Article

Single-Sensor Engine Multi-Type Fault Detection

Daijie Tang [1], Fengrong Bi [1], Jiangang Cheng [1], Xiao Yang [1], Pengfei Shen [1] and Xiaoyang Bi [2],*

[1] State Key Laboratory of Engines, Tianjin University, Tianjin 300350, China
[2] State Key Laboratory of Reliability and Intelligence Electrical Equipment, Hebei University of Technology, Tianjin 300130, China
* Correspondence: xy_bi@hebut.edu.cn

Abstract: Engine fault detection is conducive to improving equipment reliability and reducing maintenance costs. In practical scenarios, high-quality data is difficult to obtain. Usually, only single-sensor data is available. This paper proposes a fault detection method combining Variational Mode Decomposition (VMD) and Random Forest (RF). At first, the spectral energy distribution is obtained by decomposing and statistic the engine data of multiple working conditions. Based on the spectral energy distribution, the overall optimal mode number was identified, and the quadratic penalty term was optimized using SNR. The improved VMD (IVMD) improves mode aliasing and iterative efficiency and unifies feature dimensions. Decomposition of real signals demonstrates the effectiveness. The paper designs a feature vector composed of seven types of attributes, including unit bandwidth energy, center frequency, maximum singular value and so on. The feature vector is then fed to RF for classification. Features are selected in order of importance to classification to improve the training efficiency. By comparing with various algorithms, the proposed method has higher accuracy and faster training efficiency in single-speed, multi-speed and cross-speed single-sensor data diagnosis. The results show that the method has application prospects with little training data and low hardware requirements.

Keywords: fault detection; single-sensor data; variational mode decomposition; vibration; random forest

1. Introduction

As one of the critical power sources, the reliability of engines has received more attention in recent years. In time, engine fault detection can detect weak faults, which is conducive to fault prevention and repair. Data-driven approaches usually require large amounts of high-quality data for training. However, engine labelled-data is challenging to obtain and mostly comes from a single sensor due to cost constraints. Research on single-sensor engine fault detection based on small data amounts and low hardware requirements is necessary [1].

Vibration acceleration signals are widely used in fault detection research because of their rich component condition information and ease of measurement [2,3]. Ma et al. proposed a multi-channel Lanczos quaternion singular spectrum analysis to extract fault characteristic frequencies from multiple vibration sensor signals [4]. Ribeiro et al. proposed a multi-head one-dimensional convolutional neural network to diagnose six motor faults using vibration signals from two directions [5]. However, the engine vibration signal has a wide frequency band (up to about 12,000 Hz), and sensors with high sampling accuracy and a wide frequency band with good stability are usually costly. Moreover, the hardware conditions of the engine control system are ordinary, so the research on single-sensor fault detection under low hardware requirements is gradually gaining attention. Basuraj et al. proposed a single-sensor online filtering method for recursive singular spectrum analysis based on the concept of first-order feature perturbation, which proved its effectiveness in several data sets. [6]. Ayati et al. used KNN for single-sensor fault classification after

extracting features using fast Fourier transform and wavelet packet transform [7]. In general, it is challenging to diagnose faults in different cylinders of an engine separately using a single sensor.

Engine fault detection methods can be roughly divided into three categories: knowledge-driven, model-driven, and data-driven. Wang et al. proposed an aero-engine dynamic threshold fault detection based on the isolated forest method, which requires only normal data for training to achieve high accuracy [8]. Ellefsen et al. proposed an online diagnosis method for marine diesel engine degradation based on variational autoencoder and expert knowledge [9]. Knowledge-driven methods usually diagnose a single type of fault and require solid expert knowledge. Liu et al. proposed a model-based aero-engine soft fault detection method, which achieved fault diagnosis by comparing smooth residuals and preset thresholds [10]. Wang et al. established a mapping model between the shaft radial vibration average and the misalignment value based on shaft shape characteristics. A new monitoring scheme has been designed and the accuracy of detecting misalignment is greater than 90% [11]. Model-driven method research can help explore the failure mechanism, but it is usually challenging to achieve. Data-driven methods are widely used due to their ease of implementation and high accuracy [12,13]. Deep learning methods have been widely used in engine fault detection in recent years due to their powerful data mining capabilities [14,15]. These methods require less expert knowledge and more high-quality training data. However, high-dimensional and huge data processing capability leads to higher hardware requirements for deep learning methods. In practical scenarios, high-quality training data is difficult to obtain because of the dangers of engine failure simulation experiments. Under the constraints of low hardware conditions and lack of data, the combination of signal processing methods and simple pattern classification methods still has potential to be explored [16,17].

Variational Mode Decomposition (VMD) is an advanced signal processing method capable of decomposing a signal into several intrinsic mode functions (IMFs) [18]. Compared with empirical mode decomposition (EMD), VMD effectively suppresses mode aliasing and improves the quality of decomposition [19]. However, the mode number K and quadratic penalty term α, predefined in VMD, strongly influence the decomposition and are difficult to determine [20,21]. For these reasons, scholars have proposed many optimization ideas for adaptively selecting K and α [21,22]. The adaptive VMD method leads to a varying number of IMFs, so component screening is usually performed after decomposition [23,24]. The process of screening IMFs requires expert knowledge and is time-consuming and labor-intensive. In addition, many scholars optimize (K, α) through swarm intelligence optimization algorithms [25,26]. This method ignores the problem that the VMD efficiency drops sharply as K increases (as shown in Section 4).

The unsupervised clustering method is ineffective in diagnosing engine faults because there are many types of failure, complex operating conditions, and large signal noise [27,28]. Supervised pattern classification methods such as deep neural networks (DNN) are more suitable due to their powerful learning capabilities. Shahid et al. used a one-dimensional convolutional neural network (1DCNN) to identify the crankshaft angle degree of the engine and successfully diagnosed the misfire fault [29]. Zhang et al. proposed a long short-term memory recurrent neural network (LSTM-RNN) for evaluating bearing degradation and proposed waveform entropy to improve the accuracy effectively [30]. Lee et al. compared the performance of multilayer perception (MLP), residual network (ResNet), LSTM, and ResNet-LSTM in diagnosing production failure cases and found that ResNet-LSTM works best [31]. However, the effectiveness of DNN is built on sufficient high-quality labeled data. Due to the complex calculation of DNN, the training time is long, and it is challenging to optimize and retrain the model [32]. Li et al. first used a simplified DNN to extract the fault features of rotating machinery and then combined random forest (RF) for fault classification, which has higher efficiency and accuracy than advanced DNN methods [32]. RF has faster training and classification speed than DNN and may be suitable for engine fault detection.

The paper aims to propose a single-sensor, cross-speed fault detection method that is applicable to low hardware requirements and small data amounts. The work has resulted in the following contributions.

(1) A new overall K and α optimization method based on spectral energy distribution and SNR is proposed for VMD, avoiding IMF screening and unifying the feature dimension to prepare for quick diagnosis.

(2) The center frequencies are preset based on spectral energy distribution, which reduces the number of VMD iterations and mode aliasing.

(3) A feature set was designed for IVMD-RF to achieve single-sensor fault diagnosis. Further filtering of features by feature importance ranking improves efficiency. Different single-sensor datasets demonstrate the effectiveness of the method.

The rest of the article is organized as follows. Section 2 introduces the basic principles of the methods used in the paper. In Section 3, the fault data collection experiment of the diesel engine is presented. Section 4 introduces the optimization of the VMD method and the verification of its decomposition effect. In Section 5, IVMD-RF is presented and compared with various DNN methods on two diagnostic cases.

2. Theories

2.1. Variational Mode Decomposition

The purpose of VMD is to decompose an actual signal into several ideal narrowband signals while satisfying the constraint that the sum of their bandwidths is the smallest. Assume that each IMF closely surrounds its center frequency in the frequency domain. Therefore, the objective can be summarized as the following constrained variational problem:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left\{ \sum\limits_k \left\| \partial_t \left[(\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum\limits_k u_k = f \end{cases}, \tag{1}$$

where $\{u_k(t)\} = \{u_1(t), u_1(t), \ldots, u_k(t)\}$ and $\{\omega_k\} = \{\omega_1, \omega_2, \ldots, \omega_k\}$ represent the decomposed IMFs and the corresponding center frequencies, respectively. $\delta(t)$ is the shock function.

The reconstruction constraint can be addressed by introducing a quadratic penalty α and Lagrange multipliers λ. The constrained variational problem of (1) is transformed into an unconstrained one by introducing these two parameters. The obtained augmented Lagrangian is shown in (2):

$$L(\{u_k\}, \{\omega_k\}, \lambda) := \alpha \sum_k \left\| \partial_t \left[(\delta(t) + \frac{j}{\pi t}) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle, \tag{2}$$

This problem can be solved by Parseval/Plancherel Fourier isometry under the norm. The expressions of $\widehat{u}_k^{n+1}(\omega)$ and ω_k^{n+1} are shown in (3) and (4).

$$\widehat{u}_k^{n+1}(\omega) = \frac{\widehat{f}(\omega) - \sum_{i \neq k} \widehat{u}_i(\omega) + \frac{\widehat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2}, \tag{3}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \widehat{u}_k(\omega) \right|^2 d\omega}{\int_0^\infty \left| \widehat{u}_k(\omega) \right|^2 d\omega}, \tag{4}$$

where (3) is equivalent to the Wiener filter of the current residual $\widehat{f}(\omega) - \sum\limits_{i<K} \widehat{u}_i(\omega)$. IMF can be obtained by inverse Fourier transform of $\widehat{u}_k^{n+1}(\omega)$. The flow of VMD is shown in Algorithm 1. The default ε value is 1×10^{-7}.

