

基于混沌鲶鱼效应的人工蜂群算法及应用

王生生^{1,2}, 杨娟娟^{1,2}, 柴 胜^{1,2}

(1. 吉林大学计算机科学与技术学院, 吉林长春 130012; 2. 吉林大学符号计算与知识工程教育部重点实验室, 吉林长春 130012)

摘 要: 针对目前人工蜂群算法的早熟收敛、陷入局部极值等问题, 提出一种基于混沌鲶鱼效应的改进人工蜂群算法. 首先, 采用随机性更高的混沌序列初始化蜂群以扩大其遍布范围; 其次, 集成了鲶鱼效应和混沌理论提出了混沌鲶鱼蜂, 并引入了它与陷入局部极值的蜂群之间的有效竞争协调机制, 从而增进蜜蜂群体跳出局部最优解、加速收敛的能力. 支持向量机的学习能力主要取决于其惩罚因子 C 和核函数参数的合理选择, 对其参数的优化可以提升其学习效果, 然而现行算法均存在一定局限性. 基于我们提出的改进人工蜂群算法, 对支持向量机的参数进行了优化. 最后, 在 UCI(加州大学欧文分校)数据集和行为识别真实数据集上进行了测试, 验证基于改进人工蜂群算法的支持向量机具有更强的分类性能.

关键词: 人工蜂群算法; 混沌理论; 鲶鱼效应; 支持向量机; 行为识别

中图分类号: TP319 **文献标识码:** A **文章编号:** 0372-2112 (2014)09-1731-07

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2014.9.011

Artificial Bee Colony Algorithm with Chaotic Catfish Effect and Its Application

WANG Sheng-sheng^{1,2}, YANG Juan-juan^{1,2}, CHAI Sheng^{1,2}

(1. College of Computer Science and Technology, Jilin University, Changchun, Jilin 130012, China;

2. Key Laboratory of Symbolic Computing and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, Jilin 130012, China)

Abstract: There are the disadvantages of easily falling into premature convergence and local optimal solution which the elementary artificial bee colony algorithm had in some degree. Chaotic Catfish effect was hence adopted in this paper to achieve the optimum performance of artificial bee colony algorithm, in which, chaotic mechanism was conducted to instantiate each individual of the swarm firstly owing to its marvelous intrinsic randomness. Then the efficacious competition and coordination mechanism among Catfish bees which were derived from the integration of Chaos theory with Catfish effect and originals were intended to boost the capabilities of them leaping out of local optimal solution and converging expeditiously. The practicability of Support Vector Machines (SVM) is excessively affected due to the difficulty of selecting appropriate penalty factor C and kernel function parameter of SVM. Conversely, all of the common SVM parameters optimization methods have their respective disadvantages with some degree of competence. We utilized the improved artificial bee colony algorithm to optimize the two parameters of SVM, simultaneously, the public datasets from the University of California-Irvine (UCI) and the activity recognition reality data were employed for evaluating the proposed model. Experimental results demonstrate that the classification accuracy obtained by the developed SVM was higher

Key words: artificial bee colony algorithm; chaos theory; catfish effect; support vector machine; activity recognition

1 引言

人工蜂群算法(Artificial Bee Colony algorithm, ABC)是 Karaboga^[1]于 2005 年提出的一种求解函数优化问题的进化算法, 也是继蚁群算法(Ant Colony Optimization, ACO)、粒子群算法(Particle Swarm Optimization, PSO)之后的又一新型群体智能随机优化算法. 该算法控制参数少、实现简单、计算简洁且在解决某些复杂问题上比 ACO 和 PSO 颇具优势, 目前被广泛应用于多种研究领

域^[2,3]. 然而, 同 ACO 算法、PSO 算法类似, 随着迭代搜索接近尾声, 蜂群个体的多样性及搜索速度会逐渐降低而种群整体出现早熟收敛、陷入局部极值的可能性却显著提高. 鉴于混沌系统的类随机性、规律性及互不相关性等特征, 文献[4~6]特将其作为一种提高各自算法性能的有效机制.

