

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2022.0404

多输入傅里叶神经网络及其麻雀搜索优化

黎亮亮, 张著洪*, 张永丹

(贵州大学 大数据与信息工程学院, 贵州省系统优化与科学计算特色重点实验室, 贵阳 550025)

摘 要: 鉴于反向传播 (BP) 神经网络存在灵敏度高但收敛速度慢, 以及已有傅里叶神经网络不具备多输入数据特征提取能力, 借助多个傅里叶神经网络构建能接收多维数据的堆叠神经网络, 进而将其与多层感知器融合, 获得基于梯度下降的多输入傅里叶神经网络。结合此神经网络获取全局最优参数值难的因素, 通过在麻雀搜索算法中引入 Cat 混沌映射、动态种群规模调节机制及参数自适应调节方案, 提出改进型麻雀搜索算法, 并将其应用于多输入傅里叶神经网络的参数优化及高维函数优化问题的求解。理论分析可得, 所提算法的计算复杂度主要由种群规模和优化问题的维度决定。比较性的数值实验表明, 所获神经网络提取多源数据特征的能力和泛化能力强, 同时所提算法处理高维优化问题具有明显优势且收敛速度快。

关 键 词: 傅里叶神经网络; 多层感知器; 麻雀搜索; 高维函数优化; 多属性分类

中图分类号: TP301.6

文献标志码: A

文章编号: 1001-5965(2024)02-0623-11

人工神经网络已被广泛用于函数拟合、风险预测、系统控制、模式分类等^[1], 但权值与阈值多且冗余神经元难以剔除, 导致存在公认的收敛速度慢、易陷入局部搜索、稳定性差、抗噪能力弱、易梯度消失和梯度爆炸等不足^[2]。目前, 多输入神经网络主要有反向传播神经网络 (back-propagation neural network, BPNN)、极限学习机 (extreme learning machine, ELM)、径向基函数 (radial basis function, RBF) 神经网络、小波神经网络 (wavelet neural network, WNN)、循环神经网络、卷积神经网络等。前 3 种神经网络不具备噪声抑制功能, 且对输入特征的提取能力均偏弱; 小波神经网络具有强噪声抑制能力, 但不具备特征提取功能, 且收敛速度慢及求解精度不高; 循环神经网络具有记忆性、参数共享和特征提取功能, 但噪声抑制能力较弱且计算复杂度高; 卷积神经网络虽然有强的特征提取功能, 但缺乏噪声抑制能力, 致使数据分类效果受噪声干扰大, 网络训练耗时长。因此, 研究噪声抑制和特

征提取能力强, 且网络训练效果好和效率高的多输入神经网络, 已成为神经网络研究面临的重要课题。

三层傅里叶神经网络 (Fourier neural network, FNN)^[3] 提出后, 因其结构简单、权值与阈值更新速度快, 且具有特征表示、时间序列信号模拟和噪声抑制能力及数据特征提取能力强, 已被应用于时间序列分析、扰动抑制、控制器的参数实时调节、预测等问题^[3-7], 但因信号特征提取单一且仅适合于二维函数逼近、时间序列分析等, 导致应用受限。于是, 构建能同时提取多种输入信号的特征信息、求解精度高、不易陷入局部最优、参数更新速度快, 且特征提取与噪声抑制能力强的多输入 FNN (multi-input FNN, MIFNN) 已受到重视。

现有神经网络的参数优化方法大致可概括为 2 类, 即诸如梯度下降、最小二乘法、牛顿法等数值优化^[8-9] 和智能优化^[10-11]。数值优化是基于最优性条件的数值迭代算法, 其更新步长直接影响算法的全局收敛性, 常常仅能获得局部最优解。例如, 杨

收稿日期: 2022-05-21; 录用日期: 2022-11-25; 网络出版时间: 2022-12-14 09:27

网络出版地址: link.cnki.net/urlid/11.2625.V.20221213.1513.006

基金项目: 国家自然科学基金 (62063002)

* 通信作者. E-mail: zhzhzhang@gzu.edu.cn

引用格式: 黎亮亮, 张著洪, 张永丹. 多输入傅里叶神经网络及其麻雀搜索优化 [J]. 北京航空航天大学学报, 2024, 50 (2): 623-633.

LI L L, ZHANG Z H, ZHANG Y D. Multi-input Fourier neural network and its sparrow search optimization [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50 (2): 623-633 (in Chinese).

旭华^[1]提出三层单输入、多输出 FNN, 其借助最小二乘法优化网络参数, 获得的网络能有效降低噪声对样本特征提取的影响, 但仅适用于小规模数据集的函数逼近; Lin 等^[9]利用牛顿法训练网络参数, 能使网络的泛化能力相较于 FNN 有所提升, 但算法的计算复杂度高。另外, 智能优化是源于自然界中生物进化、觅食行为特性或物理现象启发的随机搜索算法, 其因结构简单、收敛速度快^[10-11], 已成为较为活跃的研究分支, 并已有众多具有强竞争力的新型优化算法被提出, 如麻雀搜索算法 (sparrow search algorithm, SSA)^[12]、 Harris 鹰优化 (Harris hawks optimization, HHO) 算法^[13]、天鹰优化 (aquila optimizer, AO) 算法^[14]、野狗优化算法 (dingo optimization algorithm, DOA)^[15]、鲸鱼优化算法 (whale optimization algorithm, WOA)^[16]等。这些算法均经由生物行为特性而提出, 属于群智能优化算法, 且求解中、低维优化问题具有强的竞争力。文献 [17] 和文献 [18] 分别将正余弦策略与柯西变异引入 SSA, 获得的 SSA 可有效平衡全局开采和局部勘测能力, 且收敛速度快及寻优精度高, 但搜索效果欠稳定且在应用于高维优化问题时, 搜索性能急剧退化, 以及早熟现象较为严重。此类算法常由已获得的最好个体引导种群进化或种群更新, 算法在进化后期出现种群多样性不足, 进而出现局部收敛现象。例如, 文献 [19] 和文献 [20] 通过将种群划分思想和精英引导的学习策略引入到粒子群优化算法中, 获得的算法能有效平衡全局开采和局部勘测能力, 可提高算法在高维函数优化问题上的求解精度, 但算法的计算复杂度较高。Zhang 等将果蝇的视觉信息处理机制与智能优化算法中种群更新策略结合^[21-22], 获得果蝇视觉进化神经网络, 其相比于元启发式随机搜索算法, 求解高维优化问题的性能有明显优势且效率高。尽管如此, 全局收敛能力强、结构简单且收敛速度快的新型优化算法仍是智能优化研究关注的重点。

综上, 本文通过将多层感知器 (multi-layer perceptron, MLP) 与 FNN 结合, 提出基于梯度下降的 MIFNN; 进而, 提出可求解高维优化问题的改进型麻雀搜索算法 (improved sparrow search algorithm, ISSA), 并将其用于 MIFNN 的结构优化。

1 多输入傅里叶神经网络模型与设计

1.1 神经网络的拓扑结构

传统的 FNN^[1,3] 是一种由 $2L+2$ 个神经节点依次连接而成的单输入单输出三层神经网络, 其中间层包含 $2L$ 个神经节点, 每个节点的激励函数在正交

基集 $\{1, \sin nx, \cos nx, 1 \leq n \leq L\}$ 中, 且输出由傅里叶级数的有限项表示, 即

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^L (a_n \cos nx + b_n \sin nx) \tag{1}$$

其中, a_0 、 a_n 、 b_n 为傅里叶系数, 且在神经网络中为神经节点的连接权值。由于此神经网络存在待定的 $2L+1$ 个权重, 当其触及多源信号拟合、多维数据挖掘时, 必然存在大量尚需确定的权值。为此, 本文将 MLP 与基于正交基集 $\{1, \cos nx, 1 \leq n \leq L\}$ 的堆叠 FNN 组合, 建立 MIFNN, 如图 1 所示。

