Vol. 44 No. 5 September, 2022

庞聪,江勇,廖成旺,等.基于 MFCC 样本熵和灰狼算法优化支持向量机的天然地震与人工爆破自动识别[J].地震工程学报, 2022,44(5):1169-1175.DOI:10.20000/j.1000-0844.20210719002

PANG Cong, JIANG Yong, LIAO Chengwang, et al. Automatic recognition of natural earthquakes and artificial blasting based on the sample entropy of the Mel frequency cepstrum coefficient and support vector machine optimized by gray wolf optimization [J]. China Earthquake Engineering Journal, 2022, 44(5):1169-1175. DOI: 10.20000/j.1000-0844.20210719002

# 基于 MFCC 样本熵和灰狼算法优化支持向量机的 天然地震与人工爆破自动识别

庞 聪<sup>1,2,3</sup>, 江 勇<sup>1,2,3</sup>, 廖成旺<sup>1,2,3</sup>, 吴 涛<sup>1,2,3</sup>, 丁 炜<sup>1,2,3</sup>

(1. 中国地震局地震研究所, 湖北 武汉 430071; 2. 地震预警湖北省重点实验室, 湖北 武汉 430071; 3. 湖北省地震局, 湖北 武汉 430071)

摘要: 针对天然地震与人工爆破波形特征相似、难以区分的情况,结合灰狼优化算法和支持向量机,提出一种地震事件性质辨识新方法。通过梅尔频率倒谱系数法对 2013 年四川芦山地震地震事件信号和人工爆破信号进行分析,进过预加重、FFT、梅尔滤波及离散余弦变换等步骤,提取静态系数样本熵、一阶差分系数样本熵和二阶差分系数样本熵等作为样本特征集。使用灰狼算法优化支持向量机径向基核函数 RBF 中的惩罚系数和核函数半径形成新的 GWO-SVM 分类器,然后对事件进行辨识。结果表明:GWO-SVM 分类器辨识效果明显优于 SVM、RobustBoost 集成学习、LDA、PLDA等分类器,其在 1000 次循环识别实验下的准确率均值相对 SVM 提高了 9.2 个百分点,标准差降低了 3.2 以上;t 检验证明 MFCC 样本熵各特征具有可靠的地震事件分类效果;GWO-SVM 与 MFCC 样本熵可作为天然地震事件与人工爆破事件的辨识方法与分类判据。

关键词:梅尔频率倒谱系数;样本熵;灰狼算法;支持向量机;径向基核函数;自动识别中图分类号:P315文献标志码:A文章编号:1000-0844(2022)05-1169-07

DOI:10.20000/j.1000-0844.20210719002

# Automatic recognition of natural earthquakes and artificial blasting based on the sample entropy of the Mel frequency cepstrum coefficient and support vector machine optimized by gray wolf optimization

PANG Cong<sup>1,2,3</sup>, JIANG Yong<sup>1,2,3</sup>, LIAO Chengwang<sup>1,2,3</sup>, WU Tao<sup>1,2,3</sup>, DING Wei<sup>1,2,3</sup>
(1. Institute of Seismology, CEA, Wuhan 430071, Hubei, China;
2. Hubei Key Laboratory of Earthquake Early Warning, Wuhan 430071, Hubei, China;

3. Hubei Earthquake Agency, Wuhan 430071, Hubei, China)

**Abstract:** The waveform characteristics of natural earthquakes and artificial blasting are similar but difficult to distinguish. Combined with gray wolf optimization (GWO) and support vector machine (SVM), a new method for identifying the nature of seismic events is proposed in this paper. The signals of seismic events during the Yushu M7.0 earthquake in Qinghai Province and

收稿日期:2021-07-19

基金项目:湖北省自然科学基金(2019CFB768);中国地震局地震研究所和应急管理部国家自然灾害防治研究院基本科研业务费专项资助项目(IS201856290,IS2018126178,IS201726156);中国大陆综合地球物理场仪器研发专项(Y201707)

