

基于改进郊狼优化算法的浅层神经进化方法研究

刘 威^{1),2)} 付 杰^{1),2)} 周定宁³⁾ 王薪予^{1),2)} 成 秘^{1),2)}
黄 敏^{1),2)} 郭直清^{1),2)} 靳 宝^{1),2)} 牛英杰^{1),2)}

¹⁾(辽宁工程技术大学理学院 辽宁 阜新 123000)

²⁾(辽宁工程技术大学数学与系统科学研究所 辽宁 阜新 123000)

³⁾(成都数联铭品科技有限公司数据服务事业群 成都 610000)

摘 要 神经进化作为一种不同于随机梯度下降的神经网络训练方法,现已成为机器学习研究领域的一个重要分支.如何设计更好的进化策略,探索新型神经网络权值的神经进化方法是当前研究的热点问题之一.本文提出了一种基于改进郊狼优化算法的浅层神经网络进化方法.该方法首先通过引入自适应影响权重因子与选择性的混沌扰动执行机制分别从收敛速度和寻优能力两个方面对传统郊狼优化算法进行了改进.其次,以改进郊狼优化算法为神经进化策略,融入到浅层神经网络的神经进化过程,并以 BP 神经网络为例,构建了一种全新的 BP 神经网络权值、阈值优化更新方法.最后,文中采用 UCI 标准数据库中几组代表性数据验证了算法的有效性.实验结果表明:改进郊狼优化算法的进化策略充分发挥了启发式优化算法在 BP 神经网络参数空间中的全局寻优能力,能够快速逼近最优解,经过神经进化后的 BP 神经网络在分类任务中表现出了优异性能,充分验证了改进郊狼优化算法作为一种新型神经进化策略的可行性和有效性.本文研究成果丰富并拓展了神经进化领域的研究内容,为构建以神经进化为主体的新型机器学习工具箱提供了重要的参考依据.

关键词 郊狼优化算法;自适应影响权重;混沌;神经进化;BP 神经网络

中图法分类号 TP18 DOI号 10.11897/SP.J.1016.2021.01200

Research on Shallow Neural Network Evolution Method Based on Improved Coyote Optimization Algorithm

LIU Wei^{1),2)} FU Jie^{1),2)} ZHOU Ding-Ning³⁾ WANG Xin-Yu^{1),2)} CHENG Mi^{1),2)}
HUANG Min^{1),2)} GUO Zhi-Qing^{1),2)} JIN Bao^{1),2)} NIU Ying-Jie^{1),2)}

¹⁾(College of Science, Liaoning Technical University, Fuxin, Liaoning 123000)

²⁾(Institute of Mathematics and Systems Science, Liaoning Technical University, Fuxin, Liaoning 123000)

³⁾(BusinessBigData Tech. Co., Ltd, Data Services Business Group, Chengdu 610000)

Abstract As a kind of Neural network training method which is different from stochastic gradient descent, neuroevolution has become an important branch in the field of machine learning research. How to design a better evolutionary strategy and explore new type of neural network weights of the neuroevolutionary method is one of the hot issues of current research. In this paper, an evolutionary method of shallow layer neural network based on improved coyote optimization algorithm is proposed. By introducing the method of adaptive weighting factor and chaos perturbation,

收稿日期:2019-10-23;在线发布日期:2020-06-18. 本课题得到国家自然科学基金(51974144,71771111,51874160)和辽宁工程技术大学学科创新团队资助项目(LNTU20TD-01,LNTU20TD-07)资助. 刘 威,博士,副教授,中国计算学会(CCF)会员,主要研究方向为机器学习、深度神经网络、矿业系统工程. E-mail: lv8218218@126.com. 付 杰,硕士,主要研究方向为机器学习、深度神经网络. 周定宁,硕士,主要研究方向为机器学习、深度神经网络. 王薪予,硕士,主要研究方向为机器学习、深度神经网络. 成 秘,硕士,主要研究方向为机器学习、深度神经网络. 黄 敏,硕士,主要研究方向为机器学习、深度神经网络. 郭直清,硕士,主要研究方向为机器学习、深度神经网络. 靳 宝,硕士,主要研究方向为机器学习、深度神经网络. 牛英杰,硕士,主要研究方向为机器学习、深度神经网络.

the traditional coyotes optimization algorithm is improved from two aspects of convergence speed and optimization ability. Secondly, the improved coyote optimization algorithm is adopted as the neural evolution strategy, which is integrated into the neural evolution process of the shallow neural network. Taking BP neural network as an example, a new updating method of BP neural network weight and threshold optimization is constructed. Finally, several representative data sets in UCI standard database are used to verify the effectiveness of the algorithm. Experimental results show that: The evolutionary strategy of improved coyote optimization algorithm gives full play to the global optimization ability of heuristic optimization algorithm in the parameter space of BP neural network, approximates the optimal solution quickly. The BP neural network that after neural evolution shows excellent performance in classification tasks, the feasibility and effectiveness of the improved coyote optimization algorithm as a new neural evolutionary strategy are fully verified. The research results in this paper enrich and expand the research content in the field of neuroevolution and provide an important reference for building a new machine learning toolbox with neuroevolution as the main body.

Keywords coyote optimization algorithm; adaptive influence weight; chaos; neuroevolution; BP neural network

1 引言

对于一些致力于机器学习研究领域的学者而言,设计一个神经网络无异于制作一项艺术作品.神经进化(NeuroEvolution)^[1]作为一种利用进化算法自动发现神经网络最适合组件的“神经网络制作方法”,近年来得到了深入而广泛的发展,并逐渐成为业界研究的热点问题.对于浅层神经网络而言,其学习过程实际上就是权值和阈值的调整过程.浅层神经网络的神经进化研究最早源自 Miller 等人^[2]的工作.关于网络训练权重的问题,Miller 提出神经进化方法,将取决于学习与结构任务的连接权值进化过程看作网络的训练过程,通过使用自适应全局训练方法对连接权值进行全局寻优,实际上采用非线性优化问题代替了网络权重问题,避免了基于梯度下降训练算法中计算梯度信息的缺点. Uber 神经进化研究所的最新研究成果表明^[3-7],使用进化策略(Evolutionary Strategy,ES)优化神经网络,能够达到,甚至超过使用随机梯度下降(Stochastic Gradient Descent,SGD)的优化结果.因此,如何设计更好的进化策略,探索新型神经网络权值进化方法成为了当前机器学习研究领域待解决的关键科学问题之一.

神经进化中的优化问题与当前工程技术和信息科学中的复杂优化问题一样,具有非凸性、不连续性、不可微性和多模态性等特点^[8],传统的牛顿方法

和基于梯度下降的方法不适用于此类优化问题求解.为解决复杂优化问题,大量以仿生学为背景的元启发式算法被提出^[9].例如,蚁群优化算法(ACO)^[10]、人工蜂群算法(ABC)^[11]、菌落觅食算法^[12]、蝙蝠启发算法(BA)^[13]、海豚回声定位算法(DE)^[14]、萤火虫算法(FA)^[15]、花授粉算法(FPA)^[16]、灰狼算法(GW)^[17]、粒子群算法(PSO)^[18]、共生有机体搜索算法(SOS)^[19]、病毒菌落搜索算法(VCS)^[20]和鲸鱼优化算法(WOA)^[21]等.基于仿生学的元启发式算法本质上是一种概率并行的全局寻优搜索方法^[22],可分为个体启发与群体智能两类,个体启发的搜索结果取决于初始解与邻域之间的映射关系,获取到的最优解为局部极值,故其全局勘探能力差;而群智能算法面向全局寻优,每次得到的解都不同,搜索时间较长,其本身具有的良好性能和易于实现等优点,使其在科学、金融和工程等领域均得到了广泛应用.

郊狼优化算法^[23](COA)是 Pierezan 于 2018 年在 Pit WC 提出的郊狼生存模型^[24]的基础上所研究的一种用于全局优化的智能仿生优化算法,与上述其他元启发式算法相比,COA 具有独特的算法结构,该结构为优化过程中探索与开发的平衡^[25]提供了新的机制.该算法的全局优化性能强,相关调节参数少并且易于实现,重要的是 COA 在保持较高多样性的同时收敛效果达到了所有元启发式算法中最优,使其在众多元启发式算法中表现出了较为优异的性能.COA 提出的时间较短,目前正处于完善和

发展阶段^[26-27],在收敛速度和全局搜索能力等方面还存在着较大提升空间,但其独特的算法结构为神经进化提供了一种全新的进化策略设计思路。

为了更好地发挥 COA 在解决复杂优化问题方面的性能优势,将郊狼种群独特的进化策略引入到浅层神经网络的神经进化过程之中,本文提出了一种基于改进郊狼优化算法(Improving Coyote Opimization Algorithm, ICOA)的浅层神经网络进化方法. ICOA 将自适应影响权重和动态混沌扰动机制引入到郊狼种群的进化过程,通过为每个郊狼个体更新增加自适应影响权重因子,改变头狼和群体对当代郊狼的影响程度,使算法在迭代后期能够进行更精确的局部开发,加快收敛速度;通过以递减方式选择性地引入混沌扰动机制,将当前代最差群体中的所有个体作为初值带入 Tent 映射,返回的新个体组成新群体取代最差群体,使算法的全局搜索能力增强,并较好地保持了郊狼种群多样性. 数值实验结果有效证明了 ICOA 具有更好的局部开发能力和全局搜索能力,提高了收敛速度. 在此基础上,本文将 ICOA 作为神经进化策略,融入到浅层神经网络的神经进化过程,设计了一种全新的 BP 神经网络权值、阈值更新方法(Improving Coyote Opimization Algorithm Back-Propagation, ICOABP),进一步的数值实验结果表明 ICOABP 有效完成了 BP 神经网络的优化,经过神经进化后的 BP 神经网络在分类机器学习任务中表现出了优异的性能,充分验证了 ICOA 作为神经进化策略的可行性和有效性。

2 郊狼优化算法

郊狼优化算法通过模拟郊狼生活中成长、生死、繁殖和变迁等社会活动完成仿生优化计算. 算法流程图如图 1 所示,具体步骤如下^[23].

