Abstract:** In view of the existing problems in the current Orthogonal Frequency Division Multiplexing (OFDM) radar signal recognition method, this paper proposes an interpretable method for identification of OFDM radar signals. The method which is based on Tree-based Pipeline Optimization Tool(TPOT) and Local Interpretable Model-agnostic Explanations(LIME) is to identify OFDM radar signals. Firstly, according to the characteristics of OFDM radar signals, the complexity features and singular value entropy of time-frequency image matrix are extracted to form the feature vectors. Then through the TPOT, the best performing machine learning process is obtained. Finally, the interpretation result is interpreted by the interpreter, and the result of the recognition is given as a risk assessment; meanwhile, according to the interpretability of OFDM radar signals, those signals difficult to distinguish are determined. The experimental results show that the recognition rate of the OFDM radar signal with  $R_{SN}=0$  dB is 91%. The interpretability given by LIME can be utilized to determine the type of radar signal that is difficult to distinguish in the data set.

**Keywords:** Orthogonal Frequency Division Multiplexing radar signal; machine learning; singular value entropy; Tree-based Pipeline Optimization Tool; Local Interpretable Model-agnostic Explanations

正确识别雷达信号调制方式不仅是雷达信号参数估计的前提, 而且对雷达辐射功能推测, 进而判断雷达的威胁等级有着重要意义<sup>[1]</sup>。随着数字射频存储技术的快速发展, 雷达信号的调制方式日益复杂化, 反侦察以及反干扰能力越来越强。传统的基于脉冲数字描述特征(脉幅、脉宽、载频、到达时间(Time Of Arrival, TOA)和到达方向(Direction Of Arrival, DOA)已经无法满足现代电子战的需求<sup>[2]</sup>。本文所考虑的 OFDM 雷达信号通过多载波调制复合而成, 使其参数空间边界严重交叠, 严重影响了传统识别方法对其多种调制方式的识别效果。因此, 能否根据 OFDM 雷达信号的特性提取出适合其调制方式识别的特征显得非常重要。

目前对于雷达信号识别的研究, 多数采用各种机器学习方法。这些方法均是在特定的范围内适用, 需要人工

收稿日期: 2018-09-06; 修回日期: 2018-12-24

作者简介: 葛 鹏(1986-), 男, 博士, 主要研究方向为雷达波形设计与信号处理。email:pengGe@swjtu.edu.cn

选择算法、模型,调整超参数等,如果输入的数据集发生变化,原来的方案和参数设定将会明显影响识别的准确性。能否找到一个可以自动选择特征、算法、模型和优化流程的机器学习方法就显得尤为重要。另一方面,在实际情况中,虽然通过传统机器学习训练的模型可以得到很好的预测效果,但是对这些机器学习方法进行交叉验证有时具有误导性以及模型的不可知性,这将明显降低模型的信任度。如果能够给出雷达信号识别结果的风险评估,将对复杂调制雷达信号识别研究的实际应用具有重要意义。针对以上问题,本文提出了一种具有可解释性的 OFDM 雷达信号识别方法,该方法从 OFDM 雷达信号脉内特征分析出发,对 OFDM 雷达信号进行脉内特征参数提取,如基于时频图矩阵的奇异值熵<sup>[1]</sup>、复杂度特征<sup>[3]</sup>。通过 TPOT 自动地对一系列数据转换和模型选择,设计和优化最佳分类模型,最大限度地提高识别准确率<sup>[4]</sup>。然后通过 LIME 对 TPOT 训练出的模型采用局部逼近全局的方法<sup>[5]</sup>,获得一个局部加权可理解的模型对预测结果做出风险评估。

## 1 信号模型

相位编码 OFDM 雷达信号的模型可以表示为<sup>[6]</sup>:

$$s(t) = \sum_{n=1}^N \sum_{m=1}^M w_n u_{n,m} \text{rect}[t - (m-1)t_b] \exp[j2\pi(n-1)\Delta f t] \quad (1)$$