第一作者简介:庞 聪(1992一),男,工程师,主要研究方向为地震信号处理和模式识别。E-mail;ponspc@foxmail.com。

some artificial blasting events were analyzed by the Mel frequency cepstrum coefficient (MFCC) method. Through the pre-emphasis, fast Fourier transform, Mel filter, and discrete cosine transform, the sample entropies of the static coefficient and first-order and second-order differential coefficients were extracted as the sample feature set. GWO was used to optimize the penalty coefficient and kernel radius in the radial basis kernel function of SVM to form a new GWO-SVM classifier. Then, the GWO-SVM classifier was used to identify events. The results show that the recognition effect of the GWO-SVM classifier is obviously better than that of other classifiers, i.e., SVM, RobustBoost ensemble learning, linear discriminant analysis (LDA), and probabilistic LDA. Under 1 000 cycles of recognition experiments, the average accuracy of the GWO-SVM classifier increased by 9.2% compared with that of SVM, and the standard deviation was reduced by more than 3.2. The t-test proves that the MFCC sample entropy has a reliable earthquake event classification effect, and the GWO-SVM and MFCC sample entropy can be used as identification methods and classification criteria for natural earthquake events and artificial blasting events.

**Keywords:** Mel frequency cepstrum coefficient (MFCC); sample entropy; gray wolf optimization (GWO); support vector machine (SVM); radial basis kernel function; automatic recognition

# 0 引言

天然地震与振动幅值较大的人工地震动事件具 有许多相似的特性,区分研究它们的事件对地震目 录生成、地震预警算法改进、强震动仪器抗干扰等至 关重要。支持向量机方法被广泛应用于天然地震与 人工爆破、塌陷等非天然地震动事件的分类识别研 究中,取得了良好的研究成果与行业应用。如陈润 航等[1]利用 SVM 识别 186 个首都圈天然地震事件 和174个人工爆破事件;范晓易等[2]以 MATLAB Lib-SVM工具箱为基础,进行天然地震、爆破及塌陷 事件的多分类研究;黄汉明等[3]从天然地震和人工爆 破事件信号中提取小波能量熵,应用支持向量机识别 上述事件。支持向量机[1-3]在处理非线性问题上具有 天然的优势,其分类机理在于选取一个核函数(线性 核函数 Linear、多项式核函数 Poly、径向基核函数 RBF 以及神经元非线性作用核函数 Sigmoid 等)将低 维样本数据映射到高维空间,构建并寻求一个最优分 类超平面,但是其分类性能受核函数参数影响较大, 采取算法优化的方式寻求一个合适的核函数参数值 是当下 SVM 算法改进的主要方向之一。

对于事件性质辨识而言,特征判据与分类模型皆是影响识别效果的关键因素,寻找良好分类性能的判据也是众多研究人员不断尝试突破的工作方向之一。梅尔频率倒谱系数是一种应用于语音识别领域的数据处理方法,在近些年已被国内外相关人员应用到地震数据处理[4-5]中,更在2018年被陈润航等[6]应用至地震事件识别中,取得了良好分类效果,但是其直接利用梅尔频率倒谱系数(Mel Frequency

Cepstral Coefficent, MFCC)各系数值组成的 40 维系数向量作为输入样本,训练过程较为复杂,存在较多的冗余或多重共线性特征参数,直接影响学习器的分类效果,限制了该特征判据的推广与应用。

本文基于灰狼算法优化 SVM 中的径向基核函数参数值,并从样本数据中提取梅尔频率倒谱系数、一阶差分系数及二阶差分系数的样本熵作为机器学习样本数据,并利用 2013 年四川芦山 7.0 级地震和人工爆破事件波形记录验证 GWO-SVM 分类器和MFCC 样本熵特征的辨识效果。

## 1 基于灰狼算法改进的支持向量机

#### 1.1 灰狼算法

灰狼算法[7](Grey Wolf Optimization Algorithm,GWO)是一种模拟狼群狩猎活动的新型群体仿生优化算法,在 2014 年由 Mirjalili 首次提出,它模仿狼群的四层等级制度,由头狼 $\alpha$ 、次级狼 $\beta$ 、三级狼 $\delta$ 及普通狼 $\omega$ 等构成,狼群狩猎活动步骤主要分为追踪目标、围捕猎物及捕捉等,狩猎过程中的指挥命令优先级为: $\alpha > \beta > \delta > \omega$ 。GWO 算法具体步骤为:

首先狼群发现并确定目标位置,对目标进行包围,二者的距离表示为:

$$\mathbf{D} = |\mathbf{C} \times \mathbf{X}_{p}(t) - \mathbf{X}(t)| \tag{1}$$

第 t+1 次迭代后灰狼的位置为:

$$\boldsymbol{X}(t+1) = \boldsymbol{X}_{p}(t) - \boldsymbol{A} \times \boldsymbol{D} \tag{2}$$

式中:X(t+1) 为第 t+1 次迭代后灰狼的最新位置 坐标; $A \setminus C$  皆为系数调节因子; $X_p(t)$  为猎物位置向量。

各个等级的灰狼实施围捕计划时,狼会带领狼 对猎物实施追捕,追捕动作直接影响狼和猎物的位 置变化。该过程的位置更新策略为:

$$\begin{cases}
\mathbf{D}_{a} = | \mathbf{C}_{1} \mathbf{X}_{a}(t) - \mathbf{X}(t) | \\
\mathbf{D}_{\beta} = | \mathbf{C}_{2} \mathbf{X}_{\beta}(t) - \mathbf{X}(t) | \\
\mathbf{D}_{\delta} = | \mathbf{C}_{3} \mathbf{X}_{\delta}(t) - \mathbf{X}(t) | \\
\mathbf{X}_{1} = \mathbf{X}_{a}(t) - \mathbf{A}_{1} \mathbf{D}_{a} \\
\mathbf{X}_{2} = \mathbf{X}_{\beta}(t) - \mathbf{A}_{2} \mathbf{D}_{\beta} \\
\mathbf{X}_{3} = \mathbf{X}_{\delta}(t) - \mathbf{A}_{3} \mathbf{D}_{\delta}
\end{cases} \tag{4}$$

$$X_{p}(t+1) = (X_{1} + X_{2} + X_{3})/3$$
 (5)

式中: $D_{\alpha}$ 、 $D_{\beta}$ 、 $D_{\delta}$  分别表示  $\alpha$ 、 $\beta$ 、 $\delta$  狼与狼群中其他 狼的距离; $X_{\rho}(t+1)$  表示最接近猎物的灰狼位置,即最优解。

追捕到何时阶段,狼群开始攻击、捕捉猎物(最优解)。攻击速度的快慢通过调节参数 a 来完成,且 A 的取值范围为[-a,a],满足迭代停止条件后即得到最优解。

#### 1.2 GWO 优化的 SVM 识别算法

在处理非线性问题时,基于径向基核函数的支持向量机(SVM)分类效果很大程度上取决于参数惩罚系数c和核函数半径 $\sigma$ 的赋值,通过引入灰狼算法优化上述两个参数,得到最优解,从而改进SVM分类机(优化流程如图1所示)。

- (1) 通过某特征提取方法提取出若干个特征向量,建立特征数据集,同时按照一定比例划分训练集和测试集:
- (2) 初始化狼群数目 N、最大迭代次数 t、目标猎物位置自变量数等参数:
  - (3) 遍历灰狼种群,计算每个个体的适应度  $f_i$ :  $f_i = m/n$  (6)

式中:m 为被 SVM 分类器准确识别的样本数量;n 为参与 SVM 分类的总样本量。将适应度值最大的 三个灰狼位置分别记为  $X_{\alpha}$ 、 $X_{\beta}$ 、 $X_{\delta}$ ;

- (4) 计算  $\alpha$ 、 $\beta$ 、 $\delta$  狼与狼群中其他狼的距离  $\mathbf{D}_a$ 、 $\mathbf{D}_\beta$ 、 $\mathbf{D}_\delta$ ,同时更新上述狼的位置,以及刷新参数 a、 $\mathbf{A}$ 、 $\mathbf{C}$  值;
- (5) 循环计算(3)  $\sim$  (4),直至迭代次数达到迭代次数临界值 t,并得到最佳目标位置,即最优解;
- (6)设置 SVM 参数,将得到的最优惩罚系数和最优核函数半径最优值替换 RBF 核函数中默认的初始参数值,以构建新的 GWO-SVM 分类器;
- (7) 应用测试集与 GWO-SVM 模型进行地震 属性辨识,得到天然地震与人工爆破的识别结果。

