processes

MDPI

Article

Unbalanced Few-Shot Diesel Engine Intelligent Fault Diagnosis Based on Successive Multivariate Variational Mode Decomposition and Optimized Deep Hybrid Kernel Extreme Learning Machine

Zichang Liu [1], Zhining Li [1], Yunqiang Zhang [1], Rongcai Wang [2], Danping Wang [1] and Yongsheng Bai [1,*]

[1] Shijiazhuang Campus, Army Engineering University of PLA, Shijiazhuang 050003, China; zc_liu1997@aeu.edu.cn (Z.L.); lgdcllzn@163.com (Z.L.); zyqgc1987@163.com (Y.Z.); wangdanping_2010@163.com (D.W.)
[2] No. 32181 Unit of PLA, Xi'an 710000, China; asd528866@126.com
* Correspondence: ys_bai1981@163.com; Tel.: +86-0311-87994617

Abstract

To address the issues of sample ratio imbalance and scarcity of available failure data in unbalanced few-shot scenarios—where models tend to favor the class with more samples and are prone to overfitting, thereby affecting the accuracy of fault diagnosis—this manuscript proposes an intelligent fault diagnosis method for diesel engines under unbalanced few-shot conditions, based on Successive Multivariate Variational Mode Decomposition (SMVMD) and the optimized Deep Hybrid Kernel Extreme Learning Machine (DHKELM). By processing the data samples using the SMVMD method, the components of the Intrinsic Mode Functions (IMF) are obtained; based on the principle of minimizing envelope entropy, the most valuable component within the sample is selected, and its spectral amplitude is used as a fault feature to improve sample quality at the data level. To optimize the feature extraction process at the model level, the Improved Sparrow Search Algorithm (ISSA) is proposed to optimize the DHKELM fault diagnosis model. The proposed ISSA is used to optimize the hyperparameters of the DHKELM model; after normalizing the fault feature samples, they are used as model inputs to perform diesel engine fault diagnosis. Case validation and comparative analysis show that, under conditions of unbalanced few-shot datasets, the proposed method can accurately and intelligently distinguish between different types of diesel engine failures.

Keywords: successive multivariate variational mode decomposition; improved sparrow search algorithm; deep hybrid kernel extreme learning machine; unbalanced few-shot; diesel engine; fault diagnosis

Check for updates

Academic Editor: José Alfredo Hernández Pérez

Received: 16 September 2026
Revised: 12 October 2026
Accepted: 3 November 2026
Published: 6 November 2026

Copyright: © 2026 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.

1. Introduction

As the power source for armored vehicles, tanks, self-propelled artillery, and other military equipment, diesel engines serve as the "heart" of these systems. With their high thermal efficiency and high torque output, they are widely used, and their operational condition is critical to the combat effectiveness of the equipment. However, due to their complex structure and harsh operating environments, which place high demands on the strength and stiffness of their components, these engines are prone to frequent malfunctions. According to statistics, diesel engine failures account for approximately 42% of all failures in self-propelled artillery chassis systems; among failure incidents involving tanks, mechanical

failures account for 35%, of which 48% are caused by diesel engine failures. To prevent these situations, the traditional approach typically relies on professional maintenance personnel using their experience or thermal parameter threshold alarms—that is, analyzing based on physical sensory observations or determining whether monitored parameters such as pressure, temperature, and rotational speed have reached preset thresholds—to manually identify the location of diesel engine failures. Although this approach is straightforward, it struggles to distinguish between certain similar diesel engine faults and has difficulty pinpointing the exact location of a fault. It suffers from slow identification, low accuracy, and high costs, and is no longer sufficient to meet the needs for optimized diesel engine maintenance and intelligent support. Failure to accurately and efficiently pinpoint the location of a diesel engine malfunction will not only hinder mission completion due to reduced engine performance but may even trigger a chain reaction leading to serious safety incidents. Therefore, to meet the demands of intelligent and information-based warfare, there is an urgent need to develop advanced diesel engine fault diagnosis methods to improve the engine's testability, maintainability, serviceability, and reliability, and to achieve standardized, precise, scientific, and intelligent fault diagnosis.

With the rapid development of sensor technology, the continuous improvement of data processing techniques, and the rise in artificial intelligence, data-driven intelligent methods have become the mainstream approach to fault diagnosis. Data-driven intelligent fault diagnosis methods do not require complex steps such as establishing physical or mathematical models for diesel engines. Instead, they rely on data collected by sensors that reflect the engine's operating status. By processing, analyzing, and mining the raw data as needed, these methods extract valuable knowledge and information embedded within the data, thereby endowing the data with "intelligence," and use the constructed models to explore solutions to problems. Compared with traditional methods, data-driven intelligent approaches utilize information processing technologies and intelligent algorithms to gain a deep understanding of maintenance data. This enables the intelligent diagnosis of diesel engine faults, reduces reliance on the professional expertise of maintenance personnel, minimizes uncertainty in the decision-making process, and effectively improves the efficiency and effectiveness of diesel engine maintenance. With their high flexibility and accuracy, these methods hold great potential for widespread application and adoption.

Generally speaking, data-driven diesel engine fault diagnosis requires sufficient data to support it. However, once certain faults occur in a diesel engine, the engine must be shut down as quickly as possible; otherwise, irreversible damage to the engine itself may result. Additionally, data sensitivity, security, or privacy concerns often lead to strict restrictions on data access, transmission, and sharing, resulting in the collection of only a very small amount of fault data. To address the issue of insufficient training data, commonly used methods can be categorized into data-based and model-based approaches. Data-based methods focus on improving data quality or expanding the dataset, while model-based methods aim to optimize classifier performance or enhance the fault feature extraction process. For example, Generative Adversarial Network (GAN) can be used to generate data similar to the original data to augment the available dataset, or transfer learning can be employed to transfer a pre-trained model from a related domain to the target domain. Huang Meng et al. [1] introduced gradient penalization and Wasserstein distance into an existing auxiliary classification GAN model. The improved model was used to generate synthetic data closely resembling the distribution of real data in the training set. The data-augmented feature dataset was then input into a one-dimensional Convolutional Neural Network (CNN) for training, enabling diesel engine fault diagnosis under conditions of scarce training data. Li et al. [2] proposed a diesel engine fault diagnosis method based on finite element simulation and deep convolutional GAN. They used the finite element method

to generate vibration data for diesel engines under different fault types and employed a deep convolutional GAN to reduce the domain gap between the simulated data and the measured data. Li Liangyu et al. [3] employed a least-squares GAN to generate training data, addressing the issue of insufficient diesel engine fault samples. After processing the training data using EEMD, they fed it into a diesel engine fault diagnosis model—which utilizes an adaptive random testing algorithm to optimize a Back Propagation (BP) neural network—to locate the fault position. To address the diagnosis of composite faults in diesel engine fuel injection systems under small-sample conditions, Cui et al. [4] proposed a diesel engine fault diagnosis method based on GAN and transfer learning. Using a convolutional attention network as the foundation, the GAN was employed to augment the original data, with the augmented data serving as the source data for transfer learning. Since data augmentation may adversely affect the original fault feature information in signals when processing one-dimensional vibration signals, Chen et al. [5] proposed a self-supervised learning framework without data augmentation for diesel engine fault diagnosis, namely domain-adaptive variance-invariant covariance regularization. This framework uses a loss function comprising three terms—variance, invariance, and covariance—to extract useful features from unlabeled data. Although GAN possess powerful sample generation capabilities, they rely on a certain scale of original samples for training. In the problem of small-sample fault diagnosis for diesel engines, the available training samples are insufficient to train an effective generator, and the diversity of generated samples is poor. Expanding the dataset with additional samples generated by this generator not only fails to improve the model's classification accuracy but may also introduce noise, leading to a decline in performance. Among commonly used data-based methods, aside from GAN, the effectiveness of transformation-based methods depends on the quality of the original signal; overlapping sampling produces multiple duplicate samples; TrAdaboost requires an auxiliary dataset that shares a certain degree of similarity with the original data; Variational Auto Encoder (VAE)-generated samples lack sufficient fault information; and synthetic minority oversampling techniques struggle to address the issue of data distribution marginalization.

Transfer learning primarily includes methods such as feature transfer, parameter transfer, and domain adaptation. To address the issues of poor generalization and susceptibility to overfitting in traditional neural network models when fault data is insufficient, Zhang et al. [6] developed a wide-kernel convolutional Long Short-Term Memory (LSTM) neural network model and proposed a deep transfer learning-based method for small-sample fault diagnosis in diesel engines. Compared with traditional deep neural networks (DNNs), this method demonstrated superior performance in fault diagnosis tasks across different equipment domains and fault domains. Xu Jia [7] proposed a transfer learning-based method for diagnosing diesel engine faults under varying operating conditions, where feature transfer from the source domain to the target domain under varying conditions is achieved through joint alignment of statistical distributions and geometric space. Xiong et al. [8] proposed a method for identifying variable operating conditions in diesel engines based on stacked autoencoders (AE) and feature transfer learning, which is of great significance for diesel engine fault diagnosis. Jia Zheyu [9] proposed a transfer learning CNN architecture based on VGG16 for diesel engine fault diagnosis. Bai et al. [10] proposed a diesel engine fault diagnosis method based on an optimized Variational Mode Decomposition (VMD), Short-Time Fourier Transform (STFT), and a ResNet18 transfer learning model. Building upon a pre-trained ResNet18 neural network model, they transferred the parameters of the pre-trained model to the target domain to identify diesel engine fault types. To address the issue of reduced diagnostic accuracy caused by inconsistent feature distributions under varying operating conditions, Zhao Zhijian [11] designed transfer learn-

ing models based on the ResNet50 neural network, deep adaptation networks, and domain adaptation, respectively, for diagnosing diesel engine misfires and valve clearance faults under different operating conditions. It should be noted that current transfer learning models still require massive amounts of data for training, and their performance depends on the correlation between source-domain and target-domain data; performance is poor when operating conditions differ significantly [12].

In addition to the methods mentioned above, researchers have also proposed other deep learning-based approaches to address the problem of fault diagnosis in diesel engines with limited data. To address the issue of low diagnostic accuracy for diesel engine valve clearance faults when the number of fault samples is insufficient, Wang Lijia et al. [13] used a Markov transition field to process vibration signals and obtain time–frequency spectra. They proposed an improved Gaussian prototype network to extract feature information from the time–frequency spectra and derive fault diagnosis results. Liao Shulang et al. [14] introduced Graph Convolutional Network (GCN) into small-sample fault diagnosis for diesel engines to address the issue of overfitting. The vibration signals were converted into graph data using a constructed adjacency matrix, and the Particle Swarm Optimization (PSO) algorithm was employed to optimize the hyperparameters of the improved GCN model, thereby enhancing the model's generalization ability. After training and optimization, the model's weight matrix was refined, and fault diagnosis results were output. Liu et al. [15] proposed a method for small-sample fault diagnosis of diesel engines based on the STFT and a proxy model-optimized CNN. Addressing the issue of suboptimal accuracy in traditional methods for diagnosing diesel engine faults with limited data, He Pengfei et al. [16] proposed a method for diagnosing diesel engine valve clearance faults based on the Continuous Wavelet Transform (CWT) and model-free learning. The results indicate that this method offers greater practical value in terms of accuracy compared with traditional deep learning methods such as CNNs. However, meta-learning suffers from high computational costs and places high demands on the quality and quantity of data from past tasks. Metric learning involves learning a metric space by optimizing a loss function to minimize the distance between samples of the same class as much as possible and maximize the distance between samples of different classes as much as possible. However, when applying metric learning to the problem of small-sample fault diagnosis in diesel engines, its performance is significantly influenced by the sampling strategy, the metric function, and prior knowledge.

Due to the complexity of operating mechanisms and mechanical structures, some diesel engine failures occur sporadically, resulting in a severe imbalance between the number of normal and fault data samples. To address the decline in diesel engine fault diagnosis performance under unbalanced datasets, Wang et al. [17] proposed a GCN-based diesel engine fault diagnosis method utilizing distance and probability topological graph models. This approach transforms the acquired condition monitoring dataset into distance and probability topological graphs and incorporates graph learning to identify correlations between adjacent samples and extract the features of the samples themselves.

In summary, although data-driven intelligent fault diagnosis methods for diesel engines have achieved breakthroughs in certain areas, they remain in the early stages of exploratory research and face challenges that urgently need to be addressed. For example, current research typically focuses only on diesel engine fault diagnosis methods for small sample sizes. Furthermore, few scholars have investigated diesel engine fault diagnosis methods for unbalanced datasets, and even fewer have simultaneously considered the impact of unbalanced datasets and insufficient training samples on the accuracy of diesel engine fault diagnosis. Furthermore, diesel engine fault diagnosis under conditions of small, unbalanced samples currently faces challenges such as a decline in diagnostic accuracy for classes with

few samples, a bias in diagnostic results toward classes with more samples, and a tendency for classification models to overfit due to insufficient parameter optimization when sample sizes are inadequate.

In light of this, this manuscript focuses on diesel engines—the primary power source for military equipment—as its research subject. Guided by the practical challenges and real-world needs in diesel engine fault diagnosis, and with the goal of achieving accurate and intelligent fault diagnosis for diesel engines, this study targets maintenance and support personnel as its end users. It proposes an intelligent fault diagnosis method for diesel engines under unbalanced few-shot conditions based on SMVMD and optimized DHKELM, thereby providing new insights for the practical application of intelligent diesel engine fault diagnosis. First, to improve sample quality at the data level, an SMVMD-based diesel engine vibration signal decomposition method is employed to process the raw signal into individual IMF components. The envelope entropy value of each component is calculated, and the component with the smallest value is selected as the most valuable component. The spectral amplitude of this component is then used as the fault feature sample to obtain the feature dataset. Second, to improve the feature extraction process at the model level, an ISSA-optimized DHKELM fault diagnosis model is proposed. By utilizing the proposed ISSA to optimize the hyperparameters in the DHKELM, the model is enabled to fully leverage its feature information mining capabilities. Next, the DHKELM is fed the normalized fault feature samples. Through learning and training with these features, the DHKELM enables diesel engine fault diagnosis under unbalanced few-shot conditions. Finally, the performance of the proposed fault diagnosis method is validated using measured test data.

The main contributions and innovations of this manuscript are as follows:

(1) To provide data support for research on diesel engine fault diagnosis methods, a data acquisition scheme for diesel engine condition monitoring was designed; data acquisition experiments were conducted; a relevant dataset was established; and the characteristics of the time-domain waveforms and spectra of the vibration signals in the dataset were analyzed.

(2) To address the issues with current vibration signal decomposition methods—including modal aliasing, endpoint effects, low decomposition accuracy, and poor computational efficiency—which adversely affect fault diagnosis performance under unbalanced few-shot conditions, this study investigates a diesel engine vibration signal decomposition method based on SMVMD and applies this method to the field of diesel engine fault diagnosis.

(3) To optimize the feature extraction process at the model level and improve the accuracy of fault diagnosis results under unbalanced few-shot conditions, an ISSA-optimized DHKELM fault diagnosis model was proposed. Using the proposed ISSA, the hyperparameters of the DHKELM model were optimized to fully enhance the model's ability to extract feature information.

(4) The proposed fault diagnosis method addresses the issues of sample class imbalance and scarcity of available fault data, as well as the tendency of models to favor the class with more samples and to overfit, which affects fault diagnosis accuracy. It plays a positive role in improving the accuracy and efficiency of diesel engine maintenance and reducing unplanned downtime.

The remainder of this manuscript is organized as follows: (1) Section 2 designs a data acquisition scheme for diesel engine condition monitoring and conducts data acquisition experiments. (2) Section 3 presents the technical approach and specific steps of the proposed intelligent fault diagnosis method for diesel engines under unbalanced few-shot conditions. (3) Section 4 elaborates on the technical principles of the proposed method, including the basic principles of the SMVMD method and the construction of the ISSA-optimized

DHKELM model. (4) Section 5 validates the performance of the proposed intelligent diesel engine fault diagnosis method using experimental data. (5) Section 6 summarizes the research findings of this manuscript and outlines potential future research directions.

2. Diesel Engine Condition Monitoring Data Acquisition Scheme Design and Experimental Implementation

The condition monitoring data collected forms the foundation and is critical to achieving intelligent fault diagnosis for diesel engines. To ensure the scientific rigor of diesel engine condition monitoring data collection, a technical approach for data acquisition was established. Following this technical approach, a diesel engine condition monitoring test bench was constructed. Through pre-set fault tests, the necessary condition monitoring data was collected under two different types of pre-set typical faults in the engine's systems, and a relevant dataset was established. The characteristics of the time-domain waveforms and frequency spectra of the obtained vibration signals were then analyzed.

2.1. Diesel Engine Condition Monitoring Data Acquisition Scheme

Based on the laboratory's existing capabilities and referring to current research findings in the field of data acquisition, a technical approach for data acquisition in diesel engine condition monitoring has been developed, as shown in Figure 1, providing a solution for data acquisition in the condition monitoring of complex reciprocating machinery such as diesel engines.

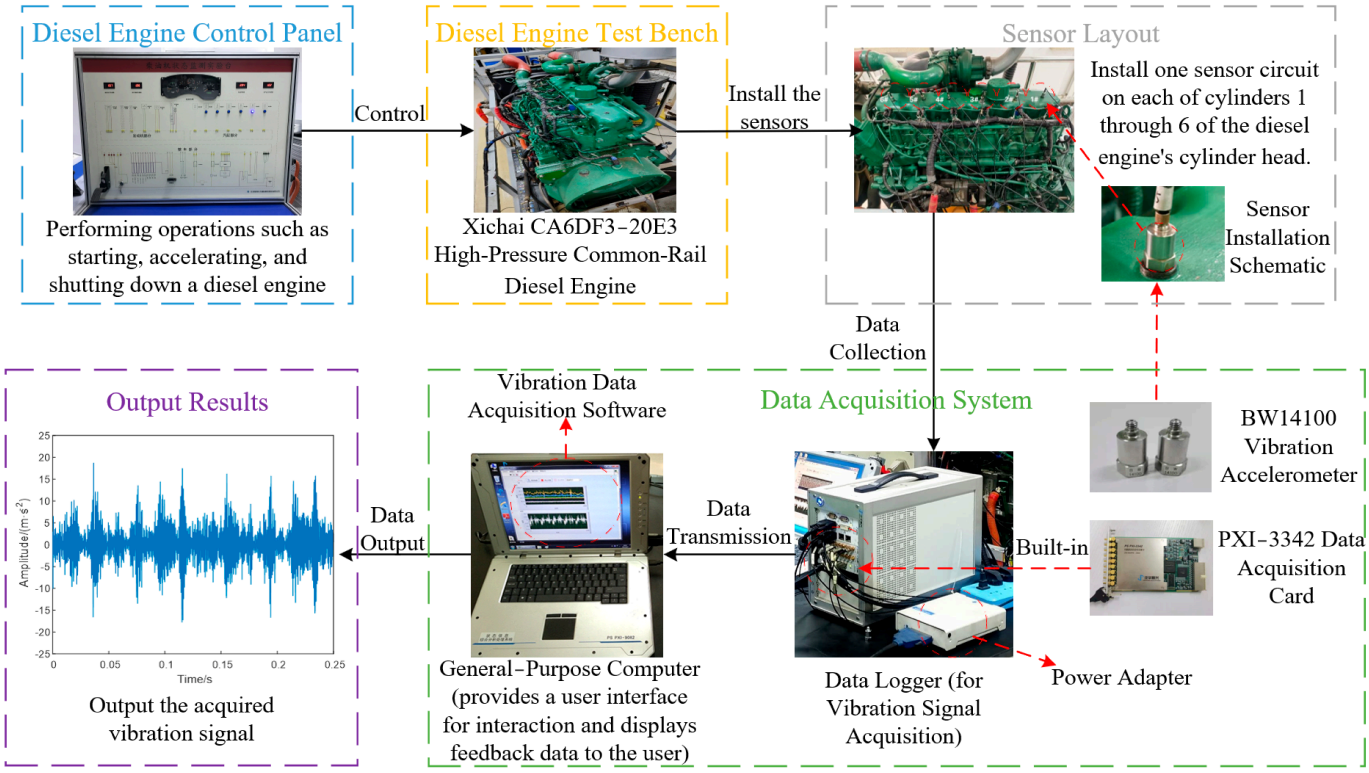

Figure 1. Diesel engine condition monitoring data acquisition technology route.

The main steps for collecting diesel engine condition monitoring data are as follows:

Step 1: Use the diesel engine control panel to preliminarily monitor certain operational parameters during engine operation, make a rough assessment of the engine's current operating status, and control routine engine operations, including ignition, acceleration, deceleration, and shutdown.

Step 2: Select the CA6DF3-20E3 diesel engine manufactured by Wuxi Diesel Engine Plant of FAW Jiefang Automobile Co., Ltd, Wuxi City, Jiangsu Province, China (hereinafter referred to as "Xichai"). Set up a diesel engine test bench in a laboratory environment that meets the functional and technical requirements for condition monitoring data acquisition testing. This test bench will be used to conduct various tests with different typical faults pre-set.

Step 3: To ensure the feasibility of the tests, a sensor layout was determined: one sensor was installed on each of the cylinder heads of cylinders 1 through 6 to capture the required diesel engine vibration data. The sampling frequency was set to 20 kHz, with each data acquisition lasting 12 s and a 30-s interval between acquisitions, resulting in a total of 300 data sets.

Step 4: Connect the sensors to the data acquisition system via wired connections. The data acquisition system is primarily used to acquire vibration signals, provide a user interface, and display feedback data to the user. It consists of piezoelectric vibration accelerometers (Model BW14100, manufactured by Shanghai Beizhi Co., Ltd. in Shanghai, China), a comprehensive condition monitoring system (PXI-9108 data acquisition chassis and PXI-9082 processing computer, manufactured by Beijing Fanhua Hengxing Technology Co., Ltd. in Beijing, China), a data acquisition card (Model PXI-3342, manufactured by Beijing Fanhua Hengxing Technology Co., Ltd. in China), and vibration data acquisition software (written in LabVIEW, the version used is Q3 2023).

Step 5: Save the acquired diesel engine cylinder head vibration signals, output the results, and create the corresponding dataset.

Based on Figure 1, a diesel engine condition monitoring test bench, as shown in Figure 2, was constructed in the laboratory. Using this test bench to simulate various typical diesel engine faults, data acquisition tests for diesel engine condition monitoring were conducted.

