

DOI: 10.11830/ISSN.1000-5013.201908011



# 运用 GA-SVM 模型的 砂石骨料分类方法

余罗兼<sup>1,2</sup>, 童昕<sup>1,2</sup>, 沈国浪<sup>1</sup>, 李占福<sup>2</sup>

(1. 华侨大学 机电及自动化学院, 福建 厦门 361021;

2. 福建工程学院 机械与汽车工程学院, 福建 福州 350108)

**摘要:** 提出一种能够直接从破碎工序上识别骨料成分的算法模型. 在花岗岩、石灰石、灰绿岩三种骨料剪切破碎实验的基础上, 获取其相关物理特征量; 然后, 采用遗传算法(GA)与支持向量机(SVM)相结合的算法, 对破碎、筛分后的人工砂石骨料物理特征量进行训练, 建立人工砂石骨料分类的 GA-SVM 模型. 研究结果表明: 所提 GA-SVM 模型具有较强的分类效果, 能够较好地识别人工砂石骨料成分; 与传统的 BP 神经网络、偏最小二乘法相比, 其分类准确率高.

**关键词:** 人工砂石; 骨料分类; 破碎工序; 遗传算法; 支持向量机

**中图分类号:** TU 528.041

**文献标志码:** A

**文章编号:** 1000-5013(2020)02-0137-05

## Research on Classification Method of Sand and Stone Aggregate Based on GA-SVM

YU Luo Jian<sup>1,2</sup>, TONG Xin<sup>1,2</sup>, SHEN Guo lang<sup>1</sup>, LI Zhan fu<sup>2</sup>

(1. College of Mechanical Engineering and Automation, Huaqiao University, Xiamen 361021, China;

2. School of Mechanical and Automotive Engineering, Fujian University of Technology, Fuzhou 350108, China)

**Abstract:** An algorithm model was proposed to directly identify the composition of aggregates in the process of crushing. According to the shear crushing experiments of three aggregates of granite, limestone, and gray greenstone, their relevant physical characteristics were obtained first; then, a genetic algorithm (GA) combined with support vector machine (SVM) algorithm was adopted to train the physical characteristics of artificial sand and gravel aggregate after crushing and sieving, and a GA-SVM model was established for artificial sand and gravel classification. The research results show that the proposed GA-SVM model has a strong classification effect and can better recognize the components of artificial sandstone aggregates. Compared with the traditional BP neural network and partial least squares method, its classification accuracy is high.

**Keywords:** artificial sandstone; aggregate classification; crushing process; genetic algorithm; support vector machine

随着我国城镇化水平及基础设施建设的快速发展,天然砂消耗量急速增长,天然河砂资源日趋紧缺.人工砂的重要性也就日趋明显,将充当混凝土细骨料的补充材料、甚至完全取代天然砂<sup>[1]</sup>.在岩石分类的研究上,最为传统的做法就是采用人工方式在偏光显微镜下对岩石薄片图像进行人工鉴定,这需要

**收稿日期:** 2019-08-07

**通信作者:** 童昕(1964-),男,教授,博士,博士生导师,主要从事机电系统动态分析与控制的研究. E-mail: xtong@fjut.edu.cn.

**基金项目:** 国家自然科学基金青年科学基金资助项目(51905100);福建省教育厅中青年教师教育科研资助项目(JT180342)

专业人员进行判读,工作量大且易受到主观因素影响<sup>[2]</sup>.刘凤英等<sup>[3]</sup>提出采用 BP 神经网络对岩石的氧化物含量为特征量建立分类模型.袁颖等<sup>[4]</sup>提出采用主成分分析的遗传算法和支持向量机相结合的算法,同样对岩石的氧化物含量为特征量建立分类模型. Harmon 等<sup>[5]</sup>采用激光诱导击穿光谱系统(LIBS)采集岩石光谱数据,结合偏最小二乘判别分析法(PLS-DA)识别火山岩的类型.

人工砂开采后需要经过机械破碎、筛分等工序制成机制砂.在破碎、筛分的过程中,自然可以形成一些表征人工砂石骨料特征的物理量.因此,本文提出采用遗传算法(GA)-支持向量机(SVM)分类模型,对破碎、筛分后的人工砂石骨料物理特征量进行训练,建立骨料分类模型.

## 1 经遗传算法优化的支持向量机基本原理

支持向量机(SVM)是由 Vapnik 等提出的一种研究小样本、小概率事件的,基于统计学习理论的神经网络模型<sup>[6]</sup>.其基本思想就是通过一个非线性映射,将数据映射到高维特征空间,将实际问题转化为一个带不等式约束的二次规划问题<sup>[7]</sup>,即寻找一个最优超平面进行分类<sup>[8]</sup>.

