Compressor fault diagnosis system based on PCA-PSO-LSSVM algorithm

Kun Zhang¹, Jinpeng Su¹,², Shaoan Sun¹, Zhixiang Liu³, Jinfeng Wang¹, Mingchao Du¹, Zengkai Liu¹ and Qiang Zhang¹
¹Key Laboratory for Robot & Intelligent Technology of Shandong Province, Shandong University of Science and Technology, Qingdao, Shandong, China
²State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai, China
³School of Mechanical Engineering, Liaoning Technical University, Fuxin, Liaoning, China

Abstract
On the basis of the principal components analysis-particle swarm optimization-least squares support vector machine (PCA-PSO-LSSVM) algorithm, a fault diagnosis system is proposed for the compressor system. The relationship between the working principle of a compressor system, the fault phenomenon, and the root cause is analyzed. A fault diagnosis model is established based on the LSSVM optimized using PSO, the compressor fault diagnosis test experimental platform is used to obtain the fault signal of various fault occurrence states, and the PCA algorithm is employed to extract the characteristic data in the fault signal as input to the fault diagnosis model. The backpropagation neural network, the LSSVM algorithm, and the PSO-LSSVM algorithm are analyzed and compared with the proposed fault diagnosis model. Results show that the PCA-PSO-LSSVM fault diagnosis model has a maximum fault recognition efficiency that is 10.4% higher than the other three models, the test sample classification time is reduced by 0.025 s, the PCA algorithm can effectively reduce the input dimension, and the PSO-LSSVM fault diagnosis model based on the PCA algorithm for extracting features has a high recognition rate and accuracy. Therefore, the proposed fault diagnosis system can effectively identify the compressor fault and improve the efficiency of the compressor.

Keywords
Compressor, fault diagnosis, least squares support vector machine, particle swarm optimization
Introduction

Air compressors are the key equipment for coal mining engineering to provide power for pneumatic tools. However, under full or over working loads and the harsh working environment, various failures will occur under the influence of numerous influencing factors during the work process. Moreover, the relationship between failures and the cause show a strong nonlinearity, and mathematical models cannot be used to express the faults in the diagnosis of mining air compressors. Thus, early warning of failures must be achieved based on the full analysis of huge data and information to discover existing risks in a timely manner. Only in this way can we ensure the stable and healthy development of intelligent coal mining enterprises.

Zhou et al. proposed an online compressor liquid floodback fault diagnosis method for the VRF system based on the back-propagation neural network (BPNN), which fills the online compressor liquid floodback fault diagnosis knowledge gap. Van Tung et al. proposed a hybrid deep belief network to diagnose the single and combined faults of suction and discharge valves in a reciprocating compressor. Zhang et al. proposed a novel convolutional deep belief network-based method and employed a novel framework that fuses multi-source information to improve the performance of fault diagnosis. Data were collected from an industrial plant to verify the proposed method. De Cooman et al. proposed a transformer fault diagnosis method based on the modified gray wolf optimization algorithm and support vector machine (SVM), realizing the optimization of the penalty factor, and the kernel parameter in SVM. Li et al. proposed a multiscale local feature learning method based on BPNN for rolling bearings fault diagnosis. Based on the local characteristics of the fault features in the time and frequency domains, the BPNN is used to locally learn meaningful and dissimilar features from signals of different scales, thus the accuracy of fault diagnosis is improved obviously. Zhang et al. proposed a fault diagnosis method based on multi-classification relevance vector machine for high voltage circuit breakers. On the basis of the ASHRAE RP-1043 data, Fan et al. extracted information from three of the factory-installed (FI) sensors along with information from all eight of the FI sensors in order to establish the SVM-3 and SVM-8 diagnostic models based on grid search and cross validation parameter optimization for seven typical faults. Wang et al. presented the results of a research on hybrid fault diagnosis techniques that utilize SVM and improved particle swarm optimization (PSO) to perform further diagnosis based on qualitative reasoning through knowledge-based preliminary diagnosis and the sample data provided by an on-line simulation model.