式中:  $w_n = |w_n|e^{j\theta_n}$  表示第  $n$  个子载波上的加权因子;  $N$  和  $M$  分别是子载波个数和每个子载波编码的码元个数;  $u_{n,m}$  为第  $n$  个子载波中的第  $m$  个码元上的相位编码;  $t_b$  表示码元时宽; 子载波频率间隔为  $\Delta f = 1/t_b$ ; 矩形窗  $\text{rect}(t) = \begin{cases} 1, & 0 \leq t \leq t_b \\ 0, & \text{other} \end{cases}$ 。带宽  $B$  由子载波数目  $N$  和子载波频率  $\Delta f$  间隔决定:  $B = N\Delta f$ , 信号持续时间  $T_s = mt_b$ 。

线性调频的 OFDM 雷达信号的模型可以表示为<sup>[7]</sup>:

$$s(t) = \sum_{n=1}^N w_n \text{rect}[t - (m-1)t_b] \exp\left\{j2\pi[(n-1)\Delta f t + \frac{1}{2}ut^2]\right\} \quad (2)$$

式中:  $u$  为调制斜率; 子载波频率间隔为  $\Delta f$ , 是  $1/mt_b$  的整数倍。带宽  $B$  由子载波数目  $N$  和子载波频率  $\Delta f$  间隔以及子载波带宽  $B_r$  决定:  $B = B_r + (N-1)\Delta f$ 。

为了减小信号的包络峰均功率比(Peak to Mean Envelope Power Ratio, PWMER), 联合子载波幅度加权和子载波初相加权。其中汉明窗  $w_n$  和 Schroeder 初相加权函数  $\theta_n$  表达式分别为:

$$|w_n| = 0.54 - 0.46 \cos \frac{2\pi(n-1)}{N}, \quad n = 1, 2, \dots, N \quad (3)$$

$$\theta_n = -2\pi \sum_{i=1}^{n-1} (n-i) |w_i|^2, \quad n = 1, 2, \dots, N \quad (4)$$

由上述信号模型可知, OFDM 雷达信号是由多路子载波复合而成, 对相位编码 OFDM 雷达信号的子载波的调制方式可采取相同编码方式, 也可采取不同编码方式, 线性调频 OFDM 信号的子载波可在整个脉冲持续时间内进行线性调频, 也可在信号整个脉冲持续时间内进行分段线性调频。

## 2 OFDM 雷达信号脉内特征提取

根据 OFDM 雷达信号的脉内特征分析发现, 由于相位编码信号相位具有随机性, 其星座密度图具有相似性, 而单码元和多码元线性调频信号的时频图区别很大, 根据以上信号特点提取了盒维数、信息维数、信息熵和时频图矩阵的奇异值熵<sup>[8]</sup>。盒维数能够度量一个几何形态的规则性, 信息维数能对疏密程度进行描述, 不同编码方式的雷达信号时频图矩阵差异性较大, 奇异值能够反映时频图矩阵的能量分布, 而奇异值熵又能够反映奇异值大小分布。因此把信号频谱的盒维数、信息维数和信号的时频图矩阵的奇异值熵作为信号调制方式的识别特征的研究具有可行性。

盒维数: 设  $A \in F(X)$ ,  $F(X)$  为一度量空间, 对每个  $\sigma > 0$ , 用  $N_\sigma(A)$  表示覆盖  $A$  的半径为  $\sigma$  最少闭球个数,

若  $D_b = \lim_{\sigma \rightarrow 0} \frac{\ln N_\sigma(A)}{\ln(1/\sigma)}$  存在, 则称  $D_b$  为集合  $A$  的分形维数。为了方便信号的盒维数计算, 采用信号重构的方法简化计算过程。用信号重构方法重构信号过程为: 用一个长度为  $N$  的信号序列  $y(n)$  进行快速傅里叶变换(Fast Fourier Transform, FFT), 得到其频谱序列  $Y(k)$  ( $k=1, 2, \dots, N$ )。将序列  $Y(k)$  置于单位正方形中, 横坐标的最小间隔

$\sigma = 1/N$ ，则有：

$$N_\sigma(A) = N + \left\{ \sum_{k=1}^{N-1} \max[Y(k), Y(k+1)]\sigma - \sum_{k=1}^{N-1} \min[Y(k), Y(k+1)]\sigma \right\} / \sigma^2 \quad (5)$$