Step1. COA 算法将种群划分为 $N_p \in \mathbb{N}^*$ 个子群和 $N_c \in \mathbb{N}^*$ 个郊狼个体,第 t 时刻第 p 个狼群的第 c 个郊狼的社会状态(决策变量集)为

$$soc_c^{p,t} = \mathbf{x} = (x_1, x_2, \dots, x_D) \quad (1)$$

式中, D 为搜索空间维数. 郊狼的初始化社会状态是通过在搜索空间内对当前郊狼的第 j 维赋值来随机设定的:

$$soc_{c,j}^{p,t} = lb_j + r_j \cdot (ub_j - lb_j) \quad (2)$$

式中, ub_j 、 lb_j 分别表示第 j 维数值的上、下界, r_j 为 $[0, 1]$ 范围内的随机实数.

Step2. 检验郊狼的适应度:

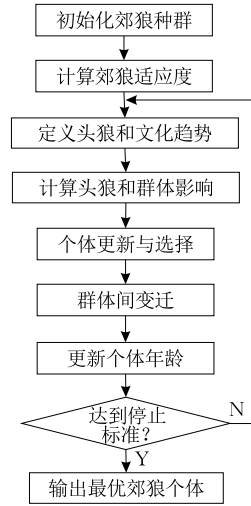


图 1 郊狼优化算法流程

$$fit_c^{p,t} = f(soc_c^{p,t}) \quad (3)$$

Step3. 群体更新过程.

(1) 定义当前郊狼群体的头狼:

$$\alpha^{p,t} = \{soc_c^{p,t} \mid \arg_{c=\{1,2,\dots,N_c\}} \min f(soc_c^{p,t})\} \quad (4)$$

(2) 郊狼群文化计算:

$$cult_j^{p,t} = \begin{cases} O_{\frac{(N_c+1)}{2},j}^{p,t}, & N_c \text{ 是奇数} \\ \frac{O_{\frac{N_c}{2},j}^{p,t} + O_{(\frac{N_c}{2}+1),j}^{p,t}}{2}, & \text{其他} \end{cases} \quad (5)$$

式中, $O_{\frac{(N_c+1)}{2},j}^{p,t}$ 为 t 时刻 p 群中所有郊狼在 N_c 为奇数时第 j 维度的社会状态的中位数.

(3) 郊狼群中郊狼个体适应度值的计算:

① 郊狼群中的郊狼个体的更新.

算法假设郊狼个体受头狼、群体的影响分别为 δ_1 、 δ_2 , 郊狼当前社会状态为

$$\delta_1 = \alpha^{p,t} - soc_{cr_1}^{p,t}, \delta_2 = cult^{p,t} - soc_{cr_2}^{p,t} \quad (6)$$

$$new_soc_c^{p,t} = soc_c^{p,t} + r_1 \cdot \delta_1 + r_2 \cdot \delta_2 \quad (7)$$

式中, r_1 和 r_2 被定义为 $[0, 1]$ 之间的随机数.

② 更新郊狼个体适应度:

$$new_fit_c^{p,t} = f(new_soc_c^{p,t}) \quad (8)$$

③ 选择郊狼适应度.

郊狼的适应能力决定了个体的新社会状态是否比旧的好, 择优选择郊狼个体:

$$soc_c^{p,t+1} = \begin{cases} new_soc_c^{p,t}, & new_fit_c^{p,t} < fit_c^{p,t} \\ soc_c^{p,t}, & \text{其他} \end{cases} \quad (9)$$

(4) 记录当前群体的出生和死亡.

COA 算法中, 离散概率 (P_s) 和关联概率 (P_a) 影响郊狼群中个体的文化多样性, 新郊狼的出生被写成:

$$P_s=1/D, P_a=(1-P_s)/2 \tag{10}$$

$$pup_j^{p,t}=\begin{cases} soc_{m_1,j}^{p,t}, & rnd_j < P_s \text{ or } j=j_1 \\ soc_{m_2,j}^{p,t}, & rnd_j \geq P_s + P_a \text{ or } j=j_2 \\ R_j, & \text{其他} \end{cases} \tag{11}$$

式中,定义 m_1 和 m_2 是第 p 个狼群的随机郊狼,定义 j_1 和 j_2 是两个随机维数, R_j 为决策变量的第 j 维的上下界内的随机数, rnd_j 为 $[0,1]$ 内的随机数.

假定 ω 表示当前幼崽的生存能力强于群体中的郊狼, φ 代表某群郊狼的数量,若 ω 成立的同时 φ 为 1 或者 φ 大于 1,则幼崽存活,群体中唯一的郊狼或者年龄最大的郊狼死亡;其余情况幼崽死亡.

Step4. 群体变迁.

起初,郊狼被随机分配到某郊狼群体中,但有时会脱离群体变为孤狼或者加入别的群体,便于实现多样化,其发生概率为 P_e :

$$P_e=0.005N_c^2 \tag{12}$$

该概率决定了被驱逐的郊狼位置.

Step5. 更新郊狼的年龄. 模拟个体随时间的推移而成长的过程,对个体年龄进行更新.

Step6. 终止条件的判定.

若满足条件,则输出适应能力最好的郊狼;否则返回 Step3 继续执行.

3 基于动态混沌扰动的自适应影响权重郊狼优化算法

3.1 郊狼优化算法的改进

鉴于 COA 在算法的收敛速度和全局搜索能力等方面还存在着较大提升空间,本文对 COA 进行了改进:一是在更新过程中对郊狼个体添加自适应影响权重,以体现头狼和群体对郊狼的影响,进而增强整个种群的局部开发能力;二是为了增大种群多样性,引入基于 Tent 映射的混沌扰动操作,更替了最差郊狼群体,使得种群的全局寻优能力增强.改进算法的具体设计思想和实现过程如下.

3.1.1 自适应影响权重机制

COA 算法在对个体进行更新时忽略了头狼与群体给个体带来的具体影响情况.该阶段是影响算法寻优结果的关键过程,不仅决定算法的收敛速度还可能导致算法陷入局部最优.因此,为了考虑到郊狼受头狼和群体影响的变化趋势,并加大郊狼跳出局部最优的概率,引入自适应影响权重因子 ω ,将更新过程中的式(7)修改为

$$new_soc_c^{p,t}=soc_c^{p,t}+\omega\cdot\delta_1+(1-\omega)\cdot\delta_2 \tag{13}$$

式中, ω 为郊狼社会状态的影响权重.

在实际进化过程中,郊狼是一种拥有独立意识的生命体,生命的前期阶段,郊狼应向着头狼的方向进化,群体影响弱于头狼影响;随着时间推移郊狼逐渐衰老,倾向于生存,大多依靠群体生存,头狼影响逐渐变小.为了算法能够符合郊狼生存过程的实际性,又为了平衡算法的全局搜索能力和局部开发能力,定义 ω 规则为:起初表示小狼受到头狼的作用偏大,受到群文化作用偏小,故值偏大,可以提高算法的全局优化性能;后期表示老狼受到头狼和群文化的作用与小狼相反,即值偏小,以此来确保算法局部开发能力;据此原则,选取 3 种权重方法^[28-32],分别为线性递减影响权重、非线性递减影响权重、线性微分递减影响权重,其定义式分别为

$$\omega_1(k)=\omega_{start}-(\omega_{start}-\omega_{end})k/T_{max} \tag{14}$$

$$\omega_2(k)=\omega_{start}-\frac{(\omega_{start}-\omega_{end})}{(2k/T_{max}-(k/T_{max})^2)} \tag{15}$$

$$\omega_3(k)=\omega_{start}-(\omega_{start}-\omega_{end})(k/T_{max})^2 \tag{16}$$

式中, ω_{start} 为初始影响权重, ω_{end} 为最大影响权重, k 与 T_{max} 分别为当前、最大迭代次数.一般来说, $\omega_{start}=0.9$ 与 $\omega_{end}=0.4$ 时算法性能最好^[33].

3.1.2 动态混沌扰动机制

郊狼种群的子群在迭代进化过程中始终保持着恒定的进化机制,该机制将会使群体相对固定于某一特定区域,并不能有效实现算法的全局搜索能力;为解决该问题,本文在群体变迁操作后以一定的条件概率选择性地引入基于 Tent 映射的混沌扰动机制.

首先引入探测机制,探测出种群中的最差群体 p .对算法设置条件限定因子 r ,若当前情况不满足限定因子 r , p 跳过混沌操作,否则 p 执行混沌操作;考虑到平衡算法全局搜索与局部搜索能力问题,设置 r 值线性减小;限定因子 r 的公式为

$$r=r_{max}-(r_{max}-r_{min})\frac{k}{T_{max}} \tag{17}$$

式中,根据文献[34]的线性递减权重原理,取 $r_{max}=0.9$ 、 $r_{min}=0.4$.算法迭代初期 r 值较大,可以保证算法效率,随着迭代次数的增加, r 值逐渐减小,执行混沌扰动的可能性增加,从而增加了种群的多样性,提高了算法的全局搜索能力.

