Research on degradation prediction of rolling bearing based on adaptive multi-GA-BP

Lingyu Hou¹, Yuanyuan Li¹, Wei Yao¹, Ming Tang¹, Qichun Sun¹ and Jiahang Chen¹

Abstract
Rolling bearings are widely used in industrial equipment. It is of great significance to study the degradation trend of rolling bearings. In this paper, an Adaptive Multi-population Genetic Algorithm (AMGA) is proposed. Firstly, Kernel Principal Component Analysis (KPCA) method is used to fuse the vibration signal in both time domain and frequency domain, which uses kernel to map the sample space to higher dimensional space and uses the higher dimensional space for linear dimensionality reduction. It can effectively reduce the dimension of nonlinear correlation variables and obtain the trend signal representing the Remaining Useful Life (RUL) characteristics of rolling bearing. Moreover, AMGA is proposed to optimize the number of neurons, initial weight, and initial threshold of the Back Propagation (BP) neural network prediction model. AMGA applies chaos algorithm to Genetic Algorithm (GA) to improve the diversity of the initial population. Meanwhile, the communication frequency between different populations is controlled by judging the similarity of the optimal solution among different populations, so as to effectively jump out of the local optimum and obtain the global optimal solution. Finally, the whole life data of the spin-up process of the rolling bearing from University of Cincinnati is taken as an example to analyze the performance of the algorithm. Compared with the traditional BP, the R2-score performance and the MAPE performance of KPCA-AMGA-BP are improved by 0.297 and 2.46% respectively. Furthermore, compared with the optimized BP, this method obtains the improved R2-score performance and the MAPE performance by 0.218 and 0.46%.

Keywords
adaptive multi-population genetic algorithm, back propagation, kernel principal component analysis, rolling bearing, degradation prediction

Date received: 3 March 2022; accepted: 13 September 2021

Introduction
With the development of technology and large-scale socialized production, reliable long-term and alternate multiple operational States are essential for modern manufacturing. Rolling bearings are widely used in these types of machinery that are normally operated in high-speed and heavy-duty environments. Once damage occurs, the consequences are often catastrophic. Therefore, it is of great significance to predict the degradation of rolling bearing for preventing equipment failure and realizing equipment maintenance based on condition.

The research on the residual life of rolling bearings was put forward at the end of the 19th century. In order to improve the reliability of predictive maintenance, the literatures use different methods to improve the prediction accuracy. At present, there are some achievements in the prediction of residual life of rolling bearings at domestic and overseas. Peng, Y., et al. used the Least Squares Support Vector Machine (LS-SVM) to predict the remaining useful life of the rolling bearing. LS-SVM involves penalty parameters, so the determination of penalty parameters has a great impact on the prediction effect. Fengtao, W., et al. used the Gray Model GM (1,1) to predict the residual life of the rolling bearing. Lu, Y., et al. used the Auto-Regressive Moving Average Model (ARMA). Kuncan, M., et al. proposed a new approach that consisted of the one-dimensional local binary pattern (1D-LBP) method and the gray relational analysis (GRA) model to classify the vibration signals. The existing studies show that the machine learning method can effectively improve the accuracy of predicting the remaining life of rolling bearings. However, with the increment of the complexity of vibration signals, the deep learning method starts to show advantages in processing complex vibration signals.

¹School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai, China

Corresponding author:
Yuanyuan Li, School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, No.333 Longteng Road, Songjiang District, Shanghai, 201620, China. Email: liyuanyuanedu@163.com
Presently, there are many works on rolling bearing fault diagnosis using different deep learning techniques, such as convolution neural network (CNN),9,10 recurrent neural network (RNN),11 transfer learning-convolutional neural network12 (TLCNN), adaptive deep transfer learning neural network13 and so on. Through the above methods, satisfactory diagnosis accuracy can be provided. When we process one-dimensional vibration signals of rolling bearing, the training speed of BP neural network is faster than those methods due to the simpler network.