于是有盒维数的计算公式为：

$$D_b = -\frac{\ln N_\sigma(A)}{\ln(\sigma)} \quad (6)$$

信息熵：设  $\{A(i)\}(i=1,2,\dots,N)$  是集合  $F$  的有限  $\sigma$ - 格型覆盖， $P_i$  表示  $F$  中的元素落入  $A(i)$  中的概率，计算公式为：

$$P_i = \frac{N(F)_i}{N(F \cap A_i)} \quad (7)$$

式中的  $N(F)_i$  和  $N(F \cap A_i)$  表示元素的个数。令信息熵  $I(\sigma)$  为：

$$I(\sigma) = -\sum_{i=1}^N P_i \lg P_i \quad (8)$$

信息维数：如果信息熵满足  $I(\sigma) \sim \lg D_i$  关系，则信息维数为  $DI = -\lim_{\sigma \rightarrow 0} \frac{I(\sigma)}{\lg(\sigma)}$ ，利用信号重构方法计算信息维数过程为：一个长度为  $N$  的信号序列  $y(n)$  进行 FFT 变换得到其频谱序列：

$$Y(k)(k=1,2,\dots,N), \quad A(i) = |Y(i+1) - Y(i)|, \quad i=1,2,\dots,N \quad (9)$$

根据式(9)令：

$$W = \sum_{i=1}^{N-1} A(i), \quad P_i = \frac{A(i)}{W}, \quad DI = \frac{\sum_{i=1}^{N-1} P_i \lg P_i}{\lg N} \quad (10)$$

奇异值熵：时频图矩阵的能量分布特征主要反映在较大的奇异值上，而较小的奇异值经常被当做噪声处理。在计算奇异值熵时，通常选取前 15 个较大的奇异值，求这 15 个奇异值的奇异值熵。其计算公式为：

$$H_q = -\sum_{i=1}^N p_i \log_2 p_i \quad (11)$$

式中  $p_i = \frac{\sigma_i}{\sum_{i=1}^N \sigma_i}$ ， $\sigma_i$  为奇异值， $N$  为奇异值的个数。

### 3 算法介绍

根据本文所提出的问题，构建了如图 1 所示的模型。

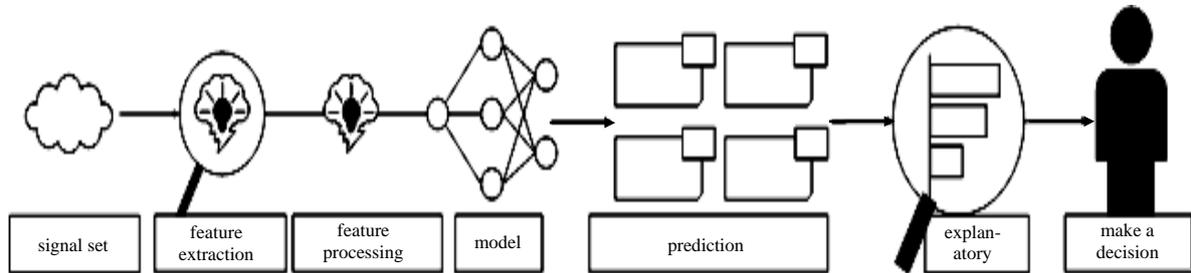


Fig.1 Algorithm model of TPOT and LIME  
图 1 TPOT 与 LIME 结合的算法模型图

通过仿真平台得到 OFDM 雷达信号集，将这些信号提取复杂度特征和基于时频图矩阵的奇异值熵特征作为 TPOT 输入，然后通过算法里面的操作算子对特征进行选择、修改和变换，再通过设置初代种群数量和遗传代数进行世代进化，优化出最佳的分类识别流程。用测试集在此模型上预测某类信号时，“解释器”首先对这个样本信号进行细微的扰动，得到扰动样本数据集，然后输入扰动样本数据集得到预测结果的相似性，从而不断地局部逼近最近开始训练的复杂模型，根据相似性以权重的方式加入到“解释器”中，得到一个简单模型，最后通

过“解释器”的简单模型解释预测信号的特征权重，为识别的信号提供是否识别准确的风险评估。

### 3.1 流程优化

目前大部分机器学习需要工程师根据经验对适用对象去选择特征，选择适合该类问题的分类器，分类器的超参数优化和模型选择以及一系列流程问题也需要工程师根据经验解决。对于不同的识别对象，流程参数设置也不尽相同。本文根据遗传编程设计的基于树结构的自动机器学习流程对信号特征集进行流程优化，自动化超参数选择，用多种算法进行建模，最后优化出一个带有参数的最佳识别流程。

