IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 22 July 2022, accepted 30 August 2022, date of publication 8 September 2022, date of current version 19 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3205105

RESEARCH ARTICLE

Fault Diagnosis of Aeroengine Control System Sensor Based on Optimized and Fused Multidomain Feature

HUIHUI LI, LINFENG GOU, YINGXUE CHEN, AND HUACONG LI
School of Power and Energy, Northwestern Polytechnical University, Xi'an 710072, China

Corresponding author: Linfeng Gou (goulinfeng07nwpu@163.com)

This work was supported in part by the National Science and Technology Major Project under Grant 2017-V-0011-0062.

ABSTRACT Improving the efficacy and dependability of aeroengines requires timely and effective sensor fault diagnosis. Deep learning-based fault diagnosis method is a current research hotspot. To overcome some of the method's existing shortcomings and improve the reliability of fault diagnosis, this paper proposes a novel intelligent fault diagnosis framework with higher quality features and more effective fault classifiers. The proposed plan includes three stages. Firstly, multidomain features (time and frequency domain features) are extracted to describe the sensor's health from several dimensions. Secondly, the advanced Henry gas solubility optimization algorithm (HGSO) is applied to improve classification accuracy through feature selection, and the operating conditions and the features extracted by the network are fused as fault indicators. Finally, an adaptive deep belief network (ADBN) with relu-softsign combination activation layers, variable learning rate, and optimized network structure is proposed as the fault identifier. The advantages of the first two stages lie in the complete utilization of information and reducing the data dimension. In addition, the detection performance and the convergence speed is enhanced by the proposed ADBN. The experimental data are derived from a combination of measured and simulated data generated from the aeroengine model. The experimental results indicate that the improved method can produce better performance and outcomes than the unimproved methods for all fault scenarios, with a higher diagnostic accuracy of 98.1% and a reduced time of 98 s. The efforts of this study provide a efficient and adaptable way to aeroengine sensor fault diagnosis.

INDEX TERMS Aeroengine control system, Henry's gas solubility optimization algorithm, adaptive deep belief network, sensor fault diagnosis, multidomain features.

I. INTRODUCTION

Aeroengine sensors must be reliable for aircraft performance and flight safety [1], [2]. Aeroengines control and health management are strongly dependent on accurate sensor measurements. However, the aeroengine sensors are prone to faults because of the extreme working conditions, making the engine run abnormally and even catastrophic events [3], [4]. To prevent accidental faults and reduce maintenance costs, an effective strategy is needed to ensure the reliability and safety of the aeroengine.

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

The great reliability contributes to the plant's stable operation and safety. Numerous efforts have been devoted to designing and evaluating complex systems with high reliability. The literature [5] developed a vectorial surrogate modeling method to accomplish the comprehensive reliability design of multi-objective structures. The literature [6] proposed an improved extremum response surface method for mechanism reliability evaluation.

Another route is to increase the reliability of systems through fault diagnosis. In reported literature, the fault diagnosis methods of engine sensors include model-based and data-driven methods [7]. The model-based technique compares available measurements to a priori information

represented by a mathematical model and treats the difference between the two as a fault indicator [8], [9]. Therefore, this method can better express the dynamic relationship within the system. The model-based approach has been widely used in many applications, such as Kalman filter [10], unknown input observer [11], sliding mode observer [12], parity space [13], and so on. Due to its high dependence on an accurate aeroengine model, this method has limitations in practical application despite its excellent performance. Specifically, an aeroengine is a complicated device with multi-field coupling, resulting in complex and difficult modeling. With growing modeling uncertainty and nonlinear complexity, diagnostic performance may decline.

In contrast, data-driven approaches do not need prior knowledge, such as esoteric engine working principles and complex modeling techniques [14]. It obtains fault information directly from the massive aeroengine running data. Data-driven methods have received much attention from researchers because of their advantages in describing nonlinear functions, minimal human involvement, and handling large amounts of data [15], [16].

The quality of the extracted features and fault classifiers is the key to data-driven fault diagnosis, directly influencing diagnosis results. Along with the increasing accuracy of fault diagnosis, the feature extraction technique is growing in importance, and many related works have been conducted in this area. Fei *et al.* [17] developed a new feature extraction method, that is, hierarchical quantum entropy, to effectively conduct the fault diagnosis of inter-shaft bearings with precision and stability. Ai *et al.* [18] suggested a fusion method based on n-dimensional characteristic parameters distance to monitor rolling bearing operating status with casings in real-time efficiently and accurately. These works demonstrate the potential of high-quality features to enhance fault diagnosis accuracy.

