

# 基于 IMODA 自适应深度信念网络的复杂 模拟电路故障诊断方法



巩彬<sup>1</sup>, 安爱民<sup>1,2</sup>, 石耀科<sup>3</sup>, 杜先君<sup>1,2\*</sup>

(1. 兰州理工大学 电气工程与信息工程学院, 兰州 730050;

2. 兰州理工大学 甘肃省工业过程先进控制重点实验室, 兰州 730050;

3. 兰州理工大学 计算机与通信学院, 兰州 730050)

**摘要** 针对传统 DBN 在无监督训练过程中预训练耗时久、诊断精度差等问题, 提出了一种基于改进多目标蜻蜓优化自适应深度信念网络 (IMODA-ADBN) 的模拟电路故障诊断方法。首先, 根据参数更新方向的异同提出了自适应学习率, 提高网络收敛速度; 其次, 传统 DBN 在有监督调优过程利用 BP 算法, 然而 BP 算法存在易陷入局部最优的问题, 为了改善该问题, 利用改进的 MODA 算法取代 BP 算法提高网络分类精度。在 IMODA 算法中, 添加 Logistic 混沌映射和基于对立跳跃以获得帕累托最优解, 增加算法的多样性, 提高算法的性能。在 7 个多目标数学基准问题上测试该算法, 并与 3 种元启发式优化算法 (MODA、MOPSO 和 NSGA-II) 进行比较, 证明了 IMODA-ADBN 网络模型具有稳定性。最后将 IMODA-ADBN 运用到二级四运放双二阶低通滤波器的诊断实验中, 实验结果表明该方法在收敛速度快的基础上保证了分类精度, 诊断率更高, 能够实现高难故障的分类与定位。

**关键词** 模拟电路; MODA 算法; 自适应学习率; 深度信念网络; 故障诊断

中图分类号 TH707

文献标志码 A

DOI 10.12178/1001-0548.2023047

## A Fault Diagnosis Method for Complex Analog Circuits Based on IMODA Adaptive Deep Belief Network

GONG Bin<sup>1</sup>, AN Aimin<sup>1,2</sup>, SHI Yaoke<sup>3</sup>, and DU Xianjun<sup>1,2\*</sup>

(1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;

2. Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou University of Technology, Lanzhou 730050, China;

3. School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

**Abstract** Aiming at the problems of time-consuming pre-training and poor diagnostic accuracy in the process of unsupervised training of traditional Deep Belief Network (DBN), In this paper, an Improved Multi-Objective Dragonfly Optimization Adaptive Deep Belief Network (IMOD-ADBN) is proposed for analog circuit fault diagnosis. Firstly, an adaptive learning rate is proposed according to the similarities and differences of parameter update directions to improve the convergence speed of the network. Secondly, traditional DBN uses Back Propagation (BP) algorithm in the supervised tuning process. However, BP algorithm has the problem that it is easy to fall into local optimum. In order to improve the problem, IMOD algorithm is used to replace BP algorithm to improve the accuracy of network classification. In the improved MODA algorithm, Logistic chaotic imprinting and oppositional jumping are added to obtain the Pareto optimal solution, which increases the diversity of the algorithm and improves its performance of the algorithm. The proposed algorithm is tested on eight multi-objective mathematical benchmark problems and compared with three meta-heuristic optimization algorithms (MODA, MOPSO, and NSGA-II), and the stability of IMOD-ADBN network model is proved. Finally, IMOD-ADBN is applied to the diagnosis experiment of a two-stage four-op-amplifier double-second-order low-pass filter. The experimental results show that the proposed IMOD-ADBN can ensure classification accuracy on the basis of fast convergence, and IMOD-ADBN has higher diagnosis rate than other methods mentioned in this paper, which can realize the classification and location of high-difficulty faults.

**Key words** analog circuit; MODA algorithm; adaptive learning rate; deep belief network; fault diagnosis

收稿日期: 2023-02-21; 修回日期: 2023-12-15

基金项目: 国家自然科学基金 (62241307, 61963025); 甘肃省科技计划 (22YF7FA166, 22YF7GA164); 甘肃省自然科学基金优秀博士生项目 (23JRRA809); 甘肃省教育厅高等学校创新基金 (2021A-027); 兰州市科技计划 (2022-RC-60)

作者简介: 巩彬, 博士生, 主要从事光伏阵列、模拟电路故障诊断方面的研究。

\*通信作者 E-mail: xdu@lut.edu.cn

模拟电路作为军事、国防、航空等领域的重要支柱<sup>[1]</sup>,其准确且高效的故障诊断已成为电路测试领域中的研究热点<sup>[2]</sup>,开发高效且精确的诊断策略成为模拟电路故障诊断领域的迫切需求<sup>[3-6]</sup>。然而由于模拟电路复杂的故障现象及多重并发故障,导致了模拟电路的故障状态是无限的,故障特征可以是连续的,同时模拟电路的输入-输出关系往往呈非线性映射关系,且电路中还存在非线性元件,使得实际应用中很难建立电路响应的精确数学模型。同时受各类因素影响,模拟电路非故障元件的实际参数值会在标称值上下随机波动,加之元件的非线性表征误差、测试误差等,使得传统的模拟电路故障诊断方法如故障字典法、参数识别法<sup>[7-8]</sup>等在实际生活中难以发挥重要作用。随着故障诊断技术的深入研究,小波变换等信号处理方法在非线性和系统故障特征提取与诊断中得到广泛应用<sup>[9-11]</sup>,但信号处理方法在特征提取的过程中容易忽略本质特征,导致非线性模拟电路故障诊断的效率和准确率过低<sup>[12-14]</sup>。

为解决上述问题,支持向量机(Support Vector Machine, SVM)、神经网络(Neural Networks, NN)<sup>[15]</sup>等数据驱动的人工智能方法广泛进入研究领域,为模拟电路故障诊断提供了可行的技术支持。文献[15]提出了一种基于控制系统脉冲响应特性的常规时域特征向量,利用改进后的向量可以提高LSSVM的分类性能。文献[16]利用马氏距离(Mahalanobis Distance, MD)的粒子群算法(Particle Swarm Optimization, PSO)对分类器进行优化,合理选择特征向量从而提高分类精度。文献[17]提出了一种基于双链量子遗传算法(Double Chains Quantum Genetic Algorithm, DCQGA)的诊断模型,DCQGA-SVM的总体最佳适应度和分类精度都有所提高。文献[18]提出了一种基于极限学习机的自动编码器进行故障诊断,可有效提高诊断准确率。文献[19]将圆模型和极限学习机进行结合,提出了一种线性模拟电路的故障诊断方法,该方法简化和加速了故障诊断的过程,为故障诊断领域提供了一种新方法。文献[20]指出ELM与SVM相比具有较少的优化约束、更好的泛化性能。然而,SVM、神经网络等分类算法的输出仅仅是故障样本对应的故障类别标签,无法输出其他诊断信息,如故障样本属于各故障类别的概率。在高维特征样本的分类器设计中,故障特征的维数约简与分类器建模是分离进行的,存在约简后输入到分类器

的特征与该分类器可能不是最佳匹配的问题,难以表征被测信号与装备健康状况之间复杂的映射关系,且面临维数灾难等问题。文献[21]指出,利用浅层学习模型无法全面彻底地揭示故障根源和信号特征之间复杂的内在关系。

近年来,卷积神经网络(Convolutional Neural Network, CNN)<sup>[22]</sup>和深度置信网络(Deep Belief Network, DBN)<sup>[23]</sup>以其强大的学习和特征提取能力在故障诊断领域中广泛应用<sup>[24-25]</sup>。文献[26]针对局部信号较弱的情况,先对信号进行变分模态分解(Variational Mode Decomposition, VMD)和奇异值分解(Singular-Value Decomposition, SVD)构造奇异值向量矩阵,然后输入到卷积神经网络可以对故障精准分类。文献[27]构造了由错位层、卷积层、子采样层和全连接层组成的错位时间序列卷积神经网络(Dislocated Time Series Convolutional Neural Network, DTS-CNN)。实验表明DTS-CNN更适用于工业信号的故障诊断。文献[28]利用DBN网络自适应提取信号的特征并自动分类,可以灵活高效地对模拟电路的故障进行诊断,为不同的诊断问题提供新的解决思路。文献[29]提出高斯-伯努利深度置信网络(Gauss-Bernoulli Deep Confidence Network, GB-DBN)可以更有效地捕捉原始输出信号中的高阶语义特征,使诊断结果更加精确。文献[30]对模拟电路中的间歇性故障,提出了一种基于总体平均经验模式分解(Ensemble Empirical Mode Decomposition, EEMD)和DBN的模拟电路间歇性故障诊断方法。实验结果表明,该方法在诊断方面的高效性要优于传统DBN。文献[31]利用小波包变换(Wavelet Packet Transform, WPT)与分层诊断网络(Hierarchical Diagnosis Network, HDN)相结合的模型,其中HDN由两层DBN组成,分别用来提取特征和进行故障诊断。文献[32]利用混沌粒子群优化算法优化受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)中的学习率,进一步提升特征提取的性能。文献[33]利用多个两层稀疏自动编码器(Sparse Autoencoder, SAE)对时域特征和频域特征进行融合,最后,利用DBN进行进一步分类。

尽管DBN在多个领域展现了广泛的应用前景,然而由于其正向训练阶段所需的时间较长,这使得DBN在应对信息爆炸式增长的大数据时代中显得不太适用。针对该问题,本文利用参数更新方向的异同动态选择学习率,加速模型收敛并提高模

型的稳定性; 其次, 传统 DBN 利用 BP 算法反向精调权重, 然而 BP 算法存在容易陷入局部最小的问题, 因此 DBN 的性能也会受此影响。为了获得一种具有强鲁棒性且同时能够避开局部极小点的 DBN 模型, 利用 IMODA 算法进行精调权重, 并验证其收敛性及稳定性。

## 1 DBN 原理

### 1.1 DBN 结构

DBN 是由多个受限玻尔兹曼机堆叠组成的概率生成模型<sup>[23]</sup>。RBM 作为两层网络, 由可视层和隐藏层双向连接, 并且同一层网络的神经元相互独立。可视层主要的作用是训练数据的输入, 而隐藏层则用于提取特征。RBM 的结构图如图 1 所示。其中,  $w$  表示连接权重,  $b$  为隐藏层的偏倚系数,  $a$  为可视层的偏倚系数。

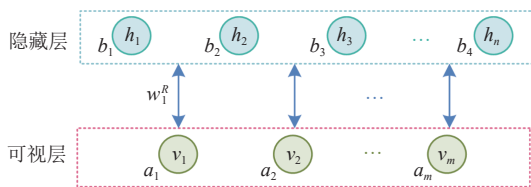


图 1 RBM 结构图

DBN 特征提取的过程分为两个阶段: 预训练阶段和微调阶段。在预训练阶段中, 先将所有的 RBM 进行逐层无监督预训练, 从而形成一个无监督学习的特征模型。再用有监督算法进行反向训练, 将所有的 RBM 初始连接权重进行微调, 从而减少训练产生的误差, 有利于 DBN 提取输入数据的本质特征, 结构如图 2 所示。

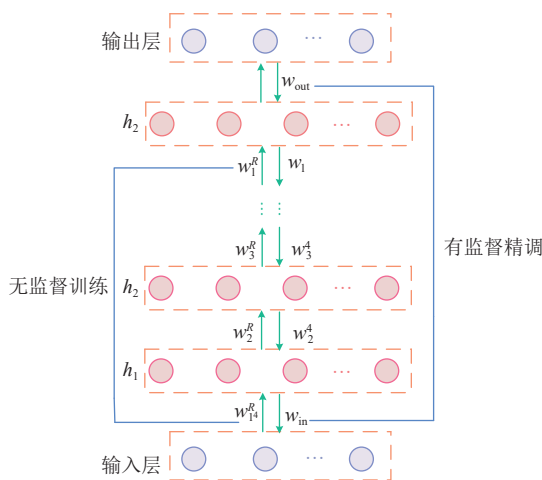


图 2 DBN 结构图

### 1.2 无监督学习

为了确定网络的初始权值, Hinton 采用无监督训练方法学习参数。1 个 RBM 中包括一个可视层和 1 个隐藏层, 分别用  $v$  和  $h$  来表示。

