

Research Article

Fault Diagnosis of Belt Conveyor Based on Support Vector Machine and Grey Wolf Optimization

Xiangong Li^{1,2}, Yu Li,² Yuzhi Zhang,² Feng Liu,² and Yu Fang²

¹Key Laboratory of Deep Coal Resource Mining, China University of Mining & Technology, Xuzhou, Jiangsu 221116, China

²School of Mines, China University of Mining & Technology, Xuzhou, Jiangsu 221116, China

Correspondence should be addressed to Xiangong Li; cumtlxg@163.com

Received 22 July 2020; Accepted 11 September 2020; Published 24 September 2020

Academic Editor: Yi-Zhang Jiang

Copyright © 2020 Xiangong Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Belt conveyor is widely used for material transportation over both short and long distances nowadays while the failure of a single component may cause fateful consequences. Accordingly, the use of machine learning in timely fault diagnosis is an efficient way to ensure the safe operation of belt conveyors. The support vector machine is a powerful supervised machine learning algorithm for classification in fault diagnosis. Before the classification, the principal component analysis is used for data reduction according to the varieties of features. To optimize the parameters of the support vector machine, this paper presents a grey wolf optimizer approach. The diagnostic model is applied to an underground mine belt conveyor transportation system fault diagnosis on the basis of monitoring data collected by sensors of mine internet of things. The results show that the recognition accuracy of the fault is up to 97.22% according to the mine site dataset. It is proved that the combined classification model has a better performance in fault intelligent diagnosis.

1. Introduction

As the widely used mechanical transportation equipment, the belt conveyor plays a critical role in transporting raw coal in the coal mining industry and other industries. In recent years, with the widespread automation in mines, accidents occurring involving mining equipment have become more and more complex. In some developing countries such as China, there have been frequent incidents of miners' casualties and property loss caused by the failure of mine equipment like belt conveyor. Generally, the accidents are mainly caused by improper use or maintenance of the belt conveyor or not timely detection of the system fault, which compounds difficulties to diagnose the equipment failure in real time when the fault occurs. Meanwhile, the development of the internet of things has brought out plenty of big data which indicate system error information in advance. Besides, timely fault detection could be a great help for belt conveyor predictive maintenance [1]. Therefore, it is especially necessary to improve the detection technology to make

a timely diagnosis of coal mine belt conveyor faults with the intelligent method.

At present, researches on fault diagnosis are basically divided into three modules, namely, (i) fault mechanisms; (ii) monitoring signal acquisition and processing; and (iii) intelligent fault diagnosis. In the fields of fault mechanisms, Miriam and Anna [2] tested different types of conveyor belts under laboratory conditions to verify the possibility and detailed condition of four degrees, and a naive Bayes classification model was adopted to predict the severity of the conveyor belt damage. Relatively, Gabriel et al. [3] conducted a series of experiments to compare the results under different types of operational conditions, thus increasing the operational reliability of conveyor belts and understanding the destructive processes in the conveyor belt better. Peter et al. [4] developed the equations that describe the dependencies of tool durability and threaded length on both observed parameters based on the experimentally obtained data to study a tap failure during the internal threads making. Papageorgiou et al. [5] studied the samples selected

and prepared for destructive testing for dimensional verification, hardness measurements, and chemical analysis to ensure the material's efficiency and to improve its operational reliability. Song [6] analysed the fault mechanism of the belt conveyor and the former focused on the deviation mechanism of the belt conveyor. With the development of Internet of Things, there are more and more fault diagnosis studies in the field of monitoring signal acquisition and processing. Furthermore, Jacek et al. [7] proposed a new approach to local damage detection in rotating machines based on an impulsive acceleration signal to extract the information on the damage of belt conveyor gearboxes, which extracted the impulses hidden in the signal in a clear visible and distinguishable way. Walter et al. [8] proposed two models of gearboxes operating under varying load conditions based on the belt conveyor and the bucket wheel excavator to find a strong correlation between load values changes in condition and the diagnostic features. Andrejiova et al. [9] applied the design of the experiment method to determine the effects of factors and the interactions between them while mathematical models were used to identify the relationship between the local peak impact forces and various values of the impact process input parameters. Tomasz et al. [10] studied the data acquired from belt core testing through a system to solve the problems related to diagnosing splices in steel-cord conveyor belts. Zhang [11] proposed a fault diagnosis method for reciprocating compressor based on multisource information fusion. Wang [12] offered a frequency domain analysis of signal processing methods for mechanical fault diagnosis. In the field of intelligent fault diagnosis, different mathematical theories emerge in an endless stream, such as the neural network method, extreme learning machine, grey clustering method, deep learning method, and support vector machine (SVM) method. Even different theories can absorb the advantages of other ways to improve their algorithms and innovate. Ravikumar et al. [13] applied the decision tree method to find the significant features for classification of self-aligning troughing rollers' faults, and the k-star algorithm was applied for fault diagnosis to raise the fault classification accuracy. Arup et al. [14] presented an automated tool called the Manufacturing Process Failure Diagnosis Tool (MPFDT), which can detect and isolate the faults and anomalies in the Programmable Logic Controller (PLC) controlled manufacturing systems effectively. Maren et al. [15] proposed a hybrid machine learning approach that blended natural language processing techniques and ensemble learning for predicting extremely rare aircraft component failure. David and Dariusz [16] developed mathematical modelling of diagnostic data with the use of selected stochastic processes—types of Wiener process and Ornstein–Uhlenbeck process, and the proposed robust functional analysis reduced bias and provided more accurate fault detection rates. Jacopo et al. [17] provided a bearing fault vibration model that took into account the mechanical design of the cart, its motion profile, the shape of the conveyor path, and so on; thus the wear condition monitoring of the Independent Cart Conveyor System's elements can be achieved. Jaroslav et al. [18] used dynamic

models to establish and predict the cause and location of the conveyor drive wear so the trouble-free operation time of the belt conveyor could be increased efficiently. John et al. [19] applied the multistream convolutional neural network (MS-CNN) for automatic feature extraction at various line frequencies to diagnose motor defects. Yang and Zhu [20] proposed a fault diagnosis method based on a probabilistic neural network, which successfully diagnosed the fault of a high voltage circuit breaker. Chen et al. [21] proposed an enhanced artificial bee colony-based support vector machine for image-based fault detection of belt conveyor which obtained a detection accuracy of 95%. Recently, Liu et al. [22] used a gradient boost decision tree as a machine learning method to detect belt conveyor idlers fault with acoustic signal and obtained a quite notable result.

