IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 19 June 2022, accepted 8 July 2022, date of publication 18 July 2022, date of current version 25 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3192005

RESEARCH ARTICLE

Fault Diagnosis of Power Plant Condenser With the Optimized Deep Forest Algorithm

YUANYUAN JU[1,2], ZILIANG CUI[1,2], AND QINGTAI XIAO[2,3]

[1]Faculty of Science, Kunming University of Science and Technology, Kunming, Yunnan 650093, China
[2]State Key Laboratory of Complex Nonferrous Metal Resources Clean Utilization, Kunming University of Science and Technology, Kunming, Yunnan 650093, China
[3]Faculty of Metallurgical and Energy Engineering, Kunming University of Science and Technology, Kunming, Yunnan 650093, China

Corresponding author: Qingtai Xiao (qingtai.xiao@kust.edu.cn)

This work was supported in part by the Postdoctoral Directional Training Funding of Yunnan Province, China, under Grant 109820210027; in part by the Scientific Research Fund Project of Yunnan Education Department, China, under Grant 2022J0059 and Grant 2021J0063; in part by the Natural Science Foundation of Yunnan Province, China, under Grant 202101AU070031; and in part by the Young Elite Scientist Sponsorship Program by the China Association for Science and Technology (CAST), China, under Grant YESS20210106.

ABSTRACT As an important component of power plant operation, condenser fault diagnosis plays a vital role in the safe and stable unit performance. However, the precision of most existing diagnostic methods is not high enough for condenser fault diagnosis. It is considerably difficult to diagnose a condenser fault even under various complicated conditions. In this study, a novel classification hybrid model (PCA-DF) combining the Principal Component Analysis (PCA) method with the Deep Forest (DF) model is proposed based on the motivation of improving the diagnosis accuracy of condenser fault. The algorithm of this hybrid model takes the dimension reduction result of the PCA method as the input to the DF model. The multigrained scanning structure and the dimension reduction method are considered to create a good effect. The experimental results verify the feasibility and effectiveness of this method on the historical fault sample data of the condenser. The focus on the work presented of this paper is to optimize the DF model based on PCA and study the fault diagnosis effect of the hybrid model. Results show that (1) the prediction accuracy for the condenser fault diagnosis can be improved by increasing the sample size with the PCA-DF method. (2) The accuracy of the results obtained by proposing the improved hybrid models is 1%-8% higher than the accuracy of the results obtained by directly introducing the DF model, when the proportion of test set is less than or equal to 30%. The modified hybrid models still have advantages over the DF model for a small sample. (3) With an increase in the proportion of training sets, the accuracy of the modified hybrid models is improved correspondingly from 88.18% to 99.23%. (4) Compared with the backpropagation neural network, convolutional neural network, relevance vector machine and kernel Fisher discriminant analysis model, the PCA-DF model has higher accuracy. In this study, the proposed models can eliminate the influence of autocorrelation between data, and condenser fault diagnosis based on modified models has the fastest convergence speed and best accuracy. Furthermore, the proposed novel models can be extended to more complex fault diagnosis in other fields.

INDEX TERMS Condenser, deep forest, fault diagnosis, principal component analysis, power plant.

I. INTRODUCTION

Thermal power plants are important for producing electric energy. Numerous components are subjected to considerably high temperatures and pressures, making the system prone to failure, which can lead to significant loss of life and property. Therefore, it is essential for a power plant to develop fault diagnosis methods for these components [1]. During the production process of thermal power plants, the thermal cycle follows the Rankine cycle. The condenser, which acts

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Wang.

as the cold source, is a crucial link in this cycle. However, the condenser often breaks down, and the faults are very complex. As is well known, the failure of a condenser, which is one of the extremely important pieces of auxiliary equipment, can result in a significant loss of life and property in the thermal power plant. Applying the fault diagnosis technique to a condenser is beneficial for reducing the energy consumption and enhancing the energy efficiency. Hence, fault diagnosis of condensers has always been a concern in industrial production. In condenser monitoring, it is very important to determine whether there is a fault or its cause. Condenser faults include mainly operating in low vacuum, condensate of high oxygen content, and condensate of subcooling, among which operating in a low vacuum is considered the most common fault. There are many reasons for operating in low vacuum, and sometimes several kinds of symptoms may appear at the same time. The most common method for diagnosing faults is utilizing the symptoms corresponding to the faults. However, the similarity among symptoms and the correlation between multiple symptoms have led to difficulties in condenser diagnosis with different faults.

In the past few years, many diagnostic methods have been investigated and applied for the fault diagnosis of thermal systems. For instance, Ma *et al.* [2] proposed an artificial neural network (ANN) combined with an optimal zoom search and used it to recognize varying degrees of faults in the thermal system operation of power plants at different load levels. Wang *et al.* [3] presented a novel nearest prototype classifier for fault diagnosis in a power plant. Ge *et al.* [4] investigated a kernel-driven semisupervised Fisher discriminant analysis model to classify nonlinear faults in industrial processes. Kang *et al.* [5] used a fuzzy inference system to investigate the feedwater heaters with performance degradation in power generation equipment. Wang *et al.* [6] proposed a hierarchical depth domain adaptive method that transfers the classifier trained on labeled data under one loading condition to the classifier trained for unlabeled data under another loading condition in a power plant thermal system. Although these methods have achieved good performance, some processes rely heavily on a great deal of prior knowledge and human labor.

Recently, machine learning has been successfully applied to fault diagnosis, which is always viewed as a discriminative problem. Examples include the backpropagation neural network (BPNN) [7], support vector machine (SVM) [8], kernel Fisher discriminant analysis (KFDA) [8], random forests (RF) [9], and deep learning (DL) [10]. Currently, research on fault diagnosis based on deep learning has focused mainly on deep neural networks. Although this method has achieved good results in the fault diagnosis of various equipment, its diagnostic performance depends on several high-quality training samples. Most diagnostic methods based on deep neural networks were developed and verified using large training samples [11], including convolutional neural network (CNN) and relevance vector machine (RVM). Large high-quality training samples are hard to be collected from many industries. Under this condition, existing methods cannot accurately classify and identify all types of faults in the training data or effectively extract fault features to identify all types of faults in the test data. At present, this problem has been investigated. Hu *et al.* [12] proposed a generative adversarial network based on trinetworks form (tnGAN) to handle leak detection problems with incomplete sensor data. Ma *et al.* [13] proposed a graph-theory-based network partitioning algorithm to realize decentralized detection in a faster response speed. Since the deep forest (DF) was proposed by Zhou and Feng [14], research on fault diagnosis and feature recognition has expanded in this direction. Pan and Chen [15] proposed an intelligent fault detection method for shipborne antennas based on multiscale inner product and local connection feature extraction. Saufi *et al.* [16] developed a fault diagnosis system for time-frequency image pattern recognition using a deep learning model. Liu *et al.* [17] applied a multiscale kernel algorithm to capture patterns and deepen the convolutional neural network depth to extract fault features from deep and hierarchical representation spaces.

DF is a new decision-tree ensemble method that introduces a multigrained cascade forest (gcForest) structure. This method generates a DF set with a cascade structure to enable gcForest to perform representational learning. Hua *et al.* [18] proposed a novel approach combining a deep Boltzmann machine with a multigrained scanning forest ensemble to effectively treat industrial fault diagnosis using big data. Liu *et al.* [19] proposed an end-to-end intelligent fault diagnosis method for hydraulic turbines based on DF. In this method, multigrained scanning is used to transform fault feature representation from the original data to enhance the fault feature learning ability. Then, cascade structures are constructed using different types of random forests, and fault characteristics are a learned step to achieve fault classification. Ding *et al.* [20] proposed a weighted cascade forest based on multigrained scanning for fault diagnosis in chemical processes. At present, there are few studies on DF models for condenser fault diagnosis. As a deep learning model of nonneural networks, DF has a strong feature learning ability and is an effective means for multitype mixed fault diagnosis problems.