#### 3.1.1 基于树的流程

在 TPOT 中，对数据处理流程如图 2 所示，其中包含了对数据处理的所有流程，在这过程中包含了许多的操作算子，把这些操作算子集成起来作为遗传编程的基元构建(Genetic Programming, GP)树，从而实现了自动机器学习。在本文中对 OFDM 雷达信号识别主要用到了以下 4 种操作算子<sup>[4]</sup>。

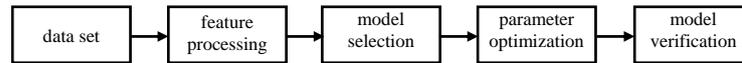


Fig.2 Flow chart of tree-based pipeline automatic machine learning

图 2 基于树结构的自动机器学习流程图

1) 特征预处理: TPOT 首先对输入的特征数据集采用样本均值、方差、样本中位数、四分位间距范围进行缩放，通过数值特征的多项式组合产生相互作用的特征。

2) 分解降维: 在经过特征预处理后，对特征用随机奇异值分解的主成分分析<sup>[9]</sup>的变体方法进行分解降维。

3) 特征选择: 根据处理后的特征信息选择 Select Kbest, Select Percentile, RFE 和 VarianceThreshold 进行特征选择，其中 Select KBest 表示选择排名前  $k$  个的特征，Select Percentile 表示选择排名前  $k\%$  的特征，RFE 表示根据递归特征消除选择，VarianceThreshold 表示设置一个方差阈值，没有达到这个方差阈值的特征将会被丢弃。

4) 模型选择: 通过以上处理后，利用遗传算法进行流程优化，找到最佳的分类模型，分类模型包括决策树分类器、随机森林分类器、Gradient Boosting 分类器<sup>[10]</sup>、支持向量机、逻辑回归和  $k$  最邻近分类算法等。

基于树的流程以原始特征数据集开始，作为树的叶子，然后通过树的节点进行 4 类主要处理: 特征预处理、分解降维、特征选择和模型选择，然后再传给下游节点。如果在某个节点输入之前有多个输入，通过特征数据集组合算子变为单一特征数据集后再通过该节点进行处理<sup>[11]</sup>。

#### 3.1.2 遗传编程

在基于树的流程优化中使用了遗传编程的方法，具体通过 DEAP 的 Python 库<sup>[12]</sup>实现。在本文中，GP 算法遵循标准进化算法程序<sup>[12]</sup>，进化过程如下:

Step1 种群初始化: 种群数量和遗传代数设置后，开始流程时就会生成固定数量的不同树形流程，以构成遗传编程中的初代种群。

Step2 个体评价: 然后根据其分类精度对这些流程进行个体评估。

Step3 个体选择: 对具有最高适应度的流程个体创建副本，然后把把这些副本放置在子代种群中，直到这些个体占种群总量的 10%。

Step4 交叉算子: 在下一代种群创建完毕后，随机地选择 2 个个体，用一个点交叉算子复制流程的百分比，在流程结构中的一个随机点进行分割，以交换彼此的内容。

Step5 突变算子: 在 Step4 中未受影响的个体按突变率选择进行替换、插入和删除突变。

在交叉和变异操作完成以后，删除上一代的个体，并以固定的代数重复这个评估—选择—交叉—变异的过程。在这个重复的过程中，GP 算法不断地优化流程。在这个优化流程中会不断地增加新的操作节点，提高适应度和去除影响效果的操作节点。

### 3.2 与模型无关的局部可理解的解释性

使用 TPOT 进行 OFDM 雷达信号的识别时，在 TPOT 训练好的模型上进行测试的结果为决策者提供依据，数据的获取方式、训练时训练数据与测试数据可能产生混合以及数据的不平衡性等会造成模型产生难以理解的结果和其他突出的问题。如果能够对预测的结果做出一定的风险评估，那将提高这个训练模型的可信度。

由于模型的复杂性及不可知性，本文用 LIME 对模型进行局部逼近得到一个简单模型，以此为预测的某类信号作出解释，从而得到是否识别准确的评估。

3.2.1 LIME 定义

LIME 的定义如下：

$$\zeta(x) = \underset{g \in G}{\operatorname{argmin}} L(f, g, \pi_x) + \Omega(g) \tag{12}$$