基于探测机制引入混沌响应机制,为了在搜索过程中尽可能保证种群多样性,可采用具有随机性、遍历性和规律性的混沌变量,不同的混沌映射算子对优化影响有所不同,最常见的混沌序列为 Logistic

和 Tent 映射,根据文献[35]指出后者相比前者的遍历均匀性更好,寻优效率更高,故本文采用 Tent 映射. Tent 映射又称帐篷映射,其原式(18)与经伯努利移位变换后得到表达式(19)如下:

$$x_{n+1} = \begin{cases} 2x_n, & x_n \in [0, 0.5] \\ 2(1-x_n), & x_n \in (0.5, 1] \end{cases} \quad (18)$$
$$x_{n+1} = (2x_n) \bmod 1 \quad (19)$$

通过探测机制寻找到最差群体 p 的位置后,建立 p 群中所有个体 $soc_c^{p,t}$ 与 Tent 映射中变量 x_i 的一一对应关系,即式(20),式中, ub_j 、 lb_j 分别表示变量第 j 维的上、下界, $x_{i,j}^t$ 为经 Tent 映射转换后的对应变量. 得到新的初值代入式(19)进行迭代,产生混沌序列,经式(21)映射到原解空间,最终产生一个新的群体取代最差群体 p .

$$x_{i,j}^t = \frac{soc_{c,j}^{p,t} - lb_j}{ub_j - lb_j}, \quad j = (1, 2, \dots, D) \quad (20)$$

$$new_soc_{i,j}^{p,t} = soc_{i,j}^{p,t} + \frac{(lb_j - ub_j)}{2} (2x_{i,j}^t - 1) \quad (21)$$

3.1.3 算法描述

ICOA 算法描述如下:

```
各参数的初始化  $D, lu, N_c, N_p, nfevalMAX$ 
适应度函数  $f(soc_c^{p,t}), soc_c^{p,t} = (soc_{c,1}^{p,t}, \dots, soc_{c,D}^{p,t})^T$ 
初始化种群  $soc$ 
由  $f(soc_c^{p,t})$  计算郊狼  $soc_c^{p,t}$  的适应能力
WHILE ( $t < nfevalMAX$ )
    year = year + 1
    FOR  $p = 1 : N_p$ 
        FOR  $i = 1 : N_c$ 
            由式(4)和(5)定义头狼  $\alpha$  及群体文化趋势  $cult$ 
            按照式(6)计算头狼和群体的影响大小  $\delta_1$  和  $\delta_2$ 
            根据式(13)和式(14)~(16)对当前郊狼进行更新,对比更新前后的郊狼,保留较优郊狼
        END  $i$ 
    END  $p$ 
    记录出生与死亡
    群体变迁
    IF ( $rand() > r$ )
        找到当前最差群体  $p$ , 记录  $p$  群中所有郊狼  $soc^{p,t}$ 
        将  $soc^{p,t}$  作为初值带入混沌映射产生新郊狼组成新郊狼  $new\_soc^{p,t}$  群体取代最差群体  $p$ 
    END IF
    年龄更新
    计算种群所有个体适应度, 确定当前最优个体
END WHILE
```

3.2 优化性能实验分析

根据第 3.1.1 节中 3 种不同的递减权重方法(式(14)、(15)、(16)),结合动态混沌扰动机制,可将 ICOA 算法分为 COA1、COA2 和 COA3 三种变体,

采用当前主流算法 PSO 的改进算法 HFPSO^[36] 及与 COA 同年提出的 FSA^[37] 算法作为对比算法,为了评估 ICOA 性能,将 6 种方法在 IEEE CEC2013^[38] 的 7 个多维单、双峰函数上对算法进行了性能测试,以此验证改进算法 ICOA 的可行性和有效性. 7 个基准测试函数的函数名称、函数区间等信息见表 1 所示. 其中, $f_1 \sim f_3$ 为单峰函数, $f_4 \sim f_7$ 为多峰函数,具有多个局部最优点,并且 f_4 是标准多维多峰函数,其自变量互相影响, f_6 有多个局部极值和障碍物;因此,这些函数的求解难度高,很适合用于测试算法的求解能力和寻优能力.

表 1 基准函数信息

函数名称	函数公式	变量区间
Bent Cigar	$f_1(x) = x_1^2 + 10^6 \sum_{i=2}^D x_i^2$	$[-100, 100]$
Schwefel01	$f_2(x) = (\sum_{i=1}^D x_i)^a, a = \sqrt{\pi}$	$[-100, 100]$
Sum Squares	$f_3(x) = \sum_{i=1}^D x_i^2$	$[-5.12, 5.12]$
Griewank	$f_4(x) = \sum_{i=1}^D \frac{x_i^2}{4000} - \prod_{i=1}^D \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	$[-600, 600]$
Rastrigin	$f_5(x) = 10D + \sum_{i=1}^D [x_i^2 - 10\cos(2\pi x)]$	$[-5.12, 5.12]$
Ackley	$f_6(x) = -20 \cdot \exp\left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2}\right) - \exp\left(\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i)\right)$	$[-32, 32]$
Powell	$f_7(x) = \sum_{i=1}^{D/4} [(x_{4i-1} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 + (x_{4i-2} + 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4]$	$[-4, 5]$

各算法选取相同的种群规模,设置 10 个子群,每个子群中个体数量为 $N_c = 10$,种群大小 $popsiz = N_p \times N_c = 100$,算法的终止条件设为 $nfevalMAX = 10000 \times D$,为保证实验的公平性与客观性,各算法在 $D = 10, D = 30$ 和 $D = 70$ 皆独立进行 30 次实验,并计算 30 次实验的平均结果,记录其最优解精度和收敛次数作为最终实验结果.

(1) 表 2 实验结果分析

多维单、双峰函数 30 次实验的平均值结果如表 2 所示,根据表 2 中实验数据结果可以看出:3 种维度下,在除了函数 f_4 、 f_5 和 f_6 以外的 4 个测试函数上,ICOA 的 3 种算法 COA1、COA2、COA3 的寻优性能对比 COA 皆表现出了较好的优越性,3 种算法的平均精度均高于 COA,而且 COA3 普遍得到了全局最优解,在函数 f_4 、 f_5 上,COA1 与 COA2 得到

表 2 改进前后算法在前 10 个函数的优化性能比较							
基准函数	算法	D=10		D=30		D=70	
		最优解	迭代次数	最优解	迭代次数	最优解	迭代次数
f_1	COA	1.31E-06	909	2.86E-07	2727	2.92E-07	6363
	COA1	5.20E-19	909	7.87E-23	2727	1.10E-23	6362
	COA2	1.46E-17	909	8.71E-22	2727	8.06E-23	6363
	COA3	0	90	0	207	0	267
	HFPSO	4.88E-08	909	2.0E+01	2727	1.01E+01	6363
	FSA	1.02E-03	223	0	286	0	383
f_2	COA	3.70E-20	909	3.61E-19	2727	5.30E-19	6363
	COA1	3.01E-42	909	1.63E-49	2727	3.80E-51	6363
	COA2	4.51E-40	909	4.53E-48	2727	3.43E-49	6363
	COA3	0	85	0	204	0	246
	HFPSO	8.23E-24	909	5.32E-13	2727	0.53E-03	6363
	FSA	0	120	0	280	0	320
f_3	COA	8.71E-14	909	1.54E-14	2727	6.89E-14	6363
	COA1	3.80E-26	909	6.17E-30	2727	7.17E-31	6363
	COA2	3.75E-25	909	4.57E-29	2727	1.07E-29	6363
	COA3	0	116	0	155	0	409
	HFPSO	1.98E-15	909	2.28E-08	2727	3.05E-05	6363
	FSA	0	224	0	321	0	412
f_4	COA	1.88E-03	909	4.78E-12	2727	1.23E-03	6363
	COA1	1.45E-02	909	2.46E-03	2255	1.73E-03	4137
	COA2	1.98E-02	909	0	1798	2.71E-03	3919
	COA3	0	121	0	134	0	460
	HFPSO	8.78E-02	909	1.37E-02	2727	7.98E-03	6363
	FSA	0.43	908	1.02	2684	6.70	6281
f_5	COA	8.19E-08	908	9.95E-02	2726	5.13E-01	6363
	COA1	1.99E-01	727	9.28E-00	2727	6.43E+01	4819
	COA2	5.50E-01	907	1.22E+01	2727	7.98E+01	5046
	COA3	0	141	0	183	0	231
	HFPSO	4.17E+00	909	3.93E+01	2727	1.38E+02	6363
	FSA	1.13E+01	872	6.37E+01	2678	1.01E+02	6274
f_6	COA	1.98E-06	909	1.91E-06	2727	2.55E-05	6363
	COA1	1.20E-12	909	11.07	2670	17.90	3440
	COA2	3.62E-12	909	11.94	2707	19.72	2022
	COA3	8.88E-16	142	8.88E-16	230	8.88E-16	266
	HFPSO	8.18E-08	909	5.81E-05	2727	1.64E-03	6363
	FSA	0.99	901	1.83	2622	3.25	5923
f_7	COA	2.04E-09	907	3.75E-04	2725	1.57E-02	6363
	COA1	3.82E-12	909	3.41E-05	2723	3.65E-03	6363
	COA2	8.94E-12	908	6.89E-05	2727	4.44E-03	6363
	COA3	0	133	0	152	0	218
	HFPSO	1.45E-05	909	3.42E-03	2727	1.34E-01	6363
	FSA	7.72E-01	908	3.46	2391	3.47E+01	6344

的最优解值虽然不优于 COA,但是均优于 HFPSO 与 FSA;在单峰函数上 FSA 同 COA3 均达到了测试函数的理论最优解,但是 FSA 达到最优解的速度弱于 COA3,且由于多峰函数具有多个局部最优的点,导致 FSA 在多峰测试函数上的寻优性能较低,故 COA3 在求解精度以及速度上性能都优于 FSA.此外,设定一定迭代次数的情况下,COA 的寻优精度随着维度的增加大幅度降低,例如在函数 f_6 下,随着维度由 10 上升到 70,COA 的寻优精度降低了近 13 倍;然而,COA3 的寻优精度随着维度增大并未呈现下降趋势,依旧保持最优值,且达到理论精度

时的迭代次数虽然有增大,但增大并不明显,故其收敛速度未出现明显的下降;7 个测试函数下,COA3 在所有测试函数上经迭代都达到了全局最优,且除了函数 f_6 的精度值外,其余最优解均为函数的理论极值;同时,观察表中数据可知 COA3 收敛速度对比 COA1、COA2 更快;不同维度下,COA 在达到最大迭代次数时均未达到理论最优解,而 COA3 明显在未达到最大迭代条件的情况下得到了函数的理论最优解,这说明 COA3 具有更好的局部搜索能力以及最大的收敛速率;据此分析,COA3 的寻优性能不会因为个体维度的变化而降低;ICOA 的 3 种改进算法 COA1、COA2、COA3 中,COA3 具以最快的速度收敛到最优精度的概率更高,故 COA3 具有更好的寻优性能.