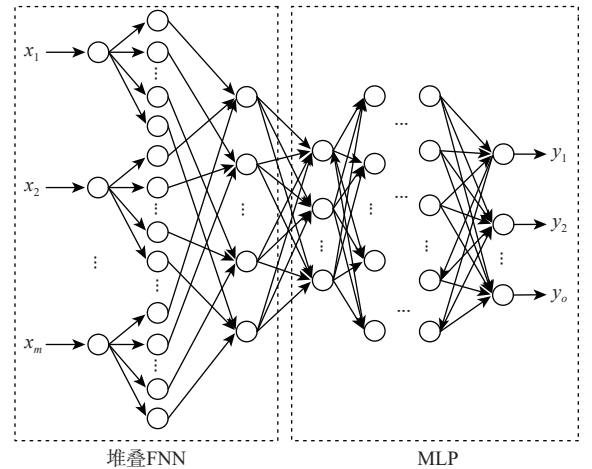


图 1 MIFNN 的结构示意图
Fig. 1 Schematic diagram of MIFNN structure

图 1 中, 输入向量为 $[x_1, x_2, \cdots, x_m]$, m 为输入维度; 输出向量为 $[y_1, y_2, \cdots, y_o]$, o 为输出维度; MIFNN 的前馈结构由堆叠 FNN 与 MLP 顺次连接而成。堆叠 FNN 的主要作用是将多源输入信息中每个分量信息经由多通道进行特征提取, 去除冗余噪声信息; 其由 m 个结构相同的单输入 FNN 堆叠而成, 包含 m 个输入神经元、 mL 个扩展神经元及 L 个输出神经元; 每个子网络的扩展神经元个数为 L ; 特别地, 第 k 个输出神经元的输出是每个 FNN 的扩展层中第 k 个神经元的输出量的线性加权组合。MLP 包含 $K+1$ 个神经层, 其输入层与堆叠 FNN 的输出层通过全连接方式连接, 且第 k 层的神经元个数为 s_k ; 每层的神经元激励函数均为 $g(\cdot)$ 。由此可知, MIFNN 首先以堆叠 FNN 的输入作为输入, 进而以 MLP 的输出为输出。

结合单输入 FNN 的固有特性, 堆叠 FNN 的主要任务是多通道提取多输入特征, 消除冗余噪声对特征提取的影响, 其每个输出神经元仅接收所有单输入 FNN 的扩展层中相同位置的神经元输出, 并经由加权方式获得输出。换言之, 其借助多通道对多源输入数据进行正交扩展, 实现多通道特征提取

和噪声抑制,继而融合多输入的属性特征;其次,每个单输入 FNN 的扩展层与堆叠 FNN 的输出层的神经元个数相等,原因在于堆叠 FNN 的每个输出神经元仅接收各单输入 FNN 对应通道的信息,此有助于利用多通道提取输入特征信息及提高神经网络的训练效率。另外, MIFNN 利用堆叠 FNN 提取多输入的特征,进而利用 MLP 实现函数拟合或数据分类,此二者的结合既可回避噪声对问题求解质量的影响,又可提升问题求解的效果与效率。

1.2 误差损失函数

假定样本集 Ω 由 M 个样例构成,即 $\{(\mathbf{x}(n), \mathbf{d}(n)) | 1 \leq n \leq M\}$, 其中 $\mathbf{x}(n)$ 为第 n 时刻的输入向量, 维度为 m ; $\mathbf{d}(n)$ 为该样例的标签或观测向量。MIFNN 接收输入样例 $\mathbf{x}(n)$ 后, 经由堆叠 FNN 中各个单输入 FNN 的余弦正交扩展, 提取样例的深度特征信息, 进而经由 MLP 产生输出信号。MIFNN 的详细设计如下:

1) 输入层: 由 m 个神经节点表示, 其节点 i 在第 n 时刻接收输入向量 $\mathbf{x}(n)$ 中第 i 个分量 $x_i(n)$; 每个神经节点与相应扩展层中各神经元的连接权值为 1。

2) 扩展层: 通过余弦正交扩展, 深度提取 $\mathbf{x}(n)$ 中每个维度的特征; 第 i 个输入分量 $x_i(n)$ 对应的 FNN 中扩展神经元 j 的输出表示如下:

$$y_{j+(i-1)L}^{(E)}(n) = \cos(jx_i(n)) \quad 1 \leq i \leq m, 1 \leq j \leq L \quad (2)$$

3) 堆叠 FNN 的第 j 个输出神经元接收 m 个 FNN 中每个网络的第 j 个扩展神经元的输出, 并经由权值 w_{ij}^F 线性组合获得其输出, 即

$$y_j^{(F)}(n) = \sum_{i=1}^m w_{ij}^F(n) y_{j+(i-1)L}^{(E)}(n) \quad (3)$$

4) 通过全连接方式将堆叠 FNN 的输出层中各神经元的输出映射为 MLP 的输入, 并经由激励函数 $g(\cdot)$ 获相应输出, 即 MLP 的输入层中第 j 个神经元的输出计算式为

$$y_j^{(0)}(n) = g(H_j^{(0)}(n)) = g\left(\sum_{i=1}^L w_{ij}^{(0)}(n) y_i^{(F)}(n) + b_j^{(0)}(n)\right) \quad (4)$$

式中: $1 \leq j \leq s_0$, $H_j^{(0)}(n)$ 为 MLP 的输入层中第 j 个神经元在第 n 时刻的诱导局部域; s_0 为 MLP 输入层的神经元个数; $b_j^{(0)}$ 为 MLP 的输入层中第 j 个神经元的阈值; $w_{ij}^{(0)}$ 为堆叠 FNN 的输出层中第 i 个神经元与 MLP 的输入层中第 j 个神经元的连接权值。

5) MLP 的第 l 个隐含层中第 j 个神经元的输出计算式为

$$y_j^{(l)}(n) = g(H_j^{(l)}(n)) = g\left(\sum_{i=1}^{s_{l-1}} w_{ij}^{(l)}(n) y_i^{(l-1)}(n) + b_j^{(l)}(n)\right) \quad (5)$$

式中: $1 \leq j \leq s_l, 1 \leq l \leq K-1$, $H_j^{(l)}(n)$ 为第 l 层中第 j 个神经元在第 n 时刻的诱导局部域; $b_j^{(l)}$ 、 $w_{ij}^{(l)}$ 和 s_{l-1} 分别为第 l 层中第 j 个神经元的阈值、该神经元与第 $l-1$ 层中第 i 个神经元的连接权值和第 $l-1$ 层的神经元个数。

6) MIFNN 的输出层中第 j 个神经元的输出计算式为

$$o_j(n) = g\left(\sum_{i=1}^{s_{K-1}} w_{ij}^{(K)}(n) y_i^{(K-1)}(n) + b_j^{(K)}(n)\right) \quad (6)$$

式中: $1 \leq j \leq s_K$, $b_j^{(K)}$ 、 $w_{ij}^{(K)}$ 和 s_K 分别为输出层中第 j 个神经元的阈值、此神经元与第 $K-1$ 个隐含层中第 i 个神经元的连接权值和此层的神经元个数。于是, 在线学习下第 n 时刻的损失函数可表示为

$$E_n(\mathbf{W}, \mathbf{b}) = \frac{1}{2} \sum_{j=1}^{s_K} (o_j(n) - d_j(n))^2 \quad (7)$$

式中: $\mathbf{W}$ 和 $\mathbf{b}$ 分别为 MIFNN 的权值和阈值构成的向量。

1.3 权值与阈值更新

传统的 BP 方法是将前馈传播产生的误差量反向传递给各神经层, 并借助梯度下降法更新网络中的 $\mathbf{W}$ 和 $\mathbf{b}$, 此易于使网络参数的更新陷入局部最优且收敛速度慢。然而, 动量项的引入却有助于提高梯度下降法的收敛速度和增强抗噪能力及网络训练的稳定性^[1-2]。为此, 建立含动量项的梯度下降 MIFNN (momentum and gradient descent based MIFNN, MGD-MIFNN), 具体如下:

1) 全连接层 (MLP) 的 BP: 此为熟知且含动量项的 BP。具体而言, 基于式(7), 在第 n 时刻下第 k 个隐含层中第 j 个神经节点的局部梯度信息计算式为

$$\delta_j^{(k)}(n) = -\frac{\partial E_n}{\partial H_j^{(k)}(n)} = -\frac{\partial E_n}{\partial y_j^{(k)}(n)} g'(H_j^{(k)}(n)) \quad (8)$$

$$1 \leq j \leq s_K$$

因此, 第 $n+1$ 时刻下 MLP 的第 k 个隐含层中第 j 个神经元节点与上一个隐含层中第 i 个神经元节点之间的连接权值由下式更新:

$$w_{ij}^{(k)}(n+1) = w_{ij}^{(k)}(n) - \eta \frac{\partial E_n}{\partial w_{ij}^{(k)}(n)} - \alpha \frac{\partial E_{n-1}}{\partial w_{ij}^{(k)}(n-1)} \quad (9)$$

$$\frac{\partial E_n}{\partial w_{ij}^{(k)}(n)} = \frac{\partial H_j^{(k)}(n)}{\partial w_{ij}^{(k)}(n)} \cdot \frac{\partial E_n}{\partial H_j^{(k)}(n)} = -y_i^{(k-1)}(n) \delta_j^{(k)}(n) \quad (10)$$

其中, η 和 α 分别为固定的学习率和动量因子。同理, 在第 $n+1$ 时刻下第 k 个隐含层中第 j 个神经元

节点的阈值经由下式更新:

$$b_j^{(k)}(n+1) = b_j^{(k)}(n) - \eta \frac{\partial E_n}{\partial b_j^{(k)}(n)} - \alpha \frac{\partial E_{n-1}}{\partial b_j^{(k)}(n-1)} \quad (11)$$

$$\frac{\partial E_n}{\partial b_j^{(k)}(n)} = \frac{\partial H_j^{(k)}(n)}{\partial b_j^{(k)}(n)} \cdot \frac{\partial E_n}{\partial H_j^{(k)}(n)} = -\delta_j^{(k)}(n) \quad (12)$$

综上, MLP 中含动量项的权值和阈值更新为

$$w_{ij}^{(k)}(n+1) = \begin{cases} w_{ij}^{(k)}(n) + \eta y_i^{(k-1)} \delta_j^{(k)}(n) - \\ \alpha \frac{\partial E_{n-1}}{\partial w_{ij}^{(k)}(n-1)} & 1 \leq k \leq K \\ w_{ij}^{(k)}(n) + \eta y_i^{(F)} \delta_j^{(k)}(n) - \\ \alpha \frac{\partial E_{n-1}}{\partial w_{ij}^{(k)}(n-1)} & k < 1, k > K \end{cases} \quad (13)$$

式中: $1 \leq j \leq s_k, 1 \leq i \leq s_{k-1}$ 。另外,

$$b_j^{(k)}(n+1) = b_j^{(k)}(n) + \eta \delta_j^{(k)}(n) - \alpha \frac{\partial E_{n-1}}{\partial b_j^{(k)}(n-1)} \quad 0 \leq k \leq K, 1 \leq j \leq s_k \quad (14)$$

2) 堆叠 FNN: 输出层中第 j 个神经元与第 i 个 FNN 中第 j 个扩展神经元的连接权值更新如下:

$$w_{ij}^F(n+1) = w_{ij}^F(n) - \eta y_i^{(k-1)} \frac{\partial E_n}{\partial w_{ij}^F(n)} - \alpha \frac{\partial E_{n-1}}{\partial w_{ij}^F(n-1)} \quad (15)$$

其中,

$$\frac{\partial E_n}{\partial w_{ij}^F(n)} = \frac{\partial y_j^{(F)}(n)}{\partial w_{ij}^F(n)} \cdot \frac{\partial E_n}{\partial y_j^{(F)}(n)} = y_{j+(i-1)N}^{(E)}(n) \frac{\partial E_n}{\partial y_j^{(F)}(n)} \quad (16)$$

$$w_{ij}^F(n+1) = w_{ij}^F(n) - \eta y_{j+(i-1)N}^{(E)}(n) \frac{\partial E_n}{\partial y_j^{(F)}(n)} - \alpha \frac{\partial E_{n-1}}{\partial w_{ij}^F(n-1)} \quad 1 \leq i \leq m, 1 \leq j \leq L \quad (17)$$

MGD-MIFNN 在 MLP 基础上增加堆叠 FNN, 可增强多属性特征提取的能力, 提升神经网络的拟合能力。该神经网络与 FNN 相比, 能处理多输入下信号拟合、数据分类等问题, 应用范围广; 与 MLP 相比, 通过堆叠 FNN 对输入信号作平滑处理, 消除隐含的噪声, 使得模型的输出受噪声干扰小, 函数逼近能力强。

MGD-MIFNN 是基于带动量项的梯度下降法获得网络的权值和阈值, 此在超参数下, 必然易于仅能获得局部最佳参数向量值, 进而较难使得损失函数达到全局最小。因此, 探讨求解高维优化问题的 ISSA, 使损失函数达到全局最小。

2 改进型麻雀搜索算法

2.1 麻雀搜索算法

SSA 由 3 种类型麻雀组成^[17], 即发现者、加入者和侦察者。发现者拥有较大的适应度值, 负责为

加入者提供觅食区域和方向; 加入者为了发现食物源, 会一直跟随发现者, 且不断监视发现者并争夺食物; 在发现者和加入者中随机选取部分麻雀作为侦察者, 其作用是当发生危险时产生反捕食行为。

在第 t 时刻, SSA 选取规模为 N 的麻雀种群 $X(t)$ 中适应度值较大的前 20% 麻雀构成发现者种群 $Y(t)$, 个体 i 的位置为 $y_i(t)$, 即 $[y_{i1}(t), y_{i2}(t), \dots, y_{iD}(t)]$ D 为函数优化问题的决策向量维度; 第 i 个发现者的位置更新描述如下:

$$y_i(t+1) = \begin{cases} y_i(t) \exp\left(-\frac{i}{\alpha T}\right) & R < G \\ y_i(t) + QL & R \geq G \end{cases} \quad (18)$$

式中: T 为最大时间步长; α 和 R 分别为区间 $[0, 1]$ 上的随机数和预警值; G 为安全值; Q 为服从标准正态分布的非线性步长因子; L 为元素均为 1 的 D 维向量。

$X(t)$ 中剩下的 80% 麻雀构成加入者种群 $Z(t)$, 加入者 i 的位置为 $z_i(t)$, 即 $[z_{i1}(t), z_{i2}(t), \dots, z_{iD}(t)]$, 其位置经由式 (19) 更新:

$$z_i(t+1) = \begin{cases} Q \exp\left(\frac{\mathbf{x}_w(t) - \mathbf{z}_i(t)}{i^2}\right) & i < \frac{N}{2} \\ \mathbf{x}_p(t+1) + |\mathbf{z}_i(t) - \mathbf{x}_p(t+1)| \mathbf{A}^* \mathbf{L} & i \geq \frac{N}{2} \end{cases} \quad (19)$$

式中: $\mathbf{x}_w(t)$ 和 $\mathbf{x}_p(t+1)$ 分别为 $X(t)$ 中适应度最低和最高的个体; $\mathbf{A}$ 表示一个与麻雀个体的位置同维度的列向量, 其内部元素由 1 和 -1 随机组成, $\mathbf{A}^+ = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1}$ 。

