A new method of power quality disturbance classification based on deep belief network

KangWang¹*, YanhuiXi²

¹School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha, Hunan, 410004, China
²School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha, Hunan, 410004, China

*Kang Wang’s e-mail: cslg@csust.edu.cn

Abstract. In view of the low accuracy of single disturbances under the problem of noise interference, a new method of power quality disturbance classification based on deep belief network was proposed. A smooth wavelet multiscale transform is performed on the power quality disturbance signal, and then the soft threshold function is used to process the estimated wavelet coefficients for reconstructing the original signal. Further, it is proposed to use deep confidence network to classify and recognize the reconstructed single disturbance signal. The simulation results demonstrate that the recognition rate of this method for seven typical single disturbances is high. Even under 20dB noise interference, the classification accuracy rate is as high as 93% or more, which proves that the method has a strong ability to resist noise interference.

1. Introduction

In recent years, the rapidly developing information and industrial society has increasingly demanded electricity consumption. The increase in impulse and non-linear load ratios has caused the issue of power quality disturbances to attract attention from all walks of life and seriously affect the safety and stability of the power system. Therefore, the key to solve the power quality problem is to quickly and accurately classify disturbances [1-2]. However, in practical applications, power quality disturbances are usually multiple power quality disturbances composed of single disturbances of different types, different amplitudes and different durations.

It is usually accompanied by noise interference, which undoubtedly causes greater difficulties in the accurate identification of disturbances. Therefore, how to remove the noise influence of disturbance signals in practical applications and achieve accurate recognition of multiple disturbances will have far-reaching significance for solving power quality problems. So far, many scholars at home and abroad have conducted in-depth research on the classification of power quality disturbances, and most of them are for a single disturbance quality classification. However, in actual power grids, multiple single disturbances often interact and combine to form multiple disturbances. Their characteristic quantities are coupled with each other, making the identification of multiple disturbances more difficult and complex. Traditional disturbance recognition methods often include feature extraction and classification. Among them, feature extraction is mainly through transforming and reconstructing the disturbance signal, and then obtaining the characteristic quantity of disturbance recognition from it. At present, the commonly used methods for extracting features include S
transform, wavelet transform, Fourier transform and Hilbert-Huang transform [3-6]. The disturbance recognition methods used today include Bayesian classifiers, support vector machines, decision trees, multi-label learning, k-nearest neighbor algorithm, and artificial neural networks [7-11]. At present, scholars have done a lot of work on the classification of single power quality disturbance and a few multiple power disturbances. Literature [12] proposed a Fischer linear discrimination method based on cross wavelet transform to classify multiple power quality disturbances. Literature [13] proposed a simple and effective method that uses S transform amplitude matrix and rule-based decision tree to detect complex power quality disturbances with different noise levels. Literature [14] proposed a method of using compressed sensing theory to obtain sparse vectors to extract signal features and then train through a deep belief network. The simulation results recognize 5 double disturbances and 2 triple disturbances, but the classification accuracy can only reach 92 %about. Aiming at the problem of low recognition accuracy of multiple disturbances (such as triple and quadruple disturbances) under noise interference in actual power grids, this paper proposes a combination of wavelet transform [15] and deep belief network (DBN) [16] Multiple disturbance recognition method. This method uses wavelet transform to reconstruct the signal waveform to retain the key and effective features in the original multiple disturbance signals, and then uses a deep belief network to classify the reconstructed disturbance signals. In order to verify the effectiveness of the method proposed in this paper, 7 single disturbances and 13 multiple disturbances are used as examples to carry out simulation analysis. The simulation results show that the deep network-based multiple disturbance classification method proposed in this paper has different noise conditions. Good robustness, high recognition accuracy, and relatively simple implementation process

2. Wavelet soft threshold denoising principle
In the process of power quality disturbance identification, it is usually affected by noise, and most of the noise is Gaussian white noise. When using wavelet to decompose the power quality signal polluted by noise, the wavelet coefficients of useful signal and noise have different characteristics, which are displayed on different scales, and the energy of useful signal and noise is distributed on the larger and smaller coefficients. It can be seen from this feature that the quantization of wavelet coefficients [17] can complete the denoising of the signal.