In practical applications, multivariate statistical analysis researches the statistical regularity of interdependence among multiple variables and has been widely applied in various industrial processes. Principal component analysis (PCA) is one of the most common multivariate techniques in the fields including signal and image processing, quality and process monitoring, and pattern recognition [21]. PCA can effectively identify a set of variables that represents the entire dataset, thus simplifying the interpretation and identifying significant characteristics. Many modifications of PCA have been developed for fault diagnosis, including dynamic PCA, nonlinear PCA, recursive PCA, sensitive PCA, slow-feature PCA, sparse PCA, and recursive transformed PCA [22]. In addition, novel improved PCA methods have been developed.

For example, Cho *et al.* [23] proposed a kernel PCA method for fault identification in process monitoring, in which practicability and performance are superior to the practicability and performance of nonkernel PCA. Zass and Shashua [24] described a nonnegative variable with sparse PCA, which created a sparse representation by constructing a low-dimensional representation from a collection of points. Chen and Role [25] presented a novel basis for sparse PCA, which approximately transforms an eigenbasis matrix into a sparse matrix by rotation. There are many methods to reduce dimensions, and PCA is just one of these methods. PCA is a well-known method for dimension reduction. Before DF fault diagnosis, using PCA to select important features can reduce the impact of unimportant features on fault diagnosis, eliminate the correlation between data points, and improve the accuracy of fault diagnosis.

In this study, a novel algorithm for improving the accuracy of condenser fault diagnosis is proposed by combining the advantages of DF and PCA. The PCA method is considered because it is a classical method, which is used to conduct a trial study, laying a foundation for future studies combined with more dimension reduction methods. The contribution of this study is that some novel PCA-DF models are presented for the accurate fault diagnosis of condensers. Except for the condenser, the pipeline system is also a potential object [26]. The advantages of the proposed models are summarized as follows. 1) The DF algorithm for condenser fault diagnosis can obtain a higher prediction accuracy than traditional methods. 2) The proposed models are simple, effective, and easy to implement, and can improve the accuracy of fault diagnosis in thermal power plants. 3) The proposed PCA-DF model can handle and eliminate the influence of autocorrelation among the condenser fault data points, and fault diagnosis performance can be significantly improved. 4) The proposed models effectively solve the contradiction between the long characteristics of fault data and the high cost of the DF model data.

The remainder of this paper is organized as follows. In Section II, the condenser system is briefly described and the common faults categories and symptoms of the condenser are summarized. In Section III, some novel hybrid models referred to as the modified PCA-DF models combining the PCA method with the DF algorithm are provided for identifying fault types. In Section IV, the historical condenser fault sample data are used to verify the feasibility and effectiveness of the proposed models. Finally, the conclusions are summarized in Section V.

II. DESCRIPTION OF CONDENSER SYSTEM

A. DESCRIPTION OF THE CONDENSER

For a condensing steam turbine, the effect of the condenser is equivalent to a ‘‘cold source’’. The main task of the condenser is that the steam turbine low-pressure cylinder exhaust condenses into water, and a vacuum is obtained in the steam turbine low-pressure cylinder exhaust port. The exhaust of the steam turbine is fully expanded. Thus, the enthalpy drop in the steam turbine exhaust can be increased. Finally, the circulating heat efficiency of the entire unit is improved. Because of the design, installation, maintenance, operating mechanism, and other reasons, the condenser often exhibits a low-vacuum phenomenon. If the condenser vacuum is too low, not only does the unit cause the steam enthalpy drop in the effective reduction and in the thermal efficiency decrease, but it also leads to an increase in faults categories, which include turbine exhaust steam temperature, exhaust hood deformation and vibration. A condenser system is an indispensable part of the modern thermal power-generating unit, which is composed mainly of a condensate pump, circulating pump, condenser, exhaust, and related valves and pipes. The condenser is the core of all parts. Fig. 1 shows the structural chart of the condenser system.

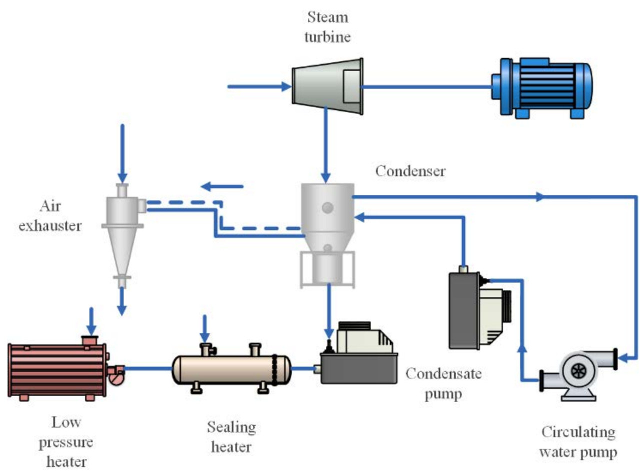

FIGURE 1. Structure diagram of condenser system.

B. COMMON FAULTS OF THE CONDENSER

The condenser can easily malfunction during the actual production process. According to [27], the common fault categories of condenser systems and the fault symptoms are listed in Table 1.

III. FAULT DIAGNOSIS METHODS

A. THE PCA FOR FEATURE EXTRACTION

The PCA model is a common linear statistical approach that is widely used in the fields of engineering and science applications, involving fault diagnosis, process monitoring, and image processing [28]. The basic concept of the PCA model is dimensionality reduction, which replaces the original multiple variables with several synthetic variables to achieve dimensionality reduction of variables. The selected comprehensive variables contain most of the information of the original variables, which can better reflect the original variables. PCA decomposes the covariance or correlation matrix obtained in the data matrix into eigenvectors, which are scaled to zero mean and unit variance and transformed into a transform subspace for dimensionality reduction [29]. Let $X \in \mathbb{R}^{n\times m}$ with monitoring fault values and n samples be a multidimensional data matrix collected from condenser equipment. The original variables in the X are considered correlated. The total numbers of original variables and principal components are represented by m and k ($k \ll m$), respectively.

TABLE 1. Set of fault categories and fault symptoms.

Fault categories	Fault symptoms
Serious fault of circulating pump	Outlet pressure of circulating water pump
	Motor current of circulating water pump
	Temperature rise of circulating water
Interruption of steam supply for rear shaft seal	Condenser terminal difference
	Pressure difference between the air extractor and the air outlet
	Degree of subcooling of condensation water
	Temperature rise of circulating water
	Condenser vacuum
Rupture of vacuum piping	Condenser terminal difference
	Pressure difference between the air extractor and the air outlet
	Condenser vacuum
	Degree of subcooling of condensation water
	Temperature rise of circulating water
Full water in condenser	Outlet pressure of circulating water pump
	Temperature rise of circulating water
	Degree of subcooling of condensation water
Incompleteness of the vacuum system	Motor current of circulating vacuum system
	Condenser terminal difference
	Degree of subcooling of condensation water
	Condenser vacuum
	Temperature rise of circulating water
Abnormal work of condensate pump	Motor current of condensation pump
	Temperature difference between pumping temperature and inlet temperature of circulating water
	Outlet pressure of condenser pump
	Conductivity of condensation water
	Degree of subcooling of condensation water
Cracking of condenser copper pipe	Outlet pressure of circulating water pump
	Condenser vacuum
	Temperature rise of circulating water
Rupture of last stage lower feed pipe	Heater water level of last stage
Squalidity of condenser copper pipe	Condenser terminal difference
	Condenser vacuum
	Temperature rise of circulating water
	Outlet pressure of circulating water pump
Small amount of circulating water	Temperature rise of circulating water
	Degree of subcooling of condensation water
Air extractor fault	Expansion difference of low-pressure cylinder
	Pressure difference between the air extractor and the air outlet

1) MATHEMATICAL DESCRIPTION OF PCA

Assuming that $X = (x_{ij})_{n\times m}$ is the original fault data matrix with $i = 1, 2, \cdots, n$ and $j = 1, 2, \cdots, m$ and $x_j = (x_{1j}, x_{2j}, \cdots, x_{nj})^{\mathrm{T}}$ is a column vector. Without loss of generality, m new comprehensive variables are obtained using a linear combination of m original variables, and the expression is as follows

$$\begin{cases} Y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m \\ Y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m \\ \vdots \\ Y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mm}x_m \end{cases} \tag{1}$$

where $R_{m\times m} = (a_{ij})^{\mathrm{T}}$ is called the coefficient matrix of the principal component with $i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, m$. When Eq.(1) meets the following three conditions including that there is no correlation between F_i and F_j, the variance of F$_i F_i$ is greater than the variance of F_j,, and $a_{i1}^2 + a_{i2}^2 + \cdots + a_{im}^2 = 1$, $F_1, F_2, \cdots, F_m$ is called the first principal component, second principal component, $\cdots$, and mth principal component (PC) respectively. To effectively extract the primary fault information, the matrix is either scaled to zero mean, forming an observation covariance matrix, or scaled to unit variance constituting a correlation matrix [30].