Although there have been research works on the fault diagnosis of mechanical equipment, diagnostic techniques for the overall failure of the belt conveyor has received little attention. Belt conveyor equipment is a kind of essential equipment in coal mine transportation, where safety and stability directly affect continuous production. In coal mines, it is significant to carry out effective fault diagnosis of the belt conveyor [23]. This study aims at providing an effective means for the fault diagnosis of belt conveyors in coal mining enterprises and strengthening the intelligent technology level of mining enterprises.

The rest of the content is organized as follows.

Section 2 introduced the model and method in this paper. Section 3 is the main part which conducts the empirical analysis. Section 4 states conclusions.

2. Model and Method

2.1. Principal Component Analysis. In this section, the principal component analysis (PCA) is introduced for data dimension reduction.

Suppose X is a set of M -dimensional data samples, and its K th principal component is expressed as

$$P_K = (P_{1k}, P_{2k}, \dots, P_{Mk})^T, \quad k = 1, 2, \dots, N, \quad (1)$$

where M represents the data dimension, N is the sample size, and P represents the principal component vector matrix. Then, the projection of the data sample set at P is represented as the principal component matrix of the sample:

$$T = XP. \quad (2)$$

A single element of the T matrix represents

$$t_{ij} = \sum_{k=1}^M p_{kj} x_{ik}, \quad (i = 1, 2, \dots, N; j = 1, 2, \dots, M). \quad (3)$$

According to formula (3), the j th principal component is

$$T_j = \sum_{K=1}^M X_K P_{Kj}, \quad (j = 1, 2, \dots, M). \quad (4)$$

To ensure the orthogonality of the principal component vectors,

$$T_j^T T_{j'} = \sum_{K=1}^M \sum_{K'=1}^M X_K X_{K'} P_{Kj} P_{K'j'} = 0 \quad (j \neq j'), \quad (5)$$

expressed in a matrix form as

$$T^T T = (XP)^T XP = P^T X^T XP = P^T DP = \Lambda. \quad (6)$$

In equation (6), the covariance matrix is represented by $D = X^T X$, and the diagonal matrix is represented by Λ , wherein

$$\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_M]. \quad (7)$$

In order to ensure that all of the above formulas are true, the sample data should meet the following constraints, and the P matrix column vectors are also guaranteed to be orthogonal to each other, namely,

$$\sum_{i=1}^M p_{ij} p_{ik} = \delta_{jk} = \begin{cases} 0, & (j \neq k), \\ 1, & (j = k). \end{cases} \quad (8)$$

Then, it obtains

$$\begin{aligned} PP^T DP &= P\Lambda, \\ DP &= P\Lambda. \end{aligned} \quad (9)$$

The above equation is the characteristic equation in the line generation. D is the characteristic matrix of the data sample. If the rank of D is M , the matrix P can be called the M -dimensional eigenvector matrix. The Jot eigenvalue in the matrix should guarantee

$$DP_j = \lambda_j P_j, \quad (10)$$

which can be rewritten as

$$|D - \lambda_j I| = 0. \quad (11)$$

The feature value is obtained by using equation (11), and the corresponding feature vector is solved.

In the actual sample data processing, the principal component contribution rate is introduced to represent the amount of information that the data sample reflects in the original sample data. The larger the proportion of the principal component contribution rate, the more representative the data sample itself and the contribution to any principal component vector. The rate is expressed as

$$g_j = \frac{\lambda_j}{\sum_{i=1}^M \lambda_i}, \quad (j = 1, 2, \dots, M). \quad (12)$$

For multidimensional sample data, principal component analysis (PCA) can be used for data dimensionality reduction. The PCA process is conducted by following the 4 steps listed below:

Step 1: input data and normalize the data to ensure that the data of the same attribute are on the same column vector.

Step 2: apply the principle of the principal component to find K principal components and obtain the feature vector and eigenvalue of the data sample.

Step 3: let the principal component vector set replace the original vector of the data sample set and arrange the principal components in order according to the size of the data sample information.

Step 4: according to the sorting result, when the accumulated contribution rate comes to 0.85, the data samples can reflect the main part of the original sample. The data samples contained what can better reflect the information amount of the data in the original sample and remove the principal component with a lower contribution rate to achieve the ultimate goal of data sample dimension reduction [24].

2.2. Support Vector Machine Classification. The support vector machine (SVM) is an advanced classification method introduced by Boser et al. [25] and then was widely used in different classification and regression. The core idea of support vector machine (SVM) classification is the optimal classification hyperplane. When using SVM to classify and identify a given sample, we find a hyperplane that meets the classification requirements, try to make the coordinate points in the sample away from the plane, and ensure that the areas on both sides of the classification hyperplane are large enough [26]. Suppose the linear separable sample set is $\{(x_i, y_i), i = 1, \dots, l\}$, where $x_i \in R^n$, $y_i \in \{+1, -1\}$, n is sample space dimension values, and the equation of the classification line P is represented as $\omega x + b = 0$ and satisfies the following formula:

$$y_i[(\omega \cdot x_i) + b] - 1 \geq 0, \quad i = 1, 2, \dots, n. \quad (13)$$

By calculating the interval between the outer lines as $2/\|\omega\|$, when the minimum value of $\|\omega\|$ is the largest, the classification hyperplane is the optimal hyperplane of the support vector machine. The training samples on the two outer lines are called support vectors.