给定模型参数  $\theta = \{w^R, a, b\}$ , 可视层和隐藏层的联合概率分布  $P(v, h; \theta)$  用能量函数  $E(v, h; \theta)$  定义为:

$$P(v, h; \theta) = \frac{1}{Z} e^{-E(v, h; \theta)} \quad (1)$$

式中,  $Z = \sum_{v, h} e^{-E(v, h; \theta)}$  是归一因子, 模型关于  $v$  的边缘分布:

$$P(v; \theta) = \frac{1}{Z} \sum_h e^{-E(v, h; \theta)} \quad (2)$$

对于一个伯努利 (可视层) 分布-伯努利 (隐藏层) 分布的 RBM, 单元联合配置的能量函数定义为:

$$E(v, h) = - \sum_{i=1}^m \sum_{j=1}^n v_i w_{ij}^R h_j - \sum_{i=1}^m a_i v_i - \sum_{j=1}^n b_j h_j \quad (3)$$

式中,  $w_{ij}^R$  是 RBM 的连接权重;  $a_i$  和  $b_j$  分别是可见层单元和隐藏层单元的偏置。

$$P(h_j = 1/v; \theta) = \sigma \left( b_j + \sum_{i=1}^m v_i w_{ij}^R \right) \quad (4)$$

$$P(v_i = 1/h; \theta) = \sigma \left( a_i + \sum_{j=1}^n w_{ij}^R h_j \right) \quad (5)$$

式中,  $\sigma$  为激活函数。

可视层和隐藏层的概率取值标准通常是设定阈值来实现, 这是因为可视层和隐藏层是伯努利二值状态。以隐藏层为例可表示为:

$$h_j = \begin{cases} 0 & p(h_j = 1/v) < \delta \\ 1 & p(h_j = 1/v) > \delta \end{cases} \quad (6)$$

式中,  $\delta$  为 0.5~1 之间的常数。

计算对数似然函数  $\lg P(v; \theta)$  的梯度, 可以得到 RBM 权值更新公式为:

$$w_{ij}^R = w_{ij}^R + \eta \Delta w_{ij}^R \quad (7)$$

$$\Delta w_{ij}^R = E_{\text{data}}(v_i h_j) - E_{\text{model}}(v_i h_j) \quad (8)$$

式中,  $\eta$  表示学习率;  $E_{\text{data}}(v_i h_j)$  是训练集中观测的数据期望;  $E_{\text{model}}(v_i h_j)$  是在模型所确定的分布上的期望, 可由吉布斯近似得到。

### 1.3 有监督学习

有监督学习是对无监督学习得到的权重  $w^R$  进行精调。根据图 2 的输出层和最后一层隐藏层为

例, 设  $F$  为模型的期望输出, 定义交叉熵函数为损失函数:

$$F = -\frac{1}{n} \sum_i [y_i \ln y'_i + (1 - y_i) \ln(1 - y'_i)] \quad (9)$$

式中,  $y_i$  表示目标输出经过 softmax 后的输出;  $y'_i$  表示期望输出经过 softmax 后的输出;  $n$  表示分类种数。

权重更新公式则可表示为:

$$w_{\text{out}}(\tau + 1) = w_{\text{out}}(\tau) - \eta \frac{\partial F(\tau)}{\partial w_{\text{out}}(\tau)} \quad (10)$$

利用这种方法从顶层输出层到底层输入层依次微调可以得到整个 DBN 网络的权重  $w = (w_{\text{out}}, w_1, w_{l-1}, \dots, w_2, w_{\text{in}})$ 。

## 2 改进的 IMODA-ADBN 学习算法

### 2.1 自适应 CD 算法

在 DBN 无监督学习过程中, 吉布斯采样作为 CD 算法的核心, 是一种马尔科夫链蒙特卡罗 (Markov Chain Monte Carlo, MCMC) 算法。在直接采样联合分布困难的情况下, 用于生成某特定多参数的概率分布的一组近似观测值。吉布斯采样主要由 3 个步骤组成。

1) 利用样本  $V$  初始化吉布斯链, 从而得到可视层输入  $v^{(0)}$ 。

2) 根据式 (4) ~ 式 (6) 分别进行采样, 其中  $h^{(t)}$  由  $P(h^{(t)}/v^{(t)}; \theta)$  采样得到,  $v^{(t+1)}$  则由  $P(v^{(t+1)}/h^{(t)}; \theta)$  采样得到,  $t$  为采样步数。

3) 重复步骤 2)。

由于每个 RBM 中需要多次迭代计算, 固定学习率  $\eta$  容易产生收敛困难的现象。因此采用自适应学习率根据参数的更新方向来确定学习率  $\eta$ 。自适应学习率的原理是连续两次迭代后的参数更新方向相同时学习率会增大; 连续两次迭代后的参数更新方向相反学习率会减小。自适应学习率  $\eta$  的更新机制如下:

$$\eta = \begin{cases} D\eta \frac{\left| (\Delta w_{ij}^R)^{(t)} \right| + \left| (\Delta w_{ij}^R)^{(t+1)} \right|}{(\Delta w_{ij}^R)^{(t)} + (\Delta w_{ij}^R)^{(t+1)} + \varepsilon} = 1 \\ d\eta \frac{\left| (\Delta w_{ij}^R)^{(t)} \right| + \left| (\Delta w_{ij}^R)^{(t+1)} \right|}{(\Delta w_{ij}^R)^{(t)} + (\Delta w_{ij}^R)^{(t+1)} + \varepsilon} \neq 1 \end{cases} \quad (11)$$

$$(\Delta w_{ij}^R)^{(t)} = v_i^{(t)} h_i^{(t)} - v_i^{(t+1)} h_i^{(t+1)} \quad (12)$$

$$(\Delta w_{ij}^R)^{(t+1)} = v_i^{(t+1)} h_i^{(t+1)} - v_i^{(t+2)} h_i^{(t+2)} \quad (13)$$

式中,  $D=1.4$ ,  $d=0.7$ ,  $\varepsilon$  的主要目的是避免分母为 0 的情况产生, 设置为  $10^{-6}$ 。

### 2.2 基于改进多目标蜻蜓优化的有监督微调

昆虫群的行为主要有以下 3 种<sup>[34]</sup>。

- 1) 分离, 指个体与相邻个体之间避免静态碰撞。
- 2) 对齐, 表示个体与相邻个体的速度匹配。
- 3) 凝聚, 指个体倾向聚集到周围群体的中心。

改进多目标蜻蜓优化算法 (Improve Multi-Objective Dragonfly Algorithm, IMODA) 3 个主要特征如下: 1) 存档集成到蜻蜓搜索中, 用来保存和检索帕累托最优解; 2) 利用拥挤距离和轮盘选择有效地管理档案种群, 其中包含在空间搜索探索过程中的最优非支配解; 3) 为缓解局部最优, 在 MODA 中添加精英个体和基于对立的跳跃。

#### 2.2.1 运动模型

昆虫群的主要行为一个是被食物吸引, 另一个则是躲避天敌。所以蜻蜓个体的位置更新主要受 5 种因素影响, 5 种因素的数学模型如下。

1) 分离

静态群中, 蜻蜓为了捕食蝴蝶和蚊子等其他飞行猎物会簇拥成群。为了避免在飞行过程中与其他蜻蜓产生碰撞, 其分离表达式如下:

$$S_i = - \sum_{j=1}^N X - X_j \quad (14)$$

式中,  $X$  表示当前个体的位置;  $X_j$  表示第  $j$  个相邻个体的位置;  $N$  表示相邻个体的数量。

2) 对齐

$$A_i = \sum_{j=1}^N V_j / N \quad (15)$$

式中,  $V_j$  表示第  $j$  个相邻个体的速度。

3) 凝聚

$$C_i = \frac{\sum_{j=1}^N X_j}{N} - X \quad (16)$$

式中,  $X$  表示当前个体的位置;  $X_j$  表示第  $j$  个相邻个体的位置;  $N$  表示相邻个体的数量。

4) 食物吸引

$$F_i = X^+ - X \quad (17)$$

式中,  $X$  表示当前个体位置;  $X^+$  表示食物来源位置。

## 5) 天敌驱散

$$E_i = X^- + X \quad (18)$$

式中,  $X$  表示当前个体的位置;  $X^-$  表示天敌的位置。

本文假设蜻蜓的行为是这 5 种纠正模式的组合。为了更新人工蜻蜓在搜索空间中的位置并模拟其运动, 需要考虑两个向量: 步长 ( $\Delta X$ ) 和位置 ( $X$ ), 定义如下:

$$\Delta X_{t+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + \omega \Delta X_t \quad (19)$$

$$X_{t+1} = X_t + \Delta X_{t+1} \quad (20)$$

式中,  $s$  表示分离权重;  $S_i$  表示第  $i$  个个体的分离;  $a$  表示对齐权重;  $A_i$  表示第  $i$  个个体的对齐;  $c$  表示内聚权重;  $C_i$  表示第  $i$  个个体的内聚;  $f$  表示食物因子;  $F_i$  表示第  $i$  个个体的食物来源;  $e$  表示敌人因子;  $E_i$  表示第  $i$  个个体敌人的位置;  $\omega$  表示惯性权重;  $t$  表示迭代次数。

## 2.2.2 食物选择

在 MODA 算法中添加一个档案, 用于存储和检索优化过程中真实帕累托最优解的最佳近似值。将食物来源的选取设置在帕累托最优前沿的人口最少的区域。该区域的识别方法是: 通过找到获得的帕累托最优解的最优目标和最差目标来分割搜索空间, 定义一个超球面和  $n$  个覆盖所有解的网格单元, 并在每次迭代中将超球划分为相等的子超球。使用轮盘赌轮机制进行选择, 利用轮盘机制可改善整个帕累托最优前沿的分布, 当帕累托最优个数越多, 则被选中几率越小。如下式所示:

$$P_i = \frac{c}{N_i} \quad (21)$$

式中,  $c=10$ ;  $N_i$  为在第  $i$  段内获得帕累托最优解的数量。

## 2.2.3 天敌选择

在迭代过程中, 食物源的存档文件每次都会更新, 并且在优化过程中有可能会达到总数上限。利用管理机制对存档文件进行筛选, 具体内容如下。

1) 如果存在一个解决方案可以对原存档中的帕累托最优解决方案产生主导作用, 则存储该解决方案并删去原存档中被主导的解决方案。

2) 如果存在一个解决方案  $a$  与原存档的解决方案两者之间不存在主导关系, 则保留原存档中的解决方案, 若存档个数没有达到上限, 则将解决方案  $a$  添加到存档文件中。

3) 若存档个数达到上限, 则会从填充最多的阶段删除解决方案, 并将解决方案  $a$  添加到存档文

件中; 若存在一个解决方案  $a$  可由原存档解决方案主导时, 则剔除解决方案  $a$ 。

将这些被删除的方案定义为天敌的位置, 选择最差 (个体最多) 的超球体, 可以有效阻止人工蜻蜓在没有前途的拥挤区域搜索。并通过轮盘赌轮机制实现, 每段的概率定义为:

$$P_i = \frac{N_i}{c} \quad (22)$$

## 4) Lévy flight

大多数飞行动物的行为可以用 Lévy flight 来描述<sup>[35]</sup>。当人工蜻蜓领域内没有解时, 利用 Lévy flight 在搜索空间飞行, 可有效改善人工蜻蜓的随机性和探索性:

$$\text{Lévy}(x) = 0.01 \frac{r_1 \sigma}{|r_2|^{\frac{1}{\beta}}} \quad (23)$$

式中,  $r_1$ 、 $r_2$  为  $[0, 1]$  内的随机值;  $\beta$  设为 1.5。

$\sigma$  的计算公式如下:

$$\sigma = \left\{ \frac{\Gamma(1 + \beta) \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left[\frac{1 + \beta}{2}\right] \Gamma\left[\frac{1 - \beta}{2}\right] \tau \times 2^{(\beta-1)/2}} \right\}^{\frac{1}{\beta}} \quad (24)$$

