



A DIAGNOSTIC ALGORITHM DIAGNOSING THE FAILURE OF RAILWAY SIGNAL EQUIPMENT

Yongcheng WU¹, Dejin CAO²

¹Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China, e-mail: dwuw91@163.com

²Lanzhou Institute of Technology, Lanzhou, Gansu 730050

Abstract

Failure of railway signal equipment can cause an impact on its normal operation, and it is necessary to make a timely diagnosis of the failure. In this study, the data of a railway bureau from 2016 to 2020 were studied as an example. Firstly, denoising and feature extraction were performed on the data; then the Adaptive Comprehensive Oversampling (ADASYN) method was used to synthesize minority class samples; finally, three algorithms, back-propagation neural network (BPNN), support vector machine (SVM) and C4.5 algorithms, were used for failure diagnosis. It was found that the three algorithms performed poorly in diagnosing the original data but performed significantly better in diagnosing the synthesized samples, among which the BPNN algorithm had the best performance. The average precision, recall rate and F1 score of the BPNN algorithm were 0.94, 0.92 and 0.93, respectively. The results verify the effectiveness of the BPNN algorithm for failure diagnosis, and the algorithm can be further promoted and applied in practice.

Keywords: failure diagnosis, railway signal equipment, denoising, feature extraction

1. INTRODUCTION

Failure diagnosis has strong engineering applications and plays a very important role in reducing maintenance cycles and improving maintenance quality. As industrial equipment become increasingly complex, more and more attention is paid to failure diagnosis [6]. The earliest failure diagnosis used traditional methods, i.e., detecting signals with some equipment and making judgments on failures according to empirical knowledge. With the development of technology, intelligent methods have emerged [4], such as acoustic diagnosis [8], vibration diagnosis [1], etc., which have been widely used in various complex systems. Jiang et al. [9] proposed a method based on an autoregressive (AR) model and fuzzy clustering for bearing failure diagnosis and found through experiments that the method could identify different types of faulty bearings. Chine et al. [3] used an artificial neural network (ANN)-based method for the failure diagnosis of new photovoltaic systems, validated it using an experimental database of climatic and electrical parameters from a PV string installed at the Renewable Energy Laboratory (REL) of the University of Jijel, and found that the method was able to accurately identify different failures. Zhao et al. [16] combined wavelet packet decomposition (WPD) with multiscale permutation entropy (MPE) to diagnose the failures of rolling bearings, conducted experiments using a data set from the Case Western Reserve University bearing data center, and found that the method was able to identify failures accurately. Cerrada et al. [2]

conducted a study on gear failure detection, used the genetic algorithm and the random forest algorithm to classify several failure types, and obtained a classification accuracy of over 97%. In order to ensure the safe transportation of railways, the diagnosis algorithm of railway signal equipment after failure has become an increasingly important topic. With the continuous development of railway construction, new requirements for the safety of railway signal equipment have been put forward, and better and faster failure diagnosis needs to be achieved in order to meet the requirements of railway operation. In this paper, the failure data of a railway bureau from 2016 to 2020 were processed and analyzed, and the diagnostic performance of different algorithms was compared, in order to make some contributions to realize better equipment fault diagnosis.

2. RAILWAY SIGNAL EQUIPMENT FAILURE DATA AND PROCESSING

Railway signal equipment includes signal machines, track circuits, etc. According to different criteria, railway signal equipment failures can be divided into different categories, as shown in Fig. 1.

In this paper, failure diagnosis was analyzed based on the failure data of the signal equipment of a railway bureau between 2016 and 2020, and ten types of failures that appear more frequently are shown in Table 1.

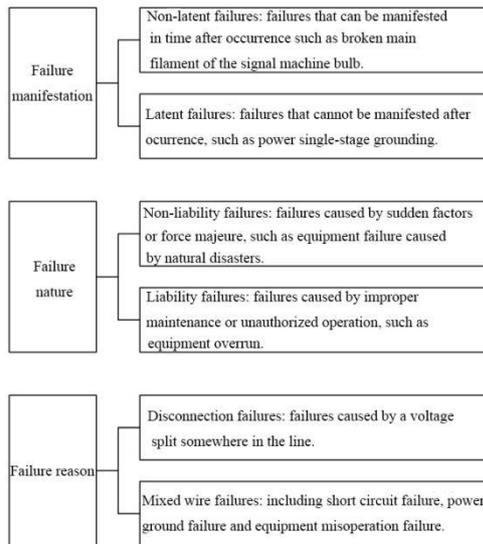


Fig. 1. Classification of railway signal equipment failures

Table 1. Failure data

| Failure type | Sample size |
|---|-------------|
| Normal | 369 |
| Signal machine failure | 95 |
| Centralized Traffic Control (CTC) equipment failure | 27 |
| Train operation monitoring device (LKJ) equipment failure | 18 |
| Train operation dispatching command system (TDCS) equipment failure | 5 |
| Occlusion equipment failure | 5 |
| On-board equipment failure | 29 |
| Turnout failure | 184 |
| Track circuit failure | 296 |
| Microcomputer interlocking failure | 6 |