遗传算法(GA)是基于达尔文生物进化过程启发的优化算法<sup>[9]</sup>.其将每一个解  $t$  看作染色体中的基因编码,通过反复的基因重组、突变等操作生产新的基因编码,最终求得  $O(t)$  的最优解  $t^*$ .

GA 优化支持向量机,主要是优化 SVM 的惩罚参数  $C$  和核参数  $g$ ,其算法流程如图 1 所示.具体有如下 5 个主要步骤.

**步骤 1** 对数据进行预处理.对训练样本数据和测试样本数据进行归一化处理,消除量纲差异.

**步骤 2** 染色体编码与种群初始化.对 SVM 的惩罚参数  $C$  和核参数  $g$  进行二进制编码<sup>[10]</sup>,并产生初始化种群.

**步骤 3** GA 判断优化是否终止.对种群进行选择、交叉、变异、计算适应度,直至满足 GA 终止条件,输出此次优化的最优参数.

**步骤 4** SVM 精度判定.将最优惩罚参数  $C$  和核参数  $g$  代入 SVM 中进行分类,判断是否符合给定精度;如不满足,重新进行步骤 3.

**步骤 5** 测试集进行分类.将满足精度的 GA-SVM 分类模型对测试集进行分类,输出结果.

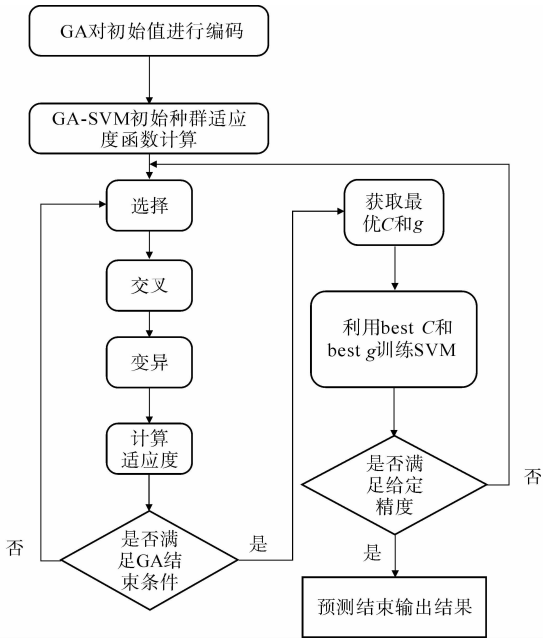


图 1 GA 优化 SVM 算法流程

Fig. 1 Flow chart of GA optimized SVM algorithm

## 2 破碎实验与数据采集

实验的人工砂石骨料分别取自宁夏的花岗岩、福鼎的石灰石和浙江的灰绿岩,在砂石颗粒破碎试验台(图 2)上进行单核破碎实验.

将花岗岩、石灰石、灰绿岩等三类骨料各选取 35~65 mm 之间粒径的试件分 30 组,通过游标卡尺测量大小径;采用密度仪测量质量与密度,并计算出其体积;在破碎试验台上进行破碎实验,并通过采集卡记录剪切力的变化及峰值;对破碎后的颗粒进行收集分类,用密度仪测量每个颗粒的质量、密度和体积.将花岗岩标记为类别 A,石灰石标记为类别 B,灰绿岩标记为类别 C,每种岩石 4 组,实验数据如表 1 所示.表 1 中: $m$  为质量; $\phi_L$  为最长粒径; $\phi_S$  为最短粒径; $V$  为体积; $\rho$  为密度; $n$  为碎石块数; $\bar{m}$  为碎块平均质量; $\bar{V}$  为碎块平均体积; $\varphi_{\max}$  为破碎后最大粒径; $\tau_{\max}$  为最大剪切力.



