International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-5, Issue-8)

Research Article

August 2016

Generalized Regression Neural Network and Wavelet Transform for Transformer Protection

Jonnavarapu Narapareddy, Harish Balaga*

Department of Electrical & Electronics Engineering, GMR Institute of Technology,
Andhra Pradesh, India

Abstract-

In this paper, a pattern classification proposed for transformer differential protection using wavelet packet transform (WPT) and artificial neural networks (ANN). WPT is a prominent signal processing technique, has a fast, reliable and computationally effective tool. WPT is use to distinguish between the magnetizing inrush and internal fault currents. The differential current data generated from a 100 MVA, 220/100kv stepdown transformer connected to a 100km transmission line and three phase RLC load. For the classification of patterns, Pattern Recognition neural network (PRNN) and Generalized Regression neural network (GRNN) are proposed. PRNN achieved the good recognition rate for few algorithms but GRNN produced effective recognition rate with less time spent. GRNN has more accurately prediction capability and better classification performance than the PRNN, in application of transformer differential protection. The classification rate is 100% by using GRNN approach. The performance of the network explained in terms of linear regression and mean squared error values.

Keywords- powertransformer protection, wavelet transform, Daubechies mother wavelet, PRNN, GRNN.

I. INTRODUCTION

In a Power Transformer, the fault diagnosis and classification is an important task for avoiding mal-operation of the relay. For pattern classification signal analysis is the one of the most useful method, for signal analysing many good techniques have been proposed such as higher order Statistics (HOS) and short time Fourier transform (STFT). Even though, they are good at signal analysing but they have the drawbacks of constant time-frequency resolution. The fault diagnosis and classification system of fault and inrush signals analysis is stressed in the variable time-frequency information domain. On considering the drawbacks of the above techniques and also this application necessity, wavelet transform (WT) is the best tool for providing a variable time-frequency resolution analysis.

The powerful wavelet transform can be divided into a continuous wavelet transform (CWT) and discrete wavelet transform (DWT). CWT has a huge operand and for usage it takes long time. Hence, the DWT has developed to overcome drawbacks of CWT. The DWT decomposes the original signal into the several resolutions. In paper[1], multi resolution analysis of DWT is applied to get the approximated and detailed coefficients at each level of decomposition. The changes in differential currents are diagnosed in [2], over a specific frequency band by applying the median absolute deviation (MAD) criteria for wavelet coefficients. Most of the researchers [1]–[5], used the Daubechies wavelet family has the optimum mother wavelet family because of its nature of transient signal adaptability. Some of the authors [4]–[6], used db4 as the optimum mother wavelet with level 4 for transformer differential current feature extraction because, most of the energy of the signal concentrated in between level 1 and level 4. In articles [3], [4] the transformer digital differential protection was presented based on the correlation coefficient (CC) of DWT. If the CC is greater than of 0.8 it represents as inrush otherwise indicates as fault case. For high accuracy transient phenomena identification, a combination of wavelet transform and support vector machine (SVM) scheme is proposed in [5].

Artificial neural networks (ANNs) use has grown drastically from last few years. The main reason is neural networks provide a novel modern approach solution to problems for which the traditional mathematics, methodologies and algorithms are unable to produce to a satisfactory solution. Such neural networks can deal effectively the complex problems. Hence, in paper [6], multilayer feed forward neural network is used for differential protection. But these feed forward networks have the drawbacks of time consuming learning. To produce superior results and to reduce the time consuming problem Genetic algorithm (GA) based neural networks are used in [7]. To reduce the processing and computing time, in paper [8], the single ANN is split into two parallel networks of master-slave method of network. Nonetheless, this parallel ANN can be applied only whenever there is a working provision of two ANNs. For solving these problems, a well suited Generalized regression neural network (GRNN) was developed in article[9]. In paper [10], the GRNN performance is compared with FFNN in case of internal engine combustion system. In power system the transient stability condition plays a vital role it is predicted with high accuracy in article [11]. In paper [12], GRNN is used for indicating whether it is the fault or normal condition of generator system.

In the present study, wavelet adaptive neural network system was proposed for feature extraction and classification of transformer differential currents. The principle component work of this article divided into the sections. Section 1 introduces about this article and the main research contents of the reference papers. In section 2 power system network simulation is described. Section 3 consists of wavelet analysis of differential current signals of transformer. Proposed PRNN and GRNN and their results are described in section 4. Conclusion conferred in section 5 followed by corresponding references.

II. POWER SYSTEMNETWORK SIMULATION

The fault and inrush currents are generated from the Simulink circuit shown in Fig. 1. The circuit consisting of a three - phase 100 MVA,220/100kv step-down transformer connected to a 160km transmission line and a three – phase 50 MVA series RLC load operating at various conditions.Fig. 2 (a)Shows the line – ground fault created at 0.04-0.08 sec in phase –A, while phases B and C are healthy.