式中,  $\Gamma(x) = (x-1)!$ 。

因此, 人工蜻蜓的位置更新公式为:

$$X_{t+1} = X_t + \text{Lévy}(x) \times X_t \quad (25)$$

## 5) 蜻蜓种群初始化

蜻蜓种群通常是随机初始化的。为了提高初始种群的多样性, 避免算法产生早熟收敛的现象, 与原本的 MODA 算法中随机初始化不同, 本文的种群初始化利用 Logistic 映射代替传统的随机初始化, 计算公式如下:

$$X_{i+1} = \eta X_i (1 - X_i) \quad 0 \leq X_0 \leq 1 \quad (26)$$

式中,  $X_i$  为第  $i$  只蜻蜓的 logistic 值;  $X_0$  为蜻蜓种群初生值;  $\eta = 4$ ,  $X_0 \in (0, 1)$ ,  $X_0 \notin \{0.0, 0.25, 0.75, 1.0\}$ 。

## 6) 基于对立跳跃的种群多样性

基于对立跳跃可以有效提高进化算法的性能<sup>[11-12]</sup>。本文基于对立跳跃对 MODA 算法进行优化, 防止陷入局部最优。

基于对立跳跃的  $X_{i'}(t)$  计算公式为:

$$X_{i'}(t) = (\text{Lb}_i + \text{Ub}_i) - X_i(t) \quad (27)$$

式中,  $\text{Lb}_i$  代表搜索空间的上界;  $\text{Ub}_i$  代表搜索空间的下界。

IMODA 算法的流程图见图 3。

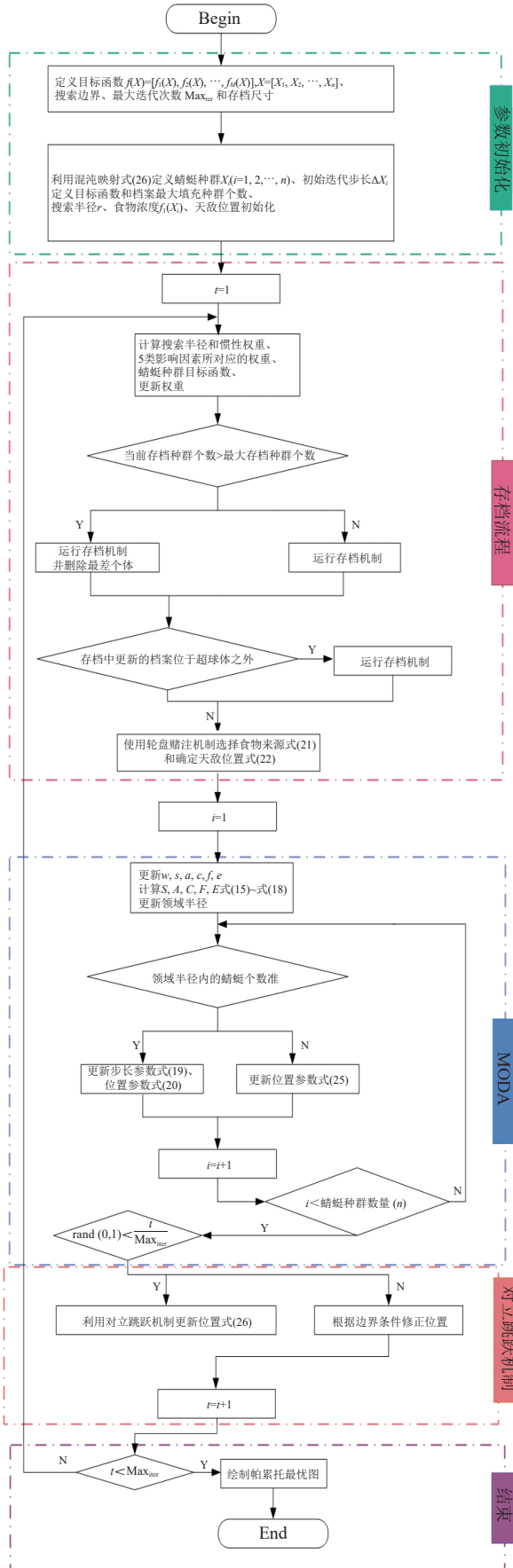


图 3 IMODA 算法流程图

### 3 算法与收敛性分析

#### 3.1 自适应学习率 CD 算法分析

收敛速度指 RBM 为了达到预期的重构误差，多次利用吉布斯采样所花费的时间。训练时间越短则表明收敛速度越快。RBM 作为一种概率模型，它的无监督学习主要是用来学习特征，这一过程称为编码。自适应学习率以变步长的方式自动调节学习因子，通过比较每两次可视层和隐含层的采样状态提高吉布斯采样的效率，从而加速 CD 算法的收敛。文献 [36] 指出，分层降维能够达到高维数据维数呈指数下降的效果。同理，由于 IMODA-ADBN 是由多个 RBM 进行的分层表述，当单个 RBM 通过自适应学习率可加速收敛的条件下，那么 DBN 的收敛速度则会提高。

自适应学习率实质上是以变步长的方式有规律地增大或减小算法对数据内在关联的学习力度，在最短的时间内收敛。但 BP 算法作为有监督学习其结论无法衡量 RBM<sup>[37]</sup>。在无监督训练阶段中，自适应学习率算法根据参数更新方向自适应地增大或减小学习率，可以准确快速寻找最优解，避免在寻优过程中陷入局部最优的情况。

#### 3.2 无监督训练阶段

在 DBN 无监督训练阶段为了快速收敛，利用自适应学习率对 RBM 依次训练。为避免特殊性，在式 (4) 和式 (5) 中将 Sigmoid 函数上下两条渐近线用  $A_H$  和  $A_L$  表示，RBM 可视层的输入信息和经过  $t$  次采样得到的重构状态分别用  $f_i^0$  和  $f_j^t$  表示。那么一次吉布斯采样过程中可视层和隐含层表达如下：

$$f_i^0 \in [A_L, A_H] \quad (28)$$

$$f_j^0 = A_L + (A_H - A_L) \sigma \left( b_j + \sum_{i=1}^m f_i^0 W_{ij} \right) \quad (29)$$

$$f_i^1 = A_L + (A_H - A_L) \sigma \left( a_i + \sum_{j=1}^n W_{ij} f_j^0 \right) \quad (30)$$

$$f_j^1 = A_L + (A_H - A_L) \sigma \left( b_j + \sum_{i=1}^m f_i^1 W_{ij} \right) \quad (31)$$

由此可以得出，经过  $t$  次吉布斯采样后有：

$$f_i^t = A_L + (A_H - A_L) \sigma \left( a_i + \sum_{j=1}^n W_{ij}^R f_j^{t-1} \right) \quad (32)$$

$$f_j^t = A_L + (A_H - A_L) \sigma \left( b_j + \sum_{i=1}^m f_i^t W_{ij}^R \right) \quad (33)$$

由上式可知, 网络输出与采样过程的中间状态有关。同时, 算法的收敛速度和精度与自适应学习率有关, 自适应学习率过大或过小都会影响到收敛速度甚至使网络不稳定。

1) 充分性证明

证明 若  $f_j^0, f_i^1 \in [A_L, A_H]$ , 根据式 (28)~式 (31) 可知, 则  $f_j^1 \in [A_L, A_H]$ 。

证明成立。

2) 必要性证明

一方面若整个网络稳定, 第一个 RBM 的输入状态满足  $f_i^0 \in [A_L, A_H]$ , 则顶层 RBM 输出状态范围满足  $f_j^1 \in [A_L, A_H]$ , 必然满足  $f_j^0, f_i^1 \in [A_L, A_H]$ 。

证明: 假设网络稳定, 那么每一层 RBM 的可视层和隐含层均满足输入输出有界。

因为 Sigmoid 函数单调递增, 且开启的神经元个数也在不断增加, 可得:

$$f_j^1 > f_i^1, f_i^1 > f_j^0 \quad (34)$$

那么可以推出:

$$f_j^0, f_i^1, f_j^1 \in [A_L, A_H] \quad (35)$$

进而推出:

$$f_j^t > f_i^t, f_i^t > f_j^0 \quad (36)$$

$$f_j^0, f_i^t, f_j^t \in [A_L, A_H] \quad (37)$$

证明成立。

推广可得如下定理。

**定理 1** 假设  $f_j^0, f_i^t, f_j^t$  分别表示 RBM 的输入状态、中间状态及输出状态, 那么网络稳定的充分必要条件则是  $f_j^0, f_i^t, f_j^t \in [A_L, A_H]$ 。

根据式 (6) 可得,  $\delta$  越大, 则神经元取 1 的概率就越小, 导致在吉布斯采样过程中可视层和隐含层神经元的稀疏性<sup>[38]</sup> 增大, 那么连续两次在吉布斯采样迭代过程中权值更新方向相同的可能性会增大。

$$P \left\{ \frac{\left| (\Delta w_{ij}^R)^{(t)} \right| + \left| (\Delta w_{ij}^R)^{(t+1)} \right|}{\left| (\Delta w_{ij}^R)^{(t)} \right| + \left| (\Delta w_{ij}^R)^{(t+1)} \right| + \varepsilon} = 1 \right\} \propto \delta \quad (38)$$

式中,  $t$  代表迭代次数。根据式 (8)~式 (12), 在调整权值过程中, 如果误差波动不太明显时, 增大学

习率可以加快权网络收敛。则有:

$$D \propto P \left\{ \frac{\left| (\Delta w_{ij}^R)^{(t)} \right| + \left| (\Delta w_{ij}^R)^{(t+1)} \right|}{\left| (\Delta w_{ij}^R)^{(t)} \right| + \left| (\Delta w_{ij}^R)^{(t+1)} \right| + \varepsilon} = 1 \right\} \quad (39)$$

根据吉布斯采样可知, 每当权值更新一次, 中间状态伴随着两次二值化采样, 主要目的是为了控制神经元的开启与关闭, 并且更新的权值与状态采样成正比, 因此可以得到  $\delta$  与学习率系数  $D$  和  $d$  的关系:

$$\begin{cases} D \approx 2\delta \\ d \approx \delta \end{cases} \quad (40)$$

式中,  $\delta$  的目的是对二值神经元状态进行判断, 一般取值为 0.7<sup>[39]</sup>。

3.3 有监督训练阶段

1) 可归约随机矩阵的稳定性

**定理 2** 设  $P$  为  $N$  阶可归约随机矩阵, 经过相同的行变换与列变换可以得  $P = \begin{bmatrix} C \cdots 0 \\ R \cdots T \end{bmatrix}$ , 其中  $C$  为  $m$  阶本原随机矩阵,  $R$  和  $T$  为  $N-M$  阶矩阵, 且均不为  $0$  矩阵。则有:

$$P^\infty = \lim_{k \rightarrow \infty} P^k = \lim_{k \rightarrow \infty} \begin{bmatrix} C^k & \cdots & 0 \\ \sum_{i=1}^{k-1} T^i R C^{k-i} & \cdots & T^k \end{bmatrix} = \begin{bmatrix} C^\infty & \cdots & 0 \\ R^\infty & \cdots & T \end{bmatrix} \quad (41)$$

式中,  $P^\infty$  为稳定随机矩阵, 且  $P^\infty = 1' P^\infty$ 、 $P^\infty = P^0 P^\infty$  唯一确定并且与初始分布无关,  $P^\infty$  满足条件:

$$P^\infty = [P_{ij}]_{N \times N} = \begin{cases} P_{ij} > 0 & 1 \leq i \leq N, 1 \leq j \leq M \\ P_{ij} = 0 & 1 \leq i \leq N, M < j \leq N \end{cases} \quad (42)$$

