

基于IDBO-SA-LSTM的冷连轧轧制力预测方法^{*}

苏九铮¹ 胡文山¹ 雷忠诚¹ 李坤杰² 刘斌斌²

(1. 武汉大学电气与自动化学院 武汉 430072; 2. 乳源东阳光优艾希杰精箔有限公司 韶关 512721)

摘要: 为了解决传统轧制力模型预测精度低和 DBO 算法易陷入局部最优解的问题,提出了一种基于改进蜣螂优化算法的结合自注意力机制的长短期记忆网络轧制力预测模型。加入黄金正弦策略和动态权重系数并引入 Circle 混沌映射得到改进蜣螂优化(IDBO)算法,通过结合长短期记忆网络(LSTM)与自注意力机制(SA),建立 IDBO-SA-LSTM 冷轧轧制力预测模型,并与其它模型进行对比。采用 6 个不同的基准函数对优化算法进行测试,仿真实验表明 IDBO 算法比麻雀搜索算法、蜣螂优化算法、灰狼搜索算法等具有更快的收敛速度和寻优精度;采用某两机架冷连轧机组 6554 次现场作业数据进行轧制力预测实验,最终结果表明 IDBO-SA-LSTM 算法预测误差指标均小于其他对比模型,预测两道次轧制力 $\pm 4\%$ 之内的命中率均大于 99%,模型精度高,泛化能力好。

关键词: 冷连轧; 轧制力预测; 改进蜣螂算法; 自注意力机制; 长短期记忆网络

中图分类号: TP183; TN06 **文献标识码:** A **国家标准学科分类代码:** 510.4

IDBO-SA-LSTM based rolling force prediction for cold continuous rolling

Su Jiuzheng¹ Hu Wenshan¹ Lei Zhongcheng¹ Li Kunjie² Liu Binbin²

(1. School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China;

2. Ruyuan Dongyangguang Fine Aluminium Foil Co., Ltd., Shaoguan 512721, China)

Abstract: In order to solve the problems of low prediction accuracy of the traditional rolling force model and the dung beetle optimizer(DBO) algorithm's tendency to fall into local optimal solutions, a rolling force prediction model based on the improved dung beetle optimisation algorithm combined with the self-attention(SA) mechanism for long and short-term memory(LSTM) networks is proposed. An improved dung beetle optimizer(IDBO) algorithm is obtained by adding the golden sine algorithm and dynamic weight coefficients and introducing the Circle chaotic mapping, and by combining the LSTM network with the SA mechanism, the IDBO-SA-LSTM cold rolling force prediction model is established, and compared with other models. Six different benchmark functions are used for testing, and simulation experiments show that IDBO algorithm has faster convergence speed and optimization accuracy than the sparrow search algorithm, the dung beetle optimization algorithm, the grey wolf search algorithm and so on. The rolling force prediction experiments are carried out using 6554 field operation data of a two-stand cold rolling unit, and the results show that the prediction error indexes of the IDBO-SA-LSTM algorithm are smaller than the other comparative models, and the IDBO-SA-LSTM algorithm can predict the rolling force within $\pm 4\%$ with the hit rate of 99%, with high model accuracy and good generalization ability.

Keywords: cold rolling; rolling force prediction; improved dung beetle algorithm; SA mechanism; long short-term memory network

0 引言

近年来,新能源产业得到了快速发展,被广泛用于锂电池生产与制造原材料的铝箔,其需求量与产量也在不断增长^[1]。冷连轧是生产铝箔的重要工序之一,在冷连轧过程

中,轧制力是控制铝箔产品厚度与板型的关键手段^[2],因此,轧制力计算模型对铝箔产品的质量起着重要影响。

在实际生产过程中,轧制力受到多种现场因素的影响,传统的机理模型往往包含较多难以确定的经验参数,难以保证模型的泛化性能与精确度。随着人工智能与优化算法

的发展,国内外有越来越多的研究开始使用深度学习与群体优化算法进行轧制力预测。文献[3]提出采用蜉蝣算法(mayfly algorithm, MA)优化支持向量机(support vector machine, SVM)建立冷连轧轧制过程的预测模型,预测误差在5%以内的概率为98.5%。文献[4]提出两种新型的粒子群优化算法,对人工神经网络的初始化权值和阈值进行优化,建立的预测模型相关系数均在0.9以上。文献[5]提出使用长短期记忆(long short term memory, LSTM)网络与相关向量机结合的方式对轧制力进行预测,模型预测准确度在3%以内。文献[6]提出使用梯度提升决策树对热连轧过程中的轧制力进行预测,误差在5%以内。从上述文献中可以看出,近年来在基于优化算法与深度学习的轧制力预测上取得了很大的进展,预测精度得到了大幅提升,但其性能仍有进一步提升的空间。

蜉蝣优化(dung beetle optimizer, DBO)算法是由Xue等^[7]于2022年提出的一种模仿蜉蝣行为的全新智能搜索算法。该算法包括滚球、繁殖、觅食和偷盗4种更新策略。该算法同时考虑了全局与局部信息,能够在保证准确率的前提下快速收敛,具有较强的全局寻优能力和搜索机制,但仍存在易陷入局部最优解和求解精度不够高的问题。为了提高DBO算法的稳定性和实际应用能力,文献[8]提出通过引入翻筋斗策略和柯西-高斯变异策略进行改进,得到一种融合多策略的蜉蝣优化(multistrategy-improved dung beetle optimizer, SDBO)算法,能够有效避免算法提前收敛。为了进一步提高DBO的全局搜索能力,文献[9]提出一种融合灰狼领袖群思想的改进蜉蝣优化算法(an enhanced dung beetle optimization algorithm, EDBO),实验表明这种融合方式能够全面提高DBO求解复杂优化问题的性能。文献[10]提出通过引入周期性突变提高DBO算法的寻优能力,对电力变压器振动信号进行精准预测。

综上所述,本文提出了一种基于改进蜉蝣优化(improved dung beetle optimizer, IDBO)算法的结合自注意力(self-attention, SA)机制与长短期记忆网络(SA-LSTM)的轧制力预测方法,构建适用于铝箔冷连轧轧制过程的高精度轧制力预测算法模型。首先,为了解决DBO算法易陷入局部最优解和求解精度不够高的问题,通过引入Circle混沌映射、黄金正弦策略(golden sine algorithm, Golden-SA)与动态权重因子来平衡其全局搜索与局部开发的能力,得到IDBO算法,并利用6种经典基准测试函数验证IDBO算法具有更好的寻优性能。然后,结合自注意力机制与长短期记忆网络建立SA-LSTM模型并利用IDBO算法对模型的隐含层神经元个数、正则化系数与初始学习率等超参数进行寻优。最后,将最优超参数代入SA-LSTM模型进行训练,并与多种其他轧制力预测模型进行对比实验,最终结果验证了IDBO-SA-LSTM模型具有更高的计算精度与泛化性能。