Algorithm 1: VMD

Input: A signal f, mode number K and quadratic penalty α.
Output: A set of IMFs

Initialize $\{\widehat{u}_k^1\}, \{\widehat{\omega}_k^1\}, \{\widehat{\lambda}^1\}, n \leftarrow 0$
repeat
　for $k \leftarrow 1$ to K **do**
Update $\widehat{u}_k$ for all $\omega \geq 0$ by (3)
Update ω_k by (4)
　end for
　Dual ascent for all $\omega \geq 0$:
$\widehat{\lambda}^{n+1}(\omega) \leftarrow \widehat{\lambda}^n(\omega) + \tau\left[\widehat{f}(\omega) - \sum\limits_k \widehat{u}_k^{n+1}(\omega)\right]$
until convergence: $\sum\limits_k \left\|\widehat{u}_k^{n+1} - \widehat{u}_k^n\right\|_2^2 / \left\|\widehat{u}_k^n\right\|_2^2 < \varepsilon$.

2.2. Random Forests

The random forest algorithm was proposed by Breiman [33], which is suitable for solving data prediction and classification. A random forest is a combination of decision tree classifiers. Each tree depends on the value of an independently sampled random vector and has the same distribution for all trees in the forest.

(1)　Suppose the original sample is $X = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, where x_i and y_i represent feature values and labels, respectively. T training samples $X_1, X_2, \ldots, X_T$ are extracted from the original dataset X by bootstrap sampling with return, and $X_i(i = 1, 2, \ldots T)$ and X have the same number of samples.

(2)　Build a decision tree $hi(X_i, \Theta_k)$ for each training sample $X_i(i = 1, 2, \ldots T)$, where i =1, 2, ... T, k =1, 2, ... The decision tree model used in the paper is shown in (5) and (6).

$$d(x_1, x_2, \ldots, x_n, h_t) = \begin{cases} label(h_t) & h_t \text{ is the leaf node} \\ d(x_1, x_2, \ldots, x_n, h_t) & h_t \text{ is the inner node} \end{cases} \tag{5}$$

$$hi(X_i, \Theta_k) = d(x_1, x_2, \ldots, x_n, root(h_t)) \tag{6}$$

where $root(h_t)$ is the root node of the decision tree. $d(x_1, x_2, \ldots, x_n, h_t)$ is the division criterion of the decision tree. The segmentation criterion consists of segmentation variables and predictions measured by the impurity function.

The Gini coefficient is proportional to the impurity level. The optimal split is to find the largest split of the Gini coefficient as follows:

$$Gini(t) = 1 - \sum_{j=1}^{J} \{p(j|t)\}^2 \tag{7}$$

where $p(j|t)$ is the probability of the jth category in node t, that is, the ratio of the jth category to the total number of sample labels J.

Before selecting attributes for each non-leaf node, randomly select m attributes from M attributes as the set of categorical attributes for the current node. Take $m = \text{int}(\sqrt{M})$, where int is the rounding function. The nodes are divided according to the optimal division method of m attributes, and a complete decision tree is established. The growth of each decision tree is not pruned until the leaf node grows.

A random forest generated from *T* decision trees is used to classify the test samples. Each tree has voting power to decide the classification result. Summarize the output categories of the decision tree, and the category with the most votes is the final classification result. The classification decision model *H(x)* is shown in (8).

$$H(x) = \underset{\gamma}{\mathrm{argmax}} \sum_{i=1}^{T} I(hi(X_i, \Theta_k) = \gamma)$$

(8)

where γ is the label variable of the output and I is the indicator function.

3. Diesel Engine Faults Simulation Experiment

To verify the effectiveness of the proposed method, our team conducted a fault simulation experiment on an in-line 6-cylinder diesel engine. The specific parameters of the engine are shown in Table 1. The experiment was performed on a bench base supported by an air spring. The engine and the dynamic dynamometer adopt a flexible connection. The photoelectric pulse speed sensor is placed at the position of the vertical connecting shaft to measure the engine speed. The vibration acceleration sensors are arranged on the cylinder head and block as shown in Figure 1. The data used in this paper are vibration acceleration signals in the Y-direction in Figure 1. The signal is input to the computer for processing and recording after passing through the acquisition front end. The models of the instruments used in the experiment are shown in Table 2.

Table 1. Parameters of diesel engine.

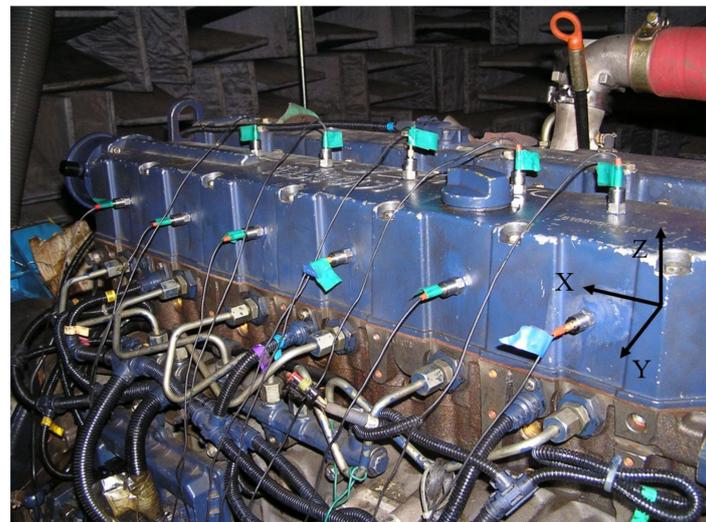| Items	Parameters
Displacement	7.14 L
Rated power/Rated speed	220 kW/2300 rpm
Maximum torque/Speed range	1250 Nm/1200–1600 rpm
Intake/Exhaust valve clearance	0.30 m/0.50 m

Figure 1. Sensor positions and coordinate direction.

Table 2. Experimental instrument parameters.

Instruments	Parameters
Dynamic dynamometer	CAC380, Xiangyi Power
Vibration acceleration sensor	621B40, PCB
Photoelectric pulse speed sensor	SPSR-115/230, Monarch
Data acquisition front end	SCADAS05, LMS

Components with the highest failure probability are fuel injection and oil supply equipment (25.1%), water leakage (13.1%), and valves and sealing (17.4%) [34]. Generally, leak failure is easy to detect by water temperature sensors. The paper focuses on two other types of failures. The paper simulates three faults for the fuel supply equipment: abnormal common rail pressure, abnormal fuel supply, and abnormal injection advance angle. Abnormal rail pressure is set to simulate the fault of the common rail system, and the insufficient fuel supply is to simulate injector failure. The weak power combustion abnormality is simulated by slightly changing the injection advance angle. In addition, the abnormal valve clearance is simulated by adjusting the opening of intake and exhaust valves with a plug gauge. The abnormal valve clearance conditions all occurred on the first cylinder only. Experiments were performed at the following rotational speeds: 700 rpm, 1300 rpm, 1600 rpm, 2000 rpm, and 2300 rpm. The parameters of normal working conditions under each speed condition are shown in Table 3. The fault settings at rated speed (2300 rpm) are shown in Table 4, where the Roman numerals represent different fault conditions. The fault conditions of other speeds are also adjusted to the same extent as those in Table 4 on the basis of the normal parameters in Table 3. The abnormal advance angle failure simulation is not carried out under 700 rpm idling conditions. The load range of the engine includes 100% and 50%.

Table 3. Normal working conditions under different speeds.

Speed (rpm)	Valve Clearance-Intake, Exhaust (mm)	Fuel Supply (mg/cyc)	Rail Pressure (bar)	Injection Advance Angle (°CA)
700	(0.30, 0.50)	60.0	405	-
1300	(0.30, 0.50)	117.0	1250	9.49
1600	(0.30, 0.50)	117.0	1350	12.98
2000	(0.30, 0.50)	117.0	1500	15.00
2300	(0.30, 0.50)	112.5	1550	18.45

Table 4. Fault type and degree parameter setting (2300 rpm).

Mark	Valve Clearance (Intake, Exhaust)/mm	Fuel Supply	Rail Pressure/bar	Injection Advance Angle/°CA
I	(0.30, 0.50)	100%	1550	18.45
II	(0.20, 0.40)	100%	1550	18.45
III	(0.35, 0.55)	100%	1550	18.45
IV	(0.40, 0.60)	100%	1550	18.45
V	(0.30, 0.50)	75%	1550	18.45
VI	(0.30, 0.50)	25%	1550	18.45
VII	(0.30, 0.50)	100%	1350	18.45
VIII	(0.30, 0.50)	100%	1150	18.45
IX	(0.30, 0.50)	100%	1550	17.45
X	(0.30, 0.50)	100%	1550	16.45
XI	(0.30, 0.50)	100%	1550	19.45
XII	(0.30, 0.50)	100%	1550	20.45

Note: The shaded green marks the location of the faulty parameter.