2) IMODA 算法全局收敛性证明

对于优化问题, 计算式如下:

$$\begin{aligned} & \min f(X) \\ & \text{s.t. } g_i(X) \geq 0 \quad i = 0, 1, 2, \dots, M \quad X \in Z \end{aligned} \quad (43)$$

式中,  $f(X)$  为目标函数;  $g_i(X)$  为第  $i$  个约束条件;  $M$  为约束条件总数;  $X$  为  $N$  维未知变量;  $Z$  为搜索空间。

设蜻蜓总数为  $N$ , 蜻蜓群用  $X'$  表示, 则  $X'$  为离散空间, 但每个蜻蜓  $X'_i$  取值连续。那么每个蜻蜓代表优化问题式 (41) 的一个候选解, 其目标函数值为  $F(X'_i)$ 。

则所有蜻蜓的状态集合表示为:

$$F = \{F(X_i^t) | X_i^t \in X^t\} \quad (44)$$

进一步令  $F = \{F_1, F_2, \dots, F_N\}, F_1 > F_2 > \dots > F_N$ ,  $\forall X \in X^t$  有  $F_N \leq F(X) \leq F_1$ , 可根据状态的区别将  $X^t$  分为若干非空子集  $X_S^i$ :

$$X_S^i = \{X | X \in X^t, F(X) = F_i\} \quad i = 1, 2, \dots, N \quad (45)$$

并且  $X_S^i$  满足下述条件:

$$\begin{aligned} \textcircled{1} \quad & \sum_{i=1}^N |X_S^i| = N, \forall i \in \{1, 2, \dots, N\}, X_S^i \neq \emptyset; \\ \textcircled{2} \quad & \forall i \neq j, \bigcup_{i=1}^N X_S^i = X^t, X_S^i \cap X_S^j = \emptyset, \bigcup_{i=1}^N X_S^i = X^t. \end{aligned}$$

令  $X^{i,j}$  满足  $i = 1, 2, \dots, N, j = 1, 2, \dots, |X_S^i|$ . 设  $X^{i,j}$  表示  $X^i$  中第  $j$  个蜻蜓. 蜻蜓在飞行过程中产生的空间位置转移记为  $X^{i,j} \rightarrow X^{m,n}$ , 发生的概率记为  $P_{ij,mn}$ ,  $X^{i,j}$  转移到第  $m$  个区域记为  $X^{i,j} \rightarrow X^m$ , 发生的概率记为  $P_{ij,m}$ ,  $X^i$  中任一蜻蜓空间位置转移至  $X^m$  则可表示  $X^i \rightarrow X^m$ , 发生的概率用  $P_{i,m}$  表示:

$$P_{ij,m} = \sum_{n=1}^{|X_S^m|} P_{ij,mn}, \quad \sum_{m=1}^N P_{ij,m} = 1, \quad P_{i,m} \geq P_{ij,m}$$

**引理 1**  $\forall X^{i,j} \in X_S^i, i = 1, 2, \dots, N, j = 1, 2, \dots, |X_S^i|$ :

$$\forall m > i, P_{i,m} = 0 \quad (46)$$

$$\exists m < i, P_{i,m} > 0 \quad (47)$$

1) 证明  $\forall m > i, P_{i,m} = 0$

设  $X^{i,j}$  为第  $t$  次演化后蜻蜓的空间位置为  $Y^t$ . 设  $Y^t$  空间中最优蜻蜓的空间位置为  $X_{\text{Best}}^t = X^*$ , 则有  $F(X_{\text{Best}}^t) = F_i$ , 根据 IMODA 算法中存档定义及在迭代过程中最好空间位置的更新可知:

$$F(X^{t+1}) \geq F(X^t)$$

可以推出:

$$\forall k > i, P_{ij,km} = 0, P_{ij,m} = \sum_{n=1}^{|X_S^m|} P_{ij,nn} = 0, P_{i,m} = 0$$

证明成立.

2) 证明  $\exists m < i, P_{i,m} > 0$ .

设  $Y^{t+1}$  空间中最优蜻蜓的空间位置为  $X_{\text{Best}}^{t+1}$ , 此时蜻蜓在探索邻域范围内随机选择行为, 有分离、对齐、聚集、捕食、避敌行为 5 种情况.

1) 蜻蜓选择分离行为. 假设存在空间位置  $F(X^{t+1})$  优于当前蜻蜓  $X$  所处空间位置  $F(X^t)$ , 即  $F(X^{t+1}) \geq F(X^t)$ , 那么蜻蜓则会更新到更优的空间位置,  $\exists m < i, P_{i,m} > 0$ , 式 (47) 得证.

2) 蜻蜓选择对齐行为. 假设存在空间位置  $F(X^{t+1})$  优于当前蜻蜓  $X$  所处空间位置  $F(X^t)$ , 即  $F(X^{t+1}) \geq F(X^t)$ , 那么蜻蜓则会更新到更优的空间位置,  $\exists m < i, P_{i,m} > 0$ , 式 (47) 得证.

3) 蜻蜓选择聚集行为. 假设存在空间位置  $F(X^{t+1})$  优于当前蜻蜓  $X$  所处空间位置  $F(X^t)$ , 即  $F(X^{t+1}) \geq F(X^t)$ , 那么蜻蜓则会更新到更优的空间位置,  $\exists m < i, P_{i,m} > 0$ , 式 (47) 得证.

4) 蜻蜓选择捕食行为. 假设存在空间位置  $F(X^{t+1})$  优于当前蜻蜓  $X$  所处空间位置  $F(X^t)$ , 即  $F(X^{t+1}) \geq F(X^t)$ , 那么蜻蜓则会更新到更优的空间位置,  $\exists m < i, P_{i,m} > 0$ , 式 (47) 得证.

5) 蜻蜓选择避敌行为. 假设存在空间位置  $F(X^{t+1})$  优于当前蜻蜓  $X$  所处空间位置  $F(X^t)$ , 即  $F(X^{t+1}) \geq F(X^t)$ , 那么蜻蜓则会更新到更优的空间位置,  $\exists m < i, P_{i,m} > 0$ , 式 (47) 得证.

由 IMODA 算法可知, 蜻蜓必选上述 5 种行为中任意一个, 因此 5 种运动行为概率总和为 1.

证明成立.

**定理 3** IMODA 算法具有收敛性.

证明:  $X_S^i (i = 1, 2, \dots, N)$  仅与当前变化有关而与历史无关, 且样本空间有限, 因此可视作有限马尔科夫链. 根据式 (2) 中引理 1 可得马尔科夫链的转移矩阵为:

$$P = \begin{bmatrix} P_{1,1} & 0 & \dots & 0 \\ P_{2,1} & P_{2,2} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ P_{N,1} & P_{N,2} & \dots & P_{N,N} \end{bmatrix} = \begin{bmatrix} C & \mathbf{0} \\ R & T \end{bmatrix} \quad (48)$$

根据式 (47) 式可得:

$$P_{2,1} > 0 \quad R = (P_{2,1}, P_{3,1}, \dots, P_{N,1})^T \quad (49)$$

$$T = \begin{bmatrix} P_{2,2} & \dots & 0 \\ \vdots & & \vdots \\ P_{N,2} & \dots & P_{N,N} \end{bmatrix} \neq \mathbf{0}, C = P_{1,1} = 1 \quad (50)$$

$P$  为  $N$  阶可归约随机矩阵, 则满足:

$$P^\infty = \lim_{k \rightarrow \infty} P^k = \lim_{k \rightarrow \infty} \left[ \sum_{i=1}^{k-1} T^i R C^{k-i} \dots T^k \right] = \begin{bmatrix} C^\infty & \dots & \mathbf{0} \\ R^\infty & \dots & T \end{bmatrix}$$

且  $C^\infty = 1, R^\infty = [1, 1, \dots, 1]^T$ .

因此  $P^\infty = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \dots & 0 \end{bmatrix}$  为稳定的随机矩阵, 可得:

$$\lim_{t \rightarrow \infty} P\{F(X_{\text{Best}}^t) = F(X^*) = F_i\} = 1 \quad (51)$$

因此 IMODA 算法具有全局收敛性。

证明成立。

### 3) IMODA 算法全局稳定性证明

由式 (2) 可知, IMODA 算法最终收敛于全局最优点, 所以  $X_0$  初始位置最终都将收敛于全局最优点  $x_{\text{max}}$ 。

假设 1  $x_{\text{max}}$  为李雅普诺夫意义下的平衡点。

证明: 设多机制水母算法的目标函数为  $f(X)$ , 那么动态方程为:

$$\dot{X} = f(X, t) \quad (52)$$

令  $x$  轴向上平移  $f(x_{\text{max}})$ , 则动态方程更新为:

$$\dot{X} = f(X, t) - f(x_{\text{max}}) \quad (53)$$

根据算法收敛性可知, 当  $t \rightarrow \infty$  时, 蜻蜓的位置状态  $X$  趋近于全局最优点  $x_{\text{max}}$ :

$$\lim_{t \rightarrow \infty} \|X(t: X_0, t_0) - X_e\| = 0 \quad (54)$$

所以对所有的  $t$ , 满足平衡状态:

$$\dot{X}_e = f(X_e, t) - f(x_{\text{max}}) = 0 \quad (55)$$

式中,  $x_{\text{max}}$  为 IMODA 算法中的平衡点;  $\dot{X}_e = f(x_{\text{max}}, t)$  为平衡状态。

所以 IMODA 算法中存在平衡点和平衡状态。

证明成立。

### 4) IMODA 算法在李雅普诺夫意义下的稳定

设 IMODA 算法的初始条件状态位于以平衡点  $x_{\text{max}}$  为圆心、 $\varepsilon$  为半径的超球域  $S(\varepsilon)$  之内, 则  $X \in S(\varepsilon)$  可表示  $S(\varepsilon) = \{X \mid \|X - x_{\text{max}}\| \leq \varepsilon\}$ , 即:

$$\|X - x_{\text{max}}\| \leq \varepsilon \quad (56)$$

如图 4 所示, 圆心为  $x_{\text{max}}$  半径为  $\gamma$  的圆  $S(\gamma)$  和半径为  $\varepsilon$  的圆  $S(\varepsilon)$ , 假设  $f(x)$  为目标函数图,  $f(x_{\text{max}})$  为函数最大值,  $f(x_{\text{max}1})$  为函数次大值,  $f(x_{\text{max}1})$  边侧两边的点的横坐标设为  $x_1, x_2$ , 圆  $S(\gamma)$  与  $f(x)$  相交点的横坐标设为  $x_3, x_4$ ,  $S_{\text{max}}$  为目标函数最大值和次大值之间区域, 称为最优区域。

证明: IMODA 算法具有李雅普诺夫意义下的稳定性。

由式 (2) 中 IMODA 全局收敛性可知, 当  $X$  处于目标函数最大值和次大值之间的区域  $S_{\text{max}}$  时,  $X$  会被食物吸引从而向着  $x_{\text{max}}$  运动, 所以方程的初始解  $X(t: X_0, t_0)$  位于  $S_{\text{max}}$  之内, 且  $S_{\text{max}}$  被包含于  $S(\varepsilon)$  与  $f(x)$  相交区域之内, 则  $X$  就不会逃离  $S(\varepsilon)$ 。

那么  $\delta$  满足:

$$\begin{cases} \delta \leq \min(\|x_{\text{max}} - x_1\|, \|x_{\text{max}} - x_2\|) \\ \delta \leq \min(f(x_{\text{max}}) - f(x_3), f(x_{\text{max}}) - f(x_4)) \end{cases} \quad (57)$$