2) MATHEMATICAL DESCRIPTION OF EPCA

Exploratory principal component analysis (EPCA) discussed by Chen and Rohe [25] is a new basis for sparse principal component analysis (SPCA). EPCA includes the introduction of SCA and SMA. SCA and SMA are new methods for sparse PCA and extension of two-way matrix analysis, respectively. Compared with alternative methods, the EPCA method is more stable and can explain more of the variance. For EPCA, the l_1-norm constraints on the loadings and the minimization of the matrix reconstruction error are shown as

$$\begin{cases} \max\limits_{Z,B,Y} & \|X - ZBY^T\|_F \\ \text{s.t} & Z \in \mathcal{V}(n,k) \\ & Y \in \mathcal{V}(n,k) \\ & \|Y\|_1 \le \gamma \end{cases} \tag{2}$$

where the columns of Y are principal component loadings, and γ is the sparsity-controlling parameter. The ZBY^{T} in Eq. (2) is viewed as an approximation of X. To solve the minimization of Eq. (2), the equivalence is expressed by

$$\begin{cases} \max\limits_{Z,Y} & \|ZBY^T\|_F \\ \text{s.t} & Z \in \mathcal{V}(n,k) \\ & Y \in \mathcal{V}(n,k) \\ & \|Y\|_1 \le \gamma \end{cases} \tag{3}$$

l_2-Inequality constraints are used to ensure the convexity of the feasible set in Eq.(3). See Eq.(4)

$$\begin{cases} \max\limits_{Z,Y} & \|ZBY^T\|_F \\ \text{s.t} & Z \in \mathcal{B}(n,k) \\ & Y \in \mathcal{B}(n,k) \\ & \|Y\|_1 \le \gamma \end{cases} \tag{4}$$

3) MATHEMATICAL DESCRIPTION OF NSPCA

Nonnegative sparse PCA (NSPCA), proposed by Zass and Shashua [24], is a nonnegative variable method for sparse PCA problem. NSPCA can construct a representation of a low dimension from a point set. Not only is the variance of the projection point maximized, but also a part of the original coordinates is used to create a sparse representation. Let the principal vectors $u_1, u_2, \cdots, u_k$ be the column vector of the matrix U. Adding a nonnegative constraint to the original PCA is represented as

$$\begin{cases} \max\limits_{U} & \frac{1}{2}\|U^T X\|_F^2 \\ \text{s.t} & U^T U = 1 \\ & U \geq 0 \end{cases} \tag{5}$$

By relaxing some conditions, a relaxed version of (5) can be obtained by

$$\begin{cases} \max\limits_{U} & \frac{1}{2}\|U^T X\|_F^2 - \frac{\alpha}{4}\|I - U^T U\|_F^2 - \beta \mathbf{1}^T U \mathbf{1} \\ s.t & U \geq 0 \end{cases} \tag{6}$$

where $\alpha \geq 0$ is a balancing parameter, $\beta \geq 0$ is a controlling parameter, and $\mathbf{1}$ is a column vector with all the elements equal to one.

4) MATHEMATICAL DESCRIPTION OF KPCA

Kernel principal component analysis (KPCA) is a popular statistical tool with nonlinear dimensionality reduction that is used to deal with a set of high-dimensional nonlinear data. Based on PCA, a nonlinear mapping named kernel function is introduced in the KPCA method, which maps data from the original space to a high-dimensional space.

B. DF MODEL

The DF model, also known as the multigrained cascade forest (gcForest) model, is a deep-learning derivative random forest [14] that inherits layer-by-layer processing, feature transformation, and a sufficiently complex model from deep learning. Compared with traditional deep learning models, the DF model has some excellent characteristics, including fast training speed, simple model architecture, and fewer superparameters and does not require large datasets. The DF model consists of two integrated components inspired by a deep neural network: cascade forest and multigrained scanning. The first integrated component is a cascade forest, which learns more diacritical representations under the supervision of input representations at each layer to provide more accurate predictions based on a set of random forests. Cascaded forests differ from deep neural networks. Numerous decision-tree forests are gathered in the cascade forest. Under the supervision of the input data in each layer, the classification characteristics are learned according to the cascaded layer. In the cascade structure, each layer is composed of an ensemble of multiple decision tree forests, which aim to encourage overall diversity by including different types of forests.

Each cascade consists of multiple forests, each of which consists of multiple decision trees. Each layer of the cascade receives feature information processed by the previous layer and inputs the processing results to the next layer. The final classification results are obtained by analyzing the results of the last layer. To encourage the diversity, each layers of cascade forest is composed of completely-random forests, which are composed of completely-random trees [31], and random forests, which are made up of multiple random trees [32]. The number of forests and trees are hyperparameters in each layer. The difference between a completely-random forest and a random forest is that the features of the completely-random forests are randomly selected in the process of random tree branching, whereas the features of the random forest are selected to branch through the best Gini coefficient. For any completely-random forest, each node of every tree in the forest randomly selects the features segmented as its branches until it finally grows into pure leaves. For any random forest, the segmentation feature is selected based on the optimal Gini value when each node of every tree is branched [33]. When an instance is input to the cascade layer, an estimate of the class distribution is produced for every forest. The class distribution outputs of all forests in the same layer form a class vector, which is then connected to the original vector and inputted to the next cascade. Simultaneously, when a new cascade layer is extended, cross validation is used to evaluate the overall performance. Once there is no significant improvement in performance, the extension process is terminated automatically [20]. As described in [14], the number of cascading layers can be determined adaptively by evaluating the performance of each layer. The cascade structure is illustrated in Fig. 2[14]. The second integrated component is multigrained scanning, which scans the local context from a high dimension by means of a variety of sliding window structures of different sizes, and learns the representation of input data according to different random forests. Inspired by the deep neural network processing feature relationship, DFs have adopted the strategy of multigrained scanning. Based on the sliding window method, the local features are extracted by scanning the original input, and a set of local low-dimensional feature vectors is generated. Then, a series of forests is trained, and the input vectors of the class distribution are obtained by using these low-dimensional vectors. A multigrained scanning process is used to enhance the cascade forest. In multigrained scanning (MGS), the purpose of the MGS is to extract helpful information from input image data or sequence data. Taking an instance of MGS, let $x_i = (x_{1i}, \cdots, x_{ni})^{\mathrm{T}}$ and $y = (y_1, \cdots, y_n)$ be the input instance and class label respectively, where $i = 1, \cdots, p$, which p is features capacity and n is sample capacity. The training datasets are $D_{train} = \{(x_1, y_1), \cdots, (x_n, y_n)\}$. The sizes of the raw input vector and sliding window are d_r-dim and d_s-dim, respectively. A d_s-dim feature vector with a step length of 1 is obtained when the sliding window slides once. A total of $(d_r - d_s + 1)$ feature vectors is generated. All the feature vectors extracted from the original instances of each category through sliding windows are treated as instances of that category. All instances are used to train the cascade forest. For each original instance, all sliding feature vectors are input into the two forests, and vectors of the class distribution are generated. These vectors are connected and converted into $2d_r(d_r - d_s + 1)$-dim vectors of the class distribution [33]. Multigrained scanning is shown in Fig. 3[14]. In Fig. 3, supposing there are three categories in the training data sample; the input data and the sliding window size are the 400 raw sequence features and 100 features respectively. Utilizing a 100-dim sliding window for the original sequence data, 301 feature vectors are extracted from each original instance. For each 100-dim vector, each forest generated a 3-dim class vector, and 301 3-dim class vectors are generated. An 1806-dim transformational class distribution vector corresponding to the original 400-dim original feature vector is obtained.