Using SVM for machine learning, the whole learning process is equivalent to finding the optimal value of the parameters. We introduce Lagrangian multipliers to solve this problem:

$$\min \frac{1}{2} \|\omega\|^2 \text{ s.t. } y_i(\omega \cdot x_i + b) \geq 1, \quad i = 1, 2, \dots, n. \quad (14)$$

Introducing the Lagrangian multiplier, this is expressed as

$$L(\omega, b, a) = \frac{1}{2} (\omega \cdot \omega) - \sum_{i=1}^n a_i \{y_i[(\omega \cdot x_i) + b] - 1\}. \quad (15)$$

Deriving the parameters w and b , respectively, and assuming zero, there is

$$\begin{cases} \frac{\partial}{\partial w} L(\omega, b, a) = 0, \\ \frac{\partial}{\partial b} L(\omega, b, a) = 0. \end{cases} \quad (16)$$

Then, it obtains

$$\begin{cases} w = \sum_{i=1}^n a_i y_i x_i, \\ \sum_{i=1}^n a_i y_i = 0. \end{cases} \quad (17)$$

Maximizing the spacing between classification planes is essentially an improvement and optimization of the SVM's generalization ability. By maximizing the classification plane spacing, the structural risk of the SVM is minimized, and the generalization ability of the algorithm is strengthened to meet the core idea of the SVM. In order to solve the problem in the above formula, the dual theory is introduced, and the classification problem is transformed into a dual problem:

$$\begin{cases} \min & \frac{1}{2} \sum_{i,j=1}^n a_i a_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^n a_i, \\ \text{s.t.} & \sum_{i=1}^n a_i y_i = 0, \end{cases} \quad (18)$$

where $a_i \geq 0, i = 1, 2, \dots, n$, a_i corresponds to the Lagrangian multiplier of the i^{th} sample, and equation (18) has the only feasible solution; then the b parameter value is obtained:

$$b = \frac{1}{N_{\text{NSV}}} \sum_{x_i \in \text{N}} \left(y_i - \sum_{x_j \in \text{J}} a_j y_j (x_i \cdot x_j) \right). \quad (19)$$

Then, the optimal classification function is finally found:

$$f(x) = \text{sgn}((w \cdot x) + b) = \text{sgn} \left\{ \sum_{i=1}^n a_i y_i (x_i \cdot x) + b \right\}. \quad (20)$$

The test sample X can be categorized by the optimal classification function.

Most of the samples classified are linear inseparable samples. Since the objective function of the dual function and the optimal classification function of the solution have only $(x_i \cdot x_j)$ and the nonlinear transformation maps to high-dimensional, the calculation of the optimal classification function and the objective function algebra in high-dimensional space only involves $\Phi(x_i) \cdot \Phi(x_j)$ [27]. So, we define the kernel function $k(x_i \cdot x_j) = \Phi(x_i) \cdot \Phi(x_j)$. At this situation, the objective function value of the dual problem after optimization is

$$\min_a \frac{1}{2} \sum_{i,j=1}^n a_i a_j y_i y_j K(x_i \cdot x_j) - \sum_{i=1}^n a_i. \quad (21)$$

The optimal classification function is transformed into

$$f(x) = \text{sgn}((w \cdot x) + b) = \text{sgn} \left\{ \sum_{i=1}^n a_i y_i K(x_i \cdot x) + b \right\}. \quad (22)$$

2.3. SVM Parameter Optimization. Suppose the population size is N , the search space for hunting prey is D , and the

position parameter of the i^{th} wolf set in the search space is $x_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{id})$. The optimal value of the population, the second fitness value, and the third fitness value are, respectively, α, β, δ , the rest of the values are ω , and the position of the prey is the optimal solution of the whole problem. The steps of the grey wolf algorithm for solving the optimization problem are as follows. Firstly, the initial population is randomly generated, and the position of the prey is judged according to the position of α, β, δ . Other individuals in the population also calculate the distance difference from the prey based on this position, on the basis of which they track the prey until they kill the prey [28]. The specific formula for this behavior is expressed as follows.

The distance D between the grey wolf individual and the prey is expressed as

$$D = |C \cdot X_p(t) - X(t)|. \quad (23)$$

In the above formula, t represents the number of iterations of the algorithm, $X_p(t)$ represents the position of the prey after t iterations, $X(t)$ represents the position of the individual grey wolf after t iterations, C is the sway factor, where $C = 2 \cdot r_1$, and r_1 is a random number generated on a closed interval of 0 to 1.

Grey wolf individual location updates are as follows:

$$X(t+1) = X_p(t) - A \cdot D, \quad (24)$$

$$A = 2 \cdot a \cdot r_2 - a. \quad (25)$$

In equations (24) and (25), A is represented as a convergence factor and r_2 is taken from a randomly generated number on a closed interval of 0 to 1 and is a value gradually decreasing from 2 to 0 as the iteration progresses.

When the grey wolf individual calculates the distance and estimates the specific position of the prey, α, β , and δ take the lead to attack the prey and locate the direction of the prey according to the position of the three, expressed as follows:

$$D_\alpha = |C_1 \cdot X_\alpha(t) - X(t)|, \quad (26)$$

$$D_\beta = |C_2 \cdot X_\beta(t) - X(t)|, \quad (27)$$

$$D_\delta = |C_3 \cdot X_\delta(t) - X(t)|, \quad (28)$$

$$X_1(t+1) = X_\alpha(t) - A_1 \cdot D_\alpha, \quad (29)$$

$$X_2(t+1) = X_\beta(t) - A_2 \cdot D_\beta, \quad (30)$$

$$X_3(t+1) = X_\delta(t) - A_3 \cdot D_\delta, \quad (31)$$

$$X_p(t+1) = \frac{X_1 + X_2 + X_3}{3}. \quad (32)$$

The distance between the individual grey wolf and the prey is calculated by (24), the position of the grey wolf after iterative updating is determined by (25), and the prey range is finally located by (26) to (32).