对比权重因子式(14)、(15)和(16),如图 2 所示,定义初始、最大权重值为 $\omega_{\text{start}}=0.4$, $\omega_{\text{end}}=0.9$,横坐标代表迭代次数 k ,假设 $T_{\text{max}}=300$,纵坐标表示 $\omega_i(k)$, $i=1,2,3$,即 3 种衰减权重方法的变化曲线.可以看出式(16)即 ω_3 的下降趋势最缓慢,故使得 COA3 避免了因为个体更新变化过快而造成的陷入局部极值问题,因此,COA3 的算法收敛性能更好.

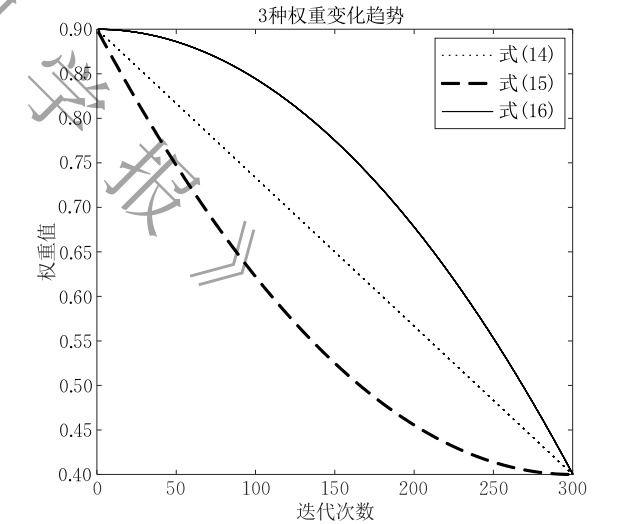


图 2 权重变化情况

(2)收敛性曲线分析

收敛曲线可以直观地显示出算法陷入局部最优中的次数和收敛速度; $f_1 \sim f_7$ 的 7 个函数中, $f_4 \sim f_7$ 为多维多峰函数,而相比单峰函数,多峰函数具有多个局部极值,比较复杂,更能说明算法本身的寻优能力,故采用这些多维多峰函数的收敛曲线图对算法收敛性能进行分析.如下所示,图 3~图 6 是 4 个多峰测试函数在维度 $D=10、30$ 及 70 时的收敛情况,横坐标为迭代次数,纵坐标为得到的最优解.

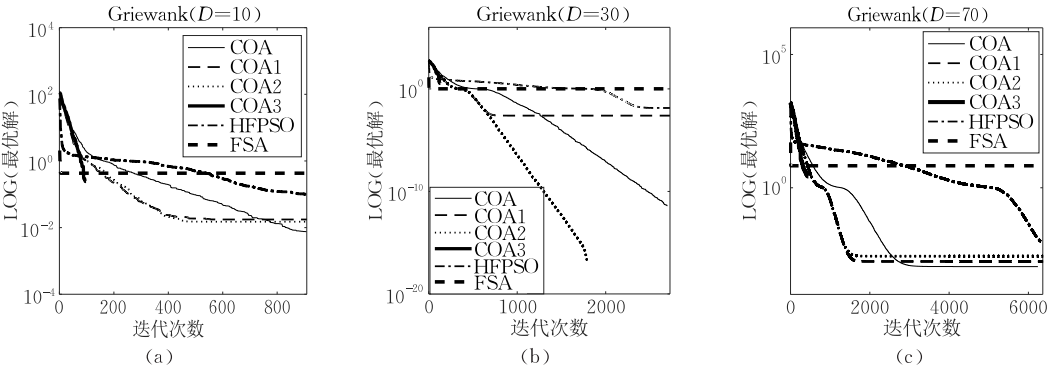


图 3 Griewank 收敛曲线

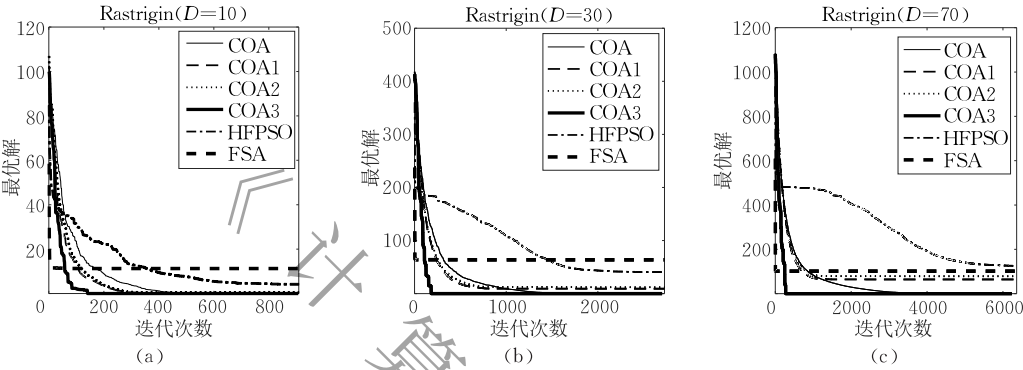


图 4 Rastrigin 收敛曲线

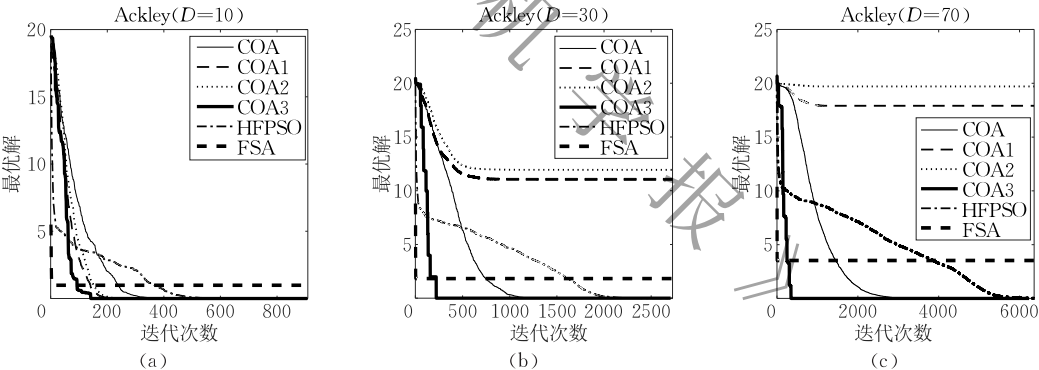


图 5 Ackley 收敛曲线

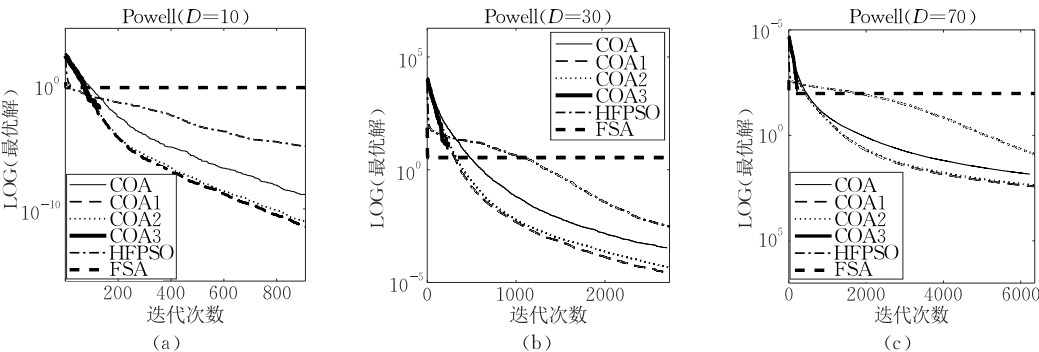


图 6 Powell 收敛曲线

由收敛曲线可以看出,当维度从 10 增加到 70,所有测试函数上的 COA3 算法的寻优精度都达到了理论最优,从图中测试函数得到的最优解的对数值曲

线可以看出,Powell 函数的 COA1、COA2、COA3 最优解结果均优于 COA,不仅最优解精度高于 COA 与另外两种对比算法,且收敛速度明显高于 COA 等

算法.此外,在 $D=30$ 及 $D=70$ 维度下,在图3、图4以及图5中,虽然COA1、COA2在算法满足终止条件时得到的最优解结果大于COA,且图3(c)中COA1陷入了局部最优, R 但是COA3的寻优精度达到了最优,且其寻优速度达到了最高.这说明COA、COA1、COA2的寻优性能都有可能受维度变化而导致影响较大,而COA3可以很好地避免这种影响.故COA3具有良好的收敛性能;因此,可以证明文中所提出的COA3算法的收敛性能具有明显的优势.