1 相关理论基础

1.1 轧制力计算模型

常用的冷轧轧制力计算模型包括采利柯夫公式、斯通公式、希尔公式等^[11],其中最常用的为勃兰德和福特简化后的希尔公式^[12],如式(1)所示。

$$P = Bl_c Q_p K_T \bar{K} \quad (1)$$

式中:P是轧件受到的平均轧制力,B是轧件的宽度,l_c是轧件与轧辊的接触弧长,Q_p是与摩擦系数相关的应力状态系数,K_T是与轧件张力相关的张力影响系数,bar{K}是与轧件化学性质有关的平均变形阻力。

铝箔冷连轧生产线布局如图1所示,实际生产过程中,轧制条件会随着轧件过程的变化而发生改变,每道次的轧制力不仅与轧件的张力、轧制速度、压下量等工艺参数有关,而且和轧辊的磨损、冷却、润滑等边界条件相关^[2]。在轧制力的理论计算模型中,摩擦系数与平均变形阻力均难以测量,导致实际生产中理论模型的计算精度与泛化性能往往偏低。

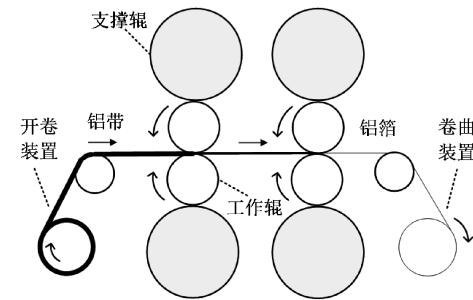


图1 铝箔冷连轧生产设备示意图

Fig. 1 Aluminium foil cold rolling production equipment

1.2 长短期记忆网络模型

为了解决传统循环神经网络(recurrent neural networks, RNN)模型反向传播过程中权重消失的问题,Yu等^[13]于1994年在传统RNN模型的基础上提出了LSTM模型,该模型能够有效处理长时间序列。随着深度学习的快速发展,LSTM及其变种模型受到了大量研究人员的关注,已经被广泛应用于语音识别、自然语言处理、时序预测、故障诊断^[14]等场景中,并在工业应用中取得良好的实际效果。LSTM模型的单元结构如图2所示,每个单元均由遗忘门、输入门和输出门以及细胞状态组成,通过精心设计的3种门控单元来增强细胞状态对信息的处理能力,能够有效保存长时间的关联信息。

其中,图2表述的LSTM模型单元结构中各个状态定义如式(2)~(7)所示。

$$\mathbf{F}_t = \sigma(\mathbf{W}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \quad (2)$$

$$\mathbf{I}_t = \sigma(\mathbf{W}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) \quad (3)$$

$$\widetilde{\mathbf{C}}_t = \tanh(\mathbf{W}_c \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c) \quad (4)$$

$$\mathbf{C}_t = \mathbf{F}_t \otimes \mathbf{C}_{t-1} + \mathbf{I}_t \otimes \widetilde{\mathbf{C}}_t \quad (5)$$

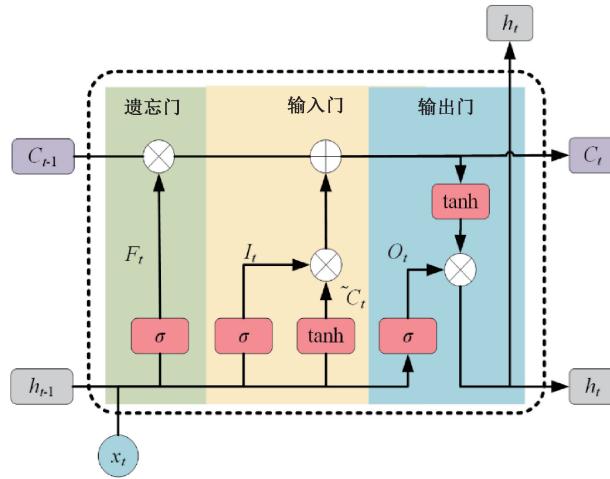


图 2 LSTM 模型单元结构

Fig. 2 Cell structure of the LSTM model

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = O_t \otimes \tanh(C_t) \quad (7)$$

式中: x_t 为 t 时刻单元结构的输入变量, h_t 为 t 时刻单元结构的隐藏状态, F_t 为 t 时刻单元结构遗忘门的计算结果, I_t 为 t 时刻单元结构输入门的计算结果, \tilde{C}_t 为 t 时刻单元结构的临时细胞状态, C_t 为 t 时刻单元结构的细胞状态, O_t 为 t 时刻单元结构输出门的计算结果, σ 为 sigmoid 激活函数, \tanh 为双曲正切激活函数, W_f, W_i, W_c 和 W_o 分别为该单元的遗忘权重、输入权重、当前状态权重和输出权重, b_f, b_i, b_c 和 b_o 分别为该单元的遗忘偏置、输入偏置、当前状态偏置和输出偏置, \otimes 代表矩阵元素逐个相乘, \oplus 代表矩阵元素逐个相加。

1.3 注意力机制

为了聚焦数据中的关键信息, Mnih 等^[15]提出向 RNN 模型中引入注意力机制进行图像分类, 取得了良好的效果。注意力机制可以高效处理大量数据, 突出重要信息, 弱化干扰信息, 常见的注意力机制包括软注意力机制、位置注意力机制、自注意力机制等^[16]。为了进一步提取输入序列内部的信息, 避免网络过拟合, 本文引入自注意力机制优化网络结构, 其原理如图 3 所示。

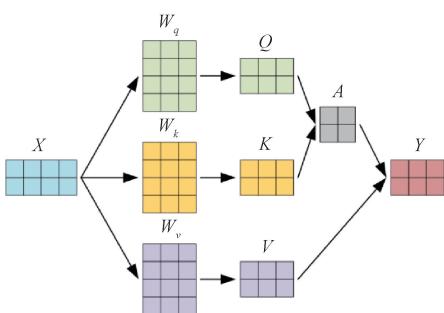


图 3 自注意力机制原理示意图

Fig. 3 Schematic diagram of the self-attention mechanism

通过进行如式(8)~(12)的计算, 即可得到输入经过自注意力层之后的结果:

$$Q = W_q X \quad (8)$$

$$K = W_k X \quad (9)$$

$$V = W_v X \quad (10)$$

$$A = \text{softmax} \left(\frac{QK^T}{\sqrt{D_k}} \right) \quad (11)$$

$$Y = AV \quad (12)$$

式中: X 为输入矩阵, Y 为输出矩阵, Q 为查询矩阵, K 为键矩阵, V 为值矩阵, A 为相关性矩阵, W_q, W_k 和 W_v 为训练得到的权重矩阵, D_k 是常量缩放因子, softmax 是归一化指数函数。

2 蜈蚣优化算法及其改进

2.1 蜈蚣优化算法

在标准 DBO 算法中, 将蜈蚣种群分为 4 种不同的行为, 主要优化流程包含以下 4 个方面:

