Defect Recognition Method for Magnetic Leakage Detection in Oil and Gas Steel Pipes Based on Improved Neural Networks

Wang Jie, Mohd. Kamal Mohd. Shah, Choong Wai Heng, Nahiyan Al-Azad*
Faculty of Engineering, Universiti Malaysia Sabah,
Kota Kinabalu, Sabah, MALAYSIA
*mkamalms@ums.edu.my

ABSTRACT

The aging infrastructure of petroleum and natural gas pipelines poses a threat to national economies, necessitating precise defect detection for safety and efficiency. To enhance the accuracy of predicting pipeline defect sizes, this study introduces a magnetic leakage detection system, employing Backpropagation (BP) neural networks optimized with genetic algorithms. Traditional BP networks face challenges, including parameter determination and slow convergence, addressed through genetic algorithms' global search capabilities. Simulated data are generated using ANSYS software by using models of semi-circular defects in steel pipes, producing magnetic leakage signals of varying intensities. MATLAB was used to construct both standard BP and genetically optimized BP neural networks. Results show that the latter significantly reduces computational errors, demonstrating improved accuracy in defect dimension prediction. The approach contributes to overcoming non-uniqueness in the recognition process and the complex nonlinear relationship between magnetic signals and defect size parameters. The study offers a guided approach for selecting BP neural network parameters, enhancing practicality. Simulations validate the method's effectiveness, indicating low workload and high reliability. This research provides a meaningful advancement in the detection of defects in long-distance pipelines, impacting the safety and efficiency of petroleum and natural gas transportation.

Keywords: *Genetic Optimization; Magnetic Leakage Detection; BP Neural Network; Steel Pipe Defects*

ISSN 1823-5514, eISSN 2550-164X

https://doi.org/10.24191/jmeche.v21i2.26256

Received for review: 2023-11-20
Accepted for publication: 2024-03-14
Published: 2024-04-15

Introduction

The petroleum and natural gas pipeline infrastructure plays a pivotal role in the national economy, emphasizing the critical need for safety and efficiency [1]. Currently, numerous petroleum steel pipes worldwide have aged, rendering them susceptible to damages such as corrosion and wear, potentially resulting in oil and gas leaks. To ensure the secure operation of these pipelines, the optimal choice is to employ magnetic leakage methods for the detection of long-distance steel pipes utilized in the petroleum and natural gas industry.

The magnetic leakage detection system comprises key components, including front-end magnetic leakage signal acquisition, data compression, and defect recognition. The defect recognition technology assesses whether the detection instrument can accurately represent the geometric parameters of steel pipe defects in a data format [2]. This provides defect detection personnel with a precise understanding of the damage and corrosion levels in the steel pipe, serving as a scientific foundation for determining the necessity of timely pipe replacement.

In recent years, neural networks have emerged as a highly dynamic interdisciplinary field known for their potent self-learning and processing capabilities in terms of defect detection. Backpropagation (BP) neural networks, recognized for their rigorous reasoning process, swift algorithm convergence, simple network model structure, and proficiency in handling classification problems, have become the most widely applied artificial neural network [3]. BP neural networks have gained popularity as defect recognition technology in recent years.

In the mid-1980s, Rumelhart and McCelland introduced the concept of Back Propagation (BP) neural networks [4]. Since then, extensive research has been conducted on BP neural networks [5]-[7]. BP neural networks were proven to predict the performance of the heat-affected zone of continuous oil pipelines [5]. Apart from that BP neural network was found to be used as Newton algorithm-optimized to predict the fatigue life of continuous pipes [6]. A previous study has also proven that BP neural networks successfully predicted the defect dimensions of steel pipes [7].

However, despite the unique advantages of BP neural networks in classification compared to other research methods, such as strong fault tolerance and non-linear mapping capabilities, and the fact that they do not require specific mathematical expressions but can memorize various relevant mappings through training sample data to derive relationships between data, they still exhibit certain limitations in practical applications. Challenges include the reliance on user experience for selecting network structures, the direct impact of the quality of training sample selection on the approximation and generalization capabilities of the network model, and the fixed training step size, which, if adjusted improperly, can lead to prolonged training times [4]-[8].

Genetic algorithms originated in the early 1960s and have evolved as a global search and optimization method, mimicking the mechanisms of biological evolution in nature. Their strength lies in their ability to adaptively control the search process and employ genetic principles, such as survival of the fittest, to iteratively generate an optimal solution among numerous alternatives during the search process [8].