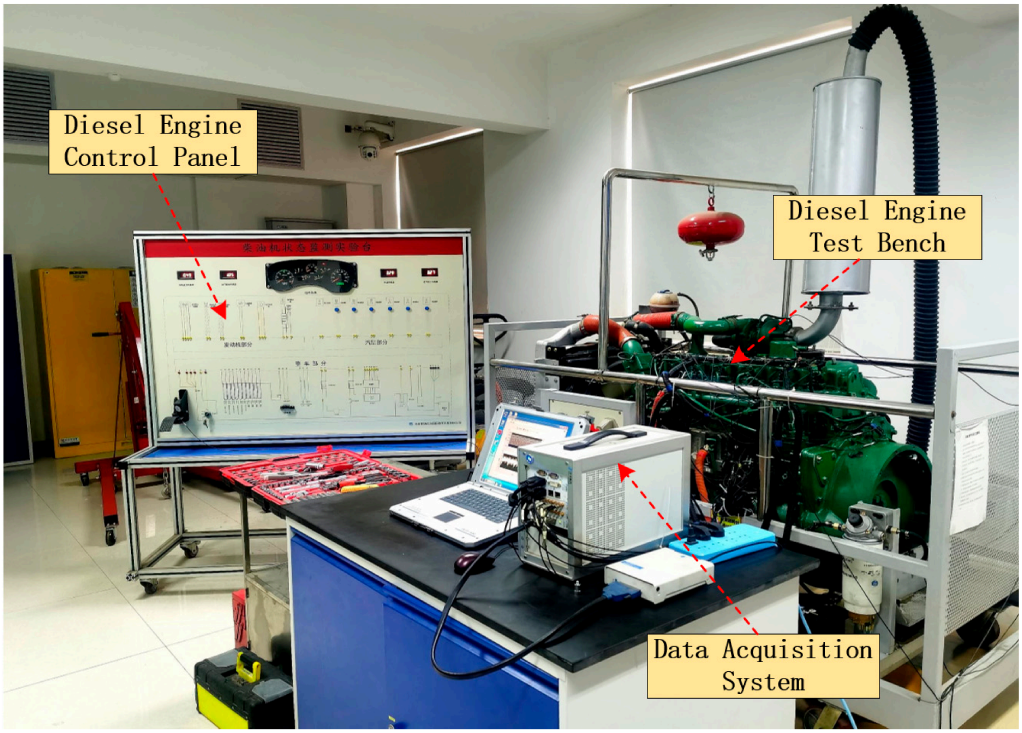

Figure 2. Diesel engine condition monitoring test bench.

The diesel engine control panel provides real-time display and control of dynamic operation. In line with the research background of this manuscript, an inline 6-cylinder,

4-stroke, water-cooled diesel engine developed by Xichai was selected as the subject of this experiment; its main technical parameters are shown in Table 1. The test bench is equipped with vibration-damping capabilities.

Table 1. Main technical parameters of diesel engine.

Category	Parameters	Category	Parameters
Displacement	6.74 L	Intake Configuration	Turbocharged with intercooler
Smoke Level	≤2.0 FSN	Maximum Torque	760
Dimensions	1330 × 970 × 1005 mm	Rated RPM	2300 r/min
Noise Level	≤95 dB(A)	Maximum Horsepower	200 horsepower
Net Weight	700 kg	Rated Power	155 kW
Number of Valves	2 valves per cylinder	Ignition Sequence	1-5-3-6-2-4
Compression Ratio	17.4	Compatibility Range	Medium- to heavy-duty trucks over 12 metric tons
Net Power	147 kW	Bore × Stroke	107 × 125 mm
Emissions Standards	China III	Minimum Fuel Consumption at Full Load	205 g/kW·h

The diesel engine test bench, following completion of its development and commissioning, meets the functional and technical requirements for data acquisition tests related to diesel engine condition monitoring and is suitable for conducting various tests on typical preset diesel engine fault types.

This test requires the collection of diesel engine status data. To ensure the test's feasibility, the sensor layout shown in Figure 3 was adopted. The diesel engine cylinder head was selected as the test point for collecting vibration signals. A total of six sensors were installed on the cylinder head; sensors 1 through 6 were mounted vertically on the cylinder heads of cylinders 1 through 6, respectively, and each sensor was secured using an adhesive mounting method (with adhesive applied to the base).

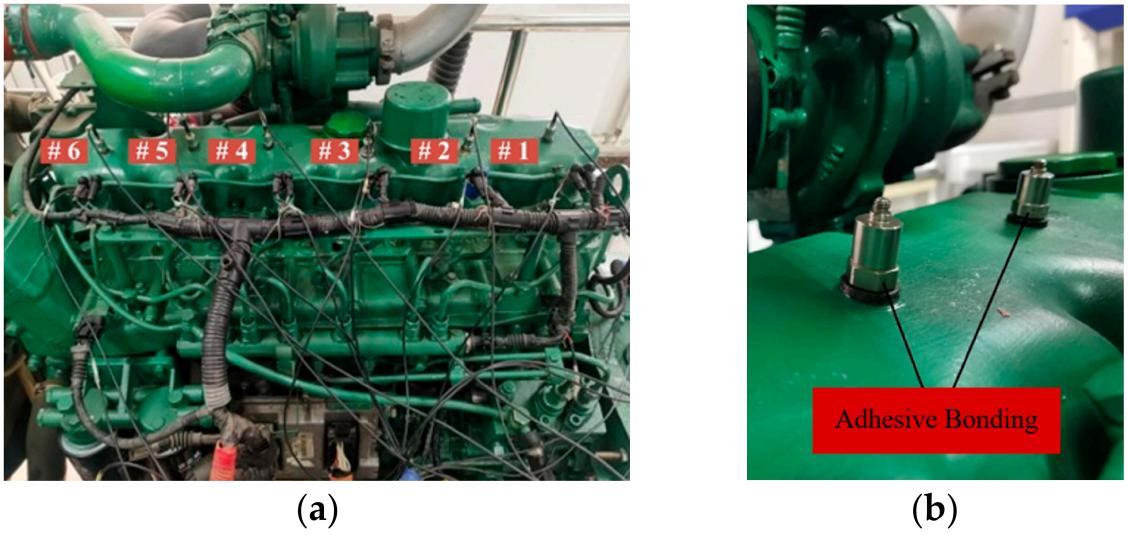

Figure 3. Sensor layout scheme: (**a**) Sensor configuration. (**b**) Sensor installation schematic.

The main parameters of the selected vibration accelerometer sensors, as shown in Table 2, meet the measurement requirements for diesel engines.

Table 2. Main parameters of the sensors.

Category	Parameters	Category	Parameters
Weight	15 g	Frequency Range	0.5 Hz–10 kHz (±10%)
Sensitivity	100 (±5%) mV/g	Transverse Sensitivity	≤5%
Power Supply	18–30 VDC	Maximum Impact Value	1000 g peak
Amplitude Linearity	1%	Broadband Resolution	0.0001 g rms
Supply Current	2–20 mA	Mounting Resonance Frequency	≥29 kHz
Measurement Range	±50 g peak	Maximum Acceleration Value	400 g peak
Output Impedance	≤100 Ω	Output Connector Type	M5/Top Output
Weight	15 g	Frequency Range	0.5 Hz–10 kHz (±10%)
Sensitivity	100 (±5%) mV/g	Transverse Sensitivity	≤5%

2.2. Diesel Engine Condition Monitoring Data Acquisition Test

Based on the results of surveys conducted at grassroots-level units and research institutes, and taking into account existing conditions, we have specifically selected the fuel supply system and the intake and exhaust system of diesel engines for pre-set fault tests, collected vibration signals, and established the relevant datasets.

2.2.1. Data Acquisition Tests on Diesel Engine Fuel Supply System

Based on the common fault types associated with relatively high fault rates in fuel supply systems under actual operating conditions, this test preset five single-fault types and two combined-fault types; the specific preset fault types are shown in Table 3.

Table 3. Diesel engine fuel supply system preset fault types.

No.	Fault Type
1	Normal
2	Misfire in one cylinder
3	Clogged air filter
4	Fuel leaking from the injector
5	Insufficient fuel supply from the fuel pump
6	Misfires in cylinders 1 and 2
7	Clogged air filter, fuel leaking from the injector

As shown in Figure 4, a misfire in a specific cylinder was achieved by disconnecting the fuel injector's power wire for that cylinder; an air filter clogging fault was achieved by installing a homemade intake hood over the air intake; a fuel injector dripping fault was achieved by replacing a functioning fuel injector with a faulty one featuring a worn needle valve; and an insufficient fuel supply from the fuel pump was achieved by replacing a functioning fuel pump with one featuring a worn internal gear.

To ensure that the data was collected while the diesel engine was in a steady state, the data acquisition tests were conducted after the engine had been warmed up, with the engine speed set at 800 r/min. Based on existing conditions and research progress, the sampling frequency was set to 20 kHz, with each acquisition lasting 12 s and containing 240,000 data points, denoted as one data set. Each data set consisted of 6 channels, with a 30-s interval between acquisitions; a total of 300 data sets were collected. Consequently, the data obtained from each channel for each fault type forms a matrix of size 240,000 × 300. Through extensive testing to evaluate the sensitivity of data collected from different channels to fault characteristic information, it was determined that, based on the test results, the sample datasets used in subsequent research refer to the last 100 sets of data from Channel 5 for each fault type. Taking the last 5000 sample points for each fault type in the sample dataset

as an example, their time-domain waveforms are shown in Figure 5, and the corresponding spectra are shown in Figure 6. The numbers 1–7 in Figures 5 and 6 correspond to the seven fault types listed in Table 3.

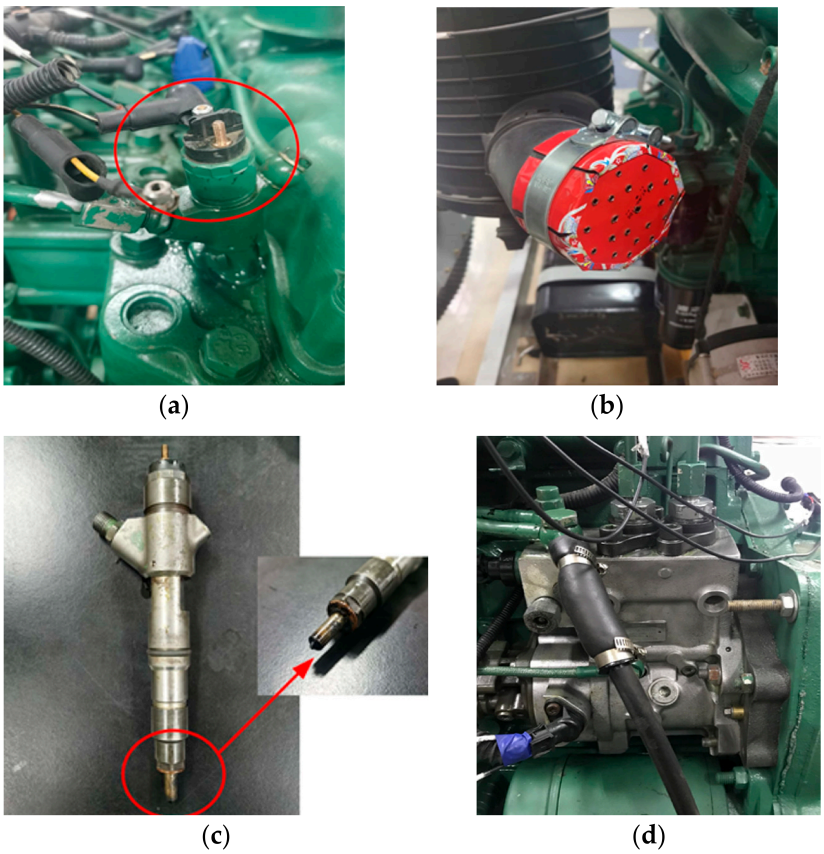

Figure 4. Diesel engine fuel supply system preset fault: (**a**) Disconnect the ignition power cord. (**b**) Add the air intake cover. (**c**) Replace the faulty injector. (**d**) Replace the faulty injection pump.

As shown in Figure 5, vibration signals under different fault types exhibit characteristics such as complexity, nonlinearity, non-stationary time-varying behavior, severe coupling, and noise interference. As a structurally complex reciprocating power machine, a diesel engine involves various forms of motion—including rotational and reciprocating movements—and interactions among its internal components. This results in vibration signals containing multiple amplitude and frequency components, making them highly complex. The output of a diesel engine's vibration signal is not always proportional to its input, and its statistical characteristics vary over time, indicating strong nonlinear and non-stationary time-varying characteristics. The interactions among the various internal components of a diesel engine result in significant coupling between vibration signals, making it difficult to extract valuable information from them. During data acquisition, the signals are inevitably affected by mechanical noise, electromagnetic interference, and other factors, which degrade the quality of the vibration signals and compromise the accuracy of vibration analysis. Furthermore, the range of amplitude variations in vibration signals across different fault types shows no significant differences, and the waveforms are similar, making it difficult to distinguish between specific fault types through manual inspection alone.

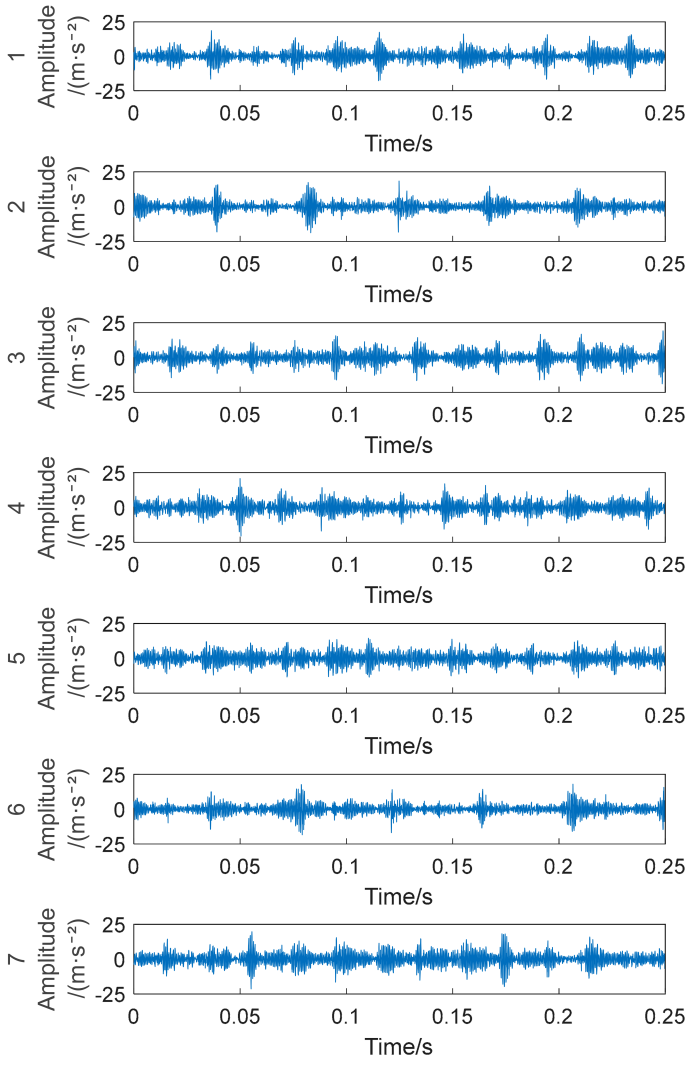

Figure 5. Time-domain waveforms of the 7 fault types of the fuel supply system.

As shown in Figure 6, the vibration signal of the diesel engine contains multiple frequency components, with energy distributed unevenly across different frequencies; the amplitude peaks at around 1.5 kHz. Although the spectral amplitudes differ slightly among different fault types, it is difficult to clearly identify characteristic fault frequencies, making it challenging to manually determine the specific fault type.

2.2.2. Data Acquisition Tests on Diesel Engine Intake and Exhaust Systems

For the diesel engine selected for this test, the intake valve clearance under normal operating conditions is 0.25–0.3 mm, and the exhaust valve clearance is 0.3–0.35 mm. Considering that valve performance degradation in a diesel engine during operation manifests as an increase in valve clearance, the data acquisition test was conducted while the engine was cold. The valve clearances were adjusted using a feeler gauge; the adjustment process is shown in Figure 7. After removing the diesel engine's cylinder head, Cylinder 3 was selected as the cylinder for simulating valve performance degradation. Different intake valve clearances were preset, while the exhaust valve clearance remained unchanged, to simulate the process by which valve performance degrades during engine operation, leading to an increase in valve clearance.

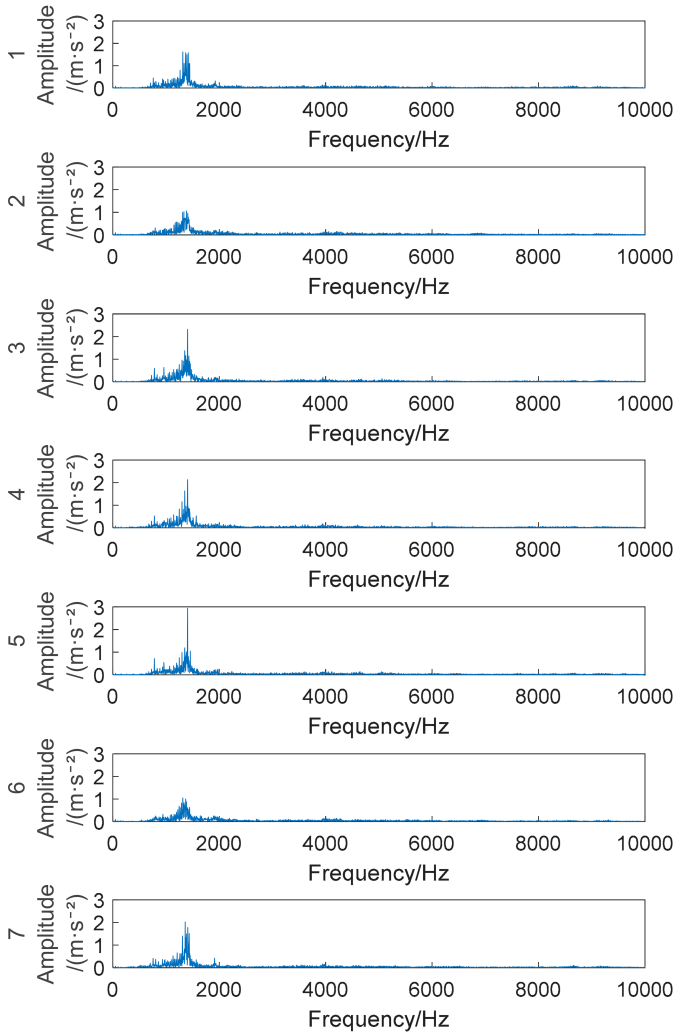

Figure 6. Spectrum of the 7 fault types of the fuel supply system.



(**a**) (**b**) (**c**)

Figure 7. Adjustment process: (**a**) Remove cylinder head. (**b**) Adjust screw. (**c**) Valve clearance.

Under the guidance of the manufacturer's service technicians, loosen the lock nut above the valve of Cylinder 3 and use a screwdriver to turn the adjustment screw to simulate 4 fault conditions: normal (healthy condition), slightly increased intake valve clearance (caution condition), increased intake valve clearance (abnormal condition), and significantly increased intake valve clearance (deteriorated condition), as shown in Table 4.

Table 4. Preset fault status for diesel engine intake and exhaust system.

No.	Intake Valve Clearance (mm)	Deterioration Process	Fault State
1	0.3	Normal	Health
2	0.4	Slightly increased valve clearance	Caution
3	0.5	Increased valve clearance	Abnormal
4	0.7	Significantly increased valve clearance	deteriorating

The parameter settings, data acquisition process, and volume of data collected for the intake and exhaust system data acquisition test were identical to those described in Section 2.2.1. Since Cylinder 3 was selected in this test to simulate valve performance degradation, the sample datasets referred to in the subsequent research all consist of the last 100 data sets from Channel 3 under each fault state. Taking the last 5000 sample points for each fault state in the sample dataset as an example, the time-domain waveforms are shown in Figure 8, and the spectra are shown in Figure 9. The numbers 1–4 in Figures 8 and 9 correspond to the four fault states listed in Table 4, respectively.

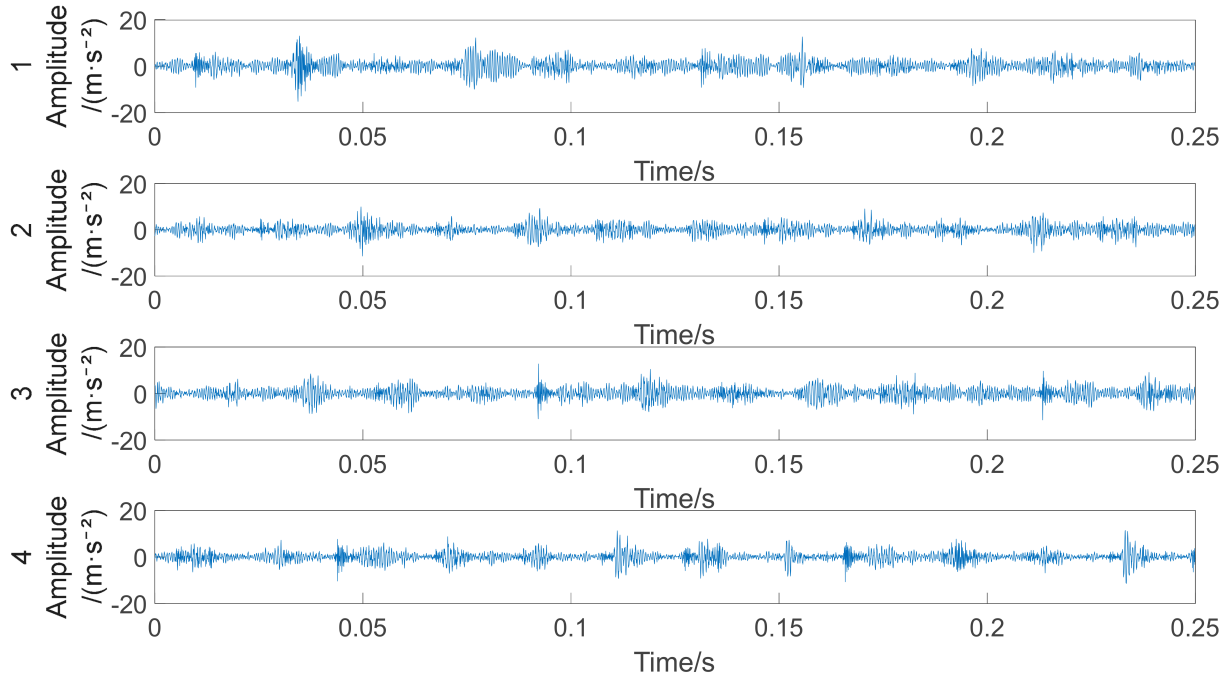

Figure 8. Time-domain waveforms of 4 fault states of the intake and exhaust system.

As shown in Figures 8 and 9, there is no significant difference in the range of amplitude variations in the time-domain waveforms, and the waveforms are similar. The spectra contain multiple frequency components; although there are slight differences in amplitude, it is difficult to clearly identify the characteristic frequencies of the faults. Therefore, it is not possible to effectively identify the various fault states or perform accurate fault diagnosis of the intake and exhaust systems based solely on the time-domain waveforms and their spectra.