Suppose the one-dimensional non-stationary signal model in the actual situation is

\[ x(t) = f(t) + \varepsilon(t) \] (1)

Where \( f(t) \) is the original signal; \( \varepsilon(t) \) is the Gaussian white noise; \( x(t) \) is the noisy signal. The discrete wavelet transform [18] is

\[ x(t) \psi_{j,k}(t)dt = f(t)\psi_{j,k}(t)dt + \varepsilon(t)\psi_{j,k}(t)dt \] (2)

Where \( \psi_{j,k}(t) \) is the discrete wavelet basis function. Equation (2) can be expressed as

\[ d_{j,k} = u_{j,k} + e_{j,k} \] (3)

Where \( d_{j,k} \) is the signal containing noise, \( x(t) \) is the wavelet coefficients of each layer after wavelet transformation; \( u_{j,k} \) is the wavelet transformation coefficient of the original signal; \( e_{j,k} \) is the wavelet transformation coefficient of the noise signal.

According to statistics of these significant characteristics of useful signal and noise wavelet coefficients, an appropriate number can be obtained as a threshold by analyzing and researching them [19]. The wavelet coefficients smaller than the threshold can be considered to be caused by noise and should be removed. This part of the coefficients; and the wavelet coefficients larger than the threshold can be considered to be mainly caused by the useful signal. Keep these coefficients or shrink these coefficients by a fixed amount to approach zero to achieve the purpose of denoising. The key to wavelet threshold denoising is the size of the threshold and the choice of threshold function. The former will affect the denoising effect. When the threshold is too large, part of the useful signal will be lost; when the threshold is too small, more noise will be retained, while the latter It is a rule for
modifying wavelet coefficients, which will directly affect the high-frequency information and smoothness of the reconstructed signal [20]. This paper uses a soft threshold function [21] to denoise, in which the wavelet basis used in wavelet decomposition is coif5 [22], and the soft threshold function is

\[
    d'_{j,k} = \begin{cases} 
    \text{sgn}(d_{j,k})\{d_{j,k} | -\lambda \} & |d_{j,k}| > \lambda \\
    0 & |d_{j,k}| < \lambda
    \end{cases} \quad (4)
\]

Where \(d'_{j,k}\) is the wavelet coefficients of each layer after soft threshold quantization, \(\lambda = \sigma \sqrt{2 \ln N}\) is the general threshold, \(N\) is the signal length, \(\sigma\) is the noise standard deviation.

3. Deep belief neural network (DBN)

DBN is a deep neural network composed of several layers of restricted Boltzmann machines and a layer of BP. The RBM contains two layers of visible layer and hidden layer, which together form a probability graph model [23-25]. The neurons in the visible layer and the hidden layer are connected by weights. Generally speaking, the visible layer units are used to describe the characteristics of the data, and then the hidden layer units are used to extract the features. Assume \(m\) is that the visible layer contains the number of neurons, \(\mathbf{v} = (v_1, v_2, \cdots, v_m)^T\) is the state vector of the visible layer, \(\mathbf{a} = (a_1, a_2, \cdots, a_n)^T\) is the bias vector of the visible layer, where \(n\) is the number of neurons in the hidden layer, \(\mathbf{h} = (h_1, h_2, \cdots, h_n)^T\) is the state vector of the hidden layer, and \(\mathbf{b} = (b_1, b_2, \cdots, b_n)^T\) is the bias vector of the hidden layer. The weight between the visible layer unit \(i\) and the hidden layer unit \(j\) is set to \(w_{ij}\). The RBM model is shown in Figure 1.

![RBM model](image)

For a given set of states \((\mathbf{v}, \mathbf{h})\), the energy of the RBM system is

\[
    E(\mathbf{v}, \mathbf{h}) = -\sum_{i=1}^{m} \sum_{j=1}^{n} v_i w_{ij} h_j - \sum_{i=1}^{m} a_i v_i - \sum_{j=1}^{n} b_j h_j \quad (5)
\]

The joint probability distribution of the visible layer and the hidden layer in this group of states is

\[
    p(\mathbf{v}, \mathbf{h}) = \frac{1}{Z} e^{-E(\mathbf{v}, \mathbf{h})} \quad (6)
\]

Where \(Z\) is the normalization coefficient, defined as

\[
    Z = \sum_{\mathbf{v}, \mathbf{h}} e^{-E(\mathbf{v}, \mathbf{h})} \quad (7)
\]

RBM layers are all fully connected, and the layers are not connected. When any one of the neurons in the visible layer and the neurons in the hidden layer is given a state, the other will be activated and
meet the conditions independent. From this we can get the probability of activating hidden layer neurons as

\[ p(h_j = 1 | v) = \sigma(\sum_{i=1}^{n} v_i w_{ij} + b_j) \] (8)

The probability that the neurons in the visible layer are activated is

\[ p(v_j = 1 | h) = \sigma(\sum_{i=1}^{n} w_{ij} h_i + a_j) \] (9)

Where \( \sigma \) is the Sigmoid activation function, namely

\[ \sigma(x) = \frac{1}{1 + \exp(-x)} \] (10)

The DBN used in this article is composed of three layers of RBM and one layer of BP neural network. The three-layer RBM performs unsupervised feature learning, trains each layer of RBM independently, and trains layer by layer from low to high until the training of all RBM layers in the DBN is completed. The last layer is connected to the BP neural network, and then the DBN is fine-tuned through the backward propagation algorithm and the category label data.