式中  $G$  是一类可解释模型，如线性模型、决策树或下降规则列表，将可解释性定义为  $g \in G$  模型。 $g$  的取值是  $\{0,1\}^{d'}$ ，即  $g$  表示是否存在可解释性。对于一些模型不能简单地解释，因此让  $\Omega(g)$  作为  $g \in G$  的一个复杂性的度量。例如，对于决策树  $\Omega(g)$  可能是树的深度，而对于线性模型， $\Omega(g)$  可能是非零权值的个数。在分类时， $f(x)$  是  $x$  属于某个分类的概率。 $\pi_x(z)$  为  $z$  到  $x$  之间的距离度量，以此表示  $x$  附近的地方。最后， $L(f, g, \pi_x)$  是一个度量  $\pi_x$  在定义的范围  $f$  与  $g$  相差多少的函数。为了保证解释性和局部保真度，当  $\Omega(g)$  比较小时，必须使  $L(f, g, \pi_x)$  最小化以产生可解释性。

3.2.2 模型的局部解释

在优化出训练模型后，在分类器上预测某类信号时，LIME 会对选取的这个样本进行细微的扰动，构成一个新的数据集，然后通过统计距离、相似性矩阵等，计算这个数据集中的样本与原始样本间的相似性。把所获得的数据集在分类模型上进行训练，得到样本间的相似性对预测效果的影响，以此得到一个局部加权的简单模型。然后使用贡献大的特征在简单模型上进行预测得到解释性，其中贡献大的特征有以下选取方法<sup>[9]</sup>：选择用分类模型进行预测时回归拟合上具有最高权重的特征；选择可以提高分类模型的预测的回归拟合的特征；在分类模型预测的基础上，选择正则化的收缩率最小的 lasso 预测拟合的特征；使用不多于已经选择的特征的节点数来构建决策树。

4 仿真实验

为验证本文方法的有效性，通过 Matlab 仿真平台对本文提出的方法进行仿真实验。OFDM 雷达信号的子载波数为 20，相位编码方式子载波间的频率间隔为 1 MHz，线性调频方式子载波间的频率间隔为 2.25 MHz，线性调频方式子载波带宽为 2 MHz，采样频率为 64 MHz，提取基于时频图矩阵的奇异值熵和复杂度特征，分别为信息维数、盒维数、奇异值熵、信息熵，用 feature1~feature4 命名表示。

实验 1 子载波调制方式不同的 OFDM 雷达信号识别，分别采用二相编码(2 Phase Shift Keying, 2PSK)、四相编码(4 Phase Shift Keying, 4PSK)、八相编码(8 Phase Shift Keying, 8PSK)对 OFDM 信号进行随机相位编码，码元个数  $M$  分别设置为 8,16,64，信噪比(Signal to Noise Ratio, SNR)在 0~10 dB 范围内每隔 2 dB 变化，在固定 SNR 下每类信号产生 100 个样本，提取特征构成数据集 1。在不同的信噪比下，重复 100 次实验 TPOT 对每类信号的平均识别率如表 1 所示，在 SNR 等于 6 dB 时，在测试集中随机选取一类编码方式的信号进行测试时，其识别是否正确风险评估如图 3 所示。

表 1 在数据集 1 子载波编码方式不同的 OFDM 信号的识别准确率

$R_{SN}/dB$	0	2	4	6	8	10
2PSK	91.4%	92%	94.4%	96%	97%	97.8%
4PSK	92.6%	93.8%	95%	96.6%	96.8%	97.6%
8PSK	92%	94.8%	95.8%	96%	97%	97.9%

由表 1 可知，在不同信噪比下，TPOT 对子载波调制方式不同的 OFDM 雷达信号识别准确率达到 91% 以上，表明该方法对新型雷达信号具有高识别精确度，可以优化出一个完整的识别流程。由图 3 可知，在对二相编码 OFDM 信号进行预测时，通过本文所引入的 LIME “解释器” 对其是否识别准确做出风险评估，每类信号都会给出特征解释，其中正数表示分到此类信号的特征贡献，负数表示不分到此类信号的特征贡献。对于分类到二相编码方式的信号权重达到了 0.9，而其他编码方式信号远远小于此值，表明此分类模型可信，决策者可根据给出的解释性对是否识别准确做出判断。