The above fault diagnosis model for compressors is mainly for equipment on the ground. Many machinery and equipment underground in coal mines will also use compressors for auxiliary operations. The underground environment makes the compressors used differ from conventional ground compressors, and the occurrence of accidents in the event of failure will cause major safety hazards. Now, more advanced deep learning and other fault diagnosis models require numerous experimental test samples. Given the complexity of the working environment of mining
machinery and the equipment failure rate, the amount of collected data fails to meet the large sample data requirements, so some advanced fault diagnosis models cannot be used for fault diagnosis in this condition. Therefore, this study aims to effectively reduce the latitude of the sample input of the fault diagnosis model based on the principal components analysis (PCA) algorithm, increase the recognition efficiency and sample classification time of the PSO-LSSVM algorithm, establish fault diagnosis based on the least squares SVM (LSSVM) model optimized using PSO, use the PCA algorithm to extract characteristic data from the fault signal as input to the fault diagnosis model, use the compressor fault diagnosis test experimental platform to obtain the fault signal when various faults occur, and use four different fault diagnosis models. The signals are compared and analyzed by training to verify the high-efficiency recognition rate and accuracy of the proposed PCA-PSO-LSSVM fault diagnosis model, which provides a new fault diagnosis method for mining compressors.

Analysis of compressor fault diagnosis system

The establishment of a diagnostic model is the key to the fault diagnosis system for air compressors. Therefore, the air compressor fault diagnosis model should be established based on a full understanding of the air compressor instructions and related documents, and on-site monitoring records and references should be fully analyzed. The relevant experience of maintenance personnel is useful for analyzing and classifying the faults that occur during the operation of air compressors and finding suitable methods to solve these faults.

The work of the air compressor is divided into three processes, the suction process, the compression and injection process, and the exhaust process. Through these three processes, the air is compressed and transmitted to the machinery that needs to provide power. According to compressor field experience and the literature, the main four failure modes of compressors are as follows: compressor cooling device failure (S2), compressor lubrication device failure (S3), compressor bearing failure (S4), and compressor power device failure (S5).

The following features appear in the failure mode: insufficient compressor discharge (A), extremely low exhaust pressure (B), extremely high compressor discharge temperature (C), extremely high compressor cooling water temperature (D), extremely low compressor cooling water pressure (E), extremely low main engine speed (F), extremely large compressor vibration (G), extremely high lubricating oil temperature (H), extremely low lubricating oil pressure (I), and extremely high bearing temperature (J).

According to the different types of faults, the corresponding fault states are as follows:

S₁ is the normal state, and the set vector is (1,0,0,0,0).
S₂ is the state when the compressor cooling water system fails, and the set vector is (0,1,0,0,0).
S₃ is the state when the compressor lubrication system fails, and the set vector is (0,0,1,0).
S₄ is the state when the compressor bearing fails, and the set vector is (0,0,0,1,0).
S₅ is the compressor power supply system failure, the set vector is (0,0,0,0,1).

The compressor’s failure status, type, and symptoms are shown in Table 1.

The fault diagnosis model studied in this work mainly includes the cooling temperature, the cooling pressure, the exhaust pressure, the bearing temperature, the lubricating oil pressure, the lubricating oil temperature, the exhaust volume, the exhaust temperature obtained by the compressor under normal and various fault conditions, the host speed, and the host vibration. The mapping relationship between the fault state and the characteristic data established by the LSSVM algorithm realizes the fault diagnosis of the compressor system.

### Fault diagnosis model algorithm

**Least squares support vector machine**

The training samples of the fault diagnosis model are set as \( \{(x_i, y_i) | x_i \in \mathbb{R}^n, i = 1, 2, \cdots, l\} \), where \( l \) is the number of training samples of the fault diagnosis model, \( x_i \) is the input vector of the fault diagnosis model, and \( y_i \) is the output of the fault diagnosis model. In high-dimensional feature space \( H \), the sample set of fault diagnosis can be expressed as the following mapping relationship:

\[
y(x) = \langle \omega \varphi(x) \rangle + b.
\]  

Relaxation factors \( \xi \) and \( \xi^* \) ensure that the fitting error obtained by the linear fitting of the fault diagnosis model is minimized as follows:

\[
\text{Min}J(\omega, \xi, \xi^*) = \frac{1}{2}(\omega \cdot \omega^T) + c \sum_{i=1}^{l} (\xi + \xi^*).
\]  