本文一方面通过使用混沌映射方法代替原一般随机化方法初始化蜜蜂种群, 在不失随机性的同时加强了蜂群的多样性; 另一方面, 则采用有别于文献[5,6]在

ABC算法中加入混沌局部搜索的做法,将鲶鱼效应^[7]施加到蜜蜂种群内,并以混沌波动的方式衍生出一种新型的混沌鲶鱼蜂,用来代替原侦查蜂负责开采新蜜源的任务,且同原蜂群个体之间形成了有效的竞争、协作机制,以此提高整个种群的搜索速率及增强个体跳出局部最优解的能力。

支持向量机(Support Vector Machines, SVM)是由Vapnik^[8]等人根据统计学习理论提出的一种非常适用于解决小样本、非线性及高维模式识别问题的新型机器学习方法。在利用训练样本建立分类模型的过程中, SVM的惩罚因子、核函数及核函数参数的选择与设置对SVM的最终分类结果均有着至关重要的影响^[9];因此,对SVM模型参数寻优问题的探究逐渐成为该领域的研究热点^[10~14],其中文献[12~14]更是巧妙地将ACO算法、PSO算法和ABC算法融入到SVM最佳参数的搜索求解过程中,它们各自提出的ACO-SVM模型、PSO-SVM模型及ABC-SVM模型在同基于网格搜索法(Grid Search Method, GSM)和遗传算法(Genetic Algorithm, GA)优化的SVM性能对比时,均表现出较强的优势。然而,群智能算法本身易受初始条件制约且都存在早熟收敛和陷入局部极值等的可能性,故目前并未获得有效的SVM参数寻优模式。本文将上述过程得到的混沌鲶鱼蜂群算法作为优化SVM参数的一种策略,并经UCI数据集和行为识别真实数据集测试验证表明该方法明显提升了对未知行为的判别准确率。

2 基于混沌鲶鱼效应的人工蜂群算法

2.1 ABC算法基本原理

ABC算法受自然界蜜蜂采蜜机理的启示,将整个蜂群分为引领蜂、跟随蜂和侦查蜂三类。其中,引领蜂和跟随蜂的个数均为种群数量的一半,侦查蜂的个数为1;引领蜂和跟随蜂负责食物源的开采过程,而侦查蜂执行新食物源的探索过程。该算法在求解优化问题时,每个食物源的位置对应优化问题的一个可行解,蜜源的含蜜量则代表求解问题的适应度值,最优食物源便是优化问题的最优解,蜜蜂搜索食物源的速度等同于优化目标的求解速度。ABC算法的整体执行过程可概括为:蜜蜂种群初始化部分、引领蜂和跟随蜂搜索求解部分和侦查蜂开采新蜜源部分。

2.1.1 蜜蜂种群初始化

ABC算法首先根据式(1),随机初始化生成NF个d维食物源向量 $F_i = (f_{i1}, f_{i2}, \dots, f_{id})$, $i = 1, 2, \dots, \text{NF}$;然后,计算出每个食物源所对应的适应度值。

$$f_{i,j} = f_{i,j}^{\min} + \text{rand}(0,1) \times (f_{i,j}^{\max} - f_{i,j}^{\min}) \quad (1)$$

式中, $j \in \{1, 2, \dots, d\}$, $f_{i,j}$ 为 F_i 的第j维分量, $f_{i,j}^{\max}$ 和 $f_{i,j}^{\min}$ 为 $f_{i,j}$ 的上、下限。

2.1.2 引领蜂和跟随蜂搜索求解

引领蜂通过式(2),对相应的食物源进行邻域搜索,若搜索到的食物源的含蜜量更高,则用新食物源位置 $S_i = (s_{i1}, s_{i2}, \dots, s_{id})$ 代替旧食物源位置 F_i ,否则,将继续保持原位置不变。

$$s_{i,j} = f_{i,j} + \eta_{i,j} \times (f_{i,j} - f_{k,j}) \quad (2)$$

式中, $k \in \{1, 2, \dots, \text{NF}\}$ 代表随机选取的不同于i的食物源, $\eta_{i,j}$ 是 $[-1, 1]$ 上的一个随机数。

当所有的引领蜂完成搜索回到蜂巢并把自身携带的信息传达给跟随蜂后,跟随蜂则按一定的概率 P_i 选择所要跟随的引领蜂,再按照式(2)在其所选择的食物源附近完成一次邻域搜索,并判断是否需要更新蜜源。

$$P_i = \text{fit}_i / \sum_{i=1}^{\text{NF}} \text{fit}_i \quad (3)$$

式中, fit_i 代表第i个食物源的适应度值且一般通过式(4)计算求得:

$$\text{fit}_i = \begin{cases} 1/(1 + \text{val}_i), & \text{if } \text{val}_i \geq 0 \\ 1 + \text{abs}(\text{val}_i), & \text{if } \text{val}_i < 0 \end{cases} \quad (4)$$