4. Algorithmsimulation ation and analy sis

4.1. Power quality disturbance model

In order to facilitate comparison and analysis, according to the IEEE power quality standards, Swell, Interruption, Sag, Harmonic, Oscillatory Transient, Impulse, Flicker and other seven typical disturbance signals are constructed separately, and their mathematical expressions are shown in Table 1.

| Disturbance type | Model |
|------------------|-------|
| Swell (C1)       | \(1 + a[u(t-t_1) - u(t-t_2)]\) \(\sin \omega_t t\) |
| Interruption (C2)| \(1 - a[u(t-t_1) - u(t-t_2)]\) \(\sin \omega_t t\) |
| Sag (C3)         | \(1 - a[u(t-t_1) - u(t-t_2)]\) \(\sin \omega_t t\) |
| Harmonic (C4)    | \(\sin \omega_t t + a_2 \sin 3\omega_t t + a_3 \sin 7\omega_t t\) |
| Oscillatory Transient (C5) | \(\sin \omega_t t + a \exp[-(t-t_i)]\) |
| Impulse (C6)     | \(\sin \omega_t t = a[u(t-t_1) - u(t-t_2)]\) |
| Flicker (C7)     | \((1 + a \sin b\omega_t t)\) \(\sin \omega_t t\) |

According to the actual working conditions, this article generates the above 7 types of single disturbances and 13 types of multiple disturbances synthesized from the above 7 types of single disturbances, each with 1800 sets of sample data, of which 1000 sets of training sample data, and the remaining 800 sets as test samples data. The single disturbance and multiple disturbances train the network separately according to the interference of different signal-to-noise ratio noise.
4.2. Network parameter setting
Set the input of the deep belief network to 1024 data points, the output layer is K-dimensional, that is, the output of K-type disturbance, and the sigmoid activation function is selected. Single and multiple power quality disturbance DBN hierarchical structures adopt 200-100-100 and 1000-100-50 respectively. Among them, the number of iterations in the unsupervised learning stage of DBN is 200, and the learning rate is 0.1. The number of iterations in the supervised learning phase is 200, and the algorithm used is gradient descent.

4.3. Disturbance recognition simulation example
For seven types of typical single disturbances, considering that the actual working conditions contain noise interference, the disturbance signals are respectively subjected to noise with a signal-to-noise ratio of 20dB, 30dB, and 40dB, and then the disturbances under various noise interference conditions are subjected to wavelet transformation. Structure. At the same time, in order to test the anti-noise ability of the method proposed in this article, when the training data is interfered by noise to a certain extent (such as 40dB), then various disturbance signals with different signal-to-noise ratios (40dB, 30dB, 20dB) are tested respectively. The training samples and test samples are 7000×1024 training data sets and 5600×1024 test data sets, respectively, which are trained and tested through the deep belief network. The classification test results are shown in Table 2-4.

It can be seen from Table 2-4 that for the seven typical single disturbances, the average accuracy under different noise conditions exceeds 93%. Therefore, the proposed method has higher classification ability and is less affected by noise. In addition, the highest accuracy in Table 2, Table 3, and Table 4 are 98.71%, 100%, and 100%, respectively. It can be seen that the highest accuracy increases with the increase of SNR. In the three cases, when the noise level of the test data is the same as the noise level of the training data, the proposed method can obtain the best performance, and when the noise level of the test data is closer to the noise level of the training data, the performance will change better.