The constraints are expressed as follows:
The LSSVM adopts the principle of regularization and transforms the above two formulas as follows through the least squares function and equality constraints:

\[
\begin{align*}
    y_i - \omega \cdot \varphi(x_i) - b & \leq \varepsilon + \xi_i \\
    \omega \cdot \varphi(x_i) + b - y_i & \leq \varepsilon + \xi_i^*
\end{align*}
\] (3)

The partial derivation of constraints is expressed as follows:

\[
\begin{bmatrix}
    0 & e^T \\
    e & K + \frac{1}{\tau}
\end{bmatrix}
\begin{bmatrix}
    b \\
    \alpha
\end{bmatrix} =
\begin{bmatrix}
    0 \\
    Y
\end{bmatrix}. 
\] (6)

The regression function for the constraints after the partial derivative is solved as follows:

\[
y(x) = \sum_{i=1}^{l} \alpha_i k(x, x_i) + b. 
\] (7)

**PSO algorithm**

As an improved SVM algorithm, the inequality constraints in the SVM algorithm can be replaced with the equality constraints in the LSSVM algorithm, the error square sum loss function is the training set loss, and the linear equations solution method can be used to solve the quadratic programming problem. The solution is simplified and the convergence accuracy and speed are improved by reducing the solution requirements. However, the network search method selected by the conventional LSSVM recognition model takes a long time to determine the parameters and can only be performed on the grid points during the search, so the parameters cannot be guaranteed when there is no more appropriate grid size effectiveness.16–18

The following steps are taken to optimize the algorithm.

First, the parameters, the maximum iterations, the weight end value, the weight initial value, the learning factor, and the population size in the LSSVM algorithm are initialized.

Second, the optimal position of the entire group and the individual optimal position in the PSO algorithm are determined. The current position of each particle is used to determine the current parameters of LSSVM, and the error after training determines the fitness of each particle. If the value of the current particle itself is optimal, then the current position is regarded as the optimal position. If the fitness value is the same as the fitness value of the entire group, then it can be replaced with the optimal position in the group.
The precocious problem in the optimization algorithm is handled by the following formula:

\[
\begin{align*}
W_{\text{end}}(k) &= w_{\text{min}} \\
W_{\text{start}}(k) &= w (w_{\text{min}}^{\text{max}} | x (k/k_{\text{max}}^2) )^{\text{max}}.
\end{align*}
\] (8)

Third, the fitness value of the progeny particles is solved. If the adaption value of the parent particle is greater than the adaption value of the particle, then the maximum value is the speed of the particle, so that the local minimum value can be prevented from appearing in the particle.

The inertial weights are solved by the following formula, so that the optimization algorithm guarantees better convergence efficiency and accuracy:

\[
\begin{align*}
W_{\text{end}}(k) &= w_{\text{min}} \\
W_{\text{start}}(k) &= w (w_{\text{min}}^{\text{max}} | x (k/k_{\text{max}}^2) )^{\text{max}}.
\end{align*}
\] (9)

Finally, the optimization algorithm is ended according to the termination condition.

**PCA principle steps**

The mean and variance of the original data matrix are solved as follows:

\[
\bar{X}_j = \frac{1}{m} \sum_{i=1}^{m} x_{ij}, S_j = \sqrt{\frac{1}{m-1} \sum_{i=1}^{m} (x_{ij} - \bar{X}_j)^2}, j = 1, 2, \ldots, n.
\] (10)

The normalized matrix of the original data matrix is solved by the following formula:

\[
Z_{ij} = \frac{x_{ij} - \bar{X}_j}{S_j}, i = 1, 2, \ldots, m, j = 1, 2, \ldots, n.
\] (11)

The covariance matrix of the standardized matrix is solved as follows:

\[
R_{n \times n} = \frac{Z^T Z}{m - 1} = (r_{ij})_{n \times n},
\] (12)

\[
r_{ij} = \frac{1}{m - 1} \sum_{k=1}^{m} Z_{ki} \times Z_{kj}, i, j = 1, 2, \ldots, n.
\] (13)

The variance contribution rate of the principal component is solved by the following formula, the number of principal components is set according to the judgment threshold of 85%, and the remaining principal components are ignored:

\[
\sum_{j=1}^{k} \lambda_j / \sum_{j=1}^{n} \lambda_j \geq 85\%.
\] (14)
Using the PCA algorithm to extract data features as input to the fault diagnosis model, the fault diagnosis model based on the LSSVM optimized using PSO is shown in Figure 1.