1) 滚球蜈蚣

滚球蜈蚣在无障碍物的场景下依靠太阳保证粪球在直线路径上滚动, 在有障碍物的场景下通过跳舞对方向进行调整。当滚球时没有遇到障碍物时, 通过式(13)~(14)进行位置更新:

$$x_i^{t+1} = x_i^t + \alpha \times k \times x_i^{t-1} + b \times \Delta x \quad (13)$$

$$\Delta x = |x_i(t) - x^{worst}| \quad (14)$$

式中: t 代表迭代次数, x_i^t 代表第 i 只蜈蚣在第 t 次迭代时的位置, α 代表是否偏离上一个方向, 随机取值为 1 或者 -1, k 和 b 都是调节常数, Δx 是为了更新位置而引入的中间变量, 用于模拟光强的变化, x^{worst} 是全局的最差位置。

当滚球时遇到障碍物的话, 通过切线函数模拟蜈蚣的跳舞行为, 如式(15)所示进行位置更新:

$$x_i^{t+1} = x_i^t + \tan(\theta) |x_i^t - x_i^{t-1}| \quad (15)$$

式中: θ 表示偏转角, 取值范围为 $[0, \pi]$, 当 $\theta = 0, \pi/2$ 或 π 时会保持上一次迭代的位置。

2) 繁殖蜈蚣

繁殖蜈蚣会在粪球到达安全地区后进行产卵行为。为了模拟这一行为, 首先根据式(16)~(18)更新优化问题边界条件的上下限:

$$Lb^* = \max(X^* \times (1 - R), Lb) \quad (16)$$

$$Ub^* = \min(X^* \times (1 + R), Ub) \quad (17)$$

$$R = 1 - \frac{t}{T_{\max}} \quad (18)$$

式中: X^* 代表当前迭代中的局部最优位置, Lb^* 和 Ub^* 分别代表可进行安全繁殖地区的上界和下界, Lb 和 Ub 分别代表待优化参数的上界和下界, R 是和迭代次数有关的中间变量, T_{\max} 代表最大迭代次数, t 代表当前迭代次数。

在确定了可繁殖的安全区域后, 雌性蜈蚣会选择在此区域内进行产卵行为, 卵球的位置由式(19)进行确定:

$$\mathbf{B}_i^{t+1} = \mathbf{X}^* + \mathbf{b}_1 \times (\mathbf{B}_i^t - \mathbf{Lb}^*) + \mathbf{b}_2 \times (\mathbf{B}_i^t - \mathbf{Ub}^*) \quad (19)$$

式中: \mathbf{B}_i^t 代表第 i 个卵球在第 t 次迭代时的位置, \mathbf{b}_1 和 \mathbf{b}_2 代表两个独立随机向量。

3) 觅食蜣螂

觅食蜣螂通过寻找最佳觅食区域进行位置更新。与繁殖蜣螂类似, 需要先通过式(20)~(21)构建可觅食区域:

$$\mathbf{Lb}^b = \max(\mathbf{X}^b \times (1 - R), \mathbf{Lb}) \quad (20)$$

$$\mathbf{Ub}^b = \min(\mathbf{X}^b \times (1 + R), \mathbf{Ub}) \quad (21)$$

式中: \mathbf{X}^b 代表所有迭代过程中的全局最优位置, \mathbf{Lb}^b 和 \mathbf{Ub}^b 分别代表最佳觅食地区的上界和下界。

在最近觅食区域更新完成后, 再根据式(22)进行计算, 得到觅食蜣螂的位置:

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + C_1 \times (\mathbf{x}_i^t - \mathbf{Lb}^b) + C_2 \times (\mathbf{x}_i^t - \mathbf{Ub}^b) \quad (22)$$

式中: \mathbf{x}_i^t 代表第 i 只觅食蜣螂在第 t 次迭代时的位置, C_1 是符合正态分布的随机常数, C_2 是取值范围为 $(0, 1)$ 随机常数。

4) 偷盗蜣螂

偷盗蜣螂偷取其他蜣螂的位置信息进行更新。其中偷盗蜣螂的位置更新策略主要和全局最优解的位置强相关, 如式(23)所示。

$$\mathbf{x}_i^{t+1} = \mathbf{X}^b + S \times \mathbf{g} \times (|\mathbf{x}_i^t - \mathbf{X}^*| + |\mathbf{x}_i^t - \mathbf{X}^b|) \quad (23)$$

式中: \mathbf{x}_i^t 代表第 i 只偷盗蜣螂在第 t 次迭代时的位置, S 是固定常数值, \mathbf{g} 代表随机向量。

2.2 蜣螂优化算法的改进

利用传统 DBO 算法对网络模型中复杂优化问题进行超参数寻优时发现其在寻优过程中容易陷入局部最优解, 导致网络模型的预测精度与泛化性能无法达到最优。为了解决这一问题, 本文针对 DBO 通过 Circle 混沌映射增加初始种群的分布均匀性, 采用黄金正弦策略增强 DBO 全局搜索性能, 引入动态权重系数对全局搜索与局部搜索之间进行平衡, 得到 IDBO 算法。

1) Circle 混沌映射

原始的 DBO 算法采用随机的方式进行种群初始化, 无法保证初始种群在搜索空间中被均衡分配, 导致初始种群的多样性较低^[17]。混沌映射具有非重复性、随机性和混沌遍历性等特点, 和依赖于概率的随机生成相比能够使种群分布更加均匀。因此, 可以利用混沌映射生成初始种群来增加潜在的多样性。

为了增强 DBO 算法的稳定性与初始种群的分布均匀性, SCDBO 算法^[18]通过结合 Logistics 映射和 Tent 映射进行改进, 但初始种群均匀性会受到映射参数的较大影响。本文使用 Circle 混沌映射代替随机数针对蜣螂种群进行初始化, 能够在增强初始种群个体多样性的同时保证映射稳定性。Circle 混沌映射具体计算方式如式(24)所示。

$$\mathbf{x}_{n+1} = \text{mod} \left(\mathbf{x}_n + a - \left(\frac{b}{2\pi} \right) \sin(2\pi \mathbf{x}_n), 1 \right) \quad (24)$$