The combination of neural networks' self-learning capabilities and genetic algorithms' global search abilities presents a promising approach. The specific operational method involves determining a parameter encoding scheme, where individuals represent the initial weights and thresholds of the network. The accuracy under cross-validation serves as the fitness function for the genetic algorithm. Through the genetic algorithm's selection, crossover, and mutation operations, the optimal individual is sought. Eventually, the best parameters for the BP neural network are obtained through genetic optimization, resulting in improved recognition performance. Therefore, leveraging the characteristics of standard BP neural networks and genetic algorithms, this study optimizes the standard BP neural networks for predicting steel pipe defect dimensions using genetic algorithms. While the use of standard BP neural networks for defect-recognition in steel pipes is not new, this research introduces a novel approach to address some of the limitations associated with the standard BP algorithm.

The defect samples utilized for testing the neural network can be acquired through three avenues: field detection, experimental processing, and finite element simulation, with the testing outcomes being essentially equivalent. While on-site testing and laboratory processing closely mimic real-world scenarios, they are hindered by drawbacks such as complex processes, extended cycles, and high costs. Moreover, the employment of finite element simulation enables the intentional avoidance of environmental factors like temperature and material surface impurities, enhancing the reference value of the neural network's predictive outcomes. Hence, this study predominantly employs ANSYS simulation to establish defect samples.

Methodology

Establishment of simulation samples

This research used ANSYS software to simulate the magnetic leakage field generated by defects of different sizes on a pipeline with a wall thickness of 25 mm. Figure 1 depicts the two-dimensional solid model of the internal defect in the pipeline, and Table 1 provides information on the materials and dimensions of various parts in the model [9].

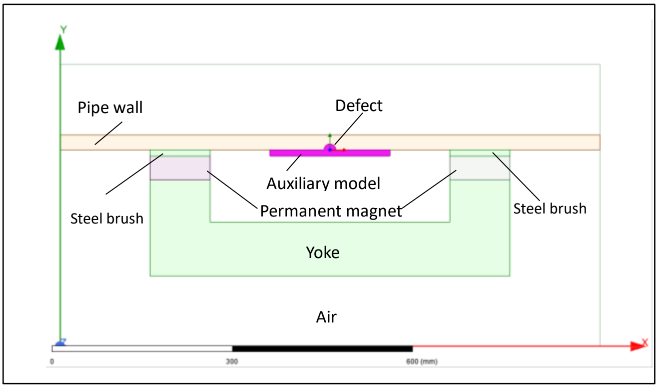


Figure 1: Two-dimensional solid model of internal defect in the pipeline

Table 1: Materials and dimensions of the model

	Material	Length	Width	Others
Pipe wall	X52	900 mm	-	Wall thickness 25 mm
Defect	Air	-	-	Radius 10 mm
Steel brush	Q235	100 mm	10 mm	-
Permanent magnet	NdFeB	100 mm	40 mm	-
Yoke	Q235	600 mm	160 mm	Groove depth 70 mm

The simulation results were extracted along the predefined path to obtain the radial magnetic flux density curve of the defect (Figure 2). In this, *a* represents the peak-to-peak value of the magnetic flux density, and *b* is the horizontal distance difference.

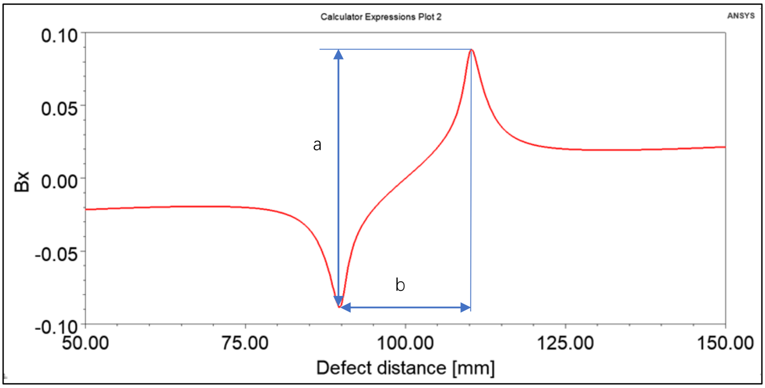


Figure 2: Curve of radial flux density

One hundred sets of defect samples with lengths and depths ranging from 1 to 10 mm were configured (Figure 3), and their magnetic flux density peak-to-peak values and horizontal distance differences were extracted using ANSYS software.