**Figure 1.** PSO-LSSVM fault diagnosis model based on PCA algorithm to extract features.

Using the PCA algorithm to extract data features as input to the fault diagnosis model, the fault diagnosis model based on the LSSVM optimized using PSO is shown in Figure 1.

**Case analysis of compressor system fault diagnosis**

*Fault diagnosis data collection*

The compressor fault diagnosis test experimental platform is adopted to collect various fault characteristic signals as shown in Figure 2. Compressor data are collected in the normal state, the compressor cooling water system failure state, the compressor lubrication device failure state, the compressor bearing failure state, and the compressor power device failure state to obtain the compressor data of the compressor system in different states.

In the test, an Embraco NJ9226GK compressor is used for the experimental test. An Audi Speed SH0002-000 Hall speed sensor is utilized to collect the speed signal of the compressor, an LD250 ultra-low frequency two-axis vibration acceleration sensor is used to collect the vibration signal of the compressor, and MIK-P300 series pressure and temperature sensors are employed to collect signals of exhaust pressure, exhaust temperature, cooling pressure, cooling temperature, bearing temperature, lubricant pressure, and lubricant temperature. The collected signals are transmitted to the signal storage computer via wired transmission. Tables 2 and 3 list the technical parameters of the main centralized signal acquisition sensors.
The amount of data under various failure states is limited due to the low frequency of the various failures of the compressor. Forty sets of data under each failure state are obtained. Among the data sets, 20 groups are selected randomly for training the fault diagnosis model, and the other 20 groups are used for the generalization ability test of the fault diagnosis model. Part of the data is shown in Table 4.

### Table 2. Main technical parameters of Sh0002-000 Hall sensor.

| Working temperature | Working voltage | Output waveform | Duty cycle | Phase difference | Working frequency |
|---------------------|-----------------|-----------------|------------|------------------|------------------|
| −40°C–150°C         | 4.5–18 V        | Two square waves| 50% ± 10%  | 90° ± 10°        | 20 Hz–15 kHz      |

### Table 3. Main technical parameters of Ld250 ultra-low frequency two-axis vibration acceleration sensor.

| Working temperature | Non linearity | Resolution | Thermal zero drift (−40°C–125°C) | Thermal zero temperature drift (−40°C–125°C) | Sensitivity temperature drift (−40°C–125°C) | Frequency response |
|---------------------|---------------|------------|-------------------------------|--------------------------------|----------------------------------|-------------------|
| −40°C–125°C         | ±0.2%         | 1 mg       | ±1%                           | ±1% FSO                     | 0–2600 Hz                      |

The amount of data under various failure states is limited due to the low frequency of the various failures of the compressor. Forty sets of data under each failure state are obtained. Among the data sets, 20 groups are selected randomly for training the fault diagnosis model, and the other 20 groups are used for the generalization ability test of the fault diagnosis model. Part of the data is shown in Table 4.