式中: \mathbf{x}_n 代表第 n 次迭代时结果的大小, a 和 b 都是常数。

Circle 混沌映射迭代 500 次的分布情况如图 4 所示, 可以看出 Circle 混沌映射在空间内的分布比较稳定且均匀, 具有较好的遍历性。

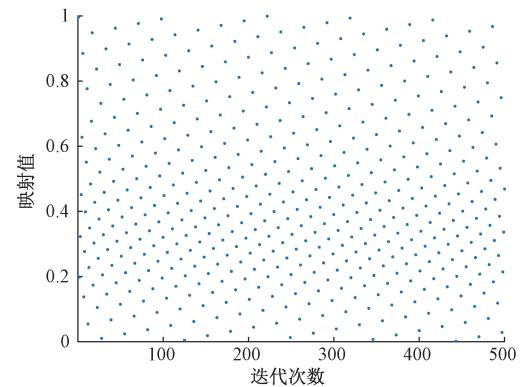


图 4 Circle 混沌映射分布情况

Fig. 4 Distribution of Circle chaotic mappings

2) 黄金正弦策略

Golden-SA 是 Tanyildizi 等^[19]于 2017 年提出的一种新型元启发式算法, 利用正弦函数进行迭代寻优, 能够加快收敛速度, 增强全局搜索能力。其更新过程如式(25)所示。

$$\mathbf{X}_i^{t+1} = \mathbf{X}_i^t \times |\sin(R_1)| + R_2 \sin(R_1) \times |\mathbf{x}_1 \times \mathbf{P}_i^t - \mathbf{x}_2 \times \mathbf{X}_i^t| \quad (25)$$

式中: \mathbf{X}_i^t 和 \mathbf{P}_i^t 分别代表第 t 次迭代时第 i 个个体的空间位置和最优位置, R_1 是代表个体移动距离的随机数, R_2 是代表个体的移动方向的随机数, x_1 和 x_2 是由黄金分割数计算得到的两个常量系数。

在滚球蜣螂的跳舞行为中, 根据当前位置周围的环境来决定下次迭代的移动方向, 但如果当前位置已经是一个局部最优点, 会导致蜣螂的跳舞动作变化极小, 容易使种群陷入局部最优解。文献[20]提出通过正弦算法扩大搜索范围, 但存在正弦策略不够先进和无法平衡局部搜索能力的问题。因此, 为了增强 DBO 算法的全局搜索能力, 本文采用 Golden-SA 扩充搜索空间, 替换蜣螂无障碍物时的跳舞更新策略。同时, 通过 R_1 和 R_2 可以有效控制跳舞蜣螂的搜索方向和搜索步长, 能够与其他行为的蜣螂形成更好的配合, 进一步平衡算法的全局搜索能力与局部搜索能力。

3) 动态权重系数

在蜣螂种群的偷盗行为中, 位置更新策略受到全局最优位置的影响比重较大。由于在迭代后期, 种群往往已经趋近最优解区间, 专注于局部搜索往往能够使得搜索结果精确度更高。为了进一步提高求解精度, 减少不必要的发散行为, 本文引入动态权重系数优化蜣螂偷盗行为, 加强迭代后期算法的局部搜索能力。优化后的偷盗蜣螂更新策略如式(26)~(28)所示。

$$\mathbf{x}_i^{t+1} = k_1 \times \mathbf{X}^b + k_2 \times \mathbf{S} \times \mathbf{g} \times (|\mathbf{x}_i^t - \mathbf{X}^*| + |\mathbf{x}_i^t - \mathbf{X}^b|) \quad (26)$$

$$k_1 = 1 - \frac{t^3}{T_{\max}^3} \quad (27)$$

$$k_2 = \frac{t^3}{T_{\max}^3} \quad (28)$$

式中: k_1 和 k_2 是由迭代次数确定的动态因子, 能够在迭代后期增强算法的局部开发能力。

将迭代次数巧妙地融入到偷盗蜣螂的行为中, 能够显著提高偷盗蜣螂的自适应能力。不仅能够在迭代初期展现出强大的全局搜索能力, 同时能够在迭代后期逐渐表现出优秀的局部优化能力。这一策略能够很好地平衡全局搜索

与局部搜索的关系, 提升算法的收敛速度和问题求解效率。

2.3 IDBO 算法性能测试

为了测试 IDBO 算法的优化性能, 本文采用基准测试函数^[21]对 IDBO 算法的性能进行测试, 并与麻雀搜索算法 (sparrow search algorithm, SSA)^[22]、灰狼优化算法 (grey wolf optimizer, GWO)^[23]、DBO 算法、SDBO 算法和 EDBO 算法进行对比实验。测试函数的基本信息如表 1 所示, 其中测试函数 $F_1 \sim F_4$ 为单峰函数, F_5 和 F_6 为多峰函数, 在运行测试函数时, 单次迭代时种群数量均设置为 30, 最大迭代次数均设置为 1 000。各搜索算法在基准函数上的寻优过程如图 5(a)~(f) 所示, 可以看出 IDBO 算法比其他五种算法具有更快的收敛速度与寻优精度。

表 1 测试函数基本信息

Table 1 Test function basic information

测试函数	搜索区间	最优解	维度
$F_1(x) = \sum_{i=1}^{30} x_i^2$	$[-100, 100]$	0	30
$F_2(x) = \sum_{i=1}^{30} x_i + \prod_{i=1}^{30} x_i $	$[-10, 10]$	0	30
$F_3(x) = \sum_{i=1}^{30} (\sum_{j=1}^i x_j)^2$	$[-100, 100]$	0	30
$F_4(x) = \max_i \{ x_i , 1 \leq i \leq 30 \}$	$[-100, 100]$	0	30
$F_5(x) = \sum_{i=1}^{30} -x_i \times \sin(\sqrt{ x_i })$	$[-500, 500]$	-12 569.5	30
$F_6(x) = -20e^{-0.2\sqrt{\sum_{i=1}^{30} x_i^2}} - e^{\frac{1}{n} \sum_{i=1}^{30} \cos 2\pi x_i} + 20 + e$	$[-32, 32]$	0	30

3 轧制力预测模型

3.1 模型建立

为了充分提取冷连轧过程中变量之间的深层耦合关系, 本文搭建如图 6 所示的深度学习模型对轧制力进行预测。

为了避免模型过拟合、提高模型运行速度, 首先使用自注意力机制对原始输入进行特征提取, 再使用 LSTM 层进一步拟合变量之间的非线性关系, 最后使用激活层、全连接层和回归层将 LSTM 层的输出结果映射到轧制力所在的维度空间, 充分拟合轧制力与工业现场参数的耦合情况。

3.2 优化流程

在 LSTM 层中, 包含的多种超参数均会对模型预测结果产生较大的影响, 其隐含层神经元个数 N 、初始学习率 a 和正则化系数 L 都是重要的超参数。隐含层神经元个数 N 决定了 LSTM 模型的容量, 如果 N 过小, 可能会导致网络无法存储足够的状态信息, 从而无法有效处理输入序列

中的非线性耦合关系, 如果 N 过大, 则会导致模型过拟合, 降低模型的精度与泛化性能。初始学习率 a 是指在训练开始时, 优化算法用来更新模型参数的初步步长, 如果 a 过小, 会导致模型训练速度过于缓慢, 可能导致优化器在局部最优解附近停滞, 如果 a 过大, 则可能导致模型无法收敛。正则化系数 L 是损失函数中增加的惩罚项系数, 用于防止模型过于复杂或过度拟合训练集数据, 如果 L 过小, 则正则化项对损失函数的影响会减弱, 容易导致模型过拟合, 如果 L 过大, 则会过分限制模型的参数拟合能力, 导致预测性能下降。手动调节参数需花费较高的人工成本。