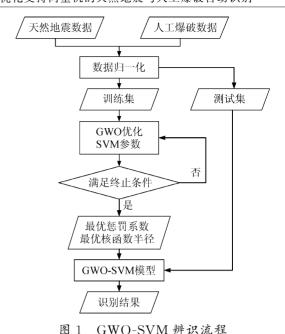


Fig.1 Identification process of GWO-SVM

# 2 数据集与特征提取

#### 2.1 数据集

选取 2013 年芦山 7.0 级地震事件的 360 条强 震动记录和 39 条人工爆破记录,其中芦山地震震中位于 30.3°N、103.0°E,震源深度 13 km;爆破事件位于中国水利水电科学研究院岩土所试验场,信号采样率为 1 000 Hz,采用 1.58 倍 TNT 当量的 RDX 球型炸药。原始记录统一截取长度为 12 000 的数据,并进行数据归一化、NaN 去除等数据预处理操作,消除台站所处环境、仪器精度、波形幅值等因素可能给数据分析带来的不利影响。图 2(a)为天然地震事件震动波形图;图 2(b)为不同人工爆破事件的固定分量信号。

#### 2.2 梅尔频率倒谱系数

梅尔频率倒谱系数<sup>[8]</sup>(MFCC)是语音识别领域中被广泛应用的特征提取算法。步骤为:

(1) 预加重:采用数字滤波器补偿初始信号中 受抑制的高频信号,该滤波器的传递函数为:

$$H(z) = 1 - \mu z^{-1} \tag{7}$$

(2)加窗:窗函数一般采用汉明窗,使得相邻两窗口之间有重叠区域,即

$$W(n,a) = (1-a) - a \times \cos\left[\frac{2\pi n}{N-1}\right],$$

$$0 \le n \le N-1 \tag{8}$$

(3) 时频域转换:对各窗口信号进行离散傅里 叶变换(DFT),将信号从时域变换到频域,其计算 公式为:

$$X_{a}(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi k/N}, 0 \leqslant k \leqslant N$$
 (9)

(4) 计算 Mel 滤波器组对数能量: 将每个窗口的频谱取模的平方得到功率谱, 并除以该窗口采样

长度,然后经 Mel 滤波器组滤波,得到第 m 个梅尔滤波器组输出能量的对数,即得到对数能量 s(m)。  $s(m) = \ln \left[ \sum_{k=0}^{N-1} |X_a(k)|^2 H_m(k) \right], 0 \leqslant m \leqslant M \ (10)$ 式中: $H_m(k)$ 为第 m 个三角滤波系数。

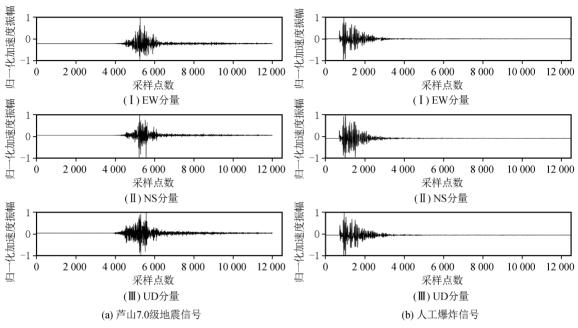


图 2 天然地震与人工爆破信号

Fig.2 Signals of natural earthquakes and artificial blasting

(5) 得到 MFCC 各阶系数: 对数能量 s(m) 经 离散余弦变换(DCT)后,即可得到第 n 阶梅尔频率 倒谱系数

$$C(n) = \sum_{m=0}^{N-1} s(m) \cos\left(\frac{\pi \times n \times (m-0.5)}{M}\right) \quad (11)$$

