An adaptive deep belief network for fault diagnosis of complex analog circuits

Bin Gong* Aimin An

College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China

* Corresponding Author Email: gong_bin01@163.com

Abstract. Aiming at the problems of long pre-training time-consuming and poor diagnostic accuracy during unsupervised training of traditional DBN, an adaptive deep belief network (ADBN) is proposed for analog circuit fault diagnosis. The adaptive learning rate is proposed according to the similarity and difference of the parameter updating direction to improve the convergence speed of the network. The ADBN is applied to the diagnosis experiments of a two-stage four-op-amp dual second-order low-pass filter, and the experimental results show that the proposed ADBN can realize the classification and localization of difficult faults by guaranteeing the classification accuracy and higher diagnosis rate based on the fast convergence speed.

Keywords: Analog circuit, adaptive learning rate, deep belief network, fault diagnosis.

1. Introduction

Analog circuits play an important role in military, defense, and aerospace fields [1], and thus their fault diagnosis has become a research focus in the field of circuit testing [2]. However, analog circuits have complex and diverse fault phenomena, infinite fault states, and possibly continuous characteristics. In addition, the presence of nonlinear components in analog circuits makes it difficult to establish accurate mathematical models. In practice, due to various factors, traditional fault diagnosis methods such as fault dictionary method and parameter identification method have limited effect in practice[3-4].

Although signal processing methods such as wavelet transform are widely used in nonlinear system fault feature extraction and diagnosis, the essential features may be neglected in the feature extraction process, resulting in lower efficiency and accuracy of nonlinear analog circuit fault diagnosis[5].

In order to solve the above problems, data-driven artificial intelligence methods such as Support Vector Machine (SVM) [6] and Extreme Learning Machine (ELM) [7] are widely entered into the research field, providing feasible technical support for analog circuit fault diagnosis. Literature [8] proposed a fault diagnosis method for linear analog circuits by combining the circular model and the Extreme Learning Machine (ELM), and experiments show that the method accelerates and simplifies the process of fault diagnosis, providing a new method in the field of fault diagnosis. Literature [9] states that ELM has better generalization performance with fewer optimization constraints compared to SVM. However, the output of classification algorithms such as SVM and neural networks is only the fault class label corresponding to the fault sample, which cannot output other diagnostic information and faces problems such as dimensionality catastrophe. Therefore, literature [10] states that the use of shallow learning models cannot fully and thoroughly reveal the complex intrinsic relationship between the root cause of the fault and the signal features.

In recent years, deep learning such as Deep Belief Network (DBN) [11] has been widely applied in the field of fault diagnosis with its powerful learning and feature extraction capabilities [12]. Literature [13] uses chaotic particle swarm optimization algorithm to optimize the learning rate in Restricted Boltzmann Machine (RBM) to further enhance the performance of feature extraction. Literature [14] utilizes multiple two-layer Sparse Autoencoder (SAE) neural networks in which time domain features and frequency domain features are fused, and finally, DBN is utilized for further classification. Experimental results show that this method can extract effective fault features and classify them accurately. These literatures provide references for this paper in the research process.

Although Deep Belief Networks (DBNs) show a wide range of applications in several fields, the long time required for their forward training phase makes them less applicable in the era of big data with explosive information growth. To solve this problem, this paper adopts a new approach, i.e., dynamically selecting the learning rate by utilizing the similarities and differences in the parameter update directions to accelerate the model convergence and improve the stability of the model.

In summary, the main contributions of this paper are as follows:

(1) An adaptive CD algorithm is proposed to dynamically select the learning rate according to whether the parameter update direction is the same after two consecutive iterations, thus avoiding the phenomenon of convergence difficulty that may be caused by using a fixed learning rate in the unsupervised learning part of the traditional DBN.

(2) A two-stage four-op-amp dual second-order low-pass filter was built on the Matlab\Simulink platform, and simulation experiments were carried out. The experimental results show that the improved ADBN model has better performance in unsupervised learning, verifying its superiority in terms of accuracy, reliability and generalization performance.

2. DBN structure

The process of DBN feature extraction is divided into two phases: the pre-training phase and the fine-tuning phase, respectively. In the pre-training phase, all RBMs are first pre-trained unsupervised layer by layer, thus forming an unsupervised learning feature model. Next, the supervised algorithm is used for reverse training to fine-tune the initial connection weights of all RBMs, so as to reduce the errors generated by the training, which is conducive to the extraction of the essential features of the input data by DBN, the structure of which is shown in Fig. 2.