In recent years, research works on the degradation prediction of rolling bearing using BP neural network have been introduced. Zhang, R., et al.14 used a dictionary learning algorithm to denoise vibration signals and used BP neural network to detect the fault state of rolling bearing. However, the results of BP neural network are often affected by the super parameters in the training process, and there is no clear criterion for the selection of the super parameters. Therefore, Liu, X., et al.15 and Wan, L., et al.16 improved the BP prediction model by using Particle Swarm Optimization (PSO), which was applied to the wind turbine blade strain prediction and rolling bearing fault prediction. This method effectively improved the nonlinear fitting ability and efficiency of the single BP prediction model. Cheng, P., et al.17 and others used Genetic Algorithm (GA) to optimize the weights and thresholds of BP, which improved the efficiency and accuracy of human body shape prediction.

In actual industrial production, the working conditions of the rolling bearing are complicated and various; therefore, how to effectively predict rolling bearing fault under various working conditions is an important problem to be solved. So, in order to improve the efficiency and accuracy of rolling bearing degradation prediction, this paper uses the time domain and frequency-domain feature extraction of the data. Then, we apply the Kernel Principal Component Analysis (KPCA) method to reduce the dimension of the characteristic signal, eliminate the multiple collinearities between variables, and obtain the health index of rolling bearing. Finally, the health index is input into the Adaptive Multi-population Genetic Algorithm (AMGA-BP) prediction algorithm. The overview of the method is outlined in Figure 1.

To summarize, the contributions of this paper are as follows:

1. KPCA method is used to fuse the time and frequency domain signals extracted from rolling bearing in order to obtain effective bearing degradation signals in this paper since that it can fuse multi-dimensional feature signals, improve the signal-to-noise ratio of the data, remove the relevant indicators, and avoid input redundancy.
2. The chaos algorithm is combined with GA to improve the diversity of the initial population of GA in this paper. Meanwhile, the similarity of the optimal solutions among different groups is judged by the multi-population parallel method to control the communication frequency between populations creatively, so as to jump out of the local optimum and obtain the global optimal solution.
3. The number of neurons, initial weight, and initial threshold of BP neural network is optimized by using AMGA, to increase the nonlinear fuzzy approximation ability of BP.

This paper is mainly arranged as follows: Bearing feature signal extraction section describes the theory of feature extraction; The prediction model of AMGA-BP section describes the theory and improvement of degradation prediction model AMGA-BP; Experimental analysis section describes the experiment and comparison between AMGA-BP, GA-BP, BP, and LSTM; Conclusion section describes the conclusion.
Bearing feature signal extraction
Time and frequency domain feature parameters extraction

The time domain characteristic indexes include the dimensional indexes such as Mean value, Skewness, and Kurtosis, as well as the dimensionless indexes such as Form Factor and Margin Factor; the frequency domain characteristic indexes involve the information such as frequency spectrum distribution, vibration energy size, and frequency band position change. To obtain the data with obvious fault features and surface bearing performance degradation, the noise in the original data is removed through feature extraction. The following calculation formulas of the time and frequency domain are shown in Tables 1 and 2.

Feature fusion
The literature shows that KPCA is an unsupervised feature fusion learning method. KPCA can effectively remove the redundant information from the original data and extract the effective information from the original data. Feature extraction based on KPCA is a nonlinear analysis method based on PCA, which mainly uses the nonlinear kernel function to map the input data to high-dimensional feature space, and transforms the nonlinear problem in low dimensional space into the linear problem in high-dimensional space, so as to carry out principal component analysis. Because of the limited degradation that can be reflected by a single index, it is necessary to construct a new index that can comprehensively reflect the deterioration of bearing performance by using the method of kernel principal component analysis.