为了进一步提高冷连轧过程中轧制力预测模型的精确度, 本文使用 IDBO 对 LSTM 模型的部分超参数进行自动寻优, 并引入 SA 机制降低网络模型的复杂度, 避免模型过度拟合, 从而提高预测精度。基于此, 搭建如图 7 所示的 IDBO-SA-LSTM 模型对轧制力进行优化与预测。在完成 IDBO-SA-LSTM 实验流程后, 使用多种其他先进优化算法进行对比实验, 并对实验的最终结果进行分析。

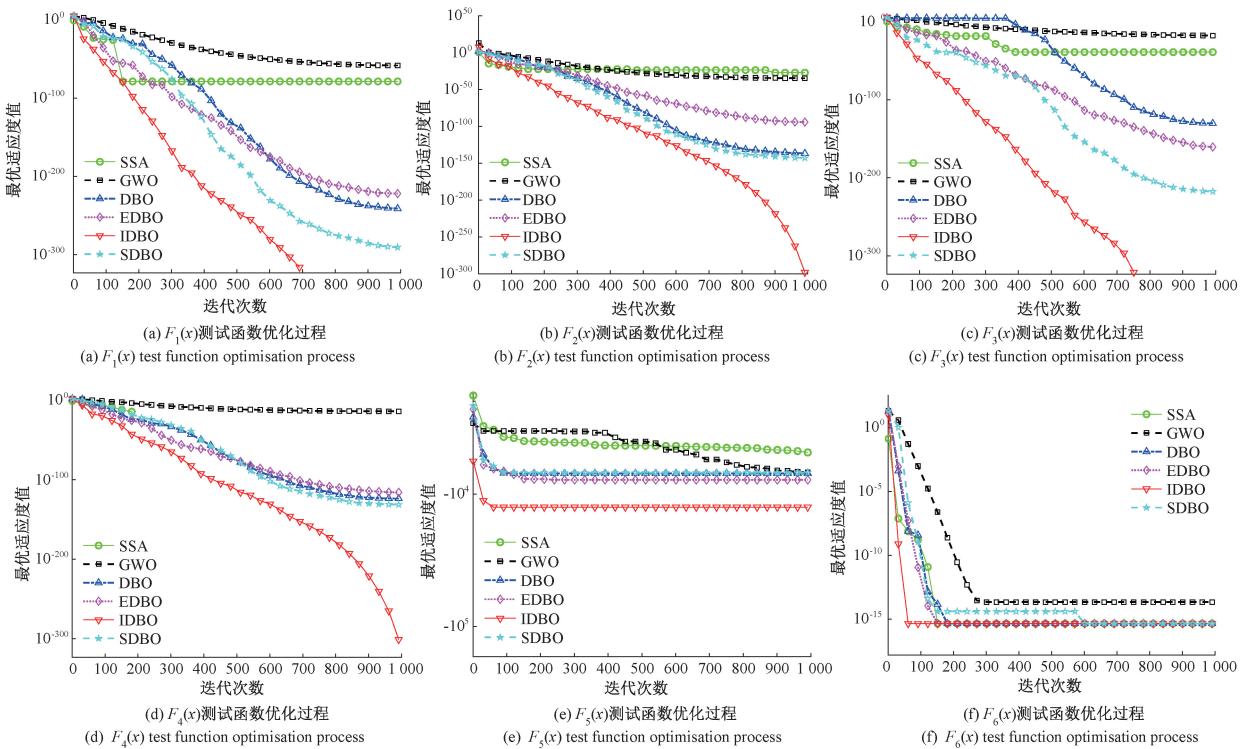


图 5 各搜索算法在不同基准函数上的优化过程曲线

Fig. 5 Optimisation process curves of each search algorithm on different benchmark functions

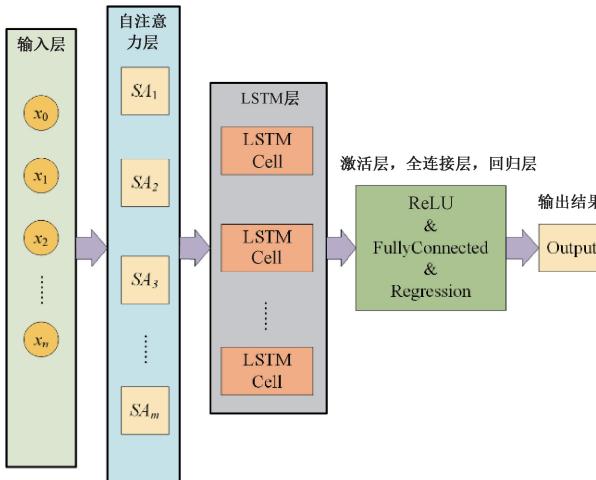


图 6 轧制力预测模型深度学习网络结构

Fig. 6 Deep learning structure of rolling force prediction model

模型预测流程的主要步骤如下:

1) 读取原始数据, 进行随机打散并归一化, 构建可训练的数据集, 并按照 8:2 的比例划分得到训练集与测试集。

2) 初始化 IDBO 算法参数, 设置种群个数为 20, 最大迭代次数为 50, 种群中滚球、繁殖、觅食和偷盗的行为比例为 2:1:1:1, 隐含层神经元个数 N 的优化范围为 [10, 100], 初始学习率 α 的优化范围为 $[1 \times 10^{-5}, 1 \times 10^{-1}]$, 正

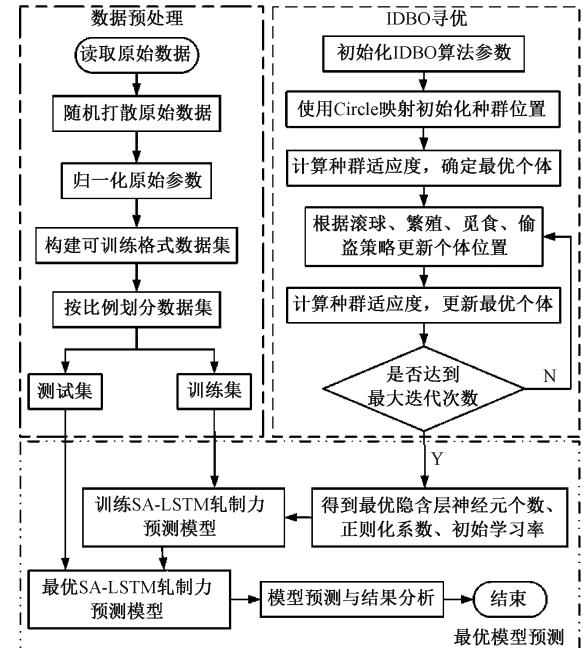


图 7 基于 IDBO-SA-LSTM 模型预测流程

Fig. 7 Flowchart of prediction based on IDBO-SA-LSTM model

则化系数 L 的优化范围为 $[1 \times 10^{-8}, 1 \times 10^{-1}]$ 。

3) 使用 Circle 混沌映射对蜣螂种群进行初始化, 每个个体代表一组可行解, 将全部个体位置代入 SA-LSTM 模型进行训练, 得到最优适应度的个体位置。