侦察者种群 $P(t)$ 由 $X(t)$ 中随机抽取的 20% 个麻雀构成; 侦察者 i 的位置为 $p_i(t)$, 即 $[p_{i1}(t), p_{i2}(t), \dots, p_{iD}(t)]$, 经由式 (20) 进行位置更新:

$$p_i(t+1) = \begin{cases} \mathbf{x}_g(t) + \beta(\mathbf{p}_i(t) - \mathbf{x}_g(t)) & f_i > f_g \\ \mathbf{p}_i(t) + \gamma \frac{(\mathbf{p}_i(t) - \mathbf{x}_w(t))}{f_i - f_g} & f_i \leq f_g \end{cases} \quad (20)$$

式中: $\mathbf{x}_g(t)$ 为到第 t 时刻为止获得的全局最好个体, 其适应度表示为 f_g ; β 和 γ 为非线性步长因子; f_i 为 $Z(t)$ 中麻雀 i 的适应度; f_w 为当前种群中最差麻雀的适应度。

2.2 改进型麻雀搜索算法

虽然 SSA 具有求解精度高、稳定性好等优点, 但面对高维或性能指标极为复杂的优化问题时, 其搜索性能和效率急剧退化, 存在的主要不足如下:

1) 初始种群中麻雀的位置因分布不均匀而导致算法搜索过早出现适应度值远超过平均水平的“超级麻雀”, 使其余的大量麻雀聚集, 丧失种群多样性;

2) 发现者和加入者种群的规模不变性导致算法在迭代前期, 发现者的数量相对较少, 而在迭代后期, 发现者的数量又相对较多, 不利于有效地从全局开采向局部勘测转变, 进而导致算法的收敛速度慢;

3) 步长因子 Q 、 β 和 γ 的取值随机性直接影响 SSA 的寻优效率和效果, 导致 SSA 易于陷入局部搜索。

2.2.1 算法改进

针对 SSA 存在的不足, 围绕初始种群的生成方式, 发现者和加入者所占比例及算法参数进行如下改进。

1) 初始种群: 在 SSA 中, 种群随机初始化致使初始个体在解空间分布不均, 容易出现聚集现象, 从而导致种群多样性不足^[17]。鉴于 Cat 混沌映射结构简单, 且在 $[0,1]$ 间产生的混沌序列分布均匀^[23], 在此借助式 (21) 产生混沌序列:

$$\begin{bmatrix} q_{n+1} \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} q_n \\ v_n \end{bmatrix} \bmod 1 \quad (21)$$

式中: $\bmod 1$ 表示取余运算符; q_n 和 v_n 为 0 到 1 之间的随机数。进而, 经由式 (22) 生成初始个体:

$$x_j = l_j + k_j(u_j - l_j) \quad 1 \leq j \leq D \quad (22)$$

式中: u_j 和 l_j 分别为优化问题的候选解 $\mathbf{x}$ 中第 j 个分量的上、下界; k_j 由式 (21) 确定依次取 q_n 的值。

2) 种群规模动态调节机制: 算法迭代初期, 发现者种群 $X(t)$ 的规模要求较大, 有利于增强处理高维优化问题的全局搜索能力。随着迭代次数的增加, 种群规模要求自适应减小, 加入者种群的规模自适应增大, 有助于增强局部搜索能力。具体而言, $X(t)$ 的规模由式 (23) 确定:

$$N_p = \lfloor rN \rfloor \quad (23)$$

式中: $\lfloor \cdot \rfloor$ 表示向下取整; r 为发现者在种群中的占比, 其由式 (24) 确定:

$$r = 0.25 + 0.3 \left(1 - \tan \frac{\pi t}{4T} \right) \quad (24)$$

$Y(t)$ 的规模由式 (25) 确定:

$$N_s = N - N_p \quad (25)$$

3) 非线性步长因子: Q 、 β 和 γ 易于导致 SSA 在迭代中、后期快速陷入局部搜索; 步长因子越大, 越容易发现全局最优解, 但收敛速度慢; 反之, 算法的局部勘测能力强, 收敛速度慢。为权衡高维优化问题求解情形下的寻优质量和搜索效率, 要求 Q 、 β 和 γ 随 t 的增大自适应地非线性递减, 即

$$Q = \frac{\exp \left(2 \left(1 - \frac{t}{T} \right) \right) - \exp \left(-2 \left(1 - \frac{t}{T} \right) \right)}{\exp \left(2 \left(1 - \frac{t}{T} \right) \right) + \exp \left(-2 \left(1 - \frac{t}{T} \right) \right)} \quad (26)$$

$$\beta = (f_g - f_i) \sin \left(\pi \exp \left(-\frac{3}{2} \tan \left(\frac{t}{T} \right)^2 \right) \right) \quad (27)$$

其中, γ 与 β 的取值相同。

2.2.2 算法描述

ISSA 的算法描述如下:

步骤 1 输入参数: 种群规模 N , 维度 D , 最大迭代数 T , 安全值 G ;

步骤 2 置 $t \leftarrow 1$, 依据式 (21) 和式 (22) 初始化麻雀种群 $X(t)$, 适应度最高、最低的个体分别记作 $\mathbf{x}_g$ 和 $\mathbf{x}_w$;

步骤 3 经由式 (23) 和式 (24) 计算发现者、加入者种群的规模 N_p 及 N_s ;

步骤 4 $X(t)$ 中适应度较高的 N_p 个个体构成 $Y(t)$, 其余的个体构成 $Z(t)$;

步骤 5 由式 (26) 和式 (27) 计算非线性步长因子 Q 、 β 及 γ ;

步骤 6 经由式 (18) 更新种群 $Y(t)$;

步骤 7 经由式 (19) 更新种群 $Z(t)$;

步骤 8 从 $X(t)$ 中随机抽取 20% 个体构成 $P(t)$, 并据式 (20) 更新 $P(t)$;

步骤 9 $Y(t) \cup Z(t) \cup P(t)$ 中适应度较高的 N 个个体构成 $X(t+1)$, 更新 $\mathbf{x}_g$ 和 $\mathbf{x}_w$;

步骤 10 置 $t \leftarrow t+1$; 若 $t < T$, 则返回步骤 4; 否则, 结束, 且输出 $\mathbf{x}_g$ 。

ISSA 的计算复杂度主要由步骤 6~步骤 8 确定。在一个迭代周期内, 步骤 6 共执行 $2N_p D + N_p$ 次算术运算; 步骤 7 共运行 $5N_s D + N_s$ 次; 步骤 8 共运行 $0.2N + 0.8ND$ 次; 因此, 步骤 6~步骤 8 的计算次数为 $3N_s D + 1.2N + 2.8ND$ 。由于 N_s 介于 $0.45N \sim 0.75N$ 之间, 因此, ISSA 的计算复杂度主要由种群规模 N 和优化问题的维度 D 决定。

3 数值实验

在配置为 Windows10 (CPU/3.7 GHz, RAM/4.0 GB)/MATLAB 2018 环境下展开数值实验。首先, 借助具有强竞争力的元启发式随机搜索算法和基准测试问题, 检测 ISSA 求解高维函数优化问题的性能; 其次, 以开源数据集为例, 测试 MGD-MIFNN 和基于 ISSA 的 MIFNN 处理数据分类问题的性能; 最后, 测试 MIFNN 的结构变化对数据分类性能的影响。

3.1 ISSA 的性能测试

与 ISSA 比较的算法包括 SSA^[12]、HHO 算法^[13]