Table 1. Time domain characteristic calculation formula.

| Characteristic | Formula |
|---------------|---------|
| T1 (Skewness) | $E((X - \text{Mean}(X))^3)/\text{Std}(X)^3$ |
| T2 (Kurtosis) | $E((X - \text{Mean}(X))^4)/\text{Std}(X)^4$ |
| T3 (Root mean square) | $\sqrt{\sum \text{abs}(X)^2}$ |
| T4 (Peak to peak) | $\text{Max}(X) - \text{Min}(X)$ |
| T5 (Form factor) | $\text{Mean}(|X|) / \text{Max}(|X|)$ |
| T6 (Crest factor) | $\text{Max}(|X|) / \text{Mean}(|X|)$ |
| T7 (Impulse factor) | $\text{Mean}(|X|)/\text{Max}(|X|)$ |
| T8 (Margin factor) | $\text{Max}(|X|)/\text{Mean}(|X|)$ |

Table 2. Frequency domain characteristic calculation formula.

| Characteristic | Formula |
|---------------|---------|
| P1 (Dispersion and concentration of the spectrum) | $\sum |Y - \text{mean}(Y)|^2$ |
| P2 (Dispersion and concentration of the spectrum) | $\sum |Y - \text{mean}(Y)|^2/\text{Max}|Y|^2$ |
| P3 (Dispersion and concentration of the spectrum) | $\sum |Y - \text{mean}(Y)|^2/\text{Max}|Y|^2$ |
| P4 (Change in the position of the main frequency band) | $\sqrt{\sum (f - f_0)^2}$ |
| P5 (Change in the position of the main frequency band) | $\sqrt{\sum (f - f_0)^2}/\text{sum} f^2$ |

For $N$ samples in the original space, $X_k ((k = 1, 2, ..., N)$, $X_k \in \mathbb{R}^M$, $M$ is the number of variables) is the covariance matrix $C^H$ of the mapping matrix on the high-dimensional feature space $H$ by means of nonlinear mapping

$$C^H = \frac{1}{M} \sum_{k=1}^{M} \phi(x_k)\phi(x_k)'$$

$$\lambda V = C^H V$$

where $\lambda$ is the characteristic value, $V$ is the eigenvector.

Define the matrix $K_{ij} = K(x_i, x_j) = \phi(x_i) \phi(x_j)$ and normalize eigenvectors $V^m$, which satisfies $(V^m, V^m) = 1$, then, the original sample data $X$ is mapped to the image $\phi(x)$ on the high-dimensional eigenspace $H$ on the $m$-th ($m = 1, 2, ...$) eigenvector matrix $t_m$ can be expressed as

$$t_m = (V^m, \phi(x)) = \sum_{i=1}^{N} \alpha_i^m K(x_i, x_i)$$

Parameter optimization based on GA

Figure 2. KPCA framework.
Particle Swarm Optimization (PSO)\textsuperscript{22} were used to generate Pareto optimal solutions for several complexity varying experiments. The basic principle of GA is derived from Darwin’s biological evolution principle. Through the simulation of biological evolution processes such as replication, exchange, and mutation, and the corresponding simulation of natural selection pressure according to the demand, the optimal solution can be obtained.\textsuperscript{23}

GA generates an initial population $X$. The $X$ population has $N$ individuals. The initial individuals $[X_1, X_2, ..., X_N]$ have three main operations in the genetic iteration process: selection, crossover, and mutation.

1. Selection of genetic factors refers to individuals according to certain laws, the selection of good individuals from the $t$-th generation population under the premise of comprehensive consideration of individual adaptability entail to the $(t+1)$-th generation population.
2. Cross-genetic factors refer to the pairing of individuals in the $t$-th generation population according to the principle of randomness. Meanwhile, the certain chromosomes of individuals exchange according to a certain probability.
3. Variant genetic factor refers to the change of some of the factors on the chromosome of an individual in the $t$-th generation population according to a lower probability of variation, by which the original gene value is converted to other allele values.

Although GA can achieve convergence globally in theory, it usually achieves local convergence in the practical process. Moreover, it is inefficient, time-consuming, and unstable, so it needs to be optimized.