First, the noise in the data needs to be removed. The method used in this paper is wavelet analysis [12], and its detailed steps are as follows.

(1) It is assumed that the nonlinear variation result of parent function $\varphi(t)$ is $\hat{\varphi}(w)$. If it satisfies $C_\varphi = \int_R \frac{|\hat{\varphi}(w)|^2}{|w|} dw < \infty$, then $\hat{\varphi}(w)$ is the basic wavelet.

(2) Parent function $\varphi(t)$ is scaled and translated to obtain the wavelet sequence $\varphi_{a,b}(t) : \varphi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \varphi\left(\frac{t-b}{a}\right)$, $a, b \in R, a \neq 0$.

(3) Wavelet transform is used to remove noise $n(t)$ in failure data $s(t) = x(t) + n(t)$: $W_{2^j} s(k) = \frac{1}{2^j} \sum_{n=0}^{N-1} s(n) \varphi(2^{-j}n - k)$, where $W_{2^j} s(k)$ refers to the coefficient of the wavelet transform. Then, the real state data $x(t)$ of railway signal equipment is obtained.

(4) For the formula in step (3), the data are processed by wavelet decomposition using recursive realization. The calculation process can be written as: $xf(j+1, k) = xf(j, k)h(j, k)$ and $Wf(j+1, k) = xf(j, k)g(j, k)$.

(5) The wavelet transform coefficient is set as 0, and then wavelet reconstruction is performed on the failure data to obtain the noise-free data, i.e., $xf(j-1, k) = xf(j, k)h_0(j, k)g_0(j, k)$.

After removing the noise from the data, feature extraction is performed using empirical modal decomposition (EMD) [5]. The detailed steps are as follows.

(1) n modal functions $c_i(t)$ ($i = 1, 2, \dots, n$) are generated.

(2) The modal moment is obtained from the modal function component: $E_i = \int_{-\infty}^{+\infty} t |c_i(t)|^2 dt$.

(3) Multiple energy moments are combined to obtain the failure feature vector: $T = [E_1 \ E_2 \ \dots \ E_n]$.

(4) Normalization is performed to obtain $T' = \left[\frac{E_1}{E} \ \frac{E_2}{E} \ \dots \ \frac{E_n}{E} \right]$, where $E = (\sum_{i=1}^n |E_i|^2)^{1/2}$.

The calculated feature vectors for different failures are shown in Table 2.

Table 2. Failure feature vectors

| Failure type | E_0 | E_1 | E_2 | E_3 | E_4 |
|---------------------------------|-------|-------|-------|-------|-------|
| Normal | 0.121 | 0.212 | 0.213 | 0.118 | 0.336 |
| Signal machine failure | 0.223 | 0.115 | 0.394 | 0.182 | 0.086 |
| CTC equipment failure | 0.056 | 0.128 | 0.723 | 0.016 | 0.077 |
| LKJ equipment failure | 0.125 | 0.264 | 0.348 | 0.152 | 0.111 |
| TDCS equipment failure | 0.121 | 0.315 | 0.088 | 0.215 | 0.261 |
| Occlusion equipment failure | 0.067 | 0.128 | 0.324 | 0.259 | 0.222 |
| On-board equipment failure | 0.178 | 0.215 | 0.321 | 0.165 | 0.121 |
| Turnout failure | 0.321 | 0.391 | 0.213 | 0.064 | 0.011 |
| Track circuit failure | 0.086 | 0.112 | 0.398 | 0.172 | 0.232 |
| Microcomputer interlock failure | 0.097 | 0.189 | 0.168 | 0.179 | 0.367 |

3. FAILURE DIAGNOSIS ALGORITHM

3.1. Generation of minority class samples

It was seen from Table 1 that some of the failures have a small amount of data, which is not conducive to failure diagnosis; therefore, this paper uses the Adaptive Comprehensive Oversampling (ADASYN) [14] method to synthesize minority class samples. Its principle is to find out a probability distribution r_i and take it as a criterion to determine how many samples need to be synthesized for every minority class sample.