C. THE PCA-DF MODEL

The fault diagnosis process is essentially a pattern classification process that determines the corresponding fault categories according to certain symptoms. On the one hand, the dimension of the training samples directly affects the training process for DFs. However, there is a correlation between the fault symptoms. Without the extraction of characteristic parameters, the efficiency and accuracy of fault diagnosis will be affected by the direct use of all parameters detected for fault diagnosis. PCA can not only reduce the dimension of the fault data space, but also reflect the relationship between fault symptoms. Therefore, the DF model based on feature extraction with PCA can improve the accuracy of fault diagnosis and the efficiency of the algorithm. In this section, novel hybrid models (PCA-DF, EPCA-DF, NSPCA-DF, and KPCA-DF) combining the PCA method and DF algorithms are proposed to identify fault types. The framework of the condenser fault diagnosis method based on the PCA-DF model is shown in Fig. 4. First, PCA is used to extract crucial fault features from the condenser diagnosis sequence data. Second, the crucial fault features are treated as the input of the DF model. In the DF model, the feature vector is converted into a feature matrix to feed the MGS. Third, the outputs of the MGS are used for training the cascade forest. Finally, the final classification results are outputted through cascade forests. The implementation steps are described in detail in Table 2.

IV. RESULTS AND DISCUSSION

In this paper, the sample data comes from Table 1 to 3 in Appendix 2 of the reference [27], which is 65 historical condenser fault sample data obtained from the information monitoring system of a power plant. Thirteen symptom parameters are considered in the data: condenser vacuum (mbar), condenser terminal difference (°C), motor current of circulating water pump (A), outlet pressure of circulating water pump (bar), expansion difference of low-pressure cylinder (mm), motor cuurrent of condensation pump (A), outlet pressure of

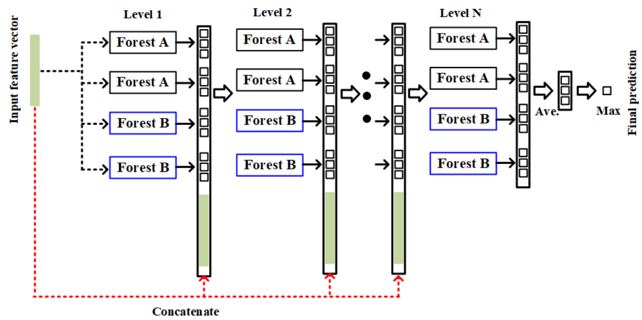

FIGURE 2. The cascade forest structure.

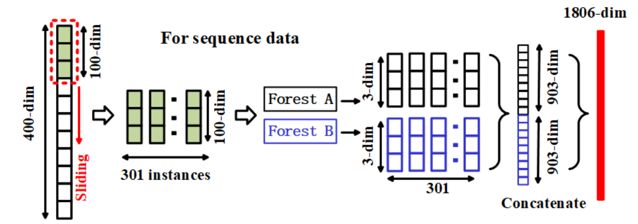

FIGURE 3. The multigrained scanning structure.

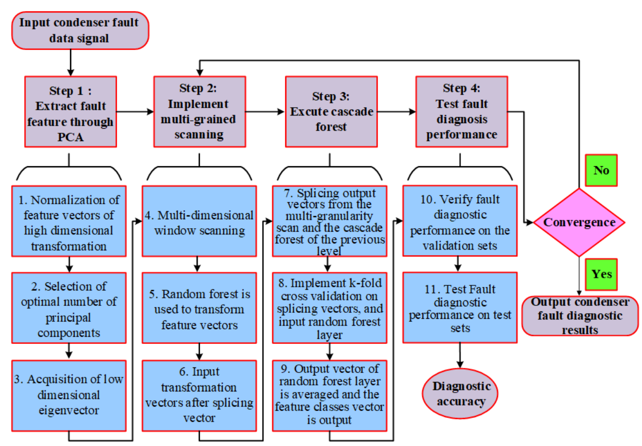

FIGURE 4. Framework of condenser fault diagnosis method based on PCA-DF model.

condenser pump (bar), degree of subcooling of condensation water (°C), heater water level of last stage (mm), conductivity of condensation water (μs/cm), temperature rise of circulating water (°C), temperature difference between pumping temperature and inlet temperature of circulating water (°C), and pressure difference between the air extractor and the air extraction port (MPa). After analyzing the data set, there are 65 samples in the data set, of which 25 samples are serious faults in the circulating water pump, 20 samples are steam supply interruption of the rear shaft seal and 20 samples are pipe break of the vacuum system. Therefore, this article uses these three main faults out of Table 1 to study the validity and feasibility of the proposed models in Section III.

In the experiment of this section, we study mainly the hybrid models of four dimensionality reduction methods (PCA, EPCA, KPCA, NSPCA) and DF, the criterion for

TABLE 2. The procedure of the modified hybrid model.

Algorithm: The procedure of modified hybrid model

Initialize parameters:

PCA method
1: Input initial fault data sequence of condenser
2: Reduce the number m with condenser fault symptoms to k dimensions through PCA method, get new sequence of fault feature

EPCA method
1: Input initial fault data sequence X of condenser and numerous of components k
2: Initialize $\hat{Z}$ (the polar of XY) with the top k left singular vectors of X and Y with the top k right singular vectors of X
3: Update $\hat{Z}$ and Y until convergence
4: Output sparse loadings$Y$

NSPCA method
1: Input initial fault data sequence of condenser
2: Start with an initial guess for U .
3: Iterate over entries (r, s) of U until convergence

KPCA method
1: Input initial fault data sequence of condenser
2: Using the kernel function to map fault sample, define the covariance matrix
3: Compute each eigenvalue of the covariance matrix and their corresponding eigenvectors
4: Reduce the number m with condenser fault symptoms to k dimensions through KPCA method, get new sequence of fault feature

Training phase:

Input: training set D_{train} after PCA
1: Setup the gcforest model
2: **if** input is training data
3: The multigrain scanning is implemented
4: Obtain the enhanced eigenvectors B_{train}
5 **for** n in **range** (amount of data in B_{train})
6: **for** i in **range** (number of layers)
7: **for** j in **range** (number of estimators for each layer)
8: Update the estimators
9: **end** (iterate all the j)
10: **end** (iterate all the i)
11: **end** (iterate all the n)
12: Show the validation accuracy on training data
13: Save the gcforest model

Test phase:

Input: test set D_{test} after PCA
1: Setup the gcforest model
2: **for** n in **range**(number of data inD_{test}):
3: Test the gcforest with D_{test}
4: The predicted value based on D_{test} is returned
5: **end** (iterate all the n)
Output: A vector V, each element corresponds to a classification label of a sample in D_{test}

selecting k features out of 13 features in each experiment is cumulative variance contribution rate. The calculation formula of cumulative contribution rate (r_k) is

$$r_k = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{13} \lambda_i} \tag{7}$$

where λ_i is an eigenvalue of the ith features. The k is the smallest integer that makes $r_k > 0.9$. All hybrid models are trained and tested with the fowllowing simulation environment: The CPU is i7-0875h. The GPU is RTX2060. The ram is 16 GB. The code implementation is based on Python 3.7 and R 4.1.1.

