


Article

Long Short-Term Memory Network-Based Normal Pattern Group for Fault Detection of Three-Shaft Marine Gas Turbine

Mingliang Bai , Jinfu Liu ^{*}, Yujia Ma, Xinyu Zhao, Zhenhua Long and Daren Yu

Harbin Institute of Technology, Harbin 150001, China; mingliangbai@outlook.com (M.B.); mayujiahit@outlook.com (Y.M.); 15B902036@hit.edu.cn (X.Z.); zhenhualong@hit.edu.cn (Z.L.); yudaren@hit.edu.cn (D.Y.)

^{*} Correspondence: jinfuliu@hit.edu.cn

Abstract: Fault detection and diagnosis can improve safety and reliability of gas turbines. Current studies on gas turbine fault detection and diagnosis mainly focus on the case of abundant fault samples. However, fault data are rare or even unavailable for gas turbines, especially newly-run gas turbines. Aiming to realize fault detection with only normal data, this paper proposes the concept of normal pattern group. A group of long-short term memory (LSTM) networks are first used for characterizing the mapping relationships among measurable parameters of healthy three-shaft gas turbines. Experiments show that the proposed method can detect all 13 common gas path faults of three-shaft gas turbines sensitively while remaining low false alarm rate. Comparison experiment with single normal pattern model verifies the necessities and superiorities of using normal pattern group. Meanwhile, comparison between LSTM network and other methods including support vector regression, single-layer feedforward neural network, extreme learning machine and Elman recurrent neural network verifies the superiorities of LSTM network in fault detection. Furthermore, comparison experiment with four common one-class classifiers further verifies the superiorities of the proposed method. This also indicates the superiorities of data-driven methods and gas turbine principle fusion to some extent.



Citation: Bai, M.; Liu, J.; Ma, Y.; Zhao, X.; Long, Z.; Yu, D. Long Short-Term Memory Network-Based Normal Pattern Group for Fault Detection of Three-Shaft Marine Gas Turbine. *Energies* **2021**, *14*, 13. <https://dx.doi.org/10.3390/en14010013>

Received: 1 December 2020

Accepted: 18 December 2020

Published: 22 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: fault detection and diagnosis; anomaly detection; three-shaft marine gas turbine; long short-term memory (LSTM) network; deep learning; normal pattern group

1. Introduction

Currently, prognostics and health management (PHM) technique of gas turbine has become a hot research topic for monitoring health condition as well as ensuring the safe and reliable operation [1–23]. PHM converts conventional “fail and fix” maintenance strategy to a more advanced conditional-based maintenance strategy. PHM can provide accurate condition monitoring, detect faults sensitively and timely, and thus avoid serious faults and significantly reduce maintenance costs.

With the boom of artificial intelligence and big data technique, data-driven methods-based intelligent PHM of gas turbines is becoming increasingly popular among various PHM methods. Many famous gas turbine companies are attempting to use artificial intelligence and big data techniques in gas turbines. Rolls-Royce company has proposed the concept of IntelligentEngine as the future development trend of gas turbine industry. Pratt and Whitney company has provided EngineWise service to provide intelligent health management and predicted maintenance for aeroengines. GE company has also established Predix platform for intelligent management of gas turbine. Through Predix platform, the health conditions of GE's gas turbines are continuously monitored to detect the need for real-time maintenance.

Data-driven fault detection and diagnosis methods extract knowledge from historical data and do not require accuracy nonlinear models [4,5]. Currently, data-driven methods are also becoming increasingly popular with researchers. Many data-driven

methods including Bayesian method [6,7], random forest [8], finite state machine [9], rough set [10–12], support vector machine [13], extreme learning machine [14], artificial neural network [15] etc., have been widely used in gas turbine fault detection and diagnosis. Mast et al. [16] proposed a Bayesian belief network-based fault diagnosis method for turbofan engines. Losi et al. [17] used Bayesian hierarchical models for gas turbine fault detection. Maragoudakis et al. [8] used random forest for fault identification of an industrial gas turbine. Li et al. [9] proposed a finite state machine-based method for fault diagnosis of a single-spool industrial gas turbine. Xu et al. [10] used fuzzy rough set for vibration fault diagnosis of aircraft engines. Wong et al. [18] used extreme learning machine for fault diagnosis of gas turbine generator systems. Fast et al. [19] used artificial neural networks for modelling and diagnosis of single-spool gas turbine-based combined heat and power plant. Orozco et al. [20] used a single-hidden layer feed-forward neural network for diagnosis of an externally fired gas turbine. Liu et al. [21] proposed a method for performance prediction of a heavy-duty gas turbine based on high dimensional model representation and artificial neural network. Wang et al. [22] used support vector machine and fuzzy c-means clustering for fault diagnosis of an industrial single-spool gas turbine. Zhou et al. [23] used support vector machine for gas turbine fault diagnosis. Loboda et al. [24] used multilayer perceptron and radial basis network for both industrial gas turbines and aircraft gas turbines. The experimental results indicate that radial basis network can obtain better performance. Loboda et al. [25] further proposed a probabilistic neural network-based method for fault diagnosis of industrial gas turbines and aircraft gas turbines and reported good detection performance. Yazdani and Montazeri-Gh [26] combined hybrid dimensionality reduction and fuzzy logic for fault diagnosis of two-shaft industrial SGT 600 gas turbine. Fentaye et al. [15] used nested artificial neural networks for gas-path fault identification of a two-shaft gas turbine. Tahan et al. [27] proposed a multiple networks artificial neural network model for an industrial 18.7-MW twin-shaft gas turbine engine. Lu et al. [28] proposed restricted-Boltzmann-based extreme learning machine for turbofan engine fault diagnosis.

Recently, deep learning [29] is enjoying a boom. Deep learning has achieved tremendous success in computer vision [30], natural language processing [31], autonomous cars [32] etc. Many researchers have begun to attempt deep learning in the field of industrial fault diagnosis [33]. Fu et al. [34] used grouped convolutional denoising autoencoders for aircraft engine fault detection and obtained good detection performance. Feng et al. [35] used information entropy and deep belief networks for aircraft turbofan engine fault diagnosis. Liu et al. [36] used convolutional neural network for fault detection of industrial gas turbines and obtained better performance than conventional artificial neural network and extreme learning machine. Mulewicz et al. [37] compared deep convolutional neural network with two conventional methods including random forest and extreme gradient boosting (XGBoost) and reported that deep convolutional neural network has better detection performance than the two conventional data-driven methods.

In the industrial scene, fault data are usually quite few or even available, especially for those gas turbines that have just been put into operation and only run for a short time. All above methods can obtain good performance when there are abundant historical fault data. However, in the case where fault data are unavailable, the above methods cannot realize fault detection due to the absence of fault information. This is the problems that this paper deals with.

Aiming to address the fault detection of three-shaft marine gas turbines in the case where only normal data are available at the beginning stage of operation, this paper proposed normal pattern group-based fault detection method for the first time. A group of long short-term memory (LSTM) networks are used for fault detection for the first time. The proposed method realizes accurate fault detection accuracy for fault data while remaining low false alarm rate for normal data, and thus effectively solves the problem of fault detection in three-shaft marine gas turbines in the case of no available historical fault data. The main contributions of this paper are summarized as follows.

- (1) Firstly, the concept of normal pattern group is proposed for three-shaft marine gas turbine fault diagnosis. Through normal pattern group, the intrinsic mapping relationships among measurable parameters of healthy three-shaft marine gas turbines are characterized by a group of normal pattern models.
- (2) Secondly, a group of long short-term memory (LSTM) networks are used in three-shaft marine gas turbine diagnosis. The superiorities of LSTM network in gas turbine fault detection are verified through comparison with other methods including support vector regression (SVR), extreme learning machine (ELM), single-hidden layer feed-forward neural network (SLFN) and Elman recurrent neural network (ERNN). To the best of our knowledge, this is the first time that LSTM network has been used in fault detection of three-shaft marine gas turbines and that the superiorities of LSTM network has been verified.
- (3) Thirdly, boxplot-based collaborative decision-making strategy for normal pattern group is proposed. Through collaborative decision-making of normal pattern group, accurate anomaly detection and low false alarm rate are realized. The normal pattern group is compared with single normal pattern models and common one-class classifiers and its superiorities are verified.

The rest of this paper is organized as follows. Section 2 elaborates the procedure of LSTM network-based normal pattern group fault detection method. Section 3 carries out detailed experiments to verify the superiorities of the proposed method. Section 4 concludes the paper and outlines the future research orientation.

2. Methods

2.1. Normal Pattern Group-Based Fault Detection

In industrial scene, the fault data of gas turbines, especially the gas turbines that have just been put into operation and only run for a short period, are quite rare or even unavailable. Current studies mainly focus on the case of abundant fault data. In the case of no available fault data, these methods cannot detect faults due to the lack of fault information. Thus, this paper will study the fault detection of three-shaft marine gas turbines in the case where only normal data are available.

The gas turbine follows the basic physical laws, such as the conservation of mass and energy etc. The gas turbine used Brayton cycle as the basic thermodynamic cycle. Thus, there exist inherent mapping relationships among all measurable parameters when the gas turbine operates normally. Thus, this paper establishes a series of normal pattern models to characterize intrinsic mapping relationships, proposes the concept of normal pattern group and detects anomaly through detecting the change of mapping relationships.

The normal pattern group is a group of normal pattern models. For a system with m input measurements (x_1, x_2, \dots, x_m) and n output measurements (y_1, y_2, \dots, y_n) . This paper establishes n normal pattern models with each model using one output measurement as its output and the rest $n - 1$ output measurements together with m inputs as its inputs. The architecture of normal pattern group is illustrated in Figure 1. Mathematically, the normal pattern group can be expressed as Equation (1).

$$\begin{cases} y_1 = f_1(x_1, x_2, \dots, x_m, y_2, y_3, \dots, y_n) \\ y_2 = f_2(x_1, x_2, \dots, x_m, y_1, y_3, \dots, y_n) \\ \dots \\ y_n = f_n(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_{n-1}) \end{cases} \quad (1)$$

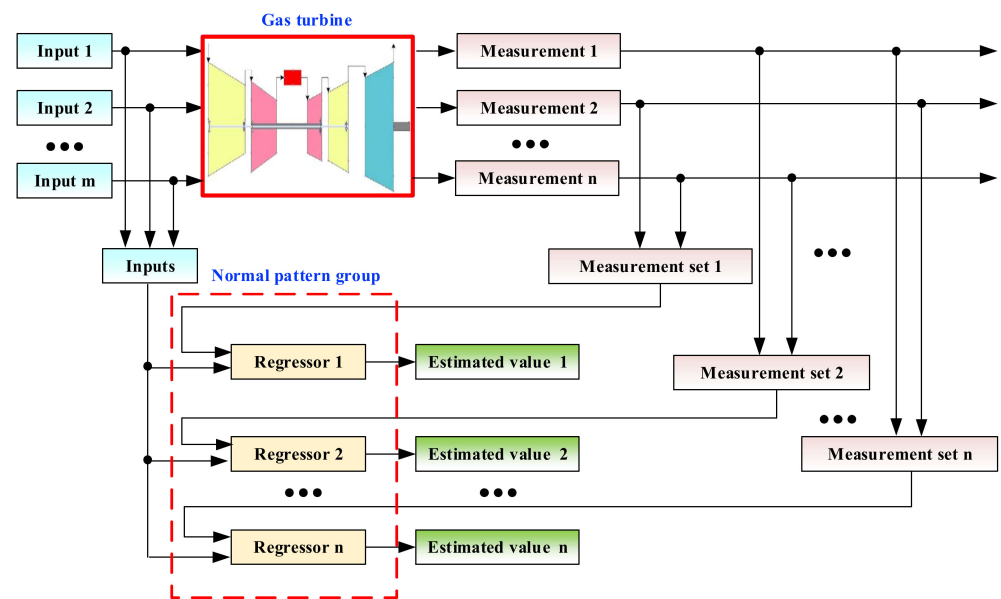


Figure 1. Architecture of normal pattern group.