Extraction of multi-domain features greatly adds to the improvement of feature quality. Chen *et al.* [19] extracted time and frequency domain features from the different sensor signals and pointed out that multi-domain features can be regarded more effectively as machine health indicators. Anam *et al.* [20] proposed a combined multidomain feature set that exploits the diversified information contained within the signals and can capture anomalies under various conditions. The multi-domain feature set can capture the fault information of the equipment more comprehensively and improve the classification accuracy.

Existing technology for feature extraction in sensor fault diagnosis is short. (1) The multi-domain features extracted need to be more concise. A large-sized feature subset composed of similar redundant features might lead to overfitting and misclassification. The compromise between fewer dimensions and better information is a question that requires careful consideration. (2) The influence of the operating conditions (different H and Ma) on the measurement results of the sensors has often been ignored in previous studies. This may hurt the accuracy of fault diagnosis.

Moreover, the performance of the fault classifier can have a direct impact on fault diagnosis. Examples include autoassociative neural networks [21], support vector machines [22], and various hybrid intelligence techniques[23], among others. The fault diagnosis of engine sensors is extremely challenging because of the increasingly complex structure of aeroengines, variable operating conditions, and noise coupling. Fortunately, deep learning (DL) provides a powerful solution for fault diagnosis. Deep learning has unique advantages in dealing with complex problems [24]. It simulates the information processing mechanism of the human brain by constructing deep neural networks capable of learning, interpreting, and analyzing input data and deciphering data knowledge [7], [25]. Several research initiatives based on DL in real applications such as aeroengines[26], wind turbines [27], bearings[19], planetary gearboxes[28], etc., have been reported in the literature to improve the reliability and safety of the considered systems.

A common DL structure is the Deep Belief Network (DBN), which has a high ability to handle complex recognition tasks, is a typical DL structure. In recent years, it has been increasingly utilized for sensor fault diagnosis. Tamilselvan and Wang [29] presented a novel sensor health diagnosis method using a deep belief network and successfully applied it to aircraft engine health diagnosis and electric power transformer health diagnosis. Liu *et al.* [30] proposed a DBN-based fault detection method for aircraft engine sensors and demonstrated that it is more precise than BP neural networks and SVM. Feng *et al.* [31] offered a new method-based DBN for engine fault diagnosis. It is worth mentioning that he extracted the entropy of the original signal and used it as an input to the DBN, and achieved a higher accuracy.

However, the performance constraints of classical DBN hinder diagnosis accuracy, and its diagnostic performance has the potential to improve. For example, constant learning rate and saturated activation layers (such as sigmoid and tanh) lead to low diagnostic accuracy and slow convergence. Again, for example, the structure of deep learning networks often comes from very time-consuming trial and error.

Motivated by the above challenges, this work aims to provide generic solutions to address the challenges. A novel intelligent fault diagnosis method is presented in this paper. The diverse fault information contained in the signal is captured from multiple dimensions (time domain and frequency domain). Then, multi-domain features should be filtered to remove redundant information and train the fault diagnosis model with the least dimensional but knowledgeable and high-quality information. Traditional feature selection schemes have the potential for local solutions. Meta-heuristic algorithms are ideally suited to solve optimization problems [32], including feature selection problems. The Henry Gas Solubility Optimization (HGSO) algorithm has been widely used in feature selection problems, because to its increasing convergence speed, reducing computational costs, and quickly eliminating local optima. The HGSO algorithm is employed to successfully filter sensitive features in this paper.

Finally, the adaptive DBN(ADBN) model is proposed and trained for fault classification. Notably, the operating conditions and the features obtained from ADBN are fused to serve as fault indicators jointly.

The main contributions of this work are summarized as follows.

(1) The HGSO algorithm, a new meta-heuristic algorithm, is introduced to self-adapt optimized features.

(2) A fusion feature is proposed. On the one hand, multi-domain features are extracted as the main features of the signal. On the other hand, the influence of operating conditions is considered and creatively used as a secondary feature. The two parts are fused to obtain complete fault information.

(3) An adaptive ADBN model for fault classification is investigated. The model has a dynamically trained activation layer relu-softsign and an adaptive learning rate for high accuracy and fast convergence. And introducing an advanced meta-heuristic algorithm, the Harris Hawk Optimisation (HHO) algorithm achieves automatic acquisition of the optimal network structure.