其中, val_i 表示与第i食物源 F_i 相对应的目标函数值。

2.1.3 侦查蜂开采新蜜源

若某个食物源 F_i 在连续经过“limit”次循环后一直都未得到替换,该处的引领蜂便会转变成侦查蜂。侦查蜂根据式(1)在解空间中随机产生一个新位置来代替 F_i ,并记录下新蜜源的适应度值。

综上所述,ABC算法主要执行了四个选择过程:

- ①跟随蜂按照式(3)发现较好食物源的全局选择过程;
- ②引领蜂和跟随蜂按照式(2)进行邻域搜索的局部选择过程;
- ③引领蜂和跟随蜂对新旧食物源进行判断,保留优异蜜源的贪婪选择过程;
- ④侦查蜂按照式(1)发现新蜜源的随机选择过程。

2.2 应用混沌鲶鱼效应改进人工蜂群算法

2.2.1 蜂群混沌初始化

种群初始化是进化算法中的一个重要环节,它不仅影响算法的收敛速度而且制约问题最终解的质量。鉴于混沌运动能够在指定范围内按其自身“规律”无重复地遍历所有状态,本文将通过混沌映射充分地提取和捕捉解空间里的信息,以增强蜜蜂种群的多样性。

在混沌理论研究中,应用较为广泛的一种映射机制是一维的Logistic映射^[15],其数学迭代方程式为:

$$\lambda_{t+1} = \mu \times \lambda_t (1 - \lambda_t), t = 0, 1, 2, \dots, T \quad (5)$$

式中, T 是预设的最大混沌迭代次数, λ_t 为区间 $[0, 1]$ 上均匀分布的随机数且 $\lambda_0 \notin \{0, 0.25, 0.5, 0.75, 1\}$, μ 是混沌控制参数,当 $\mu = 4$ 时,系统将处于完全混沌状态。

起初,利用式(5)产生一组混沌变量 λ_t ;接着,运用混沌序列依次将NF个d维食物源向量 F_i 按照式(6)映

射到混沌区间 $[F_{\min}, F_{\max}]$ 内, 其中 $F_{\max}$ 和 $F_{\min}$ 代表 F_i 的最大、最小边界; 最后, 根据式(4)计算出每个食物源对应的适应度值。

$$f_{i,j} = f_{\min,j} + \lambda_{i,t} \times (f_{\max,j} - f_{\min,j}) \quad (6)$$

2.2.2 混沌鲶鱼蜂开采新蜜源

鲶鱼效应是指在某个基本懈怠停滞的自治系统中, 通过适当引入若干异体强者而使系统整体重新恢复活力的一种负激励策略, 它出自于贩卖沙丁鱼为生的挪威商人仅凭几条鲶鱼即可救活整槽沙丁鱼的典故。这种凭借对手的牵制刺激而使自身得到提升的有效竞争机制, 在金融界和企业管理中得到了广泛应用, 文献[16]将鲶鱼效应运用到粒子群算法中, 通过鲶鱼粒子的介入, 进而成功地改善了算法求解问题的能力。

本文将鲶鱼效应融合到 ABC 算法中。若引领蜂和跟随蜂在搜索过程中遇到了历经“limit”次仍未更新的食物源, 表明蜂群已经陷入局部最优值, 下一代蜜蜂个体便会在该极值附近停滞, 这不但削弱了整个种群的搜索能力, 而且可能导致算法最终无法收敛于全局最优值。类似地, 我们将处于此状态下的蜜蜂种群视为“沙丁鱼”群, 通过引入若干混沌鲶鱼蜂来诱导引领蜂和跟随蜂向新的广阔搜索空间移动, 以增强算法获得全局最优值的几率。其具体操作如下:

(1) 升序排列当前 NF 个食物源的适应度值, 并标记出前 $10\% \times NF$ 个最差蜜源位置;

(2) 淘汰最差蜜源处的蜜蜂, 以等量混沌鲶鱼蜂代替, 设鲶鱼蜂个数为 M , 则 $M = 10\% \times NF$;

(3) 按式(7)设置混沌鲶鱼蜂初始位置向量 $V_m^d(0) = (v_m^1(0), v_m^2(0), \dots, v_m^d(0))$, $m = 1, 2, \dots, M$

$$v_m^j(0) = \begin{cases} v_{m,j}^{\max}, & \text{if } r > 0.5 \\ v_{m,j}^{\min}, & \text{if } r \leq 0.5 \end{cases} \quad (7)$$