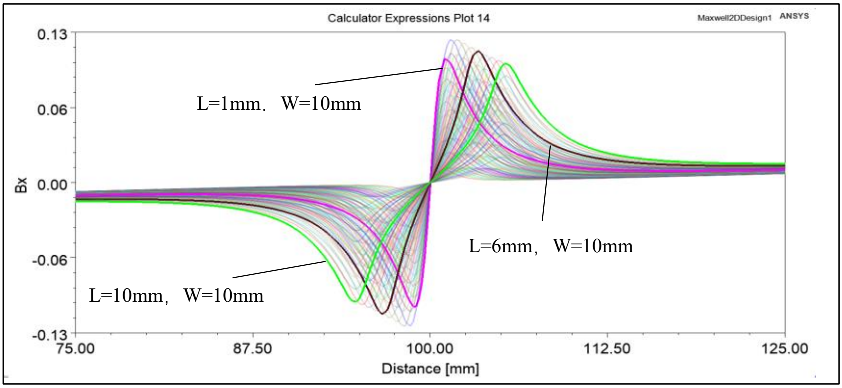


Figure 3: Radial magnetic flux density curves of 100 sets of defects

Quantification of data

The data is divided into a training set and a testing set, with the first 80 samples in the table designated as the training set and the remaining 20 samples as the testing set. A backpropagation neural network is constructed using the *newff* function in MATLAB. The network structure is defined with 2 nodes in the input layer, 4 nodes in the hidden layer, and 1 node in the output layer. The transfer functions are set as *tansig* (hyperbolic tangent function) for the hidden layer and *purelin* (linear function) for the output layer. Training is performed using the gradient descent algorithm.

The determination of the number of nodes in the hidden layer and the training function is based on the results' quality. In practical construction, specific node numbers are generally determined using an empirical formula [10].

$$S = \sqrt{n+m} + c \tag{1}$$

In the formula, *s* represents the number of nodes in the hidden layer, *n* is the number of nodes in the input layer, *m* is the number of nodes in the output layer, and *c* is a constant typically ranging from 1 to 10.

The selection of hidden layer nodes significantly impacts the network's performance; however, there is no clear analytical expression for this to date. An approach often employed is to train with different numbers of neurons and then appropriately add some margin [11].

According to the formula for calculating the number of nodes in the hidden layer, the range of node selection for the steel pipe defect magnetic leakage signal detection and prediction model is set as {2, 12}. The Root Mean Square Error (RMSE) analysis of the model results is illustrated in Figure 4.

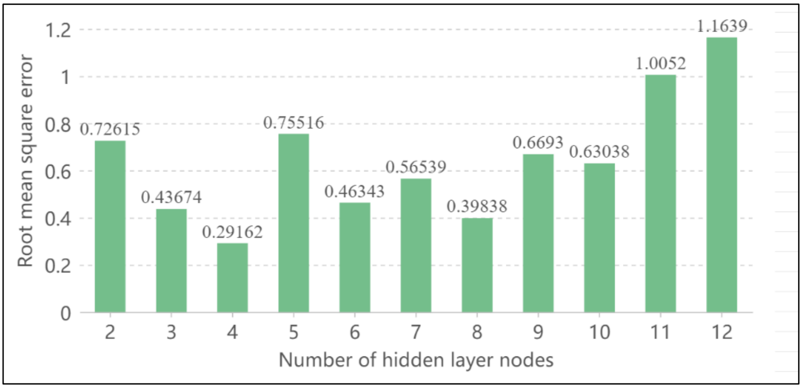


Figure 4: Root Mean Square Error (RMSE) for different hidden layer numbers in the BP neural network

Based on the training results, save the BP neural network model with the optimal performance under each parameter configuration. Determine the network structure of the BP neural network model. During the training process, export the BP neural network structure from MATLAB, as shown in Figure 5. The network structure is 2-4-1.