则有:

$$\|X(t: X_0, t_0) - X_e\| \leq \gamma, t \geq t_0 \quad (58)$$

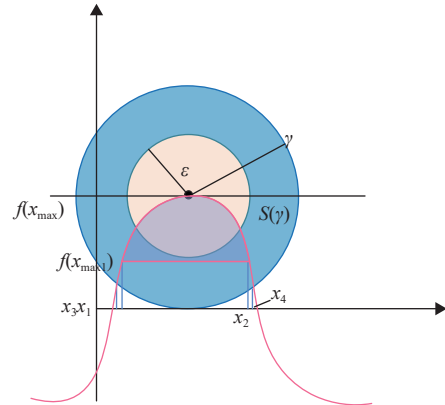


图4 Lyapunov 意义下 MODA 算法的稳定性

所以, 当  $t \rightarrow \infty$  时,  $\forall t$  使得  $X(t: X_0, t_0) \in S(\gamma)$ , 满足李雅普诺夫意义下的稳定。

因此  $\forall \gamma > 0$ ,  $\exists \varepsilon$  且  $\varepsilon$  满足式 (55), 初始状态  $x_0$  满足  $\|x_0 - x_{\text{max}}\| \leq \varepsilon$ , 那么  $x_0$  均满足  $\|X(t: X_0, t_0) - X_e\| \leq \varepsilon$ 。因此, 可得  $\varepsilon$  与  $t_0$  无关, IMODA 算法的平衡状态  $x_{\text{max}}$  是一致稳定。

证明成立。

## 4 仿真实验及研究分析

本文提出的 IMODA-ADBN 故障诊断模型主要有以下 5 个流程。

1) 被测电路两端施加脉冲信号, 采集故障数据。

2) 对数据进行故障编码。

3) 将数据集划分为训练集、测试集。

4) 在无监督学习过程中保留权重, 并利用自适应学习率加速训练过程; 在有监督学习过程中, 使用 IMODA 优化算法精调权重。

5) 模型测试: 测试集的真实编码与模型的预测编码进行对比, 若预测编码与真实编码一致, 则分类正确; 若预测编码与真实编码不一致, 则分类错误。

### 4.1 二级四运放双二阶低通滤波器诊断示例

本文用输入为幅值为 5 V、频率为 10 kHz 的正弦波信号激励的二级四运放双二阶低通滤波器 (如图 5 所示) [40], 验证本文提出的 IMODA-

ADBN 模型，将电路中电阻及电容的容差均设置为 5%。根据一般元件的参数性故障定义，当元件参数值偏离标称值 50% 时，即可判定该元件发生故障<sup>[3]</sup>。本次实验设定滤波器中单一电阻或电容

出现故障时，而其他的电阻或电容在容差范围内随意变化，即处于正常状态。并且根据灵敏度测试结果，选择电阻  $R_3$ ,  $R_{19}$  和电容  $C_1$ ,  $C_4$  这 4 种元器件作为研究对象进行实验分析，由表 1 所示。

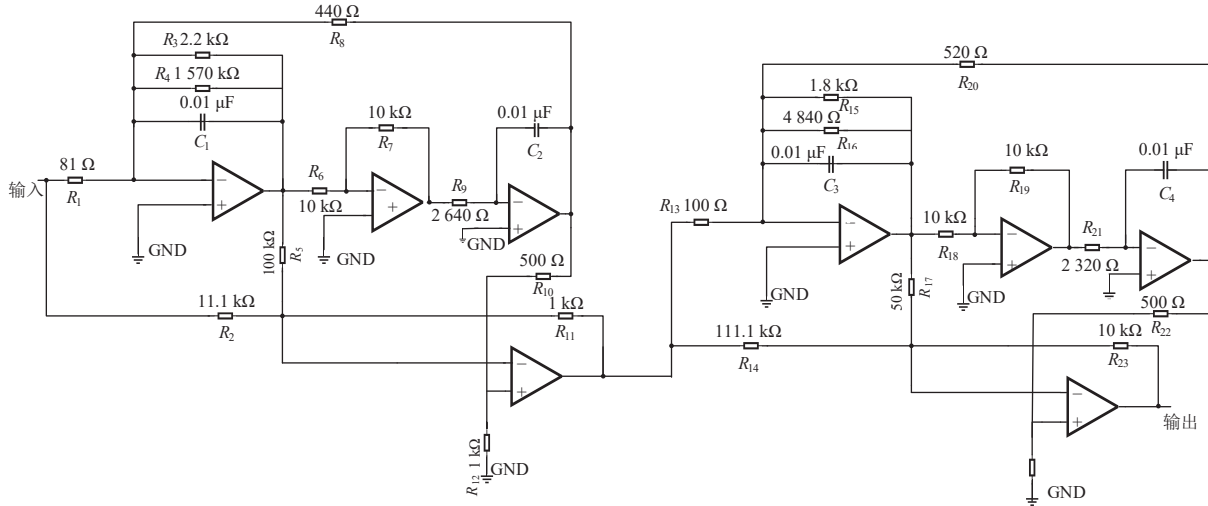


图 5 二级四运放双二阶低通滤波器

表 1 二级四运放双二阶低通滤波器电路故障模式

故障代码	故障类型	容差/%	标称值	故障值
$f_1$	无故障	—	—	—
$f_2$	$R_9$ 增大/ $\Omega$	5	2 640	3 960
$f_3$	$R_9$ 减小/ $\Omega$	5	2 640	1 320
$f_4$	$R_{19}$ 增大/ $\Omega$	5	10 k	15 k
$f_5$	$R_{19}$ 减小/ $\Omega$	5	10 k	5 k
$f_6$	$C_1$ 增大/nF	5	0.01	0.005
$f_7$	$C_1$ 减小/nF	5	0.01	0.015
$f_8$	$C_4$ 增大/nF	5	0.01	0.005
$f_9$	$C_4$ 减小/nF	5	0.01	0.015

深度学习网络多采用 3 层隐藏层叠加，可使重构误差达到最小。利用均方误差 (MSE)、绝对百分比误差 (APE) 及运行时间作为参考标准，用实验法选出最佳神经元个数。

$$APE = \frac{1}{N_t} \sum_{i=1}^{N_t} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (59)$$

$$MSE = \frac{1}{N_t} \sum_{i=1}^{N_t} (\hat{y}_i - y_i)^2 \quad (60)$$

式中， $y_i$ 和 $\hat{y}_i$ 分别表示真实值和预测分类值； $N_t$ 表示测试样本个数。

图 6 为模型误差与隐含层神经元个数关系图，

由图可知当隐含层的神经元个数为 20 时，性能效果最佳。此时 APE、MSE 分别是 0.050 7 和 0.054 2。

为证明 IMODA-ADBN 最佳的模型结构是 10-20-20-20-9，将双二阶四运放低通滤波器单次采样时间设为 0.05 s，将同种故障模式下 20 种不同时段的数据样本设置为 一组，每种故障采集 20 组，共计 9 种故障，构成 9×400 原始数据集。随机选取 280 组作为训练样本，剩余的 120 组作为测试样本。

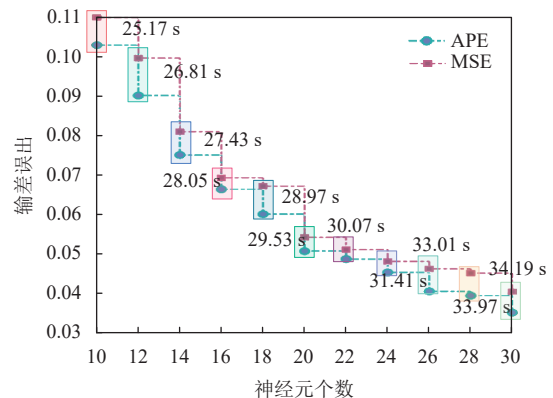


图 6 不同神经元个数模型误差图

在无监督训练阶段中每个 RBM 设置迭代 90 次，学习率系数分别设置为  $D=1.4$ ,  $d=0.7$ 。利用 KPCA (kernel principal component analysis) 提取第一个 RBM 输出特征的 3 个主成分和 DBN 最终

输出特征的 3 个主成分, 分别如图 7a 和图 7b 表示。

从图 7a 中可以看出  $f_7$ 、 $f_8$  和  $f_9$  有重叠现象产生;  $f_4$ 、 $f_7$  有轻微重叠现象;  $f_3$ 、 $f_5$  故障相聚较近; 剩余故障虽然没有重叠现象产生但同类故障之间的特征点分布较为分散。输入数据经过 IMODA-ADB N 模型会进行 4 次非线性映射并重构数据, 从图 7b 中可以看出 9 种故障经本文提出的方法可以被精准区分。并且相较图 7a 而言, 经 IMODA-ADB N 模型处理后的同类故障的主成分分布相对于图 7a 更加紧凑, 因此, 可以得出 IMODA-ADB N 模型可以通过增加 RBM 的个数提高模型的特征表达能力, 从而能够更加精确抽象地表达输入数据。

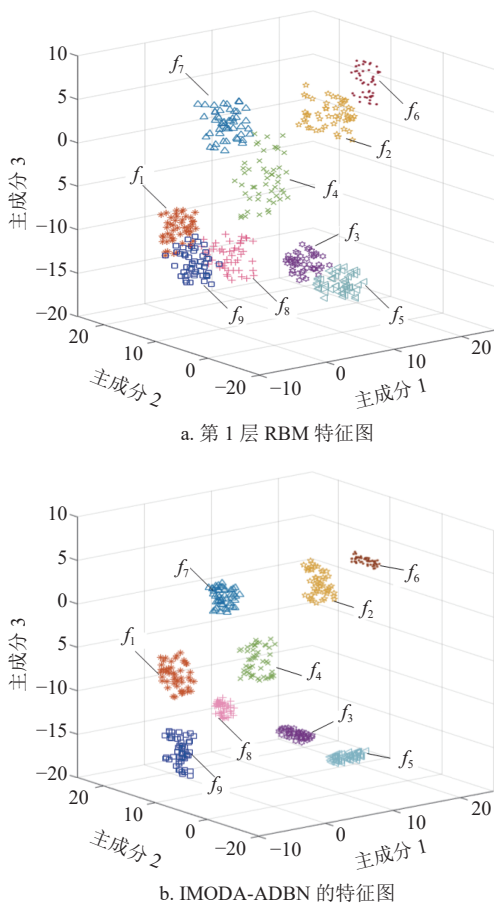


图 7 特征主成分分布图

为降低实验随机性对模型诊断性能的评判影响, 对二级四运放双二阶低通滤波器电路进行 10 次独立诊断实验。其中, 图 8a 为 IMODA-ADB N 模型的 10 次诊断故障混淆矩阵图, 其中每类故障的准确率计算 10 次均值, 而错分个数则进行 10 次实验累计统计。从图 8a 中可以看出, 将故障错分为其他故障类型的累计错分率均不超过 1%, 并且每类故障诊断准确率均在 98% 以上。表

明本文提出的 IMODA-ADB N 模型对 9 种故障可以精准识别; 图 8b 为 IMODA-ADB N 模型在 10 次诊断实验中的准确率、精确率及召回率 3 种指标的性能曲线图。从图中看出, 在 10 次实验过程中, IMODA-ADB N 模型在每一次实验过程中的精确率、准确率、召回率均在 96% 以上。因此证明本文提出的 IMODA-ADB N 模型具有较强的鲁棒性。