(4) A hybrid flexible diagnosis framework for aeroengine sensors integrating multi-domain features, feature optimization and fusion, dynamic training, and structure optimization is proposed.

The rest of this article is organized as follows. In Section 2, the theoretical background about HGSO, DBN, and HHO was introduced. In Section 3, the proposed fault diagnosis method is described. In Section 4, we verify the method's practicability through experimental comparison. Section 4 presents the main conclusions.

II. THEORETICAL BACKGROUND

A. BRIEF INTRODUCTION OF THE HGSO ALGORITHM

The Henry gas solubility optimization algorithm is a new metaheuristic algorithm proposed in 2019 [33]. The meta-heuristic algorithm is constructed based on biological or physical phenomena in nature. The stochastic factors in the evolution process enable it to escape from the optimal local solution, which has the benefits of simple operation, generalization, and good generalization [34], [35]. The famous physics law, Henry's law, inspires the HGSO algorithm, which explains the phenomenon of solubility of a gas in a liquid under a certain pressure. The concept is depicted in Figure 1 [33].

The above Henry's law is mathematically formulated as an optimization process containing 8 steps:

Step 1: Initialization. The position X_i of the i^{th} gas particle in the population *is* initialized as

$$X_i^0\,(t+1) = lb + rand\,(n, dim) \times (ub - lb) \tag{1}$$

In which *ub*, *lb* and *dim* depict the upper bound, the lower bound, and the dimension of the problem, respectively. The initial values of Henry's constant H_j for cluster j, the partial pressure $P_{i,j}$ of gas i in cluster j, and the constant value C_j of

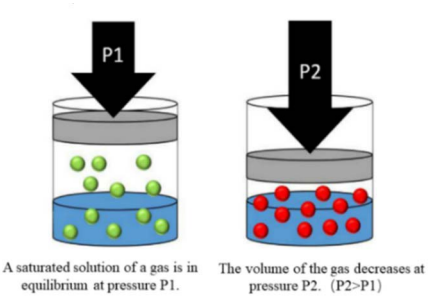

FIGURE 1. Gas particles dissolving into a liquid under partial pressure.

cluster j are formulated as,

$$H_j^0(t) = l_1 \times rand(m, 1), P_{i,j}^0 = l_2 \times rand(n, 1),$$
$$C_j^0 = l_3 \times rand(m, 1) \tag{2}$$

In which m represents the number of gas clusters. l_1, l_2, l_3 are constants equal to 5e-03, 100, 1e-02, respectively.

Step 2: Clustering. The gas particles with n population are distributed into m clusters as the gas type. Each cluster contains a similar group of candidate particles with the same Henry's coefficient H_j and constant value C_j. Each cluster gas has constant values H_j and C_j.

Step 3: Evaluation. In each cluster j, the best candidate particle $x_{j,best}$ which obtains the best fitness value in cluster j is evaluated to find the global best gas-particle $x_{g,best}$ among the population n.

Step 4: Update Henry's coefficient. In different iteration and different cluster, the Henry's coefficient is updated refer to the Henry's law as

$$H_j(t+1) = H_j\,(t) \times exp\left(-C_j \times \left(\frac{1}{T^t} - \frac{1}{T^\theta}\right)\right) \tag{3}$$

In which $T^t = \exp(\frac{-t}{iter})$ which changes in every iteration, T depicts the temperature, and T^θ is a constant set to 298.15.

Step 5: Update solubility. The solubility $S_{i,j}$ of gas particle i in cluster j is mathematically expressed as

$$S_{i,j}(t) = K \times H_j\,(t) \times P_{i,j}(t) \tag{4}$$

in which K is a constant value equal to 1.

Step 6: Update position. The next position of the i^{th} gas particle in the j^{th} cluster is updated as

$$X_{i,j}\,(t+1) = X_{i,j}\,(t) + f \times rand \times \phi_{i,j} \times \left(x_{j,best} - x_{i,j}\,(t)\right) + flag \times rand \times \alpha \times (S_{i,j}\,(t) \times x_{best}\,(t) - X_{i,j}\,(t)),$$
$$\phi_{i,j} = \beta \times exp(-\frac{F_{best}\,(t) + \varepsilon}{F_{i,j}\,(t) + \varepsilon}) \tag{5}$$

In which f is a flag index equal to -1 or 1, which is utilized to change the direction of the search agents. *rand* donates random values in (0,1), and every *rand* represents a different random value. $\phi_{i,j}$ is the ability of the gas particle i in the cluster j. α donates the influence of other gas particles on