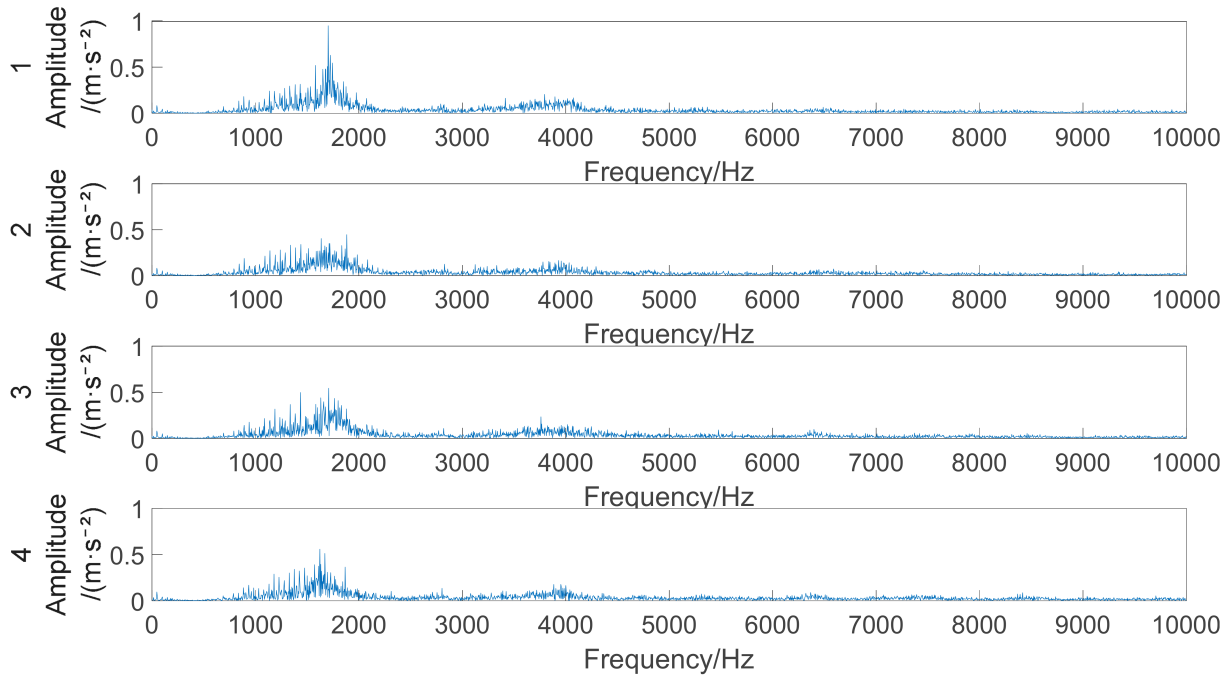

Figure 9. Spectrum of 4 fault states of the intake and exhaust system.

2.2.3. Physical Interpretation of the Collected Vibration Signals

The diesel engine used in the test was a Xichai CA6DF3-20E3 inline 6-cylinder, 4-stroke, water-cooled diesel engine, with a test speed of 800 r/min. Based on this, the following was calculated:

(1) Crankshaft rotation frequency: $f_r = 800/60 = 13.33$ Hz.
(2) Ignition frequency: $f_f = N_c \cdot n/120 = 6 \times 800/120 = 40$ Hz; the corresponding firing interval is 25 ms.
(3) Camshaft rotational speed frequency: The camshaft rotational speed is half that of the crankshaft, i.e., $n/2 = 400$ r/min, corresponding to a frequency of 6.67 Hz.

As can be seen from the time-domain waveforms (Figures 5 and 8), under normal operating conditions, there are approximately 10 pulse groups within 0.25 s, with an interval of about 25 ms between adjacent pulse groups. This is fully consistent with the calculated ignition frequency of 40 Hz, thereby verifying the physical consistency of the acquired signals.

Cylinder head vibration signals are generated by the combined action of the following four main excitation sources:

(1) Combustion pressure excitation: After reaching top dead center during compression in a diesel engine, the fuel ignites and burns, causing the pressure inside the cylinder to rise sharply in a very short time and forming a powerful pressure surge. This surge is transmitted through the bottom surface of the cylinder head as structural vibration. The energy of this excitation is primarily distributed in the 500 Hz–3 kHz frequency range, and its intensity directly reflects the quality of combustion.
(2) Valve seating impact: At the moment the intake and exhaust valves close, the valve disc collides at high speed with the valve seat, generating a broadband impact signal. The repetition frequency of valve seating is related to the camshaft speed (6.67 Hz), and the impact energy is primarily concentrated in the 1–3 kHz frequency range.
(3) Injector needle valve impact: The rapid movement of the injector needle valve at the moments of opening and closing generates mechanical impact. The characteristic

frequency range is 1–4 kHz, and the waveform characteristics are closely related to the sealing condition of the needle valve and the injection timing.

(4) Piston knocking: When the piston changes direction at the upper and lower dead centers, the change in the direction of the lateral thrust causes the piston to strike the cylinder wall laterally. This energy is transmitted through the cylinder liner to the cylinder head and is primarily distributed in the 500 Hz–2 kHz frequency range.

Physical Interpretation of the 1.5 kHz Spectrum Peak:

As shown in the spectrum plots (Figures 6 and 9), the vibration energy is concentrated in the 0–2 kHz frequency band under all operating conditions, with a dominant peak appearing near 1.5 kHz. The physical nature of this peak is a localized resonant response of the cylinder head structure: the cylinder head is a complex elastic structure with multiple natural modes. Combustion pressure surges, valve seating impacts, and fuel injector needle valve impacts are all broadband excitations; under 800 rpm idle conditions, they collectively excite the structural modes of the cylinder head in the 1–2 kHz range, thereby forming this dominant peak.

Most importantly, the frequency band near 1.5 kHz is the resonance band where the three major fault sources—combustion excitation, valve train excitation, and fuel injection excitation—converge. Any fault affecting combustion quality (misfires, insufficient fuel supply, clogged air filter), valve condition (increased valve clearance), or injection behavior (injector dripping) will alter the excitation energy in this resonance band, resulting in observable changes in amplitude and shape in the spectrum. This is precisely the physical reason why this frequency band is highly sensitive to faults, and it also serves as the physical basis for the proposed method's ability to effectively distinguish between fault types.

Analysis of the Physical Mechanisms Behind Various Fuel Supply System Failures:

All seven of the aforementioned failure types produce distinguishable characteristic changes in the vibration signal by altering specific excitation mechanisms:

(1) Single-cylinder misfire: Achieved by disconnecting the power wire to the injector of the affected cylinder. Combustion in the faulty cylinder is interrupted, and the combustion pressure excitation completely disappears, resulting in one noticeably weakened shock every six ignition cycles; simultaneously, due to the loss of combustion excitation, the total spectral energy decreases overall.

(2) Clogged air filter: Achieved by installing a homemade air intake hood at the air intake. Increased intake resistance reduces air-intake efficiency, causing the air–fuel ratio to deviate from the optimal value. Combustion becomes rough, and the rate of rise in cylinder pressure increases, manifesting as an overall increase in vibration amplitude, with energy in the low-frequency range below 1000 Hz becoming more pronounced.

(3) Fuel injector dripping: Achieved by replacing a faulty fuel injector with a worn needle valve. Poor needle valve sealing causes fuel leakage during the non-injection period, resulting in disrupted injection timing and poor atomization, which leads to irregular combustion and abnormal needle valve impacts. In the time domain, this manifests as reduced regularity in the impact pattern and the appearance of non-periodic spikes; in the frequency spectrum, it manifests as multiple discrete peaks.

(4) Insufficient fuel supply from the fuel injection pump: Simulated by replacing a faulty fuel injection pump with worn internal gears. A reduction in the cyclic fuel supply causes the peak combustion pressures of all cylinders to drop simultaneously, resulting in a significant overall reduction in vibration amplitude, a narrowing of the excitation bandwidth, and a relatively sharp, single-peak shape in the spectrum.

(5) Misfire in Cylinder 1 and Misfire in Cylinder 2: Combustion excitation from both cylinders disappears simultaneously; the density of weak knocks doubles compared to a single-cylinder misfire, and total energy decreases significantly.

(6) Clogged Air Filter and Injector Dripping: The high combustion roughness background caused by a clogged air filter overlaps with the abnormal knock characteristics caused by injector dripping, resulting in composite failure characteristics.

Analysis of the Physical Mechanisms Behind Valve Clearance Failures in the Intake and Exhaust Systems:

In tests of the intake and exhaust systems, using cylinder 3 as the test cylinder, the intake valve clearance was gradually increased from the normal range (0.25–0.3 mm) to a degraded state, while the exhaust valve clearance remained constant. The increase in valve clearance affects cylinder head vibration through the following two coupled mechanisms:

(1) Direct mechanism: Increased valve seating impact. An increased valve clearance means the effective drop distance of the valve from the open position to the seated position increases; consequently, the velocity and impact force at the moment of seating also increase. This leads to a significant increase in the mid-to-high-frequency energy (1–3 kHz) of the valve seating impact, particularly within the 1.5 kHz cylinder head resonance band, where energy rises as the clearance increases.
(2) Indirect mechanism: Reduced combustion quality. Increased valve clearance results in reduced valve lift, delayed opening, and earlier closing, which reduces the intake air charge, leads to incomplete combustion, lowers the peak combustion pressure in the cylinder, and weakens the combustion excitation.

The superposition of these two mechanisms results in a unique evolution pattern of the vibration signal near 1.5 kHz as the valve clearance progressively increases from normal → caution → abnormal → severe: the intensified seating impact causes mid-to-high-frequency energy to rise, while the deterioration of combustion weakens the combustion-related components; the relative changes in the strength of these two factors constitute the physical characteristics of valve clearance faults. This characteristic is physically distinct from the spectral variation patterns associated with fuel supply system faults.

Based on the above analysis, all types of faults in this study induce regular changes in the corresponding frequency bands and time-domain modes of the cylinder head vibration signal by altering specific physical excitation sources (combustion pressure, valve seating impact, injector needle valve impact, or intake efficiency). The SMVMD method can effectively isolate these fault-related vibration components, and based on the principle of component selection with minimum envelope entropy, it can identify the IMF components that carry the most distinctive physical information. Therefore, the proposed method extracts fault-related features with clear physical significance, rather than merely distinguishing vibration modes specific to the dataset.

Order is defined as the ratio of the vibration frequency to the crankshaft rotational frequency—that is, the normalized frequency based on one revolution of the crankshaft—which eliminates the influence of rotational speed fluctuations on spectral analysis. Using crankshaft rotation as the reference, Order = frequency/crankshaft rotational frequency f_r, f_r = 13.33 Hz. For an inline 6-cylinder, 4-stroke diesel engine, a complete power cycle corresponds to two revolutions of the crankshaft, with the six cylinders firing in sequence; therefore, three firings occur per crankshaft revolution, resulting in a primary ignition stage of 3, with a corresponding frequency of $f_{\text{fire}} = 3 \times f_r = 40$ Hz. The camshaft speed is half that of the crankshaft, so the fundamental frequency of the valve train is 0.5 orders, corresponding to a frequency of 6.67 Hz. The inertial forces resulting from the reciprocating motion of the piston in a four-stroke engine also include half-order components such as 1.5 and 2.5. Table 5 lists the key orders of interest in the analysis and their physical meanings.

Table 5. Key orders of significance in the analysis and their physical meanings.

Order	Corresponding Frequency (Hz)	Physical Meaning
0.5-order	6.67	Camshaft rotation frequency/Fundamental frequency of the operating cycle of a four-stroke engine/Valve seating frequency per cylinder
1-order	13.33	Crankshaft rotational imbalance, fundamental frequency of piston reciprocating inertial forces
1.5-order	20.0	Second-order inertial forces of a reciprocating piston (half-order system specific to the four-stroke cycle)
2-order	26.67	Second-order harmonics of crankshaft rotation, connecting rod force couple
2.5-order	33.33	Higher-order components of reciprocating inertial forces
3-order	40.0	Ignition frequency (6-cylinder, 4-stroke, primary ignition stage)
6-order	80.0	Second-harmonic of the ignition frequency
9-order	120.0	Ignition frequency 3rd-order harmonic
12-order	160.0	Ignition frequency 4rd-order harmonic

The physical mechanisms underlying the changes in each order of the disturbance are analyzed as follows:

(1) Misfire disturbances (misfire in cylinder 1, misfire in cylinder 2): The combustion pressure excitation from the misfiring cylinder disappears, causing a decrease in the total energy of the ignition order (3rd order) and its harmonics (6th, 9th, and 12th orders); the magnitude of this decrease is proportional to the number of misfiring cylinders. Furthermore, since misfires occur at a cycle of 2 revolutions (one power stroke), this is equivalent to amplitude modulation of the ignition excitation at a frequency of 0.5-order. Consequently, sideband components spaced at 0.5-order intervals appear around the 3rd order, resulting in additional spectral components at positions such as the 2.5th, 3.5th, 2.0th, and 4.0th orders. This combination of "reduced ignition-order amplitude and sidebands spaced at 0.5-order intervals" serves as a unique order identifier for misfire faults.

(2) Clogged air filter: Increased intake resistance leads to reduced air-charging efficiency, causing the air–fuel ratio to deviate from the optimal value. Combustion becomes rough, and the rate of rise in cylinder pressure increases, resulting in a general rise in vibration energy across all orders, with a more significant increase observed in the low-frequency orders (0.5th and 1st).

(3) Fuel injector dripping: Wear on the needle valve causes poor sealing, resulting in fuel leakage during the non-injection period, disrupted injection timing, poor atomization, and unstable combustion phase. This reduces the coherence between the harmonics of the ignition mode, disrupts the harmonic structure, and prevents the regular harmonic attenuation relationship observed under normal operating conditions.

(4) Insufficient fuel supply from the fuel injection pump: Wear on the internal gears of the fuel injection pump reduces the circulating fuel supply, causing the peak combustion pressures of all cylinders to decrease simultaneously. Consequently, the amplitudes of all orders exhibit a downward trend, and the magnitude of the decrease is relatively consistent.

(5) Combined failure (clogged air filter + injector dripping): The rough combustion caused by a clogged air filter coexists with the injection disruption caused by injector dripping; the order spectrum exhibits a superposition of the characteristics of these two individual failures.

An increase in valve clearance affects the amplitude of the order through two coupling mechanisms:

(1) Direct mechanism: Increased valve seating impact (rise in the 0.5th-order term): An increase in valve clearance means that the effective drop distance of the valve from the open position to the seated position increases, resulting in higher velocity and greater impact force at the moment of seating. Since valve seating occurs at the repetition period of the camshaft speed (0.5-order), the amplitudes of the 0.5-order and its related orders (1st and 1.5-order) increase monotonically as the clearance increases. At the same time, the energy of the valve seating impact is primarily concentrated in the 1–3 kHz frequency band; therefore, the energy within the 1.5 kHz cylinder head resonance band also increases accordingly.

(2) Indirect mechanism: Decline in combustion quality (Decrease in 3rd-Order Component): Increased valve clearance leads to reduced valve lift, delayed opening, and early closing, resulting in reduced air intake, incomplete combustion, and a lower peak combustion pressure within the cylinder. Consequently, the amplitude of the 3rd-order ignition fundamental frequency—which reflects combustion excitation—gradually decreases as the clearance increases.

This trend—characterized by an "increase in valve-related order (0.5-order system) and a decrease in ignition order (3rd order)"—is a physical characteristic unique to valve clearance faults: fuel system faults typically manifest as a decrease or irregularity in the 3rd order, but without a significant increase in the 0.5-order; whereas valve clearance faults manifest as opposite changes in the 0.5-order and 3rd order, forming a clear physical distinction between the two.

In summary, through order tracking and harmonic analysis, we have established a quantitative correlation between different fault types and changes in vibration order amplitudes: changes in the amplitude of the ignition order (3rd order) and its harmonics directly reflect changes in combustion excitation and serve as a core indicator for distinguishing combustion-related faults such as misfires, insufficient fuel supply, and clogged air filters; changes in the amplitude of valve-related orders (0.5-order series) directly reflect changes in valve seat impact and serve as a core indicator for diagnosing valve clearance faults; the 0.5-order sideband of misfire faults provides a unique order-based identifier for misfire types; and the opposite trends—an increase in the 0.5-order and a decrease in the 3rd-order—for valve clearance faults provide a clear physical basis for distinguishing between intake and exhaust system faults.

These quantitative order and harmonic analysis results indicate that different fault types, by altering specific physical excitation sources, produce quantifiable and distinguishable characteristic changes in the order domain, providing a solid physical foundation for the proposed intelligent diagnostic method.

3. Diesel Engine Intelligent Fault Diagnosis Technical Approach

The technical approach for the proposed intelligent fault diagnosis method for diesel engines under unbalanced few-shot conditions, based on SMVMD and optimized DHKELM, is shown in Figure 10. The specific steps are as follows:

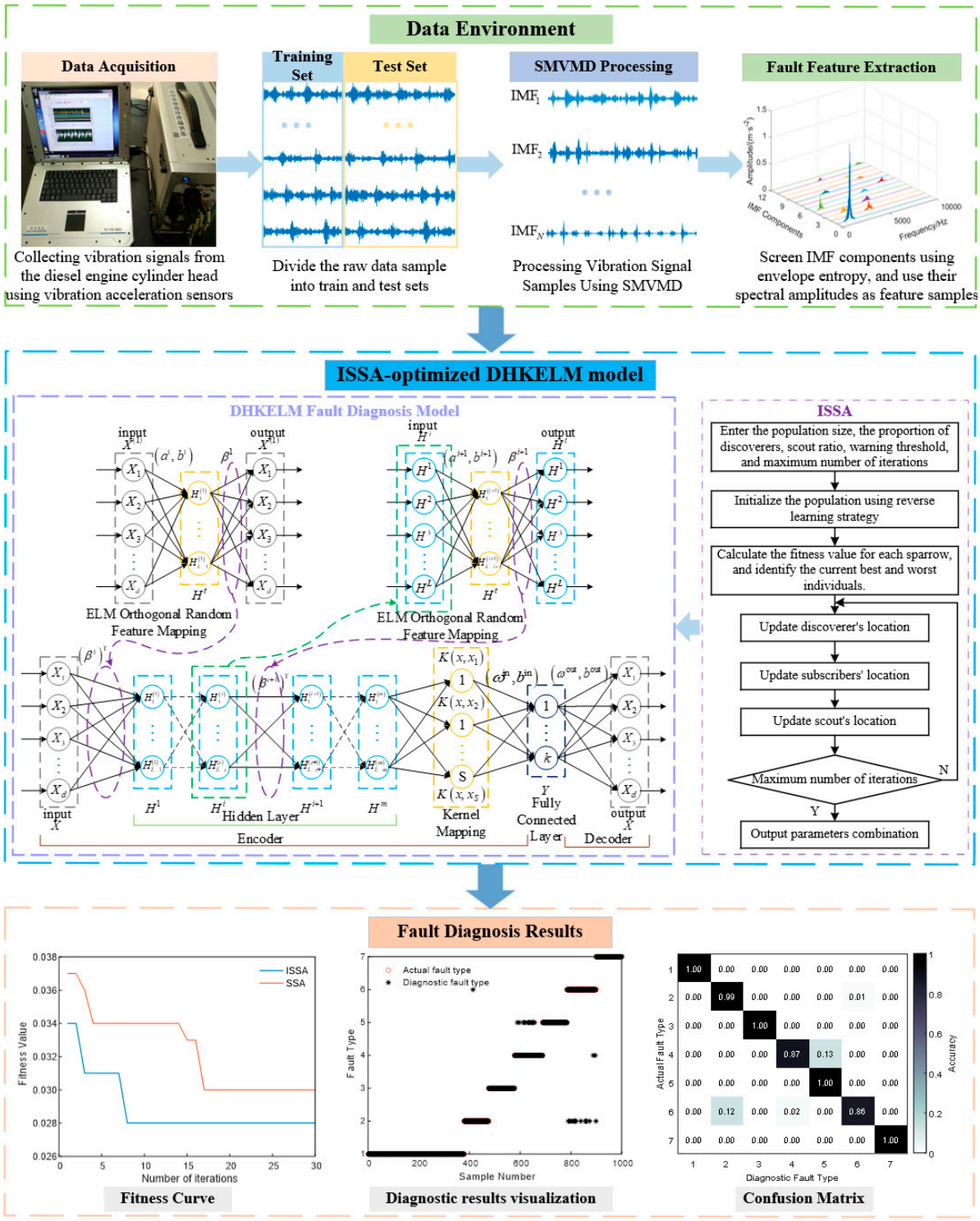

Figure 10. Technical approach of unbalanced few-shot diesel engine intelligent fault diagnosis method.

Step 1: Use the vibration acceleration sensors to collect vibration signals corresponding to different fault types from the diesel engine cylinder head, thereby obtaining the required raw data samples.

Step 2: In accordance with the requirements for unbalanced few-shot problems (detailed explanation provided in Section 5), divide the raw data samples into training and test sets, and assign labels to the data sets.

Step 3: Process the samples in the data set using the SMVMD method to obtain the IMF components of the samples.

Step 4: Calculate the envelope entropy values for all IMF components in each sample, sort them by value, and select the most valuable IMF component from each sample based on the principle of minimizing the envelope entropy value.

Step 5: Use the Fast Fourier Transform (FFT) to extract the selected IMF component. Treat its spectral amplitude as a fault feature sample; after normalization, obtain the feature dataset, and apply one-hot encoding to the fault category labels.

Step 6: Construct the DHKELM fault diagnosis model. Optimize the hyperparameters of the proposed ISSA model, using the fault diagnosis error rate during DHKELM training on the training set as the fitness function. Select the parameter combination that yields the minimum fitness function value as the optimal combination. Substitute the optimized parameter combination into the DHKELM model to obtain the trained, unbalanced few-shot diesel engine fault diagnosis model.

Step 7: Input the test set into the trained model to evaluate its performance and locate the fault.

4. Principles of Intelligent Fault Diagnosis Technology for Diesel Engines

The proposed intelligent fault diagnosis method for diesel engines processes vibration signals using the SMVMD method and employs the proposed ISSA-optimized DHKELM model to identify fault types.

4.1. Principles of SMVMD

All univariate processing methods can be regarded as special cases of multivariate signal processing [18]. The goal of SMVMD is to continuously extract K Multivariate IMF (MIMF) components $u_k(t) \in \mathbb{R}^{C\times 1}$ from input data $x(t) = [x_1(t), x_2(t), \cdots, x_C(t)]^{\mathrm{T}}$, which has C channels or dimensions.

$$x(t) = \sum_{k=1}^{K} u_k(t) \tag{1}$$

$$\begin{aligned} u_k(t) = \; & \left[u_{k,1}(t), u_{k,2}(t), \cdots, u_{k,C}(t)\right]^{\mathrm{T}} = \\ & \left\{a_{k,1}(t)\cos[\varphi(t)], a_{k,2}(t)\cos[\varphi(t)], \cdots, a_{k,C}(t)\cos[\varphi(t)]\right\}^{\mathrm{T}} \end{aligned} \tag{2}$$

In the equation: t represents time; $k = 1 : K$, where K can be determined based on convergence criteria and do not require manual setting; $c = 1 : C$; and $a(t)$ and $\varphi(t)$ represent the signal amplitude and phase functions, respectively.