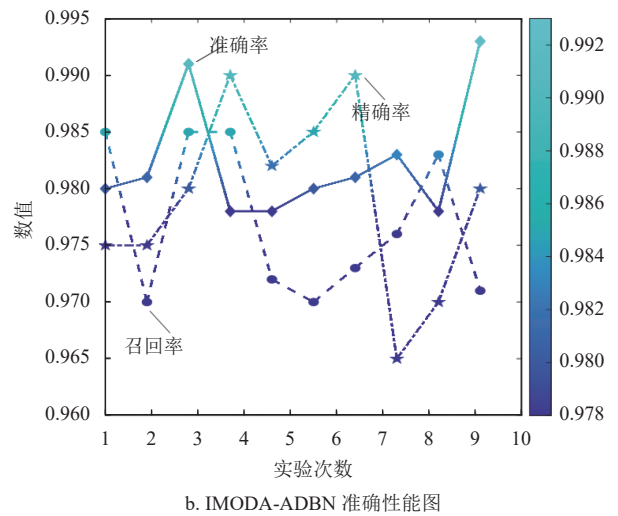
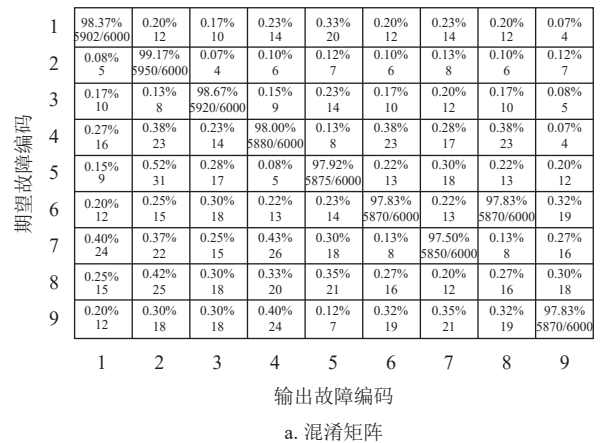


图 8 IMODA-ADB N 诊断结果图

## 4.2 对比实验

### 4.2.1 自适应学习实验验证

RBM 作为概率模型, 极易受到权重对其的影响。为了保证 RBM 准确提取特征, 需要保证合理的权重。由式 (7)~式 (10) 可以看出, DBN 在学习过程中, 权重的大小由学习率决定。然而, 以往 DBN 在学习过程中学习率都依靠人工经验设定, 缺乏合理性。因此, 为进一步地证明自适应学习率的合理性与有效性, 本文选取固定学习率与自适应学习率作为实验对象, 采用误差作为实验目标进行对比试验。其中, 固定学习率设置为 0.01、0.05、

0.1、0.5 和 1，并固定有监督学习部分并进行 10 次对比实验。将前面获得的仿真数据划分训练数据和测试数据进行实验验证。

图 9 为迭代次数和误差关系图。可以看出，模型误差会随迭代次数的增加而逐渐降低。表 2 代表不同学习率情况下模型在 10 次实验过程中的平均诊断准确率。结合图 9 和表 2 可以得出以下结果。

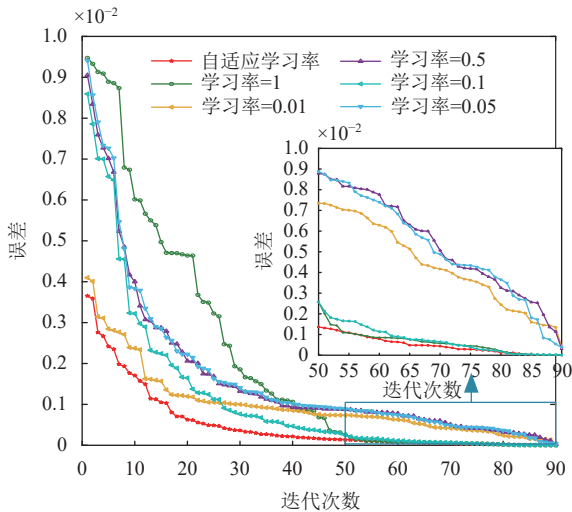


图 9 迭代次数和误差关系图

表 2 不同学习率对应的平均准确率

学习率	平均准确率/%
0.01	94.26
0.05	93.73
0.1	96.21
0.5	94.57
1	93.75
自适应	98.02

1) 以固定学习率为研究目标，以 IMODA-ADBN 模型的 10 次诊断实验平均准确率作为衡量指标，可以得出：将学习率设置为 0.1 时模型的诊断性能最好（96.21%）。而将学习率设置为 1 和 0.05 时模型的诊断性能均不太理想（93.75% 和 93.73%）。学习率设置为 0.5 时，平均诊断率为 94.57%，虽在迭代前期时模型误差下降幅度较大，但与最优固定学习率 0.1 相比，诊断准确率却下降了 1.64%；而将学习率设置为 0.01 时，虽在迭代前期时模型误差较小，准确率却下降了 1.95%。

2) 以学习率为 0.1 作为参考线，其中学习率高（学习率为 0.5 和 1）造成的准确率下降可能是因为参数更新过大，使模型无法收敛；学习率低（学习率为 0.01 和 0.05）造成的准确率下降的原

因可能是在学习率过低的情况下，模型在训练初期就停滞在一个较差的局部最小值，而没有足够的能量去跳出这个局部最小值并继续学习。

3) 自适应学习率的设计使模型在 75 次左右迭代时已经趋于稳定，在 55 次左右迭代时误差已经小于 0.01，保证模型收敛速度的同时有效降低模型的误差，从而在对比实验中获得了最高的平均诊断率（98.02%）。与固定学习率最优值 0.1 相比，准确率提高了 1.84%。可以证明所提出的基于参数更新方向设置的自适应学习率与传统经验学习率相比更具有先进性与合理性。

#### 4.2.2 IMODA 优化算法性能验证

为了增加 MODA 优化过程中的多样性，以及改善陷入局部最优的情况，在传统 MODA 优化算法中增加 Logistic 映射和对立跳跃。二者可以提高初始种群的多样性，并降低“早熟”的概率<sup>[41]</sup>，选用 ZTD 函数<sup>[42]</sup>、Schaffer<sup>[43]</sup> 和 DTLZ 函数<sup>[44]</sup> 共计 7 种测试函数验证 IMODA 算法的性能，并与 3 种多目标优化算法（MODA、MOPSO<sup>[45]</sup> 和 NSGA-II<sup>[46]</sup>）进行了比较，其中算法的参数设置见表 3，测试函数的表达式及限定范围见表 4，并根据对应的帕累托最优前沿特征进行归纳。表 5 则对不同算法的性能进行了实验验证。

表 3 不同算法的参数设置

算法	参数
IMODA	种群数量 $N = 150$
	档案大小 $N_{rep} = 150$
	网格分量 $n_{grid} = 20$
	最大迭代次数 $Max_{iter} = 2500$
MODA	种群数量 $N = 150$
	档案大小 $N_{rep} = 150$
	最大迭代次数 $Max_{iter} = 2500$
	种群数量 $N = 150$
MOPSO	档案大小 $N_{rep} = 150$
	网格分量 $n_{grid} = 20$
	最大迭代次数: $Max_{iter} = 2500$
	惯性权重（初始）: $w = 0.5$
	惯性权重（最终）: $w = 0.001$
	变异概率: $P_M = 0.1$
NSGA-II	种群数量 $N = 150$
	网格分量 $n_{grid} = 20$
	最大迭代次数: $Max_{iter} = 2500$
	变异概率: $P_M = 0.5$

表4 多目标测试函数的特性

测试函数	变量个数	目标函数个数	变量范围	表达式	特征
ZTD1	30	2	[0,1]	$g(x) = 1 + \frac{9}{N-1} \sum_{i=2}^N x_i$ $h(f_1(x), g(x)) = 1 - \sqrt{\frac{f_1(x)}{g(x)}}$ $f_1(x) = x_1$ $f_2(x) = g(x)h(f_1(x), g(x))$	帕累托最优前沿有一个凸前沿
ZTD2	30	2	[0,1]	$g(x) = 1 + \frac{9}{N-1} \sum_{i=2}^N x_i$ $h(f_1(x), g(x)) = 1 - \left(\frac{f_1(x)}{g(x)}\right)^2$ $f_1(x) = x_1$ $f_2(x) = g(x)h(f_1(x), g(x))$	帕累托最优前沿有一个非转换前沿
ZTD3	30	2	[0,1]	$g(x) = 1 + \frac{9}{29} \sum_{i=2}^N x_i$ $h(f_1(x), g(x)) = 1 - \sqrt{\frac{f_1(x)}{g(x)}} - \left(\frac{f_1(x)}{g(x)}\right) \sin(10\pi f_1(x))$ $f_1(x) = x_1$ $f_2(x) = g(x)h(f_1(x), g(x))$	帕累托最优前沿具有不连续前沿
Schaffer	1	2	[-5,5]	$f_1(x) = x_1^2$ $f_2(x) = (x-2)^2$	帕累托最优前沿有一个凸前沿
DTLZ1	7	3	[0,1]	$g(x_M) = 100 \left[  x_M  + \sum_{x_i \in x_M} (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5)) \right]$ $f_1(x) = \frac{1}{2} x_1 x_2 \cdots x_{M-1} (1 + g(x_M))$ $f_2(x) = \frac{1}{2} x_1 x_2 \cdots (1 - x_{M-1}) (1 + g(x_M))$ $f_3(x) = \frac{1}{2} x_1 x_2 \cdots (1 - x_{M-2}) x_{M-1} (1 + g(x_M))$	帕累托最优前沿有一个线性前沿
DTLZ2	12	3	[0,1]	$g(x_M) = \sum_{x_i \in x_M} (x_i - 0.5)^2$ $f_1(x) = (1 + g(x_M)) \cos\left(\frac{x_1\pi}{2}\right) \cdots \cos\left(\frac{x_{M-2}\pi}{2}\right) \cos\left(\frac{x_{M-1}\pi}{2}\right)$ $f_2(x) = (1 + g(x_M)) \cos\left(\frac{x_1\pi}{2}\right) \cdots \cos\left(\frac{x_{M-2}\pi}{2}\right) \sin\left(\frac{x_{M-1}\pi}{2}\right)$ $f_3(x) = (1 + g(x_M)) \cos\left(\frac{x_1\pi}{2}\right) \cdots \sin\left(\frac{x_{M-2}\pi}{2}\right)$	帕累托最优前沿为球面前沿
DTLZ3	12	3	[0,1]	$g(x_M) = 100 \left[  x_M  + \sum_{x_i \in x_M} (x_i - 0.5)^2 - \cos(20\pi(x_i - 0.5)) \right]$ $f_1(x) = (1 + g(x_M)) \cos\left(\frac{x_1\pi}{2}\right) \cdots \cos\left(\frac{x_{M-2}\pi}{2}\right) \cos\left(\frac{x_{M-1}\pi}{2}\right)$ $f_2(x) = (1 + g(x_M)) \cos\left(\frac{x_1\pi}{2}\right) \cdots \cos\left(\frac{x_{M-2}\pi}{2}\right) \sin\left(\frac{x_{M-1}\pi}{2}\right)$ $f_3(x) = (1 + g(x_M)) \cos\left(\frac{x_1\pi}{2}\right) \cdots \sin\left(\frac{x_{M-2}\pi}{2}\right)$	帕累托最优前沿具有多前沿

表5 不同算法测试函数性能对比

测试函数	IGD	算法			
		IMODA	MODA	MOPSO	NSGA-II
ZDT1	平均值	<b>0.005 47</b>	0.005 87	0.004 72	0.042 17
	标准差	<b>0.002 31</b>	0.003 01	0.004 18	0.006 71
	中值	<b>0.006 42</b>	0.007 70	0.004 06	0.041 33
	最优值	<b>0.002 14</b>	0.002 53	0.001 72	0.020 71
	最差值	<b>0.007 13</b>	0.008 27	0.012 24	0.059 22
ZDT2	平均值	<b>0.003 37</b>	0.004 26	0.006 72	0.040 17
	标准差	<b>0.001 85</b>	0.002 71	0.003 95	0.003 32
	中值	<b>0.002 57</b>	0.003 07	0.007 99	0.033 73
	最优值	<b>0.002 20</b>	0.002 42	0.004 43	0.025 10
	最差值	<b>0.005 93</b>	0.006 11	0.007 24	0.045 21