In Equation (1), $F_i(\cdot)$ ($i = 1, 2, \dots, n$) are nonlinear functions that need to be identified using gas turbine normal data. $F_i(\cdot)$ ($i = 1, 2, \dots, n$) almost remains unchanged when three-shaft marine gas turbines are healthy. $F_i(\cdot)$ ($i = 1, 2, \dots, n$) will change once faults occur. Thus, accurate fault detection can be realized through normal pattern group defined in Equation (1).

2.2. Long Short-Term Memory Network

Normal pattern group method requires the identification of a group of normal pattern models, namely $F_i(\cdot)$ ($i = 1, 2, \dots, n$) in Equation (1), using normal historical data of gas turbines. The gas turbine a nonlinear dynamic system with many dynamic behaviors, such as rotor inertia, heat inertia and volume dynamics etc. These dynamic behaviors usually manifest as a delay of time. Artificial neural network (ANN) has strong ability to represent nonlinearity, and thus is used to identify $F_i(\cdot)$ ($i = 1, 2, \dots, n$) in Equation (1). Among various ANN methods, long short-term memory (LSTM) network [38] is one of the most effective methods to deal with dynamic data. LSTM network can successfully address the long-term dependency problem well and effectively deal with dynamic information through introducing forget gate, input gate and output gate. LSTM network has been successfully used in various fields, such as time series forecast [39–41], remaining useful life prediction of industrial machines [42], machine translation [43], named entity recognition [44] etc. LSTM has also been widely used in identification of various dynamic systems. Literature [45–47] used LSTM to identify various dynamic systems and reported that LSTM network can characterize the dynamic systems well and obtain much better performances than conventional methods. Therefore, this paper uses LSTM network for identifying the nonlinear mapping relationships $F_i(\cdot)$ ($i = 1, 2, \dots, n$) in Equation (1).

The structure of LSTM is shown in Figure 2, which includes three gates, namely forget gate, input gate and output gate [39]. The principle of LSTM is elaborated as follows.

- (1) Forget gate f_t represents the ratio of historical information to be remained.

$$f_t = \sigma(W_f(x_t, h_{t-1}) + \delta_f) \quad (2)$$

- (2) Input gate i_t represents the ratio of current information to be inputted.

$$i_t = \sigma(W_i(x_t, h_{t-1}) + \delta_i) \quad (3)$$

- (3) Cell state C_t represents the current hidden state, which is obtained by the weighted sum of current candidate value \tilde{C}_t and historical cell state C_{t-1} with forget gate f_t and input gate i_t being their weights respectively.

$$\tilde{C}_t = \tanh(W_C(x_t, h_{t-1}) + \delta_C) \quad (4)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_t \quad (5)$$

- (4) Output gate o_t is the ratio of information to be the output of current LSTM unit. Through the output gate, the current cell state C_t is converted to the current LSTM output.

$$o_t = \sigma(W_o(x_t, h_{t-1}) + \delta_o) \quad (6)$$

$$h_t = o_t \circ \tanh(C_t) \quad (7)$$

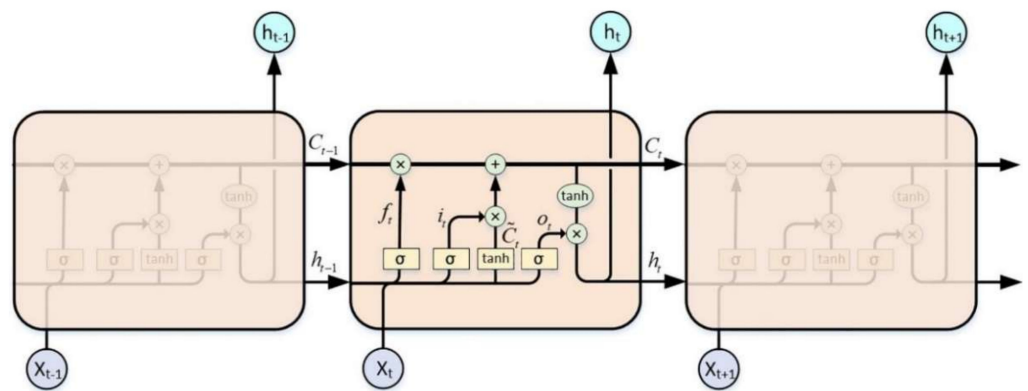


Figure 2. Structure of LSTM network [39].

In Equations (2)–(7), W_f , W_i , W_C , W_o are weight matrix, δ_f , δ_i , δ_C , δ_o are the bias term. The weight matrix and bias term of LSTM network are learned automatically from training data via the backpropagation through time (BPTT) strategy. The operation \circ is the element-wise product (also known as Hadamard product), x_t is the current input data and h_{t-1} is the LSTM unit output at the previous moment. The function $\sigma(\cdot)$ and $\tanh(\cdot)$ are nonlinear activation functions defined as follows.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (8)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (9)$$

Through LSTM network, the dynamic behaviors of three-shaft gas turbines can be effectively characterized and the nonlinear mapping relationships $F_i(\cdot)$ ($i = 1, 2, \dots, n$) in Equation (1) can be identified precisely.

2.3. Collaborative Decision for Fault Detection

After LSTM network training, normal pattern group is established. Then this section will apply boxplot to normal pattern group and design a collaborative decision strategy for fault detection.

Boxplot is a method for graphically depicting groups of numerical data through their quartiles. Its principle is shown in Figure 3, where the lower quartile Q_1 is usually the 25th percentile and the upper quartile Q_3 is usually the 75th percentile. Interquartile range (IQR) is the distance between the upper and lower quartiles and is computed by Equation (10). For a residual vector, boxplot gives the upper threshold u_{\max} and the lower threshold

u_{\min} by Equations (11) and (12) respectively. The data beyond the interval $[u_{\min}, u_{\max}]$ (the yellow solid circle in Figure 3) are regarded as outliers in boxplot.

$$IQR = Q_3 - Q_1 \tag{10}$$

$$u_{\min} = Q_1 - 1.5IQR \tag{11}$$

$$u_{\max} = Q_3 + 1.5IQR \tag{12}$$

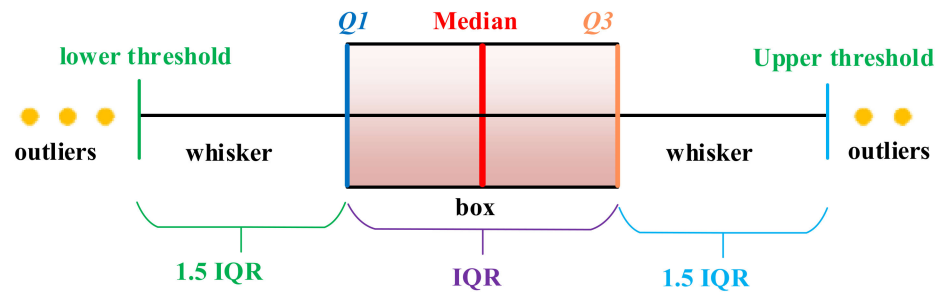


Figure 3. Principle of boxplot.

Normal pattern group includes n normal pattern models shown in Equation (1). Corresponding n residual vectors can also be obtained through the fitted values minus the corresponding real values. Each residual vector has an upper threshold and a lower threshold determined by boxplot. Let the number of samples in training set be N . Given a confidence interval CI , such as 95%, it is assumed that there are $|CI * N|$ samples that have m residuals beyond the boxplot threshold, where $|\cdot|$ is integer-valued function. Then a new instance will be normal with a confidence interval CI , if it has no more than m residual values beyond the boxplot threshold. Specifically, the fault detection process of a new sample is illustrated in Figure 4.

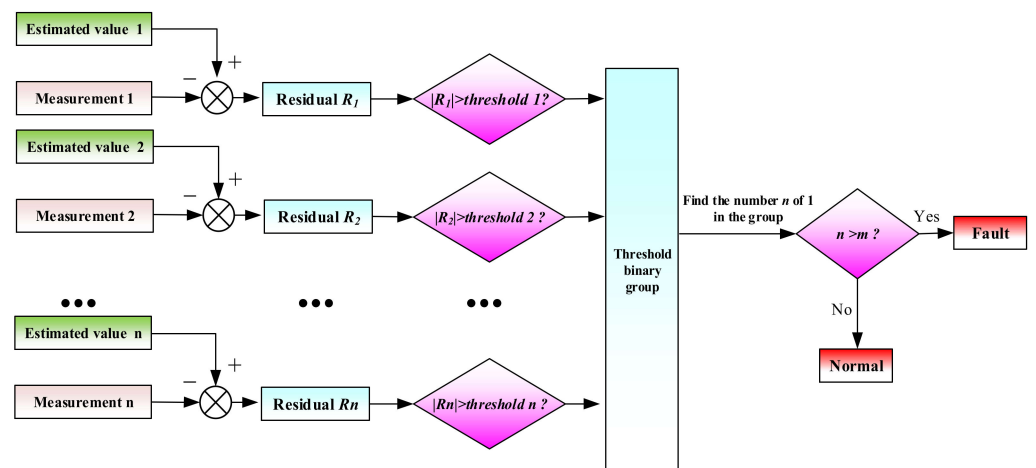


Figure 4. Collaborative decision-making strategy for fault detection.