4. Optimization of Variational Mode Decomposition

VMD's denoising ability is better than EMD [35], and the decomposed IMFs have a better signal-to-noise ratio (SNR). However, the decomposition effect of VMD is greatly affected by parameter settings, especially the mode number K and the quadratic penalty term α. Improper K value setting will lead to over-decomposition or under-decomposition. In addition, as K increases, the efficiency of the original VMD decreases drastically. Figure 2 shows the effect of different K values on the decomposition time of each IMF. The results show that the efficiency of VMD is much higher when $K \leq 3$. From Figure 2, traversing K to find the optimal value and using various swarm intelligence optimization algorithms are both inefficient. Therefore, Ref. [36] proposes an adaptive recursive variational mode

decomposition (ARVMD) that dynamically selects the *K* in recursive loops. ARVMD effectively improves efficiency and reduces recursive mode aliasing. The process of ARVMD is shown in Algorithm 2.

Algorithm 2: ARVMD

Input: A signal f_0, Sampling frequency Fs and quadratic penalty α.
Output: A set of IMFs.
$f = f_0$; IMFs = []; E_u = []; $i = 0$;
while $E_{ui} > E_{th}$ **do**
   $i = i + 1$;
   $P_f \leftarrow$ Power spectral density (f);
   $(P_{max}, F_{max}) \leftarrow$ Maximum, corresponding frequency (P_f);
   $N_{peak} \leftarrow$ Numbers of maxima points in [$F_{max} \pm 0.027 \times$ Fs];
   $\{F_1, F_2, \ldots, F_n\} \leftarrow$ Corresponding frequencies of maxima points;
   $K_i = \begin{cases} 1, N_{peak} < 2 \\ 2, N_{peak} = 2 \\ 3, N_{peak} > 2 \end{cases}$
   $\{u_1, u_2, \ldots, u_{Ki}\} \leftarrow$ VMD ($f, K_i, \alpha, \{F_1, F_2, \ldots, F_{Ki}\}$);
   $\{E_{u1}, E_{u2}, \ldots, E_{uKi}\} \leftarrow$ Unit bandwidth energy ($\{u_1, u_2, \ldots, u_{Ki}\}$);
   IMFs $\leftarrow$ IMFs $\cup \{u_1, u_2, \ldots, u_{Ki}\}$;
   $E_u \leftarrow E_u \cup \{E_{u1}, E_{u2}, \ldots, E_{uKi}\}$;
   $f = f - \sum_1^{Ki} u_{Ki}(t)$;
end while
IMFs $\leftarrow$ Selection by $E_{ui} > E_{th}$ (IMFs)
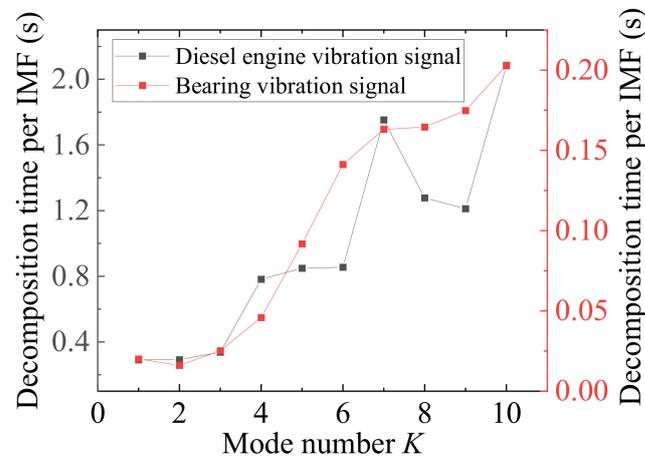**return** IMFs

---

Figure 2. Decomposition time per IMF of VMD. The curve data come from the average of normal signals of various speeds. The engine data come from the experiment of Section 3. The bearing data come from the bearing dataset of Case Western Reserve University [37].

Complex types and working conditions characterize engine faults. However, the component number obtained by ARVMD is variable, resulting in inconsistent feature vector dimensions, which is not conducive to diagnosing multi-speed engine vibration data. A *K*-value optimization method based on the energy distribution in the frequency domain is proposed to unify the feature dimension. First, ARVMD decomposes the signals of various engine working conditions and obtains many IMFs. These conditions contain data for different speeds and faults (I to XII, as shown in Table 4). Then, the unit bandwidth energy [36] of each IMFs is calculated, and the center frequency of the IMFs is recorded. The unit bandwidth energy is shown in (9):

$$E_u = \frac{e_{IMFi}}{B_{IMFi}} \tag{9}$$

where e_{IMFi} is the energy of IMF and B_{IMFi} is the bandwidth of IMF. The bandwidth is the width of the spectrum when the amplitude of the power spectral density is reduced by 99%. The frequency band [0 Hz, 12,800 Hz] is divided into 128 segments, and the width of each segment is 100 Hz. The unit bandwidth energy of the components located in each segment is counted and averaged. Figure 3 shows the unit bandwidth energy spectrum displayed on divided frequency bands. The results show that there are six prominent energy frequencies: 150 Hz, 1450 Hz, 1950 Hz, 2450 Hz, 4950 Hz, and 7050 Hz. Here, each frequency segment uses the frequency in the middle as the value of the abscissa. Therefore, the engine data will be uniformly decomposed using K equal to 6. This approach can improve the consistency of data processing and help reduce the randomness caused by adaptive decomposition. It also ensures that the dimension of the feature vectors at different speeds is uniform.
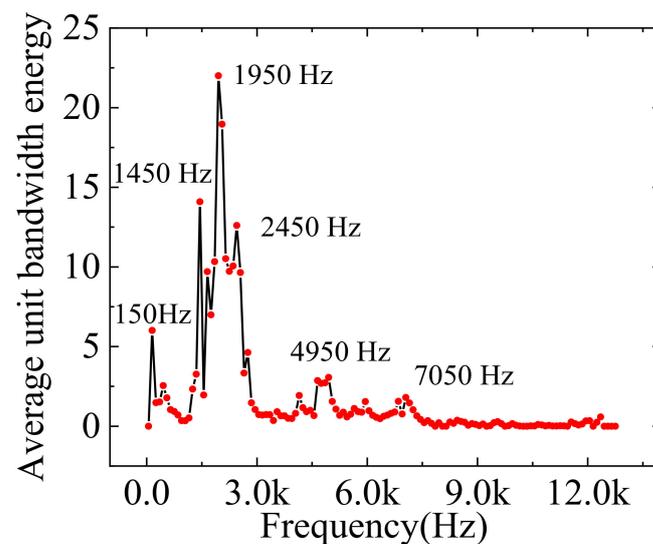

Figure 3. Frequency domain distribution of unit bandwidth energy of engine data.

Furthermore, the iterations of the center frequency of the original VMD are zero-based. It is beneficial for decomposing low-frequency components, but the decomposition time for high-frequency components is longer. Presetting suitable initial center frequencies can significantly improve the efficiency of the VMD [18]. Therefore, the six significant frequencies in Figure 3 are used as the initial center frequencies to iterate.

The quadratic penalty term α is a parameter introduced to improve the convergence when solving the variational model. The role of α in the decomposition is reflected in the noise reduction of the signal. The SNR is the best criterion for choosing a suitable α. However, it is difficult to obtain the SNR of the actual signal after decomposition. Therefore, a set of simulated signals is constructed according to the spectral energy distribution of Figure 3. The expression of the simulated signal is as (10). $\{s_1, s_2, \ldots, s_6\}$ are single-frequency components, which restore the amplitude ratio and frequency of each component in Figure 3. The amplitude of s_3 is set to 100, and the other components are reduced proportionally. s_7 is the noise component with a power of 25 dbW. S_1 is decomposed using VMD, where K is six, and the initial center frequency is preset. Set the variation range of α to [1000, 20,000], and the step size is 100. Calculate the SNR between IMFs and $\{s_1, s_2, \ldots, s_6\}$, and the results are shown in Figure 4. With the increase of α, the SNR has a trend of increasing first and then decreasing. Summing the SNR of each component, it is found that the total SNR does not change much when α is 6000 to 8000. The value of α used in the paper is 6800, and the inset of Figure 4 shows that the SNR reaches the maximum at

this value. After the optimized α is obtained, it is used in the mode number optimization for reverse verification, and the results show that it does not affect the results in Figure 3.

$$\begin{cases} s_1(t) = 27\sin(2\pi * 150t), 0 \le t \le 0.053 \\ s_2(t) = 64\sin(2\pi * 1450t), 0 \le t \le 0.053 \\ s_3(t) = 100\sin(2\pi * 1950t), 0 \le t \le 0.053 \\ s_4(t) = 57\sin(2\pi * 2450t), 0 \le t \le 0.053 \\ s_5(t) = 13\sin(2\pi * 4950t), 0 \le t \le 0.053 \\ s_6(t) = 8\sin(2\pi * 7050t), 0 \le t \le 0.053 \\ s_7(t) = \eta \\ S_1 = s_1 + s_2 + s_3 + s_4 + s_5 + s_6 + s_7 \end{cases} \tag{10}$$
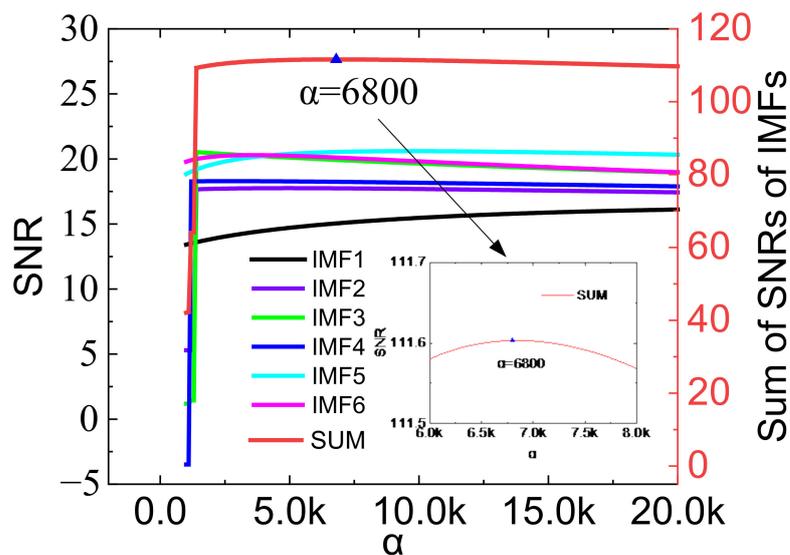

Figure 4. The effect of α on the decomposition SNR.