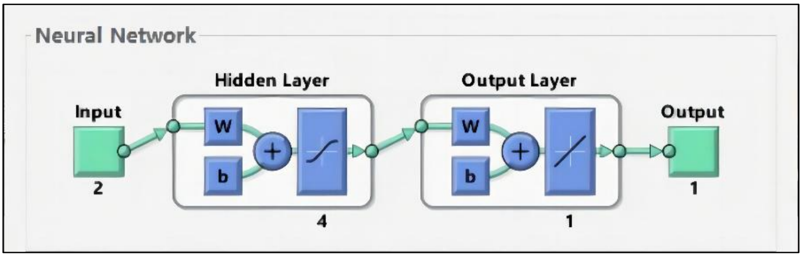


Figure 5: The network structure of the BP neural network

Train the neural network using the train function and set the training parameters. The number of epochs is set to 1000, the learning rate is set to 0.01, and the training goal is to achieve a minimum error of 0.00001.

Normalize the training set data to the range [-1, 1] using the *mapminmax* function to enhance training effectiveness. Use the sim function to predict the normalized test set data. Reverse normalizes the prediction results using the *mapminmax* function to obtain the actual values. Then, calculate the error between the predicted values and the actual values.

The Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Square Error (RMSE) are calculated to evaluate the predictive performance of the model, with the corresponding Equation (2), Equation (3), and Equation (4) as follows:

$$MAE = \frac{1}{n}\sum_{t=I}^{n} | y_t - \hat{y}_t | \tag{2}$$

$$MSE = \frac{1}{n}\sum_{t=1}^{n} (y_t - \hat{y}_t)^2 \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=I}^{n} (y_t - \hat{y}_t)^2} \tag{4}$$

Genetic algorithm of BP neural network

The network is configured as a three-layer structure, as shown in Figure 6. The input layer has 2 nodes, corresponding to two variables: the peak-to-peak value of magnetic flux density and the horizontal distance difference of radial magnetic flux density. The output layer has 1 node, representing the steel pipe defect depth (or length).

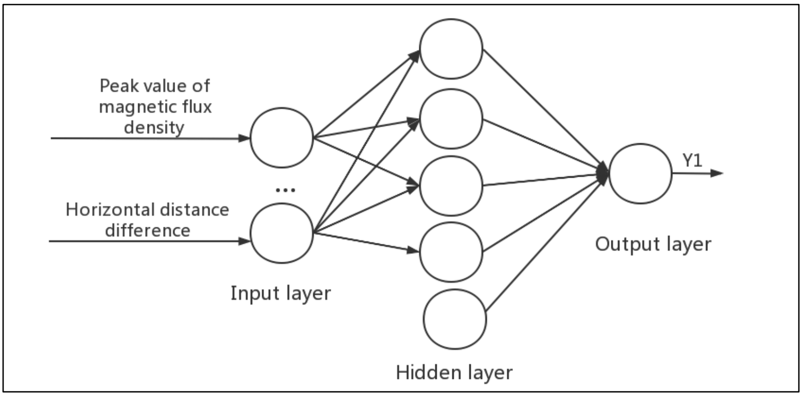


Figure 6: Neural network topology

The hidden layer nodes are obtained through model training using MATLAB software. As shown in Figure 7, the RMSE value of the steel pipe defect magnetic leakage signal detection prediction model is minimized when the number of hidden layer nodes is 5. Therefore, this study adopts 5 neurons in the hidden layer. Sigmoid transfer functions are used for full connectivity between neurons in each layer, with no connections between neurons in different layers. Figure 8 shows the network structure for the defect depth network of the genetically optimized BP neural network.

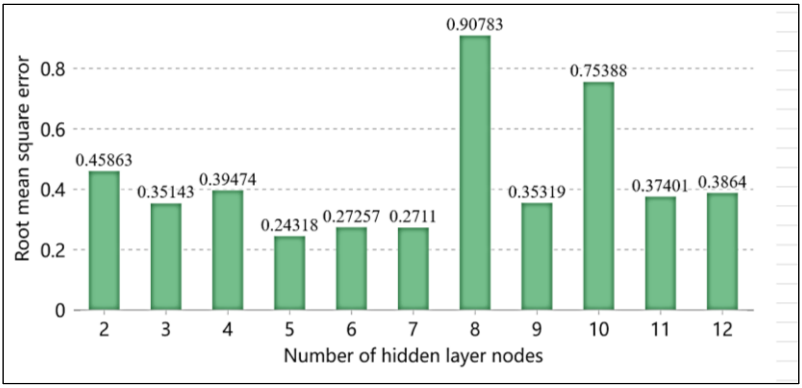


Figure 7: Root Mean Square Error (RMSE) for different hidden layer numbers in the genetic algorithm enhanced BP neural network