In Figure 4, for a new instance, n fitted values are first computed by n trained LSTM networks, namely normal pattern group, and then n residual values are computed by n fitted values minus corresponding n actual values. The n residual values are compared with the corresponding n boxplot threshold to get the threshold binary group composed of n binary values (0 or 1). An example of threshold binary group is $0, 1, 0, 0, 0, \dots, 0$. If the number of 1 in the threshold binary group exceed m , then it is detected as a fault instance, otherwise it is detected as a normal instance.

2.4. Application in Three-Shaft Marine Gas Turbine Fault Detection

This section applies the proposed LSTM-based normal pattern group to fault detection of three-shaft marine gas turbines. Three-shaft marine gas turbine has two compressors, one combustion chamber (CC) and three turbines. Two compressors are low-pressure compressor (LPC) and high-pressure compressor (HPC). Three turbines are high-pressure turbine (HPT), low-pressure turbine (LPT) and power turbine (PT). Its typical configuration is shown in Figure 5.

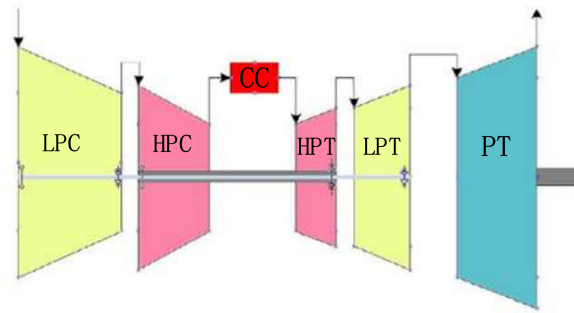


Figure 5. Typical configuration of a three-shaft marine gas turbine.

For the studied three-shaft marine gas turbines, there are 10 measurable parameters shown in Table 1 [48].

$$\begin{aligned}
 n_H &= F_1(t_1, g_f, n_L, P, plc, phc, plt, tlt, tpt) \\
 n_L &= F_2(t_1, g_f, n_H, P, plc, phc, plt, tlt, tpt) \\
 P &= F_3(t_1, g_f, n_H, n_L, plc, phc, plt, tlt, tpt) \\
 plc &= F_4(t_1, g_f, n_H, n_L, P, phc, plt, tlt, tpt) \\
 phc &= F_5(t_1, g_f, n_H, n_L, P, plc, plt, tlt, tpt) \\
 plt &= F_6(t_1, g_f, n_H, n_L, P, plc, phc, tlt, tpt) \\
 tlt &= F_7(t_1, g_f, n_H, n_L, P, plc, phc, plt, tpt) \\
 tpt &= F_8(t_1, g_f, n_H, n_L, P, plc, phc, plt, tpt)
 \end{aligned} \tag{13}$$

Table 1. Measurable parameters.

Description	Symbol
Ambient temperature	t_1
Fuel flow rate	g_f
Rotational speed of high-pressure spool	n_H
Rotational speed of low-pressure spool	n_L
Power of gas turbines	P
Outlet pressure of LPC	plc
Outlet pressure of HPC	phc
Outlet pressure of LPT	plt
Outlet temperature of LPT	tlt
Outlet temperature of PT	tpt

The ambient temperature t_1 and fuel flow rate g_f directly affect the operational state of gas turbines. Thus, the two parameters are regarded as the input measurements of three-shaft marine gas turbines. The change of the other eight measurable parameters listed in Table 1 are caused by the change of t_1 and g_f . Thus, the eight measurable parameters are regarded as the output measurements of three-shaft marine gas turbines. According to

Equation (1), we can establish the normal pattern group of three-shaft marine gas turbines in Equation (13). In Equation (13), there are eight normal pattern models, namely n equals 8. The detailed procedure of normal pattern group-based gas turbine fault detection is illustrated in Figure 6, which includes the following three steps.

Step 1: data preprocessing. This step divides the normal data into three parts. The first 70% is training set, the following 15% is the validation set and the rest 15% is the test set.

Step 2: training and validation. This step trains eight LSTM networks using training set to identify eight nonlinear mapping relationships in Equation (13). Hyperparameters of eight LSTM networks are tuned through validation set. After eight LSTM networks are trained, detection thresholds are computed through collaborative decision strategy in Section 2.3.

Step 3: test and anomaly detection. A new sample is inputted to the trained LSTM-based normal pattern group, eight residual values are obtained. Then its health condition is determined through collaborative decision strategy in Section 2.3.

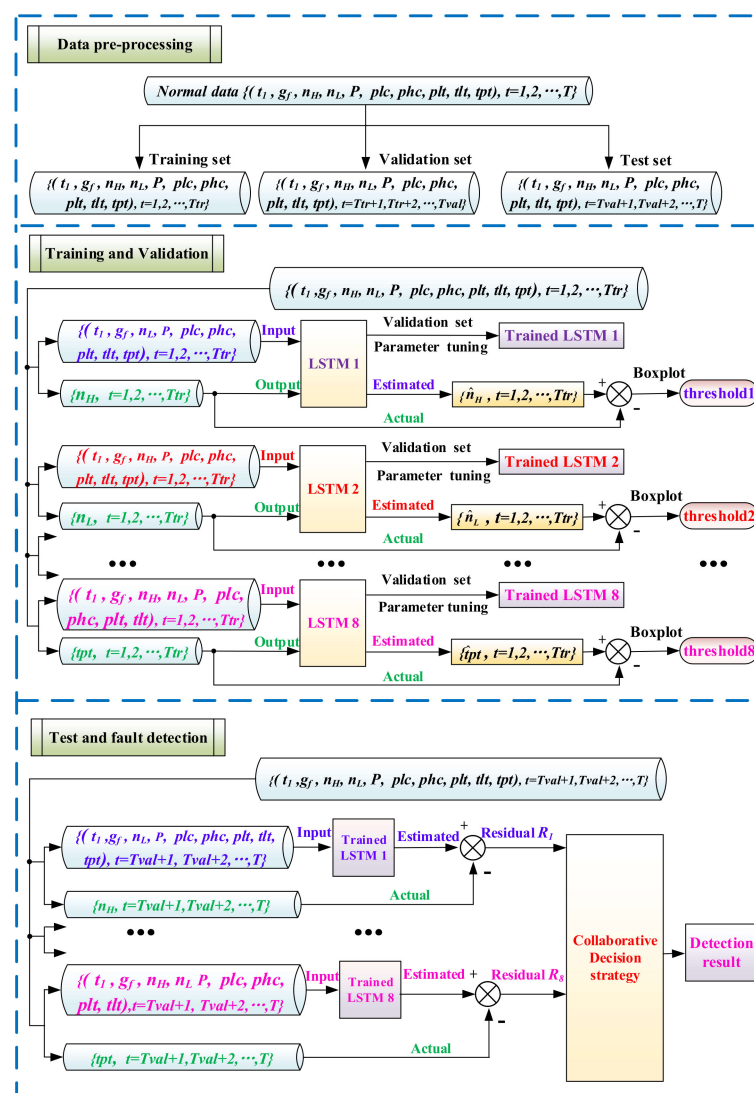


Figure 6. Technological process of this paper.

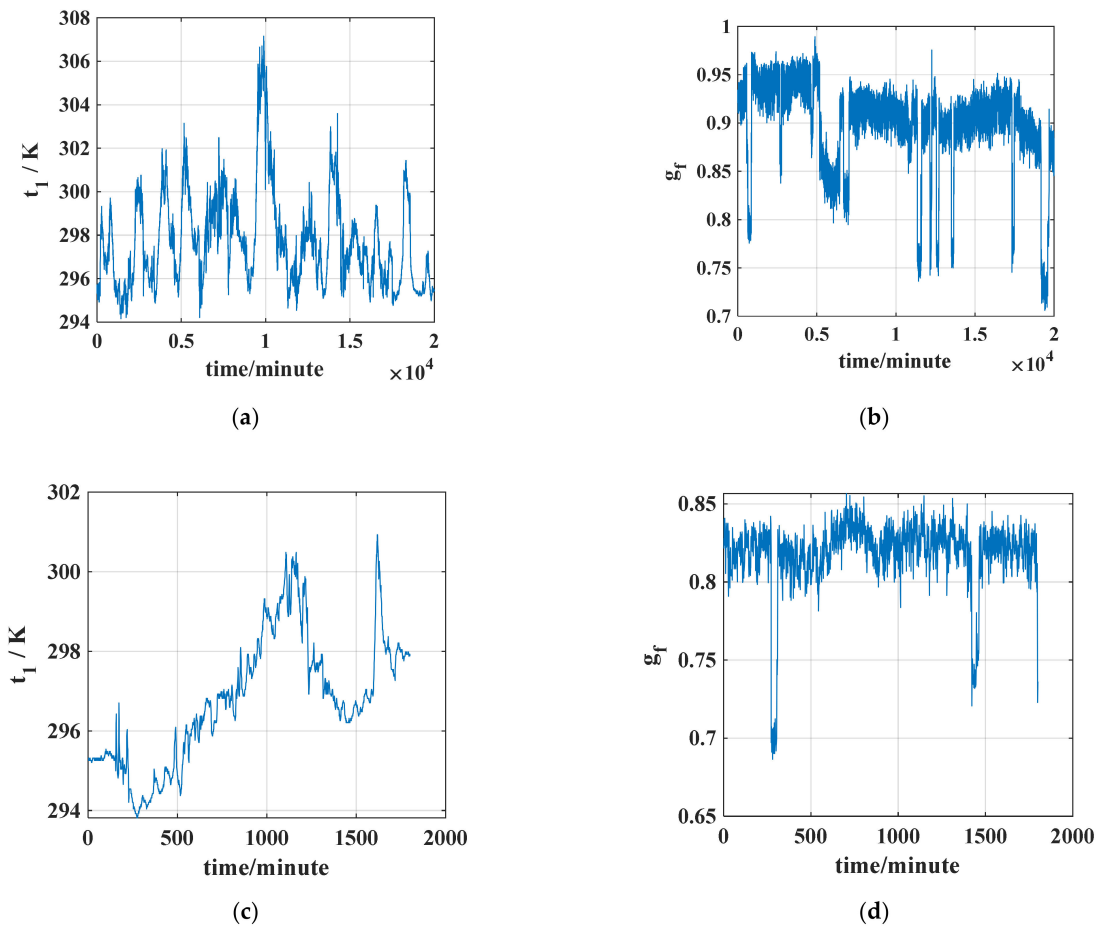


Figure 7. Input parameters of normal data and fault data: (a) Ambient temperature of normal data; (b) Fuel flow rate of normal data; (c) Ambient temperature of fault data; (d) Fuel flow rate of fault data.