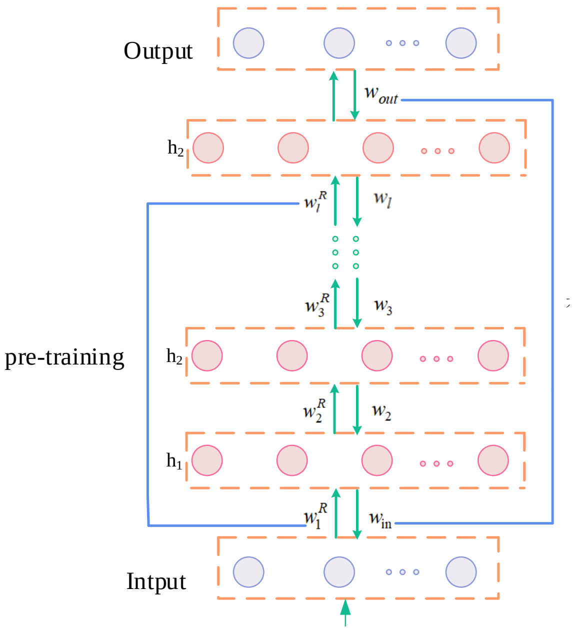

Fig. 1 DBN structure diagram

2.1. Unsupervised learning

To determine the initial weights of the network, an unsupervised training method is used to learn the parameters.1 RBM contains one visible layer and one hidden layer, denoted by v and h, respectively.

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

where $Z=\sum_{v,h}e^{-E(v,h;\theta)}$ is the normalizing factor and the model is distributed about the edges of v:

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

For an RBM with a Bernoulli (visible layer) distribution-Bernoulli (hidden layer) distribution, the energy function of the joint configuration of cells is defined as:

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

where w_{ij}^R is the connection weight of the RBM, a_i and b_j are the bias of the visible layer cell and hidden layer cell respectively.

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

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

where σ is the activation function.

The probability taking criterion for the visible and hidden layers is usually realized by setting a threshold, this is because the visible and hidden layers are Bernoulli binary states. The hidden layer, for example, can be expressed as:

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

where δ is a constant between 0.5~1.

Calculating the gradient of the log-likelihood function $\lg P(v;\theta)$ gives the formula for updating the RBM weights as:

$$w_{ij}^R=w_{ij}^R+\eta\Delta w_{ij}^R \tag{7}$$

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

where η denotes the learning rate; $E_{\text{data}}(v_i h_j)$ is the expectation of the data observed in the training set; $E_{\text{model}}(v_i h_j)$ is the expectation over the distribution determined by the model, and $E_{\text{model}}(v_i h_j)$ can be obtained by the Gibbs approximation.

3. Improved ADBN learning algorithm

3.1. Adaptive CD algorithm

Since multiple iterations of computation are required in each RBM, a fixed learning rate is prone to phenomena such as convergence difficulties. Therefore, adaptive learning rate is used to determine the learning rate according to the updating direction of the parameters. The update mechanism of adaptive learning rate is as follows:

$$\eta=\begin{cases} D\eta \quad \dfrac{\left|\left(\Delta w_{ij}^{R}\right)^{(t)}\right|+\left|\left(\Delta w_{ij}^{R}\right)^{(t+1)}\right|}{\left(\Delta w_{ij}^{R}\right)^{(t)}+\left(\Delta w_{ij}^{R}\right)^{(t+1)}+\varepsilon}=1 \\ d\eta \quad \dfrac{\left|\left(\Delta w_{ij}^{R}\right)^{(t)}\right|+\left|\left(\Delta w_{ij}^{R}\right)^{(t+1)}\right|}{\left(\Delta w_{ij}^{R}\right)^{(t)}+\left(\Delta w_{ij}^{R}\right)^{(t+1)}+\varepsilon}\neq 1 \end{cases} \tag{9}$$

$$\left(\Delta w_{ij}^{R}\right)^{(t)}=v_i^{(t)}h_i^{(t)}-v_i^{(t+1)}h_i^{(t+1)} \tag{10}$$

$$\left(\Delta w_{ij}^{R}\right)^{(t+1)}=v_i^{(t+1)}h_i^{(t+1)}-v_i^{(t+2)}h_i^{(t+2)} \tag{11}$$

where D=1.4 and d=0.7, the main purpose of which is to avoid the denominator being zero, is set to 10^{-6}.

Since multiple iterations of computation are required in each RBM, a fixed learning rate is prone to phenomena such as convergence difficulties. Therefore, adaptive learning rate is used to determine the learning rate according to the updating direction of the parameters. The update mechanism of adaptive learning rate is as follows:

$$X_{i+1}=\eta X_i\left(1-X_i\right),\ 0\leq X_0\leq 1 \tag{12}$$

where X_i is the logistic value of the ith dragonfly; X_0 is the initial value of the dragonfly population.