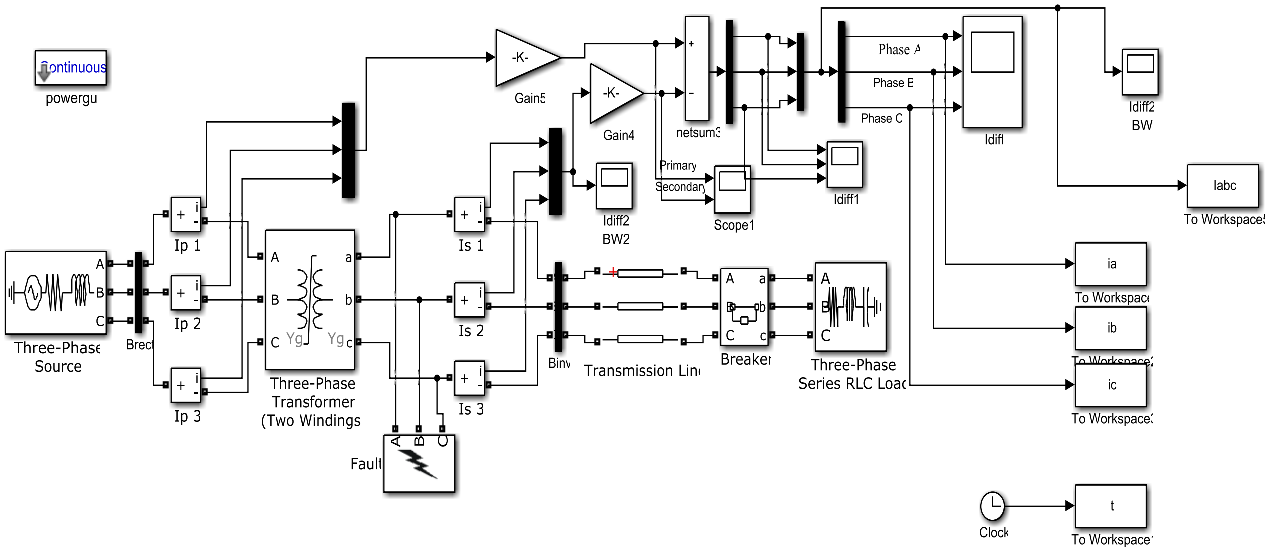


Fig. 1. Simulated model of 3 - φ Power system

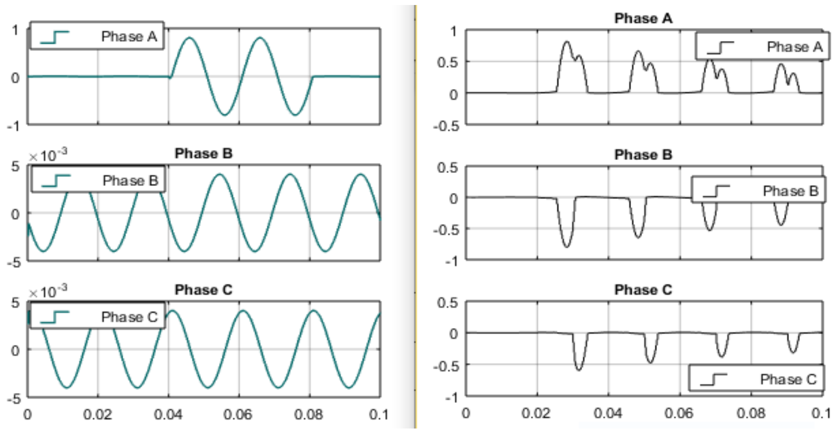


Fig. 2. Differential currents of (a) LG fault (phase-A) (b) Inrush

A. Inrush phenomena of transformer:

The magnetizing inrush current flows in primary winding when the secondary winding of the transformer is open circuited just after closing the primary side switch. Inrush current waveforms are shown in theFig. 2 (b). Inrush current is generally 10-15 times larger than the normal rated load current and flows for few cycles. The inrush current is effected by:

1. Energization voltage wave point instant.
2. Transformer core non-linear characteristics.
3. Transformer core residual flux linkage sign and value.

III. WAVELET TRANSFORM

It is difficult to use the windowed signals as the input to the ANN because of large number samples lead to risk of training. To overcome this problem wavelet decomposition is one of the superior tool for extract the features of the transient signals and to provide a better reliable data of discrimination. The wavelet transform (WT) analysis the distorted signal into different time-frequency scales. The wavelet transform can be divided into a continuous wavelet transform (CWT) and discrete wavelet transform (DWT). CWT has a huge operand and for usage it takes long time. Hence, DWT transform has developed to overcome the drawbacks of the CWT. The DWT decomposes the original signal into the several resolutions. TheDWT uses the wavelet function and scaling function to perform simultaneously decomposition and reconstruction of the signal in multiresolution analysis (MRA). The wavelet function generates the high frequency components in detailed version and the scaling function generates the low frequency components in approximated version. WT is a well-suited tool for non-periodic and non-stationary wideband signals.