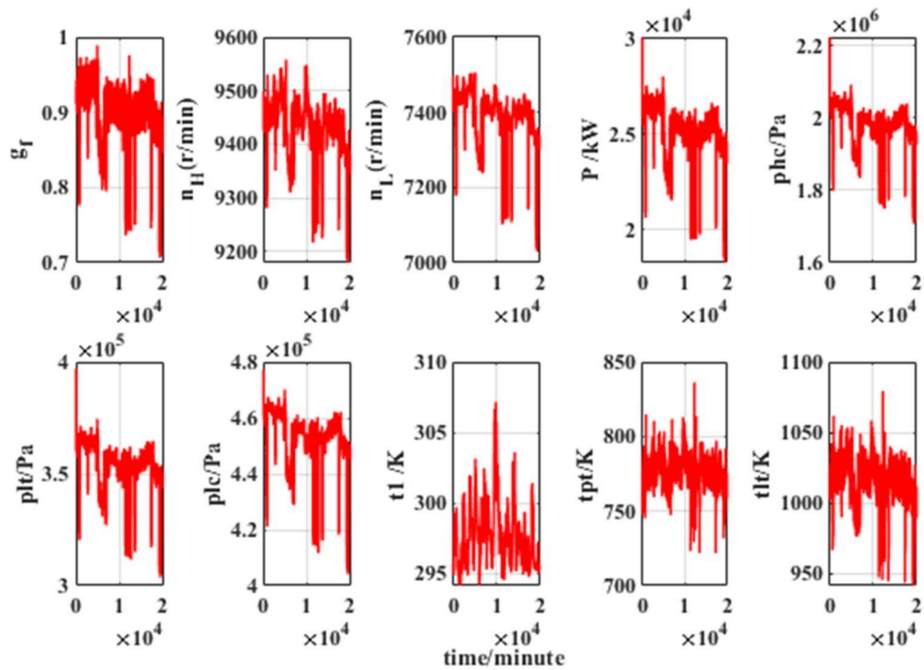


Figure 8. Normal data in the experiments.

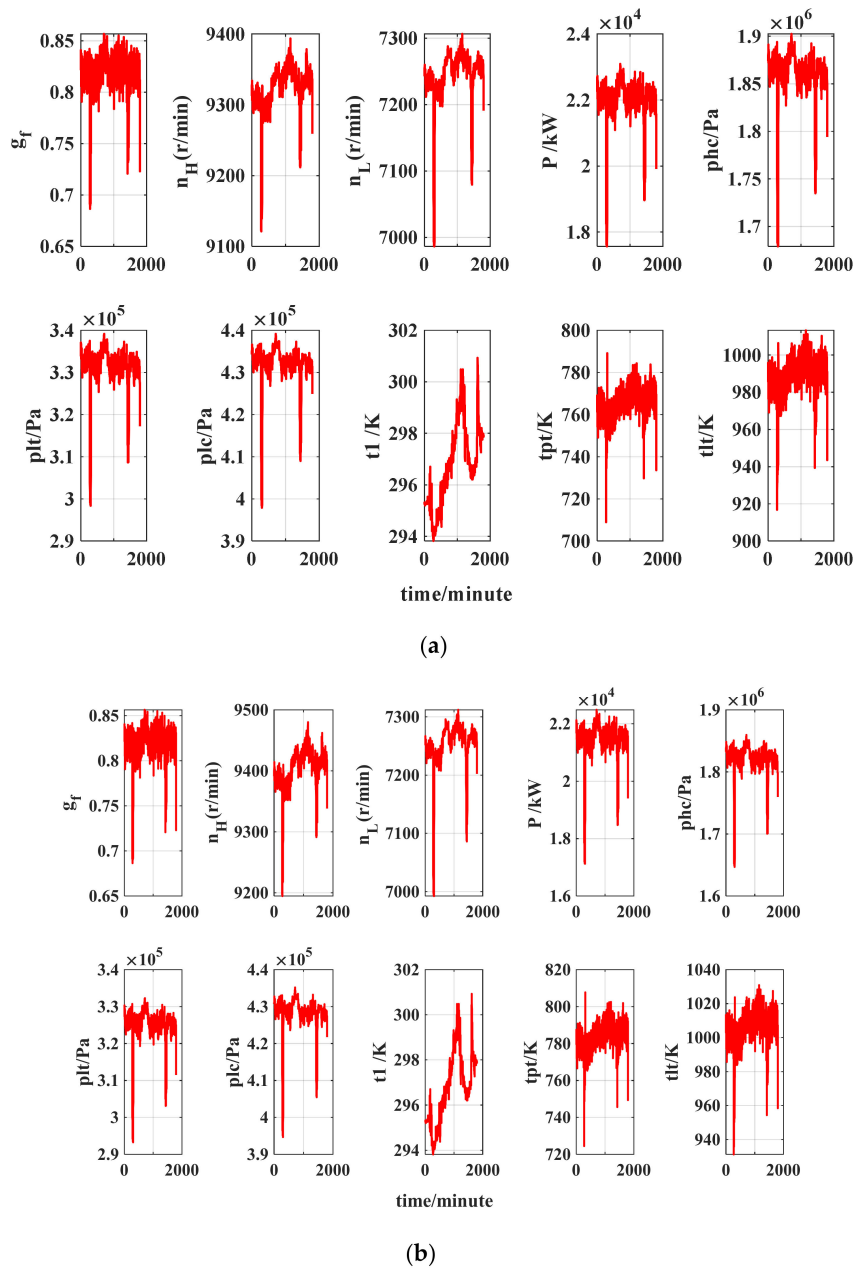


Figure 9. LPC fouling fault data in the experiments: (a) LPC fouling fault severity 1; (b) LPC fouling fault severity 5.

To evaluate the fault detection performance, the detection accuracy of normal data acc_{normal} and the detection accuracy of fault data $acc_{abnormal}$ are defined as follows.

$$acc_{normal} = \frac{1}{n_1} \sum_{t=1}^{n_1} I_t, \quad acc_{abnormal} = \frac{1}{n_2} \sum_{t=1}^{n_2} J_t, \quad (14)$$

where n_1 and n_2 are the number of normal data in test set and fault data respectively, I_t is the number of actual normal data that are detected as normal data, J_t is the number of actual fault data that are detected as fault data.

3.2. Experiment of LSTM Network-Based Normal Pattern Group

This section performed experiment of normal pattern group to verify its effectiveness. First, eight LSTM networks are trained using training data to identify eight normal pattern models in Equation (13). LSTM networks are implemented by Keras library of Python

programming language. The identification of eight normal pattern models is a typical regression task. Mean squared error is the most common loss function for regressor tasks in neural networks, and thus mean squared error is used as the loss function of LSTM network. During the training process, the validation set is used to tune the hyperparameters of LSTM networks. The fitted results are shown in Figures 10 and 11. It is observed that the fitted values are quite close to the actual values. This shows that LSTM network can characterize the normal pattern of gas turbines well.

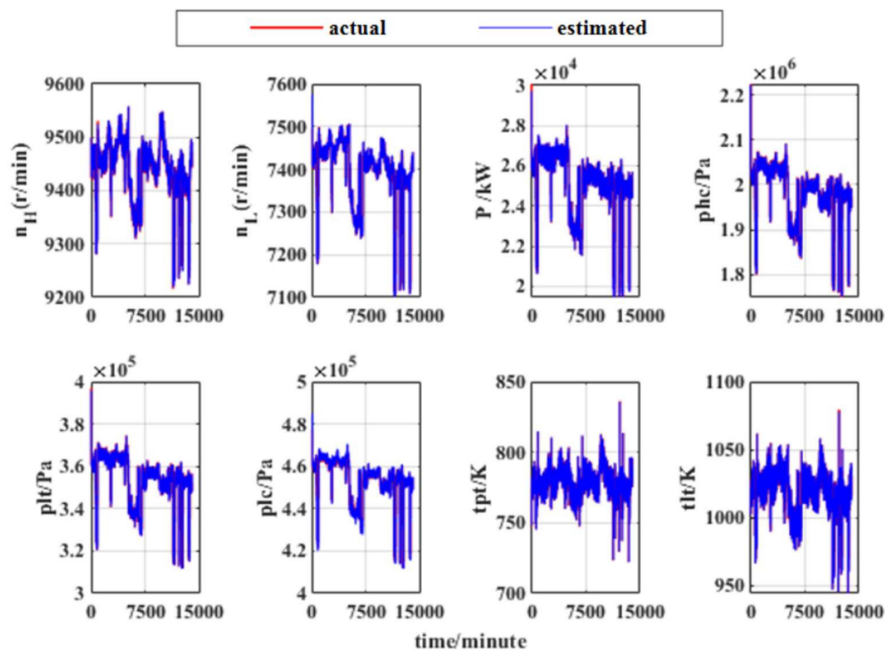


Figure 10. Actual data versus estimated data in training set.

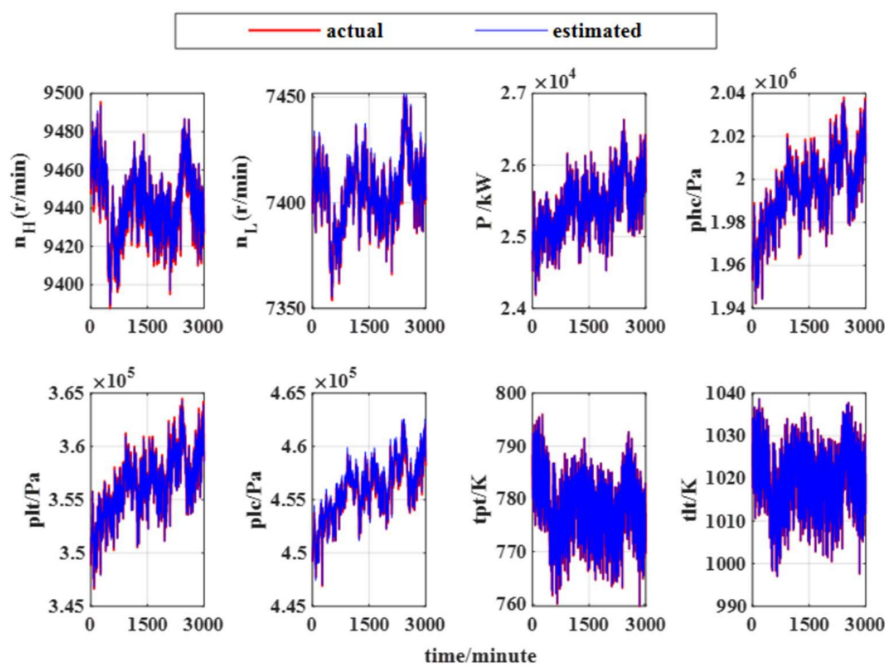


Figure 11. Actual data versus estimated data in validation set.

After network training and validation, eight residual vectors of normal pattern group are computed through the fitted values minus corresponding actual values. Boxplot is used to determine the upper threshold and the lower threshold of the eight residual vectors. The normalized threshold of each boxplot is shown in Figure 12 and corresponding detection threshold is listed in Table 3.