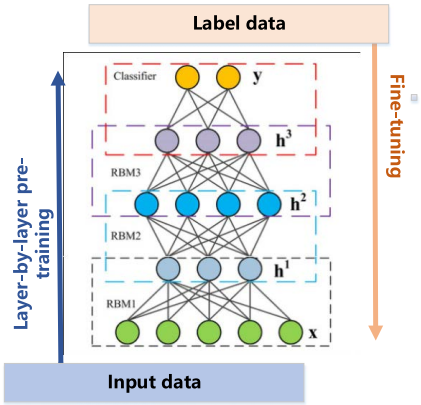

FIGURE 2. The diagram of the DBN structure and training process.

the i^{th} gas candidate which is set to 1. β and ε are constant coefficients equal to 1 and 0.05, respectively.

Step 7: Obtain the worst agent. The worst agent N_w is ranked and employed in the optimization process to avoid local optima, which is formulated as

$$N_w = n \times (rand \times (c_2 - c_1) + c_1) \tag{6}$$

In which c_1 and c_2 is constantly equal to 0.1 and 0.2, respectively. All *rand* functions in the model donate random vectors (0,1).

Step 8: Update the worst position. The position of the worst particle is updated by a random value using

$$X_w = lb + rand \times (ub - lb) \tag{7}$$

After the above mention processes, the position X_{i+1} of the $(i+1)^{th}$ gas particle is initialized

B. BRIEF INTRODUCTION OF THE DBN MODEL

DBN is one of the deep learning models that excels at feature extraction and classification ability. It is a probabilistic generative model that consists of multiple Restricted Boltzmann Machine (RBM) stacks. The.DBN model also contains a classifier (after the last RBM) to accomplish the classification task. The structure diagram of DBN is depicted in Figure 2.

The RBM is an energy-based two-layer model consisting mainly of visible and hidden layer. They are connected by weighting factors. The following functions give joint probability density and joint probability distribution.

$$E(v, h) = -\sum_{j=1}^{m} a_j v_j - \sum_{i=1}^{n} b_i h_j - \sum_{i=1}^{n}\sum_{j=1}^{m} v_j w_{ij} h_i \tag{8}$$

$$p(v, h; \theta) = \frac{1}{Z(\theta)} \exp(-E(v, h; \theta)) \tag{9}$$

where θ is the model parameters, 0 and 1 represent neurons' inactivation and activation states, respectively. v and h are the activation states of visible and hidden unit, a and b are the respective deviations, and w_{ij} is the weights between the visible and hidden unit. $Z = \sum_v \sum_h e^{-E(v,h)}$ is a normalization constant that simulates a physical system.

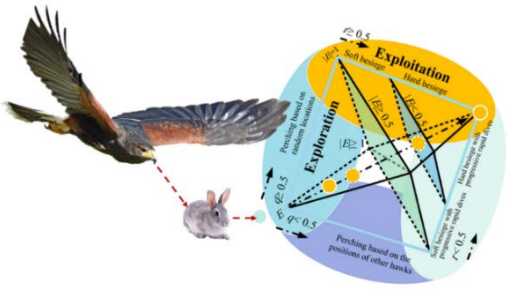

FIGURE 3. Flow chart of HHO algorithm.

The conditional probability distributions $p(h \mid v), p(v \mid h)$ are obtained by Bayesian inference.

$$p(h \mid v) = \frac{1}{\left[1 + \exp\left[-c_i - \sum_{j=1}^{m} v_j w_{i,j}\right]\right]} \tag{10}$$

$$p(v \mid h) = 1/\left[1 + \exp\left[-b_j - \sum_{i=1}^{n} h_i w_{i,j}\right]\right] \tag{11}$$

The essence of RBM is to maximize the probability that the learned RBM model matches the input sample distribution. The parameter updating process is performed using the contrast divergence algorithm [35].

$$\begin{aligned} \omega_{ij}^{n} &= \omega_{ij}^{n-1} + \gamma\left(\langle v_i h_j\rangle_{\text{data}} - \langle v_i h_j\rangle_k\right) \\ a_{ij}^{n} &= a_{ij}^{n-1} + \gamma\left(\langle v_i h_j\rangle_{\text{data}} - \langle v_i h_j\rangle_k\right) \\ b_{ij}^{n} &= b_{ij}^{n-1} + \gamma\left(\langle v_i h_j\rangle_{\text{data}} - \langle v_i h_j\rangle_k\right) \end{aligned} \tag{12}$$

where $\gamma \in [0, 1]$ is the learning rate that can be used to adjust the learning speed, n denotes the number of iterations of training, and k is the step size of the contrast divergence.