Multi Resolution Analysis (MRA):

MRA is a main characteristic in WT, it can decompose the actual signal into several other signals at different levels of resolution. The actual time domain signal can reproduce from the decomposed signal without losing any information. The DWT using MRA is shown in belowFig. 3.

MRA can be represented in mathematical form as:

$$A_i = d_{i+1} + A_{i+1} = d_{i+1} + d_{i+2} \ldots\ldots\ldots + d_{i+n} + A_n \quad (1)$$

Where

A_{i+1} isthe approximated version of signal at level i+1. n is the decomposition level,

d_{i+1} is the detailed version of the given signal that displays all transient phenomena at level i+1.

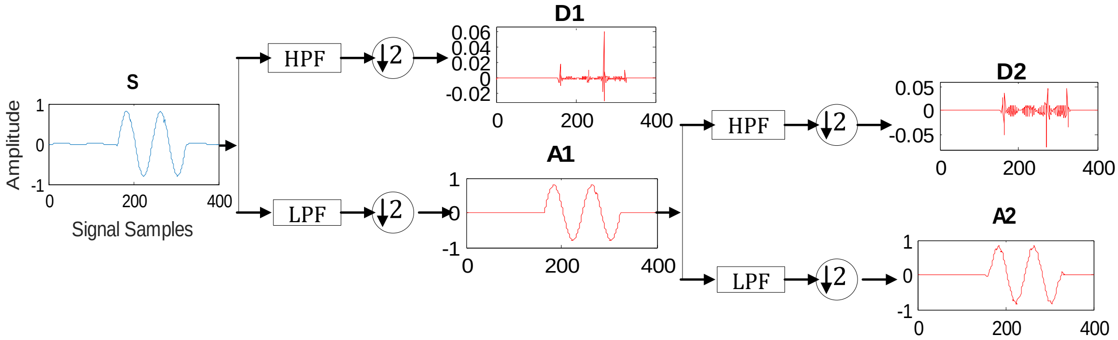


Fig. 3. DWT decomposition

In WT the optimum mother wavelet plays an important role in finding and localizing of different fault transients. Most of the researchers concluded the mother wavelet as Daubechies for transient signals. Daubechies wavelet has the characteristics of orthogonality and special low-pass and high-pass filters on compare with Haar, Coiflet and Meyer wavelets.The decomposition process can be iterative, one signal broken down into many other resolution components. It can be continued indefinitely theoretically but in practical this process can be proceeding to the individual details consist of one sample only. An optimum number of levels will be selected based on the signal nature. In Daubechies family db4 chosen as the mother wavelet over level 3 because of the maximum energy of signal localization in details was obtained at these parameters. It is important to known that each level of frequency band is directly related to sampling frequency. According to Nyquist's criteria, if a signal has a sampling frequency f_s then it can have the highest frequency component is half of the sampling frequency. The first level of decomposition can cover the frequency band from $f_s/2$ to $f_s/4$, the second detail level covers from $f_s/4$ to $f_s/8$ and this process continues up to a defined level of decomposition. The approximated and detailed frequency bandwidths are given in equations (2) and (3) respectively. The sampling frequency of the signal is 4 kHz corresponding to 80 samples per cycle.The corresponding decomposition of fault and inrush currents are shown inFig. 4. The DWT decomposition bandwidths are tabulated in TABLE I.

$$A_n = [0, \frac{f_s}{2^{n+1}}] \quad (2)$$

$$D_n = [\frac{f_s}{2^{n+1}}, \frac{f_s}{2^n}] \quad (3)$$

Table II Decomposition of Frequency Bandwidth up to Level-3

Decomposition Level	Approximation bandwidth (kHz)	Detail bandwidth (kHz)
1	A1 - (0-1000)	D1 - (1000-2000)
2	A2 - (0-500)	D2 - (500-1000)
3	A3 - (0-250)	D3 - (250-500)