在特征数据集 1 上进行实验，通过 100 次实验，得到本文提出的方法与支持向量机(Support Vector Machine,

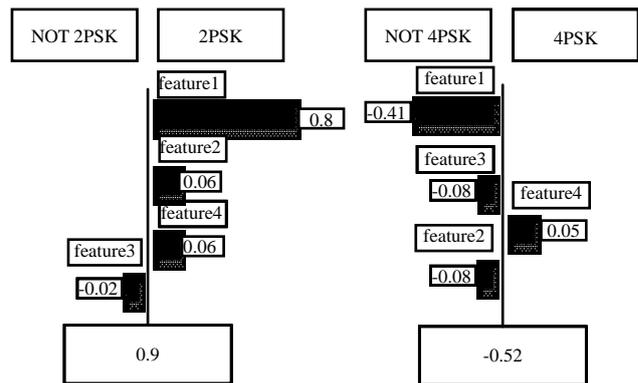


Fig.3 Risk assessment when predicting binary phase code OFDM radar signals  
图 3 预测二相编码方式的 OFDM 雷达信号的风险评估

SVM)算法识别的平均准确率对比,如图 4 所示。从图 4 可以看出,利用 TPOT 对特征数据集学习,自动地筛选出对分类效果有用的特征进行流程优化,得到最佳的分类模型,最后进行识别分类的效果要比直接采用 SVM 算法的分类效果高出 4.5 个百分点左右。说明本文采用的方法对 OFDM 信号的分类效果要好于 SVM 算法。

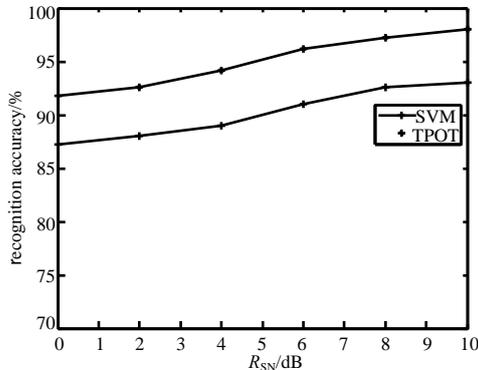


Fig.4 Comparison of recognition accuracy of TPOT and SVM algorithms  
图 4 TPOT 与 SVM 算法识别准确率对比图

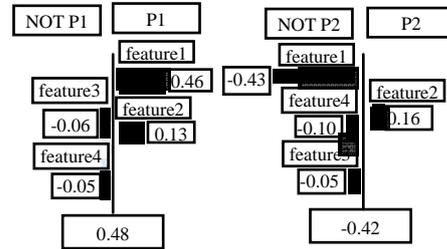


Fig.5 Risk assessment when predicting P1 coded OFDM radar signals  
图 5 预测 P1 编码方式的 OFDM 雷达信号的风险评估

实验 2 子载波调制方式相同的 OFDM 雷达信号识别,采用 P1(P1 polyphase codes), P2(P2 polyphase codes), 二相编码(Binary Phase Shift Keying, BPSK), 四相编码(Quadri-Phase Shift Keying, QPSK), 单码元线性调频(1-symbol Linear Frequency Modulation, LFM)和 8 码元分段线性调频(8-symbol Linear Frequency Modulation, 8LFM)6 种方式对子载波调制,其中 P1,P2 编码的 OFDM 雷达信号码元个数  $M$  设置为 16 和 18, BPSK 和 QPSK 编码的 OFDM 信号码元个数分别设置为 13 和 64。信噪比(SNR)在 0~10 dB 范围内每隔 2 dB 变化,在固定 SNR 下每类信号改变载频产生 100 个样本,提取特征构成数据集 2。在不同的信噪比下,重复 100 次实验 TPOT 对每类信号的平均识别率如表 2 所示,在 SNR 等于 6 dB 时,通过 LIME “解释器”对 P1 编码方式信号预测时,其识别是否正确风险评估如图 5 所示。

表 2 在数据集 2 上子载波编码方式相同的 OFDM 信号识别准确率  
Table2 Recognition accuracy of OFDM signals with the same subcarrier coding methods in data set 2