**Parameter optimization based on adaptive multi-population GA**

AMGA is proposed in this paper to balance the global optimization ability and local optimization capability of GA, and improve the accuracy of the algorithm. The specific improvement methods are as follows:

1. In order to improve the diversity of the initial population, a chaos algorithm is applied to the generation of the initial population. The chaos algorithm is expressed as:

$$X^{t+1} = \mu X^t (1 - X^t)$$

when $\mu = 4$, $X^t$ achieves a completely chaotic state, makes the population evenly distributed in the range, and effectively improves the diversity of the initial population.

2. In order to solve the problem that a single population needs to search for the global optimal solution quickly and avoid falling into the local optimal value in the iterative process, the multi-populations hybrid genetic algorithm is adopted. Its basic idea is to make the multi-populations search for the optimal solution with different crossover probability and mutation probability. In the iterative process of the populations, the optimal solution among the multi-populations is transferred to avoid premature convergence of the populations and fall into the local optimum. However, individual migration between populations in each iteration process is easy to be lead to multiple populations entering the same local optimum, so the similarity between populations in each iteration process is calculated. When the similarity between populations is too low, individual migration between populations is carried out to make the populations communicate with each other. The threshold of similarity between populations is set as $A$ (vector distance) and $B$ (difference of optimal solution). If the similarity is too low, it means that the spatial distance of the optimal solution is far between multi-population and the fitness value of the optimal solution is similar. In this case, multi-population communication can make the population jump out of the local optimum. The initial values of $A$ and $B$ are set randomly and experimentally to find the optimal values. Tables 3 and 4 show the optimal solutions of functions 1 and function 2 with different similarity thresholds. The actual optimal solutions of both functions are 0. Figures 3 and 4 show the optimization effect of function 1 and function 2 on the optimization effect of using MGA and traditional GA with different values of $A$ and $B$. The X-axis is the number of iterations, and the Y-axis is the optimal value. The curve represents the optimal value of the model as the number of iterations increases. The number of iterations required by MGA to find the best is significantly less than that of traditional GA. As shown in the figure, when $A$ is close to 1, the number of iterations to reach the optimal solution is less, indicating that the speed of finding the optimal solution is the fastest. When $B$ is set to 60, the speed of finding the optimal solution is the fastest, too. Function 1 and function 2 are defined as follows:

Function 1: Rastrigin

$$f(x) = \sum_{i=1}^{30} (x_i^2 - 10 \cos(2\pi x_i) + 10), |x_i| \leq 100 \quad (6)$$

Function 2: Sphere

$$f(x) = \sum_{i=1}^{30} x_i^2, |x_i| \leq 100 \quad (7)$$

Meanwhile, it can be seen from Tables 3 and 4 that the error between the optimal value obtained by MGA and the real optimal value is small. So, the effect of MGA is better than that of traditional GA. However, the effect of the multi-population genetic algorithm is related to the communication condition, and the best similarity judgment condition can be obtained through multiple experiments. Then parameters $A = 1$, $B = 55$ are set in this case.

3. The communication conditions of MGA are required to be obtained through many experiments, which is...
Table 3. The result of similarity A = 3, A = 2, A = 1, A = 0.8.

| Categories   | A  | B  | Optimal solution of function 1 | Optimal solution of function 2 |
|--------------|----|----|-------------------------------|-------------------------------|
| GA           | -  | -  | 38.63                         | 0.1377                        |
| Multi-GA1    | 3  | 50 | 14.64                         | 0.0036                        |
| Multi-GA2    | 2  | 50 | 12.35                         | 0.0035                        |
| Multi-GA3    | 1  | 50 | 10.78                         | 0.0047                        |
| Multi-GA4    | 0.8| 50 | 9.79                          | 0.0028                        |

Table 4. The result of similarity B = 50, B = 60, B = 70, B = 40.

| Categories   | A  | B  | Optimal solution of function 1 | Optimal solution of function 2 |
|--------------|----|----|-------------------------------|-------------------------------|
| GA           | -  | -  | 38.63                         | 0.1377                        |
| Multi-GA1    | 1  | 50 | 14.64                         | 0.0036                        |
| Multi-GA2    | 1  | 60 | 14.77                         | 0.0016                        |
| Multi-GA3    | 1  | 70 | 22.18                         | 0.0014                        |
| Multi-GA4    | 0.8| 40 | 21.34                         | 0.0035                        |

Figure 3. Optimization of function 1 with different thresholds. (a) Optimal solution under A = 3, A = 2, A = 1, A = 0.8; (b) Optimal solution under B = 50, B = 60, B = 70, B = 40.