The optimization of *K*, α and the iterative optimization of the center frequency have been completed. Next, decompose an actual signal using the improved VMD (IVMD) to verify the effect. A signal of valve clearance increase at 1600 rpm (Condition III in Table 4) was randomly selected for decomposition. The signal's time and frequency domain are shown in Figure 5a,b. Decompose this signal using VMD and IVMD. Figure 6 shows the frequency domain image of the decomposed IMFs. VMD decomposes four components in the [2000 Hz, 3000 Hz] while IVMD decomposes three. The results show that using the same *K*, VMD focuses on decomposing low-frequency components, while IVMD is more balanced. The average bandwidth aliasing ratio R_{ABA} is introduced to measure the effect of suppressing mode aliasing [36]. The expression of R_{ABA} is shown in (11):

$$R_{ABA} = \sum_{i=1}^{K} \frac{1}{K} \frac{B_A}{B_{IMFi}}, i = 1, 2, \dots, K. \tag{11}$$

where *K* is the mode number, B_A is the aliasing bandwidth of the IMFi and other components, and B_{IMFi} is the bandwidth of IMFi. The smaller the R_{ABA}, the better the effect of suppressing mode aliasing. The R_{ABA} for VMD and IVMD results is 0.13 and 0.05, respectively. IVMD suppresses mode aliasing better than VMD. In addition, the center frequency iteration curves of IMFs are shown in Figure 7. Figure 7 shows that IVMD performs 87 iterations, less than VMD's 194 iterations, effectively improving efficiency. The results show that the presetting center frequency can significantly improve the iteration efficiency.
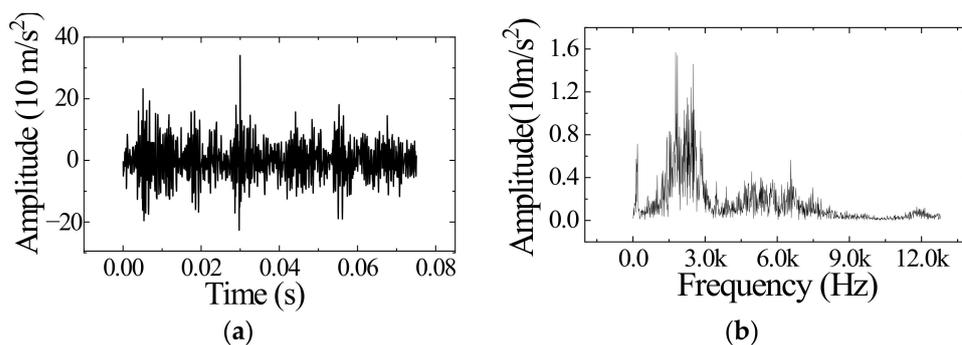
Figure 5. A signal of valve clearance increase at 1600 rpm. (**a**) The signal in time domain. (**b**) The signal in frequency domain.
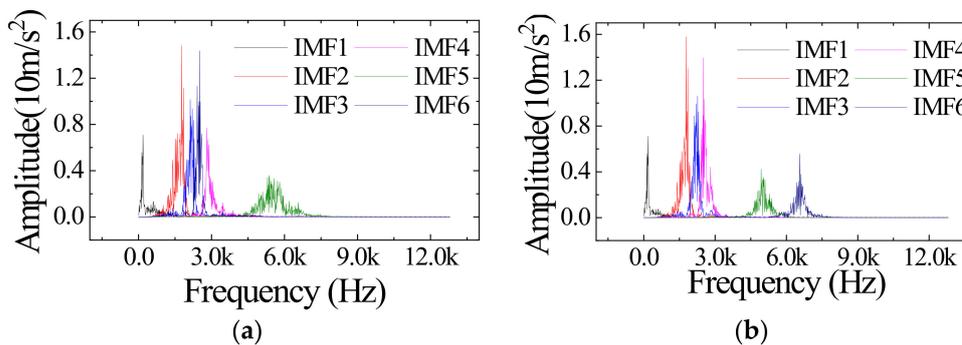

Figure 6. Frequency domain image of the decomposed IMFs. (**a**) Result of VMD. (**b**) Result of IVMD.
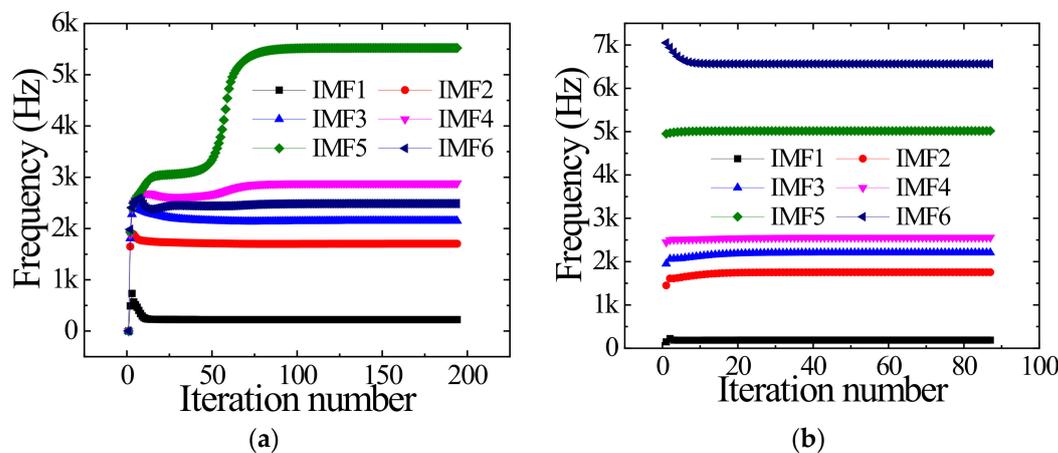

Figure 7. Center frequency iterative curves for IMFs. (**a**) Iterative curves of VMD. (**b**) Iterative curves of IVMD.

5. VMD-RF Fault Detection Method

After the IVMD decomposition of the engine signal, calculating proper features is beneficial to improve the diagnostic accuracy. The RF method can automatically select a subset of features for classification by bootstrap sampling with return. Instead of considering the feature dimension, features are required to describe the data information as comprehensively as possible. Therefore, the used features include overall features and local features. Finally, seven types of local features are selected. Namely, maximum singular value, energy, unit bandwidth energy, kurtosis, variance, root mean square value (RMS), and center frequency. The seven types of features are calculated for the six IMFs obtained by IVMD. In addition, maximum singular value, energy, RMS, and variance are calculated for the original signal. Feature names and symbols are shown in Table 5.

Table 5. Attribute names and symbols.

Attribute Name	Symbols
Maximum singular value	S_1, S_2, S_3, S_4, S_5, S_6
Energy	E_1, E_2, E_3, E_4, E_5, E_6
Unit bandwidth energy	Eu_1, Eu_2, Eu_3, Eu_4, Eu_5, Eu_6
Kurtosis	K_1, K_2, K_3, K_4, K_5, K_6
Variance	V_1, V_2, V_3, V_4, V_5, V_6
Root mean square	R_1, R_2, R_3, R_4, R_5, R_6
Center frequency	C_1, C_2, C_3, C_4, C_5, C_6
Original signal attributes	S, E, R, V

Note: The features calculated for IMF1 to IMF6 are denoted by the symbols with suffixes 1 to 6, respectively. The symbols without suffixes indicate the features calculated for the original signal.

5.1. Case 1: Diesel Engine Fault Diagnosis

Once the complete feature set is obtained, the feature set can be fed into the RF for classification. The whole flow of fault diagnosis is shown in Figure 8. The data of the first cylinder head (1H), the third cylinder head (3H), and the first cylinder block (1B) at 2300 rpm and the data of the first cylinder head at 2000 rpm were selected for preliminary verification of the algorithm's validity. Each dataset contains four types of faults in Table 4, with a total of 12 fault conditions. Each fault condition includes 200 samples, and the training/test ratio is 4:1. The number of decision trees in RF is 100. The depth of the decision tree is not limited. Then, the training samples are used to generate a random forest. The results of the diagnostic accuracy are shown in Table 6. The proposed method is compared with sequential minimum optimization for support vector machines (SMO-SVM), Multilayer Perceptron (MLP) [31], one-dimensional convolutional neural networks (1DCNN) [29], long and short term memory recurrent neural networks (LSTM-RNN) [30], and residual neural networks (ResNet) [31]. The 1DCNN and LSTM-RNN ran for 200 epochs, while ResNet ran for 30 epochs. The parameters of each algorithm are as follows:

(1) SVM: The RBF kernel is chosen, and the penalty term C is set to 1. The inverse of the radius of influence of the support vector gamma is set to 0.1.
(2) MLP: Two hidden layers are used, both with 30 neurons. The momentum is 0.2, and the learning rate is 0.3.
(3) 1DCNN: The network consists of two convolutional layers (kernel size = 5), two maximum pooling layers (kernel size = 2), and a linear layer. The activation function is ReLU, and the optimizer is Adam.
(4) LSTM-RNN: The network contains two LSTM layers with 64 nodes in each layer.
(5) ResNet: The network uses the 18-layer ResNet model, as described in Ref. [38].