**Fault diagnosis model construction**

Increasing the input dimension will increase the calculation time of the fault diagnosis model based on the LSSVM algorithm. Therefore, the MATLAB software is used to perform PCA on the data collected for compressor fault diagnosis analysis.
| Order number | Exhaust pressure (MPa) | Exhaust capacity (M³/min) | Exhaust temperature (°C) | Cooling water temperature (°C) | Speed (r/min) | Vibration (mm/s²) | Lubricating oil temperature (°C) | Lubricant pressure (MPa) | Bearing temperature (°C) | Fault type |
|--------------|------------------------|---------------------------|--------------------------|-------------------------------|--------------|----------------|-----------------------------|------------------------|------------------------|------------|
| 1            | 0.85                   | 31.6                      | 40.1                     | 21.1                          | 0.43         | 1614          | 9.5                         | 79.1                   | 0.56                   | 120.8      |
| 2            | 0.65                   | 30.7                      | 37.7                     | 23.5                          | 0.59         | 1555          | 6.5                         | 91.4                   | 0.71                   | 99.3       |
| 3            | 0.67                   | 31.3                      | 40.3                     | 22.7                          | 0.37         | 1604          | 8.9                         | 79.5                   | 0.64                   | 100.6      |
| 4            | 0.69                   | 26.4                      | 41.4                     | 21.5                          | 0.4          | 1546          | 9                           | 83.8                   | 0.71                   | 114.3      |
| 5            | 0.77                   | 29.4                      | 37.6                     | 30.1                          | 0.16         | 1589          | 9.2                         | 81.2                   | 0.48                   | 101        |
| 6            | 0.74                   | 31.4                      | 41.3                     | 22.8                          | 0.29         | 1591          | 7.2                         | 83.1                   | 0.67                   | 119.6      |
| 7            | 0.73                   | 29.3                      | 43.7                     | 37.4                          | 0.21         | 1632          | 10.2                        | 83.2                   | 0.71                   | 117.3      |
| 8            | 0.79                   | 28.4                      | 46.8                     | 33.1                          | 0.34         | 1611          | 6.9                         | 86.4                   | 0.6                    | 101        |
| 9            | 0.62                   | 29.6                      | 40.4                     | 35.7                          | 0.37         | 1476          | 12.4                        | 87.5                   | 0.33                   | 112        |
| 10           | 0.83                   | 24.3                      | 43.9                     | 38.1                          | 0.4          | 1513          | 9.2                         | 100.9                  | 0.28                   | 121.3      |
| 11           | 0.77                   | 27.7                      | 37.4                     | 32.4                          | 0.33         | 1573          | 14.5                        | 84.5                   | 0.48                   | 117        |
| 12           | 0.72                   | 30.4                      | 44.9                     | 34.6                          | 0.4          | 1500          | 8.5                         | 97.8                   | 0.51                   | 102.2      |
| 13           | 0.46                   | 18.4                      | 37.1                     | 27.6                          | 0.39         | 1433          | 10.7                        | 80.5                   | 0.62                   | 115.3      |
| 14           | 0.45                   | 20.9                      | 37.4                     | 28.4                          | 0.48         | 1318          | 16.4                        | 87.5                   | 0.59                   | 119.8      |
| 15           | 0.47                   | 17.7                      | 37.6                     | 22.8                          | 0.43         | 1474          | 8                           | 80.2                   | 0.63                   | 121        |
| 16           | 0.45                   | 27.9                      | 40.4                     | 18.5                          | 0.53         | 1588          | 17.5                        | 96.9                   | 0.38                   | 151        |
| 17           | 0.46                   | 18.3                      | 38.3                     | 26.4                          | 0.34         | 1373          | 17.2                        | 98.5                   | 0.45                   | 143.3      |
| 18           | 0.37                   | 20.4                      | 38.1                     | 20.3                          | 0.35         | 1542          | 14.5                        | 94.7                   | 0.4                    | 151.2      |
| 19           | 0.62                   | 21.6                      | 40.7                     | 24.7                          | 0.35         | 1291          | 14.8                        | 100.4                  | 0.34                   | 150.4      |
| 20           | 0.5                    | 20.8                      | 35.8                     | 22.4                          | 0.33         | 1613          | 18                          | 95.5                   | 0.4                    | 109.3      |
to obtain the characteristic values of the contribution and cumulative contribution rates of the components shown in Table 5. This processing reduces the dimensionality of the input data and the influence of redundant information on the analysis efficiency of the fault diagnosis model based on the LSSVM algorithm.

The principal component and cumulative contribution rates obtained from the analysis of the PCA algorithm show that the cumulative contribution rate of the first two principal component components reached 94.476%, which can basically express the feature matrix of the original data. Therefore, the first two principal components are selected here to form a new feature for the fault diagnosis and analysis of the compressor, reducing the input data dimension.

The principal components (1 and 2) extracted by the PCA algorithm are projected onto a two-dimensional plane, and the two-dimensional distribution of the first two principal components is obtained as shown in Figure 3.