、AO 算法^[14]、DOA^[15] 及 WOA^[16]。对比算法的参数设置除种群规模和最大迭代数外均源于相应文献；设置各算法的种群规模为 100，最大迭代次数为 100；ISSA 的安全值 G 取 0.8。测试事例包括 8 种具有代表的基准函数优化问题 $F_1 \sim F_8$ ^[24]，其中 $F_1 \sim F_4$ 是单峰值函数， $F_5 \sim F_8$ 是含大量局部最优解的多峰值函数；每个事例的维度均取为 2 000，因此， $F_1 \sim F_8$ 为高维的单峰、多峰值函数，其相应的优化问题为高维函数优化问题。对于每个测试事例，每种算法均独立运行 25 次，且将各自获得的 25 个最好目标值的均值 μ 、均方差 σ 用于算法的比较分析，如表 1 和表 2 所示；以 F_2 、 F_4 、 F_6 、 F_8 为例，各算法的搜索曲线如图 2 所示。

经由表 1 和表 2 的均值可知，DOA 和 WOA 求

解以上 8 个测试事例的性能整体上劣于其他算法，同时表中的均方差值说明，WOA 求解多模态优化问题的能力相对偏弱；HHO 算法和 AO 算法在获得的解质量和搜索效果的稳定性方面整体上劣于 SSA，且 SSA 和 AO 算法均不能有效求解 F_5 ；SSA 具有一定的优势。另外，ISSA 能成功求解以上各测试问题，特别是求解 $F_5 \sim F_8$ 获得的解质量较为理想。因此，相比对比算法，ISSA 求解高维优化问题具有明显优势。由图 2 可知，相比于对比算法，ISSA 在初始时刻能产生多样且质量较好的个体，表明 Cat 混沌映射能增强初始种群的多样性和分布的均匀性；ISSA 的收敛速度快且获得的解的精度高，此说明非线性步长因子能提升寻优效率，且种群规模动态调节有助于平衡算法的全局开采与局部勘测能力。

表 1 各算法独立运行 25 次后获得的统计结果 ($F_1 \sim F_4$)

算法	μ				σ			
	F_1	F_2	F_3	F_4	F_1	F_2	F_3	F_4
SSA	2.47×10	3.20×10^{-1}	5.14×10^{-4}	1.80×10^{-3}	1.91×10^{-1}	5.51×10^{-1}	1.10×10^{-3}	1.60×10^{-3}
HHO	6.45×10^{-1}	7.44×10^0	1.26×10^{-1}	5.79×10^{-4}	2.58×10^{-1}	9.25×10^0	1.22×10^{-1}	5.52×10^{-4}
AO	7.07×10^{-1}	7.05×10^0	2.96×10^{-2}	3.04×10^{-4}	1.24×10^{-1}	1.68×10	4.57×10^{-2}	3.42×10^{-4}
DOA	7.50×10^{-1}	2.00×10^3	4.97×10^2	1.90×10^{-3}	1.46×10^{-4}	5.80×10^{-2}	1.30×10^0	2.50×10^{-3}
WOA	7.50×10^{-1}	1.99×10^3	1.56×10^2	1.49×10^{-2}	1.49×10^{-5}	1.74×10^0	4.02×10	1.67×10^{-2}
ISSA	2.17×10^{-4}	6.90×10^{-4}	9.43×10^{-6}	6.35×10^{-4}	8.19×10^{-4}	2.80×10^{-3}	3.66×10^{-5}	5.99×10^{-4}

表 2 各算法独立运行 25 次后获得的统计结果 ($F_5 \sim F_8$)

算法	μ				σ			
	F_5	F_6	F_7	F_8	F_5	F_6	F_7	F_8
SSA	-2.10×10^5	6.72×10^{-7}	1.91×10^{-4}	1.55×10^{-2}	2.30×10^4	1.16×10^{-6}	1.98×10^{-4}	5.27×10^{-2}
HHO	-8.31×10^5	1.32×10^{-5}	1.42×10^{-2}	5.40×10^{-3}	3.46×10^4	2.12×10^{-5}	1.34×10^{-2}	9.10×10^{-3}
AO	-1.28×10^5	5.33×10^{-6}	1.77×10^{-2}	2.15×10	3.76×10^4	9.32×10^{-6}	4.21×10^{-2}	1.23×10
DOA	-4.68×10^4	1.17×10^0	2.00×10^2	4.53×10	7.49×10^3	6.90×10^{-3}	5.30×10^{-3}	2.05×10^0
WOA	-6.75×10^5	1.17×10^{-1}	7.76×10	1.05×10	1.18×10^5	4.34×10^{-2}	2.28×10	9.39×10^0
ISSA	-8.38×10^5	1.03×10^{-8}	6.26×10^{-6}	7.22×10^{-5}	2.59×10^{-2}	2.58×10^{-8}	1.87×10^{-5}	1.25×10^{-4}

3.2 数据分类

选取开源 UCI 数据集中 11 组具有代表性的数据集作为数据分类事例，即 Iris、Wine、Haberman、Hayes、Immunotherapy、Cryotherapy、Heart、Breast-Cancer、Transfusion、Audit 和 Banknote。各数据集的属性个数、样本容量及类别数构成的向量依次为 [4,150,3]、[13,178,3]、[3,306,2]、[5,132,3]、[7,90,2]、[6,90,2]、[12,299,2]、[9,116,2]、[4,748,2]、[17,773,2] 和 [4,1 372,2]。在每个数据集中随机抽取 70% 的样例构成训练集，其他样例构成测试集。

3.2.1 MIFNN 的样本分类性能测试

为测试批量损失函数下基于梯度下降的 MGD-MIFNN 和 ISSA 的 MIFNN(简记作 ISSA-MIFNN)处理以上样本分类问题的性能，一方面选取 4 种经典神经网络(BPNN、ELM、RBF、WNN)参与 MGD-MIFNN 的性能比较；另一方面，将以上参与比较的 5 种智能优化算法在批量损失函数下寻找 MIFNN 的最优权值和阈值，记此时的算法为 X-MIFNN，在此 X 意指 SSA、HHO、AO、DOA、WOA 或 ISSA。