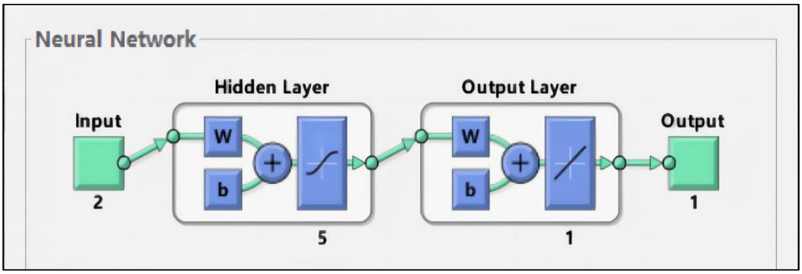


Figure 8: The network structure of the genetic algorithm enhanced BP neural network

The genetic optimization algorithm is an optimization method based on the theory of biological evolution, used to search for the optimal solution in the search space. In the above program, the genetic optimization algorithm is employed to optimize the parameters of the BP neural network.

Results and Discussion

Table 3 and Table 4 present the numerical results of predicting the depth and length of steel pipe defects using two methods: BP neural network and GA-BP neural network. "Input-test" represents the peak-to-peak value of magnetic flux density and horizontal distance difference for the last 20 groups of test data in Table 2. "Output-test" represents the depth and length of defects in the test data. "Test simulation" denotes the predicted numerical values calculated by MATLAB, and "Error" indicates the difference between "Test Simulation" and "Output-test."

Table 3: Predicted data for defect depth

		1	2	3	...	18	19	20
Input-test		0.1901	0.2143	0.2125	...	0.2072	0.2026	0.1973
		2.4	2.8	4	...	8.8	9.6	10.4
Output-test		9	9	9	...	10	10	10
Test simulation	bp	8.4868	9.402	9.0806	...	9.7864	9.7206	9.5672
	gabp	8.4144	9.2522	9.0276	...	9.8887	9.844	9.7605
Error	bp	-0.5132	0.402	0.0806	...	-0.2136	-0.2794	-0.4328
	gabp	-0.5856	0.2522	0.0276	...	-0.1113	-0.156	-0.2395
Relative error	bp	5.70%	4.47%	0.90%	...	2.14%	2.79%	4.33%
	gabp	6.51%	2.80%	0.31%	...	1.11%	1.56%	2.40%

Table 4: Predicted data for defect length

		1	2	3	...	18	19	20
Input-test		0.1901	0.2143	0.2125	...	0.2072	0.2026	0.1973
		2.4	2.8	4	...	8.8	9.6	10.4
Output-test		1	2	3	...	8	9	10
Test simulation	BP	1.1366	1.6368	3.1322	...	7.8837	8.7337	9.5037
	GA-BP	1.4949	2.0174	3.295	...	8.1822	8.9772	9.7351
Error	BP	0.1366	-0.3632	0.1322	...	-0.1163	-0.2663	-0.4963
	GA-BP	0.4949	0.0174	0.295	...	0.1822	-0.0228	-0.2649
Relative error	bp	13.66%	18.16%	4.41%	...	1.45%	2.96%	4.96%
	gabp	49.49%	0.87%	9.83%	...	2.28%	0.25%	2.65%

In a similar study conducted by past researchers [12] using the BP neural network, the average relative error for predicting the length of 16 defects was 5.98%, and for depth is 4.73%. Meanwhile, another research [13], which employed a convolutional neural network model with adaptive gradient descent for similar research, obtained an average relative error of 21.63% for predicting the length of 15 defects and 56.87% for depth. In this research, by using the BP neural network, the average relative error for predicting the depth of 20 defects was 2.55%, and for length, it was 6.46%. Using the GA-BP neural network, the average relative error for predicting the depth of 20 defects was 1.63%, and for length, it was 5.98%. The data above indicate that this research achieves smaller errors and higher accuracy compared to past research [12]-[13] in predicting steel pipe defect sizes using the genetic optimization BP neural network.

Table 5 shows the simulation results of the two defect identification methods obtained through MATLAB calculations. When the output layer data represents defect depth, the BP neural network model is trained and predicted based on the magnetic flux density peak-to-peak value and horizontal distance difference obtained from the magnetic leakage detection of steel pipes. The MAE between the predicted values and the experimental values is 0.2443, and the RMSE is 0.2916. The model exhibits good approximation and prediction capabilities, effectively meeting practical prediction requirements, indicating that the established BP neural network model is feasible.