First, the imbalance degree is calculated: $d = \frac{m_l}{m_s}$, where m_l is the number of less-category samples, m_s is the number of multi-category samples, and $d \in (0,1]$. Then, the number of samples to be synthesized is calculated: $G = (m_s - m_l) \times \beta$. $\beta = 1$ means that the sample types are perfectly balanced after adding the synthetic samples.

For a minority class sample x_i , its K -nearest neighbors in the n -dimensional space are found, and the ratio is calculated using $r_i = \frac{\Delta_i}{K}$, where Δ_i refers to the number of minority class samples among K neighbors of x_i . For $\hat{r}_i = \frac{r_i}{\sum_{i=1}^{m_l} r_i}$, r_i is regularized; then, r_i is the probability distribution, $\sum \hat{r}_i = 1$.

Finally, the failure data are supplemented by calculating the number of samples that needed to be synthesized in every class through the equation $g_i = \hat{r}_i \times G$, where g_i stands for the number of samples that need to be synthesized and G stands for the total number of the synthesized samples.

3.2. Classification algorithm

After denoising and feature extraction of the data, the failure data are diagnosed using classification algorithms. Three commonly used classification algorithms are mainly studied in this paper.

(1) Back-propagation neural network (BPNN) algorithm [15]: the input layer is the feature vector of failure data, containing five nodes. The output layer is the failure type, containing 10 nodes. The hidden layer is determined based on the empirical formula: $l = \sqrt{m + n} + a$, $1 \leq a \leq 10$. n , m and l represented the numbers of nodes in the input layer, output layer, and hidden layer, respectively. After calculation, the number of nodes in the hidden layer is 11. The L-M algorithm is used for training, and the weight adjustment rate is: $\Delta w = (J^T J + \mu I)^{-1} \cdot J^T e$, where w is a weight, J is the Jacobi matrix, e is the error vector, and μ is a scalar.

(2) Support vector machine (SVM) algorithm [13]: a hyperplane $f(x)$ was used for classification. The distance from sample point $\{x_i, y_i\}$ to the hyperplane can be written as: $d_i = \frac{|w^T x_i + b - y_i|}{\sqrt{1 + \|w\|^2}}$, where w^T refers to a weight and b refers to an intercept. In order to find the optimal hyperplane, it

is necessary to minimize $\|w\|^2$. The objective function can be written as: $\min \left(\frac{1}{2} \|w\|^2 + \frac{C}{n} \sum_{i=1}^n (\delta_i + \delta_i^*) \right)$, where δ_i and δ_i^* are the relaxation variables and C is the penalty factor.

(3) C4.5 algorithm [11]: Suppose there are k categories ($C_i (i = 1, 2, \dots, k)$) in sample D . The probability of any sample belonging to C_i is $p_i = \frac{|D_i|}{|D|}$. The classification information entropy is: $\text{entropy}(D) = -\sum_{i=1}^k p_i \log_2(p_i)$. It is assumed that attribute A has m different values (v_j), the information entropy of every value is: $\text{entropy}(A = v_j) = -\sum_{i=1}^k p_{ij} \log_2(p_{ij})$. The information entropy of A for data division is: $\text{entropy}(A) = -\sum_{j=1}^m \frac{|D_{A_j}|}{|D|} \times \text{entropy}(A = v_j)$. Finally, the information gain of attribute A is: $\text{gain}(A) = \text{entropy}(D) - \text{entropy}(A)$, the information entropy of attribute A is: $H(A) = -\sum_{j=1}^m p_j \log_2(p_j)$, and the information gain rate of attribute A is: $\text{gainratio}(A) = \frac{\text{gain}(A)}{H(A)}$. By solving the information gain rate of every attribute, the attribute with the largest value is used as the root node to generate a decision tree.

4. EXPERIMENTAL ANALYSIS

First, minority class samples were generated by the ADASYN method, and the results are shown in Table 3.

Table 3. Minority class samples generated by the ADASYN method

| Failure type | Raw data | Post-generation data |
|------------------------------------|----------|----------------------|
| Normal | 369 | 369 |
| Signal machine failure | 95 | 95 |
| CTC equipment failure | 27 | 27 |
| LKJ equipment failure | 18 | 18 |
| TDCS equipment failure | 5 | 182 |
| Occlusion equipment failure | 5 | 120 |
| On-board equipment failure | 29 | 29 |
| Turnout failure | 184 | 184 |
| Track circuit failure | 296 | 296 |
| Microcomputer interlocking failure | 6 | 151 |

It was seen from Table 3 that after the synthesis by the ADASYN method, the data volume increased for all the failures that originally had a small number of samples, for example, the data volume of TDCS equipment failure increased from 5 to 182. The

balanced data set is more conducive to the failure diagnosis, thus achieving better diagnostic results.