续表

测试函数	IGD	算法			
		IMODA	MODA	MOPSO	NSGA-II
ZDT3	平均值	<b>0.009 21</b>	0.030 17	0.036 74	0.048 33
	标准差	<b>0.009 07</b>	0.004 16	0.006 08	0.009 72
	中值	<b>0.009 51</b>	0.033 72	0.039 41	0.041 57
	最优值	<b>0.008 30</b>	0.021 51	0.020 97	0.032 70
	最差值	<b>0.009 54</b>	0.038 49	0.041 72	0.051 22
Schaffer	平均值	<b>0.002 84</b>	0.003 72	0.005 07	0.006 93
	标准差	<b>0.003 01</b>	0.003 99	0.005 00	0.007 10
	中值	<b>0.002 40</b>	0.003 57	0.004 16	0.005 93
	最优值	<b>0.001 26</b>	0.002 42	0.002 12	0.003 92
	最差值	<b>0.003 37</b>	0.004 75	0.005 01	0.007 14
DTLZ1	平均值	<b>0.005 51</b>	0.006 08	0.059 22	0.091 50
	标准差	<b>0.002 08</b>	0.004 09	0.030 61	0.060 03
	中值	<b>0.004 71</b>	0.004 07	0.050 30	0.077 01
	最优值	<b>0.002 10</b>	0.001 92	0.001 88	0.047 10
	最差值	<b>0.009 02</b>	0.009 71	0.010 27	0.122 05
DTLZ2	平均值	<b>0.008 94</b>	0.010 98	0.051 40	0.085 11
	标准差	<b>0.000 67</b>	0.009 01	0.015 25	0.057 20
	中值	<b>0.010 22</b>	0.008 70	0.020 81	0.065 87
	最优值	<b>0.009 01</b>	0.006 11	0.014 29	0.049 30
	最差值	<b>0.010 78</b>	0.012 07	0.068 14	0.102 21
DTLZ3	平均值	<b>0.003 09</b>	0.006 61	0.038 79	0.041 02
	标准差	<b>0.002 55</b>	0.005 20	0.008 19	0.009 07
	中值	<b>0.002 61</b>	0.005 87	0.040 66	0.040 99
	最优值	<b>0.002 23</b>	0.003 17	0.021 09	0.030 22
	最差值	<b>0.003 72</b>	0.009 11	0.043 07	0.050 88

通过 10 次独立实验得出的最佳帕累托最优前沿, 如图 10 所示, 并用 inverse generational distance (IGD) 作为衡量指标:

$$IGD = \frac{\sqrt{\sum_{i=1}^n d_i^2}}{n} \quad (61)$$

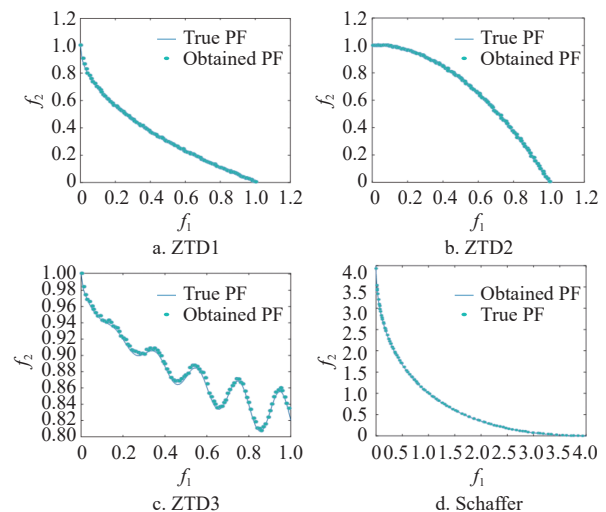
式中,  $n$  是帕累托最优解的数目;  $d_i$  表示第  $i$  个帕累托最优解与参考集中最近的帕累托最优解之间的欧氏距离。

从表 5 可以看出, IMODA 算法与 MODA 算法、MOPSO 算法和 NSGA-II 算法相比, IMODA 算法在 7 种测试函数中平均值、标准差、中值、最优值和最差值 5 种评价指标上均达到了最佳性能。图 10 为 IMODA 在上述 7 种测试函数上的 Pareto 解, 可以看出 IMODA 算法得出的解可以全覆盖 Pareto 前沿, 并且解的分布较为均匀, 表明引入 Logistic 映射和对立跳跃策略能有效改善 MODA 算法的性能。

#### 4.2.3 IMODA-ADBN 模型验证

将本文方法与一些常用模拟电路故障诊断与分类的方法进行对比实验来证明 IMODA-ADBN 模型对复杂电路故障诊断分类的有效性和优越性。在

浅层方面, 用 BP、ELM、SVM、LSSVM<sup>[16]</sup> 进行分类诊断, 在深层方面, 采用传统 DBN、CDBN<sup>[47]</sup>; 另外利用重叠采样扩充数据集后用 ECA-CNN<sup>[49]</sup> 及 CBAM-CNN 进行对比<sup>[48]</sup>, 因为选取的对比模型中部分模型并不只针对模拟电路的故障进行诊断, 所以做出以下两点约束: 1) ECA-CNN、CBAM-CNN 的输入均按照 6:2:2 划分训练集、验证集和测试集; 2) DBN、CDBN 的模型结构与 IMODA-ADBN 模型结构相同。



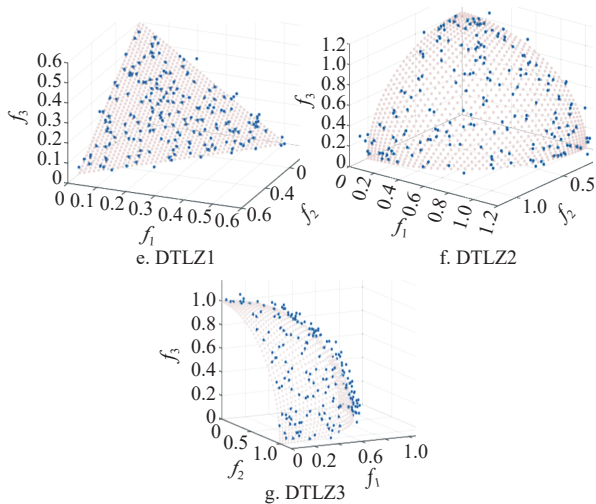


图 10 IMODA 测试函数性能图

将上述模型与本文提出的 IMODA-ADB N 进行 10 次独立诊断实验对比, 结果如表 6 所示。

表 6 不同模型性能对比

方法	F1 Measure	平均运行时间/s	平均正确率/%
BP <sup>[47]</sup>	0.796	46.12	75.92
ELM <sup>[47]</sup>	0.802	47.81	80.09
SVM <sup>[47]</sup>	0.814	53.07	82.17
LSSVM <sup>[16]</sup>	0.827	50.17	84.11
DBN <sup>[32]</sup>	0.913	43.91	91.08
CDBN <sup>[46]</sup>	0.942	40.77	92.72
ECA-CNN <sup>[48]</sup>	0.950	6.51	95.43
CBAM-CNN <sup>[47]</sup>	0.977	6.72	97.91
IMODA-ADB N	0.989	30.94	98.14

1) 仅从平均正确率角度出发, 采用的浅层机器学习方法 BP、ELM、SVM 以及 LSSVM 作为诊断模型的平均正确率均低于深层机器学习模型。在深层机器学习模型中, 以 CNN 为基础模型框架, 采用注意力机制改进的 ECA-CNN 以及 CBAM-CNN 模型的诊断准确率分别为 95.43% 和 97.91%; 以 DBN 为诊断模型的平均诊断准确率最低 (91.08%), 而 DBN 的改进模型 CDBN 和 IMODA-ADB N 模型的平均诊断准确率分别为 92.72% 和 98.14%, 并且 IMODA-ADB N 诊断模型取得了最优表现。

2) 仅从平均运行时间角度出发, 浅层机器学习模型相对于深层机器学习模型均较为耗时。在深层机器学习模型中, ECA-CNN 以及 CBAM-CNN 模型测试较快 (6.51 s 和 6.72 s), 但前期采用的重叠采样增强数据集操作增大了数据制作成本。DBN 及 CDBN 模型的测试时间为 43.91 s 和

40.77 s, IMODA-ADB N 相较于 DBN 和 CDBN 分别降低了 12.97 s 和 24.22 s。

仅从诊断准确率或运行时间单一角度出发并不能对模型性能做出最公平的判断。结合平均正确率及运行时间两个角度出发分析各类模型, 本文提出的 IMODA-ADB N 主要优势有: 1) 与浅层相比, 可以有效提取数据的本质特征; 2) 对 DBN 合理优化可以使准确和性能都有不同程度的提升; 3) 与 SENet-CNN 和 CBAM-CNN 相比, 对数据集的需求量较小。因此 IMODA-ADB N 在故障的精准识别方面有一定的优势。

#### 4.2.4 IMODA-ADB N 模型噪声验证

考虑到模拟电路中通常有白噪声的干扰, 同时, 为进一步评估 IMODA-ADB N 模型的鲁棒性, 在 4.1 小节中所获取的仿真数据基础上, 分别添加信噪比为 40、25、10 dB 的白噪声进行噪声实验。图 11 展示了正常输出情况下的部分电压曲线及添加了 3 种不同噪声的电压曲线。

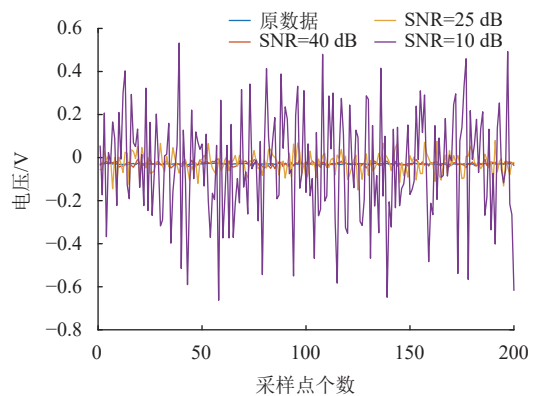


图 11 添加不同白噪声的电压曲线图

通过表 6 分析可知, 浅层机器学习与深度机器学习相比, 浅层机器学习的特征提取能力存有一定不足。鉴于此, 在噪声实验中, 选择 DBN、CDBN、ECA-CNN 以及 IMODA-CNN 作为对比模型来衡量 IMODA-ADB N 模型的性能优劣。为避免随机性对模型性能的影响, 进行 5 次独立诊断实验并取平均值作为最后的衡量指标, 实验结果如表 7 所示。

从表 7 可以看出, IMODA-ADB N 模型在 40、25、10 dB 这 3 种白噪声情况下, 均取得了最佳性能, 而传统 DBN 模型却仍有一定的改进空间。以 CNN 为基础模型采用 ECA 和 CBAM 模块改进后的 ECA-CNN 模型以及 CBAM-CNN 模型虽然在 3 种噪声扰动下也表现出不错的故障诊断性

能, 但因其 CNN 模型对数据量的要求导致这两类模型对数据成本要求过大, 经济成本也较高, 并且与 IMODA-ADBN 模型相比还是存在一定差距。结合 CNN 与 DBN 的 CDBN 模型的性能与 DBN 模型相比虽然有了一定的提高, 但由于引入 CNN 模型, 除了对数据量的要求较高以外, CDBN 模型也更加依赖参数的设置。

表 7 不同噪声下的模型性能对比

方法	SNR/dB	平均正确率/%	F1 Measure
DBN	40	89.93	0.895
	25	85.24	0.833
	10	79.27	0.790
	40	91.53	0.904
CDBN	25	85.69	0.849
	10	80.93	0.801
	40	93.21	0.922
ECA-CNN	25	89.59	0.890
	10	86.24	0.859
	40	93.85	0.934
CBAM-CNN	25	89.27	0.889
	10	87.01	0.868
	40	95.21	0.951
IMODA-ADBN	25	92.26	0.917
	10	89.14	0.884