In SMVMD, $\omega(t) = \mathrm{d}\varphi(t)/\mathrm{d}t$, $x(t)$ are split into two components: $u_k(t)$ and the residual signal $x_r(t) \in \mathbb{R}^{C\times 1}$.

$$x(t) = u_k(t) + x_r(t) \tag{3}$$

In the equation, $x_r(t)$ consists of two parts: $\sum_{i=1}^{k-1} u_i(t)$, which is the sum of the extracted MIMF, and $x_u(t) \in \mathbb{R}^{C\times 1}$, which is the unprocessed portion.

$$x_r(t) = \sum_{i=1}^{k-1} u_i(t) + x_u(t) \tag{4}$$

The constraint criteria for SMVMD include:

$$J_1 = \left\| \partial_t \left\{ \left[\delta(t) + \frac{\mathrm{j}}{\pi t} \right] * u_k(t) \right\} \mathrm{e}^{-\mathrm{j}\omega_k t} \right\|_2^2 = \sum_{c=1}^{C} \left\| \partial_t \left\{ \left[\delta(t) + \frac{\mathrm{j}}{\pi t} \right] * u_{k,c}(t) \right\} \mathrm{e}^{-\mathrm{j}\omega_k t} \right\|_2^2 \tag{5}$$

In the equation: $\|\cdot\|_2^2 = \int_{\mathrm{R}} |\cdot|^2 \mathrm{d}t$; "$*$" denotes the convolution operator; $\partial_t[\cdot]$ denotes the partial derivative with respect to the time parameter t; $\delta(t)$ denotes the unit impulse function; j denotes the imaginary unit; and ω_k denotes the center frequency of $u_k(t)$.

$$J_2 = \left\| \hat{\beta}_k(\omega)\hat{x}_r(\omega) \right\|_2^2 = \sum_{c=1}^{C} \left\| \hat{\beta}_k(\omega)\hat{x}_{r,c}(\omega) \right\|_2^2 \tag{6}$$

In the equation: $\hat{\beta}_k(\omega) = 1/\eta(\omega - \omega_k)^2$, η represent the frequency response coefficient.

$$J_3 = \sum_{i=1}^{k-1} \left\| \hat{\beta}_i(\omega)\hat{u}_k(\omega) \right\|_2^2 = \sum_{c=1}^{C} \sum_{i=1}^{k-1} \left\| \hat{\beta}_i(\omega)\hat{u}_{k,c}(\omega) \right\|_2^2 \tag{7}$$

In the equation: $\hat{\beta}_i(\omega) = 1/\eta(\omega - \omega_i)^2$, and $i = 1, 2, \cdots, k-1$.

$$x(t) = u_k(t) + \sum_{i=1}^{k-1} u_i(t) + x_u(t) \tag{8}$$

$$\min_{u_k, \omega_k, x_r} \alpha J_1 + J_2 + J_3 \text{ s.t. } u_k(t) + x_r(t) = x(t) \tag{9}$$

In the equation: α represents the regularization parameter used to balance the constraint conditions. The corresponding augmented Lagrangian function in Equation (9) is

$$\begin{aligned} L(u_k, \omega_k, x_r) = \alpha J_1 + J_2 + J_3 + \sum_{c=1}^{C} \left\| x_c(t) - \left[u_{k,c}(t) + x_{u,c}(t) + \sum_{i=1}^{k-1} u_{i,c}(t) \right] \right\|_2^2 + \\ \sum_{c=1}^{C} \left\langle \lambda_c(t), x_c(t) - \left[u_{k,c}(t) + x_{u,c}(t) + \sum_{i=1}^{k-1} u_{i,c}(t) \right] \right\rangle \end{aligned} \tag{10}$$

In this equation, $\lambda_c(t)$ is the Lagrangian operator for channel C; $\langle \cdot, \cdot \rangle$ denotes the inner product operator. Equation (10) can be solved iteratively using the alternating direction multiplier algorithm.

The iterative update process for $\hat{u}_{k,c}(\omega)$ can be expressed as

$$\hat{u}_{k,c}^{n+1}(\omega) = \frac{\hat{x}_c(\omega) - \hat{x}_{u,c}^n(\omega) - \sum_{i=1}^{k-1} \hat{u}_{i,c}(\omega) + \hat{\lambda}_c(\omega)/2}{1 + 2\alpha\left(\omega - \omega_k^n\right)^2 + \sum_{i=1}^{k-1} 1/\alpha^2\left(\omega - \omega_i^n\right)^4} \tag{11}$$

The iterative update process for ω_k can be expressed as

$$\omega_k^{n+1} = \sum_c \int_0^\infty \omega \left| \hat{u}_{k,c}^{n+1}(\omega) \right|^2 \mathrm{d}\omega / \sum_c \int_0^\infty \left| \hat{u}_{k,c}^{n+1}(\omega) \right|^2 \mathrm{d}\omega \tag{12}$$

The iterative update process for $\hat{x}_{u,c}(\omega)$ can be expressed as

$$\hat{x}_{u,c}^{n+1}(\omega) = \frac{\alpha^2\left(\omega - \omega_k^{n+1}\right)^4 \left[\hat{x}_c(\omega) - \hat{u}_{k,c}^{n+1}(\omega) - \sum_{i=1}^{k-1} \hat{u}_{i,c}(\omega) + \hat{\lambda}(\omega)/2 \right] - \sum_{i=1}^{k-1} \hat{u}_{i,c}(\omega)}{1 + \alpha^2\left(\omega - \omega_k^n\right)^4} \tag{13}$$

Substituting Equation (13) into Equation (11), $\hat{u}_{k,c}(\omega)$ can be simplified to

$$\hat{u}_{k,c}^{n+1}(\omega) = \frac{\hat{x}_c(\omega) + \alpha^2\left(\omega - \omega_k^n\right)^4 \hat{u}_{k,c}^n(\omega) + \hat{\lambda}_c(\omega)/2}{\left[1 + \alpha^2\left(\omega - \omega_k^n\right)^4\right]\left[1 + 2\alpha\left(\omega - \omega_k^n\right)^2 + \sum_{i=1}^{k-1} 1/\alpha^2\left(\omega - \omega_i^n\right)^4\right]} \tag{14}$$

The iterative update process for $\hat{\lambda}_c$ can be expressed as

$$\hat{\lambda}_c^{n+1} = \hat{\lambda}_c^n + \tau\left\{ \left[\hat{x}_c(\omega) - \hat{u}_{k,c}^{n+1}(\omega) + \hat{\lambda}_c^n/2 \right] / \left[1 + \alpha^2\left(\omega - \omega_k^{n+1}\right)^4 \right] - \hat{\lambda}_c^n/2 \right\} \tag{15}$$

In the formula: τ represents the double-rising update step size.

Let $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ both be the convergence tolerances for the SMVMD iterative process, then the condition for terminating the iteration can be expressed as

$$\sum_c \left\| \hat{u}_{k,c}^{n+1} - \hat{u}_{k,c}^n \right\|_2^2 / \left\| \hat{u}_{k,c}^n \right\|_2^2 < \varepsilon_1 \tag{16}$$

$$\frac{1}{C} \sum_{i=1}^{k} \left\| u_{k,c} \right\|_2^2 / \left\| x_c \right\|_2^2 < \varepsilon_2 \text{ or } k = K \tag{17}$$

Considering that setting the value of α too high can make it difficult for SMVMD to converge, and that setting the value of α too low can cause modal aliasing in SMVMD, this manuscript introduces the rate-of-change operator $\gamma \geq 1$, uses a variable value for α, and updates the expression to $\alpha = \min\{\gamma\alpha, \alpha_{\max}\}$, with the initial value of α set to $\alpha_{\min}$.

4.2. ISSA-Optimized DHKELM Model Construction

To optimize the feature extraction process at the model level and improve the accuracy of fault diagnosis results under unbalanced few-shot conditions, we introduce hybrid kernel functions and a deeper architecture based on the Kernel Extreme Learning Machine (KELM) to propose the DHKELM diesel engine fault diagnosis model.

KELM uses a single kernel function, making it difficult to adapt to practical application scenarios characterized by unbalanced sample ratios and a scarcity of available fault data. To effectively enhance the model's learning and generalization capabilities by combining the strengths of various kernel functions in addressing different problems, DHKELM introduces a hybrid kernel function based on KELM. The Gaussian kernel function $K_{\mathrm{RBF}}\left(x_i, x_j\right)$ and the polynomial kernel function $K_{\mathrm{Poly}}\left(x_i, x_j\right)$, as widely used examples of local and global kernel functions, respectively, possess strong learning and generalization capabilities. By combining $K_{\mathrm{RBF}}\left(x_i, x_j\right)$ and $K_{\mathrm{Poly}}\left(x_i, x_j\right)$ using a linear weighting method, we achieve more ideal fault diagnosis results than when using a single kernel function. The expression for the hybrid kernel function $K\left(x_i, x_j\right)$ is

$$K_{\mathrm{RBF}}\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = \exp\left(-\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2 / 2\sigma^2\right) \tag{18}$$

$$K_{\mathrm{Poly}}\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = \left(\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j + q\right)^p \tag{19}$$

$$K\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = v \cdot K_{\mathrm{RBF}} + (1 - v) \cdot K_{\mathrm{Poly}} \tag{20}$$

In the formula: σ represents the kernel parameter, which is used to control the radial range; the kernel parameter q represents the constant term of the polynomial; $p \geq 1$ represents the exponent; and $v \in [0, 1]$ represents the weighting coefficient, which is used to determine the relative weights of the different kernel functions.

Compared with KELM, the Deep KELM (DKELM) is better able to capture complex mapping relationships among data and possesses stronger nonlinear fitting and representation capabilities. Furthermore, DKELM eliminates the back-propagation step, significantly reducing model training time. Consequently, DHKELM incorporates deep learning concepts on top of KELM. It is a neural network comprising multiple hidden layers, with each layer functioning as an independent module. Using ELM-AE as the basic unit, these layers are cascaded and stacked to extract and combine features layer by layer, forming a multi-layer network feature extraction model.

DHKELM consists of an input layer, m hidden layers, kernel mapping layers, fully connected layers, and an output layer. The hidden nodes are selected as additive nodes with orthogonal random biases and weights. By extracting input data layer by layer through AE, the model effectively learns sample features, reduces the interference caused by unbalanced data and insufficient samples, accurately distinguishes easily confused fault types, and significantly improves the model's diagnostic accuracy, generalization ability, and stability. The structure is shown in Figure 11 and is primarily divided into two parts: the encoder and the decoder. The condition that the target input equals the output is a requirement for the AE-based DHKELM, which aims to learn meaningful representations of input features that are equal, sparse, or compressed.

In the encoder section, the initial weights $\{\boldsymbol{\beta}_i\}_{i=1}^m$ of the hidden layer are obtained through hidden layer compression and kernel mapping. During model training, the input data $\boldsymbol{X}$ is processed using a hierarchical learning approach: First, $\boldsymbol{X}$ is fed into the first ELM-AE for unsupervised AE training. After passing through the first ELM-AE, the resulting output weights β^1 are used to preprocess the data, and the preprocessed data is then passed to the lower-layer Hybrid KELM (HKELM) for learning; next, the input data for the subsequent ELM-AE is the matrix output from the first hidden layer of the HKELM.

This ELM-AE performs AE training again and generates another set of output weights β^2. Weights β^2 are used to perform a second round of data processing, after which the processed data is fed into the underlying HKELM for the next round of learning. This process repeats according to the predefined hierarchical structure until all ELM-AE layers have been trained, the output of each ELM-AE layer serves as the input to the next HKELM layer, and each HKELM layer further learns based on the results from the previous layer; finally, after all ELM-AE layers have been trained, DHKELM uses a multi-kernel weighted approach to integrate the information learned at each level, forming the kernel matrix Ω_{KELM}, which replaces the traditional random matrix $\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}}$ to obtain the final effective output of DHKELM [19].

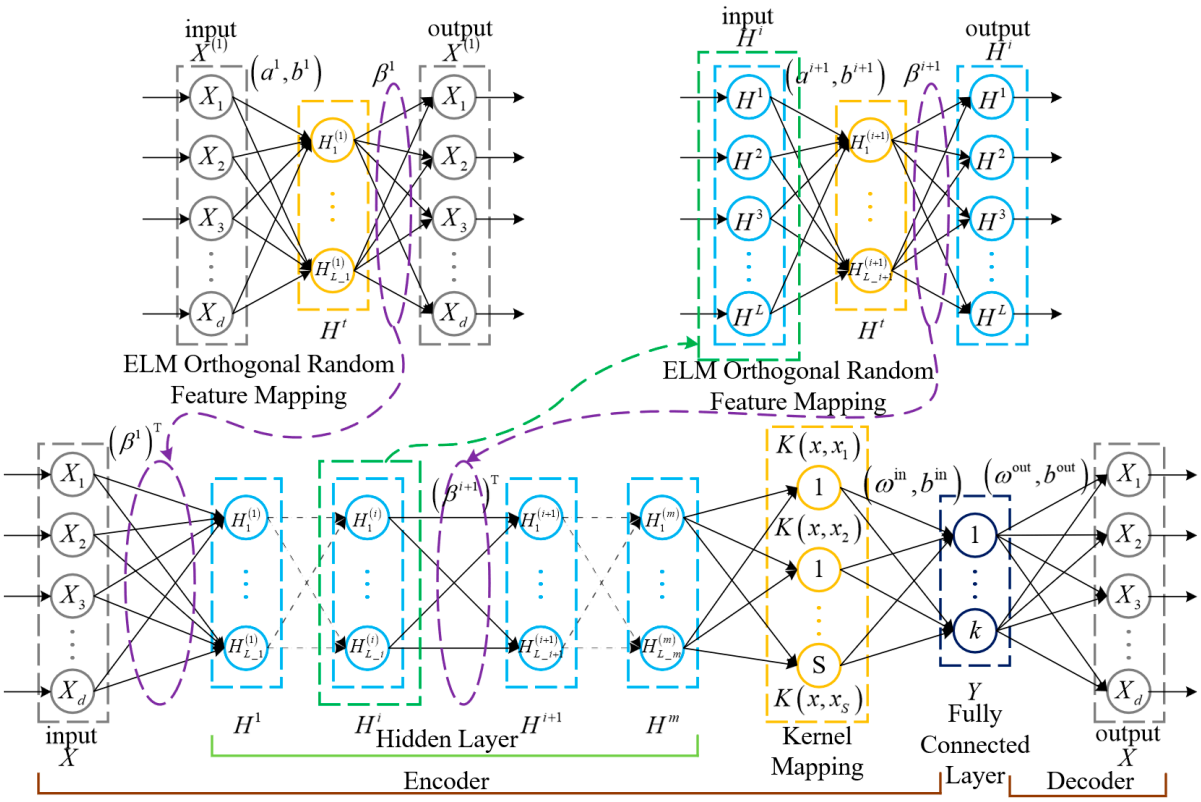

Figure 11. DHKELM model structure.

The specific steps of the initialization process are as follows:

Step 1: Randomly select the orthogonal parameter (a^i, b^i) between the input layer and the hidden layer.

Step 2: Solve Equation (21).

$$\boldsymbol{H}^t = \left(\boldsymbol{I}/\lambda + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}\boldsymbol{T} \tag{21}$$

In the equation: $\boldsymbol{H}^t$ represents the hidden layer state matrix; λ represents the penalty parameter. The Sigmoid activation function is used.

Step 3: Solve Equation (22).

$$\boldsymbol{\beta}^{i+1} = \left[\boldsymbol{I}/\lambda + \left(\boldsymbol{H}^t\right)^{\mathrm{T}}\boldsymbol{H}^t\right]^{-1}\left(\boldsymbol{H}^t\right)^{\mathrm{T}}\boldsymbol{H}^i \tag{22}$$

In the equation: $\boldsymbol{\beta}^{i+1}$ represents the weights of the $i+1$ th hidden layer.

Step 4: Stack the hidden layers to form a hierarchical structure. The input to the current layer is $\boldsymbol{H}^{i+1}$, and the output of the previous layer is $\boldsymbol{H}^i$:

$$\boldsymbol{H}^0 = \boldsymbol{X} \tag{23}$$

$$\boldsymbol{H}^{i+1} = g\left[\left(\boldsymbol{\beta}^{i+1}\right)^{\mathrm{T}}\boldsymbol{H}^{i}\right] \tag{24}$$

In the equation: $i = 0, 1, \cdots, m-1$. The output of the previous layer serves as the input to the current layer. The output of a fully connected layer can be expressed as

$$\boldsymbol{Y} = g\left(\boldsymbol{\omega}^{\mathrm{in}}\boldsymbol{H}^{m} + b^{\mathrm{in}}\right) \tag{25}$$

In the equation: $\boldsymbol{\omega}^{\mathrm{in}}$ and b^{in} represent the weights and biases, respectively, at the input to the fully connected layer; $\boldsymbol{H}^{m}$ represents the final hidden layer. The parameters between $\boldsymbol{H}^{m}$ and $\boldsymbol{Y}$ are obtained through fine-tuning.

Step 5: A fully connected layer from a traditional neural network is configured as the decoder. The reconstructed data is obtained as the output:

$$\overset{\mathrm{output}}{\boldsymbol{X}} = g\left(\boldsymbol{\omega}^{\mathrm{out}}\boldsymbol{Y} + b^{\mathrm{out}}\right) = g\left[\boldsymbol{\omega}^{\mathrm{out}}g\left(\boldsymbol{\omega}^{\mathrm{in}}\boldsymbol{H}^{m} + b^{\mathrm{in}}\right) + b^{\mathrm{out}}\right] \tag{26}$$

In the equation: $\boldsymbol{\omega}^{\mathrm{out}}$ and b^{out} represent the weights and biases, respectively, at the output of the fully connected layer; the objective function of DHKELM is to minimize the reconstruction error between the input data and the output data, and the formula is

$$F_L = \underset{\boldsymbol{\omega}^{\mathrm{in}}, b^{\mathrm{in}}, \boldsymbol{\omega}^{\mathrm{out}}, b^{\mathrm{out}}}{\operatorname{argmin}} \left\| \overset{\mathrm{output}}{\boldsymbol{X}} - \boldsymbol{X} \right\|_2^2 \tag{27}$$

The DHKELM hidden-layer state matrices and weights are determined through direct computation rather than iteration; that is, once the model structure parameters are set during the initialization phase, there is no need for time-consuming retraining. Only the weight parameters between the last two layers need to be trained using the backpropagation algorithm. This not only reduces the number of parameters requiring fine-tuning but also improves training efficiency.

The hyperparameters in the DHKELM model include the number of hidden layer nodes, the penalty coefficient, the regularization coefficient, and the weights of each kernel function. The types and numbers of kernel parameters vary depending on the kernel function. The accuracy of the model's fault diagnosis results is significantly influenced by parameter selection, and it is difficult to determine reasonable parameter combinations based solely on human experience. To fully enhance the model's ability to extract feature information and ensure that its diagnostic performance meets practical application requirements, it is essential to optimize the hyperparameters in the DHKELM using appropriate intelligent optimization algorithms. In 2020, Xue Jiankai proposed the Sparrow Search Algorithm (SSA), which demonstrates better optimization performance compared to algorithms such as Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) [20]. However, SSA suffers from issues such as low optimization accuracy and slow convergence, leaving room for improvement. Therefore, the ISSA is proposed in this manuscript to optimize the various hyperparameters in the DHKELM. The improvements over SSA include:

(1) Using a reverse learning strategy to address the issue in SSA where the randomly generated initial population is of low quality and lacks diversity, resulting in slow algorithm convergence. The core idea is as follows: based on the randomly generated initial population, a reverse population is obtained, and the better individuals from this population are selected as the next generation. The reverse learning strategy selects individuals that are closer to the optimal solution than the initial members of the population. This brings each individual one step closer to the optimal solution, thereby improving the convergence speed of all individuals in the population. The reverse learning strategy also enhances the population's diversity by exploring more viable regions, thereby strengthening the algorithm's global search capability. The reverse population $\boldsymbol{X}_i^* = \left[x_{i1}^*, \cdots, x_{id}^*, \cdots, x_{iD}^*\right]$ can be generated using Equation (28):

$$\boldsymbol{X}_i^* = \text{rand}(lb + ub) - \boldsymbol{X}_i \tag{28}$$

In the equation: $i = 1, 2, \cdots, N$; rand denote random numbers in the range (0, 1); lb and ub denote the lower and upper bounds of the search space, respectively; and $\boldsymbol{X}_i = [x_{i1}, \cdots, x_{id}, \cdots, x_{iD}]$ denotes the initial population. After merging the reverse population and the randomly generated population, a new population is obtained. By sorting the fitness function values in ascending order, the desired new initial population is obtained as the top N optimal initial solutions from the sorted list.

(2) Under the SSA framework, the position of the discoverer is updated at each iteration according to Equation (29). However, updating the position using Equation (29) can easily cause the algorithm to get stuck in a local optimum. To enhance the algorithm's global search capability, Equation (29) is improved using Equation (30).

$$\boldsymbol{X}_{i,j}^{t+1} = \begin{cases} \boldsymbol{X}_{i,j}^t \cdot \exp\left(-\frac{i}{\alpha \cdot \text{iter}_{\max}}\right) & \text{if } R_2 < ST \\ \boldsymbol{X}_{i,j}^t + Q \cdot \boldsymbol{L} & \text{if } R_2 \geq ST \end{cases} \tag{29}$$

In the equation: $\boldsymbol{X}$ represents an $n \times d$-dimensional sparrow population; n represents the total number of sparrows; d represents the number of parameters to be optimized; i represents the index of a sparrow in the sparrow population; $j = 1, 2, \cdots, d$ represents the dimension index; t represents the current iteration number; $\alpha \in (0, 1]$ represents a random number; the constant $\text{iter}_{\max}$ represents the maximum number of iterations; Q represents a random number that follows a normal distribution; $\boldsymbol{L}$ represents an $1 \times d$-dimensional matrix whose elements are all 1; $R_2 \in [0, 1]$ represents the early warning value; and $ST \in [0.5, 1]$ represents the safety value.

$$\boldsymbol{X}_{i,j}^{t+1} = \begin{cases} \eta \cdot \boldsymbol{X}_{i,j}^t \cdot \exp\left(-\frac{i}{\alpha \cdot \text{iter}_{\max}}\right) & \text{if } R_2 < ST \\ \eta \cdot \boldsymbol{X}_{i,j}^t + Q \cdot \boldsymbol{L} & \text{if } R_2 \geq ST \end{cases} \tag{30}$$

$$\eta = \omega_{\max} - (-\omega_{\min}) \times \left[\frac{2t}{\text{iter}_{\max}} - \left(\frac{t}{\text{iter}_{\max}} \right)^2 \right] \tag{31}$$

In the equation: $\omega_{\max} = 0.9$ and $\omega_{\min} = 0.4$ represent the maximum and minimum values of the weights, respectively.