Figure 4. Optimization of function 2 with different thresholds. (a) Optimal solution under A = 3, A = 2, A = 1, A = 0.8; (b) Optimal solution under B = 50, B = 60, B = 70, B = 40.

the limitation. In order to solve this problem, the similarity threshold is proposed to be adaptive. It is assumed that the similarity between multi-population is low when the initial population is generated. The similarity at the initial time can be used as the threshold to solve this problem. As shown in Figure 5, the initial value of AMGA solution is close to the best similarity threshold effect that is obtained by plenty times of experiments with other algorithms.

The flow chart of AMGA is shown in Figure 6.

**Prediction model based on BP**

BP is a typical multi-layer Back Propagation neural network. Its learning process is mainly composed of two parts: signal
forward propagation and error back-propagation. When the signal propagates forward, the signal enters the network through the input layer and passes through the hidden layer to the output layer. The state of neurons in each layer can only affect the state of neurons in the next layer. When the output layer does not get the expected signal, it enters the back-propagation mode, and the network will adjust the weight and threshold through the deviation, so that the output of BP is close to the expected output. After a certain number of iterations, the error is finally controlled in a very small range.\textsuperscript{24}

Three-layer BP is proposed to predict the residual life of the rolling bearing. As shown in Figure 7, the first layer is the data input layer, which inputs the historical degradation performance data \{X_1, X_2, \ldots, X_n\}. The second layer is the hidden layer containing neurons. The third layer is the output layer which describes the output variables which is the degraded data \{Y_1, Y_2, \ldots, Y_n\} in the future. The data is processed through the second layer \(W_{ij}\) and third layer \(W_{jk}\) with activation function.

**BP prediction model optimized by AMGA**

Although BP has a strong ability to deal with nonlinear data, its convergence speed is slow and it is easy to fall into the local optimum. The convergence of BP mainly

---

**Figure 5.** Optimal solution under AMGA.

**Figure 6.** Flow chart of AMGA.
forward propagation and error back-propagation. When the signal propagates forward, the signal enters the network through the input layer and passes through the hidden layer to the output layer. The state of neurons in each layer can only affect the state of neurons in the next layer. When the output layer does not get the expected signal, it enters the back-propagation mode, and the network will adjust the weight and threshold through the deviation, so that the output of BP is close to the expected output. After a certain number of iterations, the error is finally controlled in a very small range.

Three-layer BP is proposed to predict the residual life of the rolling bearing. As shown in Figure 7, the first layer is the data input layer, which inputs the historical degradation performance data \( \{X_1, X_2, \ldots, X_n\} \). The second layer is the hidden layer containing neurons. The third layer is the output layer which describes the output variables which is the degraded data \( \{Y_1, Y_2, \ldots, Y_n\} \) in the future. The data is processed through the second layer \( W_{ij} \) and third layer \( W_{jk} \) with activation function.

BP prediction model optimized by AMGA

Although BP has a strong ability to deal with nonlinear data, its convergence speed is slow and it is easy to fall into the local optimum. The convergence of BP mainly depends on the initial weights and thresholds. According to the characteristics of BP, it is a popular and effective method to overcome the defects of the natural heuristic algorithm.

In this paper, AMGA is used to optimize the initial weights and thresholds of BP. Each individual in the population records the ownership value and threshold value of the network. During every training, the weights, and thresholds will be updated. The individual selects, crosses, and mutates the individual population through the network feedback error value as the fitness value. After the evolution, the individual is constantly changing weights and thresholds of the network which are updated to make the BP prediction more accurate. The improved prediction model is shown in Figure 8.