In Figure 3, “*” indicates the normal state of the compressor, “□” indicates the compressor cooling water system failure, “o” indicates the compressor lubrication

### Table 5. Principal component and cumulative contribution rates.

| Main ingredient | Eigenvalues | Contribution rate % | Cumulative contribution rate % |
|-----------------|------------|---------------------|-------------------------------|
| 1               | 0.0412     | 83.261              | 83.261                        |
| 2               | 0.0045     | 6.215               | 89.476                        |
| 3               | 0.0031     | 4.321               | 93.797                        |
| 4               | 0.0004     | 0.945               | 94.742                        |
| ...             | ...        | ...                 | ...                           |

![Figure 3. Two-dimensional distribution of the first two principal elements.](image)
device failure, “△” indicates the compressor bearing failure, and “+” number indicates the fault state of the compressor power unit.

The two-dimensional distribution of the first two principal components indicates that the clustering of the principal components is pretty well, and the distribution interval of each feature is obvious, facilitating the diagnosis and analysis of the compressor system faults through the pattern recognition method.

The basic parameters of the LSSVM and PSO algorithms are set as follows: parameter $c$ in the LSSVM model is set to 50, and parameter $\sigma$ is set to 10. The population size in the PSO algorithm is set to 30; the learning factor is 1.5; the initial and final weight values are 0.9 and 0.3, respectively; and the maximum number of optimizations is 200.

**Analysis of fault diagnosis results**

Compressor fault diagnosis systems are established using conventional BPNN and LSSVM algorithms and the LSSVM algorithm optimized using PSO without feature extraction by the PCA algorithm.

The same training and test data are used to carry out model training and generalization ability test on the four established fault diagnosis models. The training and test results of the four fault diagnosis models are shown in Figures 4 to 7.

Table 6 shows the comparison of model training time, the test sample classification time, and the fault recognition accuracy rate using four fault diagnosis models for compressor system fault diagnosis.

The comparison results of the four fault diagnosis models show that the fault diagnosis model based on the BPNN has low convergence speed when few training samples are used, requires long training time, and has low fault recognition accuracy.
The LSSVM-based fault diagnosis model can use few training samples to obtain a diagnostic model with a certain generalization ability, but the random performance of parameter selection during the training process leads to a large difference in the performance of the fault diagnosis model and low overall recognition efficiency and accuracy.

The PSO-LSSVM fault diagnosis module without feature extraction through the PCA algorithm can optimize the selection of parameters in the LSSVM-based fault diagnosis model through the PSO algorithm, but redundant data and high input dimensions increase the fault diagnosis training time of the model and the accuracy of the results.

**Figure 5.** Training and test results of fault diagnosis model based on LSSVM: (a) training results and (b) test results.

**Figure 6.** Training and testing results of fault diagnosis model based on PSO-LSSVM without feature extraction by PCA algorithm: (a) training results and (b) test results.
In the PSO-LSSVM fault diagnosis model based on the features extracted through the PCA algorithm, PCA is first performed on the data collected for compressor fault diagnosis and analysis, which reduces the input data dimension and the redundant information to the fault based on the LSSVM algorithm diagnostic models analyze the impact of efficiency. Afterward, the blindness of selecting the parameters of the LSSVM fault diagnosis model is reduced through the PSO algorithm.

The comparison results indicate that the PSO-LSSVM fault diagnosis model based on the feature extraction through the PCA algorithm has a diagnosis accuracy rate of up to 10.4% higher than the other three models, and the test sample classification time is reduced by up to 0.025 s. Therefore, the proposed PSO-LSSVM fault diagnosis model diagnosis system based on PCA algorithm for extracting features has high recognition efficiency and accuracy, which can effectively identify compressor faults and improve compressor efficiency.

Figure 7. Training and testing results of fault diagnosis model based on PSO-LSSVM with feature extraction by PCA algorithm: (a) training results and (b) test results.