A. ANALYSIS OF DF MODEL

In the first study, a DF algorithm was introduced to directly diagnose condenser faults. The DF model was directly introduced to train and test the sample data. First, the data set is divided into 50% of the sample data as the training set and 50% of the sample data as the test set. Subsequently, the DF model is introduced to train the training set directly. Finally, the trained model is introduced to predict the test set, and the predicted results of ten times experiments are shown in Fig. 5(a). The test results show that 26 out of 330 sample points in the test set are predicted incorrectly, and the average prediction accuracy is 92.12%. To improve the prediction accuracy, the proportion of the training and test sets is adjusted to 6:4, and the training and prediction are repeated. The results are shown in the Fig. 5(b). The test results show that 20 out of 260 sample points in the test set are predicted incorrectly, and the average prediction accuracy is 96.15%. The proportion of the training and test sets is adjusted again to 7:3 for training and forecasting. The results are shown in Fig. 5(c). The test results show that 20 out of 200 sample points in the test set are predicted incorrectly, and the prediction accuracy decrease to 90.00%. Then the proportion of the training and test sets is adjusted to 8:2 for training and forecasting. The results are shown in Fig. 5(d). The test results show that 9 out of 130 sample points in the test set are predicted incorrectly, and the prediction accuracy becomes 93.07%. By comparing the results in Fig. 5(a), (b), (c) and (d), an interesting result was found that the prediction effect of the model was not improve with an increase in the amount of data in the training set. In the DF model, the prediction accuracy of the model should have increased with the increase in the sample size. Hence, the prediction accuracy of the DF model fluctuates significantly.

B. ANALYSIS OF HYBRID MODELS

In the experiment of this section, we study mainly the hybrid model of four dimensionality reduction methods (PCA, EPCA, KPCA, NSPCA) and DF.

In the second study, a novel hybrid model (PCA-DF) combining PCA and DF algorithm was used to diagnose condenser faults. We first used the PCA method to reduce the dimensionality of the original data by reducing the number of variables from 13 to 9. Second, based on these nine principal elements, the DF model was used for training and testing.

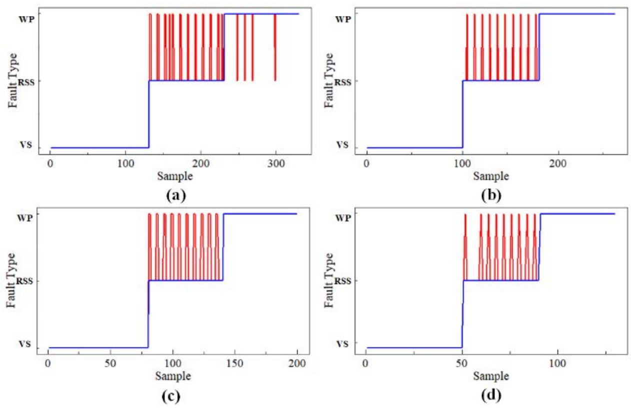

FIGURE 5. Condenser fault diagnosis with the different proportion (a) 5:5, (b) 6:4, (c) 7:3, and (d) 8:2 of the training and test sets based on DF model. WP, RSS, and VS of the ordinate represent three different serious faults in the circulating water pump, steam supply interruption of the rear shaft seal and pipe break of the vacuum system respectively.

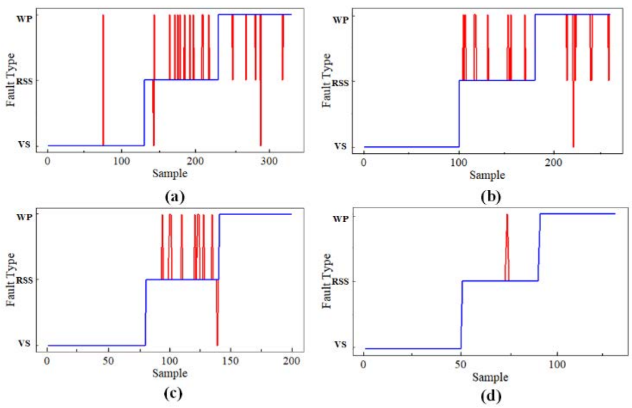

FIGURE 6. Condenser fault diagnosis with the different proportion (a) 5:5, (b) 6:4, (c) 7:3, and (d) 8:2 of the training and test sets based on PCA-DF model. WP, RSS, and VS of the ordinate represent three different serious faults in the circulating water pump, steam supply interruption of the rear shaft seal and pipe break of the vacuum system respectively.

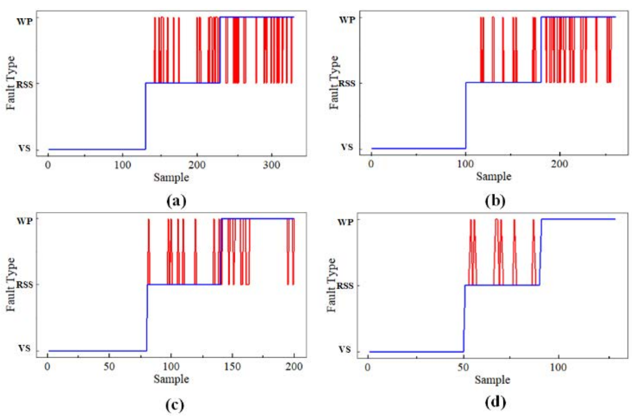

FIGURE 7. Condenser fault diagnosis with the different proportion (a) 5:5, (b) 6:4, (c) 7:3, and (d) 8:2 of the training and test sets based on EPCA-DF model. WP, RSS, and VS of the ordinate represent three different serious faults in the circulating water pump, steam supply interruption of the rear shaft seal and pipe break of the vacuum system respectively.

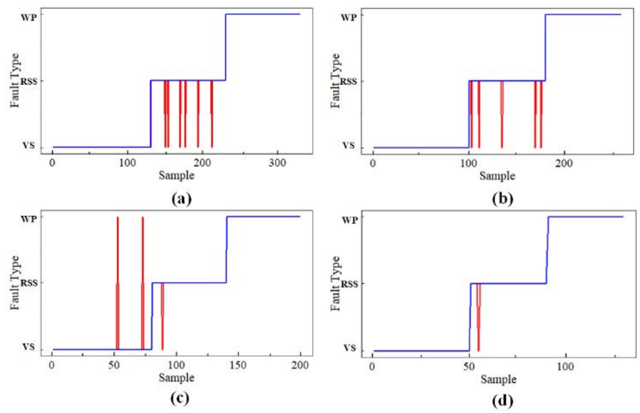

FIGURE 8. Condenser fault diagnosis with the different proportion (a) 5:5, (b) 6:4, (c) 7:3, and (d) 8:2 of the training and test sets based on NSPCA-DF model. WP, RSS, and VS of the ordinate represent three different serious faults in the circulating water pump, steam supply interruption of the rear shaft seal and pipe break of the vacuum system respectively.

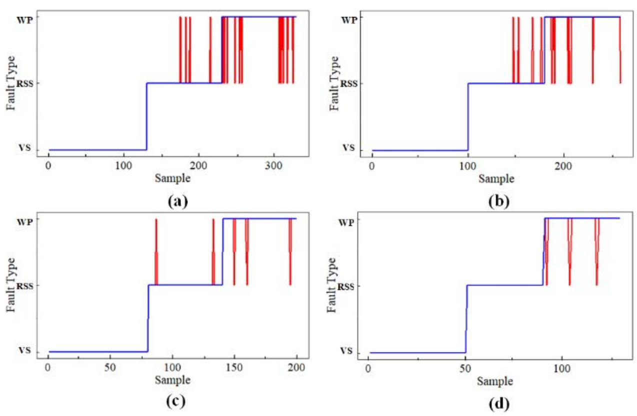

FIGURE 9. Condenser fault diagnosis with the different proportion (a) 5:5, (b) 6:4, (c) 7:3, and (d) 8:2 of the training and test sets based on KPCA-DF model. WP, RSS, and VS of the ordinate represent three different serious faults in the circulating water pump, steam supply interruption of the rear shaft seal and pipe break of the vacuum system respectively.