Three indicators were used to evaluate the diagnostic performance of the algorithms:

$$\text{precision: } \text{precision} = \frac{1}{|C|} \sum_{i \in C} \frac{(TP_i + TN_i) \times TP_i}{TP_i + FP_i},$$

$$\text{recall rate: } \text{recall} = \frac{1}{|C|} \sum_{i \in C} \frac{(TP_i + TN_i) \times TP_i}{TP_i + FN_i},$$

$$F1 \text{ score: } F_1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}},$$

where TP stands for positive samples classified as positive, TN stands for negative samples classified as negative, FP stands for negative samples classified as positive, and FN stands for positive samples classified as negative.

Seventy percent of the data set was used for the training of the algorithm and 30% for the testing of the algorithm. 10-fold cross-validation was adopted. Firstly, the original data were tested. The performance of different algorithms for failure diagnosis is shown in Fig. 2.

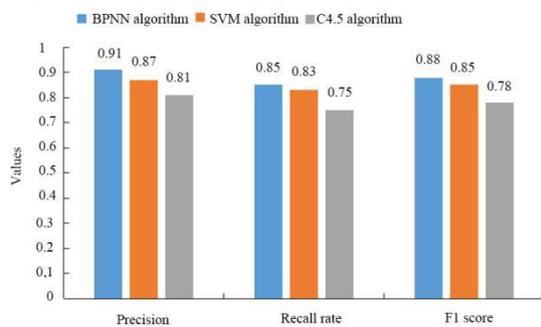


Fig. 2. Failure diagnosis results based on original data

It was seen from Fig. 2 that when experiments were conducted using the original data, the precision, recall rate and F1 score of different algorithms were lower; the highest precision was 0.91, and the lowest was 0.81; the highest recall rate was 0.85, and the lowest was 0.75; the highest F1 score was 0.88, and the lowest was 0.78. In conclusion, the BPNN algorithm had the best performance, followed by the SVM algorithm and the C4.5 algorithm.

Then, the failure data were diagnosed using the synthesized samples, and the precision of different algorithms is shown in Fig. 3.

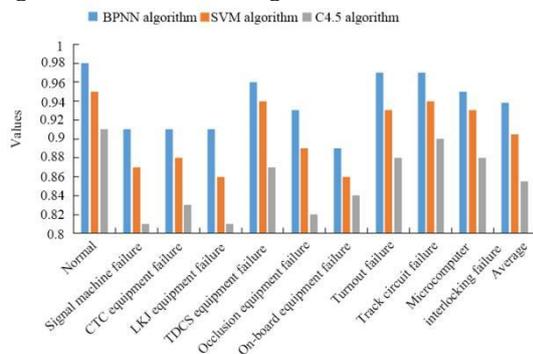


Fig. 3. Comparison of precision between different algorithms

It was seen from Fig. 3 that when using the synthesized samples for fault diagnosis, the diagnostic precision of different algorithms for different failures was above 0.8; the precision of diagnosing normal samples was the highest, and the precision of diagnosing on-board equipment failure was the lowest. In conclusion, the BPNN algorithm was better than the SVM algorithm, and the SVM algorithm was better than the C4.5 algorithm. The average precision of the three algorithms was 0.94, 0.91 and 0.86, respectively; the average precision of the BPNN algorithm was 0.03 higher than the SVM algorithm and 0.08 higher than the C4.5 algorithm.

The recall rates of different algorithms are shown in Fig. 4.

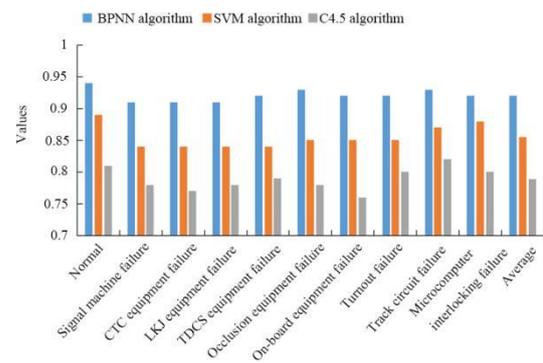


Fig. 4. Comparison of recall rates between different algorithms

It was seen from Fig. 4 that for different failures, the recall rates of different algorithms were all above 0.7; the BPNN algorithm had the highest recall rate, followed by the SVM algorithm and the lowest recall rate; the highest and lowest recall rates of the BPNN algorithm were 0.94 and 0.91, respectively; the average recall rate of the three algorithms were 0.92, 0.86, and 0.79, respectively; the average recall rate of the BPNN algorithm was 0.06 higher than the SVM algorithm and 0.13 higher than the C4.5 algorithm.