C. BRIEF INTRODUCTION OF THE HHO ALGORITHM

In 2019, the HHO algorithm was proposed as a new swarm intelligence optimization algorithm [37]. It was inspired by the collaborative foraging behavior of Harris hawks [38], [39]. The HHO algorithm comprises two phases: exploration and exploitation. The whole method flow is illustrated in Figure 3.

Phase 1: exploration. During this phase, Harris's hawk randomly perches on several locations to track and detect prey. The following update functions are available.

$$X(t+1) = \begin{cases} X_{rand}(t) - r_1 \left|X_{rand}(t) - 2r_2 X(t)\right|, & q \geq 0.5 \\ \left(X_{rabbit}(t) - X_m(t)\right) - r_3\left(lb + r_4(ub - lb)\right), & q < 0.5 \end{cases} \tag{13}$$

where, $X_{rand}(t)$ is the randomly selected individual in the current population, $X_{rabbit}(t)$ denotes the current optimal individual,, r_1, r_2, r_3, r_4 are random numbers from 0 to 1, ub and lb represent the upper and lower bounds of the population respectively, N is the population number.

$X_m(t)$ is the average position of the current population and is obtained as:

$$X_m(t) = \frac{1}{N}\sum_{i=1}^{N} X_i(t) \tag{14}$$

where, $X_i(t)$ denotes the position of the *i-th* Harris hawk in this iteration.

Phase 2: Transition. The shift of Harris Hawk from global to local search is mainly controlled by the escape energy factor E, which is calculated as follows.

$$E = 2E_0(1 - \frac{t}{t_{MAX}}) \tag{15}$$

where, E_0 is a random number from -1 to 1, t denotes the current number of iterations, and t_{MAX} is the maximum number of iterations.

Phase 3: Exploitation. The actual predation process is complex. For example, beleaguered prey may escape the harris hawks' enclosure. Harris hawks will make the necessary adjustments to the hunting strategy based on the behavior of the prey. To better simulate hunting behavior, this phase is represented in 4 modes.

(1) Soft besiege

When $|E| \geq 0.5$ and $r \geq 0.5$, the prey has enough energy to try to escape from the enclosure by random jumps but ultimately cannot escape. The Harris hawk hunts using a soft enclosure with the following equation.

$$X(t+1) = \Delta X(t) - E|JX_{rabbit}(t) - X(t)| \tag{16}$$

In which $\Delta X(t) = X_{rabbit}(t) - X(t)$ denotes the difference between the optimal individual and the current individual, $J = 2(1 - r_5)$ is the jump distance during the rabbit's escape, $r_5 \in [0, 1]$ is a random number.

(2) Hard besiege

When $|E| < 0.5$ and $r \geq 0.5$, the prey has neither enough energy to escape nor a chance to escape. The function of a hard besiege is described as follows:

$$X(t+1) = JX_{rabbit}(t) - E|X(t)| \tag{17}$$

(3) Soft besiege with progressive rapid dives

When $|E| < 0.5$ and $r \geq 0.5$, the prey has a chance to escape from the enclosure with sufficient escape energy. The hawk's location can be modified via the following equation.

$$X(t+1) = \begin{cases} Y = X_{rabbit}(t) - E|JX_{rabbit}(t) - X(t)|, F(Y) < F(X(t)) \\ Z = Y + S \times Levy(D), F(Z) < F(X(t)) \end{cases} \tag{18}$$

where D Indicates number of dimensions, S is a D-dim random vector, and Levy(·) represents the Levy flight function with the following equation.

$$Levy(x) = 0.01 \times \frac{r_6}{|r_7|^{\frac{1}{\delta}}}\left(\frac{\Gamma(1+\delta) \times \sin(\frac{\pi\delta}{2})}{\Gamma(\frac{1+\delta}{2}) \times \delta \times 2^{(\frac{\delta-1}{2})}}\right)^{\frac{1}{\delta}} \tag{19}$$

where $r_6, r_7 \in [0, 1]$ are random values and δ is a constant that can be set to 1.5.