SMO-SVM, MLP, 1DCNN, and IVMD-RF achieved high accuracy from the results of single-speed data. The LSTM-RNN had the lowest accuracy, which shows its poor classification ability for non-time series. Compared to the 1H data set, the diagnostic accuracy of the 1B and 3H datasets decreased significantly due to the increased distance of the sensor location from the combustion chamber and valve. The vibration signal may be distorted or coupled with other disturbances when it is transmitted.

The next step is to use these methods to diagnose multiple speed conditions. All types of failure data for the first cylinder head (1H) at 700 rpm, 1300 rpm, 1600 rpm, 2000 rpm, and 2300 rpm were made into one dataset. Since there are no abnormal injection advance angle faults in the 700 rpm data, a total of 56 labeled categories of data are included. The diagnostic results are shown in Table 6. Compared to the single-speed dataset for the first cylinder head (1H), the accuracy of each method decreases to varying degrees as the number of failure types increases. The accuracy of SMO-SVM dropped the most. SMO-SVM method is suitable for single-speed data classification but not as effective as other methods for multi-speed and multi-class data. The proposed method still maintains high accuracy. The results show that IVMD-RF has advantages for multi-speed and multi-type fault diagnosis scenarios.
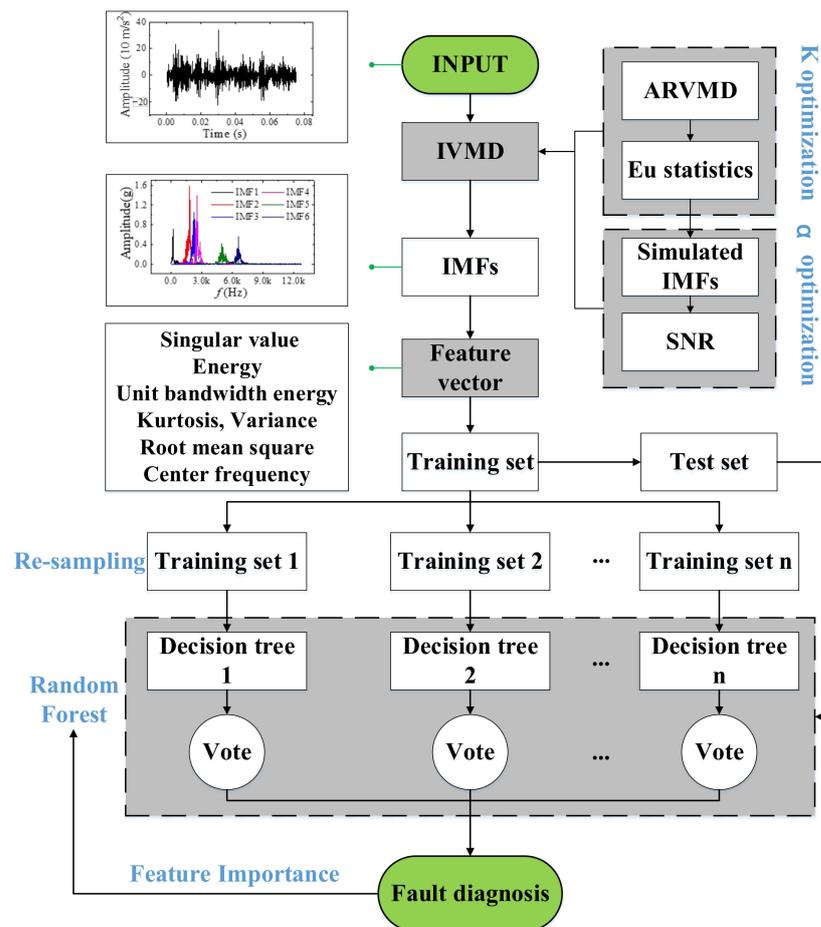
Figure 8. IVMD-RF fault detection process.

Table 6. Accuracy comparison of different algorithms (%).

Methods	1H-2000 rpm	1H-2300 rpm	1B-2300 rpm	3H-2300 rpm	1H-Multi-Speed
SMO-SVM	99.06	97.50	93.75	92.91	92.59
MLP	97.81	98.13	92.36	92.71	96.79
1DCNN	94.06	99.06	92.89	94.06	96.56
LSTM-RNN	59.06	75.16	78.12	68.75	56.23
ResNet	91.09	96.88	88.28	86.72	92.19
IVMD-RF	98.75	99.38	92.91	93.54	97.32

Note: "3H" represents the third cylinder head, "1B" represents the first cylinder block.

In addition, the proposed method requires less training time to achieve high accuracy. For comparison, all algorithms are run in the same environment (Python 3.8, Windows 11, Intel Core i7-10700 CPU @ 2.9 GHz), and the running time is recorded in Table 7 The results show that SMO-SVM has the highest training efficiency, followed closely by IVMD-RF. ResNet has the longest training time due to the deep network layers. Therefore, the proposed method has high efficiency and high accuracy. It is worth noting that deep learning may provide better diagnostic results for the original raw signal. However, the significant increase in data dimensionality leads to an increase in computation time and higher hardware requirements, which deviates from the purpose of this paper. No diagnostics were performed on the original raw data to keep the variables consistent.

Next, the 1300 rpm, 1600 rpm, 2000 rpm, and 2300 rpm data were mixed into one dataset. The data were labeled into 12 categories according to Table 4, regardless of the speed change. The cross-speed datasets include the 1H dataset, 1B dataset, and 3H dataset. The above methods are still used for classification, and the results are shown in Table 8. The results show that the accuracy of each algorithm has a certain drop compared to

Table 6, especially the SMO-SVM. The 1DCNN and IVMD-RF still maintain a relatively high accuracy rate. The overall accuracy of the 3H dataset is low because the sensor in the third cylinder head is far from the cylinder where some failures occurred. The proposed method still has some advantages over other algorithms. Table 8 also shows the classification precision, recall, and f1-score. These indicators are weighted averages, where the weights are determined by the proportion of each class sample distribution. The results show that the proposed method also performs well on these metrics. For the 3H dataset with relatively poor diagnostic results, Figure 9 shows the comparison of recall and precision of each algorithm for the 12 classes. The results showed low recall and accuracy for reduced valve clearance (condition II) and abnormal injection advance angles (condition IX to XII). The reason for the low precision and recall of Fault II is the slight increase in valve clearance and the long distance of the sensor from the cylinder where the fault occurred. Faults IX to XII, on the other hand, are due to small changes in injection advance angle, causing only minor differences in combustion conditions. The training efficiency of the proposed method is much higher than that of the deep learning method and slightly lower than that of SMO-SVM. Figure 10 shows the confusion matrix for the 1H dataset, indicating that most of the misclassified samples are data of the same type but with different failure levels, which proves the effectiveness of the proposed method.

Table 7. Comparison of training time of various algorithms (s).

Methods	1H-2000 rpm	1H-2300 rpm	1B-2300 rpm	3H-2300 rpm	1H-Multi-Speed
SMO-SVM	0.05	0.06	0.05	0.07	2.84
MLP	9.60	9.91	9.55	9.53	274.21
1DCNN	13.02	14.10	14.65	14.87	167.38
LSTM-RNN	17.85	18.21	20.26	20.48	224.95
ResNet	49.32	47.66	54.53	51.66	446.92
IVMD-RF	0.26	0.28	0.31	0.29	4.30

Note: "3H" represents the third cylinder head, "1B" represents the first cylinder block.

Table 8. Comparison of fault diagnosis results of each algorithm for cross-speed dataset.

/	Methods	SMO-SVM	MLP	1DCNN	LSTM-RNN	ResNet	IVMD-RF
First cylinder head Y-direction (1H)	Accuracy	0.81	0.92	0.93	0.68	0.87	0.97
	Precision	0.82	0.92	0.94	0.65	0.89	0.97
	Recall	0.81	0.92	0.92	0.63	0.88	0.96
	F1-score	0.82	0.92	0.92	0.62	0.88	0.97
	Time (s)	0.99	70.19	130.02	189.29	219.77	2.20
First cylinder block Y-direction (1B)	Accuracy	0.81	0.89	0.93	0.67	0.86	0.92
	Precision	0.81	0.89	0.94	0.75	0.87	0.92
	Recall	0.80	0.89	0.91	0.72	0.85	0.92
	F1-score	0.80	0.89	0.91	0.72	0.86	0.92
	Time (s)	1.13	70.69	145.99	192.59	224.21	2.74
Third cylinder head Y-direction (3H)	Accuracy	0.65	0.75	0.83	0.47	0.78	0.94
	Precision	0.66	0.76	0.83	0.42	0.72	0.93
	Recall	0.65	0.76	0.80	0.42	0.71	0.94
	F1-score	0.65	0.76	0.81	0.40	0.71	0.93
	Time (s)	1.03	70.28	143.31	203.47	221.34	2.92