图 2 砂石颗粒破碎试验台

Fig. 2 Gravel crushing test equipment

表 1 三种岩石破碎物理特征量  
Tab. 1 Three physical characteristics of rock fragmentation

类别	$m/g$	$\phi_L/mm$	$\phi_S/mm$	$V/cm^3$	$\rho/g \cdot cm^{-3}$	$n$	$\bar{m}/g$	$\bar{V}/cm^3$	$\varphi_{max}/mm$	$\tau_{max}/N$
A1	53.16	45	28	20.33	2.62	3	17.03	6.51	22.60	548
A2	54.07	42	37	20.52	2.64	2	22.78	8.65	23.17	983
A3	57.97	40	23	21.40	2.71	2	27.72	10.23	26.85	810
A4	64.25	57	26	24.88	2.58	2	31.53	12.21	25.81	809
B1	37.12	55	24	13.38	2.77	2	14.59	5.26	15.90	205
B2	31.95	52	17	11.80	2.71	2	15.93	5.88	25.49	254
B3	25.76	40	25	9.41	2.74	2	12.12	4.43	18.64	367
B4	22.01	45	18	8.11	2.71	2	10.86	4.00	23.48	227
C1	25.67	44	21	9.48	2.71	3	8.32	3.07	15.41	354
C2	20.91	40	17	7.71	2.71	2	10.04	3.70	19.88	331
C3	23.85	43	13	8.89	2.68	2	11.79	4.39	14.22	305
C4	23.85	43	13	8.89	2.68	2	11.79	4.39	14.22	305

3 GA-SVM 分类模型建立与实验结果分析

3.1 模型数据配置

模型建立硬件环境为 Win10,i5,8 GB 内存,软件环境基于 MATLAB2018a 运行. GA-SVM 分类模型数据采用表 1 的数据,每种石头试件样品做 30 组破碎实验,每组破碎实验数据中有 10 个特征量,三种石头试件样品总共 90 组数据;然后,随机选择 63 组作为训练数据集,剩余 27 组测试数据集,并对其进行归一化处理.

3.2 SVM 核函数的选取

核函数的选择对支持向量机的性能有着至关重要的作用,如果核函数选择不合适,意味着将样本映射到一个不合适的特征空间,将会导致其性能不佳<sup>[11]</sup>. 因此,需要进行核函数选择的对比分析. 常见的核函数有线性核函数、多项式核函数、径向基核函数,以及 Sigmoid 核函数. 选择不同的核函数对分类结果进行对比分析,结果如表 2 所示.

表 2 不同核函数下的 SVM 分类效果  
Tab. 2 SVM classification effect under different kernel functions

数据集规模	核函数类型	正确率/%			正确率 平均值/%	训练平均 时间/s
		样本分类 1	样本分类 2	样本分类 3		
90×10	线性核函数	62.963	55.556	70.370	62.963	36.30
	多项式核函数	55.556	59.259	51.852	55.556	21.97
	径向基核函数	62.963	59.259	77.778	66.667	11.83
	Sigmoid 函数	59.259	51.852	51.852	54.321	8.96

从表 2 可知:选用线性核函数的分类正确率平均值为 62.963%,训练平均时间为 36.30 s;而选用径向基核函数的分类正确率平均值为 66.667%,训练平均时间为 11.83 s. 在正确率平均值上,两者相差不是很大,但在训练平均时间上,径向基核函数训练更快. 径向基核函数相比线性核能够处理分类标注和属性的非线性关系,同时具有简单实用、普适性好的优点<sup>[12]</sup>. 其他两种核函数的正确率平均值远远低于径向基核函数. 因此,确定 SVM 的核函数选择径向基核函数.

3.3 K 折交叉验证的选择

支持向量机可以以交叉验证法的准确率作为适应度函数. 交叉验证是消除取样随机性造成训练偏差的方法,使用交叉验证能够有效评价训练模型性能,提升模型稳定性与泛化能力<sup>[13-14]</sup>. 采用 K 折交叉验证方法,核函数选择径向基核函数,支持向量机暂未经过 GA 优化,分别将 K 取 5 和 10,验证分类效果,结果如表 3 所示.

从表 3 可知:当 K 取 5,8,10 时,其正确率平均值分别为 70.370%,56.790%,58.025%,训练时间分别为 11.969,21.034,26.401 s. 因此,选用五折交叉验证. 未经 GA 优化的选择径向基核函数、五折交

叉验证的 SVM 分类结果,如图 3 所示.

表 3 不同  $K$  值下的 SVM 分类结果  
Tab. 3 SVM classification results at different  $K$  values

$K$	正确率/%			正确率平均值/%	训练平均时间/s
	样本分类 1	样本分类 2	样本分类 3		
5	62.963	77.778	70.370	70.370	11.969
8	59.259	55.556	55.556	56.790	21.034
10	59.259	51.852	62.963	58.025	26.401