(3) In the SSA, the update expression for the position of the sparrow that detects danger (the scout) at each iteration is shown in Equation (32). Parameters β and K are random numbers that play a critical role in balancing local exploration and global search capabilities; arbitrary random values for these parameters can easily cause the algorithm to get stuck in a local optimum. Therefore, by refining parameters β and K through Equations (33) and (34), respectively, we achieve dynamic adjustment of the step size control parameters, balance the algorithm's local exploration and global search capabilities, and improve optimization accuracy.

$$\boldsymbol{X}_{i,j}^{t+1} = \begin{cases} \boldsymbol{X}_{\text{best}}^t + \beta \cdot \left| \boldsymbol{X}_{i,j}^t - \boldsymbol{X}_{\text{best}}^t \right| & \text{if } f_i \neq f_g \\ \boldsymbol{X}_{i,j}^t + K \cdot \left\{ \left| \boldsymbol{X}_{i,j}^t - \boldsymbol{X}_{\text{worst}}^t \right| / [(f_i - f_w) + \varepsilon] \right\} & \text{if } f_i = f_g \end{cases} \tag{32}$$

In the equation: $\boldsymbol{X}_{\text{best}}$ and $\boldsymbol{X}_{\text{worst}}$ represent the current global optimum and worst positions, respectively; β represents the random number used to control the step size, which follows a normal distribution; $K \in [-1, 1]$ represents a random number used to control the step size and movement direction; ε represents a constant used to avoid a zero

denominator; and f_i, f_w, and f_g represent the current individual fitness, worst fitness, and best fitness, respectively.

$$\beta = fitness_{\text{best}} - (fitness_{\text{best}} - fitness_{\text{worst}}) \cdot [(\text{iter}_{\max} - t)/\text{iter}_{\max}]^{1.5} \tag{33}$$

$$K = (fitness_{\text{best}} - fitness_{\text{worst}}) \cdot \exp\left[-20 \cdot \tan(t/\text{iter}_{\max})^2\right] \cdot (2 \cdot \text{rand} - 1) \tag{34}$$

In the equation: $fitness_{\text{best}}$ and $fitness_{\text{worst}}$ represent the optimal and worst fitness function values, respectively.

(4) SSA is prone to getting stuck at local optima, making it difficult to find the global optimum; therefore, the Levy strategy is used to optimize Equation (38). During the search for the optimal solution, the Levy flight can perform both short-range local searches and long-range global searches, effectively addressing the problem of SSA getting stuck at local optima. The expression is

$$\boldsymbol{X}_{i,j}(t+1) = \begin{cases} Levy(d) \cdot \boldsymbol{X}_{\text{best}}(t) + \beta \cdot \left|\boldsymbol{X}_{i,j}(t) - Levy(d) \cdot \boldsymbol{X}_{\text{best}}(t)\right| & if \quad f_i \neq f_g \\ \boldsymbol{X}_{i,j}(t) + K \cdot \left\{\left|\boldsymbol{X}_{i,j}(t) - \boldsymbol{X}_{\text{worst}}(t)\right|/[(f_i - f_w) + \varepsilon]\right\} & if \quad f_i = f_g \end{cases} \tag{35}$$

$$Levy(d) = 0.01 \cdot r_1 \cdot \sigma/|r_2|^{1/\beta} \tag{36}$$

$$\sigma = \left\{\frac{\Gamma(1+\beta) \cdot \sin(\pi\beta/2)}{\Gamma[(1+\lambda)/2] \cdot \beta \cdot 2^{(\beta-1)/2}}\right\}^{1/\beta} \tag{37}$$

In the equation: d represents the vector dimension; $r_1, r_2 \in [0, 1]$ represents a random number; β represents a constant; Γ represents the gamma function; and $Levy(d)$ represents the formula for the Levy strategy.

Algorithms such as GA, PSO, and SSA were selected for comparison with ISSA, with a maximum number of iterations set to 1000 and a population size of 30. The optimization performance of each algorithm was evaluated using the nine benchmark test functions shown in Table 6, including three unimodal test functions ($F_1 \sim F_3$), three multimodal test functions ($F_4 \sim F_6$), and three multimodal test functions with fixed dimensions ($F_7 \sim F_9$). The convergence curves of each algorithm for the different test functions are shown in Figure 12.

Table 6. Benchmarking functions.

No.	Test Functions	Optimal Value	Domain	Dimension
F_1	$F_1(x) = \sum_{i=1}^{n}\|x_i\| + \prod_{i=1}^{n}\|x_i\|$	0	$[-10, 10]$	30
F_2	$F_2(x) = \max_i\{\|x_i\|, 1 \le i \le n\}$	0	$[-100, 100]$	30
F_3	$F_3(x) = \sum_{i=1}^{n}([x_i + 0.5])^2$	0	$[-100, 100]$	30
F_4	$F_4(x) = \sum_{i=1}^{n}\left[x_i^2 - 10\cos(2\pi x_i) + 10\right]$	0	$[-5.12, 5.12]$	30
F_5	$F_5(x) = \frac{1}{4000}\sum_{i=1}^{n}x_i^2 - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	0	$[-600, 600]$	30

Table 6. *Cont.*

No.	Test Functions	Optimal Value	Domain	Dimension
F_6	$F_6(x) = 0.1\left\{\sin^2(3\pi x_1) + \sum_{i=1}^{n-1} \begin{array}{l}(x_i-1)^2[1+\sin^2(3\pi x_{i+1})] \\ +(x_n-1)^2[1+\sin^2(2\pi x_n)]\end{array}\right\} + \sum_{i=1}^{n} u(x_i,5,100,4)$	0	$[-50,50]$	30
F_7	$F_7(x) = \sum_{i=1}^{11}\left[a_i - x_1(b_i^2+b_i x_2)/b_i^2+b_i x_3+x_4\right]^2$	0.000 3	$[-5,5]$	4
F_8	$F_8(x) = -\sum_{i=1}^{4} c_i \exp\left[-\sum_{j=1}^{3} a_{ij}\left(x_j - p_{ij}\right)^2\right]$	−3.862 8	$[0,1]$	3
F_9	$F_9(x) = -\sum_{i=1}^{5}\left[(x-a_i)(x-a_i)^{\mathrm{T}} + c_i\right]^{-1}$	−10.153 2	$[0,10]$	4

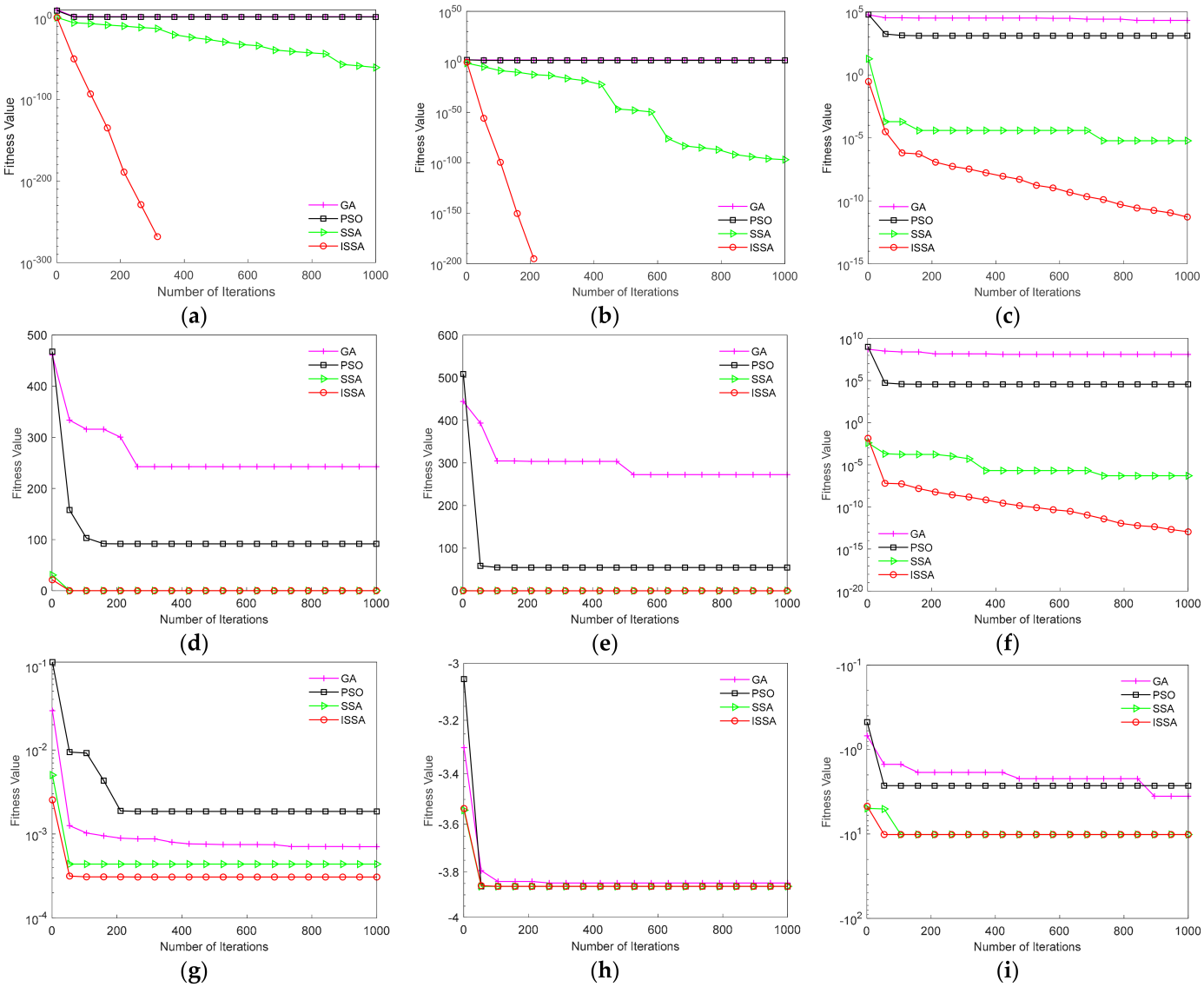

Figure 12. Compare convergence curves: (**a**) F_1. (**b**) F_2. (**c**) F_3. (**d**) F_4. (**e**) F_5. (**f**) F_6. (**g**) F_7. (**h**) F_8. (**i**) F_9.

Table 7 shows the optimal value R_b, mean R_m, and standard deviation R_{std} obtained by optimizing the objective function using different algorithms. These values are pre-

sented to demonstrate ISSA's convergence efficiency and optimization performance; the best optimization results R_b for each benchmark function are highlighted in bold.

Table 7. Compare search results.

Functions	Indicators	Values			
		GA	PSO	SSA	ISSA
F_1	R_b	50.10	21.98	2.36×10^{-61}	0.00
	R_m	1.89×10^{7}	1.69×10^{6}	5.81×10^{-3}	5.68×10^{-3}
	R_{std}	4.73×10^{8}	4.52×10^{7}	1.46×10^{-1}	1.65×10^{-1}
F_2	R_b	58.04	26.93	1.23×10^{-97}	0.00
	R_m	69.03	27.22	1.21×10^{-4}	5.26×10^{-4}
	R_{std}	6.40	2.45	1.65×10^{-3}	1.59×10^{-2}
F_3	R_b	2.14×10^{4}	1.30×10^{3}	5.99×10^{-6}	5.17×10^{-12}
	R_m	3.03×10^{4}	1.61×10^{3}	2.40×10^{-2}	6.82×10^{-4}
	R_{std}	6.31×10^{3}	3.14×10^{3}	6.13×10^{-1}	1.34×10^{-2}
F_4	R_b	242.59	91.66	0.00	0.00
	R_m	261.44	100.44	3.88×10^{-2}	2.16×10^{-2}
	R_{std}	38.75	32.63	1.00	6.76×10^{-1}
F_5	R_b	272.40	54.43	0.00	0.00
	R_m	294.62	56.09	1.88×10^{-4}	5.30×10^{-5}
	R_{std}	33.96	18.48	4.60×10^{-3}	1.67×10^{-3}
F_6	R_b	1.30×10^{8}	3.63×10^{4}	4.96×10^{-7}	1.11×10^{-13}
	R_m	1.69×10^{8}	1.93×10^{6}	7.20×10^{-5}	1.41×10^{-4}
	R_{std}	8.29×10^{7}	3.97×10^{7}	2.67×10^{-4}	1.33×10^{-3}
F_7	R_b	7.07×10^{-4}	1.86×10^{-3}	4.38×10^{-4}	3.07×10^{-4}
	R_m	1.06×10^{-3}	3.27×10^{-3}	4.54×10^{-4}	3.26×10^{-4}
	R_{std}	2.03×10^{-3}	4.68×10^{-3}	2.18×10^{-4}	1.46×10^{-4}
F_8	R_b	−3.85	−3.86	−3.86	−3.86
	R_m	−3.84	−3.86	−3.86	−3.86
	R_{std}	5.00×10^{-2}	4.61×10^{-2}	1.57×10^{-2}	2.52×10^{-2}
F_9	R_b	−3.57	−2.68	−10.15	−10.15
	R_m	−2.17	−2.67	−9.78	−10.08
	R_{std}	5.69×10^{-1}	1.13×10^{-1}	1.32	5.49×10^{-1}

As shown in Figure 12 and Table 7, regardless of whether the test function is single-peaked or multi-peaked, the ISSA demonstrates significantly better optimization performance than the comparison algorithms and offers greater practical value. It exhibits faster convergence rates across different test functions, with the optimized values coming very close to the theoretical values and the standard deviation approaching zero. This indicates that the optimization performance of the algorithm has been significantly improved through the enhancements made to the SSA. In summary, the effectiveness and superiority of the proposed ISSA algorithm have been validated.

5. Case Studies

The performance of the proposed unbalanced few-shot diesel engine intelligent fault diagnosis method (SMVMD-ISSA-DHKELM) was validated using experimental data. The comparison methods included ISSA-DHKELM (which extracts spectral amplitude fault features directly from the data and is trained using ISSA-DHKELM), SMVMD-DHKELM (which processes data using SMVMD to extract spectral amplitude features from filtered

components and is trained using DHKELM), SMVMD-KELM (which processes data using SMVMD to extract spectral amplitude features from filtered components and is trained using KELM), and 1DCNN (which directly extracts spectral amplitude fault features from the data and is trained using 1DCNN).

Properly partitioning a dataset to meet the requirements of unbalanced few-shot scenarios is a prerequisite and foundation for validating the performance of the proposed method. According to Reference [21], in fault diagnosis tasks, an unbalanced sample refers to a scenario where there is a significant disparity in the number of training samples across different states. Typically, an unbalanced sample problem occurs when the ratio of the majority class to the minority class in the sample is significantly greater than 1:1; in the context of diesel engine fault diagnosis, this corresponds to a ratio of normal to fault data that is significantly greater than 1:1. According to References [22,23], in the field of fault diagnosis, a few-shot scenario refers to a situation where the number of samples for each fault type in the training set is only in the tens, and the proportion $\alpha \leq 0.5$ of the training samples relative to the total sample size is small.

5.1. Fuel Supply System Data Validation

The dataset includes seven types of faults: normal, misfire in one-cylinder, clogged air filter, fuel injector dripping, insufficient fuel supply from the fuel pump, misfires in one and two cylinders, and clogged air filter combined with fuel injector dripping, for a total of 1500 samples. Specifically, the training set contains 200 normal samples, with 50 samples for each of the remaining six fault types; the test set contains 400 normal samples, with 100 samples for each of the remaining six fault types. In this case, the training set accounts for a proportion of 1/3 of the total sample size, and the ratio of the majority class to the minority classes is 4:1.

It should be noted that the total number of samples in the final training and test sets was fixed (500; 1000), while the number of samples in each category (200, 50; 400, 100) was approximate. This is because, during the splitting of the dataset, to ensure the randomness of the split, the MATLAB 2023b command "n = randperm (N)" was used to generate a random permutation of integers from 1 to $N = 1500$ with no duplicate elements, and the resulting permutation was mapped to the indices of the dataset. The samples corresponding to the first 500 indices were designated as the training set, and those corresponding to the last 1000 indices were designated as the test set. The MATLAB commands used were "train_X = data (n (1:500), :); test_X = data (n (501: end), :);", where "data" represents the dataset. Since samples belonging to different fault types appear completely randomly across indices 1 through 1500, the specific number of samples for each type in the resulting training and test sets may deviate slightly from the preset values. However, this still ensures the reasonableness of the sample division, and the resulting division meets the requirements for an unbalanced few-shot dataset.

SMVMD processes samples from the training and test sets, with the parameter settings shown in Table 8. Taking the first sample from the training set as an example, the decomposition results are shown in Figure 13.

Table 8. Parameters setting for SMVMD method.

Method	$\alpha_{\min}$	$\alpha_{\max}$	ε_1	ε_2	γ	τ	K
SMVMD	1	2×10^4	1×10^{-10}	1×10^{-10}	$\sqrt{2}$	0	–

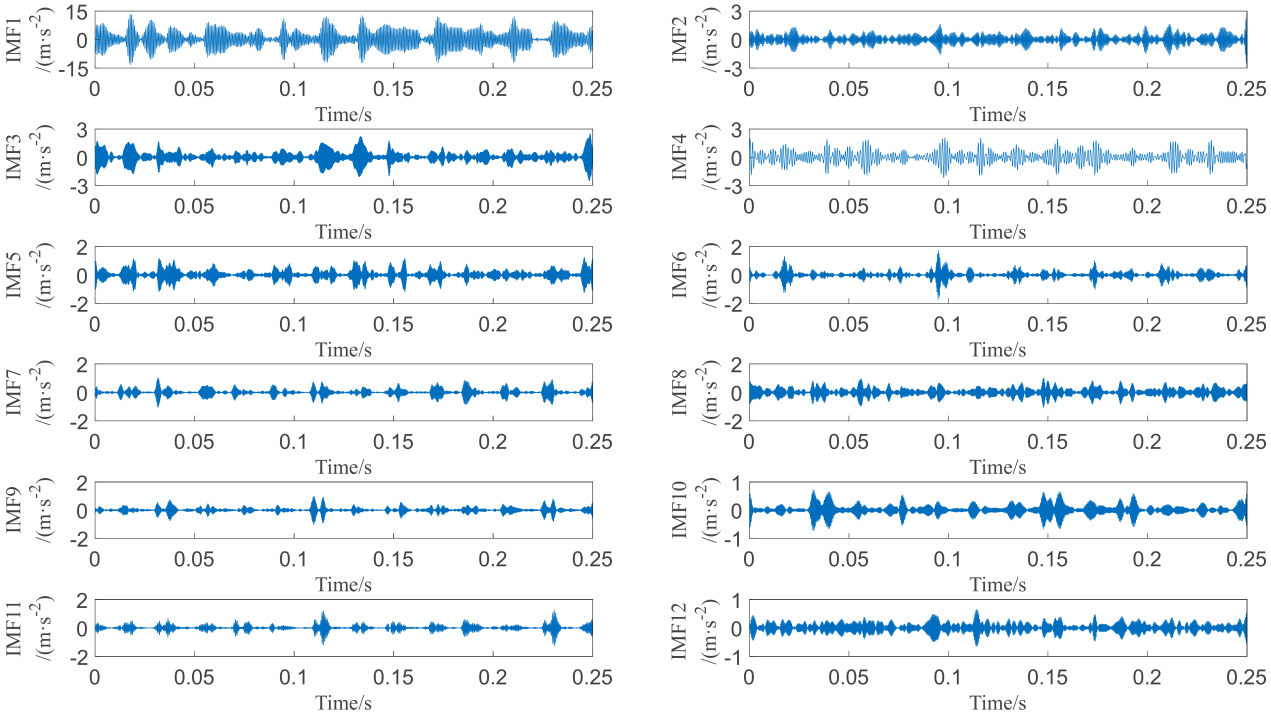

Figure 13. Result after processing the 1st sample in the training set using SMVMD.

As shown in Figure 13, a total of 12 IMF components were obtained after processing with SMVMD. Next, we need to select the most valuable component from among them.

After SMVMD processing of the samples, if it is difficult to identify regular impact features related to faults in a component, the envelope entropy value is relatively high. Conversely, if a component contains abundant fault information, the envelope entropy value is relatively low. The envelope entropy values for the 12 components in Figure 13 were calculated, and the results are shown in Figure 14.

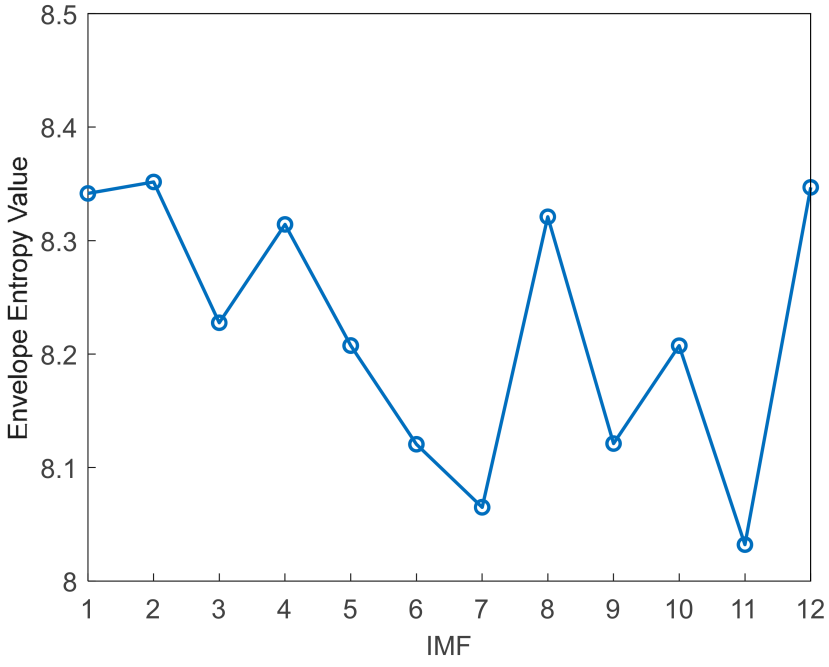

Figure 14. Envelope entropy values of the components in the 1st sample of the training set.

As shown in Figure 14, compared with the other components, IMF11 has the lowest envelope entropy value, at 8.032. Therefore, IMF11 is selected as the desired component. The remaining samples in the dataset are processed in the same manner until the most valuable component among all samples is identified.

Feature extraction is performed on the selected components using the Fast Fourier Transform (FFT), and their spectral amplitudes are used as fault feature samples. Taking the first sample of each fault type in the training set as an example, the spectral amplitudes of the selected components are shown in Figure 15.

As shown in Figure 15, the spectral amplitudes of the most valuable IMF components are intuitively presented in the form of histograms. Although there are slight differences in the spectral amplitudes among different fault types, it is difficult for humans to effectively distinguish between them.