However, with the increase in the number of iterations, the grey wolf population will have a population difference in some areas of the search space. This phenomenon has a great negative impact on the optimization performance of the algorithm. Aiming at the limitations of the standard form of grey wolf algorithm, this paper introduces a differential evolution of the grey wolf algorithm [29]. The specific operations are as follows:

- (1) The mutation operation of the differential evolution algorithm can significantly enhance the global search ability of the algorithm. The main point is that two different individuals in the arbitrarily selected population can differentially scale their position vectors, thereby obtaining a series of differential information and assigning it to another among individuals who are not mutated. The mutation operator needs a mutation operator. This algorithm refers to the adaptive mutation operator proposed by FAN. The introduced adaptive operator is

$$F = 1 - \sqrt{\frac{f(X_{\text{best}})}{f(X_i)}}. \quad (33)$$

The above formula $f(X_{\text{best}})$ is the current optimal fitness value, expressed as the fitness of the head wolf $f(X_\alpha)$, and the fitness value of any individual wolf X_i in the population is expressed as $f(X_i)$.

The complete mutation operation is

$$X_i(t+1) = X_i(t) + F_1 * (X_\alpha - X_i(t)) + F_2 * (Y_i(t) - X_i(t)). \quad (34)$$

In the variation operation formula (34), the adaptive operator is represented by F_1 , and the difference between the selected individual fitness value and the current population optimal fitness value is adjusted mainly by changing the size so that the algorithm jumps out of the local optimum to achieve optimal global results. F_2 is a specific real value, generally 0.1, in order to protect the individual species diversity of the population and reduce the impact of individuals with low fitness values on the overall situation. $Y_i(t)$ is expressed as the individual fitness value of the wolves in the population.

If the individual resulting from the cross-variation of equation (34) exceeds the set search range, the constraint is specified by the following constraint method:

$$\begin{cases} X_i = X_i - (X_i - L) * F_1, \\ X_{i+1} = X_{i+1} + (U - X_{i+1}) * F_1. \end{cases} \quad (35)$$

In equation (35), the upper boundary of the search space is represented as U , and the lower boundary is represented as L .

- (2) The cross-operation of the differential evolution algorithm mainly improves the improvement effect of mutation data by exchanging related data elements of mutated individuals and unmutated individuals. The algorithm applies the basic cross-strategy of the differential evolution algorithm. The populations initialized by the optimal point set are equally divided into two groups according to the fitness value and are processed by the mutation operation, followed by the intersection of the elements. The specific formula is as follows:

$$u_{i,j}(t+1) = \begin{cases} x_{i,j}(t+1)\text{rand} \leq Cr, & j = j_{\text{rand}}, \\ x_{i,j}(t), & \text{others.} \end{cases} \quad (36)$$

Among them, CR represents the probability value of cross mutation and is the random generation number, and j_{rand} represents the dimension of random variation; that is, a number is randomly generated, and then, this value is multiplied by the dimension. A simple cross is shown schematically in Figure 1.

- (3) Selection operation of differential evolution using greedy thought to select the next generation of grey wolf population; the operation mode is expressed by the following formula:

$$X_i(t+1) = \begin{cases} u_i(t+1), & f(u_i(t+1)) \leq f(x_i(t)), \\ x_i(t), & \text{others.} \end{cases} \quad (37)$$

The penalty factor and kernel function parameters of the SVM are randomly selected by the algorithm. Under this condition, the classification accuracy of the algorithm is restricted [30]. In this paper, the differential evolution of the grey wolf algorithm for SVM hyperparameter optimization not only ensures that the algorithm can obtain the global optimal solution but also improves the accuracy of the model.

3. Empirical Analysis

3.1. Data Preparation and Dimensionality Reduction. The dataset is obtained from the monitoring information system of a coal mine which is transferring into an intelligent mine on the basis of IOTs (Internet of Things) and AI (artificial intelligence) in eastern China; the typical fault conditions and the normal running state of the belt conveyor are selected as the research objects. The six typical faults are belt slip and belt tear, belt deviation, motor failure, main belt overload, and belt fire accident [31]. According to the field visit and the processing of sample data class imbalance, this paper selects a total of 126 sets of sample data with state information tags for simulation experiments and 90 sets of training sample data for training of fault diagnosis model,

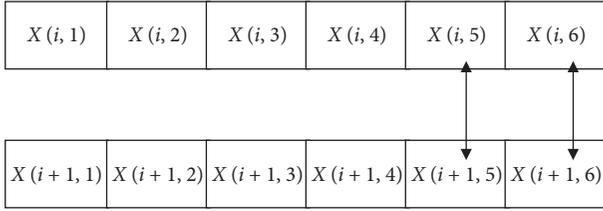


FIGURE 1: Population crossing rule diagram.

and the remaining 36 sets of data as test set samples are used to verify the accuracy of the fault diagnosis model. Each sample data contain 19 fault characteristics, followed by motor power, motor temperature, belt speed, motor bearing temperature, coal bunker position, CST (controlled start transmission) power, CST temperature, CST pressure, belt tension, drive to drum temperature, belt temperature, drive drum speed, roadway temperature, redirection drum speed, motor current, motor voltage, reversing drum temperature, belt offset, and smoke concentration, respectively, setting the fault label for each sample data and then establishing a classification model. The composition of the training sample and the fault sample is shown in Table 1. Label 1 to label 7 represent the normal state, the state of the belt slip, the state of the belt tear, the belt deviation, the state of the belt fire, the belt overload state, and the belt fire condition, respectively.