The BP neural network model optimized by genetic algorithm (GA) has an MAE of 0.1562 and an RMSE of 0.2432 between the predicted values and experimental values. It can well meet the practical prediction requirements, indicating that the GA-optimized BP neural network model has high accuracy in predicting defect sizes.

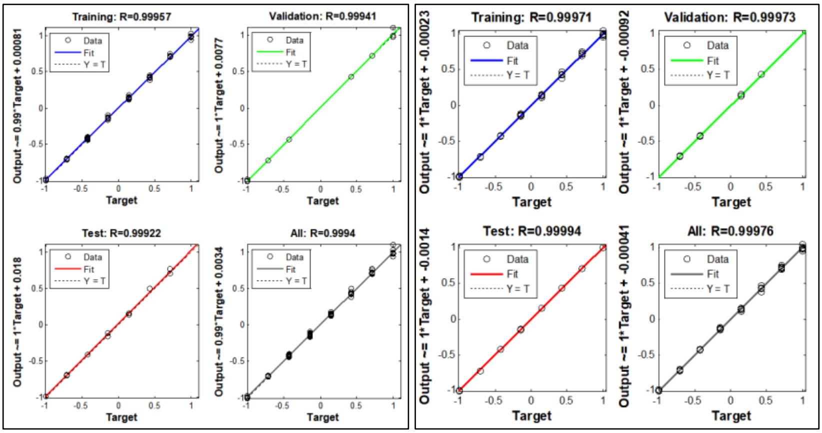
The comparative analysis of the prediction models between the GA-BP neural network and the BP neural network indicates that in the field of magnetic leakage detection data prediction of the GA-optimized BP network model exhibits smaller relative errors, higher accuracy, and improved convergence speed compared to the BP model [14].

Table 5: Network simulation results of two defect identification methods

	BP	GA-BP
MAE	0.24431	0.15617
MSE	0.085041	0.059139
RMSE	0.29162	0.24318

The data correlation relationships of the training set, validation set, testing set, and overall results after network training are illustrated in Figure 9. The horizontal axis represents the number of data groups, and the vertical axis represents sample values. The regression value R represents the correlation between predicted output and target output.

By utilizing the BP neural network model to train and predict the relationship between the magnetic flux density peak-to-peak value, horizontal distance difference obtained from the magnetic leakage detection of steel pipes, and the size of defects, the correlation coefficients R for the four datasets range from 0.99922 to 0.99957. The R values obtained using the GA-optimized BP neural network range from 0.99971 to 0.99994. The comparison indicates that the GA-optimized BP neural network provides better training results for the samples.



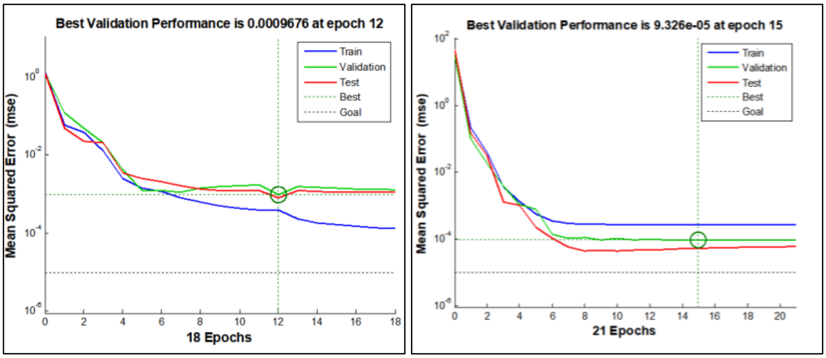
(a) BP (b) GA-BP

Figure 9: Neural network training regression

Figure 10 illustrates the training mean squared error curve for defect depth. In the graph, the green line represents predicted values, while the red line represents sample values. It is noticeable that both the BP neural network's sample data and the model-predicted data converge to the optimal training value around the 12th cycle, with the mean squared error dropping below 10^{-3}. The genetically optimized BP neural network's sample data and model prediction achieve their best training value in the 15th cycle, with the mean squared error decreasing below 10^{-4}.