For comparison with the first study, the proportion of the training and test sets was still 5:5. The results are shown in Fig. 6 (a). The prediction accuracy of the DF model was found to be improved based on dimensionality reduction data after PCA. In the test set of 330 sample points, only 18 samples were predicted incorrectly, and the prediction accuracy of the entire model was improved to 94.55%. We also adjusted the proportion of training and test sets to 6:4. In the test set of 260 sample points, only 14 samples were incorrect, with an accuracy of 94.62%. The results are shown in Fig. 6 (b). The ratio of the training set to the test set was further adjusted to 7:3, and the training and prediction were reproduced to achieve 95.00% accuracy. The results are shown in Fig. 6 (c). Then the accuracy increased to 99.23% when the ratio of the training set to the test set was further adjusted to 8:2. The results are shown in Fig. 6 (d). From the above analysis results, except when the proportion of the test set is 40%, the accuracy of the results obtained by proposing the PCA-DF combined model is 2%-6% higher than the accuracy of the results obtained by directly introducing the DF model. In addition, the prediction accuracy of the PCA-DF model increases with the increase in sample size, so the PCA-DF model overcomes the problem of low data quality. If the PCA-DF hybrid model is trained using large datasets, its performance will improve.

In the third study, a novel hybrid model (EPCA-DF) combining EPCA and a DF algorithm was used to diagnose the condenser fault. We first used the EPCA method to reduce the dimensionality of the original data by reducing the number of variables from 13 to 5. Second, based on these five principal elements, the DF model was used for training and testing. For comparison with the first study, the proportion of the training and test sets was still 5:5. The results are shown in Fig. 7 (a). The prediction accuracy of the DF model was found not to be improved based on dimensionality reduction data after the EPCA. In the test set of 330 sample points, 39 samples were predicted incorrectly, and the prediction accuracy of the entire model was 88.18%. We also adjusted the proportion of training and test sets to 6:4. In the test set of 260 sample points, 28 samples were predicted incorrectly, and the accuracy rate was 89.23%. The results are shown in Fig. 7 (b). The ratio of the training set to the test set was further adjusted to 7:3, and the training and prediction were reproduced to achieve 91.00% accuracy. The results are shown in Fig. 7 (c). When the ratio of the training set to the test set was adjusted to 8:2, the accuracy increased to 94.62%. The results are shown in Fig. 7 (d). From the above analysis results, when the proportion of test set is less than or equal to 30%, the accuracy of the results obtained by proposing the EPCA-DF combined model is to be approximately 1% higher than the accuracy of the results obtained by directly introducing the DF model. In addition, the prediction accuracy of the EPCA-DF model also increases with the increase in sample size. If the EPCA-DF hybrid model is trained using large datasets, its performance will improve.

In the fourth study, a novel hybrid model (NSPCA-DF) combining NSPCA and the DF algorithm was used to diagnose the condenser fault. We first used the NSPCA method to reduce the dimensionality of the original data, reducing the number of variables from 13 to 7. Secondly, based on the 7 principal elements, the DF model was used for training and testing. To compare with the first study, the proportion of training set and test set are was still 5:5. The results are shown in the Fig. 8 (a). The prediction accuracy of the DF model was improved based on the dimensionality reduction data after NSPCA. In the test set of 330 sample points, only 7 samples were predicted incorrectly, and the prediction accuracy of the whole model was improved to 97.88%. Then, we also adjusted the proportion of the training set and the test set to 6:4. In the test set of 260 sample points, only 5 samples were predicted incorrectly, and the prediction accuracy of the whole model was improved to 98.08%. The results are shown in Fig. 8 (b). The ratio of the training set to the test set was further adjusted to 7:3 and reproduced the training and prediction, achieving 98.50% accuracy. The results are shown in Fig. 8 (c). When the ratio of the training set to the test set was adjusted to 8:2, the accuracy increased to 99.23%, and only one of the 130 sample points were predicted incorrectly. The results are shown in Fig. 8 (d). From the above analysis results, the accuracy of the results obtained by proposing the NSPCA-DF combined model are 2%-8.5% higher than the results obtained by introducing the DF model directly. In addition, the size of the sample data set is small in this paper, but the NSPCA-DF hybrid model can still gain advantages. If the NSPCA-DF hybrid model is trained with large data sets, its performance will be better.

In the fifth study, a novel hybrid model (KPCA-DF) combining KPCA and DF algorithm was used to diagnose the condenser fault. We first used the KPCA method to reduce the dimensionality of the original data, reducing the number of variables from 13 to 12. Secondly, based on the 12 principal elements, the DF model was used for training and testing. To compare with the first study, the proportion of training set and test set was still 5:5. The results are shown in Fig. 9 (a). The prediction accuracy of the DF model was found to be improved based on the dimensionality reduction data after KPCA. In the test set of 330 sample points, only 16 samples were predicted incorrectly, and the prediction accuracy of the whole model was improved to 95.15%. Then we also adjusted the proportion of the training set and the test set to 6:4. In the test set of 260 sample points, only 14 samples were predicted incorrectly, and the prediction accuracy of the whole model was improved to 95.77%. The results are shown in Fig. 9 (b). The ratio of the training set to the test set was further adjusted to 7:3, and reproduced the training and prediction, achieving 97.50% accuracy. The results are shown in Fig. 9 (c). When the ratio of the training set to the test set was adjusted to 8:2, the accuracy increased to 97.69%, and only 5 of the 130 sample points were predicted incorrectly. The results are shown in Fig. 9 (d). From the above analysis results, except when the proportion of the test set is 40%, the accuracy of the results obtained by proposing the KPCA-DF combined model is 3%-7% higher than the accuracy obtained by introducing the DF model directly. In addition, the size of the sample data set is small in this paper, but the KPCA-DF hybrid model can still gain advantages. If the KPCA-DF hybrid model is trained with large data sets, its performance will get better.

C. ANALYSIS OF HYBRID MODELS UNDER DIFFERENT PROPORTIONS OF TEST SET

In the sixth study, we considered the classification accuracy of various methods with DF, PCA-DF, EPCA-DF, NSPCA-DF and KPCA-DF under different proportions of training set and test set. BPNN, CNN, RVM and KFDA were tested to compare with the proposed hybrid model. Table 3 shows the average results of 10 experiments. The following conclusions can be drawn from Table 3: 1) when the proportion of the test set is less than or equal to 30%, the DF model has a lower accuracy than all the hybrid models, indicating that there is a correlation between different fault symptoms. Reducing the dimensions of symptom data by using PCA methods can significantly improve the accuracy of fault classification. 2) For the five models, different proportions of the training and test sets led to different fault classification accuracies. The larger the proportion of training and test sets is, the higher the prediction accuracy of the hybrid models. When the ratio of the training set to the test set is 8:2, the prediction accuracies of the hybrid models are the highest, which means that the hybrid model will show more advantages in large data sets. 3) When the ratio of the training set to the test set is 5:5, the prediction accuracy of the NSPCA-DF model is the highest at 97.88%. The accuracy of this model is 5.76% higher than the accuracy of the DF model. When the ratio of the training set to the test set is 6:4, the NSPCA-DF model exhibits the highest prediction accuracy, achieving 98.08% accuracy, which is 1.93% higher than the accuracy of the DF model. When the ratio of the training set to the test set is 7:3, the NSPCA-DF model still achieves the highest prediction accuracy, which is 98.5%. The accuracy of the NSPCA-DF model is 8.5% higher than the accuracy of the DF model. When the ratio of the training set to the test set is 8:2, the prediction results of the PCA-DF and NSPCA-DF models achieves 99.23%, which is 6.16% higher than the accuracy of the DF model. 4) With the increase in the proportion of the training sets, the accuracy of both the PCA-DF, EPCA-DF, NSPCA-DF, and KPCA-DF models improved accordingly. 5) The accuracy of the DF-based model is 8%-22% higher than the accuracy of the BPNN model, and 13%-26% higher than the accuracy of the CNN model, and 8%-27% higher than the accuracy of the KFDA model. The accuracy of RVM model is the lowest among all models under the different proportion of test set. These results show that the DF-based model performs better on low-quality data sets.