The 19 parameter indices such as motor power, motor temperature, belt speed, and bearing temperature are standardized, then notated by X1 to X19, and imported into SPSS software. Finally, the total variance between the output parameter indicators is explained, and the total variance is explained. The contribution rate of each principal component to the population can be arranged, the principal component whose total contribution rate of the principal component reaches 0.85 or above is selected, and the above main component is used instead of the original sample data for classification processing, thereby achieving the purpose of dimension reduction of the characteristic index.

The total variance interpretation chart output in the SPSS software is shown in Table 2. When the fifth principal component is extracted, the cumulative contribution rate of the principal component has exceeded 0.85. Therefore, this paper selects five principal components as feature vectors and derives the feature vector normalization to carry out the next fault identification.

3.2. Simulation Verification. The number of feature samples is imported into the fault diagnosis model, and 126 sets of samples after the dimension reduction using the PCA algorithm are selected. The training set and the test set are divided according to the ratio of the test set and the training set of about 3:1. Among them, 90 sets of training sets are used for the establishment of the multiclass fault diagnosis model and the optimization operation of SVM kernel parameters. The remaining 36 sets of data are used as test set samples [32].

Set the number of wolves to twenty and the maximum number of iterations is sixty. Choose RBF kernel function,

TABLE 1: Training set and test set partition table.

Type label	Fault type	Train set	Test set
Label 1	Normal status	20	8
Label 2	Belt slip	17	7
Label 3	Belt tear	8	3
Label 4	Belt deviation	15	6
Label 5	Motor failure	15	6
Label 6	Belt overload	10	4
Label 7	Belt fire	5	2

TABLE 2: Interpretation of the total variance.

Component	Extraction sums of squared loadings		
	Total	Percentage of variance	Cumulation (%)
1	8.522	44.852	44.852
2	3.082	16.220	61.072
3	2.129	11.207	72.279
4	1.577	8.303	80.581
5	1.069	5.625	86.208

Extraction method: principal component analysis

the superparameter values range is $[2^{-8}, 2^8]$, and the crossover probability CR is 0.2. When the model is verified by simulation, the improved grey wolf algorithm is used to iteratively search for key parameters. The penalty factor of the fault classification model is 25.3684; the width of the kernel function is 8.8765. Finally, the classification effect diagrams of the training set and the test set of the belt conveyor failure are shown in Figures 2 and 3, respectively. Therefore, the prediction accuracy of the training set is 97.78%, and it is 97.22% for the test set. It can be seen that the faults of the belt conveyor are effectively classified based on the model built in this paper.

To verify the performance of the hybrid fault diagnosis model proposed in this paper, the fault classification efficiency is compared with GWO-SVM, PCA-SVM, and PCA-GWO-SVM. The specific setting of the control experiment is as follows: the data used in the GWO-SVM model are 19-dimensional feature sample data. Model parameters optimization method with a standard grey wolf optimizer (GWO) and the setting of the parameters in accordance with the setting of the hybrid grey wolf optimizer for the same settings: the PCA-GWO-SVM model was used to reduce the dimension of the data, and then it was imported into the standard grey wolf optimizer optimized SVM classifier for classification and recognition. The PCA-SVM model aims to reduce the dimension of sample data, after which the data are imported to the SVM model for classification, while the parameter of the model is set by experience. Table 3 shows the values of key parameters processed by different models.

As shown in Figures 4 and 5, the classification result can be seen by PCA-SVM classification of the training sets and test set.

As shown in Figures 6 and 7, the classification result can be seen by GWO-SVM classification of the training sets and test set.

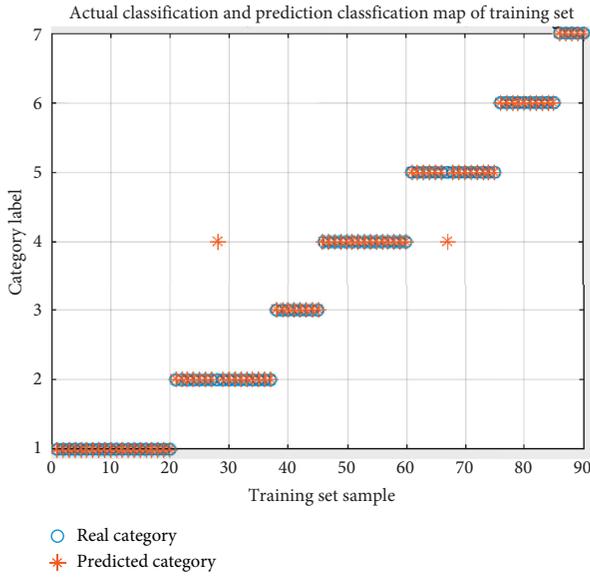


FIGURE 2: Training set classification identification.

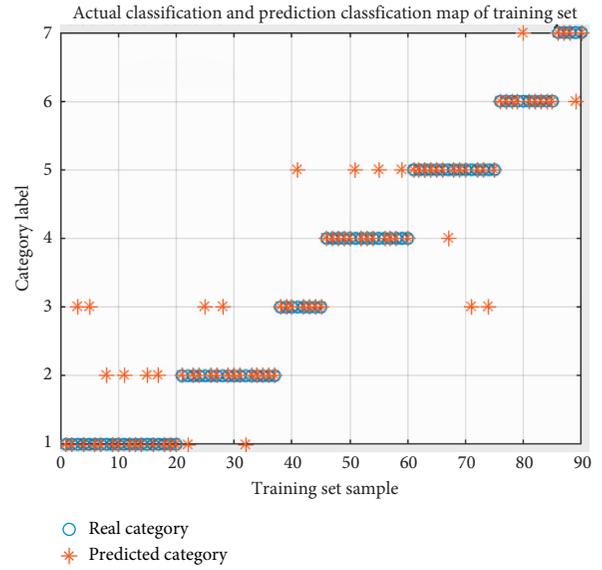


FIGURE 4: PCA-SVM classification of the training sets.