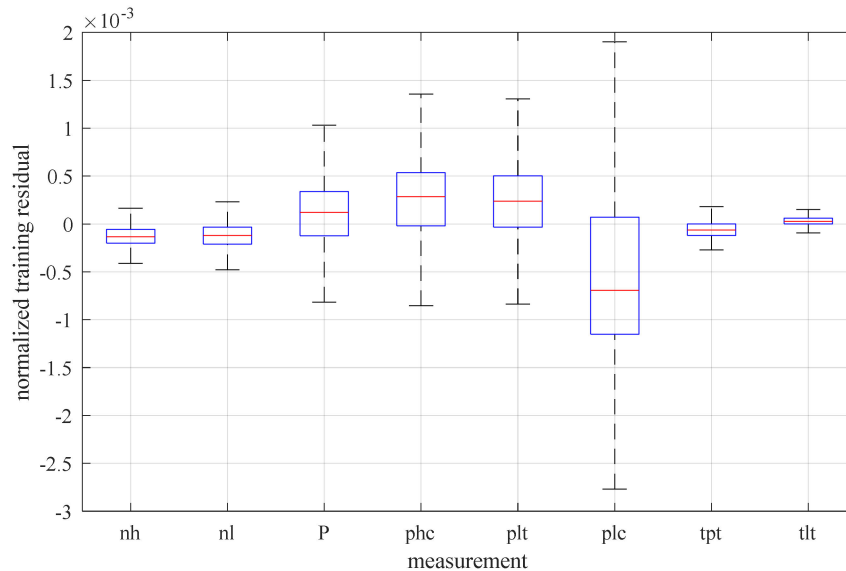


Figure 12. Boxplot of residuals in training set.

Table 3. Detection threshold of normal pattern group.

	<i>n_H</i>	<i>n_L</i>	<i>P</i>	<i>phc</i>	<i>plt</i>	<i>plc</i>	<i>tpt</i>	<i>tlt</i>
Upper limit	−3.9774	−3.6015	−23.8765	−1850.8100	−325.4210	−1443.6000	−0.2329	−0.0944
Lower limit	1.5426	1.7531	30.1565	2964.3100	507.2070	920.9380	0.1400	0.1559

In the trained normal pattern group, each training instance has eight residual values. According to the boxplot threshold of training set, we can count the number of samples that has no residual values beyond the corresponding boxplot threshold. Similarly, we can count the number of samples that has $z(z = 1, 2, \dots, 8)$ residual values beyond the corresponding boxplot threshold. Furthermore, the percentage of these samples is computed through being divided by the number of all samples in training set, which is shown in Figure 13.

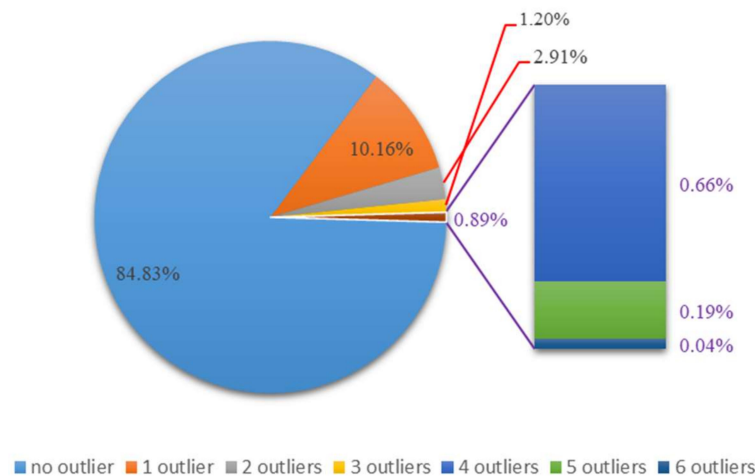


Figure 13. Percentage of different threshold overshoot numbers in training set.

It is observed from Figure 13 that the percentage of threshold overshoot number 0 and threshold overshoot number 1 are the largest. The threshold overshoot number of as many as 94.99% training samples is no more than 1. After that, although the threshold overshoot number increases, the percentage does not increase much and the ability to detect fault samples can decrease significantly. Therefore, we set the parameter m in Figure 4 to be 1, which can ensure about 95% training samples to be classified correctly. For a new instance, it is first inputted to the trained normal pattern group to obtain eight residual values. If more than one of the eight residual values exceed corresponding boxplot threshold, this instance is detected as a fault instance, otherwise it is detected as a normal instance.

After establishing LSTM-based normal pattern group and determining the fault detection strategy, test set of normal data and 13 categories of fault data are used for fault detection. First, the test set of normal data are used to evaluate the fault detection performance of normal data. The fitting results and residuals of test data of normal data are shown in Figure 14. To evaluate the fitting performance better, mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE) are used. Their definitions are given in Equations (15)–(17). MAE describes the mean fitting errors and RMSE is sensitive to extreme fitting errors and MAPE describes the mean percentage error. For all three metrics, the smaller the better. These three describe the fitting performance from different perspectives. MAE, MAPE and RMSE of training set, validation set and test set are shown in Table 4.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2} \quad (15)$$

$$MAE = \sqrt{\frac{1}{n} \sum_{i=1}^n |y - \hat{y}|} \quad (16)$$

$$MAPE = \sqrt{\frac{1}{n} \sum_{i=1}^n (|y - \hat{y}| / |y|)} \times 100\% \quad (17)$$

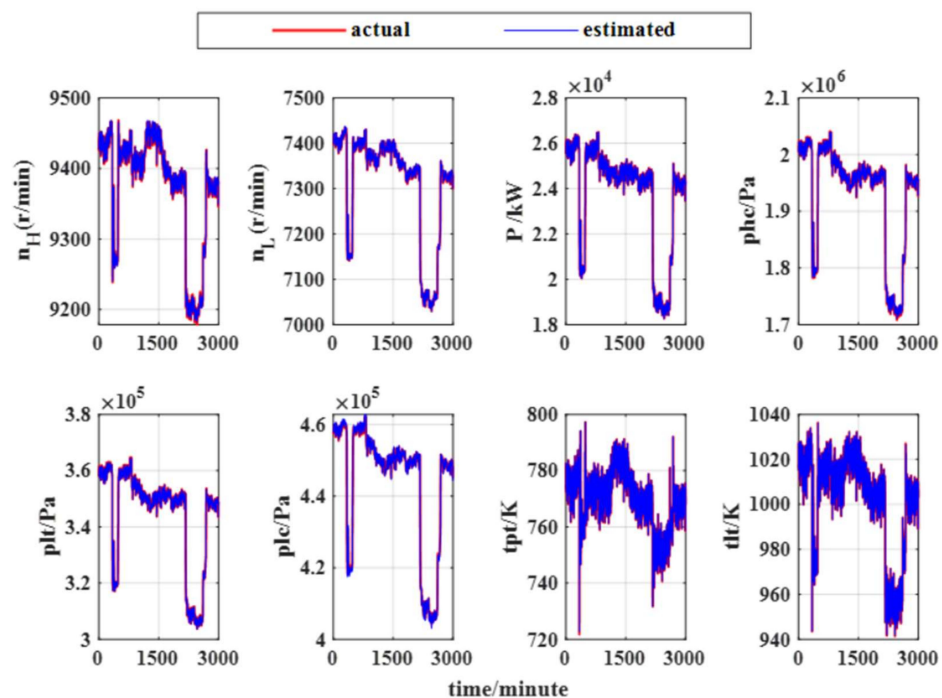


Figure 14. Actual data versus estimated data in test set.

Table 4. RMSE, MAE and MAPE of normal pattern group.

		n_H	n_L	P	phc	plt	plc	tpt	tlt
RMSE	train	1.6394	1.5979	12.5409	1149.7546	183.9000	516.1266	0.0844	0.0669
	validation	1.5571	1.4210	10.9332	751.4320	161.6139	387.0732	0.0752	0.0490
	test	1.6005	1.6448	12.2254	1053.9267	163.5448	374.5660	0.0962	0.0790
MAE	train	1.4006	1.1706	9.2172	972.4260	147.8990	449.2510	0.0698	0.0488
	validation	1.3595	1.1526	8.3614	626.3500	133.0760	349.0990	0.0676	0.0384
	test	1.3902	1.4262	9.8911	823.7550	135.3260	287.7390	0.0844	0.0667
MAPE	train	0.0149%	0.0158%	0.0371%	0.0493%	0.0414%	0.0992%	0.0090%	0.0048%
	validation	0.0144%	0.0156%	0.0329%	0.0313%	0.0372%	0.0765%	0.0087%	0.0038%
	test	0.0148%	0.0195%	0.0428%	0.0440%	0.0392%	0.0654%	0.0110%	0.0067%

From Figure 14 and Table 4, it is observed that the fitted values are close to the actual values in test set. RMSE, MAE and MAPE are all very small, which means that LSTM network can fit the normal data of three-shaft marine gas turbine well. Next, collaborative fault detection strategy is used for fault detection. Fault detection accuracy of LSTM-based normal pattern group in normal data and fault data is summarized in Table 5.

Table 5. Fault detection accuracy of the proposed normal pattern group method.

	Normal Data			LPC Fault		HPC Fault				
	Train	Validation	Test	Fouling	FOD	Fouling	FOD			
Proposed Method	0.9499	0.9867	0.9383	1.0000	1.0000	1.0000	1.0000			
	HPT Fault			LPT Fault			PT Fault			
	Fouling	Erosion	FOD	Fouling	Erosion	FOD	Fouling	Erosion	FOD	
Proposed Method	1.0000	1.0000	0.9936	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	

The results in Table 5 show that the proposed method can detect all 13 categories of faults sensitively and remain low false alarm rate for normal data. Thus, through the proposed LSTM network-based normal pattern group and designed collaborative decision-making strategy, faults can be sensitively detected and the robustness to normal data is maintained simultaneously.

3.3. Comparison with Single Normal Pattern Methods

The proposed normal pattern group is a combination of eight normal pattern models in Equation (13). To verify the necessities and superiorities of normal pattern group, this section compared it with eight single normal pattern models in Equation (13). Comparison results are shown in Table 6, n_H , n_L , P , phc , plt , plc , tpt and tlt denote the normal pattern model that uses n_H , n_L , P , phc , plt , plc , tpt and tlt as the output of LSTM network respectively. The bold values denote the best detection accuracy in Table 6.