其中, $v_{m,j}^{\max}$ 和 $v_{m,j}^{\min}$ 分别表示第 m 个鲶鱼蜂第 j 维上的最大、最小值; r 是 $[0, 1]$ 内的一个随机数, 并令 r 以 0.5 为界点, 以使得 v_m^j 在两极值处的取值分布近似。

(4) 混沌鲶鱼蜂根据式(8)进行位置更新, 其中, T 为混沌迭代阈值;

$$\begin{cases} V_m^d(t) = V_{\min} + (V_{\max} - V_{\min}) \times V_m^d(t), & V_m^d \in [V_{\min}, V_{\max}] \\ V_m^d(t+1) = \mu \times V_m^d(t) \times (1 - V_m^d(t)), & t = 0, 1, \dots, T \end{cases} \quad (8)$$

(5) 利用式(3)与式(4)计算各混沌鲶鱼蜂所处食物源的适应度值, 以替换原蜂群中适应度最小的 M 个值, 并同引领蜂和跟随蜂一起进入下一次循环搜索过程。

3 SVM 参数优化应用

3.1 支持向量机

SVM 的核心思想就是构造一个最优分类超平面,

进而能够将两类待分样本集准确地区分开。设输入样本集为 (x_i, y_i) , $x_i \in R^d$, $y_i \in \{+1, -1\}$; 就线性可分问题而言, 求解最优分类超平面所对应的二次规划问题为:

$$\begin{aligned} \min_{w, b} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & \begin{cases} y_i(w^T x_i + b) + \xi_i - 1 \geq 0 \\ \xi_i \geq 0 \end{cases} \end{aligned} \quad (9)$$

式中, w 为可调整的系数向量, b 为偏移量; ξ_i 为松弛变量, 用以衡量样本被错分的程度; C 为非负惩罚因子, 用来获取最少错分样本与最大分类间隔 $2/\|w\|$ 两者间的平衡。可见, C 是决定 SVM 学习能力和经验风险协调度的一个关键因素; 若 C 过大, 就会引起过学习, 致使分类器的泛化能力降低, 反之, 则会导致分类器的分类正确率过低, 甚至使整个分类器模型失效。

利用 Lagrange 法求解式(9), 最终得到的最优分类超平面函数为:

$$\begin{aligned} f(x) &= \text{sgn}\{(w^T x) + b\} \\ &= \text{sgn}\left\{\sum_{i=1}^n \alpha_i^* y_i (x_i^T x) + b^*\right\} \end{aligned} \quad (10)$$

式中, $\text{sgn}(x)$ 为符号函数, α_i^* 为支持向量所对应的 Lagrange 系数; b^* 为分类阈值, 它既可以用任一个满足式(9)的支持向量求得, 也可通过两类中任意一对支持向量取中值求得。

通过在式(9)中引入核函数 $K(x_i, x)$, 以将低维空间内的线性不可分样本映射到高维空间内而使之成为线性可分的, 便可得到非线性分类问题下的最优分类超平面函数:

$$f(x) = \text{sgn}\left\{\sum_{i=1}^n \alpha_i^* y_i K(x_i, x) + b^*\right\}, \quad (11)$$

$$K(x_i, x) = f(x_i)^T \times f(x)$$

其中, 径向基核函数 RBF 是最常使用的核函数^[17], 相应的数学表达式为:

$$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2), \gamma > 0 \quad (12)$$

综上所述, 惩罚参数 C 和 RBF 核函数参数 γ 的设置是影响 SVM 分类器性能好坏的主要因素。

3.2 混沌鲶鱼蜂群算法优化 SVM 参数模型 (CCABC-SVM)

把 SVM 惩罚因子 C 和 RBF 核参数 γ 的组合变量 (C, γ) 作为改进 ABC 算法的寻优搜索对象, 以找到使 SVM 分类正确率最高化的组合变量值, 是本文混沌鲶鱼蜂群算法优化 SVM 参数模型的最终目标。该模型所需参数及适应度函数的设置如下:

首先, 初始化种群规模; 其中包括: 蜜源数目 NF、食物源位置 (C, γ) 、混沌序列的最大迭代次数 T 、蜜源最

大更新次数 limit 和循环终止次数 MC.