MIFNN 与 BP 均包含 2 个隐含层，各自的第 1、

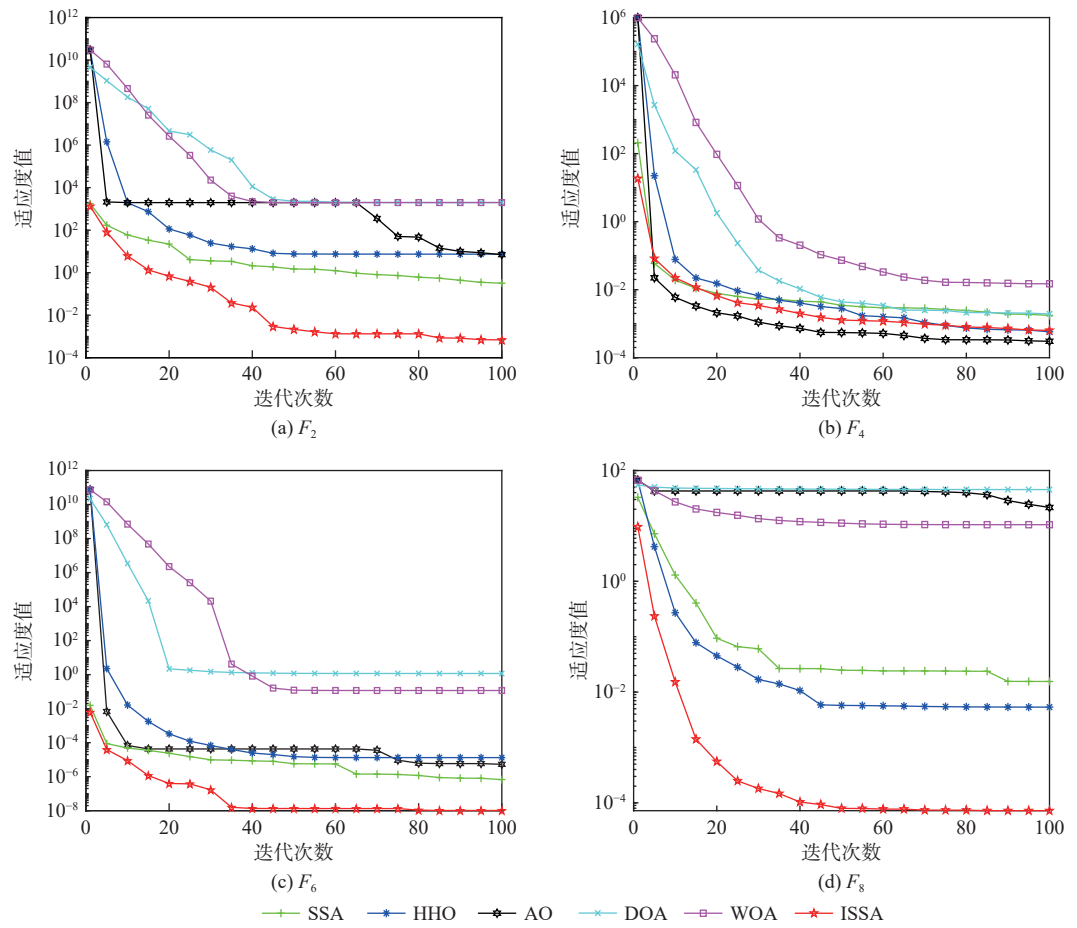


图2 各算法的平均搜索曲线

Fig. 2 Average search curves of algorithms

第2 隐含层均分别含 15 个和 5 个神经元,同时 MIFNN 的每个扩展子层包含的神经元个数 L 取 15 且 MLP 的输入层神经元个数为 5; ELM、RBF 与 WNN 均仅含 1 个隐含层,其神经元个数为 20; 设置梯度下降学习率为 0.08, 动量因子为 0.05, 非线性激活函数为 tanh 函数, 网络训练的最大迭代次数为 300。另外, ISSA-MIFNN、SSA-MIFNN、HHO-MIFNN、AO-MIFNN、DOA-MIFNN 及 WOA-MIFNN 的种群大小设置为 30, 网络的权值和阈值所在区间为 $[-10,10]$, 其他参数设置同上。各基于梯度下降、智能优化的神经网络获得的训练分类准确率 TrAcc 及测试分类准确率 TeAcc 分别如表 3 和表 4 所示。以 Haberman、Immunotherapy、Cryotherapy 和 Transfusion 为例, 各神经网络的损失函数值变化曲线如图 3 所示。

由表 3 和表 4 可知, MGD-MIFNN 与 BP、ELM、RBF、WNN 相比, 其对其中 9 种数据集的训练分类性能最好, 且对其中 8 种数据集的测试性能具有明显优势。然而, 其他神经网络至多求解 3 种数据集的分类问题有优势, 因此, MGD-MIFNN 的优势较为明显, 且适用范围较广。尽管 MGD-MIFNN 已展

现出优势, 但其求解诸如 Wine、Immunotherapy、Transfusion 等分类问题的测试分类精度有待提高, 根源在于动量项并不能完全使梯度下降法摆脱陷入局部搜索的困境, 进而导致对特定的分类问题出现过拟合或欠拟合。

与 MGD-MIFNN 相比, ISSA-MIFNN 在数据集的训练分类准确率方面, 对于 Haberman、Heart、Transfusion 分类问题的训练分类精度得到显著提升, 且二者对于 Immunotherapy、Audit 的训练分类准确率均为 100%。ISSA-MIFNN 比 MGD-MIFNN 的测试分类性能有显著性的提升, 且对于 8 种数据集, 获得的测试分类准确率提升较大, 仅求解数据集 Banknote 测试集获得的测试分类准确率提升 1.95%。

由此可知, ISSA 能有效改善 MIFNN 的数据分类性能, 且 ISSA-MIFNN 能有效地回避 MGD-MIFNN 出现欠拟合或过拟合。ISSA-MIFNN 与以上其他基于智能优化的 MIFNN 相比, 数据分类能力较强, 因而其泛化能力较强。具有能求解高维优化问题的 ISSA 有助于提升神经网络的分类性能。由图 3 可知, MGD-MIFNN 的搜索速度快, 收敛精度高, 但容

易致使模型过拟合; BP、ELM、RBF 及 WNN 收敛速度慢, 易陷入局部最优; ISSA-MIFNN 能有效克服梯度下降易陷入局部搜索的不足, 具有平衡局部勘测与全局开采的能力, 能快速地获得较为理想的分类效果。基于 SSA、HHO、AO 及 WOA 的 MIFNN

易陷入局部搜索, 且局部搜索速度较慢; DOA-MIFNN 的收敛速度快, 但获得的数据分类效果不理想。由此可知, ISSA-MIFNN 能有效兼顾全局与局部搜索, 得到的分类效果较好, 模型泛化能力强, 数据分类能力强。

表 3 各神经网络的训练分类准确率比较

Table 3 Comparison of training classification accuracy of neutral networks

%

神经网络	TrAcc										
	Iris	Wine	Haberman	Hayes	Immunotherapy	Cryotherapy	Heart	BreastCancer	Transfusion	Audit	Banknote
MGD-MIFNN	100	100	81.30	100	100	100	85.17	100	67.50	100	74.27
BPNN	95.24	95.16	73.36	70.65	80.65	93.55	75.42	88.89	72.08	100	100
ELM	98.10	100	74.77	64.13	85.48	96.77	84.69	90.12	74.95	100	100
RBF	95.24	95.97	73.36	64.13	82.26	80.65	68.42	82.72	77.44	46.03	91.35
WNN	95.24	99.19	78.50	77.17	96.77	98.39	83.73	88.89	67.50	100	99.48
SSA-MIFNN	100	89.52	78.50	86.96	95.16	98.39	74.64	97.53	82.22	100	70.83
HHO-MIFNN	93.33	79.30	74.30	90.22	96.77	95.16	78.95	88.89	77.44	100	56.15
AO-MIFNN	96.19	87.10	82.71	90.22	95.16	93.55	77.51	91.36	74.28	100	66.15
DOA-MIFNN	100	100	83.18	98.91	100	100	86.6	100	86.23	100	70.52
WOA-MIFNN	100	91.13	80.84	94.57	98.39	100	83.73	86.42	84.32	100	67.71
ISSA-MIFNN	99.05	98.39	86.45	97.83	100	98.39	87.08	95.06	87.95	100	73.13

表 4 各神经网络的测试分类准确率比较

Table 4 Comparison of test classification accuracy of neutral networks

%

神经网络	TeAcc										
	Iris	Wine	Haberman	Hayes	Immunotherapy	Cryotherapy	Heart	BreastCancer	Transfusion	Audit	Banknote
MGD-MIFNN	100	87.04	82.60	82.50	71.43	89.29	85.56	82.86	48.00	100	56.55
BPNN	100	92.59	79.35	42.50	75.00	85.71	83.33	71.43	76.00	100	100
ELM	100	96.30	77.17	65.00	82.14	89.29	81.11	82.86	78.22	99.14	100
RBF	100	87.04	77.17	67.50	75.00	82.14	81.11	68.57	78.67	71.12	92.96
WNN	100	88.89	71.74	52.50	42.86	82.14	76.67	62.86	70.22	100	98.54
SSA-MIFNN	100	77.78	80.43	80.00	75.00	92.86	82.22	78.57	76.00	99.57	57.04
HHO-MIFNN	100	68.52	84.87	85.00	78.57	92.86	87.78	68.57	76.89	99.14	56.80
AO-MIFNN	97.78	83.33	80.43	80.00	78.57	85.71	75.56	64.19	80.44	99.14	55.34
DOA-MIFNN	97.78	88.89	83.70	77.50	82.14	96.43	85.56	82.86	73.78	100	56.07
WOA-MIFNN	97.78	68.52	81.52	75.00	75.00	96.43	78.89	65.71	76.00	100	57.04
ISSA-MIFNN	100	96.30	85.78	90.00	89.29	100	91.11	85.71	80.89	100	58.50