From Table 6, it is observed that none of the eight normal pattern models can detect all 13 faults sensitively. The eight normal pattern models obtain the accuracy of less than 0.8 for some categories of faults. For example, n_H normal pattern model, n_L normal pattern model and plc normal pattern model both obtain very bad accuracy (accuracy less than 0.6) for HPT erosion fault. P normal pattern model and plt normal pattern model both obtain accuracy less than 0.6 for HPT FOD fault. The proposed normal pattern group method effectively improves the detection performance of fault data while remaining low false alarm rate for normal data through the collaborative decision of normal pattern group. Thus, the proposed normal pattern group significantly improves the fault detection performance compared with the eight normal pattern methods.

Table 6. Fault detection accuracy comparison with single normal pattern method.

Method	Normal Data			LPC Fault		HPC Fault	
	Train	Validation	Test	Fouling	FOD	Fouling	FOD
n_H	0.9630	0.9990	0.9253	0.9967	0.9991	0.9998	0.9997
n_L	0.9730	0.9857	0.9897	0.9997	0.7469	0.9977	0.6737
P	0.9692	0.9840	0.9950	0.9993	0.9996	1.0000	0.9999
phc	0.9605	0.9997	0.9230	0.9999	1.0000	1.0000	1.0000
plt	0.9876	0.9970	0.9977	0.9940	0.9874	1.0000	0.9926
plc	0.9821	1.0000	0.9800	0.7523	0.8147	0.9857	0.8328
tpt	0.9648	0.9997	0.9953	0.8624	0.3632	0.8513	0.4309
tlc	0.9654	0.9897	0.9550	0.9674	0.7670	0.9994	0.7251
Proposed Method	0.9499	0.9867	0.9383	1.0000	1.0000	1.0000	1.0000

Method	HPT Fault			LPT Fault			PT Fault		
	Fouling	Erosion	FOD	Fouling	Erosion	FOD	Fouling	Erosion	FOD
n_H	0.9998	0.5959	0.8703	0.9996	0.9994	0.8328	0.9423	0.5896	0.9999
n_L	0.9977	0.0828	0.9941	0.9976	0.7324	0.9992	0.9998	0.9998	1.0000
P	1.0000	0.9996	0.5136	0.9997	0.7533	0.6366	0.4061	0.9997	1.0000
phc	1.0000	1.0000	0.9841	1.0000	0.7742	0.9997	0.9978	0.9958	0.9966
plt	1.0000	0.9787	0.5734	0.9974	0.1662	0.6247	1.0000	1.0000	0.8572
plc	0.9857	0.3821	0.8039	0.4059	0.7179	0.9138	0.9624	0.9999	0.9973
tpt	0.8513	0.9967	0.6548	0.9968	0.8373	0.4809	0.8447	1.0000	0.9999
tlc	0.9994	0.9990	0.8954	0.9652	0.7858	0.9904	0.9993	1.0000	0.9690
Proposed Method	1.0000	1.0000	0.9936	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

3.4. Comparison between LSTM Network and Other Methods

To verify the superiorities of LSTM network, this section compares LSTM network with some widely used regressors. The compared methods include support vector regression (SVR) [57], single-layer feedforward neural network (SLFN) [58], extreme learning machine (ELM) [59] and Elman recurrent neural network (ERNN) [60]. All of these methods are widely used in pattern recognition, regression, time series forecast, industrial fault diagnosis, etc. [61–63].

SVR uses kernel method to map the original data to a high dimensional space, so that an approximately linear regression can be used for regression in this space. Radial basis function (RBF) kernel is the most common kernel function. SLFN, ELM and ERNN are three kinds of neural networks. SLFN is a static neural network with three layers, namely input layer, hidden layer and output layer. SLFN is usually trained through backpropagation strategy. Its structure is shown in Figure 15. ELM has the same structure as SLFN. ELM random generates weights and bias between input layer and hidden layer and determines the weights of output layer by computing Moore–Penrose generalized inverse matrix instead of iteratively learning through error backpropagation. ELM can be trained faster than SLFN. ERNN introduces a one-step time delay to characterize the dynamic behaviors and its structure is shown in Figure 16. ERNN is also trained through error backpropagation.

In this paper, these methods are used to identify the normal pattern group in Equation (13). They are trained using data from training set, and their hyperparameters are tuned through validation set. Their fault detection performances are shown in Table 7.

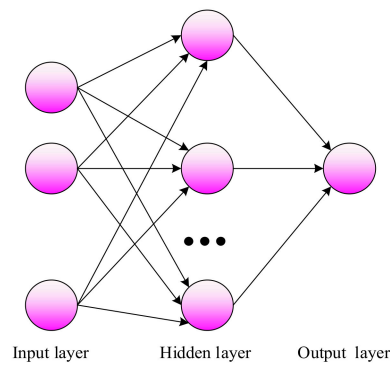


Figure 15. Structure of SLFN [63].

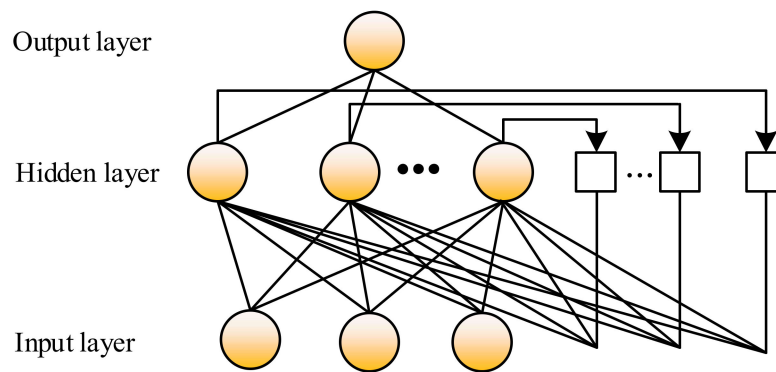


Figure 16. Structure of ERNN [63].

Table 7. Fault detection accuracy comparison with other methods.

Method	Normal Data			LPC Fault		HPC Fault			
	Train	Validation	Test	Fouling	FOD	Fouling	FOD		
ELM	0.9295	0.9853	0.8787	1.0000	1.0000	1.0000	1.0000		
ERNN	0.9631	0.9963	0.9213	0.9989	1.0000	1.0000	1.0000		
SLFN	0.9336	0.9943	0.9000	1.0000	1.0000	1.0000	1.0000		
SVR	0.9292	1.0000	0.7850	0.6741	0.9806	0.7998	0.8739		
Proposed Method	0.9499	0.9867	0.9383	1.0000	1.0000	1.0000	1.0000		
Method	HPT Fault			LPT Fault			PT Fault		
	Fouling	Erosion	FOD	Fouling	Erosion	FOD	Fouling	Erosion	FOD
ELM	1.0000	1.0000	1.0000	1.0000	0.9946	1.0000	1.0000	1.0000	1.0000
ERNN	0.9869	0.9987	0.9999	1.0000	1.0000	1.0000	1.0000	0.9966	1.0000
SLFN	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
SVR	0.6688	0.7628	0.8176	0.7968	1.0000	0.9032	0.9441	0.9067	1.0000
Proposed Method	1.0000	1.0000	0.9936	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

From Table 7, it is observed that the proposed LSTM network-based fault detection method significantly outperforms other methods. For test set of normal data, LSTM improves the accuracy by 0.0596 when compared to ELM, improves the accuracy by 0.1533 when compared to SVR, improves the accuracy by 0.0383 when compared to SLFN and improves the accuracy by 0.0170 when compared to ERNN. For the fault data, LSTM can ensure the fault detection accuracy of each fault class to be at least 0.9936. By contrast, ELM, SLFN and ERNN obtain almost as high accuracy as LSTM, but SVR obtains the accuracy of less than 0.9 for some categories of faults including LPT fouling fault, HPT

fouling fault, HPT erosion fault, LPC fouling fault and HPC fouling fault. Thus, LSTM can ensure the fault detection accuracy of normal data and fault data to be more than 0.9 and have more reliable fault detection performance. It is also observed that ELM, ERNN and SLFN outperforms the SVR method in the test set of normal data and some types of faults. This shows that neural network-based methods including ELM, ERNN and SLFN has better fault detection performance. Meanwhile, ERNN outperforms SLFN and ELM. This is because that ERNN considers the time-delayed relationship among gas turbine measurements to some extent. Compared with ERNN, LSTM considers time-delayed relationship better through introducing input gate, forget gate and output gate, and can characterize time-delayed relationship with much longer time lags. Thus, the proposed LSTM-based method obtains significantly better fault detection performances than ERNN and other methods in Table 7.

3.5. Comparison with One-Class Classifiers

Currently, one-class classifiers in machine learning field have also been widely used for industrial anomaly detection in the case of only requires normal data. These methods have been widely used in industrial fault detection [64,65], spam detection [66], etc. Thus, this paper compared the proposed normal pattern group method with these common one-classifiers to further verify its supercities. The compared methods include one-class support vector machine (OCSVM) [67,68], local outlier factor (LOF) [69], isolation forest [70], principal component analysis (PCA) [71].

OCSVM uses the kernel function to map the original normal data to a high-dimensional space, where OCSVM tries to find a hyperplane that enables the normal data can be as far from the origin as possible. Let the distance between the hyperplane and the origin be ρ , then the samples whose distance from the origin is smaller than ρ is detected as abnormal samples. Common kernel functions include RBF kernel, linear kernel, sigmoid kernel etc., and RBF kernel is the most widely used one. LOF detects anomaly through comparing the density of the given sample and the sample density in its neighborhood. If its density is obviously smaller than the density in its neighborhood, then this sample is detected as an abnormal sample. Isolation forest isolates fault samples through constructing trees and abnormal samples are usually isolated first. PCA detects anomaly through the compression and reconstruction of data. PCA is trained using normal data, and it can ensure that the reconstruction errors of normal samples are small and that the reconstruction errors of fault samples are large. Square prediction error (SPE) and T^2 statistics [71] are two common ways for determining thresholds in PCA-based fault detection method.

In this paper, OCSVM, LOF and isolation forest were implemented by scikit-learn library [72,73] of Python programming language. PCA-based fault detection method was coded through Numpy library [74] of Python programming language. This paper uses three kernel functions including radial basis function (RBF) kernel, linear kernel and sigmoid kernel for OCSVM method. For PCA-based fault detection method, square prediction error (SPE) and T^2 statistics [71] are both used in the experiment. The parameters of these methods were selected by the validation set. Corresponding comparison results are shown in Table 8.

Table 8. Fault detection accuracy comparison with one-class classifiers.