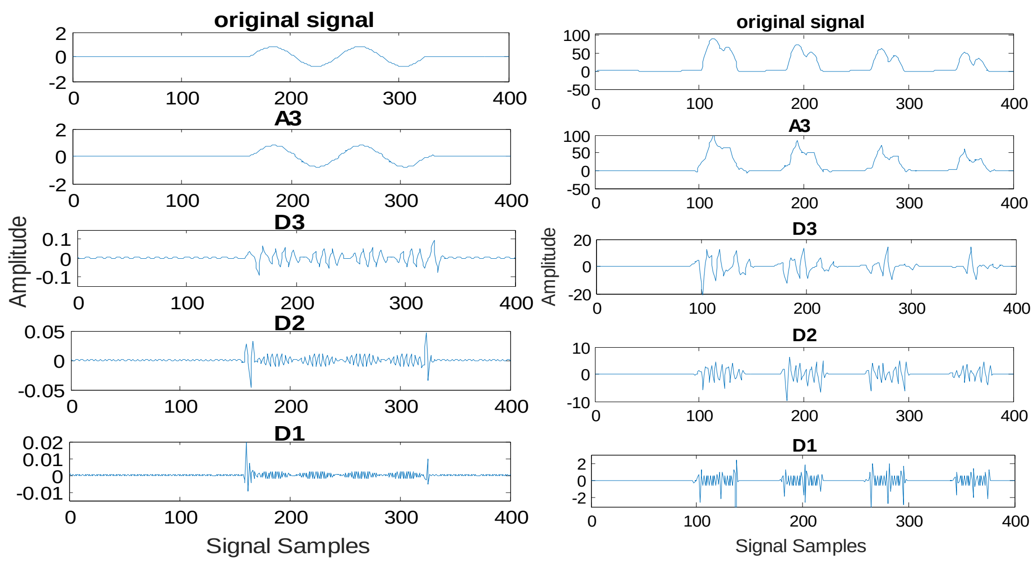


Fig. 4. Wavelet decomposition of (a) fault (b) inrush currents

IV. PROPOSED ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANNs) use has grown drastically from last few years. The main reason is neural networks provide a novel modern approach solution to problems for which the traditional mathematics, methodologies and algorithms are unable to produce to a satisfactory solution. In this work, Pattern recognition neural network and Generalized regression neural network are proposed for pattern classification. These two models have the topology of 12 inputs and 3 outputs. By using the extracted data, the training and testing of neural networks are conducted. The output codification of neural networks is mentioned in belowTABLE IIII.The total data is arranged in moving window format, have the 644 sets of data. This data is employed commonly for both neural networks.

TABLE IVI Output Codification of ANNs

Output case	Output pattern (binary)
Normal	1 0 0
Fault	0 1 0
Inrush	0 0 1

A. Pattern Recognition neural network:

Pattern recognition neural networks (PRNN) are feed forward neural networks that can be used to classify the inputs with corresponding target classes. Neurons organization of pattern network can be categorized into three parts. The foremost part is the input layer, receives inputs from real world. The second part is generally known as hidden part with one, two or multiple layers of neurons exist depending upon the problem severity and generalization requirements. In second part each input is forward to its succeeding layer where it is treated. The last layer is output layer produces the outputs to the real world. The architecture of pattern neural network is shown in below figure. The optimal process of ANN is that when the inputs are employed to input neurons, the network executes a summation of the weight factors and activates one or more particular neurons which are most suitable solution for the given problem. The architecture of pattern recognition neural network is shown inFig. 5consisting of input, hidden and output layers.

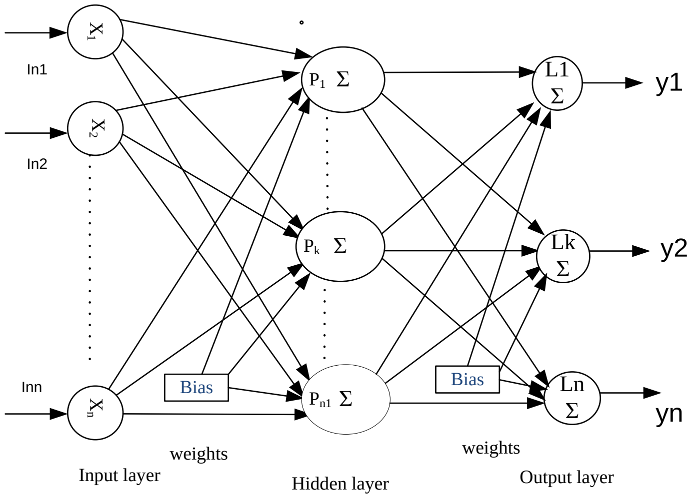


Fig. 5. Architecture of pattern recognition neural network

The training of a supervised network can be done by applying typical input patterns along with corresponding expected output patterns. The connection weights between the neurons are modified by an error measurement between produced and expected results, until the network results are satisfactory. For this process a group of algorithms, (a) Gradient descent algorithm (Gd) (b) Scaled conjugate gradient backpropagation algorithm (Scg) (c) Bayesian Regularization back propagation algorithm (BR) (d) Levenberg-Marquardt back propagation algorithm (LM)(e) Resilient Backpropagation algorithm (Rprop) are used. They compute the weight modification to improve the actual network results in order to provide a most accurate solution to the problem.