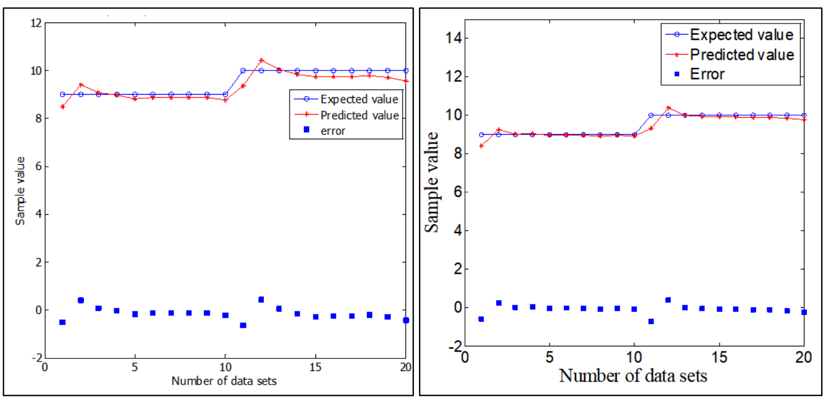
According to statistical definitions [15], a smaller MSE between predicted values and outcomes indicates a smaller disparity between the estimated value and the sample data, which in turn suggests higher accuracy of the trained model. The genetically optimized BP neural network exhibits a smaller MSE, indicating more precise model training compared to the standard BP neural network.

Figure 11 is a comparison of the predicted and actual values of the test set. As can be seen from the figure, the output values of 20 samples given by the two methods are almost the same as the real values.



(a) BP (b) GA-BP

Figure 10: Defect depth training mean square error curve



(a) BP (b) GA-BP

Figure 11: Comparison of the predicted and actual values

Data sets with higher fitness are selected for the next cycle using genetic algorithms. After defining the BP neural network structure, the encoding of weights and thresholds is done. Adaptive genetic algorithms are employed for optimization. The change in fitness across generations is depicted in Figure 12. The fitness curve stabilizes after approximately three generations of genetic simulation. At this point, the best fitness value corresponds to optimal weights and thresholds.

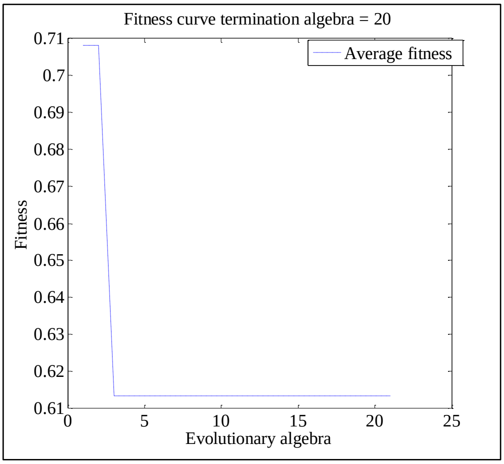


Figure 12: Fitness change with genetic generations curve

Conclusion

This research addresses the issue of dimension recognition for defects in long-distance oil and gas pipelines. To overcome challenges such as the non-uniqueness in the recognition process and the complex nonlinear relationship between defect leakage magnetic signals and defect size parameters, a novel recognition method based on a genetic algorithm-optimized backpropagation (BP) neural network is proposed. This approach targets the difficulty in determining BP neural network parameters. By employing improved genetic operations (selection, crossover, mutation) to optimize the parameters of the BP neural network, the optimal parameters are obtained. This method overcomes the arbitrariness in manually selecting BP neural network parameters, providing a meaningful and guided approach to the selection of BP neural network parameters.

To validate the effectiveness of the method, a semi-circular defect model was established using ANSYS software, and simulations were conducted to obtain corresponding leakage magnetic data as learning samples. Both the backpropagation (BP) neural network and the genetically optimized BP neural network were trained and tested. Experimental results demonstrate that applying the genetically optimized BP neural network to the identification of steel pipe defects is not only feasible but also yields desirable outcomes. It has advantages such as low workload and high reliability, indicating significant practical applications.

Although this research has achieved the goal of predicting defect sizes, it is well-known that real-world situations can be more complex than laboratory experiments. Irregular defect sizes, fluctuating temperatures and pressures, insufficient training samples, and low-quality data can all impact the accuracy of the model predictions. Additionally, considerations such as the hardware and software requirements of the prediction model, the real-time demands of enterprises for the prediction model, and issues like network security cannot be overlooked. In future research, a comprehensive approach is required to address these issues, optimizing aspects such as model structure and data processing in a timely manner to ensure the feasibility of the model in practical scenarios.