经过上式计算,只能得到 MFCC 静态 L 阶系数,包含原始信号的静态特性,而研究信号的动态特性就需要单独提取动态系数(一阶差分系数和二阶差分系数等),即

$$d_{t} = \begin{cases} C_{t+1} - C_{t}, & t < k \\ \frac{\sum_{k=1}^{K} k (C_{t+k} - C_{t-k})}{\sqrt{2 \sum_{k=1}^{K} k^{2}}}, & \sharp \& \\ C_{t} - C_{t+1}, & t \geqslant Q - K \end{cases}$$
(12)

式中: $d_t$  为第t 个一阶差分; $C_t$  即第t 个倒谱系数;Q 为倒谱系数  $C_t$  的阶数;K 是一阶导数时差值。

以天然地震和爆破事件中随机各取出一条记录为例,得到的 MFCC 静态系数、MFCC 一阶差分系数、MFCC 二阶差分系数、其中 Mel 滤波器的阶数设置为 24,离散余弦变换系数维度为 12,信号长度为 8 000,预加重滤波器参数取为 0.937 5,窗函数采

用汉明窗。由于 MFCC 一阶差分系数及其二阶差分系数的首尾两帧皆为 0,这里作删除处理。同时,由于 MFCC 系数的幅值与维数一般呈现显著的负相关性,维数越低,系数的幅值越大,为降低地震事件类型辨识所需的特征向量维数与特征矩阵复杂度,提高辨识效率,可只提取 MFCC 静态系数矩阵、MFCC 一阶差分系数矩阵及 MFCC 二阶差分系数矩阵的第一维系数,具体如图 3 所示。

#### 2.3 提取 MFCC 各阶系数样本熵

通过 MFCC 分析法,将得到的 MFCC 静态系数向量组、MFCC 一阶差分系数向量组及 MFCC 二阶差分系数向量组及 MFCC 二阶差分系数向量组的首维系数向量分别记为  $\{M_1(i), i=1,2,\cdots,N\}$ 、 $\{M_1'(i), i=1,2,\cdots,N\}$  及 $\{M_1'(i), i=1,2,\cdots,N\}$ ,其中 N 为 MFCC 每一维系数向量的长度。

样本熵<sup>[9]</sup>(Sample Entropy, 简称 SampEn)是一种可表征一维离散样本内部混乱状态的特征值,由学者 Richman 等于 2000 年提出。由于天然地震事件与人工爆破事件的震源机制、传播波形、瞬时最大能量、能量衰减规律等皆有明显差异,将样本熵应用至 MFCC 各系数向量的状态特征描述中是一种新的尝试。样本熵计算步骤为:

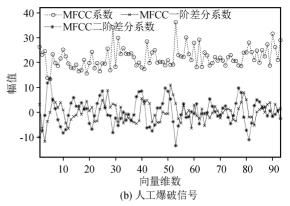


图 3 天然地震与爆破信号的 MFCC 系数首维特征提取结果

Fig.3 The first-dimensional extraction results of MFCC from signals of natural earthquakes and blasting

(1) 给定嵌入维数 m 和阈值 r ,将长度为 N 的原始信号 x(n) 重组为 m 维向量  $\mathbf{X}_m(i) = \{x(i), x(i+1), \cdots, x(i+m-1)\}$  ,并定义向量  $\mathbf{X}_m(i)$  与  $\mathbf{X}_m(j)$  的距离  $d_{ij}^m$  为对应元素 x(i+k) 与 x(j+k) 中最大差值的绝对值,其中  $k=0,1,\cdots,m-1$  ;  $i,j=1,2,\cdots,N-m+1$  且  $i\neq j$  ,即

$$d_{ij}^{m} = \max |x(i+k) - x(j+k)|$$
 (13)

(2) 针对每一个  $i(i \neq j)$  值,计算得到 N-m 个  $d_{ij}^m$  值,并组成距离向量  $\mathbf{D}^m$ , $\mathbf{D}^m$  中值小于阈值 r 的元素数目与 N-m 的比值定义为  $B_i^m(r)$ ,即

$$B_{i}^{m}(r) = \frac{\sup\{D^{m} < r\}}{N - m}$$
 (14)

对 N-m+1 个  $B_i^m(r)$  求取平均值即可得到

$$B^{m}(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_{i}^{m}(r)$$
 (15)