Table 6. Comparison results of four fault diagnosis models.

|                     | BPNN model | LSSVM model | PSO-LSSVM model | PCA-PSO-LSSVM model |
|---------------------|------------|-------------|-----------------|---------------------|
| Model training time/s | 2.602      | 1.935       | 8.124           | 3.651               |
| Test sample classification time/s | 0.0412     | 0.0211      | 0.0512          | 0.0162              |
| Fault recognition accuracy/% | 88.7       | 93.2        | 96.5            | 99.1                |

In the PSO-LSSVM fault diagnosis model based on the features extracted through the PCA algorithm, PCA is first performed on the data collected for compressor fault diagnosis and analysis, which reduces the input data dimension and the redundant information to the fault based on the LSSVM algorithm diagnostic models analyze the impact of efficiency. Afterward, the blindness of selecting the parameters of the LSSVM fault diagnosis model is reduced through the PSO algorithm.

The comparison results indicate that the PSO-LSSVM fault diagnosis model based on the feature extraction through the PCA algorithm has a diagnosis accuracy rate of up to 10.4% higher than the other three models, and the test sample classification time is reduced by up to 0.025 s. Therefore, the proposed PSO-LSSVM fault diagnosis model diagnosis system based on PCA algorithm for extracting features has high recognition efficiency and accuracy, which can effectively identify compressor faults and improve compressor efficiency.
Conclusion

This work studies the compressor system fault diagnosis system. First, the compressor fault diagnosis system is analyzed, a fault diagnosis algorithm model based on the type of fault is established, and an example analysis is conducted to verify the reliability of the proposed fault diagnosis method. The conclusions are as follows:

1. According to the working principle and characteristics of the compressor system, the connection between the fault phenomenon and the root cause is analyzed, and the characteristic signal is extracted for fault diagnosis.
2. A fault diagnosis model based on the LSSVM optimized using PSO is established, and the PCA algorithm is used to extract data features as input to the fault diagnosis model and reduce the input dimension.
3. The comparative analysis of various fault diagnosis models indicates that the proposed PSO-LSSVM fault diagnosis model based on the feature extraction through the PCA algorithm has the highest fault recognition accuracy rate that is 10.4% higher than the that of the other three fault diagnosis models, and the test sample classification time is greatly reduced. Only 0.0162 s is consumed to verify the identification efficiency and accuracy of the proposed fault diagnosis model, providing a new method for the fault diagnosis of mine air compressors.

Declaration of conflicting 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 receipt of the following financial support for the research, authorship, and/or publication of this article: "This work was supported by the China Postdoctoral Science Foundation (No.2020M682268), the Major Science and Technology Innovation Projects in Shandong Province (2019SDZY04), National Natural Science Foundation of China (51974159), and Tiandi Technology Innovation and Entrepreneurship Funding Project (2020-TD-ZD005).

ORCID iDs

Kun Zhang https://orcid.org/0000-0001-7339-5056
Jinpeng Su https://orcid.org/0000-0003-0122-8771