4) 利用黄金正弦策略和动态权重系数对 DBO 算法进行优化得到 IDBO 算法, 根据 IDBO 算法更新个体位置。

5) 循环迭代至最大迭代次数, 将得到的最优解参数代入 SA-LSTM 模型进行训练, 得到最优 SA-LSTM 轧制力预测模型, 使用模型进行结果预测并进行数据分析。

4 实验与分析

4.1 硬件平台设置

在操作系统与计算机硬件方面, 本实验在 Windows 10 专业版 64 位的操作系统下进行模型训练, 处理器为 Intel (R) Xeon (R) Gold 5218 CPU @ 2.30 GHz, 内存为 256 GB, 显卡为 NVIDIA TITAN RTX。在软件方面, 本模型使用 MATLAB R2024b 搭建网络结构并进行实验。

4.2 数据选择

本文通过监控某工厂 3 000 系列铝箔生产线, 采用 1 950 mm 的两机架六辊冷连轧机平稳轧制过程中的实际数据作为数据源^[24], 选取 109 个不同生产状态的平稳轧制过程进行数据采集, 得到 6 554 组原始数据, 部分原始数据如表 2 所示。本文采用出口速度、空气系统压力、工作辊弯辊力、支撑辊弯辊力、窜移位置、入口厚差、入口张力、中间张力、出口张力、绝对位置、厚度误差等变量作为输入变量, 轧制力作为输出变量。

表 2 部分冷连轧过程现场数据

Table 2 Field data for selected cold rolling processes

参数	样本 1	样本 2	样本 3
出口速度/(m·s ⁻¹)	15.833 3	11.965 5	13.830 7
空气系统压力/bar	5.551 1	6.375 1	5.689 3
工作辊弯辊力 1/%	19.270 7	-2.214 1	-9.143 0
工作辊弯辊力 2/%	18.287 3	-2.300 1	-5.295 6
支撑辊弯辊力 1/%	37.211 8	15.006 2	15.074 9
支撑辊弯辊力 2/%	26.952 1	14.876 5	16.142 7
窜移位置 1/mm	70.0	10.2	3.6
窜移位置 2/mm	48.07	12.85	7.04
入口厚差/um	-1.503 0	1.698 3	5.397 1
入口张力/N	37 071.7	52 176.9	52 938.5
中间张力/N	30 382.9	43 821.4	43 237.2
出口张力/N	12 228.4	17 260.3	17 584.1
绝对位置 1/mm	2 131.0	1 527.5	1 849.0
绝对位置 2/mm	1 809.5	1 263.0	1 605.5
厚度误差 1/um	0.697 334	-0.752 274	-0.700 386
厚度误差 2/um	1.499 96	0.099 18	0.099 18
一道次轧制力/kN	4 448.37	7 571.62	6 882.83
二道次轧制力/kN	3 341.4	5 491.8	4 500.4

4.3 归一化方法与适应度计算方法

本文采用最大最小值归一化方法将原始数据映射到

[0,1]区间, 具体计算方法如式(29)所示。

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (29)$$

式中: x 是原始数据, x_{\max} 是原始数据中的最大值, x_{\min} 是原始数据中的最小值, x_{norm} 代表原始数据经归一化处理后的值。

在使用 IDBO 寻优过程中, 本文使用均方根误差(root mean squared error, RMSE)作为适应度计算指标, 具体计算方法如式(30)所示。

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{x}_i - x_i)^2}{n}} \quad (30)$$

式中: \hat{x}_i 代表轧制力预测值, x_i 代表轧制力实际值。

4.4 实验结果

经过实验, 得到 IDBO-SA-LSTM 模型对两道次的轧制力预测情况如图 8 和 9 所示, 可以看出模型在两道次上轧制力预测误差在 4% 误差带以内的概率均大于 99%。

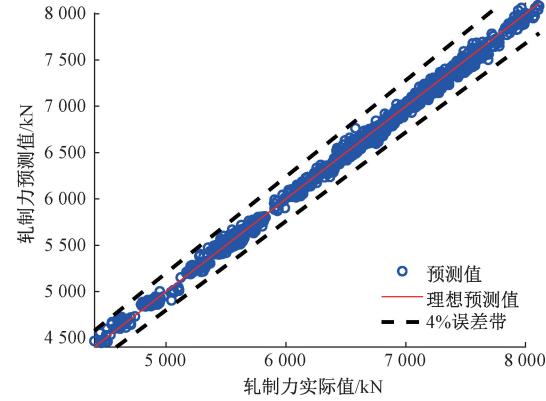


图 8 一道次轧制力预测值与实际值对比结果

Fig. 8 Comparison of rolling force result in pass one

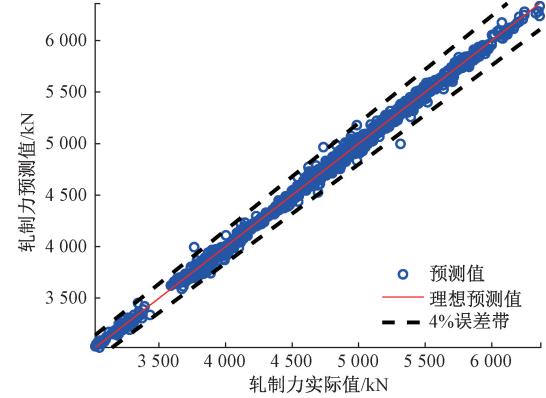


图 9 二道次轧制力预测值与实际值对比结果

Fig. 9 Comparison of rolling force result in pass two

4.5 对比实验

为了验证 IDBO-SA-LSTM 模型的有效性和优越性, 本文构建 DBO-SA-LSTM、SSA-SA-LSTM、SDBO-SA-LSTM 和 EDBO-SA-LSTM 模型进行对比实验, 计算不同

预测模型的决定系数 (R-square, R^2)、均方误差 (mean squared error, MSE)、平均绝对误差 (mean absolute error, MAE) 和均方根误差。使用不同模型对两道次的轧制力进行预测, 得到误差评价指标如表 3 和 4 所示。

表 3 各模型一道次轧制力预测误差评估

Table 3 Evaluation result for one pass of each model

模型	R^2	MSE	MAE	RMSE
IDBO-SA-LSTM	0.996 5	2 612	40.67	51.11
DBO-SA-LSTM	0.988 0	8 017	59.41	89.54
SSA-SA-LSTM	0.970 1	20 351	111.21	142.65
SDBO-SA-LSTM	0.994 8	3 578	46.01	59.82
EDBO-SA-LSTM	0.976 2	17 682	101.42	132.97