*1) Gradient-descent algorithm:*This algorithm updates the weights and biases according to gradient descent method. The results of the network on using this algorithm are mentioned in 0 and the best of network results produced when network consisting of 100 hidden neurons as shown in Fig. 6(a).

*2) Scaled conjugate gradient backpropagation algorithm:*This algorithm (SCG) can train any network as long as its net input, weights and transfer functions are having derivative functions. It uses backpropagation to calculate the derivatives of the performance with respect to weights and bias variables. The results of the network are mentioned in the0The network providing the best results having hidden neurons of ten, and shown in Fig. 6 (b).

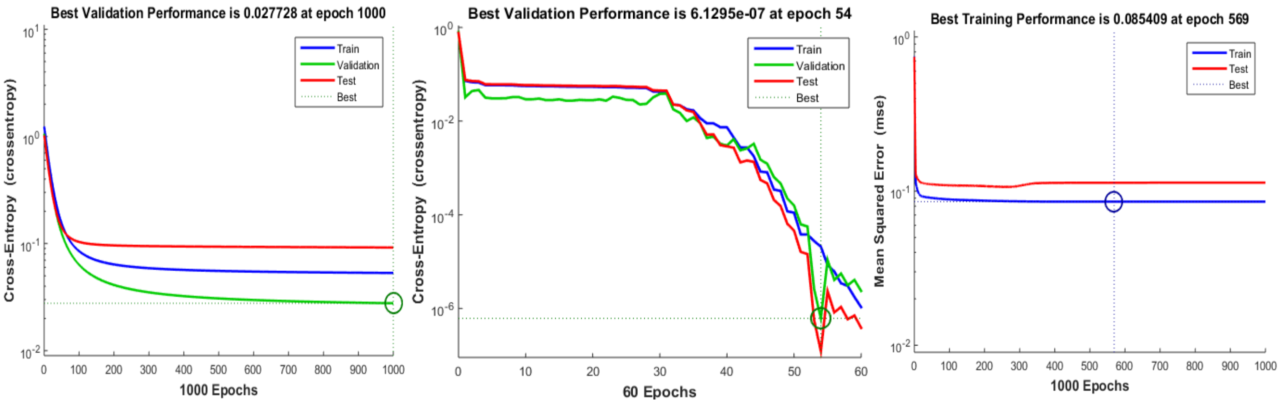
*3) Bayesian Regularization back propagation algorithm:*This algorithm (BR) updates the weights and biases according to Levenberg – Marquardt optimization. It minimizes the combination of weights and squared errors then predicts the correct combination so as to produce a well generalized network. The results of the pattern network when employed with BR algorithm are mentioned in 0 and the best of the network results on using this algorithm shown in Fig. 6 (c).

*4) Levenberg-Marquardt back propagation algorithm:*This algorithm (LM)does not require more memory than other algorithms. Based on Levenberg -Marquardt optimization it updates the weights and biases of the network. The results of pattern network when having the LM algorithm as training algorithm are mentioned in 0The best optimized results are produced when the network having 10 hidden neurons and it is shown in Fig. 6 (d).

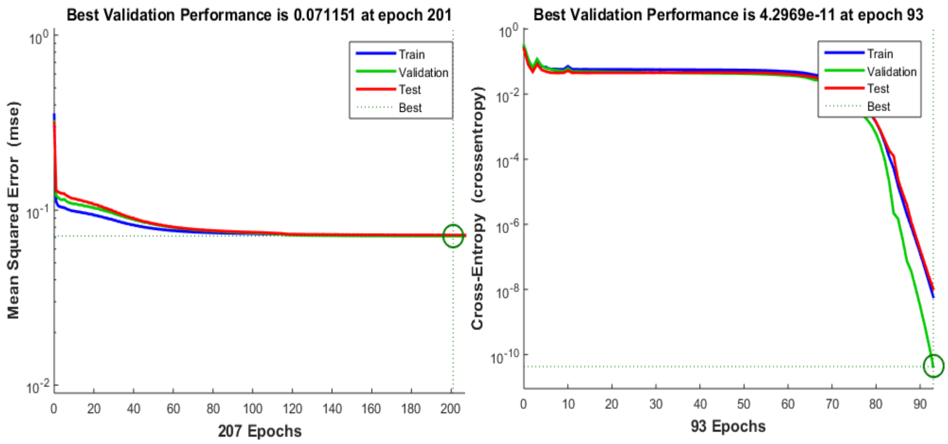
*5) Resilient Backpropagation algorithm:*In hidden layers, multilayer networks use sigmoid transfer functions, they squash an infinite range of inputs into finite range of outputs. Sigmoid transfer functions are characterized by the fact of their slopes must and should approach to zero. To train a multilayer network this causes a problem while using steepest descent with sigmoid functions because, the gradient descent has the smaller magnitude can cause small changes only in weights and biases even in case of weights and biases are far from the optimal values. To overcome the disadvantages of GD algorithm, Rprop algorithm was pioneered by Martin Reidmiller. The purpose of Rprop is to eliminate these harmful effects of partial derivative magnitudes. In Rprop the weights updating direction can only determine by the derivative sign, but not the effect of derivative magnitude. Separate update value can determine the changes in the weights size.The results of Rprop algorithm at different neurons are mentioned in 0 and the best one of results are shown in Fig. 6 (e).