(3) 将维数增加至m+1,重复步骤(1)  $\sim$  (2), 得到 $B^{m+1}(r)$ 

$$B^{m+1}(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^{m+1}(r)$$
 (16)

(4) 样本熵的计算公式为

Samp
$$En(m,r) = -\ln \frac{B^{m+1}(r)}{B^m(r)}$$
 (17)

在实际应用中,嵌入维数 m 常取 2,阈值 r 为原始样本的标准差  $\times$  自定义权值,其中权值常取 0.15 或 0.2。

# 3 实验与分析

本实验硬件条件为: Intel(R) Core(TM) i5-8400 @ 2.80GHz,内存为 8 G 的 Windows10 64 位操作系统,实验平台为 MATLAB 2019a。采用国家地震科学数据共享中心(data.earthquake.cn)提供的 2013 年芦山 7.0 级地震事件的 360 条强震动三

分量记录(EW、NS、UD),及中国水利水电科学研究院岩土工程研究所(www.geoeng.iwhr.com)与陈祖煜院士团队提供的39条人工爆破事件数据,共399条加速度幅值记录,并根据本文特征提取方法从长度一致的归一化三分量数据中得到一个399×3的MFCC系数样本熵特征矩阵(图4),其特征从分别为MFCC静态系数首维向量样本熵、MFCC一阶差分系数首维向量以及样本熵MFCC二阶差分系数首维向量样本熵。

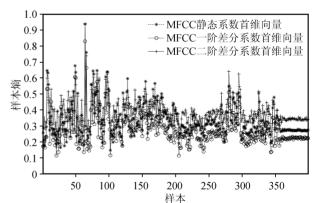


图 4 MFCC 系数样本熵特征提取结果

Fig.4 Feature extraction results for MFCC coefficient sample entropy

由图 4,人工爆破信号(样本号:361~399)与天然地震信号(样本号:1~360)存在较大的区分度,爆破信号波形特征较为一致,样本熵变化幅度较小,天然地震信号成分较复杂,不同信号波形特征差异较大,其 MFCC 系数样本熵值变化明显。MFCC 静态系数首维向量样本熵(记作 Mfcc0\_SE)、MFCC 一阶差分系数首维向量样本熵(记作 Mfcc1\_SE)以及MFCC 二阶差分系数首维向量样本熵(记作 Mfcc2\_SE)的 t 检验结果如表 1 所列,显著性水平设置为

0.05,3 个特征参数的假设检验结果均为 1,符合特征参数有明显地震事件区分能力的预期目标。

表 1 天然地震与爆破信号的单一特征向量 t 检验结果

Table 1 The t-test results of single eigenvector of signals of natural earthquakes and blasting

or meaning and ordering						
—— 特征	假设	显著差异标准	置信区间	 统计		
向量	检验	$/(\times 10^{-4})$	直行区門	偏差		
Mfcc0_SE	1	1.182 3	(0.033 6,0.102 3)	0.103 6		
Mfcc1_SE	1	3.474 1	(0.024 0,0.081 3)	0.086 5		
$Mfcc2\_SE$	1	1.096 0	(0.026 2,0.079 3)	0.080 1		

实验方案设计为:特征矩阵按 360: 39 比例随机分成训练集  $Tr_{360\times3}$ 与测试集  $Te_{39\times3}$ ,对应的标签值向量为训练集标签列向量  $Lr_{360\times1}$ 与测试集标签列向量  $Le_{39\times1}$ ,共含 2 个辨识子实验,其循环次数分别为 100 次和 1 000 次,实验结果如图 5 和表 2。

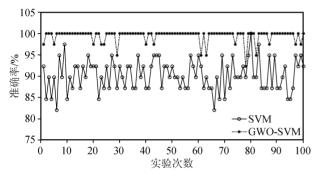


图 5 GWO-SVM 与 SVM 模型辨识结果对比 Fig.5 Comparison between identification results of GWO-SVM and SVM models