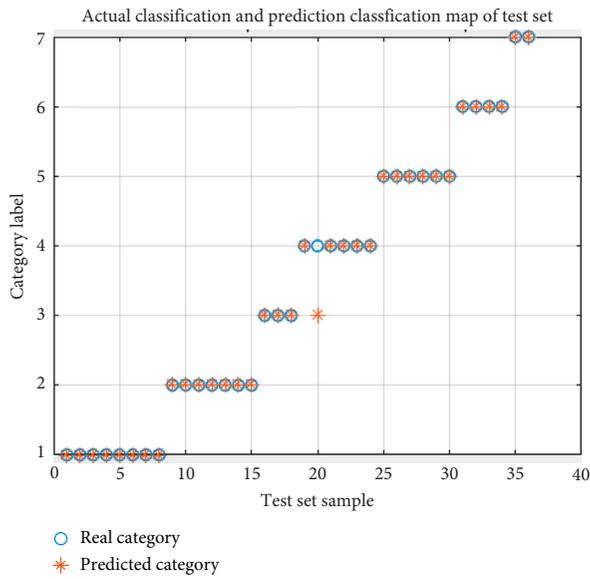


FIGURE 3: Test set classification identification.

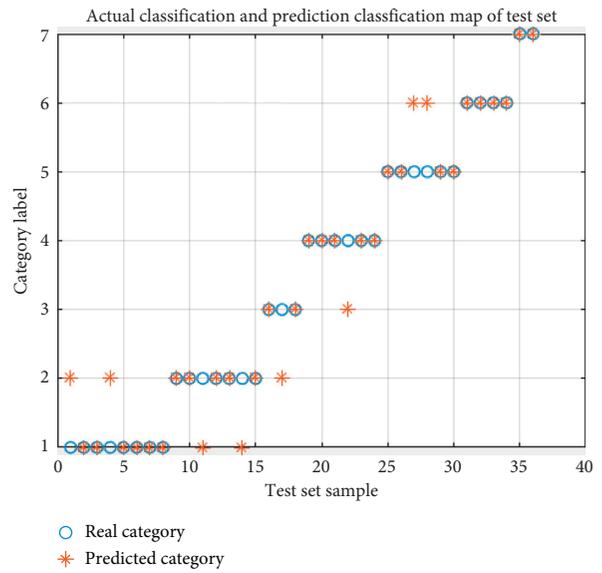


FIGURE 5: Test set classification diagram.

TABLE 3: Kernel parameter table.

Model building form	Penalty factor	Kernel function width
PCA-SVM	2	0.05
GWO-SVM	22.6545	12.4765
PCA-GWO-SVM	36.9754	4.9578
PCA-HGWO-SVM	25.3684	8.8765

As shown in Figure 8, the classification result can be seen by PCA-GWO-SVM classification of the training sets and test set.

Figure 9 shows the classification of PCA-GWO-SVM models for belt conveyor fault diagnosis, which proves it is much better than other models. Table 4 shows the recognition and diagnostic of different models.

According to Table 4, the verification and comparison of the model are conducted. Firstly, by comparing the PCA-SVM model with the PCA-GWO-SVM model, it can be seen from the data that the running time of the PCA-GWO-SVM model is slightly longer than that of the PCA-SVM model, but the classification accuracy of the model is greatly improved, which indicates that the optimization algorithm can improve the accuracy of the diagnostic model by optimizing the key parameters. Then, the GWO-SVM and PCA-GWO-SVM models are compared. According to the data, compared with the former one, the running time of the latter model was reduced from 11.9546 s to 5.1547 s, indicating that the complexity of the model was reduced after dimension reduction by principal component analysis.

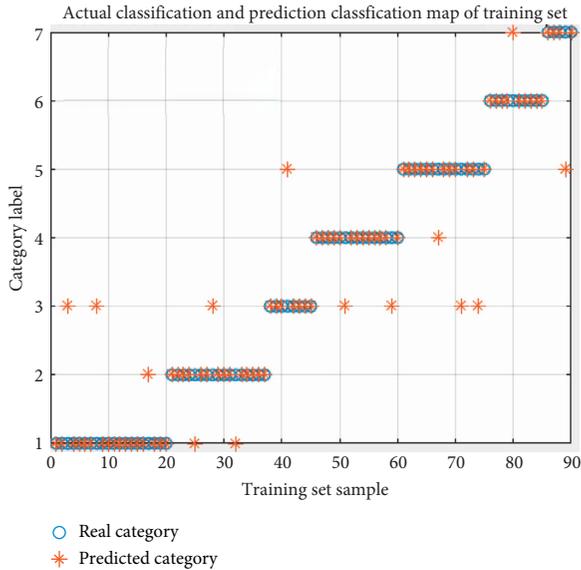


FIGURE 6: GWO-SVM training set classification recognition map.

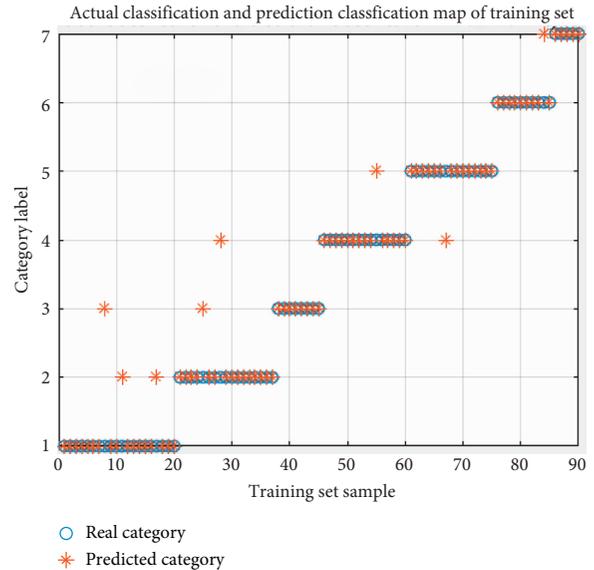


FIGURE 8: Classification identification of the PCA-GWO-SVM training set.