The datasets with different training/testing ratios are set up for classification to verify the diagnostic effectiveness of various methods for the small sample case. Figure 11 shows each algorithm's accuracy and time consumption curves for the 1H dataset at different training test ratios (0.1 to 4). When the training test ratio <0.25, the accuracy of 1DCNN, RNN, and MLP significantly decrease, while SMO-SVM and IVMD-RF decrease more smoothly. When the training test ratio is 0.1, IVMD-RF has the highest accuracy of 88.77%. Figure 11b shows that the training efficiency of each algorithm increases as the training/test

ratio decreases. The efficiency of IVMD-RF and SMO-SVM remains higher than the other methods. The results illustrate the good diagnostic effect of the proposed method for small samples of single-sensor data. Figure 12 shows the accuracy and training loss curves when the three deep learning methods are applied to the 1H dataset. Figure 12 indicates that the 1DCNN has converged while the RNN clearly shows over-fitting, which is the reason for its low accuracy. Continuing to train ResNet may improve the accuracy, but the training efficiency is too low compared to other methods. Therefore, IVMD-RF has a high fault diagnosis accuracy and high efficiency for cross-speed data. It is worth noting that deep learning methods still have more advantages and potential when the amount of labeled data and computational resources are sufficient.
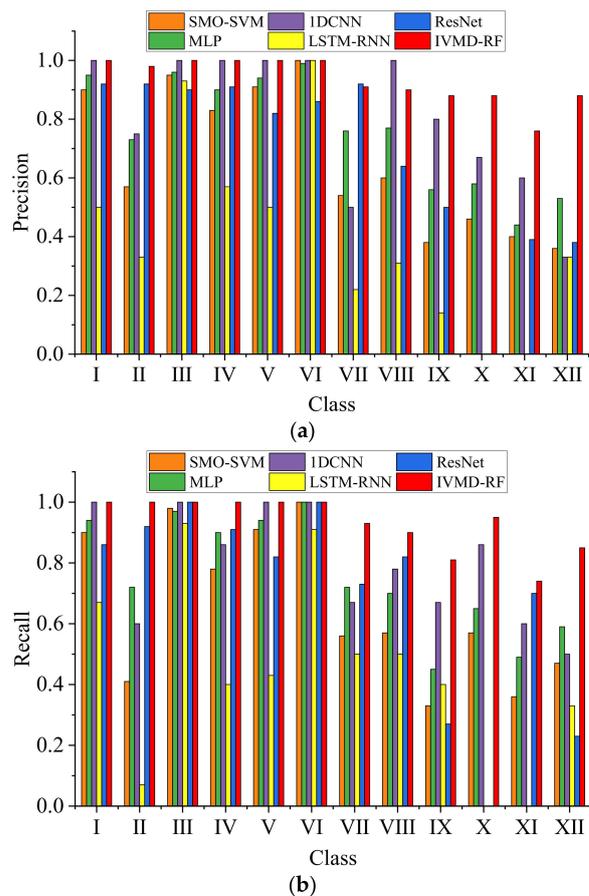

Figure 9. Precision and recall results of each algorithm for the 3H dataset. (**a**) Precision result. (**b**) Recall result.
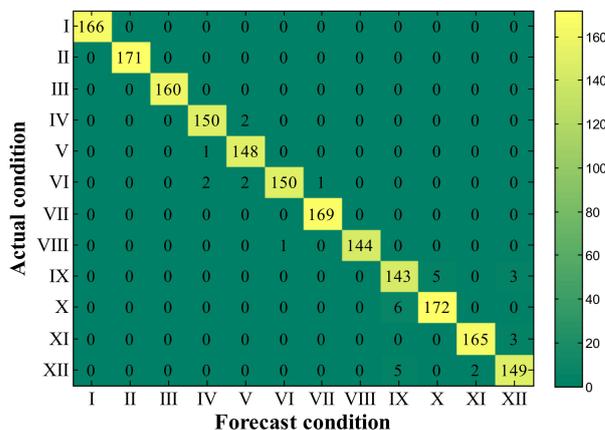

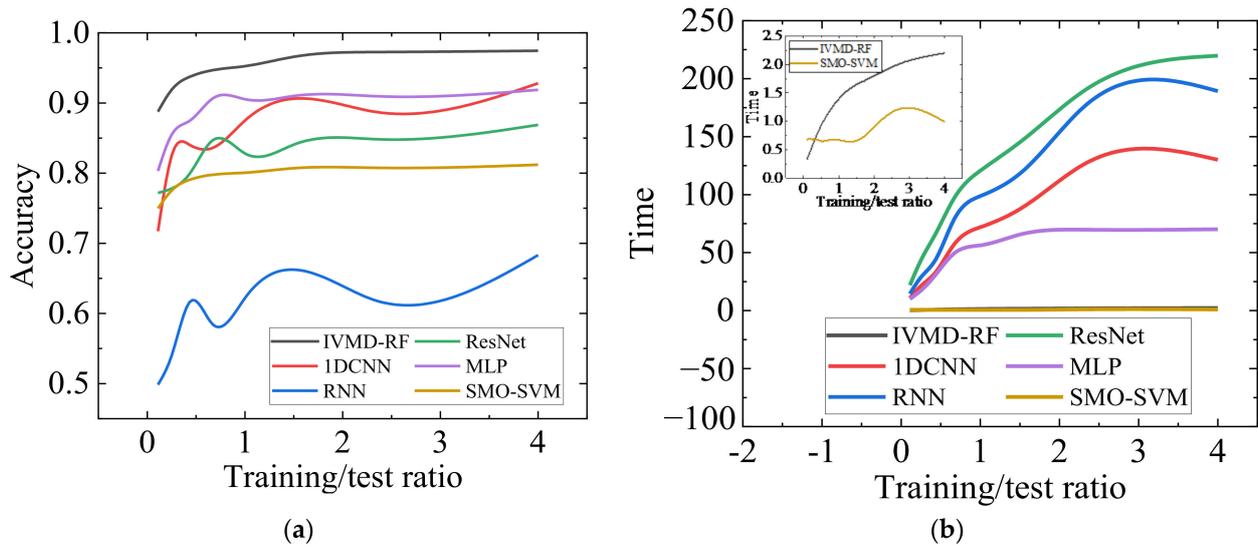**Figure 10.** Confusion matrix for 1H dataset of IVMD-RF.

Figure 11. Comparison of each algorithm with different training/testing ratios (0.1 to 4). (**a**) Comparison of accuracy. (**b**) Comparison of the training time.
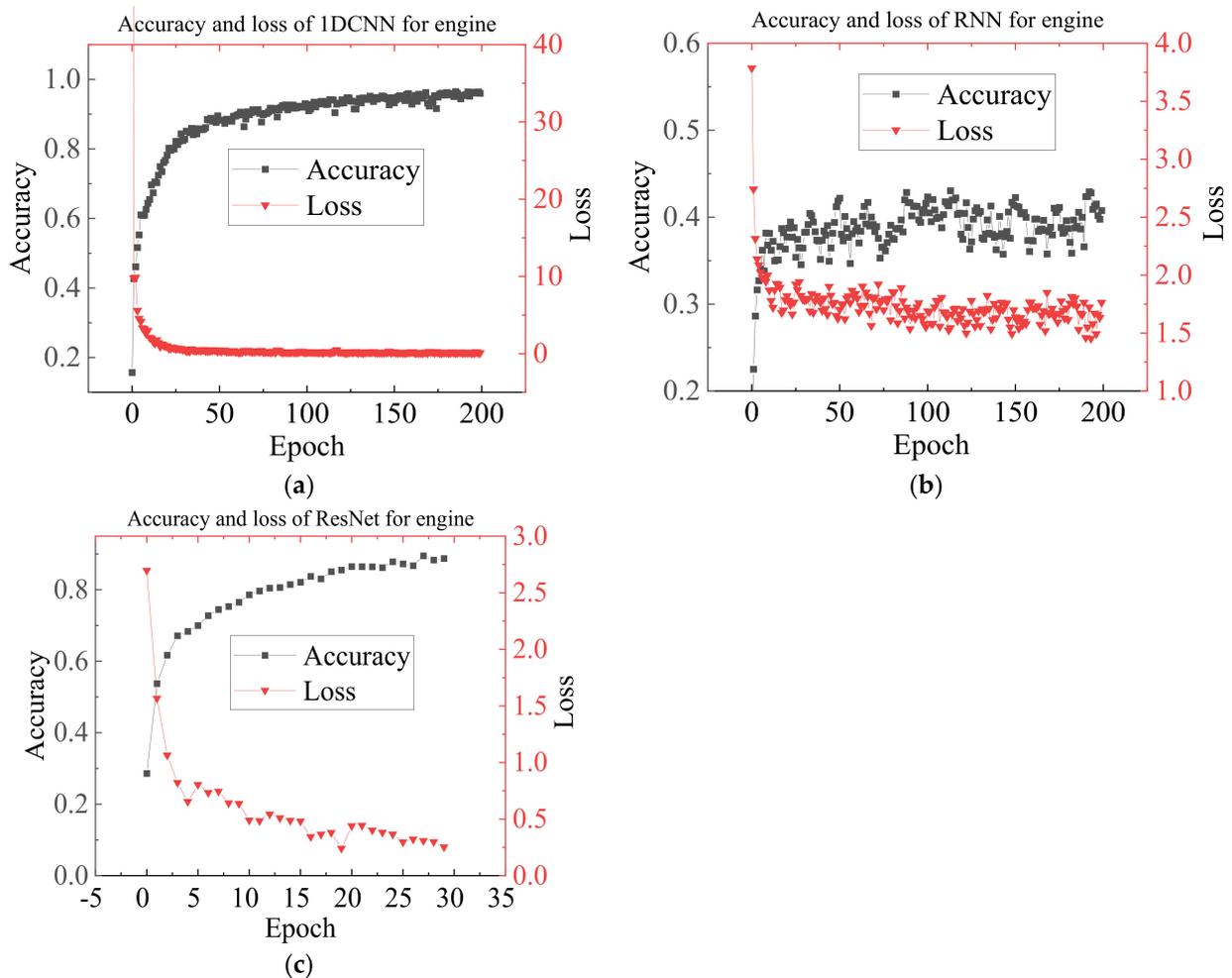

Figure 12. Accuracy and training loss of deep learning methods for 1H dataset. (**a**) Accuracy and training loss of 1DCNN. (**b**) Accuracy and training loss of LSTM-RNN. (**c**) Accuracy and training loss of ResNet.