Method	Normal Data			LPC Fault			HPC Fault		
	Train	Validation	Test	Fouling	FOD	Fouling	FOD		
Isolation Forest	0.9310	0.9993	0.7167	0.7914	0.9482	0.8731	0.9948		
LOF	0.9916	0.9997	0.8277	0.6670	0.5676	0.7564	0.6839		
PCA(SPE)	0.9893	0.9743	0.9738	0.5632	0.8501	0.7119	0.9637		
PCA(T^2)	0.9714	0.8052	0.8262	0.0700	0.3461	0.0778	0.5513		
OCSVM(RBF)	0.9892	0.9993	0.7810	0.7318	0.8498	0.8608	0.8864		
OCSVM (linear)	0.9501	1.0000	0.7873	0.1131	0.1066	0.0874	0.0838		
OCSVM (Sigmoid)	0.9497	1.0000	0.7877	0.1249	0.1200	0.0944	0.1359		
Proposed Method	0.9499	0.9867	0.9383	1.0000	1.0000	1.0000	1.0000		

Method	HPT Fault			LPT Fault			PT Fault		
	Fouling	Erosion	FOD	Fouling	Erosion	FOD	Fouling	Erosion	FOD
Isolation Forest	0.8918	0.9750	0.9763	0.9616	0.9999	0.9998	0.9560	0.8981	1.0000
LOF	0.8053	0.9650	0.9451	0.9301	0.9240	0.9842	0.4769	0.8046	1.0000
PCA(SPE)	0.7458	0.8801	0.9033	0.8807	0.9993	0.9856	0.8502	0.7102	1.0000
PCA(T^2)	0.1031	0.1519	0.2891	0.1276	0.1090	0.6497	0.1150	0.1871	1.0000
OCSVM (RBF)	0.8459	0.9927	0.9692	0.9648	1.0000	0.9843	0.8591	0.8587	1.0000
OCSVM (linear)	0.1247	0.2469	0.1457	0.1102	0.2203	0.1074	0.2004	0.1978	0.0486
OCSVM (Sigmoid)	0.1103	0.1923	0.1007	0.0900	0.3264	0.5196	0.1716	0.2766	0.9996
proposed method	1.0000	1.0000	0.9936	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

From Table 8, it is observed that the four one-class classifiers are not sensitive to some categories of faults. Isolation Forest, PCA, LOF and OCSVM all have bad performance (accuracy less than 0.8) for some fault categories. Meanwhile, isolation Forest, LOF, OCSVM and PCA with T^2 statistics obtains the accuracy of only about 0.8 for test set of normal data. Among these one-class classifiers, only PCA with SPE statistics obtains good accuracy for test set of normal data. The proposed method can ensure the detection accuracy of each fault class to be at least 0.9936 while remaining the accuracy of more than 0.9 for normal data. The proposed method incorporates the gas turbine prior knowledge, and thus is sensitive to all faults and remains low false alarm rate for normal data. Thus, the proposed method significantly outperforms common one-class classifiers in fault detection performance of three-shaft marine gas turbines. This can also indicate that incorporating gas turbine prior knowledge can improve the gas turbine fault detection performance to some extent.

4. Conclusions and Future Work

Fault detection of three-shaft marine gas turbines has great significance in increasing operational reliability and reducing maintenance costs. Current researches mainly focus on the situation where abundant fault data are available. However, fault data are quite few or even unavailable, especially for newly-run gas turbines. Aiming at the case where only normal data are available, this paper proposes long short-term memory (LSTM) network-based normal pattern group for fault detection of three-shaft gas turbines. Through experiments in a three-shaft marine gas turbine, the following conclusions can be drawn.

Firstly, this paper characterizes the healthy state of three-shaft marine gas turbines using normal pattern group composed of a group of normal pattern models. A group of long short-term memory (LSTM) networks are used to identify these normal pattern models and detect anomalies. Experimental results show that the proposed method can detect all 13 common gas path faults of three-shaft gas turbines sensitively while remaining low false alarm rate simultaneously.

Secondly, the proposed normal pattern group method is compared with eight single normal pattern models to verify its superiorities. Experimental results show that the proposed method significantly outperforms the eight normal pattern models in terms of fault detection performance.

Thirdly, the proposed normal pattern group method is compared with some common one-class classifiers including one-class support vector machine, principal component analysis, isolation forest and local outlier factor to further verify its superiorities. Experimental results show that the proposed method significantly outperforms all one-class classifiers to some extent. This can also indicate that introducing appropriate prior knowledge can improve the fault detection performance of gas turbines compared with purely data-driven one-class classifiers to some extent.

In the future, the proposed normal pattern group method can be applied in other types of gas turbines after analyzing the mapping relationships among corresponding measurement parameters. Besides, more data-driven methods will also be explored in fault detection of gas turbines. Additionally, the authors hope that LSTM network-based normal pattern group can be applied to fault detection of other industrial systems except gas turbines, such as diesel engines, steam turbines, nuclear power plants, wind turbines, chillers, pumps, photovoltaic arrays etc.

Author Contributions: Conceptualization, M.B. and J.L.; Methodology, M.B., Y.M., X.Z. and Z.L.; Supervision, D.Y. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by National Natural Science Foundation of China (Grant No. 51976042), and National Science and Technology Major Project of China (Grant No. 2017-I-0007-0008).

Acknowledgments: The authors would like to thank reviewers and editors for their valuable suggestions to refine this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tahan, M.; Tsoutsanis, E.; Muhammad, M.; Karim, Z.A.A. Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: A review. *Appl. Energy* **2017**, *198*, 122–144. [[CrossRef](#)]
2. Lu, S.; He, Q.; Wang, J. A review of stochastic resonance in rotating machine fault detection. *Mech. Syst. Signal Process.* **2019**, *116*, 230–260. [[CrossRef](#)]
3. Lei, Y.; Li, N.; Guo, L.; Li, N.; Yan, T.; Lin, J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mech. Syst. Signal Process.* **2018**, *104*, 799–834. [[CrossRef](#)]
4. Pan, Z.; Meng, Z.; Chen, Z.; Gao, W.; Shi, Y. A two-stage method based on extreme learning machine for predicting the remaining useful life of rolling-element bearings. *Mech. Syst. Signal Process.* **2020**, *144*, 106899. [[CrossRef](#)]
5. Wang, B.; Lei, Y.; Li, N.; Yan, T. Deep separable convolutional network for remaining useful life prediction of machinery. *Mech. Syst. Signal Process.* **2019**, *134*, 106330. [[CrossRef](#)]
6. Zaidan, M.A.; Mills, A.R.; Harrison, R.F.; Fleming, P., J. Gas turbine engine prognostics using Bayesian hierarchical models: A variational approach. *Mech. Syst. Signal Process.* **2016**, *70*, 120–140. [[CrossRef](#)]
7. Asr, M.Y.; Etefagh, M.M.; Hassannejad, R.; Razavi, S., N. Diagnosis of combined faults in Rotary Machinery by Non-Naive Bayesian approach. *Mech. Syst. Signal Process.* **2017**, *85*, 56–70. [[CrossRef](#)]
8. Maragoudakis, M.; Loukis, E.; Pantelides, P.P. Random forests identification of gas turbine faults. In Proceedings of the 2008 19th International Conference on Systems Engineering, Las Vegas, Nevada, 19–21 August 2008; pp. 127–132.
9. Li, F.; Wang, H.; Zhou, G.; Yu, D.; Li, J.; Gao, H. Anomaly Detection in Gas Turbine Fuel Systems Using a Sequential Symbolic Method. *Energies* **2017**, *10*, 724. [[CrossRef](#)]
10. Li, N.; Zhou, R.; Hu, Q.; Liu, X. Mechanical fault diagnosis based on redundant second generation wavelet packet transform, neighborhood rough set and support vector machine. *Mech. Syst. Signal Process.* **2012**, *28*, 608–621. [[CrossRef](#)]
11. Liu, J.; Bai, M.; Jiang, N.; Yu, D. A novel measure of attribute significance with complexity weight. *Appl. Soft Comput.* **2019**, *82*, 105543. [[CrossRef](#)]
12. Liu, J.; Bai, M.; Jiang, N.; Yu, D. Structural risk minimization of rough set-based classifier. *Soft Comput.* **2020**, *24*, 2049–2066. [[CrossRef](#)]
13. Liu, R.; Yang, B.; Zhang, X.; Wang, S.; Chen, X. Time-frequency atoms-driven support vector machine method for bearings incipient fault diagnosis. *Mech. Syst. Signal Process.* **2016**, *75*, 345–370. [[CrossRef](#)]
14. Chen, Z.; Gryllias, K.; Li, W. Mechanical fault diagnosis using convolutional neural networks and extreme learning machine. *Mech. Syst. Signal Process.* **2019**, *133*, 106272. [[CrossRef](#)]
15. Fentaye, A.D.; Baheta, A.T.; Gilani, S.I.U.H. Gas turbine gas-path fault identification using nested artificial neural networks. *Aircr. Eng. Aerosp. Technol.* **2018**, *90*, 992–999. [[CrossRef](#)]