总的实验结果表明,本文所提出的改进郊狼优化算法COA3不论是在低维条件下还是在高维条件下皆具有较好的收敛性能和寻优性能,并且COA3的求解精度、收敛速度以及稳定性均是最优的.因此,可以体现出改进的郊狼优化算法ICOA的3种方法中,COA3找到理论最优值的概率明显高于另外两种算法,并且COA3的寻优精度及其稳定性相对来说更好一些.在维度 $D=10,30,70$ 的条件下,COA3比COA算法具有更高的寻优性能,这是

因为COA在更新过程中采用自适应影响权重体现了头狼与群体对郊狼影响大小,有利于局部极值的跳出,提高了收敛速度,且算法后期引入混沌Tent映射扰动,提高了算法的全局搜索能力,很大程度上做到了全局搜索与局部开发的平衡,使算法更容易找到理论最优值,证明了改进后的郊狼优化算法具有可行性和有效性,寻优性能较高.

(3) 复杂优化问题能力分析

为进一步体现改进的COA算法处理复杂优化问题的能力,取 $N_c=N_p=5$, $popsize=N_p\times N_c=25$,算法的终止条件设置为 $nfevalMAX=10\,000\times D$,将HFPSO、FSA作为对比算法,采用IEEE CEC 2017上的10个混合函数、10个复合函数^[39]共20个测试函数(具体函数信息见文献[39]),在设置维度为 $D=30$ 的情况下,将COA与COA3(以下统称ICOA)HFPSO、FSA等4种算法在20个测试函数上进行了函数收敛性测试,得到结果如表3.根据表中数值结果得到的表3的具体分析如下:

表3 D=30 算法在 IEEE CEC2017 基准函数下的优化性能比较

基准 算法	算法	最优解	迭代 次数	基准 函数	算法	最优解	迭代 次数	基准 函数	算法	最优解	迭代 次数
f_8	COA	1150.13	9993	f_{15}	COA	42439.63	9728	f_{22}	COA	2888.63	6948
	ICOA	1137.47	9998		ICOA	41771.10	9989		ICOA	2888.74	7235
	HFPSO	1183.83	9991		HFPSO	453365.00	10000		HFPSO	2888.08	9935
	FSA	21663.02	9063		FSA	6.97E+7	6972		FSA	5278.66	5473
f_9	COA	36954.27	9999	f_{16}	COA	2286.64	10000	f_{23}	COA	4593.10	9991
	ICOA	16817.22	9999		ICOA	2041.10	9998		ICOA	4566.01	9764
	HFPSO	201228.10	9999		HFPSO	6278.44	9998		HFPSO	2900.01	9999
	FSA	6.54E+10	9865		FSA	2.16E+10	9972		FSA	12631.23	8024
f_{10}	COA	47040.78	10000	f_{17}	COA	2514.64	8482	f_{24}	COA	3232.60	9988
	ICOA	27623.50	10000		ICOA	2486.23	8820		ICOA	3217.71	9970
	HFPSO	22489.43	9987		HFPSO	2324.68	9986		HFPSO	3267.74	9997
	FSA	1.72E+10	9985		FSA	3402.36	9814		FSA	4941.08	9299
f_{11}	COA	1525.70	10000	f_{18}	COA	2398.23	5853	f_{25}	COA	3218.18	9999
	ICOA	1488.48	9890		ICOA	2379.36	9985		ICOA	3214.91	9999
	HFPSO	12038.52	9994		HFPSO	2384.82	9629		HFPSO	3215.99	10000
	FSA	733254.50	7313		FSA	2824.89	6836		FSA	7569.19	9128
f_{12}	COA	2210.04	9985	f_{19}	COA	6187.00	9986	f_{26}	COA	3638.18	9992
	ICOA	1686.94	9989		ICOA	5096.73	9980		ICOA	3594.51	10000
	HFPSO	19328.32	9987		HFPSO	2300.01	9999		HFPSO	3611.42	9998
	FSA	1.59E+10	5637		FSA	11444.61	9162		FSA	13424.80	7371
f_{13}	COA	2648.72	10000	f_{20}	COA	2782.67	9936	f_{27}	COA	9177.89	9996
	ICOA	2561.25	9996		ICOA	2734.16	9960		ICOA	5605.99	9986
	HFPSO	2737.98	9993		HFPSO	2813.61	9985		HFPSO	2888.06	6948
	FSA	4783.88	6354		FSA	3876.10	8971		FSA	2888.74	7235
f_{14}	COA	2322.93	9999	f_{21}	COA	2951.18	5626				
	ICOA	1983.84	9279		ICOA	2914.87	7019				
	HFPSO	1943.06	9999		HFPSO	2985.48	9504				
	FSA	8971.27	6743		FSA	3872.03	9669				

在 $f_8\sim f_{27}$ 的混合函数和复合函数上,除了在 f_{22} 上ICOA的收敛性能稍差于COA外,其余函数上ICOA算法的收敛性能都优于COA的收敛性

能,这证明了在处理复杂优化问题的能力上,ICOA整体优于COA.此外,从表中可以看到,在测试函数 f_{10} 、 f_{14} 、 f_{17} 、 f_{19} 、 f_{22} 和 f_{23} 上,虽然COA与ICOA二

者的表现均不如 HFPSO,且在函数 f_{27} 上弱于 HFPSO 与 FSA,但是在二者其它的测试函数上表现都是偏好的,同时,观察表中函数达到最优解时的算法迭代次数,可以看出,FSA 算法在函数 $f_{11} \sim f_{15}$ 、 f_{18} 、 f_{22} 、 f_{23} 、 f_{26} 和 f_{27} 10 个测试函数上,均早早达到了收敛,且得到的最优解非理论最优解,说明 FSA 易陷入函数的局部极值,故整体分析来说,ICOA 与 COA、HFPSO 及 FSA3 种方法相比,在所有算法同达到最大迭代次数时,ICOA 的收敛性能表现更优,证明了在处理复杂优化问题的能力上,ICOA 的整体表现优于 COA、HFPSO 和 FSA.

结合表 2 和收敛曲线的分析结果,从总体上来说,不管是单峰测试函数和多峰测试函数,还是混合测试函数与复合测试函数,ICOA 都具有更好的收敛性能与更强的寻优能力.

4 基于 ICOA 的浅层神经进化方法

神经进化针对神经网络的早期工作主要在于优化初始权值以及加快收敛速度两方面,该方法只能在有限程度上提高网络的训练精度,而网络训练的泛化性能却无法得以提升.对于浅层 BP 神经网络来说,传统的 SGD 训练方法存在容易使其陷入局部极小点、收敛速度慢等缺点,而神经进化在对 BP 神经网络的权值与阈值进行训练时会自动搜索更优的网络,是一种不同于 SGD 的神经网络训练方法,其进化策略和全局寻优的设计思想在解决传统 SGD 训练方法存在的问题时具有一定优势.在神经网络权值空间上,应尽可能地将神经进化全局搜索能力强这一优点体现出来,以便获取最优网络权值进而得到最优神经网络.

鉴于 ICOA 结合了自适应影响权重与动态混沌扰动机制,这一操作增大了算法局部最优解的跳出概率,并使得算法的全局勘探与局部搜寻能力得以均衡体现.本文将 ICOA 在解决复杂优化问题方面的性能优势和郊狼种群独特的进化策略引入到浅层神经网络的神经进化过程之中,提出了一种基于 ICOA 方法的浅层神经网络进化方法——ICOABP,具体设计思想和实现过程如下.

4.1 设计原理与思想

传统神经网络在样本空间上进行网络训练,首先对初始权值与阈值进行随机选择,再经过训练对

样本进行更新,最终确定网络最优权重.本文提出的浅层神经进化中,在浅层神经网络上采取了“权值”替代“样本”的空间思想,在新的空间思想上执行网络训练并通过进化算法完成神经进化操作.即将权值空间上的网络训练问题定义为最优化问题再通过使用进化算法求取优化目标的最优解,进而达到了神经网络与进化算法二者融合的目的.

现将问题的优化目标定义为 BP 神经网络模型的权值 ω 与阈值 b ,即 ω 与 b 的共同组合成实数空间中的空间变量 x ,该实数空间即为网络训练的权值空间.采用分类准确率最大原则为优化方向,定义神经网络分类问题数学模型为

$$\begin{aligned} &\max F(\omega, b) \\ \text{s. t. } &\begin{cases} \omega \in [\omega_{\min}, \omega_{\max}]; \\ b \in [b_{\min}, b_{\max}] \end{cases} \end{aligned} \tag{22}$$

式中, ω 与 b 分别为 BP 神经网络的权值与阈值.

利用 BP 神经网络算法训练权值与阈值的过程可以形容为上述的优化模型,传统 BP 算法与神经进化方法在空间上搜索最优解的过程如图 7,左图为 BP 算法通过梯度下降方式单点搜索的寻优路线,这种搜索方式求取到的最优解陷入局部极值的概率很大;右图为神经进化方法的搜索过程,通过采用进化算法,实现了在空间上的多出发点搜索,对比单点寻优,该方法在更大程度上可以找到全局最优解,故进化算法的网络搜索能力以及权值与阈值计算能力更强.