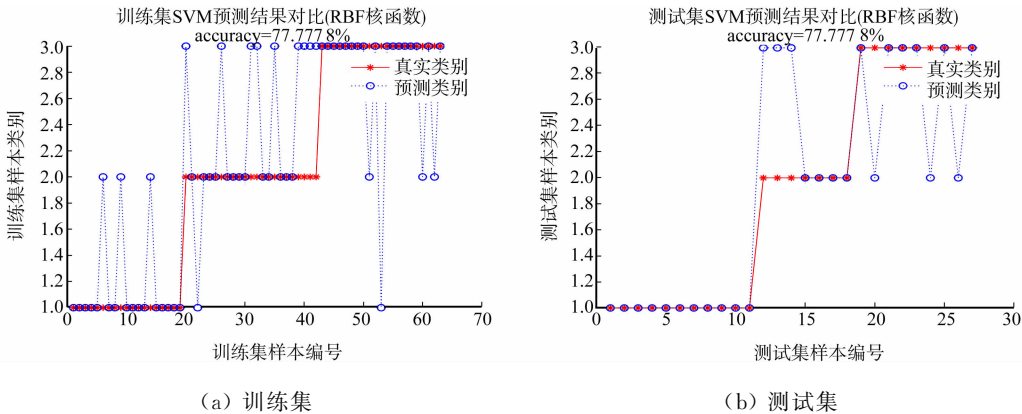


图 3 未经 GA 优化的选择径向基核函数的 SVM 分类结果

Fig. 3 SVM classification results of RBF kernel function without GA optimization

3.4 GA 优化的 SVM 分类效果分析

选用径向基核函数的 SVM,惩罚参数  $C$  和核参数  $g$  对 SVM 的泛化性能有很大的影响<sup>[15]</sup>,因此,采用遗传算法 GA 对其进行寻优. 配置 SVM 核函数为径向基核函数,进行五折交叉验证,参数组合寻优空间范围为  $[0, 100]$ ,种群规模选择为 50,进化迭代次数为 300 次,交叉概率和变异概率选择默认值,结果分别如图 4,5 所示.

从图 4 可知:遗传算法适应度随着迭代次数的增加而增加,当迭代到第 38 次的时候收敛至稳定值,可得最优惩罚参数  $C_{best}=1.3195, g_{best}=0.2176$ .

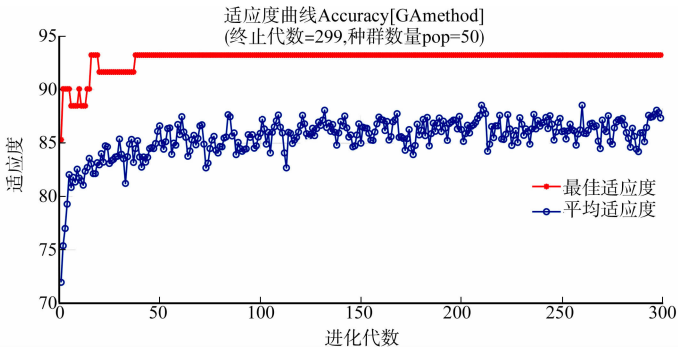


图 4 GA 适应度变化曲线  
Fig. 4 GA fitness curve

对比图 3 和图 5 可知:在人工砂石骨

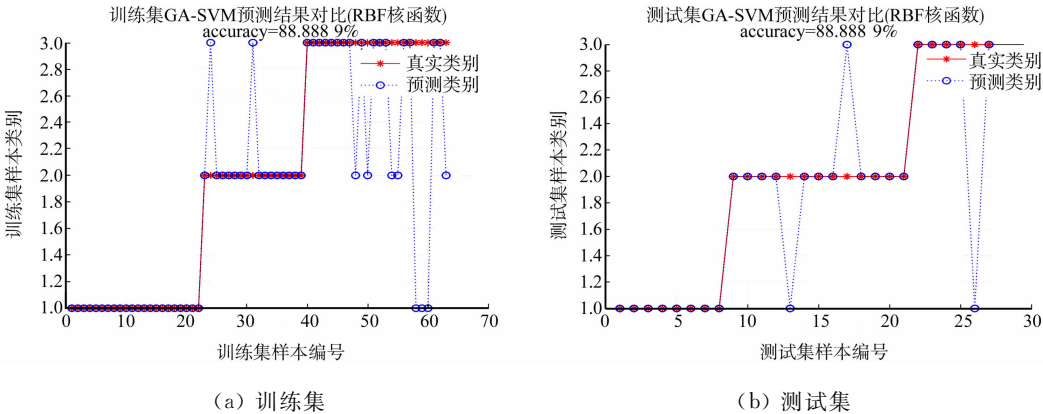


图 5 经 GA 优化的选择径向基核函数的 SVM 分类结果

Fig. 5 SVM classification results with RBF kernel function optimized by GA

料分类的模型建立基础上,与未经 GA 优化的 SVM 相比,经过 GA 优化的 SVM 的岩石识别正确率提升了约 11%,正确率达到了 88.89%。在测试集样本中,花岗岩的识别正确率同为 100%,识别错误的地方都存在于石灰石与辉绿石上。