4. Simulation experiments and research analysis

In this paper, a two-stage quad-op-amp dual second-order low-pass filter (Fig. 2) is used with an input of a sinusoidal signal with an amplitude of 5 V and a frequency of 10 kHz for excitation. To verify the ADBN model proposed in this paper, the tolerances of the resistors and capacitors in the circuit are both set to 5%, respectively. According to the definition of parametric failure of a general component: when the parameter value of a component deviates from the nominal value by 50%, it can be determined that the component has failed. According to the results of the sensitivity test, four components, resistors R_9 and capacitors C_4, were selected as the research objects for experimental analysis, as shown in Table 1.

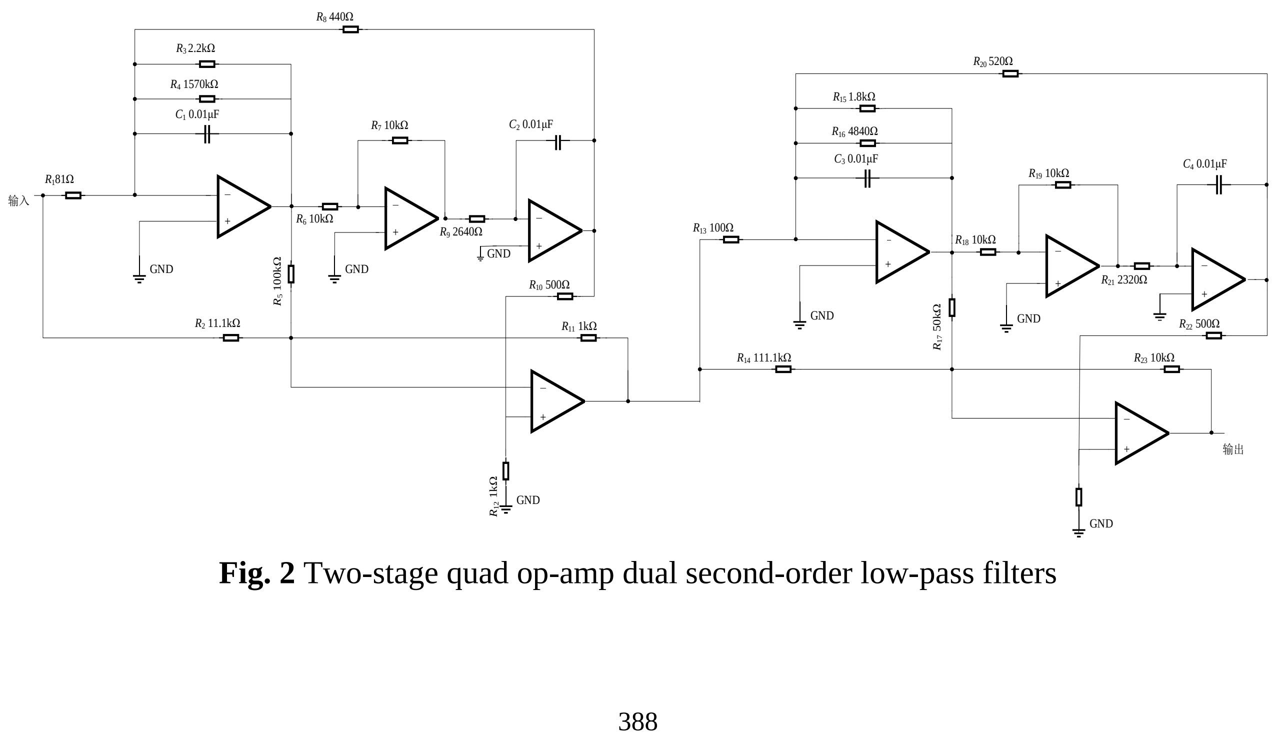

Fig. 2 Two-stage quad op-amp dual second-order low-pass filters

Table 1. Failure modes of a two-stage quad op-amp dual second-order low-pass filter circuit

Code	Type	Tolerance/(%)	Standard value	Fault value
f1	Norm	-	-	-
f2	R_9↑	5	2640Ω	3960Ω
f3	R_9↓	5	2640Ω	1320Ω
f4	C_4↑	5	0.01nF	0.005nF
f5	C_4↓	5	0.01nF	0.015nF

4.1. Comparison experiment

4.1.1. Experimental validation of adaptive learning

In order to further prove the rationality and effectiveness of the adaptive learning rate, this paper selects the fixed learning rate and the adaptive learning rate as the experimental object, and adopts the error as the experimental target for the comparison test. Among them, the fixed learning rate is set to 0.1. Table 2 represents the average diagnostic accuracy of the model in the course of 10 experiments under different learning rates. Combined with Table 2, the following results can be obtained: the accuracy rate is improved by 1.84% compared with the optimal value of 0.1 for the fixed learning rate. It can be proved that the proposed adaptive learning rate based on the setting of parameter update direction is more advanced and reasonable compared with the traditional empirical learning rate.