其次,确定模型适应度函数形式;本文算法是将 SVM 分类器的分类正确率作为评价准则,故适应度函数定义为:

$$\text{fit} = \frac{1}{1 - \text{Accuracy}}$$

(13)

式中,分类正确率 Accuracy 由 k 折交叉验证法求得,平均交叉验证识别率获得. k 折交叉验证法是指:将样本数据集随机划分成 K 个子集,依次把其中的一个子集作为测试集,剩余的 $K-1$ 组子集作为训练集,重复进行 K 次训练,而得到 K 个分类器模型的过程,再利用得到的这 K 个分类器的分类准确率的平均值作为整个数据集分类器的性能指标,即式(13)中的 Accuracy 值.

最后,设定惩罚因子 C 及 RBF 核参数 γ 的搜索范围;用以减少模型求解的时间^[13].

CCABC-SVM 模型的整体运行过程,如图 1 所示:

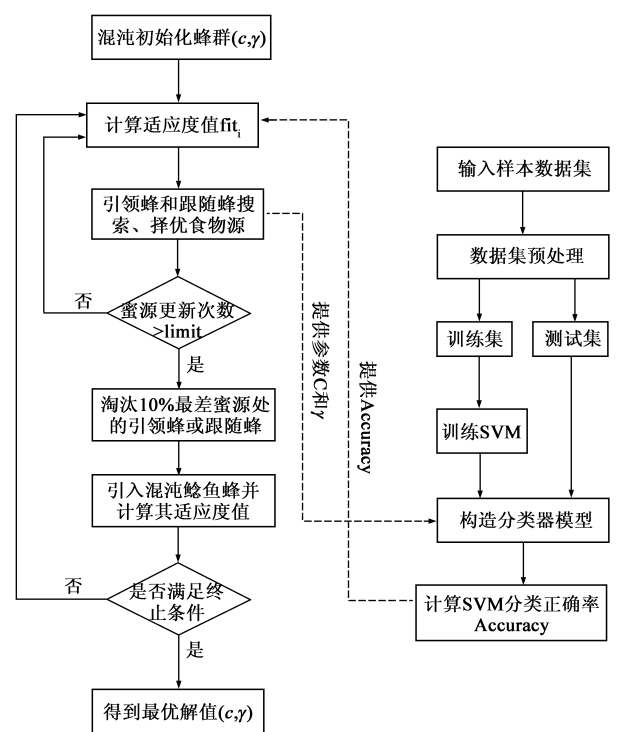


图1 模型整体流程图

4 CCABC-SVM 模型实验及应用

4.1 UCI 数据测试

本文提出的 CCABC-SVM 参数优化模型通过 MatlabR2013a 编译实现,算法中的 SVM 由 LIBSVM^[18] 开源软件包提供,为了验证该模型的有效性和评估其性能,使用来自 UCI 中的 8 个标准数据集来完成实验测试,数据集的相关信息如表 1 所示.实验平台在 Intel Core2、2.8GHz CPU、2.0GB 内存、操作系统为 Windows7 的计算

机上搭建.

表 1 UCI 数据集说明

数据集名	样本种类数	样本特征数	样本总数	训练样本数	测试样本数
Australian	2 类	15	690	350	340
Heart	2 类	13	270	170	100
Iris	3 类	4	150	80	70
Ionosphere	2 类	34	351	200	151
Glass	6 类	9	214	120	94
Sonar	2 类	60	208	108	100
Vowel	11 类	10	528	278	250
Wine	3 类	13	178	90	88

为方便 LIBSVM 对数据的处理以及消除因数据尺度不统一对分类准确率所造成的影响,数据集中的数据均先按照式(14),归一化到区间 $[-1, +1]$ 上.

$$x' = 2 \times \frac{x - L_i}{U_i - L_i} - 1$$

(14)

其中, x' 为处理后的值, x 为原数据值, U_i 和 L_i 则分别表示第 i 个特征取值的上、下界.