The study of feature importance can further improve the performance of the method. Using the RF to rank the importance of features, the results for the 1H dataset are shown in Figure 13. Singular values, energy, and center frequency contributed more to the classification, followed by variance and unit bandwidth energy. However, kurtosis has little contribution to the classification results. Figure 13 shows that features with suffixes 1 and 4 contribute significantly to the classification, i.e., IMF1 and IMF4 contribute the most to the classification, followed by IMF5 and IMF6. For different working conditions, the difference in the body surface vibration is mainly reflected in the low-frequency (IMF1) and high-frequency components (IMF4~6). IMF2 and IMF3 have high energy but weak contribution. This conclusion is valuable for the study of unsupervised engine fault diagnosis. Figure 14 shows the impact of using different numbers of features in order of importance on training time and accuracy. Finally, we found an optimal point. When using fifteen features, it only takes 1.23 s to train and can achieve 97% diagnostic accuracy as marked in Figure 14. The selected fifteen categories of features are marked in Figure 13. Feature selection significantly improves training time with little change in accuracy.
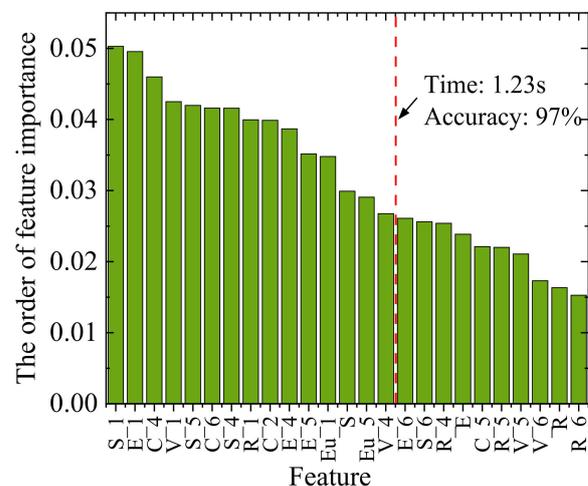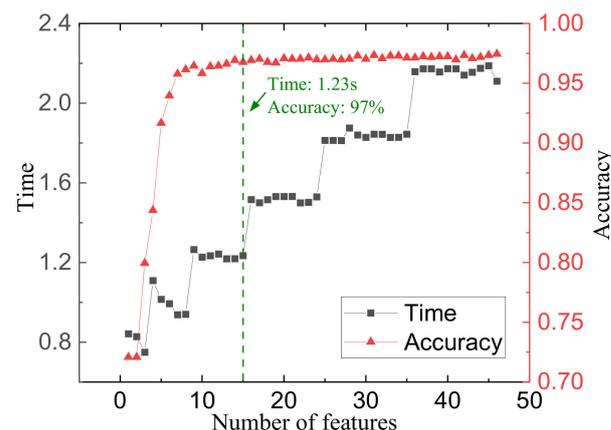

Figure 13. The order of feature importance.

Figure 14. The impact of changing the number of features used.

5.2. Case 2: Gasoline Engine Fault Diagnosis

To verify the effectiveness of the proposed method on different engines, the gasoline engine fault data will be diagnosed in the following. The fault data came from a two-cylinder, two-stroke gasoline engine with the specific engine parameters shown in Table 9. The sensor locations and coordinate system for the engine are shown in Figure 15. The data used are from the Y-direction of cylinder 1 and the X-direction of cylinder 2 (the two sensors connected by the white wire in Figure 15). Three common faults were simulated:

abnormal injection advance angle, abnormal air-fuel ratio, and misfire. Among them, the abnormal injection advance angle occurs in both cylinders, while the other two failures occur only in cylinder 1. The specific fault level settings are shown in Table 10. Use Roman numerals I through VII to indicate individual faults. Signals from 5000 rpm and 7000 rpm were collected. Each working condition contains 200 samples, each containing data from one working cycle.

Table 9. Parameters of gasoline engine.

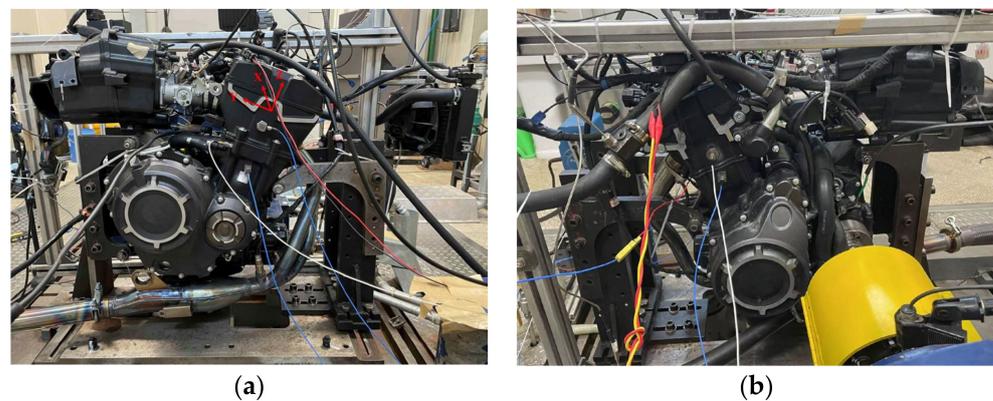| Items	Parameters
Displacement	0.294 L
Rated power/Rated speed	35.5 kW/8500 rpm
Maximum torque/Speed range	44.5 Nm/7000 rpm

(a)

(b)

Figure 15. Sensor positions and coordinate direction of the engine. (**a**) Position of sensor 1Y. (**b**) Position of sensor 2X.

Table 10. Fault type and degree parameter setting.

Mark	Injection Advance Angle	Air/Fuel Ratio	Misfire Rate
I	10 °CA	1	0
II	5 °CA	1	0
III	15 °CA	1	0
IV	10 °CA	1.1	0
V	10 °CA	1.2	0
VI	10 °CA	1	0.05
VII	10 °CA	1	0.1

Note: The shaded green marks the location of the faulty parameter.

The 5000 and 7000 rpm data were mixed to form the cross-speed dataset. Various algorithms diagnose the fault data of 1Y and 2X sensors separately. The settings of each algorithm are shown in Section 5.1. A comparison of the diagnostic results for each algorithm is shown in Table 11. The results show an overall decrease in the diagnostic accuracy of each algorithm due to the increase in signal noise as the two-cylinder, two-stroke engine vibrates more than the diesel engine. The higher speed is also one of the reasons. SMO-SVM is still the fastest, but its accuracy is low. The proposed method works best for fault diagnosis of 1Y data, and 1DCNN works best for 2X data. The difference in accuracy between the two is not significant. The proposed method is more efficient and suitable for low hardware conditions. The results show that IVMD-RF can be used for gasoline engine fault diagnosis.

Table 11. Comparison of fault diagnosis results of each algorithm for gasoline engine data.

	Methods	SMO-SVM	MLP	1DCNN	LSTM-RNN	ResNet	IVMD-RF
First cylinder head Y-direction (1Y)	Accuracy	0.67	0.75	0.82	0.52	0.78	0.85
	Precision	0.68	0.77	0.81	0.69	0.78	0.85
	Recall	0.67	0.75	0.82	0.51	0.79	0.84
	F1-score	0.67	0.72	0.82	0.52	0.78	0.84
	Time (s)	0.33	15.83	9.45	6.44	11.02	1.16
Second cylinder head X-direction (2X)	Accuracy	0.72	0.74	0.82	0.59	0.79	0.79
	Precision	0.72	0.74	0.82	0.58	0.77	0.78
	Recall	0.72	0.74	0.81	0.58	0.76	0.79
	F1-score	0.72	0.74	0.82	0.58	0.74	0.79
	Time (s)	0.25	15.80	9.61	6.51	11.05	1.12

6. Conclusions and Discussion

This article proposes an IVMD-RF for single-sensor multi-fault detection of the engine. In IVMD, the engine data spectral energy distribution is obtained through multiple decompositions and statistics. The alpha value was chosen based on the spectral distribution and the SNR. By presetting the center frequency and the optimal K and α values, the efficiency is improved, the mode aliasing is reduced, and the feature size is unified. The effectiveness of IVMD is proved by decomposing the engine signals. Seven types of attributes are calculated to form a feature group for IMFs, which is input into RF for classification. Compared with various machine learning and deep learning algorithms, it is proved that the proposed method has advantages in training efficiency and accuracy. Through the feature importance study, it is found that the high-frequency and low-frequency IMFs contribute more to the classification. Fifteen optimal features have been selected to improve the efficiency of RF. The IVMD-RF method has application prospects in engine single-sensor multi-class fault detection.

Author Contributions: Conceptualization, D.T. and F.B.; software, J.C.; validation, D.T., X.Y. and P.S.; writing—original draft preparation, D.T.; writing—review and editing, J.C. and X.B.; project administration, F.B.; funding acquisition, X.B. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are not publicly available at this time but may be obtained upon reasonable request from the authors.