表 2 1000 次循环辨识实验结果

Table 2 Identification results of experiment under 1 000 cycles

实验	分类机 -	辨识准确率/%			
次数/次		均值	标准差	极差	变异系数
1 000	SVM	90.184 6	4.382 5	23.076 9	0.048 6
	GWO-SVM	99.407 7	1.180 2	7.692 3	0.011 9
100	SVM	90.461 5	3.960 2	20.512 8	0.043 8
	GWO-SVM	99.282 1	1.367 5	7.692 3	0.013 8

由图 5 和表 2 可看出,GWO-SVM 分类机的辨识性能各方面都优于 SVM 分类机,不论是小型实验和较大型实验,GWO-SVM 算法都保持了较为稳定的辨识效果:以 1 000 次循环辨识实验为例,GWO-SVM 的识别准确率稳健性指标—标准差、极差、变异系数分别为 1.180 2、7.692 3、0.011 9,均远低于 SVM 的 4.382 5、23.076 9、0.048 6;在 100 次辨识实验条件下,GWO-SVM 的辨识曲线与 SVM曲线边界区分明显,大部分的连续点位连线体现为水平分布趋势。这证明:GWO-SVM 算法具有更强的非线性目标求解能力与稳健性,实现了灰狼算法

优化支持向量机关键参数惩罚系数 c 与核函数半径  $\sigma$  的预定辨识目标。优化参数列于表 3。

表 3 GWO-SVM 辨识结果中的 c 与  $\sigma$ (部分)

Table 3 c and  $\sigma$  in GWO-SVM identification results

SVM 辨识率 /%	GWO-SVM 辨识率/%	С	σ
82.051 2	100	35.959 5	95.048 7
87.179 4	100	42.782 8	92.207 4
87.179 4	97.435 8	89.470 1	98.827 0
82.051 2	100	90.828 0	91.453 4
97.435 8	100	43.873 5	96.547 4
89.743 5	97.435 8	85.700 9	82.125 1
89.743 5	100	61.264 6	93.189 5
89.743 5	94.871 7	97.931 5	85.196 6
79.487 1	100	56.774 0	78.565 4
84.615 3	97.435 8	86.014 9	54.704 9
84.615 3	97.435 8	4.413 1	45.545 4
94.871 7	97.435 8	87.030 0	88.963 5

表3数据来自1000次循环辨识实验结果中随机抽取的12份惩罚系数与核函数半径优化结果,由于每一个辨识过程所需的样本数据都是随机组合的,这导致灰狼算法优化得到的惩罚系数最优值与核函数半径最优值也相差较大,但这丝毫不影响GWO-SVM分类机的辨识性能。优化后的SVM分类器辨识性能得到明显的提升,识别结果基本稳定在95%以上,证明支持向量机中惩罚系数与核函数半径的最优值选取与训练集数据息息相关,也验证了本文辨识方案的有效性和可行性。

为了验证 GWO-SVM 在地震事件类型识别中的优越性,在同样的训练集与测试集基础上,本实验采用 RobustBoost 集成学习、LDA、PLDA 等 3 种机器学习分类模型与 GWO-SVM 模型进行辨识效果对比实验,4 个模型的辨识次数与识别效果如图 6 所示。

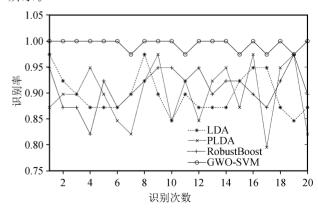


图 6 4 个辨识模型的识别率图 g.6 Identification rates of four models

由图 6 可知, GWO-SVM 模型除了明显改善

SVM 分类机识别性能,还在同类型的机器学习方法中有一定优势,具有较突出的曲线平稳特性。

### 4 结论

本文引用灰狼优化算法优化传统支持向量机, 并采用 MFCC 三维系数样本熵作为特征样本,提出 一种新型地震事件性质辨识方法。核心要点在于使 用灰狼算法优化支持向量机径向基核函数(RBF)中 的惩罚系数和核函数半径,使之最大程度地匹配训 练集和支持向量机模型,形成新的 GWO-SVM 分 类器,然后对测试集进行辨识实验。辨识结果表明:

- (1) GWO-SVM 模型性能明显优于 SVM 分类器,也优于 RobustBoost、LDA、PLDA 等学习机,灰狼算法的参数优化效果高度匹配预期目标;
- (2) MFCC 静态系数、MFCC 一阶差分系数、MFCC 二阶差分系数样本熵作为地震事件类型辨识判据,积极拓展了天然地震与人工爆破辨识研究在机器学习判据上的空白。

本文研究可作为一种探索性方法应用到中小型 地震预警系统或地震学研究中,以进行天然地震与 其他非天然地震事件的分类识别,从而降低地震事 件误判概率。但是在实际监测中,天然地震事件性 质识别除了会受到人工爆破事件的干扰外,也存在 监测仪器受到电磁干扰后生成异常波形、塌陷地震 等各种干扰事件类型。并且,学习样本数量和数据 种类如果较少,只能说明在当前实验数据下是积极 可行的,缺乏普适性实验论证。鉴于以上只是新方 法的初步探讨,下一步将重点收集各类相关数据,研 判该方法在不同震级、不同震中区域、多种地震动事 件下的辨识差异性与有效性,挖掘更多有现实意义 的地震事件类型辨识方法。

**致谢:**感谢中国地震局工程力学研究所和中国水 利水电科学研究院岩土工程研究所提供数据支撑。

#### 参考文献(References)

[1] 陈润航,黄汉明,施佳朋,等.天然地震与人工爆破地震波形的

实时分类研究[J].地球物理学进展,2019,34(5):1721-1727. CHEN Runhang, HUANG Hanming, SHI Jiapeng, et al. Study on real-time identification of natural earthquakes and artificial blasting seismic waveforms[J]. Progress in Geophysics, 2019, 34(5):1721-1727.

- [2] 范晓易,曲均浩,曲保安,等.支持向量分类机 LIBSVM 方法识别天然地震、爆破与塌陷[J].大地测量与地球动力学,2019,39 (9):916-918.
  - FAN Xiaoyi, QU Junhao, QU Baoan, et al. Support vector machine LIBSVM method for identifying natural earthquakes, blasting and collapse[J]. Journal of Geodesy and Geodynamics, 2019, 39(9):916-918.
- [3] 黄汉明,边银菊,卢世军,等.天然地震与人工爆破的波形小波特征研究[J].地震学报,2010,32(3):270-276.

  HUANG Hanming, BIAN Yinju, LU Shijun, et al. A wavelet feature research on seismic waveforms of earthquakes and explosions[J]. Acta Seismologica Sinica, 2010, 32(3):270-276.
- [4] AMENDOLA A, GABBRIELLINI G, DELL' AVERSANA P, et al. Seismic facies analysis through musical attributes[J]. Geophysical Prospecting, 2017, 65:49-58.
- [5] 周海军,李磊.地震波形的 HHT 特征提取和 GMM 识别研究 [J].黑龙江工业学院学报(综合版),2018,18(4):69-73. ZHOU Haijun,LI Lei.Feature extraction of HHT and GMM recognition of seismic waveform [J]. Journal of Heilongjiang University of Technology (Comprehensive Edition),2018,18 (4):69-73.
- [6] 陈润航,黄汉明,柴慧敏.地震和爆破事件源波形信号的卷积神 经网络分类研究[J].地球物理学进展,2018,33(4):1331-1338. CHEN Runhang, HUANG Hanming, CHAI Huimin. Study on the discrimination of seismic waveform signals between earth-quake and explosion events by convolutional neural network [J]. Progress in Geophysics, 2018, 33(4):1331-1338.
- [7] ZHU Aiijun, XU Chuanpei, LI Zhi, et al. Hybridizing grey wolf optimization with differential evolution for global optimization and test scheduling for 3D stacked SoC[J]. Journal of Systems Engineering and Electronics, 2015, 26(2):317-328.
- [8] MURTY K S R, YEGNANARAYANA B.Combining evidence from residual phase and MFCC features for speaker recognition [J].IEEE Signal Processing Letters, 2006, 13(1):52-55.
- [9] RICHMAN J S, LAKE D E, MOORMAN J R. Sample entropy [J]. Methods in Enzymology, 2004, 384:172-184.