To prevent the accuracy of fault diagnosis from being affected by differences in the orders of magnitude among various variables, the spectral amplitudes of the components selected after filtering are normalized by uniformly scaling all data samples to $[0, 1]$, resulting in a feature dataset. The expression is

$$x' = (x - x_{\min})/(x_{\max} - x_{\min}) \tag{38}$$

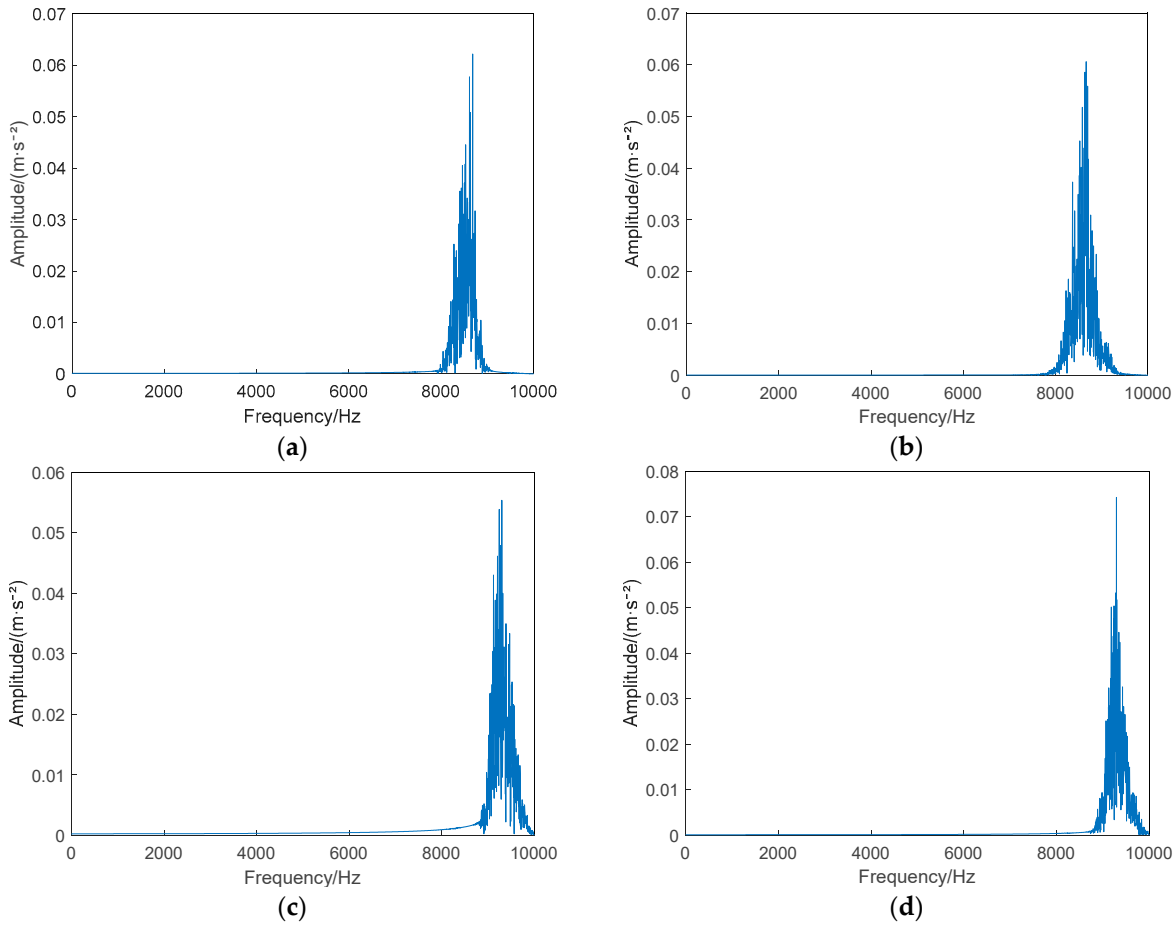

Figure 15. *Cont.*

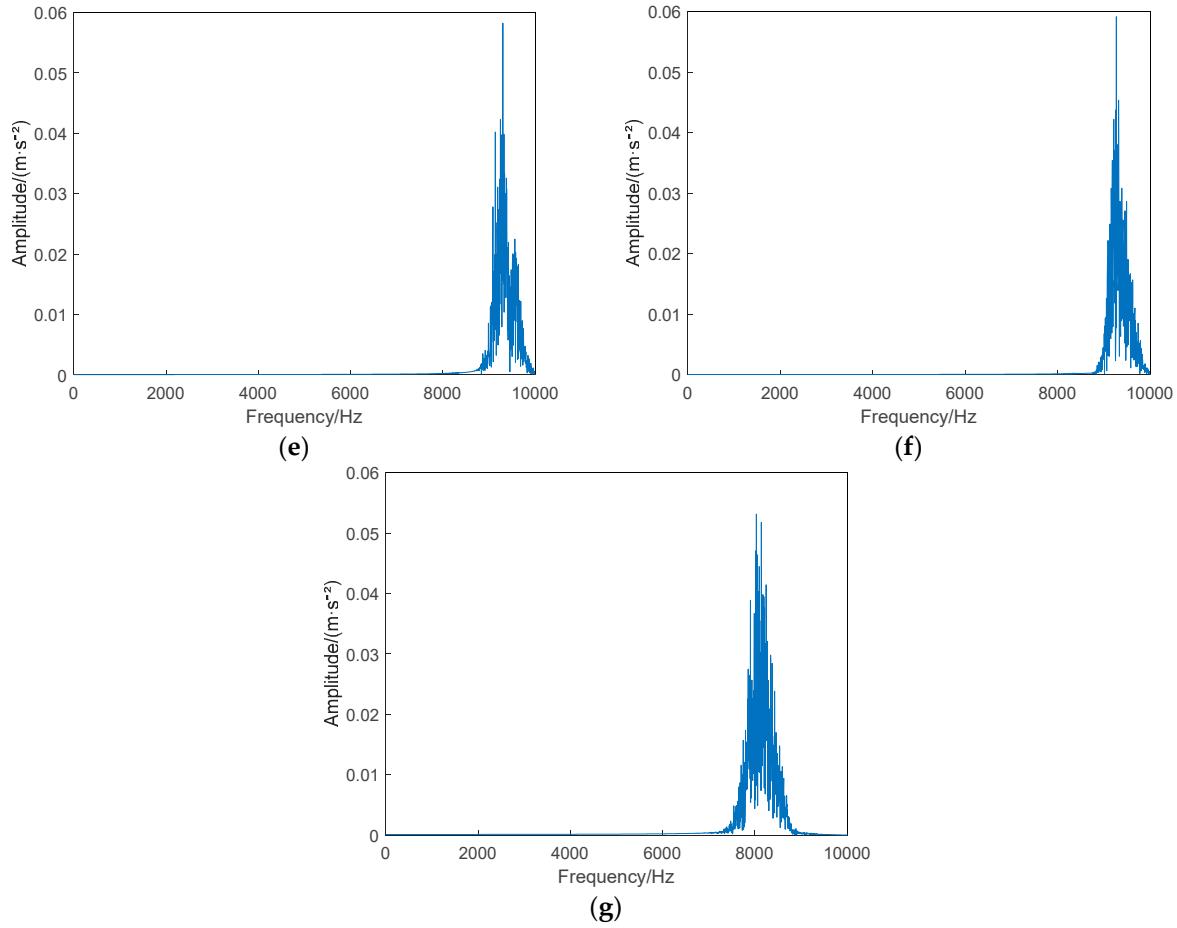

Figure 15. Spectral amplitude of selected components: (**a**) Normal. (**b**) One-cylinder misfire. (**c**) Clogged air filter. (**d**) Dripping injector. (**e**) Insufficient fuel supply to the injection pump. (**f**) One-cylinder misfire and two-cylinder misfire. (**g**) Clogged air filter and dripping injector.

In the equation: the result x' is the normalized value, and the minimum and maximum values of the unnormalized sample x are $x_{\min}$ and $x_{\max}$, respectively. x' is used as the input data for DHKELM, and the fault category labels are encoded using one-hot encoding.

Hyperparameter optimization for DHKELM was performed using ISSA, with SSA as the comparison algorithm. The error rate of fault diagnosis on the training set was used as the fitness function. Figure 16 shows the curves of the fitness function values for different optimization algorithms as a function of the number of iterations. Based on practical considerations, the following parameters were set for both ISSA and SSA: a sparrow population size of 10, an early warning value of 0.8, a discoverer ratio of 0.7, a scout ratio of 0.2, and a maximum number of iterations of 30.

As shown in Figure 16, as the optimization algorithm optimizes the hyperparameters in DHKELM, it yields more ideal hyperparameter combinations, gradually improving the model's fault diagnosis accuracy. Compared with SSA, ISSA exhibits higher optimization accuracy and faster convergence, indicating that the proposed ISSA performs better.

Based on the results obtained from ISSA optimization, the hyperparameters in DHKELM were set as follows: the number of hidden layers was 3, with 80, 95, and 88 nodes in each layer, respectively; the regularization coefficient was 380.793; the penalty coefficient was 705.92; the weight of the Gaussian kernel was 0.57; the kernel parameter was 753.91; and the constant and exponential parameters of the polynomial kernel were 770.33 and 9, respectively.

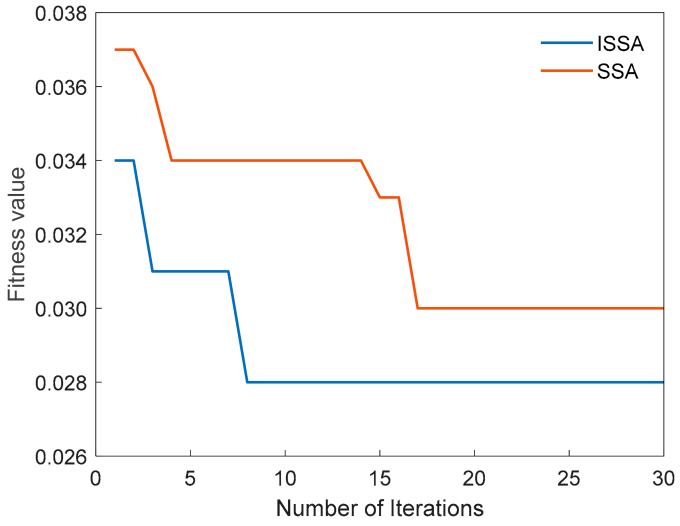

Figure 16. Hyperparameter optimization in DHKELM model using ISSA and SSA respectively.

Considering the impact of different numbers of DHKELM network layers on diagnostic results, DHKELM models with Layer M were constructed separately, $M = 2, 3, 4, 5, 6$. The fault diagnosis accuracy rates on the training set for different numbers of network layers are shown in Figure 17.

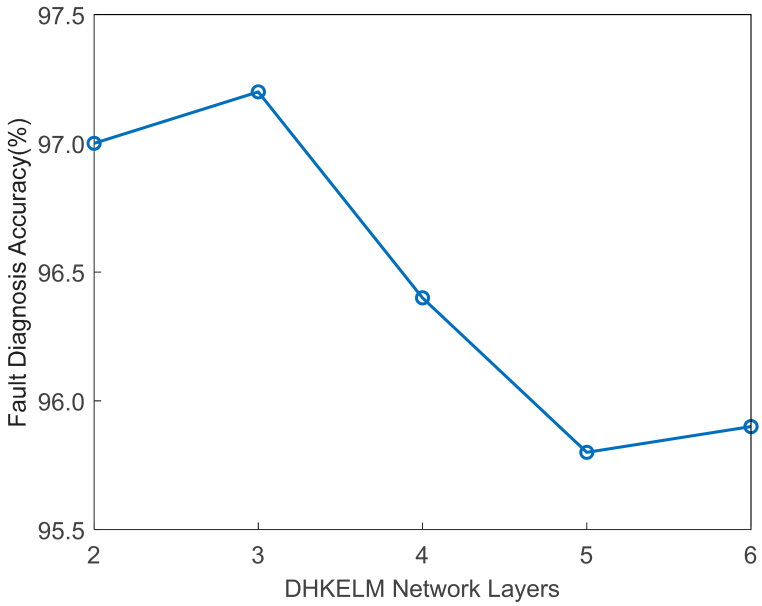

Figure 17. Fault diagnosis accuracy at different network layers.

As shown in Figure 17, the accuracy reaches its maximum when the network has 3 layers. As the network depth increases, the accuracy fluctuates to some extent, and when the network has 5 layers, the accuracy drops to its minimum. The results indicate that increasing or decreasing the number of layers does not necessarily improve diagnostic accuracy; the most appropriate number of layers must be selected based on actual conditions. In this case, a 3-layer network is optimal, which is consistent with the optimization results.

The proposed SMVMD-ISSA-DHKELM intelligent fault diagnosis method for diesel engines under unbalanced few-shot conditions, as well as methods such as ISSA-DHKELM, SMVMD-DHKELM, SMVMD-KELM, VMD-DHKELM and 1DCNN, were used to identify various fault types in the test set. To minimize the impact of randomness and ensure the reliability and reproducibility of the experimental results, after partitioning the dataset,

10 independent and mutually exclusive experiments were conducted for all methods, yielding different trained fault diagnosis models. The test set was input into the various trained models, and the average of the experimental results was taken as the final performance metric for each method. The standard deviation of the results was also calculated to reflect stability. The experimental results are shown in Table 9, and the visualization of the diagnostic results and the confusion matrix are shown in Figure 18. Fault types 1 through 7 correspond to the seven fault types defined in the dataset.

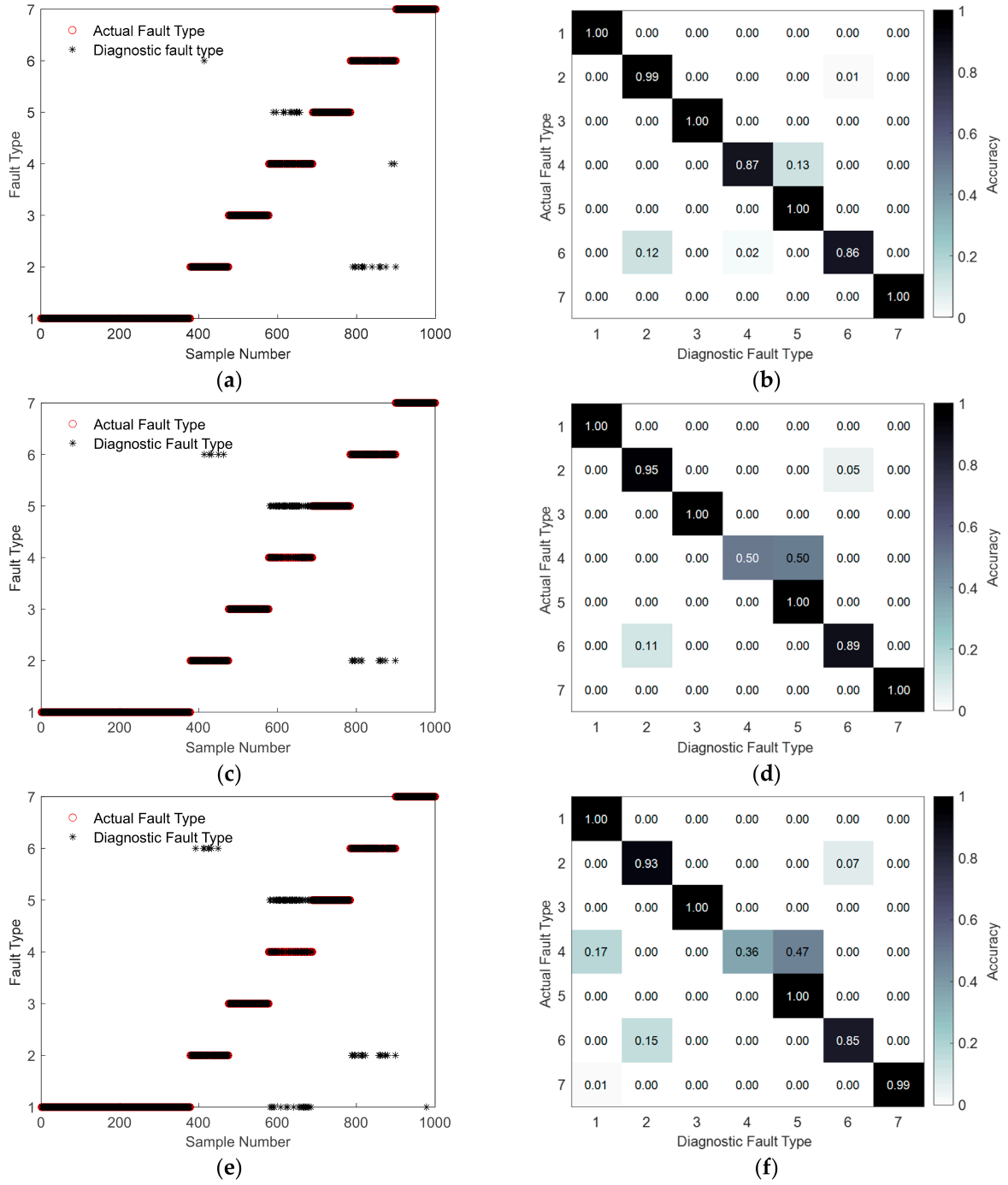

Figure 18. *Cont.*

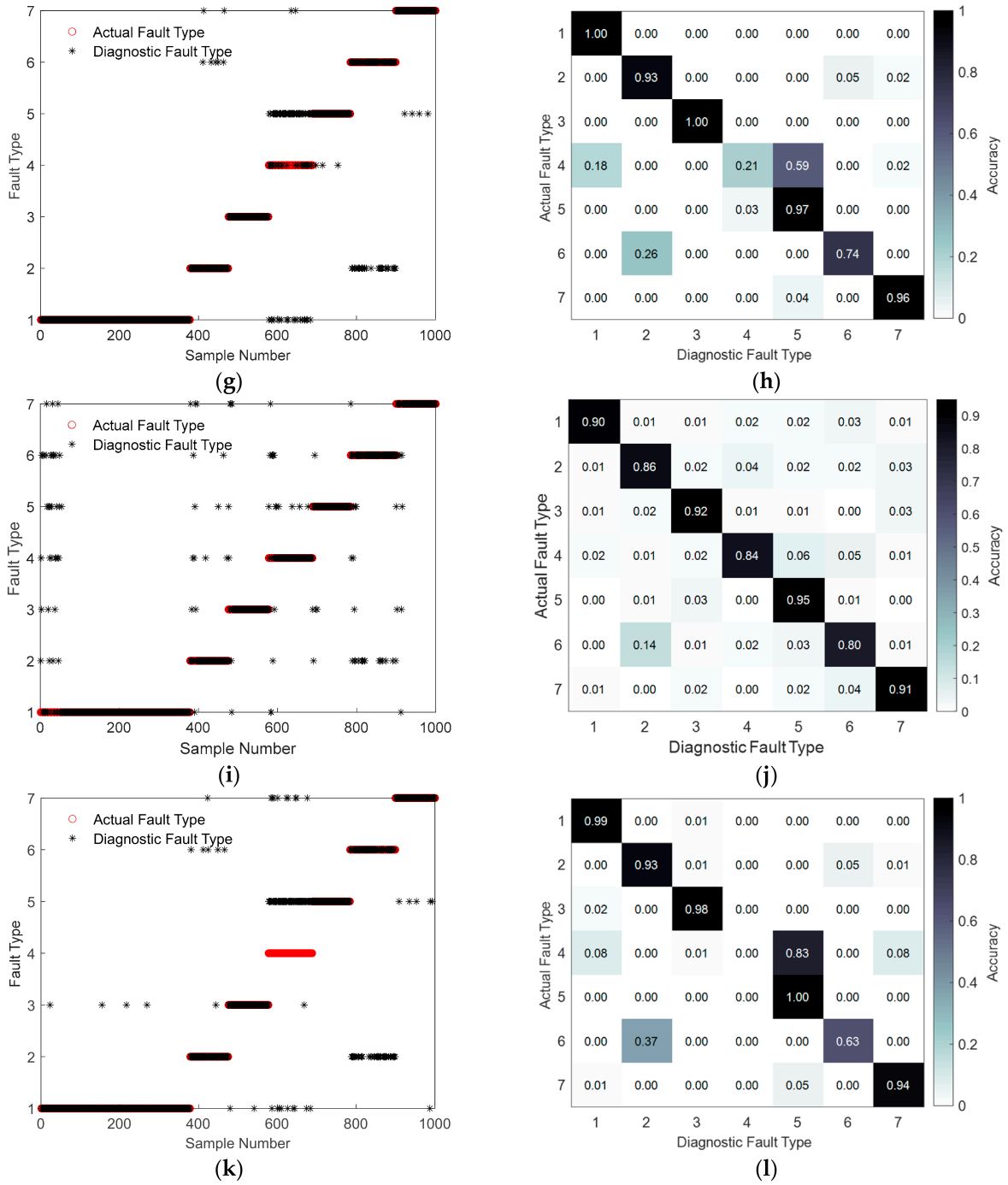

Figure 18. Fault type identification result and confusion matrix for different methods: (**a**) SMVMD-ISSA-DHKELM identification result. (**b**) SMVMD-ISSA-DHKELM confusion matrix. (**c**) ISSA-DHKELM identification result. (**d**) ISSA-DHKELM confusion matrix. (**e**) SMVMD-DHKELM identification result. (**f**) SMVMD-DHKELM confusion matrix. (**g**) SMVMD-KELM identification result. (**h**) SMVMD-KELM confusion matrix. (**i**) VMD-DHKELM identification result. (**j**) VMD-DHKELM confusion matrix. (**k**) 1DCNN identification result. (**l**) 1DCNN confusion matrix.

As shown in Figure 18 and Table 9, under unbalanced few-shot conditions, the proposed diesel engine fault diagnosis method achieved an overall identification accuracy of 96.90% on the test set, demonstrating the highest accuracy compared to the baseline methods. This is because the proposed method not only improves sample quality at the data level but also optimizes the feature extraction process at the model level. Specifically, the diagnostic accuracy reached 100% for four fault types: normal, clogged air filter, insufficient

fuel supply from the fuel pump, and injector dripping; 99% for misfire in one cylinder; and 87% and 86% for injector dripping and misfires in one and two cylinders, respectively. Compared with the ISSA-DHKELM method, the proposed method improved diagnostic accuracy by 4.20%, demonstrating the necessity and effectiveness of using SMVMD to process vibration signals and select the most valuable components within them. Compared with the SMVMD-DHKELM method, the proposed method improved diagnostic accuracy by 6.50%, demonstrating that the hyperparameter optimization in DHKELM via the proposed ISSA is essential for fully leveraging the model's fault diagnosis performance. Compared with the VMD-DHKELM method, the SMVMD-DHKELM method improved diagnostic accuracy by 1.70%, demonstrating that the SMVMD method outperforms the traditional VMD method. Compared with the SMVMD-KELM method, the SMVMD-DHKELM method improved diagnostic accuracy by 3.60%, demonstrating that DHKELM—which builds upon KELM by incorporating deep learning concepts to form a multi-layer network feature extraction model and using a linear weighting method to combine different kernel functions to compensate for the shortcomings of a single kernel function—can more effectively extract features from data samples and achieve better fault diagnosis results. Compared with the 1DCNN method, the proposed method improves diagnostic accuracy by 14.10%. This indicates that, under unbalanced few-shot conditions, traditional methods struggle to meet the practical application requirements of diesel engine fault diagnosis, necessitating the development of new fault diagnosis methods with high accuracy and strong generalizability. The method proposed in this paper not only achieves high accuracy but also demonstrates superior performance in this application scenario.

Table 9. Accuracy of each method under test set (fuel supply system).