Table VII Pattern Network Results of Different Algorithms

ANN topology	Rprop (mse)	SCG (mse)	LM (mse)	GD (mse)	BR (mse)
12-10-3	4.29e-11	6.12e-7	0.0711	0.0285	0.0856
12-20-3	3.05e-7	4.28e-5	0.0813	0.0341	0.0874
12-30-3	5.22e-10	3.07e-5	0.0730	0.0367	0.0854
12-40-3	4.53e-5	4.10e-6	0.0781	0.0287	0.0863
12-50-3	1.81e-8	5.04e-5	0.0704	0.0309	0.0868
12-70-3	6.96e-9	8.69e-6	0.0702	0.0313	0.0867
12-100-3	2.55e-7	1.73e-6	0.0730	0.0277	0.0885



(a) GD with 100 HN (b)SCG with 10 HN (c)BR with 30 HN



(d) Lm with 10 HN (e)Rprop with 10 HN

Fig. 6. PRNN results at respective algorithms (HN=hidden neurons)

B. Generalized regression neural network:

In the ANN so far, pattern recognition neural network (PRNN) is the most representative and basic type of the neural network. In PRNN, the error is used to adjust with recurring learning till the error is smaller than the threshold or expecting value. PRNN has the fast recall, high learning precision and wide application. Albeit, it has these merits PRNN has some flaws such as: improper learning rate, local minimum existence. To over come these drawbacks, few ANNs have been developed. In 1991 specht, proposed a neural network called as Generalized regression neural network (GRNN). The GRNN is very fast, because it is no need of an iterative training for converging the wanted solution. Fig. 7shows the architecture of GRNN contains of three layers: Input layer, Hidden layer (Radial basis layer) and Output layer (Special linear layer). Hidden layer units are connected through weights to two summation neurons namely, S- summation and D- summation. The S- summation neuron evoluates the sum of weighted outputs of hidden layer where as the D- summation neuron computes the weighted outputs of pattern neurons. From input cell to the hidden layer, GRNN adopts the direct mapping between the hidden and output layers adopts the mapping mode as linearly weighted sum of hidden layers. GRNN needs only the method of probability density function to be presented which is different from traditional regression analysis method that needs to suppose exact function.

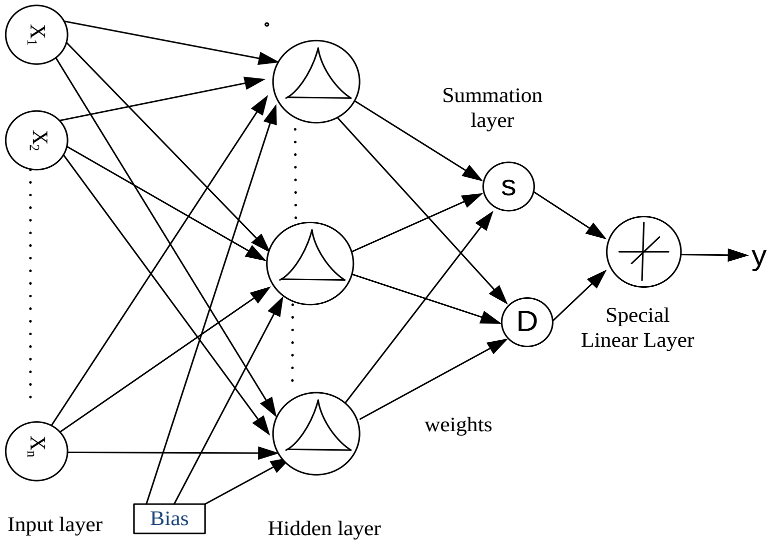


Fig. 7. Architecture of GRNN

C. Results of GRNN:

The proposed WPT-ANN, fault classification systemis useful for pattern classification in various working conditions of Power Transformer. The total number of pattern data sample stes are 644. In those 544 sets are used to training and the remaining 100 sets are choosen randomly by network for testing, of the recognition rate of the proposed network. The recognition rate is defined as the total number of tests and the correct classification. PRNN achieved the good recognition rate for few algorithms but GRNN produced effective recognition rate with less time spent. The training recognition rate of GRNN is 100% where as, the testing recognition rate is 99.99% accurately. The training and testing classification of GRNN is shown in Fig. 8