## 5 结束语

DBN 作为深度学习中心中最典型的模型已经在不同领域中得到了研究与应用。DBN 在面对大量数据时, 无监督预训练耗时也会急剧增加, 利用 graphic processing unit (GPU) 硬件加速收敛已成为近几年的主要手段, 然而该方法需要较大的经济成本; 另外在反向精调过程中, BP 算法由于本身性质易陷入局部最优。针对上述问题, 本文提出了一种基于 IMODA-ADBN 的模拟电路故障诊断方法, 以应对 DBN 在大数据背景下的训练耗时和反向精调中的挑战。研究创新点如下。

1) 对 MODA 算法进行优化, 加入 Logistic 混沌映射和对立跳跃增加多样性, 并对改进的 IMODA 算法与传统 MODA 算法、MOPSO 算法和 NSGA-II 算法比较, 利用测试函数进行实验验证, 实验表明 IMODA 算法产生的 Pareto 解的分布更加均匀。

2) 引入自适应学习率以加速无监督预训练阶段的网络收敛, 并证明优化后的无监督部分具有稳定性。

3) 利用 IMODA 算法对有监督部分进行精细调整, 证明其具有全局收敛性, 并验证优化系统在李雅普诺夫意义下的稳定性。

最终将 IMODA-ADBN 利用双二阶四运放低通滤波器进行模拟实验验证。仿真实验和噪声实验结果表明, IMODA-ADBN 模型能够实现对故障高效分类与定位, 在模拟电路故障诊断等领域中, 可以作为一种新的解决方案。

## 参考文献

- [1] SAI S V A, LONG B, PECHT M. Diagnostics and prognostics method for analog electronic circuits[J]. IEEE Transactions on Industrial Electronics, 2013, 60(11): 5277-5291.
- [2] AO Y, SHI Y, WEI Z, et al. An approximate calculation of ratio of normal variables and its application in analog circuit fault diagnosis[J]. Journal of Electronic Testing Theory & Applications, 2013, 29(4): 555-565.
- [3] YANG C L. Parallel-series multi objective genetic algorithm for optimal tests selection with multiple constraints[J]. IEEE Transactions on Instrumentation and Measurement, 2018, 67(8): 1-18.
- [4] TANG X, XU A. Practical analog circuit diagnosis based on fault features with minimum ambiguities[J]. Journal of Electronic Testing, 2016, 32(1): 1-13.
- [5] TADEUSIEWICZ M, HALGAS S. A method for local parametric fault diagnosis of a broad class of analog integrated circuits[J]. IEEE Transactions on Instrumentation and Measurement, 2018, 67(2): 328-337.
- [6] AMINIAN M, AMINIAN F. A modular fault-diagnostic system for analog electronic circuits using neural networks with wavelet transform as a preprocessor[J]. IEEE Transactions on Instrumentation and Measurement, 2007, 56(5): 1546-1554.
- [7] LAURAIN V, TÓTH R, PIGA D, et al. An instrumental least squares support vector machine for nonlinear system identification[J]. Automatica, 2015, 54: 340-347.
- [8] HU Q L, XIAO B. Adaptive fault tolerant control using integral sliding mode strategy with application to flexible spacecraft[J]. International Journal of Systems Science, 2013, 44(12): 2273-2286.
- [9] 段晨东, 何正嘉, 姜洪开. 非线性小波变换在故障特征提取中的应用[J]. 振动工程学报, 2005(1): 129-132.
- [10] DUAN C D, HE Z J, JIANG H K. Fault feature extraction using nonlinear wavelet transform[J]. Journal of Vibration Engineering, 2005(1): 129-132.
- [11] COHEN A, TIPLICA T, KOBI A. Design of experiments and statistical process control using wavelets analysis[J]. Control Engineering Practice, 2016, 49(4): 129-138.
- [12] TADEUSIEWICZ M, HALGAS S. A new approach to

- multiple soft fault diagnosis of analog BJT and CMOS circuits[J]. *IEEE Transactions on Instrumentation and Measurement*, 2015, 64(10): 2688-2695.
- [12] SPYRONASIOS A D, DIMOPOULOS M G, HATZOPOULOS A A. Wavelet analysis for the detection of parametric and catastrophic faults in mixed-signal circuits[J]. *IEEE Transactions on Instrumentation and Measurement*, 2011, 60(6): 2025-2038.
- [13] GAN X S, GUO W M, DAI Z, et al. Research on WNN soft fault diagnosis for analog circuit based on adaptive UKF algorithm[J]. *Applied Soft Computing*, 2017, 50: 252-259.
- [14] XIAO Y Q, FENG L G. A novel linear ridgelet network approach for analog fault diagnosis using wavelet-based fractal analysis and kernel PCA as preprocessors[J]. *Measurement*, 2012, 45(3): 297-310.
- [15] LONG B, LI M, WANG H J, et al. Diagnostics of analog circuits based on LS-SVM using time-domain features[J]. *Circuits Systems & Signal Processing*, 2013, 32(6): 2683-2706.
- [16] LONG B, XIAN W M, LI M, et al. Improved diagnostics for the incipient faults in analog circuits using LSSVM based on PSO algorithm with mahalanobis distance[J]. *Neurocomputing*, 2014, 133: 237-248.
- [17] PENG C, YUAN L F, HE Y G, et al. An improved SVM classifier based on double chains quantum genetic algorithm and its application in analogue circuit diagnosis[J]. *Neurocomputing*, 2016, 211: 202-211.
- [18] RAJAGOPAL M. Multiple parametric fault diagnosis using computational intelligence techniques in linear filter circuit[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2020, 11(2): 5533-5545.
- [19] GUO S, WU B, ZHOU J, et al. An analog circuit fault diagnosis method based on circle model and extreme learning machine[J]. *Applied Sciences*, 2020, 10(7): 2386-2400.
- [20] HUANG G B, DING X, ZHOU H. Optimization method based extreme learning machine for classification[J]. *Neurocomputing*, 2010, 74(1-3): 155-163.
- [21] GUO X J, CHEN L, SHEN C Q. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis[J]. *Measurement*, 2016, 93: 490-502.
- [22] JIA F, LEI Y G, LIN J, et al. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data[J]. *Mechanical Systems & Signal Processing*, 2016, 72-73: 303-315.
- [23] SHAO H D, JIANG H K, WANG F, et al. Rolling bearing fault diagnosis using adaptive deep belief network with dual-tree complex wavelet packet[J]. *ISA Transactions*, 2017, 69: 187-201.
- [24] FANG F, LI L Y, GU Y, et al. A novel hybrid approach for crack detection[J]. *Pattern Recognition*, 2020, 107: 107474.
- [25] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 39(4): 640-651.
- [26] LIU C, CHENG G, CHEN X, et al. Planetary gears feature extraction and fault diagnosis method based on VMD and CNN[J]. *Sensors*, 2018, 18(5): 1523-1542.
- [27] LIU R, MENG G, YANG B, et al. Dislocated time series convolutional neural architecture: An intelligent fault diagnosis approach for electric machine[J]. *IEEE Transactions on Industrial Informatics*, 2017, 13(3): 1310-1320.
- [28] ZHAO G Q, LIU X Y, ZHANG B, et al. A novel approach for analog circuit fault diagnosis based on deep belief network[J]. *Measurement*, 2018, 121: 170-178.
- [29] LIU Z, JIA Z, VONG C M, et al. Capturing high-discriminative fault features for electronics-rich analog system via deep learning[J]. *IEEE Transactions on Industrial Informatics*, 2017, 1(3): 1213-1226.
- [30] ZHONG T, QU J F, FANG X Y, et al. The intermittent fault diagnosis of analog circuits based on EEMD-DBN[J]. *Neurocomputing*, 2021, 436: 74-91.
- [31] MENG G, CONG W, ZHU C A. Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings[J]. *Mechanical Systems & Signal Processing*, 2016, 72-73: 92-104.
- [32] 张朝龙, 何怡刚, 杜博伦. 基于 DBN 特征提取的模拟电路早期故障诊断方法[J]. *仪器仪表学报*, 2019, 40(10): 112-119.
- ZHANG C L, HE Y G, DU B L. Analog circuit incipient fault diagnosis method based on DBN feature extraction[J]. *Chinese Journal of Scientific Instrument*, 2019, 40(10): 112-119.
- [33] CHEN Z Y, LI W H. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network[J]. *IEEE Transactions on Instrumentation and Measurement*, 2017, 66(7): 1693-1702.
- [34] MIRJALILI S. Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems[J]. *Neural Computing and Applications*, 2016, 27(4): 1053-1073.
- [35] PAVLYUKEVICH I. Lévy flights, non-local search and simulated annealing[J]. *Journal of Computational Physics*, 2007, 226: 1830-1844.
- [36] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.
- [37] 乔俊飞, 王功明, 李晓理, 等. 基于自适应学习率的深度信念网设计与应用[J]. *自动化学报*, 2017, 43(8): 1339-1349.

- QIAO J F, WANG G M, LI X L, et al. Design and application of deep belief network with adaptive learning rate[J]. *Acta Automatica Sinica*, 2017, 43(8): 1339-1349.
- [38] JIA N N, ZHANG J S, ZHANG C X. A sparse-response deep belief network based on rate distortion theory[J]. *Pattern Recognition*, 2014, 47(9): 3179-3191.
- [39] 王功明, 李文静, 乔俊飞. 基于 PLSR 自适应深度信念网络的出水总磷预测[J]. *化工学报*, 2017, 68(5): 1987-1997.
- WANG G M, LI W J, QIAO J F. Prediction of effluent total phosphorus using PLSR-based adaptive deep belief network[J]. *CIESC Journal*, 2017, 68(5): 1987-1997.
- [40] GONG B, DU X J. Fault diagnosis of analog circuit based on multi-input convolution[J]. *Journal of Electronics and Information Science*, 2022, 7(1): 89-93.
- [41] CHOU J S, NGO N T. Modified firefly algorithm for multidimensional optimization in structural design problems[J]. *Structural and Multidisciplinary Optimization*, 2017, 55(6): 2013-2038.
- [42] ZITZLER E, DEB K, THIELE L. Comparison of multiobjective evolutionary algorithms: Empirical results [J]. *Evolutionary Computation*, 2000, 8(2): 173-195.
- [43] CHOU J S, TRUONG D N. Multiobjective optimization inspired by behavior of jellyfish for solving structural design problems[J]. *Chaos, Solitons & Fractals*, 2020, 135: 109738.
- [44] 张屹, 万兴余, 郑小东, 等. 基于正交设计的元胞多目标遗传算法[J]. *电子学报*, 2016, 44(1): 87-94.
- ZHANG Y, WAN X Y, ZHENG X D, et al. Cellular genetic algorithm for multiobjective optimization based on orthogonal design[J]. *Acta Electronica Sinica*, 2016, 44(1): 87-94.
- [45] COELLO C, PULIDO G T, LECHUGA M S. Handling multiple objectives with particle swarm optimization[J]. *IEEE Trans Evolutionary Computation*. 2004, 8(3): 256-279.
- [46] DEB K, MOHAN M, MISHRA S. Evaluating the  $\epsilon$ -Domination based multiobjective evolutionary algorithm for a quick computation of pareto-optimal solutions[J]. *Evolutionary Computation*, 2005, 13(4): 501-525.
- [47] SHAO H D, JIANG H K, ZHANG H Z, et al. Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network[J]. *IEEE Transactions on Industrial Electronics*, 2017, 65(3): 2727-2736.
- [48] 杜先君, 巩彬, 余萍, 等. 基于 CBAM-CNN 的模拟电路故障诊断[J]. *控制与决策*, 2022, 37(10): 2609-2618.
- DU X J, GONG B, YU P, et al. Research on CBAM-CNN based analog circuit fault diagnosis[J]. *Control and Decision*, 2022, 37(10): 2609-2618.
- [49] SHI Y K, WANG Z W, DU X J, et al. Research on the membrane fouling diagnosis of MBR membrane module based on ECA-CNN[J]. *Journal of Environmental Chemical Engineering*, 2022, 10(3): 107649.

编辑 叶芳