D. ANALYSIS OF HYBRID MODEL FOR THE THREE TYPES OF FAULTS

In the seventh study, the prediction results of the condenser fault diagnosis were elaborated for five models. Table 4 lists the results of the condenser fault diagnosis with various hybrid models based on four representative test set proportions (20%, 30%, 40% and 50%). The diagnostic effects of the three types of faults (serious fault of circulating water pump, steam supply interruption of rear shaft seal, and pipe break of vacuum system, represented by WP, RSSβ and VS) are displayed in the Table 4. Through the analysis of Table 4, the following conclusions can be drawn: 1) The lower the proportion of test sets is, the better the diagnostic effect of the three types of faults. 2) The correct diagnosis of the all five models occurred in a serious fault of the circulating water pump, when the proportion of test sets is 40% and 20%. For the DF models, incorrect predictions occurred in the steam supply interruption of the rear shaft seal. 3) For the PCA-DF, EPCA-DF, NSPCA-DF, and KPCA-DF models, the diagnostic performance for the three types of faults was obviously better than the diagnostic performance of the DF model. Eliminating the influence of data autocorrelation was found to be able to significantly improve the diagnosis effect of the three types of faults. If the sample size of the training set is further increased, this difference will be more obvious, and the prediction accuracy of the methods proposed in this study can even reach more than 99%.

TABLE 3. Classification accuracy of various methods under different proportions of test sets.

Classification accuracy	20.00%	30.00%	40.00%	50.00%
DF	93.07%	90.00%	96.15%	92.12%
PCA-DF	99.23%	95.00%	94.62%	94.55%
EPCA-DF	94.62%	91.00%	89.23%	88.18%
KPCA-DF	97.69%	97.50%	95.77%	95.15%
NSPCA-DF	99.23%	98.50%	98.08%	97.88%
BPNN	84.62%	78.50%	76.15%	75.76%
CNN	78.46%	76.50%	73.08%	71.21%
RVM	46.92%	44.21%	39.62%	37.81%
KFDA	72.31%	82.00%	79.62%	73.03%

TABLE 4. Condenser fault diagnosis results with using various model for the three types of faults under different proportion of test set.

Classification accuracy	50.00%			40.00%		
	WP	RSS	VS	WP	RSS	VS
DF	100%	78.00%	96.00%	100%	87.50%	100%
PCA-DF	99.23%	88.00%	95.00%	100%	90.00%	92.50%
EPCA-DF	100%	79.00%	82.00%	100%	88.75%	76.25%
NSPCA-DF	100%	93.00%	100%	100%	93.75%	100%
KPCA-DF	100%	95.00%	89.00%	100%	95.00%	91.25%
Classification accuracy	30.00%			20.00%		
	WP	RSS	VS	WP	RSS	VS
DF	100%	66.66%	100%	100%	77.50%	100%
PCA-DF	100%	83.33%	100%	100%	97.50%	100%
EPCA-DF	100%	86.67%	83.33%	100%	82.50%	100%
NSPCA-DF	97.50%	98.33%	100%	100%	97.50%	100%
KPCA-DF	100%	96.67%	95.00%	100%	100%	92.50%

E. ANALYSIS OF TIME REQUIRED FOR THE FAULT DIAGNOSIS PROCESS

In the study, the time required for the fault diagnosis process was discussed. Table 5 lists the time required for the DF, PCA-DF, EPCA-DF, NSPCA-DF, KPCA-DF, BPNN, CNN, RVM and KFDA models for the fault diagnosis process. Through Table 5, the following conclusions can be drawn: 1) The BPNN model takes the longest time to complete the fault diagnosis process, while KFDA model takes the shortest time. 2) If the time of training and testing is considered at the same time, the time of the EPCA-DF and NSPCA-DF models is shorter than the time of the DF model. If only the test time is considered, the PCA-DF, EPCA-DF and NSPCA-DF models all need less time than the DF model. 3) The time required for KPCA-DF is longer than other DF-based models. 4) The time of fault diagnosis process of DF-based models is between 0.85-1.13, while the time taken by other comparative models including BPNN, CNN, RVM and KFDA is between 0.04-4.99 shows a great fluctuation. These results show that the hybrid models can reduce the time required for fault diagnosis to a certain extent and have a stable time requirements.

TABLE 5. Time required for the fault diagnosis process.

Time (s)	Train and Test	Test
DF	0.862906	0.017952
PCA-DF	0.869673	0.015958
EPCA-DF	0.858674	0.015958
NSPCA-DF	0.845778	0.015957
KPCA-DF	1.128024	0.032912
BPNN	4.987936	0.063217
CNN	2.556557	0.069039
RVM	0.139762	0.001137
KFDA	0.037321	0.000000

V. CONCLUSION

The purpose of this work is to diagnose a condenser fault by utilizing modified DF models including PCA-DF, EPCA-DF, NSPCA-DF, and KPCA-DF. The main conclusions of this study are briefly summarized as follows. 1) When the DF algorithm was introduced for the fault diagnosis of the condenser, the prediction accuracy improved as the sample size increased. 2) The accuracy of the results obtained by proposing the improved hybrid models was 1%-8% higher than the accuracy obtained by directly introducing the DF model, when the proportion of test set is less than or equal to 30%. In the case of small-sample datasets, the modified hybrid model still has advantages over the DF model. 3) With an increase in the proportion of training sets, the accuracy of all the four modified hybrid model is improved accordingly. The accuracy of the hybrid models was higher than the accuracy of the DF model when the proportion of the training set was more than 40%. 4) Through comparative analysis, the lower the proportion of test sets is, the better the diagnostic effect of the three types of faults. The accuracy of fault diagnosis of a hybrid model can reach more than 99%. For the DF models, incorrect predictions all occurred in the steam supply interruption of the rear shaft seal. 5) For the PCA-DF, NSPCA-DF, and KPCA-DF models, the diagnostic performance of the three types of faults is obviously better than the diagnostic performance of the DF model. 6) The existence of autocorrelation affects the fault diagnosis of the condenser, and the introduction of PCA methods can eliminate the influence of autocorrelation between data points, thus significantly improving the fault diagnosis performance. In conclusion, the proposed hybrid models including PCA-DF, EPCA-DF, NSPCA-DF and KPCA-DF models, have the advantage of prediction accuracy in both large and small data sets. However, EPCA-DF and NSPCA-DF models are recommended because of their short runtimes.

Our proposed hybrid models can not only successfully achieve condenser faults diagnosis but can also be extended to faults diagnosis in other fields, such as pipeline system fault and bearing fault diagnosis. However, when analyzing the original DF framework, PCA is found to be able to be tried before multigranularity scanning. In addition, for dimensionality reduction of fault features, variable selection and feature screening in statistics can be considered, which makes it possible to improve the model. In the future, the methods in this study will be further optimized.

REFERENCES

[1] A. Ali and D. Mahdi, "Independent component analysis approach for fault diagnosis of condenser system in thermal power plant," *J. Central South Univ.*, vol. 21, no. 1, pp. 242–251, Mar. 2014.

[2] L. Ma, Y. Ma, and K. Y. Lee, "An intelligent power plant fault diagnostics for varying degree of severity and loading conditions," *IEEE Trans. Energy Convers.*, vol. 25, no. 2, pp. 546–554, Jun. 2010.