3.2.2 MIFNN 的结构变化对分类性能的影响

设置 MIFNN 的 MLP 中第 1、第 2 隐含层的神经元个数构成的结构向量为 [10,5]、[15,10] 或 [20,15]。每种设置及其他参数设置不变, 以 Transfusion 和 Banknote 为例, 各神经网络获得的训练、测试分类准确率如表 5 所示。由表 3~表 5 可知, 在 MGD-MIFNN 中 MLP 的 2 个隐含层的神经元个数为 15 和 5 的前提下, 若减少神经元个数, 如 10 和 5, 则在 Transfusion 上的训练和测试分类准确率分别下降 12.24% 和 10.67%, 导致欠拟合现象; 若神经元个

数依次增加到 20 和 15, 则在 Banknote 的训练集上的分类准确率提高 0.21%, 但在测试集上的分类准确率却下降 0.97%, 且计算复杂度显著提高, 以及出现过拟合现象。另外, 随着隐含层的神经元数目逐渐增加, ISSA-MIFNN 在 Transfusion、Banknote 上的训练和测试分类准确率均无明显提升, 暗示随着神经元个数的不断增多, 将存在冗余神经元, 且导致计算成本高。因此, 合理设置 MIFNN 中 MLP 的隐含层神经元个数, 将有助于高效处理多属性数据分类问题。

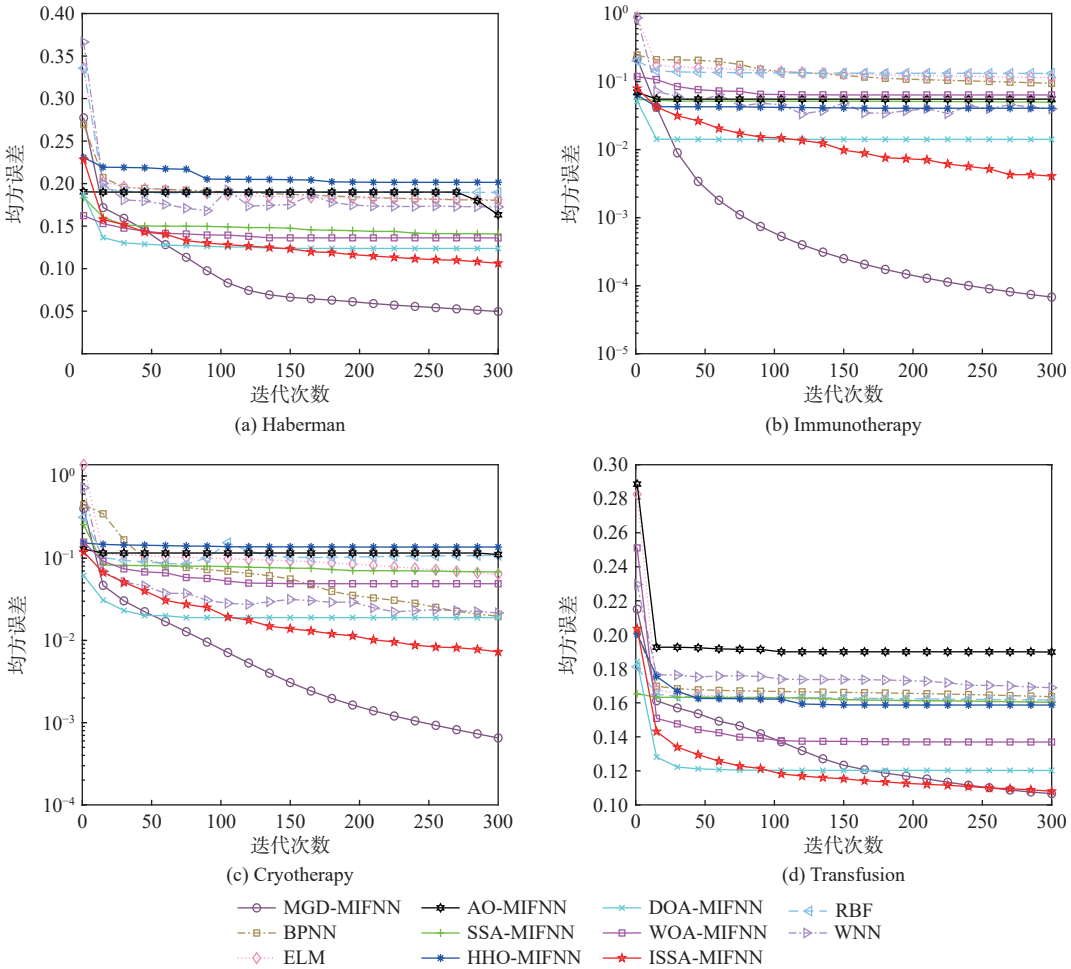


图 3 各神经网络的损失函数变化曲线

Fig. 3 Loss function variation curves of neural networks

表 5 各神经网络在不同网络结构下的分类准确率比较

Table 5 Comparison of neural networks' classification accuracy rates based on different network structures

%

神经网络	TrAcc				TeAcc			
	Transfusion		Banknote		Transfusion		Banknote	
	[10,5]	[20,15]	[10,5]	[20,15]	[10,5]	[20,15]	[10,5]	[20,15]
MGD-MIFNN	55.26	86.81	72.25	77.50	37.33	61.33	56.31	54.13
SSA-MIFNN	81.26	80.88	72.19	76.56	75.56	78.22	56.31	52.67
HHO-MIFNN	81.07	74.19	58.13	70.63	76.44	79.11	56.80	55.83
AO-MIFNN	79.35	73.80	69.17	57.40	75.56	76.89	56.55	57.28
DOA-MIFNN	80.88	87.38	73.33	73.85	76.00	74.22	56.80	57.52
WOA-MIFNN	80.31	83.37	73.75	71.04	75.56	77.33	53.16	59.47
ISSA-MIFNN	82.79	87.57	72.25	77.50	76.44	79.11	57.28	58.50

4 结 论

鉴于单输入 FNN 的应用范围窄及多层感知器对含噪声的输入数据的灵敏度高等问题, 本文通过设计能提取多源信息特征的堆叠 FNN, 并将其与多层感知器融合, 获得 MIFNN。进而, 采用含动量项的梯度下降法或 ISSA 优化网络的权值与阈值。理论分析和实验验证表明:

1) MGD-MIFNN 的搜索速度快, 收敛精度高,

在中小规模数据集上的分类准确率高;

2) ISSA 相较于参与比较的元启发式随机搜索算法, 处理高维优化问题具有明显优势, 其收敛速度快, 得到的解质量高;

3) ISSA-MIFNN 的模型泛化能力强且数据分类准确率高, 能有效回避过拟合和欠拟合。

总之, MGD-MIFNN 是一种具有明显竞争力的新型神经网络, 为数据分类问题的解决提供了又一新途径, 但对于大样本分类问题, 仍存在模型过拟

合且分类准确率不高的不足; ISSA 求解高维优化问题已有明显优势, 其与 MIFNN 结合获得的 ISSA-MIFNN 求解数据分类问题, 已展现独特优势, 但处理个别数据分类问题的能力有待提高。下一步将进一步研究 MGD-MIFNN 及 ISSA-MIFNN 处理大样本图像分类、高维曲面拟合的能力。

参考文献 (References)

[1] 杨旭华. 神经网络及其在控制中的应用研究[D]. 杭州: 浙江大学, 2004: 24-32.
YANG X H. Study on neural networks machine and its application in control[D]. Hangzhou: Zhejiang University, 2004: 24-32(in Chinese).