Table 2. Average accuracy corresponding to different learning rates

Rate	Average accuracy/%
0.1	95.12
Proposed	97.25

4.1.2. ADBN model validation

Comparison experiments are conducted between the method of this paper and some commonly used methods for analog circuit fault diagnosis and classification to demonstrate the effectiveness and superiority of the ADBN model for diagnosing and classifying complex circuit faults. On the shallow side, BP and SVM are used for classification and diagnosis, and on the deep side, traditional DBN and CDBN are used[47]. The above models are compared with the ADBN proposed in this paper in 1 independent diagnosis experiment, as shown in Table 3, where the comparison metrics include F1-Measure and average accuracy.

Table 3. Performance comparison of different models

Method	F1 Measure	Average accuracy/%
BP	0.796	75.92
SVM	0.814	82.17
DBN	0.913	91.08
CDBN	0.942	92.72
Proposed	0.9602	94.26

In terms of average correct rate, the shallow machine learning methods BP and SVM as diagnostic models have lower average correct rates than deep machine learning models. Among the deep machine learning models, DBN as a diagnostic model has the lowest average correct rate (91.08%), while the improved model of DBN, CDBN and ADBN models have average correct rates of 92.72% and 94.26%, respectively, and the ADBN diagnostic model has the best performance.

5. Summary

The unsupervised pre-training time-consumption of DBN also increases dramatically when facing a large amount of data. Aiming at the above problems, this paper proposes an analog circuit fault diagnosis method based on ADBN, which is validated by simulation experiments using a dual second-order four-operational-amplifier low-pass filter. The results of the simulation experiments show that the ADBN model is capable of efficient fault classification and localization, and can be used as a new solution in areas such as analog circuit fault diagnosis.

Acknowledgements

The authors gratefully acknowledge the financial support from the Excellent Doctoral Program of Gansu Provincial Science and Technology Plan Project (23JRRA809).

References

[1] Sai Sarathi Vasan Arvind, Long Bing, Pecht Michael. Diagnostics and Prognostics Method for Analog Electronic Circuits. IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, 2013, 60(11): 5277-5291.

[2] Ao Yongcai, Shi Yibing, Zhang Wei, et al. An approximate calculation of ratio of normal variables and its application in analog circuit fault diagnosis. Journal of Electronic Testing Theory & Applications, 2013, 29 (4): 555-565.

[3] Laurain Vincent, Tóth Roland, Piga Dario, et al. An instrumental least squares support vector machine for nonlinear system identification. Automatica, 2015, 54: 340-347.

[4] Hu Qinlei, Xiao Bing. Adaptive fault tolerant control using integral sliding mode strategy with application to flexible spacecraft. International Journal of Systems Science, 2013, 44 (12): 2273-2286.

[5] Cohen Achraf, Tiplica Teodor, Kobi A , Abdessamad. Design of experiments and statistical process control using wavelets analysis. Control Engineering Practice, 2016, 49(4): 129-138.

[6] Long Bing, Li Min, WANG Houiun, et al. Diagnostics of analog circuits based on LS-SVM using time-domain features. Circuits Systems & Signal Processing, 2013, 32(6): 2683-2706.

[7] Varatharaj Kalpana, Maheswar, R., Nandakumar, E. 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.

[8] Guo Sumin, Wu Bo, Zhou Jingyu, et al. An analog circuit fault diagnosis method based on circle model and extreme learning machine[J]. Applied Sciences, 2020, 10(7): 2386-2400.

[9] Huang GuangBin, Ding Xiaojian, Zhou Hongming. Optimization method based extreme learning machine for classification. Neurocomputing, 2010, 74(1-3): 155-163.

[10] Guo Xiaojie, Chen Liang, Shen Changqing. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. Measurement, 2016, 93: 490-502.

[11] Shao Haidong, Jiang Hongkai, Wang Fuan, et al. Rolling bearing fault diagnosis using adaptive deep belief network with dual-tree complex wavelet packet. ISA Transactions, 2017, 69: 187-201.

[12] Shelhamer Evan, Long Jonathan, Darrell Trevor. Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651

[13] Zhang Chaolong, He Yigang, Du Bolong. Analog circuit incipient fault diagnosis method based on DBN feature extraction. Chinese Journal of Scientific Instrument, 2019, 40(10): 112-119.

[14] Chen Zhuyun, Li Weihua. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. IEEE Transactions on Instrumentation and Measurement, 2017, 66(7): 1693-1702.

[15] Coello Carlos A. coello, Pulido Gregorio Toscano, Lechuga Maximino Salazar. Handling multiple objectives with particle swarm optimization[J]. IEEE Trans Evolutionary Computation. 2004, 8(3): 256-279.