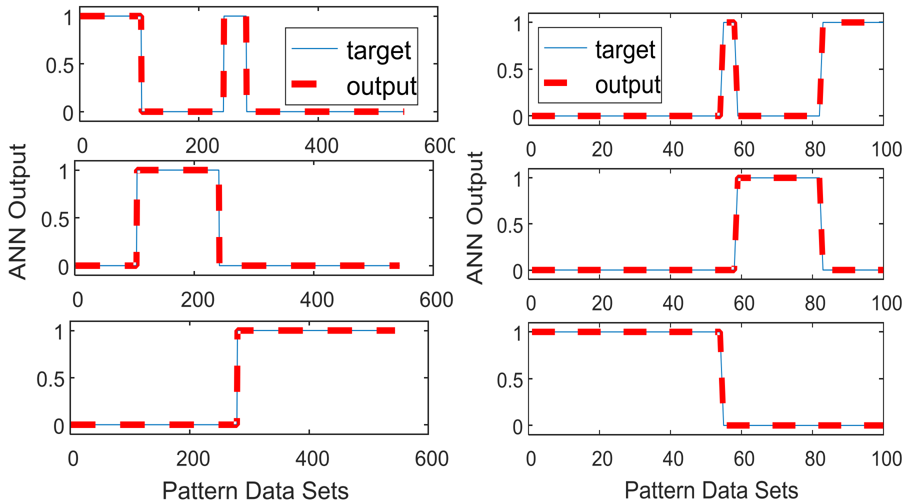


Fig. 8. GRNN pattern classification during (a) training case (b) testing case

*1) Error Histogram:*The error histogram represents the mean square error (mse) values of the network drawn between the network targets and outputs. In case of training the number of instances are 1632 (544 sample sets x 3 output), where as in case of testing the instances are 300 (100 x 3). Fig. 9(a) shows the training error histogram with 30 bins and (b) represents the testing error histogram with 20 bins respectively. The mse values for both training and testing cases are 8.881e^{-16} and 2.22e-^{16} respectively.

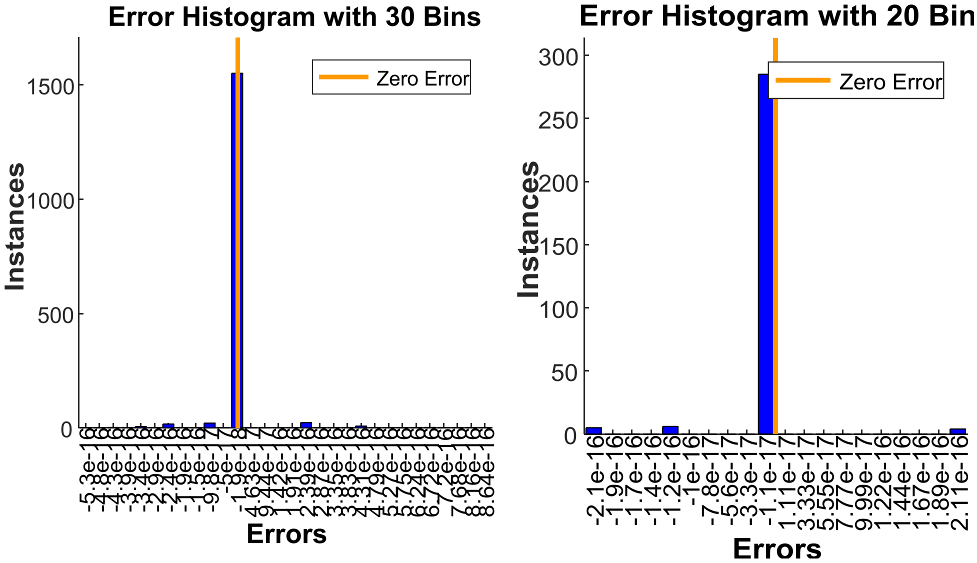


Fig. 9. (a) Training Error Histogram (b) Testing Error Histogram

*2) Regression analysis:*The linear regression analysis is the commonly used basic predictive analysis to represent the fitness relationship between the output and input variables. Generally linear regression value varies in betwwen 0-1.Fig. 10 represents the regression analysis of GRNN for both (a) training and (b) testing respectively. The regression value for trining is 1 i.e. the output and target are best fits for this application and incase of testing is 0.99992 it means that it mal operate at just only one or two sample sets of data. The results proved that the GRNN is a best choice for solving power transformer normal, inrush and fault cases with high reliability.