Model	Accuracy/%
SMVMD-ISSA-DHKELM	96.90 ± 0.40
ISSA-DHKELM	92.70 ± 1.10
SMVMD-DHKELM	90.40 ± 2.80
VMD-DHKELM	88.70 ± 2.60
SMVMD-KELM	86.80 ± 3.30
1DCNN	82.80 ± 5.30

To further validate the performance of the method proposed in this paper, we compare it with existing research results in the literature, including References [24,25]. To minimize the impact of randomness and ensure the reliability and reproducibility of the experimental results, after partitioning the dataset, we conducted 10 independent and mutually exclusive experiments for all methods, yielding different pre-trained fault diagnosis models. We input the test set into each of the pre-trained models, took the average of the experimental results as the final performance metric for each method, and calculated the standard deviation of the results to reflect stability. The comparison results of each method on the test set are shown in Table 10.

Table 10. Comparing the method proposed in this paper with existing results in the literature.

Model	SMVMD-ISSA-DHKELM	Reference [24]	Reference [25]
Accuracy/%	96.90 ± 0.40	87.00 ± 1.20	89.50 ± 1.30

As shown in Table 10, compared with existing research results in the literature, the method proposed in this paper still achieves the highest accuracy on the test set.

To further validate the performance of the proposed SMVMD-ISSA-DHKELM method under few-shot conditions, we maintained a 4:1 ratio between the majority class (normal) and the minority class (fault) and kept the proportion of the training set relative to the total sample size constant. We then randomly selected samples from the remaining six fault types—10, 30, and 70 samples, respectively—to serve as the model's training set for experimentation. The visualization results of accuracy rates under different numbers of fault samples are shown in Figure 19, and the accuracy of each method on the test set is presented in Table 11. In the following figures, Methods 1 through 5 represent SMVMD-ISSA-DHKELM, ISSA-DHKELM, SMVMD-DHKELM, SMVMD-KELM, and 1DCNN, respectively.

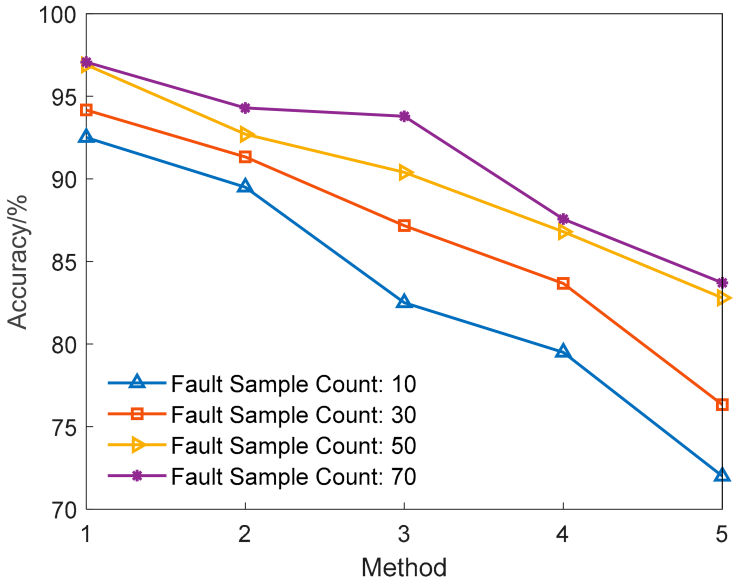

Figure 19. Accuracy of each method for different number of fault samples.

Table 11. Accuracy of each method for different sample sizes (fuel supply system).

The Number of Samples of Each Failure Type in the Training Set	Methods	Accuracy/%
10	SMVMD-ISSA-DHKELM	92.50
	ISSA-DHKELM	89.50
	SMVMD-DHKELM	82.50
	SMVMD-KELM	79.50
	1DCNN	72.00
30	SMVMD-ISSA-DHKELM	94.17
	ISSA-DHKELM	91.33
	SMVMD-DHKELM	87.17
	SMVMD-KELM	83.67
	1DCNN	76.33
70	SMVMD-ISSA-DHKELM	97.07
	ISSA-DHKELM	94.29
	SMVMD-DHKELM	93.79
	SMVMD-KELM	87.57
	1DCNN	83.71

As shown in Figure 19 and Table 11, as the number of samples increased from 10 to 70, the model was trained more thoroughly, and the fault diagnosis accuracy of each method improved. Regardless of the sample conditions, the proposed method is still able to learn the distribution of real samples relatively well and demonstrates the best fault diagnosis

performance. Even when there were only 10 samples of each fault type in the training set, the proposed method still achieved an accuracy of 92.50% on the test set. Comparative experiments further demonstrate that the proposed method can meet the requirements for fault diagnosis under few-shot conditions.

To further validate the performance of the proposed SMVMD-ISSA-DHKELM method under unbalanced sample conditions, we kept the training set at one-third of the total sample size and maintained the number of samples for each of the remaining six fault types at 10 for the model's training set. We then conducted experiments by setting the majority-to-minority class ratio to 2:1, 10:1, and 30:1, respectively. The visualization results of the accuracy of each method on the test set are shown in Figure 20, and the accuracy values are presented in Table 12.

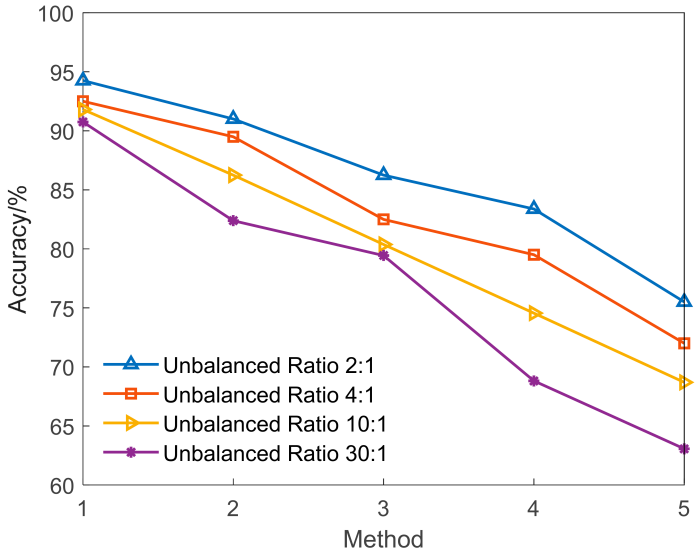

Figure 20. Visualization of the accuracy of each method for unbalanced dataset.

Table 12. Accuracy of each method for unbalanced dataset (fuel supply system).

Unbalanced Dataset	Methods	Accuracy/%
2:1	SMVMD-ISSA-DHKELM	94.25
	ISSA-DHKELM	91.00
	SMVMD-DHKELM	86.25
	SMVMD-KELM	83.38
	1DCNN	75.50
10:1	SMVMD-ISSA-DHKELM	91.81
	ISSA-DHKELM	86.25
	SMVMD-DHKELM	80.38
	SMVMD-KELM	74.56
	1DCNN	68.69
30:1	SMVMD-ISSA-DHKELM	90.75
	ISSA-DHKELM	82.39
	SMVMD-DHKELM	79.44
	SMVMD-KELM	68.81
	1DCNN	63.06

As shown in Figure 20 and Table 12, as the unbalanced ratio between the majority and minority classes increased from 2:1 to 30:1, the fault diagnosis accuracy of the proposed method decreased by only 3.50%, whereas the fault diagnosis accuracies of the comparison

methods decreased by 8.61%, 6.81%, 14.57%, and 12.44%, respectively. This indicates that the proposed method exhibits the best stability. Under any class unbalanced ratio, the proposed method demonstrates the best fault diagnosis performance, indicating that it retains strong adaptability even when sample ratios are unbalanced. Even when the number of fault samples per class in the training set was only 10 and the imbalance ratio reached as high as 30:1, the proposed method still achieved an accuracy of 90.75%, far exceeding that of the other four methods. This comparative test demonstrates that the proposed method can meet the requirements for fault diagnosis under unbalanced sample conditions.

5.2. Intake and Exhaust System Data Validation

The dataset includes vibration data from four fault conditions of a diesel engine with valve clearances of 0.3 mm, 0.4 mm, 0.5 mm, and 0.7 mm, for a total of 1050 samples. The training set contains 200 samples in the normal state and 50 samples each for the remaining three fault states; the test set contains 400 samples in the normal state and 100 samples each for the remaining three fault states. That is, the training set accounts for one-third of the total samples, with a majority-to-minority class ratio of 4:1. It should be noted that while the total number of samples in the actual training and test sets is fixed (350; 700), the number of samples per state (200, 50; 400, 100) is approximately equal. The specific reasons for this are similar to those in Section 5.1, but this still ensures the rationality of the sample partitioning, and the resulting partition meets the requirements for unbalanced few-shot datasets.

SMVMD processes samples from the training and test sets using the same parameter settings as in Section 5.1. Taking the first sample from the processed training set as an example, the results are shown in Figure 21.

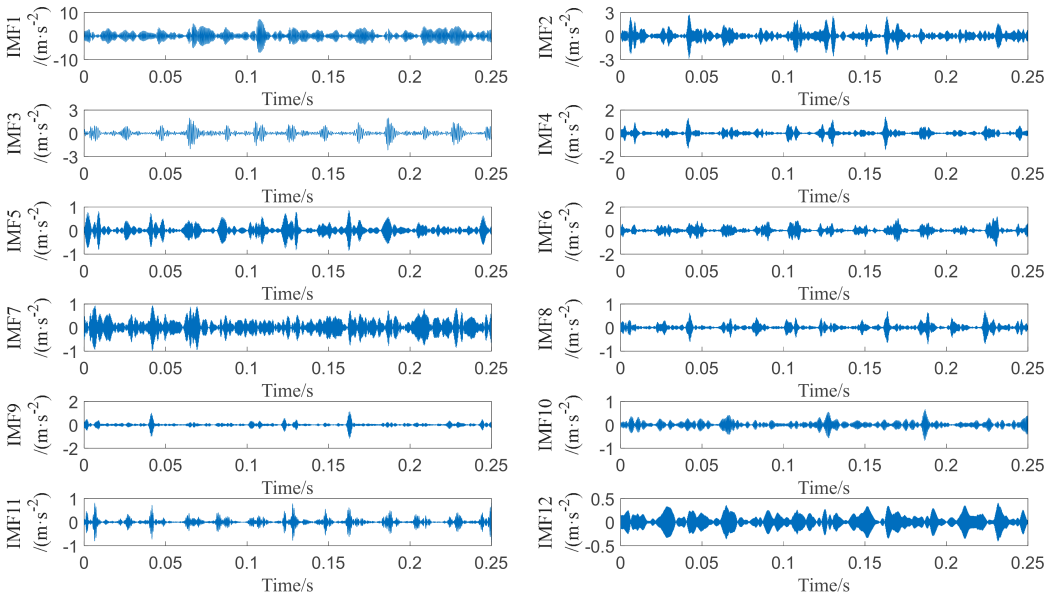

Figure 21. Result after SMVMD for the 1st sample in the training set.

As shown in Figure 21, a total of 12 IMF components were obtained after processing with SMVMD. To identify the most valuable component from the decomposition results, the envelope entropy of each component was calculated. As shown in Figure 22.

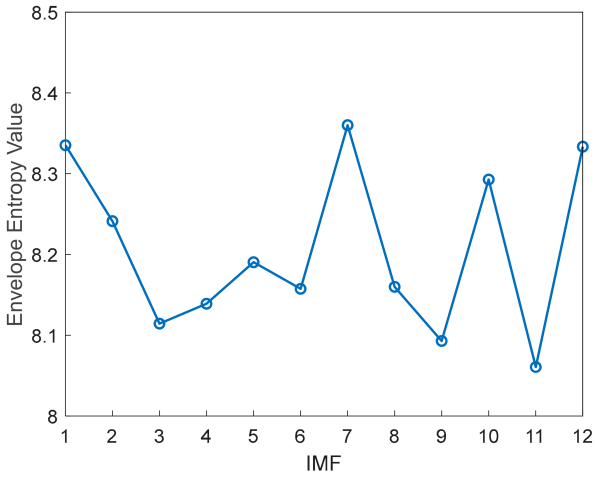

Figure 22. Envelope entropy value for each component of the 1st sample of the training set.

As shown in Figure 22, compared with the other 11 components, IMF11 has the lowest envelope entropy value, at 8.061. Therefore, IMF11 is selected as the desired component. The remaining samples in the dataset are processed in the same manner until the most valuable component among all samples is identified. The components selected through feature extraction using the FFT are used, with the spectral amplitudes of each component serving as fault feature samples. Taking the first sample representing different fault states in the training set as an example, the spectral amplitudes of the selected components are shown in Figure 23.

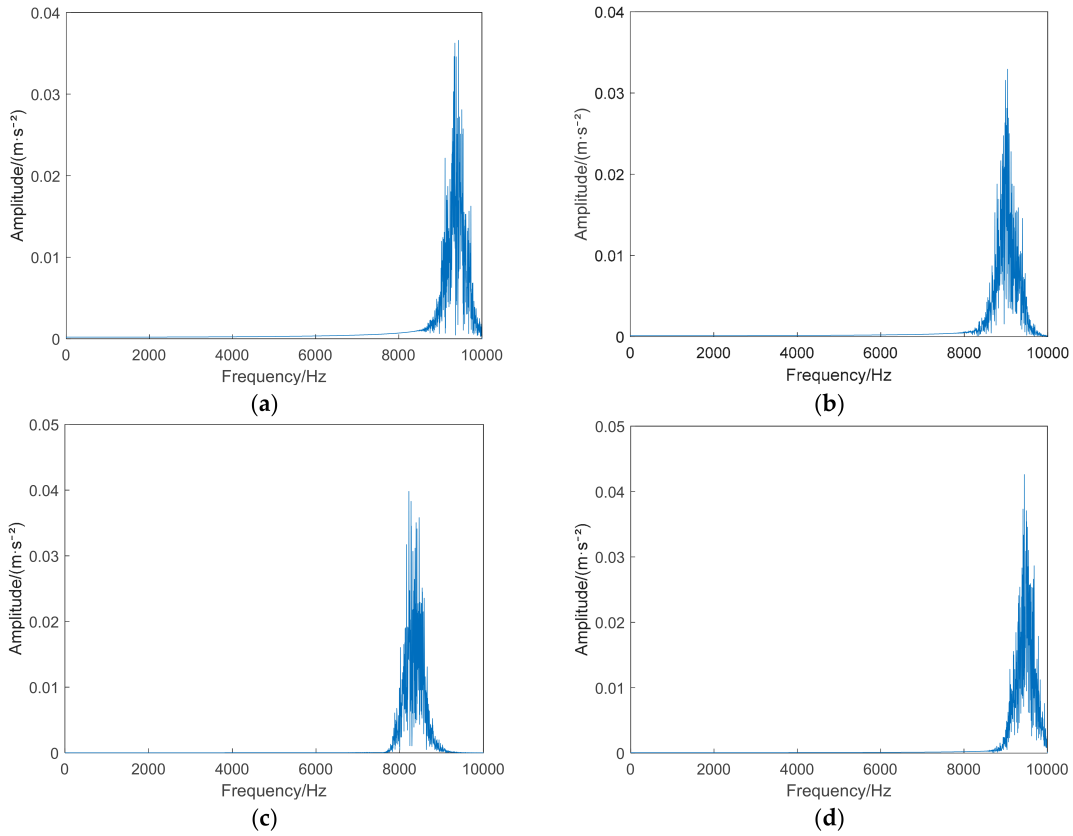

Figure 23. Spectral amplitude of the most valuable component: (**a**) Healthy state. (**b**) Attention state. (**c**) Anomaly state. (**d**) Deteriorated state.

As shown in Figure 23, the spectral amplitudes of the most valuable components are intuitively presented in the form of histograms. The spectral amplitudes of the selected components are normalized as shown in Equation (38), uniformly restricting all data samples to the interval $[0, 1]$, thereby yielding the feature dataset.

ISSA was used for hyperparameter optimization of DHKELM, with SSA serving as the comparison algorithm. The error rate in fault diagnosis using the training set was used as the fitness function. Figure 24 shows the curves of the fitness function values for different optimization algorithms as a function of the number of iterations. The parameter settings for SSA and ISSA were the same as those in Section 5.1.

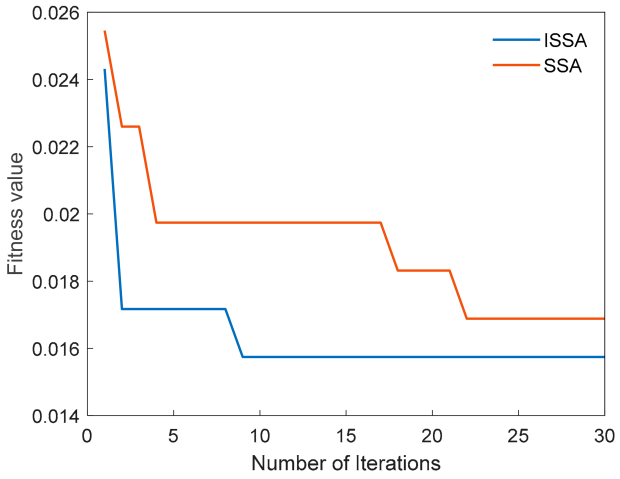

Figure 24. Hyperparameters in the DHKELM model optimized by ISSA and SSA respectively.

As shown in Figure 24, ISSA exhibits a faster convergence rate and higher optimization accuracy compared with SSA. Based on the optimization results from ISSA, the hyperparameters for DHKELM were set as follows: three hidden layers with 41, 55, and 38 nodes, respectively; a regularization coefficient of 614.203; a penalty coefficient of 566.699; a Gaussian kernel weight of 0.77; a kernel parameter of 713.818; and constant and exponential polynomial kernel parameters of 721.114 and 3, respectively.

Considering the impact of different numbers of DHKELM network layers on diagnostic results, M-layer DHKELM models were constructed, $M = 2, 3, 4, 5, 6$. The fault diagnosis accuracy rates for the training set under different numbers of network layers are shown in Figure 25.

As shown in Figure 25, the accuracy reaches its maximum when the network has 3 layers. As the network depth increases, the accuracy fluctuates to some extent; when the network has 6 layers, the diagnostic accuracy drops to its minimum. The results indicate that increasing or decreasing the number of network layers does not necessarily improve diagnostic accuracy; the most appropriate number of layers must be selected based on actual conditions. In this case, a 3-layer network is optimal, which is consistent with the optimization results.

The proposed SMVMD-ISSA-DHKELM intelligent fault diagnosis method for diesel engines under unbalanced few-shot conditions, as well as methods such as ISSA-DHKELM, SMVMD-DHKELM, SMVMD-KELM, and 1DCNN, were used to identify the various fault states in the test set. To minimize the impact of randomness and ensure the reliability and reproducibility of the experimental results, after partitioning the dataset, 10 independent and mutually exclusive experiments were conducted for all methods, yielding different trained fault diagnosis models. The test set was input into the various trained models, and the average of the experimental results was taken as the final performance metric for each method. The standard deviation of the results was also calculated to reflect stability. The

experimental results are shown in Table 13, and the visualization of the diagnostic results and the confusion matrix are shown in Figure 26. Fault states 1–4 correspond to diesel engine valve clearances of 0.3 mm, 0.4 mm, 0.5 mm, and 0.7 mm, respectively.

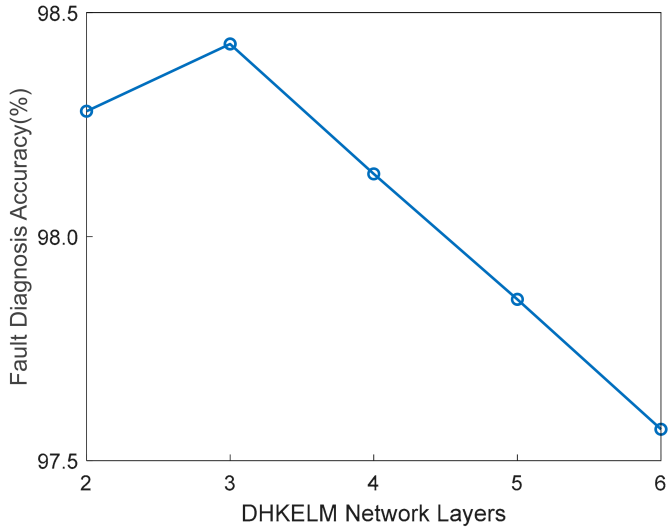

Figure 25. Fault diagnosis accuracy results with the number of network layers.