## 4 结束语

文中研究了人工砂石骨料分类识别的方法。在分析了传统岩石分类工作量任务大、主观因素强,所提岩石分类方法预处理成本高等缺点的基础上,提出了一种能够直接从破碎工序上识别骨料成分的算法模型。实验结果表明:基于 GA-SVM 的人工砂石骨料分类模型能够较好地识别人工砂石骨料成分,相比 BP、PLS 算法具有优势,为人工砂石骨料特征上的研究提供了一种新思路。在未来的研究工作中,需要将 GA-SVM 模型运用在其他常见的人工砂石骨料上,进行分类模型的建立,进一步验证此方法的有效性,为砂石骨料的智能加工打下基础。

## 参考文献:

- [1] 李俊杰,陈家全,唐宇,等. 人工砂胶砂流动度评价方法的研究[J]. 广东建材,2019,35(7):6-8.
- [2] 吉春旭. 基于卷积神经网络的岩石图像分类研究与应用[D]. 西安:西安石油大学,2017.
- [3] 刘凤英,张楠,郎伟民. BP 网络在火成岩岩石化学判别分类中的应用[J]. 新疆地质,2007,25(2):140-143. DOI:10.3969/j. issn. 1000-8845. 2007. 02. 003.
- [4] 袁颖,李绍康,周爱红. 基于 PCA-GA-SVM 的火成岩分类方法研究[J]. 数学的实践与认识,2017,47(12):121-128.
- [5] HARMON R S, THROCKMORTON C S, HARK R R, *et al.* Discriminating volcanic centers with handheld laser-induced breakdown spectroscopy (LIBS) [J]. Journal of Archaeological Science, 2018, 98: 112-127. DOI: 10. 1016/j. jas. 2018. 07. 009.
- [6] 温廷新,孙红娟,徐波,等. 矿区采空塌陷危险性预测的 RS-SVM 模型[J]. 中国安全科学学报,2015,25(10):16-21. DOI:10. 16265/j. cnki. issn1003-3033. 2015. 10. 003.
- [7] 龙文,梁昔明,龙祖强,等. PSO-LSSVM 灰色组合模型在地下水埋深预测中的应用[J]. 系统工程理论与实践,2013,33(1):243-248. DOI:10. 3969/j. issn. 1000-6788. 2013. 01. 031.
- [8] 邱昌桂,孔兰芬,杨式华,等. 基于 GA-SVM 算法的烤烟香型自动识别研究[J]. 烟草科技,2019,52(2):101-108. DOI:10. 16135/j. issn1002-0861. 2018. 0311.
- [9] ZHANG Guohui, GAO Liang, SHI Yang. An effective genetic algorithm for the flexible job-shop scheduling problem [J]. Expert Systems with Applications, 2011, 38(4): 3563-3573. DOI: 10. 1016/j. eswa. 2010. 08. 145.
- [10] 温廷新,于凤娥,邵良杉,等. 基于 GA-SVM 的隧道围岩分类研究[J]. 公路交通科技,2018,35(9):63-70. DOI:10. 3969/j. issn. 1002-0268. 2018. 09. 010.
- [11] 梁礼明,陈明理,刘博文,等. 基于图论的支持向量机核函数选择[J]. 计算机工程与设计,2019,40(5):1316-1321. DOI:10. 16208/j. issn1000-7024. 2019. 05. 022.
- [12] 孙波,梁勇,汉牟田,等. 基于 GA-SVM 的 GNSS-IR 土壤湿度反演方法[J]. 北京航空航天大学学报,2019,45(3):486-492. DOI:10. 13700/j. bh. 1001-5965. 2018. 0417.
- [13] 吐松江·卡日,高文胜,张紫薇,等. 基于支持向量机和遗传算法的变压器故障诊断[J]. 清华大学学报(自然科学版),2018,58(7):623-629. DOI:10. 16511/j. cnki. qhdxxb. 2018. 25. 032.
- [14] 李亦滔. 基于支持向量机的改进分类算法[J]. 计算机系统应用,2019,28(10):145-151. DOI:10. 15888/j. cnki. csa. 007080.
- [15] 朱超岩,姚晓东. 基于遗传算法优化的支持向量机在变压器故障诊断中应用[J]. 仪表技术,2019(3):21-23,32. DOI:10. 19432/j. cnki. issn1006-2394. 2019. 03. 006.

(责任编辑:黄仲一 英文审校:崔长彩)