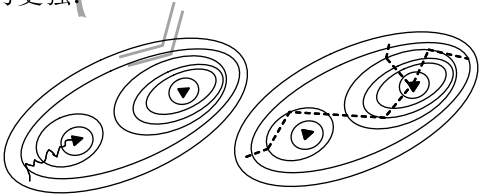


图 7 BP 算法与神经进化的寻优过程图

4.2 ICOABP 实现过程

传统仿生优化算法的初始种群是随机设定的,适应度函数即为优化问题的目标函数.ICOABP 则是通过获取到 BP 算法神经网络训练产生的分类器,并将每个分类器作为进化算法的郊狼个体,进而得到所需的初始种群,ICOABP 中的适应函数即为分类问题的最大准确率函数.ICOA 算法通过对权值与阈值进行不断的更新与优化,最后经神经进化获取最优网络分类器.ICOABP 的主要操作流程见图 8.

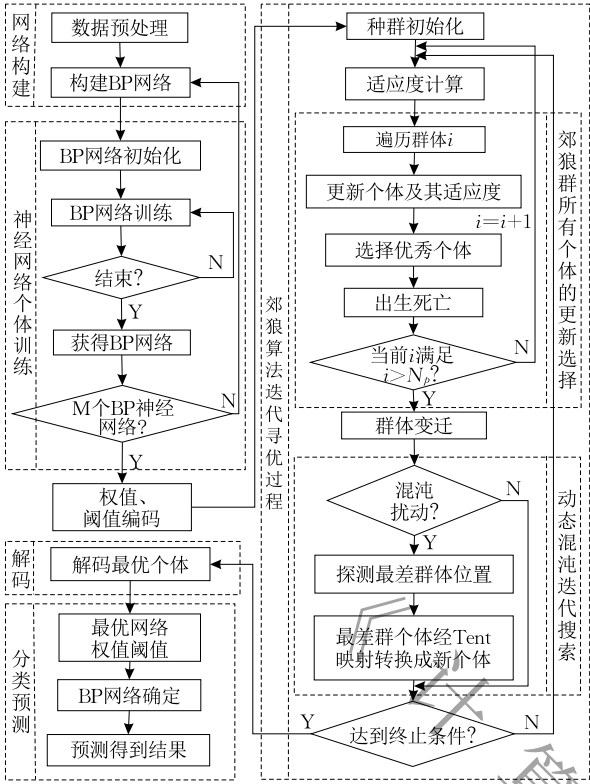


图8 基于ICOA的浅层神经进化方法流程

假定神经网络的输入层、隐藏层以及输出层的节点数分别为 inp 、 hid 、 out ，算法的个体维度为 $D=inp\times hid+hid\times hid\times out+out$ ，其余种群、适应函数等相关参数同郊狼优化算法的表示相同，在ICOA算法的性能分析中，得到COA3的寻优性能较高，故采用式(16)作为更新过程的自适应影响权重；具体的实现步骤如下：

Step1. 数据的处理与参数的设定. 考虑到分类问题的分类准确率，将数据集按照比例分为训练与测试两部分；某郊狼个体 $soc_c^{p,t}$ 可以表示成 D 维行向量： $soc_c^{p,t}=[soc_{c,(1)}^{p,t}, soc_{c,(2)}^{p,t}, \dots, soc_{c,(D)}^{p,t}]\in R^D$ ；模型中需要设置的参数有迭代终止条件 $nfevalMAX$ 、个体与群体数目 N_c 和 N_p 、个体维度 D 以及搜索空间 lu 。

Step2. 种群初始化. 采用实数值方式编码初始的权值与阈值并作为初始个体组成种群. 假定当前时刻为 t ，则定义此时的第 p 群体 $soc^{p,t}$ 与种群 soc^t 矩阵如下：

$$soc^{p,t}=[soc_1^{p,t}, soc_2^{p,t}, \dots, soc_{N_c}^{p,t}]^T, \tag{23}$$

$$soc^t=[soc^{1,t}, soc^{2,t}, \dots, soc^{N_p,t}]^T$$

式中， $soc^{p,t}$ 矩阵表示 p 群中的 N_c 个个体，即BP神经网络的待预测参数， soc^t 矩阵表示种群中的 N_p 个群。

Step3. 个体适应度的计算. 假定总样本长度为 l ， rv_i 为测试样本 i 的实际结果， cv_i 为样本 i 经分类预测得到的计算结果，定义适应度值 f_i 等价于分类准确率大小，其计算公式如下：

$$Y_i=\begin{cases} 1, & cv_i=tv_i \\ 0, & cv_i\neq tv_i \end{cases}, f_i=\sum_{i=1}^l Y_i/l \tag{24}$$

Step4. 更新群体.

- (1) 结合1中Step3计算群体中分类准确率最大个体 $\alpha^{p,t}$ 与群体文化趋势 $cult^{p,t}$.
- (2) 结合改进后的自适应影响权重方法，根据式(12)和(15)更新个体，选择适应度最优的作为当前个体.
- (3) 记录当前群体的出生和死亡.

Step5. 若当前随机数大于驱逐概率 P_e ，则发生变迁行为；否则不发生.

Step6. 动态混沌探查机制. Step4的群体变迁结束后，通过检测当前迭代是否满足更新系数 r ，若满足，则找出当前分类准确率最差群体 p ，通过混沌Tent映射对 p 群个体进行迭代搜索以获得较优的个体取代原 p 群体对应个体继续算法的迭代寻优.

Step7. 以年作为单位，每更新完郊狼群，郊狼个体年龄随之增长1个单位.

Step8. 如果终止条件不满足，跳转至Step3. 否则程序终止，给出最优解结果.

4.3 数值实验

为了更好地分析基于ICOA的浅层神经进化方法的泛化性能，采用数值实验方式对其进行性能评估. 实验挑选了UCI库中的10个分类数据集，并对每个数据集进行分层取样，将各数据集按照相同的比例分为训练、验证与测试样本，以便在测试的时候可以尽可能地获取到网络的真实性能，相关数据名称及属性见表4.

表4 UCI中10组数据集物理属性

数据名称	特征大小	分类大小	训练集长度	验证集长度	测试集长度	数据长度
DRD	19	2	384	384	383	1151
CMC	9	3	491	491	491	1473
Seeds	7	3	70	70	70	210
WFN	40	3	1666	1666	1668	5000
Iris	4	3	50	50	50	150
WF	21	3	1666	1666	1668	5000
Music	24	4	666	666	668	2000
Wine	13	3	60	60	58	178
IP	34	2	60	60	60	180
Glass	9	6	71	71	72	214

实验时存在量纲等差异的数据集的不同属性指标可能会影响分类器性能,因此,需要对数据集进行归一化处理^[40],计算公式为

$$X_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \quad (25)$$

式中, $x_j^{\max}$ 和 $x_j^{\min}$ 分别表示第 j 列最大值和最小值, X_{ij} 为归一化后的样本。

实验对比算法设置:标准的 BP 神经网络是利用单一的网络进行优化,本文提出的浅层神经网络在进化过程中产生了多个神经网络进行学习优化,故为了全面公平地分析 ICOA 算法优化 BP 神经网络参数的性能,本文将基于 ICOA 的浅层神经进化方法分别与基于 COA 的浅层神经进化方法、基于单一神经网络分类器的标准神经网络方法 SNN、分别采用 Bagging 与 Boosting 集成方法作为策略将决策树设定为基分类器的集成方法 BDTE 与 ADTE 以及采用 Boosting 集成方法作为策略的神经网络集成方法 ANNE 5 种算法进行了对比分析^[41]。

训练过程中,针对同一数据集,将 6 种算法的模型参数设置为相同的值,其中,BP 迭代次数设置为 3000,ICOA 训练终止条件设置为 3000,种群规模设置为 25,学习率设置为 0.2,不同数据集的隐层数目设置如表 5 所示。

表 5 不同数据集参数设置情况

数据集	DRD	Seeds	CMC	Iris	WFN	WF	Music	Wine	IP	Glass
隐层数	14	10	6	12	6	10	10	12	12	12

利用 6 种算法优化 BP 神经网络参数并分别对 10 组实验数据进行 30 次实验,并以 30 次实验的测试集平均值、最大值、标准差以及最小值的平均值作为实验最终评价指标,实验结果如表 6 所示。

由表 6 可知,30 组数据实验的平均结果中,首先分析 COABP 相对 SNN、BDTE、ADTE 和 ANNE 4 种算法的性能优势,其平均值指标在 6 组数据集上较高,3 组数据集上仅次于 SNN,因此 COABP 的平均分类准确率最大;最大值指标在 7 组数据集上达到了最优,故其最大分类准确率最高;标准差结果在 DRD、IP 数据集达到了最高,在 Seeds、WFN、Wine 数据集上仅次于 SNN,在 CMC 数据集上低于 SNN 与 ANNE,说明 COABP 的稳定性能低于 SNN 位于第二;同时,最小值指标在 6 组数据集上得到了最优结果,在 CMC、Iris 上仅次于 SNN,即 COABP 的最坏分类效能也是 5 种算法