$R_{SN}/dB$	0	2	4	6	8	10
P1	91.2%	92%	93.4%	96.6%	96%	96.8%
P2	91%	92.6%	93.6%	95.2%	95%	96.2%
QPSK	91.5%	92%	94.8%	95.8%	97%	97.8%
BPSK	92.8%	93.6%	94%	95.2%	96.4%	97.2%
LFM	93.6%	94.2%	94.8%	95.8%	96%	96.8%
8LFM	94%	95.6%	96%	96.8%	96.4%	97%

由表 2 可知,在不同信噪比下,TPOT 对子载波调制方式相同的 OFDM 雷达信号识别准确率达到 91% 以上。子载波编码方式是否相同,对本文方法识别效果影响很小。根据图 5 中每个特征对分到每类信号所给出的权重可知,对于测试的 P1 编码方式的 OFDM 雷达信号能够识别正确的特征权重达到了 0.48,而其他编码方式的分类识别权重远远低于此值,表明此分类识别结果可信。经多次试验得到的效果相同,验证了本文引入的“解释器”可对预测信号是否识别准确做出评估。

实验 3 在实验 2 中仿真信号时,在 P2 编码方式的样本中随机混入 30 个 P1 编码方式样本,对信号提取特征构成数据集 3。在不同的信噪比下,重复 100 次实验 TPOT 对每类信号的平均识别率如表 3 所示,在  $R_{SN}=6$  dB 时,通过 LIME “解释器”对随机选取的某一类信号预测时,其识别是否正确风险评估如图 6 所示。

表 3 在数据集 3 上子载波编码方式相同的 OFDM 信号识别准确率  
Table3 Recognition accuracy of OFDM signals with the same subcarrier coding methods in data set 3

$R_{SN}/dB$	0	2	4	6	8	10
P1	68%	72%	75.4%	76.6%	77%	77.2%
P2	69.5%	74.6%	75.6%	77.2%	77.4%	78.1%
QPSK	91.6%	92%	93.8%	94.8%	96%	96.8%
BPSK	92.8%	93.6%	95%	95.2%	96.4%	97%
LFM	93.4%	94.9%	95.8%	96.8%	97%	97.2%
8LFM	93.8%	93.6%	95%	95.8%	97.4%	97.8%

由表 3 可得,在仿真信号时,在 P2 编码方式的样本中随机混入了 30 个 P1 编码方式样本。P1 编码方式的 OFDM 信号和 P2 编码方式的 OFDM 信号的识别准确率比较低,其他 OFDM 信号的识别率都在 91% 以上。图 6 中,对于预测 P1 信号时,通过对“解释器”给出的信号特征权重可知,在数据集 3 中 P1 和 P2 编码方式的 2 类信号存在接近的识别权重,表明通过“解释器”可以对预测的信号判断出当前识别的信号存在识别错误的风险。

险,可得到哪些类型信号不易区分。

根据以上实验可得本文提出的方法不仅可以自动优化出最佳流程,减少对超参数的选择等,而且对不同类型的 OFDM 信号的识别有较高的识别准确率。同时还可以对识别信号判断出是否存在识别错误的风险,当数据集中存在不易区分的 OFDM 雷达信号时,能够判断出哪些 OFDM 雷达信号类型易区分错误。

## 5 结论

本文根据目前 OFDM 雷达信号调制方式识别存在的问题,首先将 OFDM 信号提取复杂度特征和基于时频图矩阵的奇异值熵作为特征向量,然后采用自动生成树形结构的机器学习流程和与模型无关的局部可理解的解释性相结合的方法对雷达信号进行脉内调制方式识别。该方法不仅能够对特征数据集优化出一个最佳的分类流程,而且可通过“解释器”对识别的结果做出风险评估。当数据集中存在不可区分的雷达信号时,可根据解释性来判断出不可区分的是哪些类型雷达信号。仿真实验结果表明,使用该方法能够对 OFDM 雷达信号的调制方式进行有效的识别,并且能够对预测的结果做出评估,从而判断出存在哪些不可区分的雷达信号。