表 4 各模型二道次轧制力预测误差评估

Table 4 Evaluation result for pass two of each model

模型	R^2	MSE	MAE	RMSE
IDBO-SA-LSTM	0.996 7	1 702	30.64	41.26
DBO-SA-LSTM	0.986 9	5 886	50.90	76.72
SSA-SA-LSTM	0.991 1	4 153	48.58	64.45
SDBO-SA-LSTM	0.995 1	2 404	37.16	49.03
EDBO-SA-LSTM	0.969 2	15 396	96.15	124.08

由表 3 可知, 在一道次轧制力预测实验中, IDBO-SA-LSTM 模型与其他模型相比, R^2 平均提高 1.4%、MSE 平均下降 66.7%、MAE 平均下降 41.6%、RMSE 平均下降 45.8%。由表 4 可知, 在二道次轧制力预测实验中, IDBO-SA-LSTM 模型与其他模型相比, R^2 平均提高 1.1%、MSE 平均下降 62.1%、MAE 平均下降 40.6%、RMSE 平均下降 41.1%, 充分证明 IDBO 相对比传统 DBO 有了较大的性能改进, 同时相对于 EDBO 和 SDBO 算法的均有提升。在五个轧制力预测模型中, IDBO-SA-LSTM 模型在两道次实验中的误差评价指标均最小, 决定系数均最大, 说明相对于其他预测模型, 本文提出的轧制力预测方法误差最小。

考虑到工业现场会同时依据轧制力预测误差命中率进行模型可用性评估, 给出各模型两道次轧制力预测误差范围情况如表 5 和 6 所示, 在一道次轧制力预测实验中, IDBO-SA-LSTM 模型相对于其他 4 种对比模型, 预测误差在 $\pm 1\%$ 、 $\pm 2\%$ 、 $\pm 3\%$ 、 $\pm 4\%$ 之内的命中率分别平均提升了 55.7%、24.1%、9.5%、3.7%, 在二道次轧制力预测实验中, 预测误差在 $\pm 1\%$ 、 $\pm 2\%$ 、 $\pm 3\%$ 、 $\pm 4\%$ 之内的命中率分别平均提升了 54.5%、23.2%、9.5%、4.4%。同时, IDBO-SA-LSTM 模型在两道次轧制力预测值与实际

值的误差在 $\pm 4\%$ 之内的命中率均大于 99%, 模型精度高, 泛化能力好。

表 5 各模型一道次轧制力预测误差范围命中表

Table 5 Hitting table of the error range of pass one for each model

模型	$\pm 1\%$	$\pm 2\%$	$\pm 3\%$	$\pm 4\%$
IDBO-SA-LSTM	80.09	98.55	100.00	100.00
DBO-SA-LSTM	68.95	91.96	95.96	97.56
SSA-SA-LSTM	37.22	67.05	85.51	94.89
SDBO-SA-LSTM	75.21	97.10	99.54	99.92
EDBO-SA-LSTM	43.25	70.02	85.81	93.36

表 6 各模型二道次轧制力预测误差范围命中表

Table 6 Hitting table of the error range of pass two for each model

模型	$\pm 1\%$	$\pm 2\%$	$\pm 3\%$	$\pm 4\%$
IDBO-SA-LSTM	79.48	96.95	99.16	99.77
DBO-SA-LSTM	62.32	87.49	94.28	96.87
SSA-SA-LSTM	58.35	87.11	95.96	98.55
SDBO-SA-LSTM	71.85	93.21	98.78	99.77
EDBO-SA-LSTM	32.57	58.05	76.74	87.87

箱型图以及概率密度分布曲线能够进一步衡量模型预测误差分布情况, 各模型两道次轧制力预测误差箱型图及概率密度分布函数如图 10 和 11 所示, 在两道次轧制力预测实验中, IDBO-SA-LSTM 模型的误差分布范围最小, 且误差整体分布更加集中, 整体分布范围相较于其他模型更趋近于零。

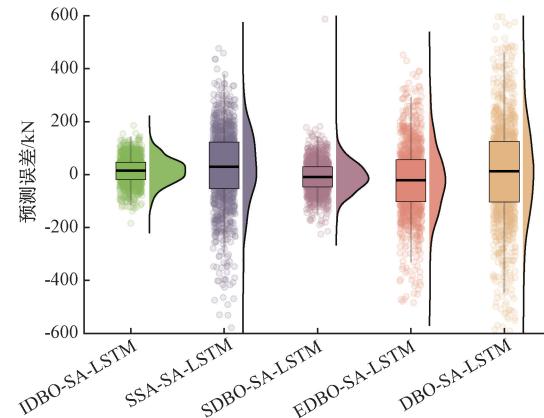


图 10 各模型一道次轧制力预测误差评估图

Fig. 10 Evaluation of the error result for pass one of each model

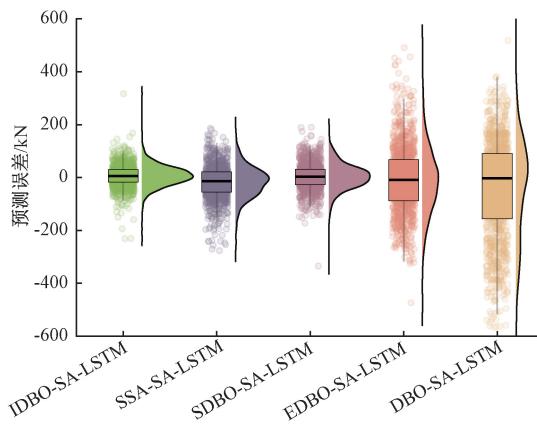


图 11 各模型二道次轧制力预测误差评估图

Fig. 11 Evaluation of the error result for pass two of each model

5 结 论

轧制力是铝箔冷连轧生产过程中的重要参数,为了解决传统机理模型预测精度低、泛化性能差的问题,本文提出了 IDBO-SA-LSTM 轧制力预测模型。

针对传统 DBO 算法种群初始化稳定性较差、局部搜索与全局寻优能力不均衡的问题,本文通过加入黄金正弦策略和动态权重系数并引入 Circle 混沌映射得到 IDBO 算法,能够增强滚球蜣螂跳舞行为的全局搜索能力和偷盗蜣螂在迭代后期局部搜索能力,提高种群初始化时的稳定性。通过基准函数仿真实验对比发现 IDBO 具有更快的搜索速度和更高的寻优精度。结合自注意力机制与 LSTM 构建深度学习网络模型对轧制力进行预测,考虑到 LSTM 网络中隐含层神经元个数 N 、初始学习率 a 和正则化系数 L 对模型效果的影响,使用 IDBO 优化网络模型参数,建立 IDBO-SA-LSTM 模型,实现了轧制力的高精度预测。使用 DBO-SA-LSTM、SSA-SA-LSTM、EDBO-SA-LSTM 和 SDBO-SA-LSTM 模型进行对比实验,最终实验结果表明 IDBO-SA-LSTM 模型效果均优于上述模型,在轧制力预测上有较高的实用性与可行性。