16. Mast, T.A.; Reed, A.T.; Yurkovich, S.; Ashby, M.; Adibhatla, S. Bayesian belief networks for fault identification in aircraft gas turbine engines. In Proceedings of the 1999 IEEE International Conference on Control Applications (Cat. No. 99CH36328), Kohala Coast, HI, USA, 22–27 August 1999; Volume 1, pp. 39–44.
17. Losi, E.; Venturini, M.; Manservigi, L.; Ceschini, G.F.; Bechini, G. Anomaly Detection in Gas Turbine Time Series by Means of Bayesian Hierarchical Models. *J. Eng. Gas Turbines Power* **2019**, *141*. [[CrossRef](#)]
18. Wong, P.K.; Yang, Z.; Vong, C.M.; Zhong, J. Real-time fault diagnosis for gas turbine generator systems using extreme learning machine. *Neurocomputing* **2014**, *128*, 249–257. [[CrossRef](#)]
19. Fast, M.; Palme, T. Application of artificial neural networks to the condition monitoring and diagnosis of a combined heat and power plant. *Energy* **2010**, *35*, 1114–1120. [[CrossRef](#)]
20. Orozco, D.J.; Venturini, O.J.; Palacio, J.C.; Olmo, O.A. A new methodology of thermodynamic diagnosis, using the thermo-economic method together with an artificial neural network (ANN): A case study of an externally fired gas turbine (EFGT). *Energy* **2017**, *123*, 20–35. [[CrossRef](#)]
21. Liu, Z.; Karimi, I.A. Gas turbine performance prediction via machine learning. *Energy* **2020**, *192*, 116627. [[CrossRef](#)]
22. Wang, Z.; Zhao, N.; Wang, W.; Tang, R.; Li, S. A fault diagnosis approach for gas turbine exhaust gas temperature based on fuzzy c-means clustering and support vector machine. *Math. Probl. Eng.* **2015**, *2015*, 1–11. [[CrossRef](#)]
23. Zhou, D.; Zhang, H.; Weng, S. A new gas path fault diagnostic method of gas turbine based on support vector machine. *J. Eng. Gas Turbines Power* **2015**, *137*, 102605. [[CrossRef](#)]
24. Loboda, I.; Feldshcheyn, Y.; Ponomaryov, V. Neural networks for gas turbine fault identification: Multilayer perceptron or radial basis network? *Int. J. Turbo Jet Engines* **2012**, *29*, 37–48. [[CrossRef](#)]
25. Loboda, I.; Robles, M.A.O. Gas turbine fault diagnosis using probabilistic neural networks. *Int. J. Turbo Jet Engines* **2015**, *32*, 175–191. [[CrossRef](#)]
26. Yazdani, S.; Montazeri-Gh, M. A novel gas turbine fault detection and identification strategy based on hybrid dimensionality reduction and uncertain rule-based fuzzy logic. *Comput. Ind.* **2020**, *115*, 103131. [[CrossRef](#)]
27. Tahan, M.; Muhammad, M.; Karim, Z.A.A. A multi-nets ANN model for real-time performance-based automatic fault diagnosis of industrial gas turbine engines. *J. Braz. Soc. Mech. Sci. Eng.* **2017**, *39*, 2865–2876. [[CrossRef](#)]
28. Lu, F.; Wu, J.; Huang, J.; Qiu, X. Restricted-Boltzmann-based extreme learning machine for gas path fault diagnosis of turbofan engine. *IEEE Trans. Ind. Inform.* **2019**, *16*, 959–968. [[CrossRef](#)]
29. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
30. Prakash, R.M.; Thenmozhi, N.; Gayathri, M. Face Recognition with Convolutional Neural Network and Transfer Learning. In Proceedings of the 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 27–29 November 2019; pp. 861–864.
31. Vaswani, A.; Bengio, S.; Brevdo, E.; Chollet, F.; Gomez, A.N.; Gouws, S.; Jones, L.; Kaiser, L.; Kalchbrenner, N.; Parmar, N. Tensor2tensor for neural machine translation. *arXiv* **2018**, arXiv:1803.07416.
32. Gao, H.; Cheng, B.; Wang, J.; Li, K.; Zhao, J.; Li, D. Object classification using CNN-based fusion of vision and LIDAR in autonomous vehicle environment. *IEEE Trans. Ind. Inform.* **2018**, *14*, 4224–4231. [[CrossRef](#)]
33. Jia, F.; Lei, Y.; Lu, N.; Xing, S. Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization. *Mech. Syst. Signal Process.* **2018**, *110*, 349–367. [[CrossRef](#)]
34. Fu, X.; Hui, L.U.O.; Zhong, S.; Lin, L. Aircraft engine fault detection based on grouped convolutional denoising autoencoders. *Chin. J. Aeronaut.* **2019**, *32*, 296–307. [[CrossRef](#)]
35. Feng, D.; Xiao, M.; Liu, Y.; Song, H.; Yang, Z.; Hu, Z. Finite-sensor fault-diagnosis simulation study of gas turbine engine using information entropy and deep belief networks. *Front. Inf. Technol. Electron. Eng.* **2016**, *17*, 1287–1304. [[CrossRef](#)]
36. Liu, J.; Liu, J.; Yu, D.; Kang, M.; Yan, W.; Wang, Z.; Pecht, M. Fault Detection for Gas Turbine Hot Components Based on a Convolutional Neural Network. *Energies* **2018**, *11*, 2149. [[CrossRef](#)]
37. Mulewicz, B.; Marzec, M.; Morkisz, P.; Oprocha, P. Failures prediction based on performance monitoring of a gas turbine: A binary classification approach. *Schedae Inform.* **2018**, *26*, 9–21.
38. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
39. Li, F.; Ren, G.; Lee, J. Multi-step wind speed prediction based on turbulence intensity and hybrid deep neural networks. *Energy Convers. Manag.* **2019**, *186*, 306–322. [[CrossRef](#)]
40. Wang, K.; Qi, X.; Liu, H. A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network. *Appl. Energy* **2019**, *251*, 113315. [[CrossRef](#)]
41. Wang, Y.; Gan, D.; Sun, M.; Zhang, N.; Lu, Z.; Kang, C. Probabilistic individual load forecasting using pinball loss guided LSTM. *Appl. Energy* **2019**, *235*, 10–20. [[CrossRef](#)]
42. Zhang, Y.; Xiong, R.; He, H.; Pecht, M. Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE Trans. Veh. Technol.* **2018**, *67*, 5695–5705. [[CrossRef](#)]
43. Cui, Y.; Wang, S.; Li, J. LSTM Neural Reordering Feature for Statistical Machine Translation. In Proceedings of the NAACL-HLT, San Diego, CA, USA, 12–17 June 2016; pp. 977–982.
44. Chiu, J.P.C.; Nichols, E. Named entity recognition with bidirectional LSTM-CNNs. *Trans. Assoc. Comput. Linguist.* **2016**, *4*, 357–370. [[CrossRef](#)]

45. Feng, C.; Chang, L.; Li, C.; Ding, T.; Mai, Z. Controller optimization approach using LSTM-based identification model for pumped-storage units. *IEEE Access* **2019**, *7*, 32714–32727. [[CrossRef](#)]
46. Gonzalez, J.; Yu, W. Non-linear system modeling using LSTM neural networks. *IFAC PapersOnLine* **2018**, *51*, 485–489. [[CrossRef](#)]
47. Wang, Y. A new concept using lstm neural networks for dynamic system identification. In Proceedings of the 2017 American Control Conference (ACC), Seattle, WA, USA, 24–26 May 2017; pp. 5324–5329.
48. Cheng, X. Research on Simulation of Marine Gas Turbine Gas Path Component Degradation. Master's Thesis, Harbin Institute of Technology, Harbin, China, 2014.
49. Simon, D.L.; Borguet, S.; Léonard, O.; Zhang, X. *Aircraft Engine Gas Path Diagnostic Methods: Public Benchmarking Results*; American Society of Mechanical Engineers: New York, NY, USA, 2013.
50. Cao, Y.; Zhang, B.; Wang, H.; Bai, Y. Gas Path Fault Diagnosis of Aeroengine Based on Soft Square Pinball Loss ELM. *IEEE Access* **2020**, *8*, 131032–131046. [[CrossRef](#)]
51. Childers, S.A. Methods Relating to Gas Turbine Control and Operation. U.S. Patent 8,355,854, 15 January 2013.
52. Jones, S.M. An Introduction to Thermodynamic Performance Analysis of Aircraft Gas Turbine Engine Cycles Using the Numerical Propulsion System Simulation Code. 2007. Available online: <https://ntrs.nasa.gov/citations/20070018165> (accessed on 11 December 2020).
53. Park, Y.; Choi, M.; Li, X.; Jung, C.; Na, S.; Chou, G. Prediction of operating characteristics for industrial gas turbine combustor using an optimized artificial neural network. *Energy* **2020**, *213*, 118769. [[CrossRef](#)]
54. Camporeale, S.M.; Fortunato, B.; Mastrovito, M. A modular code for real time dynamic simulation of gas turbines in simulink. *J. Eng. Gas Turbines Power* **2006**, *128*, 506–517. [[CrossRef](#)]
55. Song, C. Model-Based Gas Path Diagnosis for Three-Shaft Gas Turbine. Master's Thesis, Harbin Institute of Technology, Harbin, China, 2014.
56. Cai, D. Method for Underdetermined Fault Diagnosis with the Prior Knowledge of Gas Turbine. Master's Thesis, Harbin Institute of Technology, Harbin, China, 2015.
57. Drucker, H.; Burges, C.J.C.; Kaufman, L.; Smola, A.; Vapnik, V. Support vector regression machines. *Adv. Neural Inf. Process. Syst.* **1997**, *9*, 155–161.
58. Liu, J.; Jin, X.; Dong, F.; He, L.; Liu, H. Fading channel modelling using single-hidden layer feedforward neural networks. *Multidimens. Syst. Signal Process.* **2017**, *28*, 885–903. [[CrossRef](#)]
59. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: A new learning scheme of feedforward neural networks. In Proceedings of the 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541), Budapest, Hungary, 25–29 July 2004; Volume 2, pp. 985–990.
60. Elman, J.L. Finding structure in time. *Cogn. Sci.* **1990**, *14*, 179–211. [[CrossRef](#)]
61. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mech. Syst. Signal Process.* **2020**, *138*, 106587. [[CrossRef](#)]
62. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2018**, *108*, 33–47. [[CrossRef](#)]
63. Bai, M.; Liu, J.; Chai, J.; Zhao, X.; Yu, D. Anomaly detection of gas turbines based on normal pattern extraction. *Appl. Therm. Eng.* **2020**, *166*, 114664. [[CrossRef](#)]
64. Chen, J.; Xu, X.; Zhang, X. Fault Detection for Turbine Engine Disk Based on Adaptive Weighted One-Class Support Vector Machine. *J. Electr. Comput. Eng.* **2020**, *2020*, 1–10. [[CrossRef](#)]
65. Cao, J.; Dai, H.; Lei, B.; Yin, C.; Zeng, H.; Kummert, A. Maximum Correntropy Criterion-Based Hierarchical One-Class Classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**. [[CrossRef](#)] [[PubMed](#)]
66. Annadatha, A.; Stamp, M. Image spam analysis and detection. *J. Comput. Virol. Hacking Tech.* **2018**, *14*, 39–52. [[CrossRef](#)]
67. Schölkopf, B.; Williamson, R.C.; Smola, A.J.; Shawe-Taylor, J.; Platt, J. Support vector method for novelty detection. *Adv. Neural Inf. Process. Syst.* **2000**, *12*, 582–588.
68. Schölkopf, B.; Platt, J.C.; Shawe-Taylor, J.; Smola, A.J.; Williamson, R.C. Estimating the support of a high-dimensional distribution. *Neural Comput.* **2001**, *13*, 1443–1471. [[CrossRef](#)]
69. Breunig, M.M.; Kriegel, H.P.; Ng, R.T.; Sander, J. LOF: Identifying density-based local outliers. In Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, Dallas, TX, USA, 16–18 May 2000; pp. 93–104.
70. Liu, F.T.; Ting, K.M.; Zhou, Z.H. Isolation forest. In Proceedings of the 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 15–19 December 2008; pp. 413–422.
71. Zhou, F.; Park, J.H.; Liu, Y. Differential feature based hierarchical PCA fault detection method for dynamic fault. *Neurocomputing* **2016**, *202*, 27–35. [[CrossRef](#)]
72. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
73. Buitinck, L.; Louppe, G.; Blondel, M.; Pedregosa, F.; Müller, A.; Grisel, O.; Niculae, V.; Prettenhofer, P.; Gramfort, A.; Grobler, J.; et al. API design for machine learning software: Experiences from the scikit-learn project. *arXiv* **2013**, arXiv:1309.0238.
74. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Wieser, E.; Taylor, J.; Berg, S.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [[CrossRef](#)]