Contributions of Authors

The authors confirm the equal contribution in each part of this work. All authors reviewed and approved the final version of this work.

Funding

This work received no specific grant from any funding agency.

Conflict of Interests

All authors declare that they have no conflicts of interest.

Acknowledgment

Special thanks to the staff and members of the Faculty of Engineering, Universiti Malaysia Sabah who provided the facilities for this research in order to complete this work.

References

[1] Yan, L. I., & Fei, X., "Prediction of Residual Life of Corroded Pipeline Based on Improved Neural Network", *Contemporary Chemical Industry*, vol. 49, no. 11, pp. 2629–2632, 2020. http://doi.org/10.3969/j.issn.1671-0460.2020.11.059

[2] Yu, L., Si-quan, Z., Chang, Y., & Chuan, Q., "Summary of two pipeline electromagnetic nondestructive testing methods", *Mechanical &*

Electrical Engineering Magazine, vol. 31, no. 7, pp. 844–848, 2014. http://doi.org/10.3969/j.issn.1001-4551.2014.07.006

[3] Zou, Y., Yang, L., Li, B., Yan, Z., Li, Z., Wang, S., & Guo, Y., "Prediction Model of End-Point Phosphorus Content in EAF Steelmaking Based on BP Neural Network with Periodical Data Optimization", *Metals*, vol. 12, no. 9, p. 1519, 2022

[4] Kan, Z., & Wanlin, Z., "Under the background of the "dual carbon" strategy, thermal power enterprises BP neural network optimized by genetic algorithm", *Financial Management Research*, vol. 5, pp. 74–81, 2023.

[5] Li, J. H., Li, L., Bi, Z. Y., & Zhang, M., "Prediction of mechanical properties of welding HAZ of coiled tubing based on BP neural network", *Ordnance Material Science And Engineering*, vol. 35, no. 03, pp. 15–18, 2012.

[6] Yihua, D., Jiaqiang, Z., & Guoliang, L., "Fatigue Life Prediction of Coiled Tubing Based on Optimized BP Neural Network", *China Petroleum Machinery*, vo. 51, no. 10, pp. 144–149, 2023.

[8] Jiang, Q., Huang, R., Huang, Y., Chen, S., He, Y., Lan, L., & Liu, C., "Application of BP neural network based on genetic algorithm optimization in evaluation of power grid investment risk", *IEEE Access*, vol. 7, pp. 154827–154835, 2019.

[9] Liu, Y., Wu, W., Ren, X., Qin, L., & Wang, Y., "Research on Geometric Parameter Prediction Algorithm for Oil and Gas Pipeline Defects". *Academic Journal of Science and Technology*, vol. 7, no. 3, pp. 34–39, 2023.

[10] Guoliang, L., Jiaqiang, Z., Liang, W., Yinping, C., & Yihua, D., "The fatigue life prediction of coiled tubing under different internal pressure based on BP neural network", *Machine Design And Manufacturing Engineering*, vol. 52, no. 4, pp. 97–101. 2023. http://doi.org/10.3969/j.issn.2095-509X.2023.04.020

[11] Yan-juan, H. U., Zhan-li, W., & Dan, Z., "The prediction of milling force based on linear regression and BP neural network", *Manufacturing Automation*, vol. 18, pp. 96–99, 2013. http://doi.org/10.3969/j.issn.1009-0134.2013.18.029

[12] Ling, J., Feng, K., Wang, T., Liao, M., Yang, C., & Liu, Z., "Data Modeling Techniques for Pipeline Integrity Assessment: A State-of-the-Art Survey", *IEEE Transactions on Instrumentation and Measurement*, 2023.

[13] Feng, J., Zhang, X., Lu, S., & Yang, F., "A single-stage enhancement-identification framework for pipeline MFL inspection", *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–13, 2022.

[14] Zhan-li, W., Ping, X. I., Jing, L., & Dan, Z., "Genetic algorithm optimize BP network in the application of the milling force prediction".

Manufacturing Automation, vol. 14, pp. 64–68, 2014. http://doi.org/10.3969/j.issn.1009-0134.2014.14.016

[15] Ming, H., "*Probability Theory and Mathematical Statistics*", Tongji University Press, 2014.