## 参考文献:

- [1] 曲志昱,毛校洁,侯长波. 基于奇异值熵和分形维数的雷达信号识别[J]. 系统工程与电子技术, 2018(2):303-307. (QU Zhiyu, MAO Xiaojie, HOU Changbo. Radar signal recognition based on singular value entropy and fractal dimension[J]. System Engineering and Electronic, 2018(2):303-307.)
- [2] ALAKHUTDINOV R, HINTON G E. Deep Boltzmann machines[C]// Proceedings of the 12th International Conference on Artificial Intelligence and Statistics. 2009:448-455.
- [3] 韩俊,何明浩,朱振波,等. 基于复杂度特征的未知雷达辐射源信号分选[J]. 电子与信息学报, 2009,31(11):2552-2556. (HAN Jun, HE Minghao, ZHU Zhenbo, et al. Sorting unknown radar emitter signal based on the complexity characteristics[J]. Journal of Electronics and Information Technology, 2009,31(11):2552-2556.)
- [4] OLSON R S, BARTLEY N, URBANOWICZ R J, et al. Evaluation of a tree-based pipeline optimization tool for automating data science[C]// Genetic and Evolutionary Computation Conference. [S.l.]:ACM, 2016:485-492.
- [5] LETHAM B, RUDIN C, MCCORMICK T H, et al. Interpretable classifiers using rules and Bayesian analysis: building a better stroke prediction model[J]. Annals of Applied Statistics, 2013,9(3):1350-1371.
- [6] 郑明峰. 相位编码 OFDM 雷达关键技术的仿真实现及抗干扰性能研究[D]. 成都:电子科技大学, 2016. (ZHENG Mingfeng. Research on key techniques and anti-jamming performance and implementation of simulation software system for phase coded OFDM radar[D]. Chengdu, China: University of Electronic Science and Technology of China, 2016.)
- [7] 陈正辉,严济鸿,何子述. MIMO 雷达 OFDM-LFM 波形设计与实现[J]. 雷达科学与技术, 2013,11(1):77-81. (CHEN Zhenghui, YAN Jihong, HE Zishu. Design for MIMO radar and implementation of OFDM-LFM waveform[J]. Radar Science and Technology, 2013,11(1):77-81.)
- [8] 徐伟,余建宇,陈矛. 基于多维特征的雷达信号脉内调制识别技术[J]. 太赫兹科学与电子信息学报, 2018,16(1):81-86. (XU Wei, YU Jianyu, CHEN Mao. Intra-pulse modulation recognition of radar signal based on multi-dimensional features [J]. Journal of Terahertz Science and Electronic Information Technology, 2018,16(1):81-86.)
- [9] MARTINSSON P G, ROKHLIN V, TYGERT M. A randomized algorithm for the decomposition of matrices[J]. Applied & Computational Harmonic Analysis, 2011,30(1):47-68.
- [10] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system[C]// Association for Computing Machinery. 2016:785-794.
- [11] OLSON R S, MOORE J H. TPOT: a tree-based pipeline optimization tool for automating machine learning[J]. Journal of Machine Learning Research, 2016(64):66-74.
- [12] PARIZEAU M. DEAP: evolutionary algorithms made easy[J]. Journal of Machine Learning Research, 2012,13(1):2171-2175.

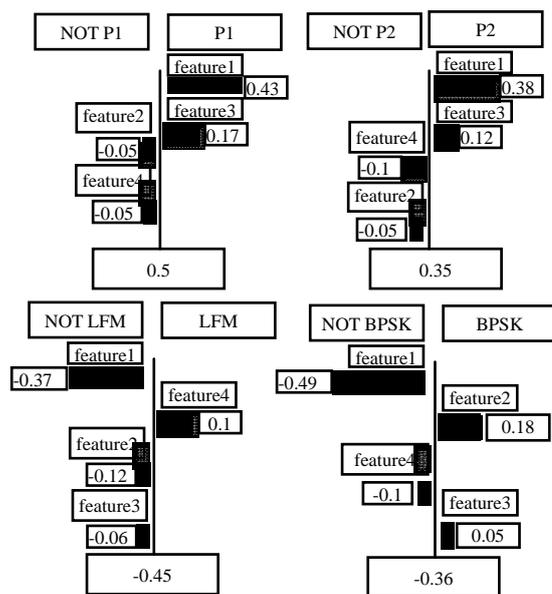


Fig.6 Risk assessment when predicting P1 coded OFDM radar signals  
图 6 预测 P1 编码方式的 OFDM 雷达信号的风险评估