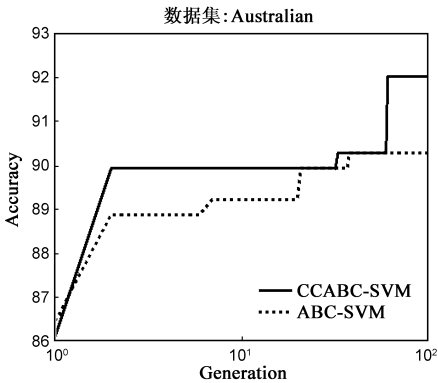
便于对比其他模型的测试结果,经实验将本模型中的各参数设置为:蜜源数目 $NF = 20$,食物源最大更新次数 $\text{limit} = 5$,最大循环次数 $MC = 100$,最大混沌迭代次数 $T = 300$,惩罚因子 C 的取值范围为 $[0.01, 35000]$,RBF 核函数参数 γ 的范围为 $[0.0001, 32]$,交叉验证度 $k = 10$.然后,利用 CCABC-SVM 模型和 ABC-SVM 模型共同在上述 8 个 UCI 数据集上进行分类测试实验,并将它们各自得到的 Accuracy 同特征染色体遗传算法 (Genetic Algorithm with Feature Chromosomes, FCGA) 优化 SVM 参数模型 (FCGA-SVM)^[11]、蚁群算法优化 SVM 参数模型 (ACO_R-SVM)^[12] 及粒子群算法优化 SVM 参数模型 (PSO-SVM)^[13] 依次对比:

由表 2 可知,CCABC-SVM 模型较其他四种模型在处理分类问题上更具优势,在提高 SVM 分类正确率的同时也强化了 SVM 的学习性能.为便于更加具体地说明 CCABC 算法拥有比基本 ABC 算法更好的求解能力,将它们在处理 2 类样本集 Australian、3 类样本集 Iris、6 类样本集 Glass 和 11 类样本集 Vowel 的分类问题中,所得到的循环迭代次数 Generation 和每代最优分类正确率 Accuracy 之间的关系以曲线图的形式分别绘制如图 2.

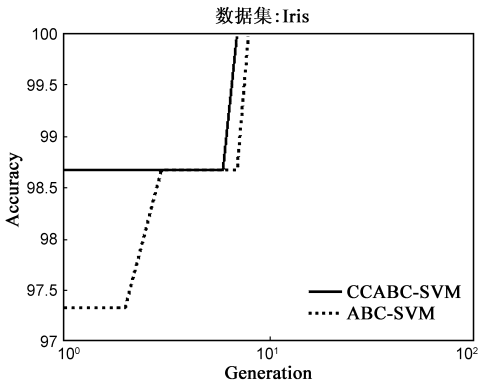
图 2 中,曲线的倾斜程度反映了算法寻优的速度,水平线表示种群已陷入局部最优解,曲线的拐点代表种群已跳出局部极值.从图中可明显看出,与 ABC 算法相比,混沌鲶鱼蜂群算法具有更高效的搜索寻优速度,即使进入迭代后期,算法同样具有很强的摆脱局部停滞的能力;尽管在数据集 Iris 上两个算法都获得了 100% 的分类正确率,但是 CCABC 算法比基本 ABC 算法更早收敛于全局最优值.

表 2 模型的分类正确率

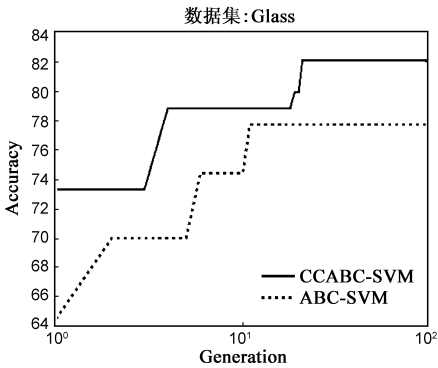
数据集名	CCABC-SVM	ABC-SVM	ACO _R -SVM	PSO-SVM	FCGA-SVM
Australian	92.02%	90.21%	91.90%	88.09%	86.81%
Heart	96.25%	93.75%	94.87%	88.17%	91.11%
Iris	100.00%	100.00%	99.87%	98.00%	96.00%
Ionosphere	99.34%	98.17%	97.76%	97.50%	98.57%
Glass	82.23%	77.78%	—	79.00%	80.80%
Pima-Indian Diabetes	88.42%	82.31%	88.00%	80.19%	81.97%
Sonar	98.78%	95.56%	98.36%	88.32%	95.00%
Vowel	100.00%	99.51%	—	99.27%	98.89%
Wine	100.00%	100.00%	—	99.56%	100.00%