表 6 6 种算法实验结果对比

数据集	评价标准	6 种算法结果					
		SNN	BDTE	ADTE	ANNE	COABP	ICOABP
DRD	Avg	72.15	63.24	62.89	71.42	72.10	73.42
	Max	75.19	68.40	69.20	74.41	75.72	76.50
	Std	1.87	2.62	3.04	1.82	1.53	1.68
	Min	66.06	58.49	57.70	66.58	69.71	69.71
CMC	Avg	53.65	51.65	46.19	51.24	53.47	54.17
	Max	57.06	56.03	51.33	55.21	58.08	59.92
	Std	1.45	1.99	2.45	1.79	1.96	1.83
	Min	51.12	47.24	42.74	47.65	50.72	50.51
Seeds	Avg	92.68	88.60	88.28	92.73	93.58	93.69
	Max	96.97	95.45	96.97	98.48	98.48	98.48
	Std	2.36	4.30	4.47	3.36	2.40	2.32
	Min	86.36	77.27	77.27	84.85	89.39	89.39
WFN	Avg	84.45	80.79	72.43	84.47	84.59	84.90
	Max	85.89	82.34	76.21	86.97	86.07	86.85
	Std	0.85	0.89	1.76	0.88	0.80	0.78
	Min	82.77	78.87	68.31	82.89	83.25	83.43
Iris	Avg	95.00	94.24	94.65	93.75	95.35	95.42
	Max	100.00	100.00	100.00	100.00	100.00	100.00
	Std	2.70	3.53	3.03	7.28	3.60	2.70
	Min	89.58	85.42	87.50	58.33	87.50	89.58
WF	Avg	84.37	81.38	72.61	84.09	84.71	84.75
	Max	85.34	83.17	75.84	85.58	86.42	86.66
	Std	0.70	0.71	1.64	0.80	0.92	0.75
	Min	82.75	79.69	68.39	81.97	82.93	83.47
Music	Avg	89.05	88.77	50.62	88.82	89.33	89.55
	Max	92.17	90.36	51.66	90.81	92.62	92.47
	Std	1.16	1.01	0.49	1.19	1.37	1.16
	Min	86.90	86.90	49.85	86.75	86.60	86.75
Wine	Avg	96.03	94.02	88.62	96.03	96.89	96.95
	Max	100.00	100.00	94.83	100.00	100.00	100.00
	Std	2.44	2.99	4.14	2.76	2.65	2.17
	Min	89.66	87.93	79.31	89.66	89.66	91.38
IP	Avg	67.41	68.22	66.32	66.72	67.53	68.48
	Max	81.03	75.86	82.76	77.59	79.31	77.59
	Std	6.20	6.04	6.92	5.80	4.71	4.82
	Min	51.72	51.72	46.55	48.28	56.90	58.62
Glass	Avg	63.48	67.78	51.21	65.30	61.87	66.09
	Max	74.24	80.30	65.15	77.27	72.73	77.27
	Std	4.99	5.97	8.17	5.95	8.63	5.21
	Min	54.55	57.58	36.36	54.55	36.36	56.06

中最优的;根据结果总体进行分析,COABP 的综合表现略优于 SNN、BDTE、ADTE 和 ANNE 4 种算法.其次对 ICOABP 与包含 COABP 在内的 5 种算法进行对比分析:ICOABP 在除 Glass 数据集外的 9 组数据集上的平均分类准确率都位列第一,远远优于另外 5 种算法,其最大分类准确率在 6 组数据集上达到了所有算法的最高值,仅在 WFN、Music、IP、Glass 4 组数据集分别略低于 ANNE、COABP、ADTE、COABP 与 BDTE 等算法,其最小分类准确率在 7 组数据集上达到了最优,在 Seeds 数据集弱于 COABP,在 Glass 数据集上低于 BDTE,在 CMC 和 Music 数据集上排第三;针对 ICOABP 的标准差

这一指标,为了分析清楚其与另外 5 种算法的性能优势对比,分别计算 ICOABP 与 SNN、ADTE、BDTE、ANNE 和 COA 在所有数据集上的 5 个最优值占比:5:3:8:2:9:1:8:2:7:3;以此确定 ICOABP 算法的稳定性能高于 SNN、ADTE、BDTE、ANNE 和 COA;结合 4 个指标总体进行分析,ICOABP 的综合表现优于 SNN、BDTE、ADTE、ANNE 和 COA 5 种算法.通过表 6 结果,我们可以发现在 Glass 数据集上 COABP、ICOABP 的表现情况都不是最好的,通过对数据集的性质分析,可能是由于 Glass 分类类别数目大,导致相对于小分类数据集分类操作更复杂,同时,该数据集数据数目较少,数目少会存在一定偶然性,在进行分类时,高分类的数据集若数目少就可能分类不准确;但是根据表中结果可发现在该数据集上,ICOA 各项指标都排名第二.综上所述,得到以下结果:

(1) 对比测试数据集上分类准确率的均值和最大值,总体上 COABP、ICOABP 比 SNN、BDTE、ADTE 和 ANNE 方法具有更高的平均分类准确率(Avg)和最大分类准确率(Max),且 ICOABP 性能优于 COABP;在实际工程应用中,该结果更可能给工程带来最大的经济效益;说明基于 COA、ICOA 的浅层神经进化方法是有效且分类泛化能力好的神经进化方法.

(2) 对比 SNN、多神经网络个体结合的学习方法以及 COABP,ICOBP 在测试数据集上具有更高的稳定性(Std)与最坏分类效能(Min),即 ICOABP 的算法稳定性更强,鲁棒性高,更适用于实际工程领域,尽可能保证算法的最小错判率,降低工程生产的经济损失.

综合分析可知,基于 ICOA 的浅层神经进化在能力表现上具有一定的优越性,原因在于 ICOA 在 COA 的基础上添加了自适应影响权重因子并引入了混沌 Tent 映射,使算法的收敛速度以及局部搜索能力得到很大的提高,同时还保证了较好的全局搜索能力和种群多样性,提高了算法的泛化性能.

5 结 论

本文以神经进化策略为主旨,从提升智能仿生优化算法性能、寻找新型神经进化方法两个方面展开研究,通过理论分析和数值实验形成结论如下:

(1) 为改善传统郊狼优化算法的性能,本文提出了一种基于混沌扰动的自适应影响权重郊狼优化算法(ICOA),该方法通过引入自适应影响权重和动

态混沌机制,改进了郊狼个体的更新过程,并利用混沌映射的遍历性和均匀性实现了群体位置的更新,数值实验结果表明,ICOA 有效提高了郊狼优化算法的收敛速度与全局寻优性能.

(2) 为更好地发挥 ICOA 求解复杂优化问题优势,本文将 ICOA 中的郊狼种群进化策略引入到浅层神经网络的神经进化过程之中,提出了一种基于 ICOA 的 BP 神经网络进化方法(ICOABP),数值实验对比结果表明:ICOABP 有效实现了 BP 神经网络权值和阈值的优化,与 SNN、BDTE、ADTE、ANNE、COABP 等算法相比具有更高的分类性能.

本文所做工作是对新型神经进化方法探索的一次有益尝试,数值实验结果充分证明了 ICOA 和 ICOABP 的有效性和性能优势.但是基于改进郊狼算法的浅层神经进化在提升了分类准确性能的同时,也使得算法的运行时间增加,因此,本文方法主要适合对算法的运行时间效率要求不高的预测分类问题,包括科研或者实际工程领域的二分类、多分类预测等.本文从权值空间的角度提出浅层网络训练方法,后续可基于结构空间的角度,将郊狼优化算法用于深度学习网络,如 CNN 等的训练,围绕基于 ICOA 的深度神经网络超参数的神经进化方法,以及 ICOABP 机器学习工具箱的研发等方向继续深入开展,该研究在深层网络的研究中具有重要意义.

参 考 文 献

[1] Stanley K O, Clune J, Lehman J, et al. Designing neural networks through neuroevolution. *Nature Machine Intelligence*, 2019, 1(1): 24-35

[2] Miller G F, Todd P M, Hegde S U. Designing neural networks using genetic algorithms//*Proceedings of the 3rd International Conference on Genetic Algorithms*. Fairfax, USA, 1989: 379-384

[3] Such F P, Madhavan V, Conti E, et al. Deep neuroevolution: genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning. *arXiv preprint arXiv:1712.0 6567*, 2017

[4] Lehman J, Chen J, Clune J, et al. Safe mutations for deep and recurrent neural networks through output gradients. *arXiv preprint arXiv:171 2.06563*, 2017

[5] Zhang X, Clune J, Stanley K O, et al. On the relationship between the open-AI evolution strategy and stochastic gradient descent. *arXiv preprint arXiv:1712.06564*, 2017

[6] Lehman J, Chen J, Clune J, et al. ES is more than just a traditional finite-difference approximator. *arXiv preprint arXiv:1712.06568*, 2017

[7] Conti E, Madhavan V, Such F P, et al. Improving exploration in evolution strategies for deep reinforcement learning via a