| Disturbance type | Classification accuracy (%) | SNR=40dB | SNR=30dB | SNR=20dB |
|------------------|-----------------------------|----------|----------|----------|
| C1               | 100                         | 100      | 100      | 100      |
| C2               | 100                         | 100      | 100      | 100      |
| C3               | 92.5                        | 86.5     | 89       | 100      |
| C4               | 100                         | 100      | 100      | 100      |
| C5               | 98.5                        | 100      | 100      | 100      |
| C6               | 100                         | 100      | 100      | 98       |
| C7               | 100                         | 100      | 100      | 100      |
| Average          | 98.71                       | 98.07    | 98.42    | 99.71    |

| Disturbance type | Classification accuracy (%) | SNR=40dB | SNR=30dB | SNR=20dB |
|------------------|-----------------------------|----------|----------|----------|
| C1               | 100                         | 100      | 100      | 99       |
| C2               | 100                         | 100      | 100      | 87       |
| C3               | 100                         | 100      | 100      | 100      |
| C4               | 100                         | 100      | 100      | 100      |
| C5               | 99.5                        | 100      | 100      | 94.5     |
| C6               | 100                         | 100      | 100      | 76       |
Table 4. Classification accuracy when training data is disturbed by 40dB noise.

| Disturbance type | Without noise | SNR=40dB | SNR=30dB | SNR=20dB |
|------------------|---------------|----------|----------|----------|
| C1               | 100           | 100      | 100      | 98       |
| C2               | 100           | 100      | 100      | 84.5     |
| C3               | 100           | 100      | 100      | 99       |
| C4               | 100           | 100      | 100      | 100      |
| C5               | 99.5          | 100      | 100      | 97       |
| C6               | 100           | 100      | 100      | 73.5     |
| C7               | 100           | 100      | 100      | 100      |

Average 99.92 100 100 93.78

It can be seen from Table 2-4 that for the seven typical single disturbances, the average accuracy under different noise conditions exceeds 93%. Therefore, the proposed method has higher classification ability and is less affected by noise. In addition, the highest accuracy in Table 2, Table 3, and Table 4 are respectively 98.71%, 100%, and 100%. It can be seen that the highest accuracy increases with the increase of SNR. In the three cases, when the noise level of the test data is the same as the noise level of the training data, the proposed method can obtain the best performance, and when the noise level of the test data is closer to the noise level of the training data, the performance will change better.

4.4. Comparison of disturbance recognition accuracy

In order to verify the effectiveness of the method in this paper, the classification results are compared with those in the literature [1] (see Table 5). Among them, the document [1] used compressed sensing to extract signal features and then combined with BP neural network for training.

Table 5. Comparison of disturbance recognition accuracy.

| Methods            | Classification accuracy (%) |
|--------------------|---------------------------|
|                    | Without noise | SNR=40dB | SNR=30dB | SNR=20dB |
| the document [1]   | 99.43          | 98.86    | 98.57    | 96.43    |
| Proposed method 20dB| 98.71          | 98.07    | 98.42    | 99.71    |
| 30dB               | 99.92          | 100      | 100      | 93.78    |
| 40dB               | 99.92          | 100      | 100      | 93.14    |

It can be seen from Table 5 that the method proposed in this paper is better than other methods, especially in the case of high noise (20dB and 30dB), the accuracy of single disturbance and other methods is relatively high. In particular, in Table 5, literature [1] does not give the noise intensity of the training data. The method in this paper has carried out simulation verification on the training data of different noise intensity (20dB, 30dB, 40dB).

5. Conclusion

This paper proposes a deep belief neural network multi-disturbance classification method based on wavelet transform. In this method, the original disturbance signal is reconstructed by wavelet transform and classified and recognized by DBN. This method is suitable for single disturbance recognition after pollution with different noise levels. Simulation comparison analysis results show
that the depth classification method has high accuracy and good robustness to noise even in a high-noise environment.

Acknowledgments
I hope to thank the National Natural Science Foundation of China (51507015, 61673388), Hunan Natural Science Foundation (2018JJ2439), Hunan Provincial Department of Education Outstanding Youth Project (18B130) for financial support.