Finally, the F1 scores of different algorithms were compared, and the results are shown in Fig. 5.

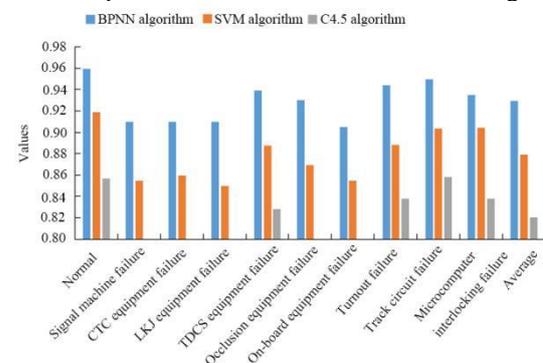


Fig. 5. Comparison of F1 scores between different algorithms

It was seen from Fig. 5 that the F1 score of all the three algorithms was above 0.8; the highest and lowest F1 score of the BPNN algorithm was 0.96 and

0.91, respectively, the F1 scores of the SVM algorithm were below 0.92, and the F1 scores of the C4.5 algorithm were below 0.9, indicating that the BPNN algorithm had the best performance. The average F1 score of the three algorithms was 0.93, 0.88 and 0.82. The average F1 score of the BPNN algorithm was 0.05 larger than the SVM algorithm and 0.11 larger than the C4.5 algorithm.

5. DISCUSSION

Railway signal equipment needs to be kept in good working condition in order to maintain the normal operation of the railway; therefore, the railway needs to carry out regular and irregular strict maintenance to ensure the safety of traffic. However, the maintenance cannot predict the failure of the equipment, and repeated maintenance and inspection also need to consume a lot of time and energy [10]. The real-time nature of fault diagnosis is receiving increasing attention [7]. It is very important to improve the maintenance level and ensure system reliability by providing information about the equipment and achieving timely mastery of the equipment status through modern technology.

In this paper, failures of railway signal equipment were briefly analyzed and introduced, the fault data were processed by denoising and feature extraction, and failure diagnosis experiments were conducted using three algorithms. The results showed that the BPNN algorithm had the best performance when diagnosing the original samples, and the precision, recall rate and F1 score were 0.91, 0.85 and 0.88, respectively. After synthesis by the ADASYN method, the samples were more balanced, and the effect of failure diagnosis significantly improved. The BPNN algorithm performed best in the precision, recall rate and F1 score, followed by the SVM algorithm and the C4.5 algorithm, indicating that the BPNN algorithm showed the best performance in failure diagnosis and was more capable of identifying different types of failures. The average precision of BPNN, SVM and C4.5 algorithms were 0.94, 0.91 and 0.86, respectively, the recall rates were 0.92, 0.86 and 0.79, respectively, and the F1 scores were 0.93, 0.88 and 0.82, respectively, which verified the effectiveness of the BPNN algorithm in fault diagnosis.

Although this paper has obtained some results involving failure diagnosis of railway signal equipment, there are many shortcomings that need further research in the future, such as:

- (1) analyzing a wider range of failure types;
- (2) comparing and studying more failure diagnosis algorithms.

6. CONCLUSION

This study designed the data denoising, feature extraction, minority class sample synthesis, and diagnosis algorithm based on the data of a railway bureau between 2016 and 2020 and conducted the

experimental analysis. The experiment found that the diagnosis performance of different algorithms effectively improved after the synthesis of minority class samples, and the BPNN algorithm had the best performance in failure diagnosis in the comparison of BPNN, SVM, and C4.5 algorithms. The experimental results verify that the algorithm designed in this study is reliable and effective in diagnosing the failure of railway signal equipment; thus, the BPNN algorithm can be further promoted and applied in practice to further improve the diagnostic performance of failures in railway signal equipment and satisfy the demand of safe railway operation.