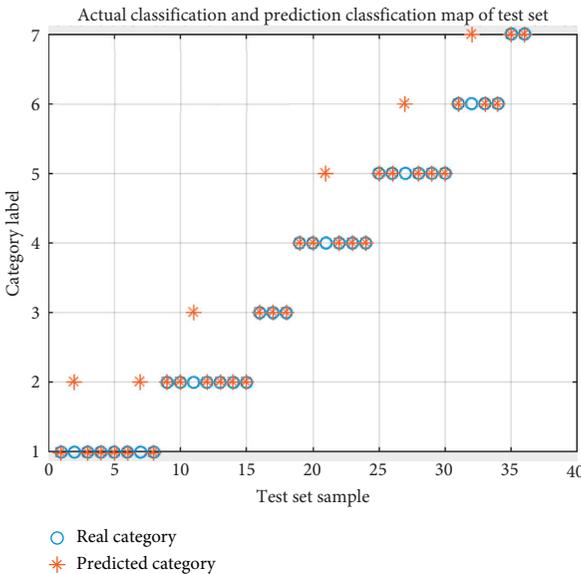


FIGURE 7: GWO-SVM test set classification identification.

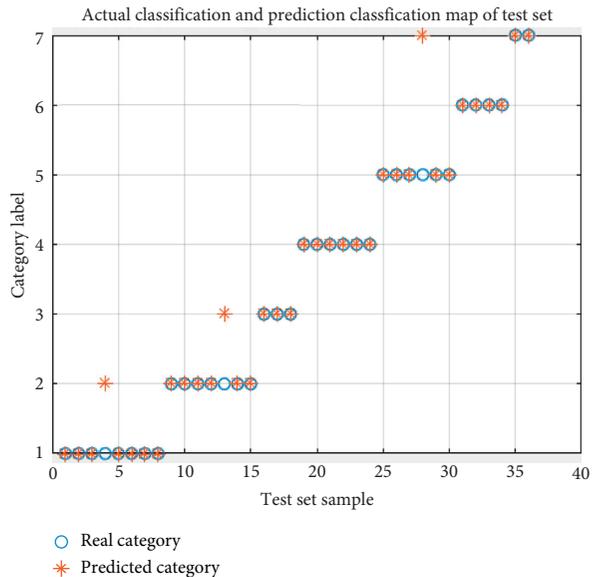


FIGURE 9: Classification identification of the PCA-GWO-SVM test set.

At the same time, the data also showed that the classification accuracy of the PCA-GWO-SVM model was greatly improved compared with that of the former one, indicating that PCA does not only simplified the complexity of the fault model but also eliminated the redundant information of the model and further improved the accuracy of the model. Finally, to verify the effectiveness of the proposed hybrid model, the comparative analysis was conducted with several other methods for model verification. According to Table 3, it is proved in this paper that the recognition accuracy of the proposed model is the highest. Moreover, the running time of the model is shorter than that of the PCA-GWO-SVM model, which is a relatively better

fault diagnosis model. This shows that the improved GWO of differential evolution adopted solves the problem that the model falls into the optimal local solution through a series of cross mutation and selection operations, which significantly improves the performance of the algorithm and ensures the global optimization effect. Based on the above analysis, the hybrid model optimized by the hybrid grey wolf optimizer proposed in this paper is an effective fault diagnosis method for belt conveyor. With the application of this method, the fault of the belt conveyor could be timely detected in advance

TABLE 4: Classification accuracy table of different models.

Model type	Train set accuracy (%)	Test set accuracy (%)	Operation time (s)
PCA-SVM	78.89 (71/90)	77.78 (28/36)	0.965
GWO-SVM	84.44 (76/90)	80.55 (29/36)	11.9546
PCA-GWO-SVM	91.11 (82/90)	91.67 (33/36)	5.1547
PCA-HGWO-SVM	97.78 (88/90)	97.22 (35/36)	4.4567

for treatment to avoid equipment loss or human injury accidents.

4. Conclusion

In this paper, a hybrid diagnosis model is developed based on the support vector machine model, which combines principal component analysis and grey wolf optimizer to apply to the fault classification of belt conveyor. Aiming at the limitations of the standard grey wolf optimizer, we proposed a method of hybrid wolf optimizer for parameter optimization. The experimental results show that the fault diagnosis model proposed in this paper has a higher overall diagnosis and recognition efficiency compared with the single method model, and its fault classification accuracy is up to 97.22%, which could help to improve the reliability of the belt transport system of the coal mine.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Key R&D Program of China (Grant no. 2017YFC0804408).