参考文献

- [1] 吴燕翔. 国内电池铝箔的供求格局和发展建议[J]. 有色金属加工, 2024, 53(3): 4-6.
- [2] 葛林恒. 可逆冷轧机轧制力智能预测[D]. 西安: 西安理工大学, 2019.
- [3] 陈树宗, 侯佳琦, 白芸松, 等. 基于 MA-SVM 算法的冷轧轧制力预测模型[J]. 燕山大学学报, 2023, 47(5): 428-432, 440.
- [4] CHEN SH Z, HOU J Q, BAI Y S, et al. Rolling force prediction model for cold rolling based on MA-SVM[J]. Journal of Yanshan University, 2023, 47(5): 428-432, 440.
- [5] 夏军勇, 卢奇, 张子健, 等. 基于混合人工神经网络的冷连轧水平力预测[J]. 锻压技术, 2024, 49(3): 86-93.
- [6] XIA J Y, LU Q, ZHANG Z J, et al. Prediction on horizontal force in cold continuous rolling based on hybrid artificial neural network [J]. Forging & Stamping Technology, 2024, 49(3): 86-93.
- [7] 孙浩, 赵明达, 李静, 等. 基于 LSTM-JITRVM 的冷轧轧制力建模方法研究[J]. 计量学报, 2023, 44(9): 1409-1416.
- [8] SUN H, ZHAO M D, LI J, et al. Research on method of cold rolling force based on LSTM-JITRVM[J]. China Journal of Econometrics, 2023, 44(9): 1409-1416.
- [9] WANG Q, SONG L, ZHAO J, et al. Application of the gradient boosting decision tree in the online prediction of rolling force in hot rolling [J]. The International Journal of Advanced Manufacturing Technology, 2023, 125(1): 387-397.
- [10] XUE J, SHEN B. Dung beetle optimizer: A new meta-heuristic algorithm for global optimization[J]. The Journal of Supercomputing, 2023, 79 (7): 7305-7336.
- [11] 孙仟硕, 王英博. 融合多策略的改进蜣螂优化算法及其应用[J]. 信息与控制, 2024, 53(5): 631-641, 651.
- [12] SUN Q SH, WANG Y B. Improved dung beetle optimization algorithm with multi-strategy and its application [J]. Information and Control, 2024, 53(5): 631-641, 651.
- [13] 刘文强, 李涛. 基于改进蜣螂优化算法的无人机航迹规划[J]. 电子测量技术, 2024, 47(15): 64-72.
- [14] LIU W Q, LI T. UAV trajectory planning based on enhanced dung beetle optimization algorithm [J]. Electronic Measurement Technology, 2024, 47(15): 64-72.
- [15] 周亚中, 何怡刚, 邢致恺, 等. 基于 IDBO-ARIMA 的电力变压器振动信号预测[J]. 电子测量与仪器学报, 2023, 37(8): 11-20.
- [16] ZHOU Y ZH, HE Y G, XING ZH K, et al. Power transformer vibration signal prediction based on IDBO-ARIMA[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(8): 11-20.
- [17] KALKER J J. Three-dimensional elastic bodies in rolling contact [M]. Springer Science & Business

- Media, 2013.
- [12] TAO L, WANG Q, WU H. Establishment and numerical analysis of rolling force model based on dynamic roll gap [J]. Applied Sciences, 2023, 13(13): 7394.
- [13] YU Y, SI X, HU C, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural computation, 2019, 31(7): 1235-1270.
- [14] VAN-HOUDT G, MOSQUERA C, NAPOLES G. A review on the long short-term memory model [J]. Artificial Intelligence Review, 2020, 53 (8): 5929-5955.
- [15] MNIH V, HEESS N, GRAVES A. Recurrent models of visual attention[J]. Advances in Neural Information Processing Systems, 2014, 27, DOI:10.48550/arXiv.1406.6247.
- [16] 朱张莉, 饶元, 吴渊, 等. 注意力机制在深度学习中的研究进展[J]. 中文信息学报, 2019, 33(6):1-11.
ZHU ZH L, RAO Y, WU Y, et al. Research progress of attention mechanism in deep learning[J]. Journal of Chinese Information Processing, 2019, 33(6):1-11.
- [17] LIU J, LYU Z, ZHAO L. A dual-optimization building energy prediction framework based on improved dung beetle algorithm, variational mode decomposition and deep learning [J]. Energy and Buildings, 2025, 328: 115143.
- [18] 刘微, 任腾腾, 韩广雨, 等. 多策略改进蜣螂优化算法及其应用 [J]. 电子测量技术, 2024, 47 (12): 109-121.
LIU W, REN T T, HAN G Y, et al. Multi strategy improvement of dung beetle optimization algorithm and its application [J]. Electronic Measurement Technology, 2024, 47(12): 109-121.
- [19] TANYILDIZI E, DEMIR G. Golden sine algorithm: A novel math-inspired algorithm [J]. Advances in Electrical & Computer Engineering, 2017, 17 (2): 1-78.
- [20] 潘劲成, 李少波, 周鹏, 等. 改进正弦算法引导的蜣螂优化算法[J]. 计算机工程与应用, 2023, 59 (22): 92-110.
PAN J CH, LI SH B, ZHOU P, et al. Dung beetle optimization algorithm guided by improved sine algorithm [J]. Computer Engineering and Applications, 2023, 59(22): 92-110.
- [21] JAMIL M, YANG X S. A literature survey of benchmark functions for global optimisation problems [J]. International Journal of Mathematical Modelling and Numerical Optimisation, 2013, 4(2): 150-194.
- [22] XUE J, SHEN B. A novel swarm intelligence optimization approach: Sparrow search algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.
- [23] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69: 46-61.
- [24] 苏九铮, 胡文山, 刘国平, 等. 融合云边协同的分布式铝箔厂设备监控管理系统[J]. 现代制造工程, 2024(9): 119-126.
SU J ZH, HU W SH, LIU G P, et al. Cloud-edge collaboration integrated distributed aluminum foil plant equipment monitoring management system [J]. Modern Manufacturing Engineering, 2024 (9): 119-126.

作者简介

苏九铮,硕士研究生,主要研究方向为智能优化算法、工业互联网。

E-mail:jiuzheng.su@whu.edu.cn

胡文山(通信作者),博士,教授,博士研究生导师,主要研究方向为网络控制系统、远程网络实验室和智能电网。

E-mail:wenshan.hu@whu.edu.cn

雷忠诚,博士,副研究员,硕士研究生导师,主要研究方向为网络化控制系统、远程和虚拟实验系统、数字孪生控制系统和电力系统。

E-mail:zhongcheng.lei@whu.edu.cn

李坤杰,博士,高级工程师,主要研究方向为企业信息化、智能制造。

E-mail:likunjie@hotmail.com

刘斌斌,本科,工程师,主要研究方向为铝箔轧制技术。

E-mail:canglong1079@163.com