**Experimental analysis**

**Experimental data**

The experimental data are provided by the University of Cincinnati, USA. The Dataset is retrieving from https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository. Four Rexnord ZA-2115 double row bearings are installed on the shaft as shown in Figures 9 and 10. Each bearing is equipped with
acceleration sensors in the direction perpendicular to each other, so it can meet the basic conditions of full vector fusion. The motor speed is 2000 r/min, the sampling rate is set to 20 kHz, the data acquisition interval is 20480 points, and the acquisition time is 984*10 min. Bearing 1 is selected to collect the outer ring failure vibration signal, as shown in Figure 11, which is the time domain waveform of the whole life process of the bearing from normal operation to outer ring failure.

**Verification results and discussion**

**Time and frequency feature extraction.** Because of the large amount of original data, the time-frequency domain feature extraction can effectively extract features with a large amount of recognition information, which is reliable and can improve the accuracy of the training and testing of the prediction model.

9 time domain signals and 5 frequency domain signals shown in Figures 12 and 13 are extracted from the rolling bearing. It can be seen that each index reflects the bearing degradation trend in different ways. Among them, Kurtosis, Root Mean Square and P1 have obvious degradation trends and Form Factor, Crest Factor, Margin Factor, and P2 are too noisy to be sensitive to early faults, so they require for further feature extraction.

**Feature fusion.** The process of bearing degradation is complex and mutable, so it is necessary to use multiple characteristic parameters to characterize the whole degradation process, so as to eliminate the parameters which are insensitive to bearing degradation. Compared with the traditional feature extraction methods, the feature extraction method combining time domain and frequency domain features improve the accuracy, adaptability, and robustness of bearing degradation trend prediction.

Through time frequency feature extraction, 9 time domain features and 5 frequency domain features can be obtained, but the contribution of each feature to the life characteristics of rolling bearing is different. KPCA is used to obtain the principal component after time and frequency domain feature fusion. The first principal component with the highest correlation of rolling bearing life is used as the trend signal of the rolling bearing life feature. The first principal component is shown in Figure 14.

**Degradation prediction.** The obtained trend signal of rolling bearing life characteristics is input into the BP prediction
model improved by AMGA, which is proposed in a new communication strategy to improve the accuracy and efficiency of optimization.

Figures 15–17 respectively show that 30 points are used to predict the next 60 points under BP, GA-BP, and AMGA-BP. The blue line is the actual value and the red line is the predicted value. Obviously, the predicted value is much closer to the actual value of using the AMGA-BP prediction. Table 5 shows the comparison of MAE, MAPE, RMSE, and R2-Score under different models, where MAE is mean square error, MAPE is mean absolute percentage, and RMSE is root mean square error. The values are much closer to 0 when the model is more accurate. Moreover, the prediction model achieves accurate performance when R2-Score is close to 1. MAE, MAPE, RMSE, and R2-Score are obtained from equations (8-11). In the equation $y_i$ is the predicted value, $\bar{y}$ is the actual value and $\bar{y}$ is the mean of actual value. Among them, the evaluation index results of LSTM are obtained by trying to replicate. The results from AMGA-BP of these evaluation indexes are better than other methods, which show that this method is indeed effective.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$ (8)
MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{|y_i|} \quad (9)
\begin{equation}
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (10)
\end{equation}
R2 – Score = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (11)

Conclusion

In this paper, the KPCA method is used to fuse the time domain and frequency domain signals of rolling bearing vibration signals obtained in complex environments, and obtain the whole life trend of rolling bearing effectively. Moreover, KPCA-AMGA-BP prediction model is proposed, which is higher than BP, GA-BP, and LSTM method in prediction accuracy. AMGA uses the distance and fitness comparison of the optimal solution in the iterative process of two populations to determine whether the two populations are allowed to carry out gene exchange, which limits the exchange conditions and increases the effectiveness of the exchange between populations in the optimization process. The proposed prediction model uses AMGA method to optimize the number of neurons, the initial weights, and thresholds of BP, which increases the prediction accuracy of BP.