References

- [1] X. Liu, D. He, G. Lodewijks, Y. Pang, and J. Mei, "Integrated decision making for predictive maintenance of belt conveyor systems," *Reliability Engineering & System Safety*, vol. 188, pp. 347–351, 2019.
- [2] A. Miriam and G. Anna, "Classification of impact damage on a rubber-textile conveyor belt using Naïve-Bayes methodology," *Wear*, vol. 414-415, pp. 59–67, 2018.
- [3] F. Gabriel, M. Vierslav, G. Anna et al., "Failure analysis of irreversible changes in the construction of rubber-textile conveyor belt damaged by sharp-edge material impact," *Engineering Failure Analysis*, vol. 39, pp. 135–148, 2014.
- [4] M. Peter, M. Katarina, M. Vladimir et al., "Study of a tap failure at the internal threads machining," *Engineering Failure Analysis*, vol. 100, pp. 25–36, 2019.
- [5] D. G. Papageorgiou, G. Manolas, C. Medrea, and I. Chicinaş, "Failure analysis of a lancing tool used in a bridge slot filter pipe machine," *Engineering Failure Analysis*, vol. 105, pp. 237–245, 2019.
- [6] W. B. Song, "Study on the mechanism and countermeasures of the deviation fault of coal mine belt conveyor," *Shandong Coal Science and Technology*, vol. 05, pp. 121-122, 2015.
- [7] W. Jacek, Z. Rafał, W. Agnieszka et al., "Nonnegative factorization of spectrogram for local damage detection of belt conveyor gearboxes," *IFAC Papers Online*, vol. 50, no. 1, pp. 4714–4718, 2017.
- [8] B. Walter, C. Fakher, Z. Radoslaw et al., "Modelling of gearbox dynamics under time-varying nonstationary load for distributed fault detection and diagnosis," *European Journal of Mechanics-A Solids*, vol. 29, no. 4, pp. 637–646, 2010.
- [9] M. Andrejiova, A. Grincova, and D. Marasova, "Monitoring dynamic loading of conveyor belts by measuring local peak impact forces," *Measurement*, vol. 158, Article ID 107690, 2020.
- [10] K. Tomasz, B. Ryszard, L. Jurdzia et al., "Magnetic methods in monitoring changes of the technical condition of splices in steel cord conveyor belt," *Engineering Failure Analysis*, vol. 104, pp. 462–470, 2019.
- [11] M. Zhang, "Reciprocating compressor fault diagnosis technology based on multi-source information fusion," *Journal of Mechanical Engineering*, vol. 53, no. 23, pp. 46–52, 2017.
- [12] J. F. Wang, "Signal processing method for mechanical fault diagnosis: frequency domain analysis," *Noise and Vibration Control*, vol. 33, no. 01, pp. 173–180, 2013.
- [13] S. Ravikumar, H. Kanagasabapathy, and V. Muralidharan, "Fault diagnosis of self-aligning troughing rollers in belt conveyor system using *k*-star Algorithm," *Measurement*, vol. 133, pp. 341–349, 2019.
- [14] G. Arup, G. N. Wang, and L. Jooyeoun, "A novel automata and neural network-based fault diagnosis system for PLC controlled manufacturing systems," *Computers & Industrial Engineering*, vol. 139, pp. 135–148, 2020.
- [15] D. Maren, S. Zakwan, and K. J. Ian, "An integrated machine learning model for aircraft components rare failure prognostics with log-based dataset," *ISA Transactions*, 2020.
- [16] V. David and M. Dariusz, "Application of selected Levy processes for degradation modelling of long range mine belt using real-time data," *Archives of Civil and Mechanical Engineering*, vol. 18, no. 4, pp. 1430–1440, 2018.
- [17] M. Jacopo, C. Luca, R. Riccardo et al., "A bearing fault model for Independent Cart Conveyor System and its validation," *Applied Acoustics*, vol. 159, Article ID 107069, 2020.
- [18] H. Jaroslav, G. Robert, and K. Peter, "Removal of systematic failure of belt conveyor drive by reducing vibrations," *Engineering Failure Analysis*, vol. 99, pp. 192–202, 2019.
- [19] G. John, J. J. Zhang, and P. Wang, "Multi-stream convolutional neural network-based fault diagnosis for variable frequency drives in sustainable manufacturing systems," *Procedia Manufacturing*, vol. 43, pp. 511–518, 2020.
- [20] L. X. Yang and Y. L. Zhu, "Fault diagnosis of high voltage circuit breaker based on probabilistic neural network," *Power System Protection and Control*, vol. 43, no. 10, pp. 62–67, 2015.
- [21] G. Chen, X. Zhang, Z. J. Wang, and F. Li, "An enhanced artificial bee colony-based support vector machine for image-

- based fault detection,” *Mathematical Problems in Engineering*, vol. 2015, pp. 1–12, 2015.
- [22] X. Liu, D. Pei, G. Lodewijks, Z. Zhao, and J. Mei, “Acoustic signal based fault detection on belt conveyor idlers using machine learning,” *Advanced Powder Technology*, vol. 31, no. 7, pp. 2689–2698, 2020.
- [23] Q. R. Li, “Common fault analysis and countermeasure treatment of coal mine belt conveyor,” *Coal Mine Modernization*, vol. 01, pp. 102–104, 2016.
- [24] G. Z. Geng, *Application of Support Vector Machine in Fault Diagnosis of Power Generation Equipment*, Tianjin University, Tianjin, China, 2016.
- [25] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” in 5th Annual ACM Workshop on COLT, D. Haussler, Ed., pp. 144–152, ACM Press, Pittsburgh, PA, USA, 1992.
- [26] X. G. Zhang, “On statistical learning theory and support vector machine,” *Acta Automatica Sinica*, vol. 01, pp. 36–46, 2000.
- [27] D. Sun, *Research on Support Vector Machine Classification and Regression Method*, Central South University, Changsha, China, 2004.
- [28] H. Chen, *Optimization of Turbine Governor Parameters Based on Grey Wolf Algorithm*, Huazhong University of Science and Technology, Wuhan, China, 2016.
- [29] X. Jin, Z. C. Shao, and S. H. Wang, “A hybrid optimization algorithm based on differential evolution and grey wolf algorithm,” *Science Technology and Engineering*, vol. 17, no. 16, pp. 266–269, 2017.
- [30] H. N. Qi, “Overview of support vector machines and their applications,” *Computer Engineering*, vol. 10, pp. 6–9, 2004.
- [31] H. X. Li, “Failure analysis and treatment of coal mine belt conveyor,” *Energy and Energy Conservation*, vol. 12, pp. 7–8, 2017.
- [32] S. B. Hou, *Fault Diagnosis of Pumping Unit Working Condition Based on Firefly Support Vector Machine*, Northeast Petroleum University, Daqing, China, 2017.