References

1. Zhang K. Design and realization of energy-saving intelligent monitoring system of compressor. MA Thesis. Shandong University, Shandong, 2017.
2. Li S. Study on protection and intelligent diagnosis system for air compressor. MA Thesis, Tongji University, Shanghai, 2008.
3. Xu Q and Hu J. Fault feature extraction method for compressor based on improved incremental Laplacian Eigen map algorithm. *Chin J Sci Instrum* 2013; 34(4): 791–796.
4. Liu HG and Dong Z. Comprehensive fault diagnose technology research of aircraft generator. *Electron Meas Technol* 2014; 37(3): 125–127.
5. Zhou Z, Wang J, Chen H, et al. An online compressor liquid floodback fault diagnosis method for variable refrigerant flow air conditioning system. *Int J Refrig* 2020; 111: 9–19.
6. Van Tung T, AlThobiani F, Tinga T, et al. Single and combined fault diagnosis of reciprocating compressor valves using a hybrid deep belief network. *Proc Inst Mech Eng C J Mech Eng Sci* 2018; 232(20): 3767–3780.
7. Zhang Y, Ji J and Ma B. Fault diagnosis of reciprocating compressor using a novel ensemble empirical mode decomposition-convolutional deep belief network. *Measurement* 2020; 156: 107619.
8. De Cooman T, Vandecasteele K, Varon C, et al. Personalizing heart rate-based seizure detection using supervised SVM transfer learning. *Front Neurol* 2020; 11: 145.
9. Li J, Yao X, Wang X, et al. Multiscale local features learning based on BP neural network for rolling bearing intelligent fault diagnosis. *Measurement* 2020; 153: 107419.
10. Zhang Y, Jiang Y, Chen Y, et al. Fault diagnosis of high voltage circuit breaker based on multi-classification relevance vector machine. *J Electr Eng Technol* 2020; 15(9): 413–420.
11. Fan Y, Cui X, Han H, et al. Feasibility and improvement of fault detection and diagnosis based on factory-installed sensors for chillers. *Appl Therm Eng* 2020; 164: 114506.
12. Wang H, Peng MJ, Hines JW, et al. A hybrid fault diagnosis methodology with support vector machine and improved particle swarm optimization for nuclear power plants. *ISA Trans* 2019; 95: 358–371.
13. Yushu C. Nonlinear dynamical principle of mechanical fault diagnosis. *Chin J Mech Eng* 2007; 43(1): 25–34.
14. You J. *Fault diagnosis of single-screw air compressor*. MA Thesis, Liaoning Technical University, Liaoning, 2006, pp. 36–50.
15. Li J. *Research and development of remote monitoring early-warning and performance test system for the reciprocating air compressor*. MA Thesis, Beijing University of Technology, Chaoyang, 2009, pp. 15–46.
16. Ahmadi H, Baghban A and Baghban M. Implementing PSO-LSSVM model to approximate vaporization enthalpy of pure hydrocarbons and petroleum fractions. *Pet Sci Technol* 2019; 37(16): 1891–1897.
17. Ipakchi H, Masoud Rezadoust A, Esfandeh M, et al. Modeling and optimization of electrospinning conditions of PVB nanofiber by RSM and PSO-LSSVM models for improved interlaminar fracture toughness of laminated composites. *J Compos Mater* 2020; 54(3): 363–378.
18. Jun C. *The application of improved PSO-LSSVM in hourly urban water consumption forecasting*. MA Thesis, Shanghai Jiaotong University, Shanghai, 2013.

**Author biographies**

Kun Zhang received the Ph.D. degree from Liaoning Technical University, China, in 2019. He is currently with the Shandong University of Science and Technology, Qingdao, China. His current research interests include vibration characteristics of large-scale equipment intelligent manufacturing and coal mine mechanical mechanics.
Jinpeng Su received the Ph.D. degree in mechanical science from Shanghai Jiao Tong University, China, in 2018. He is currently a Postdoctoral Researcher with the Shandong University of Science and Technology, Qingdao, China. His current research interests include dynamic modeling, vibro-acoustic analysis, and control of mechanical systems.

Shaoan Sun received the M.S. degree from Liaoning Technical University, China, in 2020. His main research direction is mechatronics engineering and fluid power transmission and control engineering.

Zhixiang Liu received the Ph.D. degree from Liaoning Technical University, China, in 2017. He is currently with the Liaoning Technical University, Fuxin, China. His current research interests include vibration characteristics of large-scale equipment intelligent manufacturing and coal mine mechanical mechanics.

Jinrui Wang received the Ph.D. degree in Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2019. He is currently working as an associate Professor at Shandong University of Science and Technology, Qingdao, China. His current research interests focus on intelligent fault diagnosis of machines.

Mingchao Du received the bachelor's degree from Shandong University of Science and Technology, Qingdao, China, in 2015. He is currently a doctoral candidate at China University of Petroleum (East China) Qingdao, China. His current research interests include coal mine mechanical mechanics and special machining technology.

Zengkai Liu received the M.S. degree from the Shandong University of Science and Technology, China, in 2019. He is currently pursuing the Ph.D. degree. His current research interest includes analysis of mechanical characteristics of coal mine equipment.

Qiang Zhang received the Ph.D. degree from Liaoning Technical University, China, in 2010. He completed the Postdoctoral Research from China Coal Energy Company Ltd. (China Coal Energy), Beijing, China, in 2012. His research interests include intelligent manufacturing and precision mining of coal mining machinery. He is a member of the International Intelligent Manufacturing Society.