REFERENCES

1. Cerrada M, Sánchez RV, Cabrera DR, Zurita GV, Li C. Multi-Stage Feature Selection by Using Genetic Algorithms for Fault Diagnosis in Gearboxes Based on Vibration Signal. *Sensors* 2015; 15(9): 23903-23926. <http://dx.doi.org/10.3390/s150923903>
2. Cerrada M, Zurita G, Cabrera D, Sánchez R, Artés M, Li C. Fault diagnosis in spur gears based on genetic algorithm and random forest. *Mechanical Systems & Signal Processing* 2016; 70-71: 87-103.
3. Chine W, Mellit A, Lughri V, Malek A, Sulligoi G, Pavan AM. A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. *Renewable Energy* 2016; 90: 501-512. <http://dx.doi.org/10.1016/j.renene.2016.01.036>
4. Chine W, Mellit A, Lughri V, Malek A, Sulligoi G, Pavan AM. A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. *Renewable Energy* 2016; 90: 501-512. <http://dx.doi.org/10.1016/j.renene.2016.01.036>
5. Daubechies I, Lu J, Wu H T. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Applied and Computational Harmonic Analysis*. 2011;30(2):243-261. <http://dx.doi.org/10.1016/j.acha.2010.08.002>
6. Gao Z, Cecati C, Ding S X. A Survey of Fault Diagnosis and Fault-Tolerant Techniques—Part I: Fault Diagnosis With Model-Based and Signal-Based Approaches. *IEEE Transactions on Industrial Electronics*. 2015;62(6):3757-3767. <http://dx.doi.org/10.1109/TIE.2015.2417501>
7. Gao Z, Ding S X, Cecati C. Real-time fault diagnosis and fault-tolerant control. *IEEE Transactions on Industrial Electronics* 2015; 62(6): 3752-3756. <http://dx.doi.org/10.1109/TIE.2015.2417511>
8. Glowacz A. Acoustic based fault diagnosis of three-phase induction motor. *Applied Acoustics* 2018; 137(AUG.):82-89. <http://dx.doi.org/10.1016/j.apacoust.2018.03.010>
9. Jiang L, Liu Y, Li X, Chen A. Degradation Assessment and Fault Diagnosis for Roller Bearing Based on AR Model and Fuzzy Cluster Analysis. *Shock & Vibration* 2015;18(1-2):127-137. <http://dx.doi.org/10.3233/SAV-2010-0572>
10. Li B, Wu S, Wang Z, et al. Railway Track Circuit Signal State Check Using Object Detection. *Journal of Physics: Conference Series* 2020; 1486(4): 042018 (7pp).
11. Ngoc P V, Ngoc C, Ngoc T, Dat ND. A C4.5 algorithm for English emotional classification. *Evolving*

- Systems. 2017:1-27.
<http://dx.doi.org/10.1007/s12530-017-9180-1>
12. Pu H, Xie A, Sun D W, Kamruzzaman M, Ma J. Application of Wavelet Analysis to Spectral Data for Categorization of Lamb Muscles. *Food & Bioprocess Technology*. 2015;8(1):1-16.
<http://dx.doi.org/10.1007/s11947-014-1393-8>
 13. Shi J, Lee W J, Liu Y, Yang Y, Wang P. Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines. *IEEE Transactions on Industry Applications* 2015; 48(3): 1064-1069.
<http://dx.doi.org/10.1109/TIA.2012.2190816>
 14. Tang B, He H. Kernel ADASYN: Kernel based adaptive synthetic data generation for imbalanced learning. *Evolutionary Computation* 2015: 664-671.
<http://dx.doi.org/10.1109/CEC.2015.7256954>
 15. Zhang Y, Gao X, Katayama S. Weld appearance prediction with BP neural network improved by genetic algorithm during disk laser welding. *Journal of Manufacturing Systems* 2015; 34: 53-59.
<http://dx.doi.org/10.1016/j.jmsy.2014.10.005>
 16. Zhao L Y, Wang L, Yan R Q. Rolling Bearing Fault Diagnosis Based on Wavelet Packet Decomposition and Multi-Scale Permutation Entropy. *Entropy* 2015; 17(9):6447-6461.
<http://dx.doi.org/info:doi/10.3390/e17096447>
-

Received 2021-08-06

Accepted 2021-11-14

Available online 2021-11-20



Yongcheng WU, born in 1983, has received the master's degree from Lanzhou Jiaotong University in 2013. He is a lecturer in Lanzhou Jiaotong University. He is interested in traffic information engineering and control.



Dejin CAO, born in 1983, received the bachelor's degree from China University of Geosciences. He is an engineer. He is interested in industrial automatic control.