References
[1] Cao Sivang, Dai Chaohua, Zhu Yunfang, et al. (2017) A novel compressed sensing-based recognition method for power quality disturbance signals. Power System Protection & Control, 45(3): 7-12.
[2] WANG Renming, WANG Hongyang, ZHANG Yunning, et al. (2020) Composite power quality disturbance recognition based on segmented modified S-transform and random forest. Power System Protection & Control, 48(7): 19-28.
[3] YANG Jianfeng, JIANG Shuang, SHI Gege. (2019) Classification of composite power quality disturbances based on piecewise-modified S transform. Power System Protection and Control, 5(1): 64-71.
[4] GUAN Weiguo, YAO Qingzhi, GAO Yang, et al. (2015) Transient power quality detection and location of distribution network based on db4 wavelet transform. Power System Protection and Control, 43(8): 765-774.
[5] Chen Heyang, Zheng Wendi, Wu Wenxuan. Harmonic analysis of DC system using DC component recovery method. Electrical Measurement & Instrumentation, 1-8 [2020-09-23]. http://kns.cnki.net/kcms/detail/23.1202.TH.20191021.0906.004.html.
[6] Zhang Yang, Liu Zhigang, (2011) Application of EEMD in power quality disturbance detection. Electric Power Automation Equipment, 31(12): 86-91.
[7] WU Junfeng, LI Yuzhe, QUEVEDODE, et al. (2015) Data-driven power control for state estimation: a Bayesian inference approach. Automatica, 54: 332-339.
[8] Zhao Luoyin, ZHUANG Lei, DING Jianshun, et al. (2020) Identification of power quality hybrid disturbances based on S&TT transform and PSO-SVMs. Electrical Measurement & Instrumentation, 57(04): 78-86.
[9] Ma Xiangyun, Bao Yanhong, Zhang Jinlong, et al. (2019) Transient stability assessment based on support vector machine and decision function. Electrical Measurement & Instrumentation, 56(23): 48-53.
[10] ZHANG Qiaoge, LIU Zhigang, ZHU Ling, et al. (2013) Recognition of multiple power quality disturbances using multi-label wavelet support vector machine. Proceedings of the CSEE, 33(28): 114-120.
[11] KUMAR R, SINGH B, SHAHANI D T, et al. (2015) Recognition of power-quality disturbance using S-transform-based ANN classifier and rule-based decision tree. IEEE Transactions on Industry Applications, 51(2): 1249-1258.
[12] Dalai S, Dey D, Chatterjee B, et al. (2015) Cross-spectrum analysis-based scheme for multiple power quality disturbance sensing device. IEEE Sensors Journal, 15(7): 3989-3997.
[13] Kubendran A K, Loganathan A K. (2017) Detection and classification of complex power quality disturbances using S-transform amplitude matrix-based decision tree for different noise levels. International Transactions on Electrical Energy Systems, 27(4).
[14] Chen Wei, He Jiahuan, Pei Xiping, et al. (2018) Application of deep belief network in power quality compound disturbance identification. Proceedings of the CSU-EPSA, 30(9): 75-82.
[15] Wu Zhaogang, Li Tangbing, Yao Jiangang, et al. (2014) Power quality disturbance classification based on a wavelet and improved neural tree. Power System Protection and Control, 42(24): 86-92.

[16] Ma Jian, Chen Kexu, Xiao Luxin, et al. (2016) Classification on mixed disturbances of power quality based on restricted Boltzmann machine. Journal of Nanchang University (Natural Science), 40(1): 30-34.

[17] Huang Jianzhao, Xie Jian, Gao Qinhe, et al. (2012) Research of pressure fluctuation signal denoising based on translation invariance wavelet and improved threshold method. Machine Tool & Hydraulics, 40(15): 161-164, 144.

[18] Wu Chunming, Xie Nina. (2012) Application of improved wavelet threshold in power quality signal denoising. Computer Engineering and Applications, 48(3): 114-116.

[19] Liu Weidong, Liu Shanghe, Hu Xiaofeng, et al. (2007) Analysis of modified methods of wavelet threshold de-noising function. High Voltage Engineering, 33(10): 59-63.

[20] Wu Zhivu, Zhu Yunfang, Hou Yishuang, et al. (2019) Power quality disturbance recognition method based on wavelet compressive sensing. Proceedings of the CSU-EPSA, 31(5): 1-7.

[21] Wang Weibo, Dong Ruiying, Zeng Wenru, et al. (2019) A wavelet de-noising method for power quality based on an improved threshold and threshold function. Transactions of China Electrotechnical Society, 34(2): 409-418.

[22] Wang Hongwei, Rong Hang, Sun Peng. (2014) Research on application of Coif5 wavelet to locate acoustic emission source in coal-rock. Computer Engineering and Applications, 50(23): 215-219.

[23] Zhu Yunfang, Wu Zhiyuan, Gao Yan, et al. A multi-sparse power quality disturbance depth recognition method. Journal of Southwest Jiaotong University.

[24] Li Danqi, Mei Fei, Zhang Chenyu, et al. (2020) Deep belief network based method for feature extraction and source identification of voltage sag. Automation of Electric Power Systems, 44(04): 150-160.

[25] Xu Chunhua, Chen Kexu, Ma Jian, et al. Recognition of power loads based on deep belief network. Transactions of China Electrotechnical Society.