[3] X. Wang, L. Ma, and T. Wang, "An optimized nearest prototype classifier for power plant fault diagnosis using hybrid particle swarm optimization algorithm," *Int. J. Electr. Power Energy Syst.*, vol. 58, pp. 257–265, Jun. 2014.

[4] Z. Ge, S. Zhong, and Y. Zhang, "Semisupervised kernel learning for FDA model and its application for fault classification in industrial processes," *IEEE Trans. Ind. Informat.*, vol. 12, no. 4, pp. 1403–1411, Aug. 2016.

[5] Y. K. Kang, H. Kim, G. Heo, and S. Y. Song, "Diagnosis of feedwater heater performance degradation using fuzzy inference system," *Expert Syst. Appl.*, vol. 69, pp. 239–246, Mar. 2017.

[6] X. Wang, H. He, and L. Li, "A hierarchical deep domain adaptation approach for fault diagnosis of power plant thermal system," *IEEE Trans. Ind. Informat.*, vol. 15, no. 9, pp. 5139–5148, Sep. 2019.

[7] P. Zhang, Z. Cui, Y. Wang, and S. Ding, "Application of BPNN optimized by chaotic adaptive gravity search and particle swarm optimization algorithms for fault diagnosis of electrical machine drive system," *Electr. Eng.*, vol. 104, no. 2, pp. 819–831, Apr. 2022.

[8] L. H. Chiang, M. E. Kotanchek, and A. K. Kordon, "Fault diagnosis based on Fisher discriminant analysis and support vector machines," *Comput. Chem. Eng.*, vol. 28, no. 8, pp. 1389–1401, Jul. 2004.

[9] T. G. Dietterich, "An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization," *Mach. Learn.*, vol. 40, no. 2, pp. 139–157, 2000.

[10] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[11] S. R. Saufi, Z. Ahmad, M. S. Leong, and M. H. Lim, "Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review," *IEEE Access*, vol. 7, pp. 122644–122662, 2019.

[12] X. Hu, H. Zhang, D. Ma, and R. Wang, "A tnGAN-based leak detection method for pipeline network considering incomplete sensor data," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–10, 2021.

[13] D. Ma, X. Hu, H. Zhang, Q. Sun, and X. Xie, "A hierarchical event detection method based on spectral theory of multidimensional matrix for power system," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 4, pp. 2173–2186, Apr. 2019.

[14] Z. Zhou and F. Ji, "Deep forest: Towards an alternative to deep neural networks," *Nat. Sci. Rev.*, vol. 6, no. 1, pp. 74–86, Oct. 2019.

[15] T. Pan, J. Chen, Z. Zhou, C. Wang, and S. He, "A novel deep learning network via multiscale inner product with locally connected feature extraction for intelligent fault detection," *IEEE Trans. Ind. Informat.*, vol. 15, no. 9, pp. 5119–5128, Sep. 2019.

[16] S. R. Saufi, Z. A. B. Ahmad, M. S. Leong, and M. H. Lim, "Gearbox fault diagnosis using a deep learning model with limited data sample," *IEEE Trans. Ind. Informat.*, vol. 16, no. 10, pp. 6263–6271, Dec. 2020.

[17] R. Liu, F. Wang, B. Yang, and S. J. Qin, "Multiscale kernel based residual convolutional neural network for motor fault diagnosis under nonstationary conditions," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 3797–3806, Jun. 2020.

[18] G. Hu, H. Li, Y. Xia, and L. Luo, "A deep Boltzmann machine and multi-grained scanning forest ensemble collaborative method and its application to industrial fault diagnosis," *Comput. Ind.*, vol. 100, pp. 287–296, Sep. 2018.

[19] X. Liu, Y. Tian, X. Lei, M. Liu, X. Wen, H. Huang, and H. Wang, "Deep forest based intelligent fault diagnosis of hydraulic turbine," *J. Mech. Sci. Technol.*, vol. 33, no. 5, pp. 2049–2058, May 2019.

[20] J. Ding, Q. Luo, L. Jia, and J. You, "Deep forest-based fault diagnosis method for chemical process," *Math. Problems Eng.*, vol. 2020, pp. 1–15, Jan. 2020.

[21] L. M. Elshenawy, T. A. Mahmoud, and C. Chakour, "Simultaneous fault detection and diagnosis using adaptive principal component analysis and multivariate contribution analysis," *Ind. Eng. Chem. Res.*, vol. 59, no. 47, pp. 20798–20815, Nov. 2020.

[22] Y. Du and D. Du, "Fault detection and diagnosis using empirical mode decomposition based principal component analysis," *Comput. Chem. Eng.*, vol. 115, pp. 1–21, Jul. 2018.

[23] J. H. Cho, J. M. Lee, S. W. Choi, D. Lee, and I. B. Lee, "Fault identification for process monitoring using kernel principal component analysis," *Chem. Eng. Sci.*, vol. 60, no. 1, pp. 279–288, 2005.

[24] R. Zass and A. Shashua, "Nonnegative sparse PCA," in *Proc. Adv. Neural Inf. Process. Syst.*, Jan. 2006, pp. 1561–1568.

[25] F. Chen and K. Rohe, "A new basis for sparse principal component analysis," 2020, *arXiv:2007.00596*.

[26] X. Hu, H. Zhang, D. Ma, and R. Wang, "Hierarchical pressure data recovery for pipeline network via generative adversarial networks," *IEEE Trans. Automat. Sci. Eng.*, vol. 19, no. 3, pp. 1960–1970, Jul. 2022.

[27] S. Zou, "Fault diagnosis of condenser based on principal component analysis & RBF neural network," M.S. thesis, School Energy Power Eng., Huazhong Univ. Sci. Technol., Hubei, China, 2014.

[28] J. E. Jackson, *A User's Guide to Principal Components*. New York, NY, USA: Wiley, 2005.

[29] S. Gajjar, M. Kulahci, and A. Palazoglu, "Real-time fault detection and diagnosis using sparse principal component analysis," *J. Process Control*, vol. 67, pp. 112–128, Jul. 2018.

[30] L. M. Elshenawy, T. A. Mahmoud, and C. Chakour, "Simultaneous fault detection and diagnosis using adaptive principal component analysis and multivariate contribution analysis," *Ind. Eng. Chem. Res.*, vol. 59, no. 47, pp. 20798–20815, Nov. 2020.

[31] F. T. Liu, M. T. Kai, Y. Yu, and Z. H. Zhou, "Spectrum of variable-random trees," *J. Artif. Intell. Res.*, vol. 32, no. 1, pp. 355–384, May 2008.

[32] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[33] Z. Jia, Z. Liu, Y. Gan, C.-M. Vong, and M. Pecht, "A deep forest-based fault diagnosis scheme for electronics-rich analog circuit systems," *IEEE Trans. Ind. Electron.*, vol. 68, no. 10, pp. 10087–10096, Oct. 2021.

YUANYUAN JU was born in Kunming, Yunnan, China, 1984. She received the B.S. degree in mathematics and applied mathematics and the M.S. degree in probability theory and mathematical statistics from Yunnan Normal University, Kunming, in 2007 and 2010, respectively, and the Ph.D. degree in applied statistics from Yunnan University, Kunming, in 2018.

ZILIANG CUI was born in Binzhou, China, in 1997. He received the B.S. degree from Binzhou University, Binzhou, in 2020. He is currently pursuing the M.S. degree with the Kunming University of Science and Technology, Kunming, China. His current research interests include image processing and deep learning.

QINGTAI XIAO was born in Heze, China, in 1989. He received the B.S. degree from Qufu Normal University, Qufu, China, in 2012, and the M.S. and Ph.D. degrees from the Kunming University of Science and Technology, Kunming, China, in 2016 and 2019, respectively. Since 2020, he has been a Distinguished Associate Professor with the Kunming University of Science and Technology. His research interests include energy and artificial intelligence.