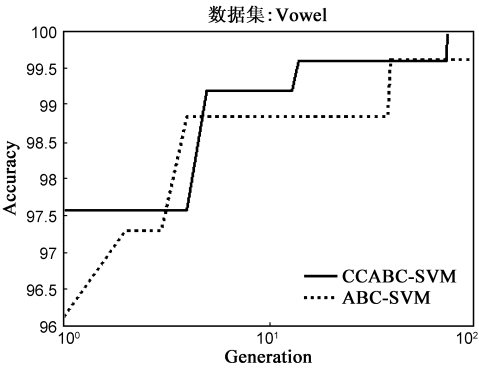
(a) 数据集Australian的性能对比图



(b) 数据集Iris的性能对比图



(c) 数据集Glass的性能对比图



(d) 数据集Vowel的性能对比图

图2 不同人工蜂群算法在4个数据集性能曲线

4.2 行为识别真实数据测试

随着智能手机与人们日常生活联系的日益紧密,智能手机的功能也逐渐多样化,其内嵌的各式传感器不仅能为用户提供语音识别、位置锁定及健康监测等服务,而且也因它们具有比其他用于行为姿态识别研究的专用硬件设备更低的成本且携带便捷、不易受外界环境限制等优势,目前已被广泛应用于对特定的行为动作进行实效数据采集上^[19-21].然而,基于智能手机传感器的行为识别模型,均需要附加额外的分类器对这些已有的行为数据进行学习,以便通过预测或判别某未知行为来方便人们做出最佳决策.

通过利用智能手机里的加速度传感器和陀螺仪传感器对10名志愿者的7种日常行为进行信号采集,这7

种日常行为分别是:站立、行走、就坐、平躺、上楼、下楼和慢跑.记录下各行为所对应的加速度等传感器值并附上行为标签,在经过降噪等处理后,最终获得了行为识别真实数据集.

将上述收集的数据集输入到CCABC-SVM模型中进行行为识别测试,并与文献[19]中基于网格搜索的SVM(GSM-SVM)方法及传统SVM方法的学习结果进行对照,如表3所示.其中,三分之二的记录作为训练集,剩余的用作测试集;鉴于数据集的特性及便于结果对比,现将蜜源的数量NF设为40,食物源最大更新次数limit为50,最大循环次数MC为1000,最大混沌迭代次数T为300,惩罚因子C的取值范围为[0.1,1000],RBF核函数参数γ的范围为[10⁻⁵,10],交叉验证度k

为 10.

表 3 三种 SVM 的识别结果

行为活动	CCABC-SVM(%)	GSM-SVM(%)	SVM(%)
站立	100	97.8	97.1
行走	93.7	87.2	87.2
就坐	99.4	94.7	95.6
平躺	100	100	100
上楼	94.6	81.7	87.2
下楼	92.8	74.0	72.6
慢跑	90.4	—	—
平均识别率	95.9	89.0	89.3

由表 3 所示,本文所提的 CCABC-SVM 在 7 种行为活动中均获得了较高的识别率,特别是明显提高了对动态活动行为的辨别能力,这在一定程度上也预示了本模型亦可用于更复杂的行为活动分辨上.

5 结论

惩罚因子 C 和 RBF 核函数 γ 是决定支持向量机学习能力的两个重要因素,本文将混沌鲶鱼蜂群算法应用于求解此问题,进而提出一种新的 CCABC-SVM 寻优参数模型,并从可行性和实用性两方面对其有效性加以验证.测试结果表明本模型具有比基本 ABC 算法更快的收敛速度和更强的跳出局部极值的能力,明显地改善了 SVM 的学习能力,提高了其分类性能.在未来研究中,将考虑采用特征选择理论对样本数据集进行预处理,以消除其中的冗余特征,使得 SVM 的学习能力进一步提升并做出更准确的判断.另外,本文提出的混沌鲶鱼蜂群算法同样适用于解决其他核函数的 SVM 参数优化问题.

参考文献

[1] Karaboga D. An Idea Based on Honey Bee Swarm for Numerical Optimization[R]. Kayseri: Erciyes University, Engineering Faculty, Computer Engineering Department, 2005.

[2] 高卫峰,刘三阳,黄玲玲.受启发的人工蜂群算法在全局优化问题中的应用[J].电子学报,2012,40(12):2398 – 2403.

Gao Wei-feng, Liu San-yang, Huang Ling-ling. Inspired artificial bee colony algorithm for global optimization problems[J]. Acta Electronica Sinica, 2012, 40(12): 2398 – 2403. (in Chinese)

[3] 张银雪,田学民,邓晓刚.基于改进人工蜂群算法的盲源分离方法[J].电子学报,2012,40(10):2026 – 2030.