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

References

1. Zhong, J.H.; Wong, P.K.; Yang, Z.X. Fault diagnosis of rotating machinery based on multiple probabilistic classifiers. *Mech. Syst. Signal Process.* **2018**, *108*, 99–114. [CrossRef]
2. Delvecchio, S.; Bonfiglio, P.; Pompoli, F. Vibro-acoustic condition monitoring of Internal Combustion Engines: A critical review of existing techniques. *Mech. Syst. Signal Process.* **2018**, *99*, 661–683. [CrossRef]
3. Goyal, D.; Pabla, B.S. The Vibration Monitoring Methods and Signal Processing Techniques for Structural Health Monitoring: A Review. *Arch. Comput. Methods Eng.* **2016**, *23*, 585–594. [CrossRef]
4. Ma, Y.L.; Cheng, J.S.; Wang, P.; Wang, J.; Yang, Y. A novel Lanczos quaternion singular spectrum analysis method and its application to bevel gear fault diagnosis with multi-channel signals. *Mech. Syst. Signal Process.* **2022**, *168*, 108679. [CrossRef]
5. Ribeiro, R.F.; Areias, I.A.D.; Campos, M.M.; Teixeira, C.E.; da Silva, L.E.B.; Gomes, G.F. Fault detection and diagnosis in electric motors using 1d convolutional neural networks with multi-channel vibration signals. *Measurement* **2022**, *190*, 110759.

6.   Bhowmik, B.; Panda, S.; Hazra, B.; Pakrashi, V. Feedback-driven error-corrected single-sensor analytics for real-time condition monitoring. *Int. J. Mech. Sci.* **2022**, *214*, 106898. [CrossRef]

7.   Ayati, M.; Shirazi, F.A.; Ansari-Rad, S.; Zabihihesari, A. Classification-Based Fuel Injection Fault Detection of a Trainset Diesel Engine Using Vibration Signature Analysis. *J. Dyn. Syst. Meas. Control* **2020**, *142*, 051003. [CrossRef]

8.   Wang, H.F.; Jiang, W.; Deng, X.Y.; Geng, J. A new method for fault detection of aero-engine based on isolation forest. *Measurement* **2021**, *185*, 110064. [CrossRef]

9.   Ellefsen, A.L.; Han, P.H.; Cheng, X.; Holmeset, F.T.; Æsøy, V.; Zhang, H.X. Online Fault Detection in Autonomous Ferries: Using Fault-Type Independent Spectral Anomaly Detection. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 8216–8225. [CrossRef]

10.   Liu, J.X.; Yang, L.D.; Xu, M.J.; Zhang, Q.; Yan, R.Q.; Chen, X.F. Model-based detection of soft faults using the smoothed residual for a control system. *Meas. Sci. Technol.* **2021**, *32*, 015107. [CrossRef]

11.   Wang, Z.J.; Zhang, J.J.; Jiang, Z.N.; Mao, Z.W.; Chang, K.; Wang, C.G. Quantitative misalignment detection method for diesel engine based on the average of shaft vibration and shaft shape characteristics. *Measurement* **2021**, *181*, 109527. [CrossRef]

12.   Naderi, E.; Khorasani, K. Data-driven fault detection, isolation and estimation of aircraft gas turbine engine actuator and sensors. *Mech. Syst. Signal Process.* **2018**, *100*, 415–438. [CrossRef]

13.   Chao, M.A.; Adey, B.T.; Fink, O. Implicit supervision for fault detection and segmentation of emerging fault types with Deep Variational Autoencoders. *Neurocomputing* **2021**, *454*, 324–338. [CrossRef]

14.   Elmasry, W.; Wadi, M. EDLA-EFDS: A Novel Ensemble Deep Learning Approach for Electrical Fault Detection Systems. *Electr. Power Syst. Res.* **2022**, *207*, 107834. [CrossRef]

15.   Feng, Y.; Liu, Z.J.; Chen, J.L.; Lv, H.X.; Wang, J.; Zhang, X.W. Unsupervised Multimodal Anomaly Detection with Missing Sources for Liquid Rocket Engine. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**. [CrossRef]

16.   Bi, F.R.; Ma, T.; Wang, X. Development of a novel knock characteristic detection method for gasoline engines based on wavelet-denoising and EMD decomposition. *Mech. Syst. Signal Process.* **2019**, *117*, 517–536.

17.   Wang, X.; Cai, Y.P.; Li, A.H.; Zhang, W.; Yue, Y.J.; Ming, A.B. Intelligent fault diagnosis of diesel engine via adaptive VMD-Rihaczek distribution and graph regularized bi-directional NMF. *Measurement* **2021**, *172*, 108823. [CrossRef]

18.   Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [CrossRef]

19.   Huang, S.L.; Sun, H.Y.; Wang, S.; Qu, K.F.; Zhao, W.; Peng, L.S. SSWT and VMD Linked Mode Identification and Time-of-Flight Extraction of Denoised SH Guided Waves. *IEEE Sens. J.* **2021**, *21*, 14709–14717. [CrossRef]

20.   Miao, Y.H.; Zhao, M.; Lin, J. Identification of mechanical compound-fault based on the improved parameter-adaptive variational mode decomposition. *ISA Trans.* **2019**, *84*, 82–95. [CrossRef]

21.   Nazari, M.; Sakhaei, S.M. Successive variational mode decomposition. *Signal Process.* **2020**, *174*, 107610. [CrossRef]

22.   Lian, J.J.; Liu, Z.; Wang, H.J.; Dong, X.F. Adaptive variational mode decomposition method for signal processing based on mode characteristic. *Mech. Syst. Signal Process.* **2018**, *107*, 53–77. [CrossRef]

23.   Mei, L.; Li, S.Y.; Zhang, C.; Han, M.X. Adaptive Signal Enhancement Based on Improved VMD-SVD for Leak Location in Water-Supply Pipeline. *IEEE Sens. J.* **2021**, *21*, 24601–24612. [CrossRef]

24.   Li, J.M.; Wang, H.; Zhang, J.F.; Yao, X.F.; Zhang, Y.G. Impact fault detection of gearbox based on variational mode decomposition and coupled underdamped stochastic resonance. *ISA Trans.* **2019**, *95*, 320–329. [CrossRef]

25.   Diao, X.; Jiang, J.C.; Shen, G.D.; Chi, Z.Z.; Wang, Z.R.; Ni, L.; Mebarki, A.; Bian, H.T.; Hao, Y.M. An improved variational mode decomposition method based on particle swarm optimization for leak detection of liquid pipelines. *Mech. Syst. Signal Process.* **2020**, *143*, 106787. [CrossRef]

26.   Sahani, M.; Dash, P.K. Deep Convolutional Stack Autoencoder of Process Adaptive VMD Data with Robust Multikernel RVFLN for Power Quality Events Recognition. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 9001912. [CrossRef]

27.   Luo, X.J.; Fong, K.F.; Sun, Y.J.; Leung, M.K.H. Development of clustering-based sensor fault detection and diagnosis strategy for chilled water system. *Energy Build.* **2019**, *186*, 17–36. [CrossRef]

28.   Ni, Y.R.; Zeng, X.J.; Liu, Z.L.; Yu, K.; Xu, P.B.; Wang, Z.; Zhuo, C.; Huang, Y. Faulty feeder detection of single phase-to-ground fault for distribution networks based on improved K-means power angle clustering analysis. *Int. J. Electr. Power Energy Syst.* **2022**, *142*, 108252. [CrossRef]

29.   Shahid, S.M.; Ko, S.; Kwon, S. Real-time abnormality detection and classification in diesel engine operations with convolutional neural network. *Expert Syst. Appl.* **2022**, *192*, 116233. [CrossRef]

30.   Zhang, B.; Zhang, S.H.; Li, W.H. Bearing performance degradation assessment using long short-term memory recurrent network. *Comput. Ind.* **2019**, *106*, 14–29. [CrossRef]

31.   Lee, J.; Lee, Y.C.; Kim, J.T. Fault detection based on one-class deep learning for manufacturing applications limited to an imbalanced database. *J. Manuf. Syst.* **2020**, *57*, 357–366. [CrossRef]

32.   Li, H.F.; Hu, G.Z.; Li, J.Q.; Zhou, M.C. Intelligent Fault Diagnosis for Large-Scale Rotating Machines Using Binarized Deep Neural Networks and Random Forests. *IEEE Trans. Autom. Sci. Eng.* **2022**, *19*, 1109–1119. [CrossRef]

33.   Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

34.   Nahim, H.M.; Younes, R.; Shraim, H.; Ouladsine, M. Oriented review to potential simulator for faults modeling in diesel engine. *J. Mar. Sci. Technol.* **2016**, *21*, 533–551. [CrossRef]

35.   Yu, S.W.; Ma, J.W. Complex Variational Mode Decomposition for Slop-Preserving Denoising. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 586–597. [CrossRef]

36. Tang, D.J.; Bi, F.R.; Lin, J.W.; Li, X.; Yang, X.; Bi, X.Y. Adaptive Recursive Variational Mode Decomposition for Multiple Engine Faults Detection. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3513111. [CrossRef]
37. The Bearing Data Source. Available online: https://csegroups.case.edu/bearingdatacenter (accessed on 10 April 2018).
38. He, K.M.; Zhang, X.Y.; Ren, S.Q.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 27–30 June 2016; pp. 770–778.