The experimental results show that the proposed method can accurately predict the future degradation trend of bearings within the allowable error range. Compared with other similar methods, this method has certain advantages under MAE, MAPE, RMSE, and R2-Score, and can provide new ideas for the related research in this field.

Acknowledgements

We thank Shanghai University of Engineering Science for assisting with the measurement setups.

Author Contributions

Conceptualization, L.H and Y.L; methodology, L.H and Y.L; software, L.H, Q.S and M.T; validation, L.Y, L.C, M.T, J.H and W.Y; writing-original draft preparation, L.H and Y.L; writing-review and editing, L.H, Y.L and W.Y; supervision, Y.L and W.Y. All authors have read and agreed to the published version of the manuscript.

Declaration of competing interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed the following financial support for the research, authorship, and/or publication of this article: This article was supported by the National Key R&D Program of China (grant number: 2020AAA0109300).

ORCID iD

Yuanyuan Li https://orcid.org/0000-0002-5719-1018

References

1. Wang F, Wang B, Dun B, et al. Remaining life prediction of rolling bearing based on PCA and improved logistic regression model. J Vibroengineering 2016; 18(8): 5192–5203.
2. Tallian TE. Data fitted bearing life prediction model for variable operating conditions. ASLE Trans 1999; 42(1): 241–249.
3. Karasu S, Altan A, Sara Z, et al. Prediction of wind speed with non-linear autoregressive (NAR) neural networks. In: 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, Turkey, 15–18 May 2017.
4. Aslimani N, El-Ghazalii T and Elliaa R. A new chaotic-based approach for multi-objective optimization. Algorithms 2020; 13(9): 204.
5. Peng Y, Cheng J, Liu Y, et al. An adaptive data-driven method for accurate prediction of remaining useful life of rolling bearings. Front Mech Eng 2018; 13(2): 301–310.
6. Fengtao W, Xiaofei L, Chenxi L, et al. Remaining useful life prediction method of rolling bearings based on Pchip-EEMD-GM(1, 1) mode. Shock and Vibration 2018; 2018: 1–10.
7. Lu Y, Li Q, Pan Z, et al. Prognosis of bearing degradation using gradient variable forgetting factor rbbs combined with time series model. IEEE Access 2018; 6: 10986–10995.
8. Kuncan M. An intelligent approach for bearing fault diagnosis: combination of 1D-LBP and GRA. IEEE Access 2020; 8: 137517–137529.
9. Dya B, Bla B, Hla B, et al. Remaining useful life prediction of roller bearings based on improved 1D-CNN and simple recurrent unit. Measurement 2021; 175: 109166.
10. Wang R, Shi R, Hu X, et al. Remaining useful life prediction of rolling bearings based on multiscale convolutional neural network with integrated dilated convolution blocks. Shock and Vibration 2021; 2021: 1–11.
11. Guo L, Li N, Jia F, et al. A recurrent neural network based health indicator for remaining useful life prediction of bearings. Neurocomputing 2020; 349: 96–109.
12. Ma P, Zhang H, Fan W, et al. A novel bearing fault diagnosis method based on 2-D image representation and transfer learning–convolutional neural network. Meas Sci Technol 2019; 30(5).
13. Wu Z, Jiang H, Zhao K, et al. An adaptive deep transfer learning method for bearing fault diagnosis. Measurement 2019; 151: 107227.
14. Zhang R, Fang Y and Zhou Z. Fault diagnosis of rolling bearing based on k-svd dictionary learning algorithm and BP neural network. In: the 2019 International Conference. 12–13 July, 2019, Wuhan, China.
15. Liu X, Liu Z, Liang Z, et al. PSO-BP neural network-based strain prediction of wind turbine blades. Materials 2019; 12(12): 1889.
16. Wan L, Li H, Chen Y, et al. Rolling bearing fault prediction method based on QPSO-BP neural network and dempster–shafer evidence theory. Energies 2020; 13(5): 1094.