Zhang Yinxue, Tian Xuemin, Deng Xiaogang. Blind source separation based on modified artificial bee colony algorithm[J]. Acta Electronica Sinica, 2012, 40(10): 2026 – 2030. (in Chinese)

[4] Xie Chunli, Shao Cheng, Zhao Dandan. Parameters optimization

of least squares support vector machines and its application[J]. Journal of Computers, 2011, 6(9): 1935 – 1941.

[5] Yu Jieyue, Lin Jian. The ink preset algorithm based on the model optimized by chaotic bee colony[A]. Proceedings of the 5th International Congress on Image and Signal Processing. [C]. America: IEEE Computer Society, 2012. 547 – 551.

[6] 李志勇,李玲玲,等.基于 Memetic 框架的混沌人工蜂群算法[J].计算机应用研究,2012,29(11):4045 – 4049.

Li Zhi-yong, Li Ling-ling, et al. Chaos artificial bee colony based on Memetic framework[J]. Application research of Computers, 2012, 29(11): 4045 – 4049. (in Chinese)

[7] Zhang Likang. An analysis of common search in Chinese of Google the meta library[J]. Data Science Journal, 2007, 6 (Supplement): 813 – 823.

[8] Vapnik V. The Nature of Statistical Learning Theory[M]. New York: Springer Science Business Media, 2000. 10 – 60.

[9] Keerthi S S, Lin C J. Asymptotic behaviors of support vector machines with Gaussian kernel[J]. Neural Computation, 2003, 15(7): 1667 – 1689.

[10] Friedrichs F, Igel C. Evolutionary tuning of multiple SVM parameters[J]. Neurocomputing, 2005, 64(Special): 107 – 117.

[11] Zhao Mingyuan, Fu Chong, et al. Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes[J]. Expert Systems with Applications, 2011, 38(5): 5197 – 5204.

[12] Alwan H B, Ku-Mahamud K R. Solving support vector machine model selection problem using continuous ant colony optimization[J]. International Journal of Information Processing and Management, 2013, 4(2): 86 – 97.

[13] Lin S W, Ying K C, et al. Particle swarm optimization for parameter determination and feature selection of support vector machines[J]. Expert Systems with Applications, 2008, 35(4): 1817 – 1824.

[14] 刘路,王太勇.基于人工蜂群算法的支持向量机优化[J].天津大学学报,2011,44(9):803 – 809.

Liu Lu, Wang Taiyong. Support vector machine optimization based on artificial bee colony algorithm[J]. Journal of Tianjin University, 2011, 44(9): 803 – 809. (in Chinese)

[15] Schuster H G, Wolfram J. Deterministic Chaos: An Introduction[M]. Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA, 2005. 19 – 68.

[16] Chuang L Y, Tsai S W, Yang C H. Fuzzy adaptive catfish particle swarm optimization[J]. Artificial Intelligence Research, 2012, 1(2): 149 – 170.

[17] Muller K, Mike S, et al. An introduction to kernel-based learning algorithms[J]. IEEE Transactions on Neural Networks, 2001, 12(2): 181 – 201.

[18] Chang C C, Lin C J. LIBSVM: A library for support vector machines[J]. ACM Transactions on Intelligent Systems and

Technology, 2011, 2(3): 27:1 – – 27:27.

[19] Davide A, Alessandro G, et al. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine[A]. International Workshop of Ambient Assisted Living[C]. Berlin Heidelberg: Springer Verlag, 2012. 216 – 223.

[20] Hamm J, Stone B, et al. Automatic Annotation of Daily Activity from Smartphone-Based Multisensory Streams[M]. Berlin Heidelberg: Springer Verlag, 2013. 328 – 342.

[21] Pekka S, Juha R. Recognizing human activities user-independently on smartphones based on accelerometer data[J]. International Journal of Artificial Intelligence and Interactive Multimedia, 2012, 1(5): 38 – 45.

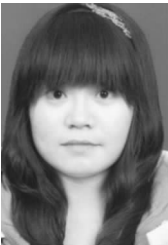
作者简介



王生生 男, 1974 年 8 月出生于吉林省长春市. 博士、教授、博士生导师. 主要研究领域为时空推理、机器视觉、计算智能、数据挖掘等.
E-mail: wss@jlu.edu.cn



柴胜(通信作者) 男, 1976 年 4 月出生于吉林省, 现为吉林大学计算机学院讲师. 研究方向为网络安全、数据挖掘等.
E-mail: chaisheng@jlu.edu.cn



杨娟娟 女, 1989 年 4 月出生于河南省信阳市. 现为硕士研究生. 主要研究领域为计算智能.
E-mail: hengyuklxts@163.com