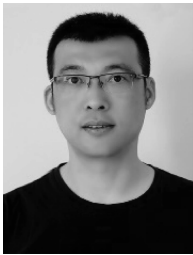
- population of novelty-seeking agents. arXiv preprint arXiv: 1712.06560, 2017
- [8] Wei Y H, Xu C, Hu Q Y. Transformation of optimization problems in revenue management, queueing system, and supply chain management. *International Journal of Production Economics*, 2015, 146(2): 588-597
- [9] Yang X-S. *Nature-Inspired Optimization Algorithms*. Amsterdam, Netherlands: Elsevier, 2014
- [10] Socha K, Dorigo M. Ant colony optimization for continuous domains. *European Journal of Operational Research*, 2008, 185(3): 1155-1173
- [11] Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of Global Optimization*, 2007, 39(3): 459-471
- [12] Chen H, Niu B, Ma L, et al. Bacterial colony foraging optimization. *Neurocomputing*, 2014, 137(5): 268-284
- [13] Yang X-S. A new metaheuristic bat-inspired algorithm. *Nature Inspired Cooperative Strategies for Optimization*. Berlin, Germany: Springer, 2010, 284(1): 65-74
- [14] Kaveh A, Farhoudi N. A new optimization method: Dolphin echolocation. *Advances in Engineering Software*, 2013, 59(5): 53-70
- [15] Yang X-S. Firefly algorithms for multimodal optimization. *Lecture Notes in Computer Science*, 2009, 5792(1): 169-178
- [16] Yang X-S. Flower pollination algorithm for global optimization. *International Conference on Unconventional Computation*, 2012, 7445(1): 240-249
- [17] Mirjalili S, Mirjalili S M, Lewis A, et al. Grey wolf optimizer. *Advances in Engineering Software*, 2014, 69(3): 46-61
- [18] Kennedy J, Eberhart R. Particle swarm optimization// *Proceedings of the IEEE International Conference on Neural Networks*. Perth, Australia, 1995: 1942-1948
- [19] Cheng M Y, Prayogo D. Symbiotic organisms search: A new metaheuristic optimization algorithm. *Computers and Structures*, 2014, 139(1): 98-112
- [20] Mu D L, Hui Z, Xing W W, et al. A novel nature-inspired algorithm for optimization. *Advances in Engineering Software*, 2016, 92(C): 65-88
- [21] Mirjalili S, Lewis A. The Whale optimization algorithm. *Advances in Engineering Software*, 2016, 9(5): 51-67
- [22] Chen Zhe, Geng Guo-Sheng, Cui Xing-Yue, et al. Optimization algorithm based on bionics and its application in signal processing. *Data Collection and Processing*, 2018, 33(4): 662-682(in Chinese)
(陈喆, 耿国胜, 崔行悦等. 基于仿生学的优化算法及其在信号处理中的应用. *数据采集与处理*, 2018, 33(4): 662-682)
- [23] Pierezan J, Coelho L S. Coyote optimization algorithm: A new metaheuristic for global optimization problem// *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*. Rio de Janeiro, Brazil, 2018: 2633-2640
- [24] Pitt W C, Box P W, Knowlton F F. An individual-based model of canid populations: Modelling territoriality and social structure. *Ecological Modelling*, 2003, 166(1/2): 109-121
- [25] Salcedo-Sanz S. Modern meta-heuristics based on nonlinear physics processes: A review of models and design procedures. *Physics Reports*, 2016, 655(1): 1-70
- [26] Zhang Xin-Ming, Fu Zi-Hao, Chen Hai-Yan, et al. Global guidance and interaction of coyote optimization algorithm and its application. *Computer Application Research*, 2020, 37(9): 2711-2717(in Chinese)
(张新明, 付子豪, 陈海燕等. 全局引导和相互作用的郊狼优化算法及其应用. *计算机应用研究*, 2020, 37(9): 2711-2717)
- [27] Zhang Xin-Ming, Wang Dou-Dou, Chen Hai-Yan, et al. Coyote optimization algorithm and its QAP application for strengthening optimal and worst wolves. *Computer Application*, 2019, (10): 1-8(in Chinese)
(张新明, 王豆豆, 陈海燕等. 强化最优和最差狼的郊狼优化算法及其 QAP 应用. *计算机应用*, 2019, (10): 1-8)
- [28] Eberhart R C. Empirical study of particle swarm optimization// *Proceedings of the International Conference on Evolutionary Computation*. Washington, USA, 1999: 1945-1950
- [29] Hu Jian-Xiu, Zeng Jian-Chao. Adjustment strategy of inertia weight in particle swarm optimization. *Computer Engineering*, 2007, (11): 193-195(in Chinese)
(胡建秀, 曾建潮. 微粒群算法中惯性权重的调整策略. *计算机工程*, 2007, (11): 193-195)
- [30] Zhang Yan-Mei, Jiang Shu-Juan, Chen Ruo-Yu, et al. Class integration test sequence determination method based on particle swarm optimization algorithm. *Chinese Journal of Computers*, 2018, 41(4): 931-945(in Chinese)
(张艳梅, 姜淑娟, 陈若玉等. 基于粒子群优化算法的类集成测试序列确定方法. *计算机学报*, 2018, 41(4): 931-945)
- [31] Xie Xiao-Feng, Zhang Wen-Jun, Yang Zhi-Lian. Overview of particle swarm optimization. *Control and Decision*, 2003, 18(2): 129-134(in Chinese)
(谢晓锋, 张文俊, 杨之廉. 微粒群算法综述. *控制与决策*, 2003, 18(2): 129-134)
- [32] Cui Guang-Zhao, Li Xiao-Guang, Zhang Xun-Cai, et al. DNA coding sequence optimization based on improved particle swarm optimization. *Chinese Journal of Computers*, 2010, 33(2): 311-316(in Chinese)
(崔光照, 李小广, 张勋才等. 基于改进的粒子群遗传算法的 DNA 编码序列优化. *计算机学报*, 2010, 33(2): 311-316)
- [33] Tang Yang, Wang Zi-Dong, Fang Jian-An. Feed back learning particle swarm optimization. *Applied Soft Computing*, 2011, 11(8): 4713-4725
- [34] Ren Xiao-Bo, Yang Zhong-Xiu. Improvement of particle swarm optimization algorithm. *Computer Engineering*, 2010, 36(7): 205-207(in Chinese)
(任小波, 杨忠秀. 粒子群优化算法的改进. *计算机工程*, 2010, 36(7): 205-207)
- [35] Zhao Xin. Comparative study on optimization performance of different one-dimensional chaotic maps. *Computer Application Research*, 2012, 29(3): 913-915(in Chinese)
(赵欣. 不同一维混沌映射的优化性能比较研究. *计算机应用研究*, 2012, 29(3): 913-915)
- [36] Aydilek I B. A hybrid firefly and particle swarm optimization

algorithm for computationally expensive numerical problems. *Applied Soft Computing*, 2018, 66(5): 232-249

[37] Elsisi M. Future search algorithm for optimization. *Evolutionary Intelligence*, 2019, 12(1): 21-31

[38] Li X, Tang K, Omidvar M N, et al. Benchmark functions for the CEC 2013 special session and competition on large-scale global optimization. *Gene*, 2013, 7(33): 8

[39] Wu G, Mallipeddi R, Suganthan P N. Problem definitions and evaluation criteria for the CEC 2017 competition on constrained real-parameter optimization. National University of Defense Technology, Changsha, Hunan, PR China and Kyungpook National University, Daegu, South Korea and Nanyang Technological University, Singapore, Technical



FU Jie, M. S. Her research interests include machine learning, deep neural network.

ZHOU Ding-Ning, M. S. His research interests include machine learning, deep neural network.

WANG Xin-Yu, M. S. His research interests include

Background

Neuroevolution is an important branch in the field of machine learning research. At present, the research on shallow nerve evolution mainly focuses on two aspects of neural network parameter evolution and hyperparameter evolution. Among them, the shallow neural network parameter evolution research focuses on with genetic algorithm as the main body of the bionic evolution, and evolutionary strategy adopts the strategy of “survival of the fittest” which is dominated by replication, crossover and variation. The new research results of Uber neuroevolutionary research institute show that using Evolutionary Strategy (ES) to optimize neural network can achieve or even surpass the optimization results of using Stochastic Gradient Descent(SGD). Therefore, how to design a better evolutionary strategy and explore a new neural evolution method of neural network weight is still one of the hot issues in current research.

This paper introduces the currently popular coyote optimization algorithm into the process of parameter evolution of shallow neural network, and proposes a new neural evolutionary strategy: A neural evolutionary strategy based on improved Coyote Optimization Algorithm. On the basis of the traditional Coyote Optimization Algorithm, this strategy effectively improves the convergence speed and optimization ability of the traditional Coyote Optimization Algorithm by

Report, 2017

[40] Zhao Shi-Jie, Gao Lei-Fu, Yu Dong-Mei, Tu Jun. Ant lion optimization algorithm with chaotic detection mechanism to optimize SVM parameters. *Computer Science and Exploration*, 2016, 10(5): 722-731(in Chinese)
(赵世杰, 高雷阜, 于冬梅, 徒君. 带混沌侦查机制的蚁狮优化算法优化 SVM 参数. *计算机科学与探索*, 2016, 10(5): 722-731)

[41] Zhou Zhi-Hua, Chen Shi-Fu. Neural network integration. *Chinese Journal of Computers*, 2002, 25(1): 1-8(in Chinese)
(周志华, 陈世福. 神经网络集成. *计算机学报*, 2002, 25(1): 1-8)

machine learning, deep neural network.

CHENG Mi, M. S. Her research interests include machine learning, deep neural network.

HUANG Min, M. S. Her research interests include machine learning, deep neural network.

GUO Zhi-Qing, M. S. His research interests include machine learning, deep neural network.

JIN Bao, M. S. His research interests include machine learning, deep neural network.

NIU Ying-Jie, M. S. His research interests include machine learning, deep neural network.

introducing the adaptive weighting factor and selective chaotic disturbance execution mechanism. Taking the improved Coyote Optimization Algorithm as the evolutionary method of weights and thresholds of BP neural network, this paper form a new shallow evolutionary neural network method-ICOABP, theoretical analysis and numerical experiment results show that: ICOABP is effective and has performance advantages. To a certain extent, it solve the problem of slow convergence and local minimum caused by BP neural network training by stochastic gradient descent.

This work is supported by the National Natural Science Foundation of China under Grant Nos.51974144, 71771111 and 51874160 and Project supported by discipline innovation team of Liaoning Technical University LNTU20TD-01, LN-TU20TD-07. The relevant theories and methods provide a solution to the basic scientific problems such as automatic search for the best driving path of mining trucks and intelligent simulation of the open-pit mining process system in the construction of intelligent open-pit mines.

The research results in this paper enrich and expand the research content in the field of neural evolution and provide an important reference for building a new machine learning toolbox with neural evolution as the main body.