Table 5. Comparison of different prediction models.

| Model       | MAE  | MAPE, % | RMSE       | R2-score |
|-------------|------|---------|------------|----------|
| BP prediction | 1.7*10^-3 | 2.67%   | 1.5*10^-3  | 0.5381   |
| GA-BP       | 7.87*10^-4 | 0.67%   | 1.2*10^-3  | 0.6172   |
| LSTM network | 4.93*10^-4 | 0.84%   | 1.2*10^-3  | 0.5159   |
| AMGA-BP     | 3.98*10^-4 | 0.21%   | 2.8969*10^-4 | 0.8348   |

Measurement and Control 55(5-6)
MAE, MAPE, RMSE, and R²-Score, and can provide new insights within the allowable error range. Compared with other methods, the proposed prediction model uses AMGA method to optimize the fitness between populations in the optimization process. The process allows to carry out gene exchange, which limits the exchange two populations to determine whether the two populations are compared. The comparison of the optimal solution in the iterative process of predicting accuracy. AMGA uses the distance and which is higher than BP, GA-BP, and LSTM method in vibration signals obtained in complex environments, and the KPCA method is used to fuse the time.

Table 5.

| Model                  | MAE      | MAPE     | RMSE     | R²-Score |
|------------------------|----------|----------|----------|----------|
| AMGA-BP                | 3.98     | 0.84     | 1.2      | 0.76     |
| BP prediction          | 1.7      | 0.67     | 1.2*10⁻² | 0.8      |
| LSTM prediction        | 2.7      | 0.51     | 1.2*10⁻³ | 0.9      |
| BP                     | 2.1      | 0.45     | 1.2*10⁻⁴ | 0.95     |

Declaration of conflict of interest

W.Y. All authors have read and agreed to the published version of the manuscript.

Author Contributions

Conceptualization, L.H and Y.L; methodology, L.H and Y.L; formal analysis, L.H and Y.L; data curation, L.H and Y.L; supervision, Y.L and W.Y; writing-original draft preparation, L.H and Y.L; writing-review and editing, L.H, Y.L and W.Y; visualization, Y.L; funding acquisition, W.Y.

References

1. Cheng P, Chen D and Wang J. Clustering of the body shape of the adult male by using principal component analysis and genetic algorithm–BP neural network. Soft Comput 2020; 17: 13219–13237.
2. Cheng Z and Cai B. Predicting the remaining useful life of rolling element bearings using locally linear fusion regression. J Intell Fuzzy Syst 2018; 34(6): 1–12.
3. Hu X, Xiao Z, Liu D, et al. KPCA and AE based local-global feature extraction method for vibration signals of rotating machinery. Math Probl Eng 2020; 2020: 5804509.
4. Wang F, Liu X, Deng G, et al. Remaining life prediction method for rolling bearing based on the long short-term memory network. Neural Process Lett 2019; 50: 2437–2454.
5. Kumar Ghadai R, Kalita K, Mondal SC, et al. Genetically optimized diamond-like carbon thin film coatings, materials and manufacturing processes. Mater Manufacturing Process 2019; 34(13): 1476–1487.
6. Kalita K, Mukhopadhyay T, Dey P, et al. Genetic programming assisted multi-scale optimization for multi-objective dynamic performance of laminated composites: the advantage of more elementary-level analyses. Neural Comput Appl 2020; 32(2): 7969–7993.
7. Dou K and Sun X. Long-term weather prediction based on GA-BP neural network. IOP Conf Ser Earth Environ Sci 2021; 668(1): 012015.
8. Lin JW, Chao CT and Chiou JS. Determining neuronal number in each hidden layer using earthquake catalogue data as training data in training an embedded back propagation neural network for predicting earthquake magnitude. IEEE Access 2018; 6: 52582–52597.
9. Professor WS, Professor CC and Professor CHC. Multifractional brownian motion and quantum-behaved particle swarm optimization for short term power load forecasting: an integrated approach. Energy 2020; 194: 116847.1–116847.8.