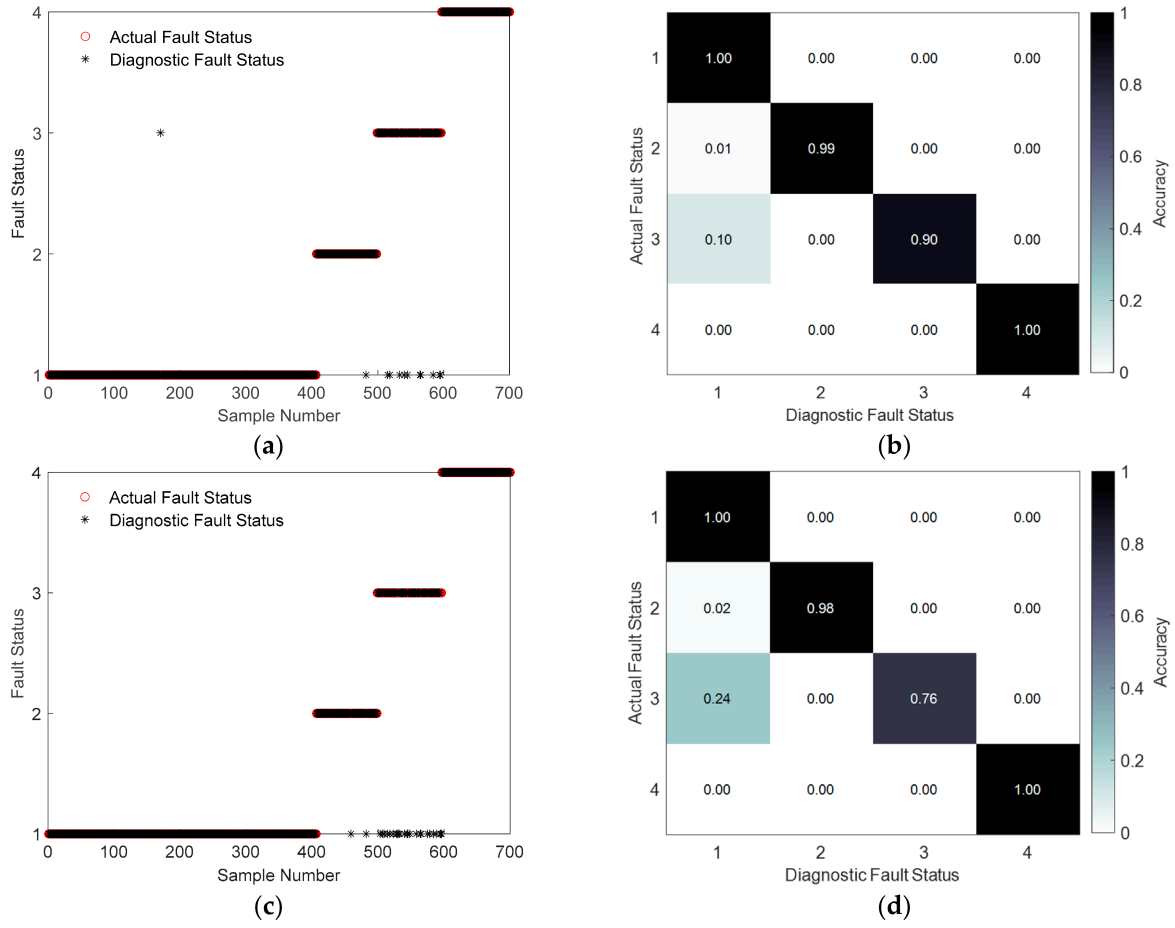

Figure 26. *Cont.*

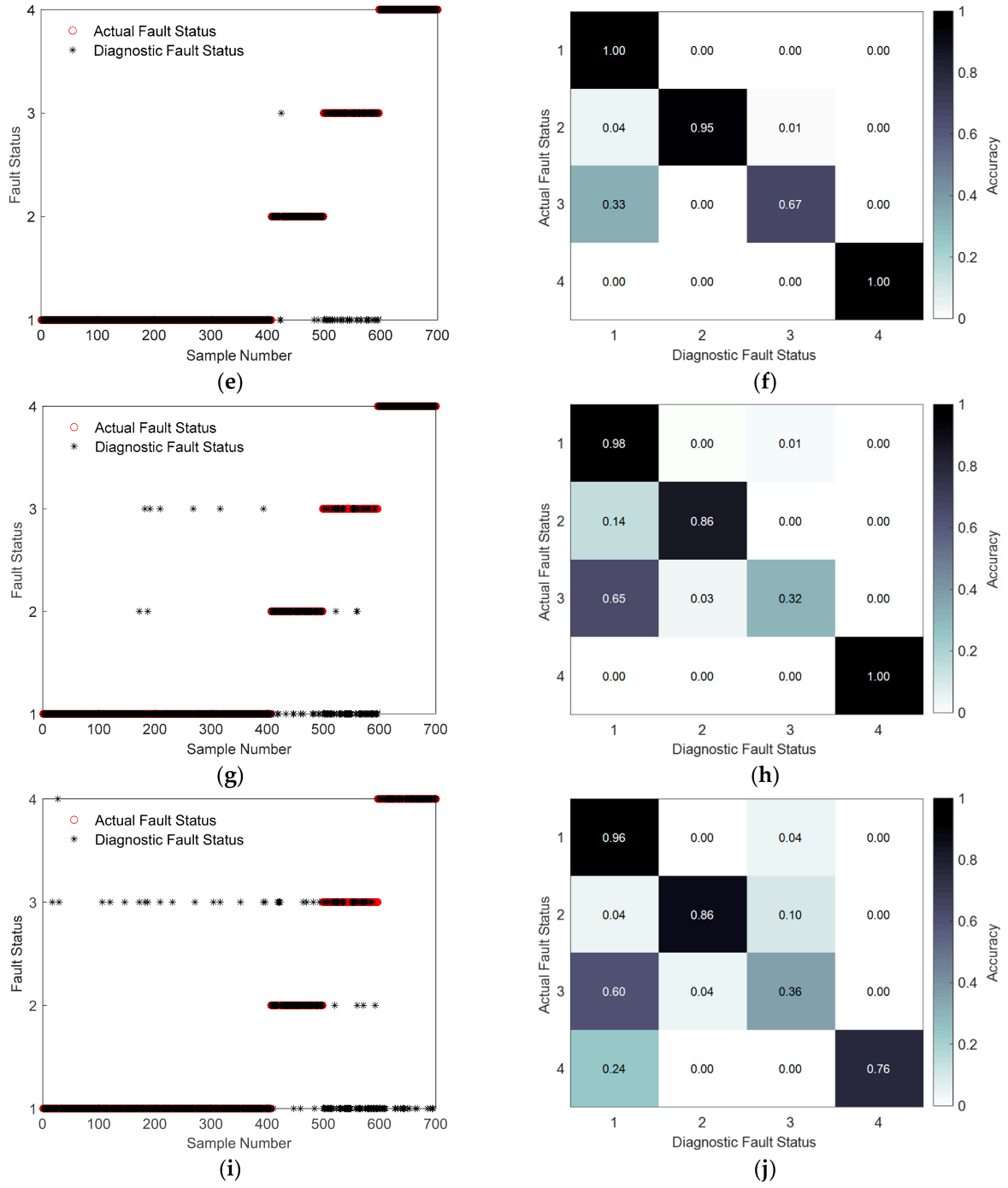

Figure 26. Fault state identification result and confusion matrix for different methods: (**a**) SMVMD-ISSA-DHKELM identification result. (**b**) SMVMD-ISSA-DHKELM confusion matrix. (**c**) ISSA-DHKELM identification result. (**d**) ISSA-DHKELM confusion matrix. (**e**) SMVMD-DHKELM identification result. (**f**) SMVMD-DHKELM confusion matrix. (**g**) SMVMD-KELM identification result. (**h**) SMVMD-KELM confusion matrix. (**i**) 1DCNN identification result. (**j**) 1DCNN confusion matrix.

As shown in Figure 26 and Table 13, the proposed unbalanced few-shot intelligent fault diagnosis method for diesel engines achieves an overall fault state recognition accuracy of 98.29% using intake and exhaust system data, which is the highest accuracy among the five methods. Specifically, the diagnostic accuracy for the "healthy" and "deteriorating" fault states reached 100.00%, while that for the "attention" fault state was 99.00%, and the diagnostic accuracy for the "abnormal" fault state was 90.00%. Compared with the ISSA-DHKELM method, the proposed method improves diagnostic accuracy by 2.00%, further

demonstrating the necessity and effectiveness of using the SMVMD method to process vibration signals and extract the most valuable component from the signal. Compared with the SMVMD-DHKELM method, the proposed method improved diagnostic accuracy by 3.58%, further demonstrating that optimizing the hyperparameters in DHKELM via the proposed ISSA can fully leverage the model's fault diagnosis performance. Compared with the SMVMD-KELM method, the SMVMD-DHKELM method improved diagnostic accuracy by 7.28%, further demonstrating that DHKELM—which builds upon KELM by incorporating deep learning concepts to form a multi-layer network feature extraction model and using a linear weighting method to combine different kernel functions to compensate for the shortcomings of a single kernel function—can more effectively extract features from data samples and achieve better fault diagnosis results. Compared with the 1DCNN method, the proposed method improves diagnostic accuracy by 15.15%. This further illustrates that, under unbalanced few-shot conditions, traditional methods struggle to meet the practical demands of diesel engine fault diagnosis, necessitating the development of new methods with high accuracy and strong generalization capabilities. The method proposed in this manuscript not only achieves high accuracy but also demonstrates superior performance in this application scenario.

Table 13. Accuracy of each method under test set (intake and exhaust system).

Model	Accuracy/%
SMVMD-ISSA-DHKELM	98.29 ± 0.61
ISSA-DHKELM	96.29 ± 1.85
SMVMD-DHKELM	94.71 ± 3.07
SMVMD-KELM	87.43 ± 4.30
1DCNN	83.14 ± 4.92

To further validate the performance of the proposed SMVMD-ISSA-DHKELM method under few-shot conditions, we maintained a 4:1 ratio between the majority and minority classes and kept the training set at one-third of the total sample size. We then randomly selected 10, 30, and 70 samples from the remaining three fault states, respectively, to serve as the training set for the model. Figure 27 visualizes the accuracy of each method on the test set for different numbers of fault samples, and Table 14 shows the precision results.

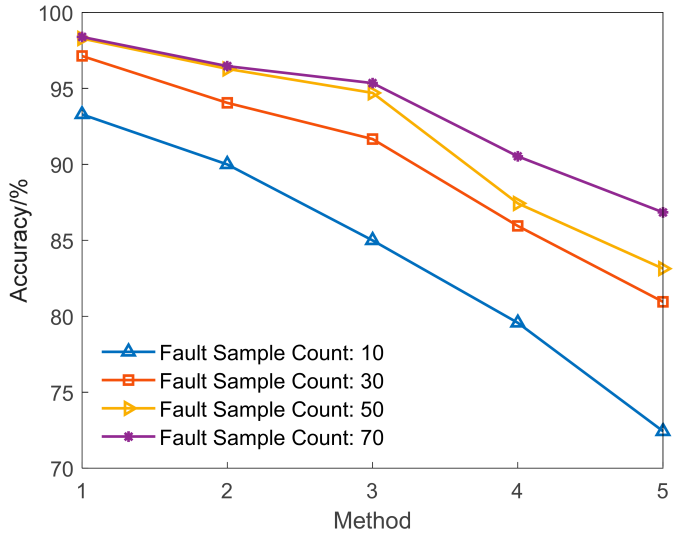

Figure 27. Accuracy of each diagnosis method for different numbers of fault samples.

Table 14. Accuracy of each method for different sample sizes (intake and exhaust system).

The Number of Samples of Each Failure Type in the Training Set	Methods	Accuracy/%
10	SMVMD-ISSA-DHKELM	93.29
	ISSA-DHKELM	90.00
	SMVMD-DHKELM	85.00
	SMVMD-KELM	79.57
	1DCNN	72.43
30	SMVMD-ISSA-DHKELM	97.14
	ISSA-DHKELM	94.05
	SMVMD-DHKELM	91.67
	SMVMD-KELM	85.95
	1DCNN	80.95
70	SMVMD-ISSA-DHKELM	98.39
	ISSA-DHKELM	96.47
	SMVMD-DHKELM	95.35
	SMVMD-KELM	90.53
	1DCNN	86.84

As shown in Figure 27 and Table 14, as the number of samples increased from 10 to 70, the models were trained more thoroughly, and the fault diagnosis accuracy of the various methods improved, indicating that the accuracy of fault diagnosis methods is directly related to the number of training samples. Regardless of the sample size, the proposed method can still learn the distribution of real samples relatively well and demonstrates the best fault diagnosis performance. Even when there are only 10 samples of each fault type in the training set, the proposed method can still achieve an accuracy rate of 93.29%. This comparative experiment further demonstrates that the proposed method can meet the requirements for fault diagnosis under conditions of few-shot sizes.

To further validate the performance of the proposed SMVMD-ISSA-DHKELM method under unbalanced sample conditions, we kept the training set at one-third of the total sample size and maintained the number of samples for the remaining three fault states at 10 each as the model's training set. We then conducted experiments by setting the majority-to-minority class ratio to 2:1, 10:1, and 30:1, respectively. The visualization results of the accuracy rates for each method on the test set are shown in Figure 28, and the accuracy values are presented in Table 15.

As shown in Figure 28 and Table 15, as the proportions of majority-class and minority-class samples increased from a 2:1 ratio to a 30:1 ratio, the fault diagnosis accuracy of the proposed method decreased by only 3.73%, whereas the fault diagnosis accuracies of the other four methods decreased by 8.64%, 6.06%, 9.76%, and 10.82%, respectively. This indicates that the proposed method exhibits the best stability. Under any class unbalance ratio, the proposed method demonstrates the best fault diagnosis performance, indicating that it retains strong adaptability even when sample ratios are unbalanced. Even when the training set contained only 10 samples per fault class and the unbalance ratio reached as high as 30:1, the proposed method still achieved an accuracy of 91.27% on intake and exhaust system data, far exceeding that of the comparison methods. This comparative test further demonstrates that the proposed method can meet fault diagnosis requirements under unbalanced sample conditions.

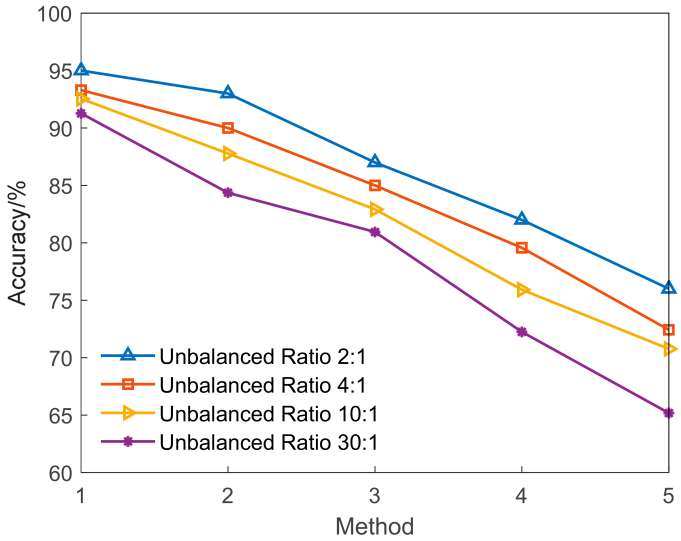

Figure 28. Accuracy of different diagnosis methods under each unbalanced data set.

Table 15. Accuracy of each method for unbalanced dataset (intake and exhaust system).

Unbalanced Dataset	Methods	Accuracy/%
2:1	SMVMD-ISSA-DHKELM	95.00
	ISSA-DHKELM	93.00
	SMVMD-DHKELM	87.00
	SMVMD-KELM	82.00
	1DCNN	76.00
10:1	SMVMD-ISSA-DHKELM	92.54
	ISSA-DHKELM	87.77
	SMVMD-DHKELM	82.92
	SMVMD-KELM	75.92
	1DCNN	70.77
30:1	SMVMD-ISSA-DHKELM	91.27
	ISSA-DHKELM	84.36
	SMVMD-DHKELM	80.94
	SMVMD-KELM	72.24
	1DCNN	65.18

In summary, the proposed method for diesel engine fault diagnosis under unbalanced few-shot conditions has been validated. Compared with other methods, it can accurately and intelligently distinguish between different types of diesel engine faults under unbalanced few-shot conditions.

6. Summary and Outlook

Given the severe unbalanced ratio between normal and fault data samples for diesel engines and the scarcity of historical fault data in practical applications, this manuscript proposes an intelligent fault diagnosis method for diesel engines under unbalanced few-shot conditions based on SMVMD and optimized DHKELM. The following conclusions can be drawn:

(1) This study marks the first application of the SMVMD method to the field of diesel engine fault diagnosis, which can improve sample quality at the data level. By applying this method to processed vibration signal samples, the most valuable IMF components are identified, enabling the model to extract valuable fault feature information embed-

ded in the signals. Compared with directly extracting fault features from raw signals, processing the data using SMVMD effectively improves fault diagnosis performance.

(2) An ISSA-optimized DHKELM diesel engine fault diagnosis model is proposed, which optimizes the feature extraction process at the model level. Building upon KELM, DHKELM incorporates deep learning concepts and combines different kernel functions using a linear weighting method to achieve better fault diagnosis results. ISSA is proposed to optimize the hyperparameters in DHKELM, effectively improving fault diagnosis accuracy.

(3) The proposed fault diagnosis method addresses the issues of sample class imbalance and scarcity of available fault data, as well as the tendency of models to favor the class with more samples and to overfit, which affects fault diagnosis accuracy. This method further enriches and advances research on fault diagnosis for diesel engines under unbalanced few-shot conditions. It holds significant theoretical significance and practical value, and can further improve maintenance efficiency and reduce maintenance and support costs.

Although this manuscript has yielded preliminary research results, due to limitations in experimental conditions, research time, and individual capabilities, further exploration and research are needed on fault diagnosis methods suitable for variable-speed and variable-load operating conditions, as well as diesel engine fault diagnosis methods based on multi-source heterogeneous data. Building on this foundation, further exploration and research are needed on deploying the proposed methods to embedded systems or general-purpose platforms.

Author Contributions: Conceptualization, Z.L. (Zichang Liu) and Y.B.; methodology, Z.L. (Zichang Liu); software, Z.L. (Zichang Liu) and Y.Z.; validation, Z.L. (Zichang Liu), D.W. and Y.Z.; formal analysis, Z.L. (Zhining Li); resources, Z.L. (Zhining Li) and R.W.; data curation, Z.L. (Zhining Li); writing—original draft preparation, Z.L. (Zichang Liu); writing—review and editing, Z.L. (Zichang Liu), R.W. and D.W.; visualization, Y.B.; supervision, Y.B.; project administration, Z.L. (Zichang Liu); funding acquisition, Z.L. (Zichang Liu). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Equipment Pre-research Common Technology Project, grant number 50904020201.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Huang, M.; Bi, X.Y.; Yang, X.; Li, X.; Tang, D.J.; Bi, F.R. Diesel engine fault data augmentation method based on artificial data fusion. *J. Vib. Shock* **2023**, *42*, 278–286. [CrossRef]
2. Li, B.Y.; Yu, Y.H.; Wang, W.C.; Cao, B.X.; Xu, D.F.; Yao, Y.F. Sample generation method for marine diesel engines based on FEM simulation and DCGAN. *J. Mech. Sci. Technol.* **2024**, *38*, 2335–2345. [CrossRef]
3. Li, L.Y.; Su, T.X.; Ma, F.K. Small-sample fault diagnosis method for high pressure common rail system. *Trans. CSICE* **2023**, *41*, 255–262. [CrossRef]
4. Cui, Z.Q.; Lu, Y.L.; Yan, X.; Cui, S.Y. Compound fault diagnosis of diesel engines by combining generative adversarial networks and transfer learning. *Expert Syst. Appl.* **2024**, *251*, 123969. [CrossRef]
5. Chen, T.Y.; Xiang, Y.; Wang, J.X. DA-VICReg: A data augmentation-free self-supervised learning approach for diesel engine fault diagnosis. *Meas. Sci. Technol.* **2024**, *35*, 086109. [CrossRef]
6. Zhang, J.H.; Pei, G.B.; Zhu, X.L.; Guo, X.; Deng, L.L.; Gao, L.; Liu, Z.W.; Ni, Q.; Lin, J.W. Diesel engine fault diagnosis for multiple industrial scenarios based on transfer learning. *Measurement* **2024**, *228*, 114338. [CrossRef]
7. Xu, J. Research on Fault Diagnosis of High-Pressure Fuel Pump of Marine Common Rail Diesel Engine Under Complicated Working Conditions. Master's Thesis, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2021. [CrossRef]

8. Xiong, G.Q.; Ma, W.S.; Zhao, N.Y.; Zhang, J.J.; Jiang, Z.N.; Mao, Z.W. Multi-type diesel engines operating condition recognition method based on stacked auto-encoder and feature transfer learning. *IEEE Access* **2021**, *9*, 31043–31052. [CrossRef]
9. Jia, Z.Y. Research on Diesel Engine Fault Diagnosis Method Based on Machine Learning. Master's Thesis, Jiangsu University of Science and Technology, Zhenjiang, China, 2022. [CrossRef]
10. Bai, H.J.; Zhan, X.B.; Yan, H.; Wen, L.; Jia, X.S. Combination of optimized variational mode decomposition and deep transfer learning: A better fault diagnosis approach for diesel engines. *Electronics* **2022**, *11*, 1969. [CrossRef]
11. Zhao, Z.J. Research on Intelligent Diagnosis Method of Diesel Engine Valve Clearance Abnormal and Misfire Fault. Master's Thesis, Beijing University of Chemical Technology, Beijing, China, 2021. [CrossRef]
12. Li, Y.F.; Han, T. Deep learning based industrial equipment prognostics and health management: A review. *J. Vib. Meas. Diagn.* **2022**, *42*, 835–847. [CrossRef]
13. Wang, L.J.; Huang, X.Y. Air valve fault diagnosis of diesel engine based on MTF and CBAM-IGPN with small samples. *Mech. Sci. Technol. Aerosp. Eng.* **2025**, *44*, 1143–1150. [CrossRef]
14. Liao, S.L.; Bi, F.R.; Tian, C.F.; Yang, X.; Li, X.; Tang, D.J. Typical fault diagnosis method of Diesel engine based on QPSO-MC-GCNN. *J. Vib. Shock* **2022**, *41*, 268–275. [CrossRef]
15. Liu, Y.S.; Kang, J.S.; Guo, C.M.; Bai, Y.J. Diesel engine small-sample transfer learning fault diagnosis algorithm based on STFT time-frequency image and hyperparameter autonomous optimization deep convolutional network improved by PSO-GWO-BPNN surrogate model. *Open Phys.* **2022**, *20*, 993–1018. [CrossRef]
16. He, P.F.; Huang, G.Y.; Ruan, A.G. Study on valve fault diagnosis of compression ignition piston engine based on continuous wavelet transform and model agnostic meta learning. *Chin. Intern. Combust. Engine Eng.* **2024**, *45*, 57–63. [CrossRef]
17. Wang, R.H.; Chen, H.; Guan, C. DPGCN model: A novel fault diagnosis method for marine diesel engines based on imbalanced datasets. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3504011. [CrossRef]
18. Liu, Z.C.; Bai, Y.S.; Wang, R.C.; Li, S.Y.; Jia, X.S. Diesel engine vibration signal decomposition method based on successive multivariate variational mode decomposition. In *Proceedings of the 14th International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE2024)*; Institution of Engineering and Technology: Harbin, China, 2024; Volume 2024, pp. 916–923. [CrossRef]
19. Liu, Z.C.; Bai, Y.S.; Han, Y.M.; Jia, X.S. Application of optimized DHKELM to fault diagnosis of diesel engine. *J. Army Eng. Univ. PLA* **2024**, *3*, 77–85.
20. Xue, J.K. Research on the Application of Novel Swarm Intelligent Optimization Algorithms. Ph.D. Thesis, Donghua University, Shanghai, China, 2024. [CrossRef]
21. Liu, J.; Tan, Y.T.; Gu, Y.L.; Yang, N. LSGAN-Swin Transformer diagnosis method of bearing fault under unbalanced samples. *J. Vib. Eng.* **2025**, *38*, 1775–1787. [CrossRef]
22. Chen, X.D.; Fu, Z.S.; Wu, Z.H.; Chen, Q.Y.; Guo, W.K. Small-Sample Fault Diagnosis Method Based on Multi-Head Convolution and Differential Self-Attention. *J. South China Univ. Technol. (Nat. Sci. Ed.)* **2023**, *51*, 21–33. [CrossRef]
23. Si, W.W.; Cen, J.; Wu, Y.B.; Hu, X.L.; He, M.Z.; Yang, Z.H.; Chen, H.H. Review of Research on Bearing Fault Diagnosis with Small Samples. *Comput. Eng. Appl.* **2023**, *59*, 45–56. [CrossRef]
24. Li, S.Y.; Liu, Z.C.; Yan, Y.B.; Wang, R.C.; Dong, E.Z.; Cheng, Z.H. Research on diesel engine fault status identification method based on synchro squeezing s-transform and vision transformer. *Sensors* **2023**, *23*, 6447. [CrossRef] [PubMed]
25. Liu, Z.C.; Bai, Y.S.; Li, S.Y.; Zhang, K.; Liu, M.; Jia, X.S. Assessment of diesel engine valve performance degradation status based on synchroextracting enhanced generalized S-transform. *Acta Armamentarii* **2024**, *45*, 2003–2016. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.