[2] 杨晓帆, 陈廷槐. 人工神经网络固有的优点和缺点[J]. 计算机科学, 1994, 21(2): 23-26.
YANG X F, CHEN T H. Advantages and disadvantages inherent in artificial neural networks[J]. Computer Science, 1994, 21(2): 23-26(in Chinese).

[3] 邹阿金, 沈建中. 傅立叶神经网络建模研究[J]. 湘潭大学自然科学学报, 2001, 23(2): 23-26.
ZOU A J, SHEN J Z. A neural network model structure study based on Fourier[J]. Natural Science Journal of Xiangtan University, 2001, 23(2): 23-26(in Chinese).

[4] GASHLER M S, ASHMORE S C. Modeling time series data with deep Fourier neural networks[J]. Neurocomputing, 2016, 188(5): 3-11.

[5] 党选举, 潘登. 基于电流注入法的开关磁阻电机转矩脉动抑制[J]. 组合机床与自动化加工技术, 2021(5): 115-119.
DANG X J, PAN D. Torque ripple reduction of switched reluctance motor based on current injection method[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2021(5): 115-119(in Chinese).

[6] BENRABAH M, KARA K, AITSAHED O, et al. Adaptive Fourier series neural network PID controller[J]. International Journal of Control, Automation and Systems, 2021, 19(10): 3388-3399.

[7] PENG P, XIE L, WEI H. A deep Fourier neural network for seizure prediction using convolutional neural network and ratios of spectral power[J]. International Journal of Neural Systems, 2021, 31(8): 2150022.

[8] 张雨浓, 李钧, 张智军, 等. SIMO 傅里叶三角基神经网络的权值直接确定法和结构自确定算法[J]. 信息与控制, 2011, 40(4): 507-513.
ZHANG Y N, LI J, ZHANG Z J, et al. A weights-direct-determination method and structure-automatic-determination algorithm for SIMO trigonometrically-activated Fourier neural networks[J]. Information and Control, 2011, 40(4): 507-513(in Chinese).

[9] LIN L, WU X H, QI J, et al. Power quality disturbance classification based on a novel Fourier neural network and hyperbolic S-transform[J]. International Journal of Signal Processing, Image Processing and Pattern Recognition, 2016, 9(1): 111-124.

[10] 龚鹏, 邓方, 陈杰. 基于自适应差分进化算法和傅里叶神经网络的电子罗盘的标定[C]//中国自动化学会控制理论专业委员会 B 卷. 北京: 中国自动化学会控制理论专业委员会, 2011: 741-746.

GONG P, DENG F, CHEN J. Calibration of electronic compass based on adaptive differential evolution algorithm and Fourier neural network[C]// Technical Committee on Control Theory, Chinese Association of Automation, Vol. B. Beijing: Technical Committee on Control Theory, Chinese Association of Automation, 2011: 741-746(in Chinese).

[11] 陈诗雨, 李小勇, 杜杨杨, 等. Fourier 神经网络非线性拟合性能优化研究[J]. 武汉大学学报 (工学版), 2020, 53(3): 277-282.
CHEN S Y, LI X Y, DU Y Y, et al. Optimization study of Fourier neural network nonlinear fitting performance[J]. Engineering Journal of Wuhan University, 2020, 53(3): 277-282(in Chinese).

[12] 薛建凯. 一种新型的群智能优化技术的研究与应用[D]. 上海: 东华大学, 2020.
XUE J K. Research and application of a novel swarm intelligence optimization technique[D]. Shanghai: Donghua University, 2020(in Chinese).

[13] HEIDARI A A, MIRJALILI S, FARIS H, et al. Harris hawks optimization: Algorithm and applications[J]. Future Generation Computer Systems, 2019, 97(8): 849-872.

[14] ABUALIGAH L, YOUSRI D, ABD ELAZIZ M, et al. Aquila optimizer: A novel meta-heuristic optimization algorithm[J]. Computers & Industrial Engineering, 2021, 157: 107250.

[15] BAIRWA A K, JOSHI S, SINGH D. Dingo optimizer: A nature-inspired metaheuristic approach for engineering problems[J]. Mathematical Problems in Engineering, 2021, 2021(26): 2571863.

[16] MIRJALILI S, LEWIS A. The whale optimization algorithm[J]. Advances in Engineering Software, 2016, 95: 51-67.

[17] 李爱莲, 全凌翔, 崔桂梅, 等. 融合正余弦和柯西变异的麻雀搜索算法[J]. 计算机工程与应用, 2022, 58(3): 91-99.
LI A L, QUAN L X, CUI G M, et al. Sparrow search algorithm combining sine-cosine and cauchy mutation[J]. Computer Engineering and Applications, 2022, 58(3): 91-99(in Chinese).

[18] 付华, 刘昊. 多策略融合的改进麻雀搜索算法及其应用[J]. 控制与决策, 2022, 37(1): 87-96.
FU H, LIU H. Improved sparrow search algorithm with multi-strategy integration and its application[J]. Control and Decision, 2022, 37(1): 87-96(in Chinese).

[19] YANG Q, CHEN W N, GU T, et al. A distributed swarm optimizer with adaptive communication for large-scale optimization[J]. IEEE Transactions on Cybernetics, 2020, 50(7): 3393-3408.

[20] LI D, GUO W, LERCH A, et al. An adaptive particle swarm optimizer with decoupled exploration and exploitation for large scale optimization[J]. Swarm and Evolutionary Computation, 2021, 60(7): 100789.

[21] ZHANG Z H, LI L, LU J X. Gradient-based fly immune visual recurrent neural network solving large-scale global optimization[J]. Neurocomputing, 2021, 454: 238-253.

[22] ZHANG Z H, XIAO T Y, QIN X C. Fly visual evolutionary neural network solving large-scale global optimization[J]. International Journal of Intelligent Systems, 2021, 36(11): 6680-6712.

[23] 徐辰华, 李成县, 喻昕, 等. 基于 Cat 混沌与高斯变异的改进灰狼优化算法[J]. 计算机工程与应用, 2017, 53(4): 1-9.
XU C H, LI C X, YU X, et al. Improved grey wolf optimization algorithm based on chaotic Cat mapping and Gaussian mutation[J]. Computer Engineering and Applications, 2017, 53(4): 1-9(in

Chinese).

[24] 杨彤. 改进鲸鱼算法在函数优化中的应用[D]. 西安: 西安电子科技大学, 2020: 15-17.

YANG T. Applications of improved whale optimization algorithm in function optimization[D]. Xi'an: Xidian University, 2020: 15-17(in Chinese).

Multi-input Fourier neural network and its sparrow search optimization

LI Liangliang, ZHANG Zhuhong*, ZHANG Yongdan

(Guizhou Provincial Characteristic Key Laboratory of System Optimization and Scientific Computation, College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

Abstract: In engineering applications, the back-propagation (BP) neural network often encounters many limitations due to its slow convergence and high noise sensitivity, and the reported Fourier neural networks cannot extract the features of multi-attribute input data. Hereby, this work proposes a gradient descent-based multi-input Fourier neural network after integrating the multi-layer perceptron with an overlapping Fourier neural network. Then to address the difficulty in deciding the global optimal parameter settings, the Cat chaotic map and the mechanisms of population-size adjustment and parameter adaptiveness are designed to promote the sparrow search algorithm’s ability to balance global exploration and local exploitation. An improved sparrow search algorithm is thus developed, optimizing the parameter settings and solving high dimensional function optimization problems. The theoretical analysis shows that the improved algorithm’s computational complexity is decided by its population size and the optimization problem dimension. Numerically comparative experiments have validated that the acquired Fourier neural network can effectively extract the features of multi-attribute data with strong generalization ability, and that the improved algorithm has significant advantages in coping with high dimensional function optimization problems.

Keywords: Fourier neural network; multi-layer perceptron; sparrow search; high dimensional function optimization; multi-attribute classification