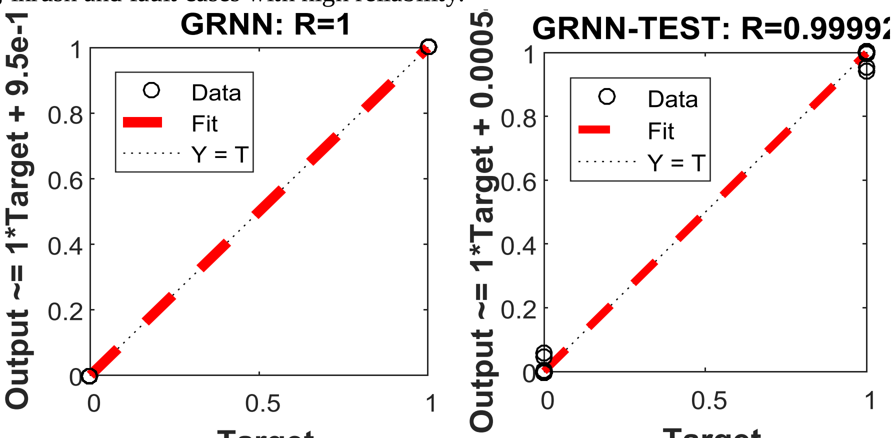


Fig. 10. Linear Regression during (a) Training (b) Testing

V. CONCLUSION

A fault diagnosis system of transformer inrush and internal fault conditions, based on a combination of wavelet and neural network has been developed. By using multi resolution analysis in this system, the differential currents features are extracted with out losing their original properties. For pattern classification of differential currents, both PRNN and GRNN are used. The neural network system using GRNN provided the superior classification time i.e. 10ms than the PRNN. The results proveed that the GRNN provides the best results in terms of linear regression and mean square error values. Thus, for power transformer differential protection GRNN is a promising network in both relaibility and accuracy.

REFERENCES

[1] S. A. Saleh, M. A. Rahman, and L. Fellow, "Testing of a Wavelet-Packet-Transform-Based Differential Protection for Resistance-Grounded Three-Phase Transformers," vol. 46, no. 3, pp. 1109–1117, 2010.

[2] A. A. H. Eldin and M. A. Refaey, "A novel algorithm for discrimination between inrush current and internal faults in power transformer differential protection based on discrete wavelet transform," *Electr. Power Syst. Res.*, vol. 81, no. 1, pp. 19–24, Jan. 2011.

[3] B. Vahidi, N. Ghaffarzadeh, and S. H. Hosseinian, "Electrical Power and Energy Systems A wavelet-based method to discriminate internal faults from inrush currents using correlation coefficient," *Int. J. Electr. Power Energy Syst.*, vol. 32, no. 7, pp. 788–793, 2010.

[4] M. Rasoulpoor and M. Banejad, "A correlation based method for discrimination between inrush and short circuit currents in differential protection of power transformer using Discrete Wavelet Transform: Theory,

simulation and experimental validation," *Int. J. Electr. Power Energy Syst.*, vol. 51, pp. 168–177, Oct. 2013.

[5] S. Jazebi, B. Vahidi, and M. Jannati, "A novel application of wavelet based SVM to transient phenomena identification of power transformers," *Energy Convers. Manag.*, vol. 52, no. 2, pp. 1354–1363, Feb. 2011.

[6] S. K. Bhasker, P. K. Bera, V. Kumar, and M. Tripathy, "Differential Protection of Indirect Symmetrical Phase Shift Transformer and Internal Faults Classification using Wavelet and ANN," vol. 247667, 2015.

[7] H. Balaga, N. Gupta, and D. N. Vishwakarma, "Electrical Power and Energy Systems GA trained parallel hidden layered ANN based differential protection of three phase power transformer," *Int. J. Electr. Power Energy Syst.*, vol. 67, pp. 286–297, 2015.

[8] H. Balaga and D. N. Vishwakarma, "Artificial Neural Network Based Backup Differential Protection of Generator-Transformer Unit," *Int. J. Electron. Electr. Eng.*, vol. 3, no. 6, pp. 482–487, 2015.

[9] D. Tomandl and A. Schober, "A Modi ® ed General Regression Neural Network (MGRNN) with new , ef ® cient training algorithms as a robust ` black box ' -tool for data analysis," vol. 14, 2001.

[10] J. Wu and C. Liu, "Expert Systems with Applications An expert system for fault diagnosis in internal combustion engines using wavelet packet transform and neural network," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 4278–4286, 2009.

[11] A. M. A. Haidar, M. W. Mustafa, F. A. F. Ibrahim, and I. A. Ahmed, "Transient stability evaluation of electrical power system using generalized regression neural networks," *Appl. Soft Comput. J.*, vol. 11, no. 4, pp. 3558–3570, 2011.

[12] J. Wu and J. Kuo, "Expert Systems with Applications An automotive generator fault diagnosis system using discrete wavelet transform and artificial neural network," *Expert Syst. Appl.*, vol. 36, no. 6, pp. 9776–9783, 2009.