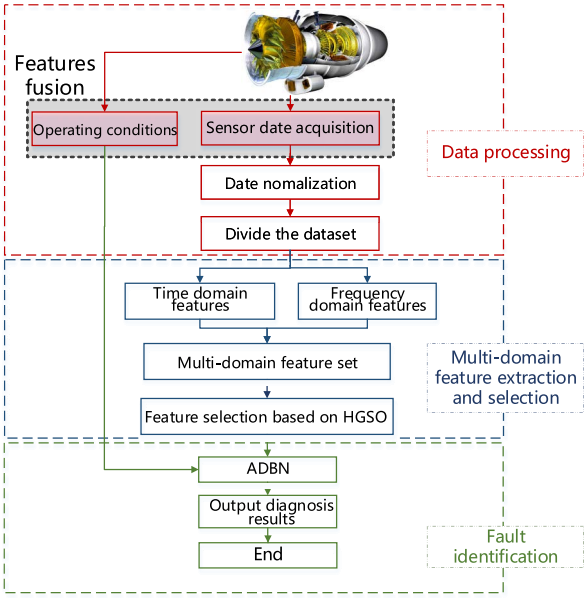

FIGURE 4. Flowchart of the proposed method.

(4) Hard besiege with progressive rapid dives

When $|E| < 0.5$ and $r < 0.5$, the prey has a chance to escape but not enough escape energy. Harris's hawks will use the following strategy for hunting.

$$X(t+1) = \begin{cases} Y = X_{rabbit}(t) - E|JX_{rabbit}(t) - X_m(t)|, F(Y) < F(X(t)) \\ Z = Y + S \times Levy(D), F(Z) < F(X(t)) \end{cases} \tag{20}$$

III. THE PROPOSED INTELLIGENCE FAULT DIAGNOSIS METHOD

To perform accurate fault diagnosis of aeroengine sensors, an intelligent fault diagnosis framework based on optimized and fused multidomain feature and ADBN is presented. The flow chart of the proposed algorithm is shown in Figure 4.

The overall process of this algorithm is as follows.

(1) **Data acquisition**: The sensor history data from the aeroengine control system is collected under various operating conditions. The dataset is then split into two categories, with the first 70% of the data used for training and the rest for testing.

(2) **Signal to preprocess**: The signal is normalized and rescaled to the range [0,1].

(3) **Multidomain feature extraction**: Multiple dimensions knowledge (time and spectral domain) provide excellent and comprehensive system understanding.

(4) **Feature selection**: In this step, the fault features are optimized by HGSO to obtain a subset of significant features after filtering.

(5) **Feature fusion:** Two kinds of features appear in this method. A subset of multidomain features is input to ADBN, and the obtained network extracted features are the primary features. The operational conditions are the secondary

TABLE 1. Multidomain feature set.

Number	Feature	Number	Feature
1	Maximum	14	Peak factor
2	Minimum	15	Pulse factor
3	Mean value	16	Square root amplitude
4	Median	17	Margin factor
5	Peak value	18	Kurtosis factor
6	Peak-peak value	19	Gravity frequency
7	Rectified mean value	20	Mean square frequency
8	Variance	21	RMS frequency
9	Standard deviation	22	Frequency variance
10	Kurtosis	23	Frequency standard deviation
11	Skewness	24	Spectral entropy
12	RMS	25	Signal complexity
13	Waveform factor		

features. To fully use both types of features, they are connected to form a complete feature set.

(6) **ADBN classification**: A series of enhancements are made to the DBN model to produce a more general and robust ADBN. The classification layer of the ADBN is performed to classify the fused feature vectors, which eventually automatically gives the final fault diagnosis results. It is worth mentioning that the method has two stages: offline training and online diagnosis.

The intelligent fault diagnosis method suggested in this paper outperforms previous methods in terms of diagnostic accuracy. It avoids the problem of insufficient fault information sources brought by single-domain features and minimizes the interference of irrelevant information by optimizing multidomain features. Furthermore, features fusion and DBN improvement boost the model's diagnostic accuracy.

The whole technique specifications are discussed below.

A. SENSOR SIGNAL PRE-PROCESSING

In practice, aeroengine variables are generally measured in different units. Because it is critical to eliminate mistakes in signal capture, measurement data should be standardized. In general, normalizing a signal allows it to be treated at the mean level (Eq.(21)). This can lower the complexity of computation and processing time in subsequent steps.

$$y(k) = \frac{x(k) - \bar{x}(k)}{\sigma(x(k))} \tag{21}$$

where $y(k)$ is the processed signal and $x(k)$ is the sensor signal to be processed.

$$\bar{x}(k) = \frac{1}{n}\sum_{i=1}^{n} x_i(k)$$

$$\sigma(x(k)) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i(k) - \bar{x}(k))^2}$$

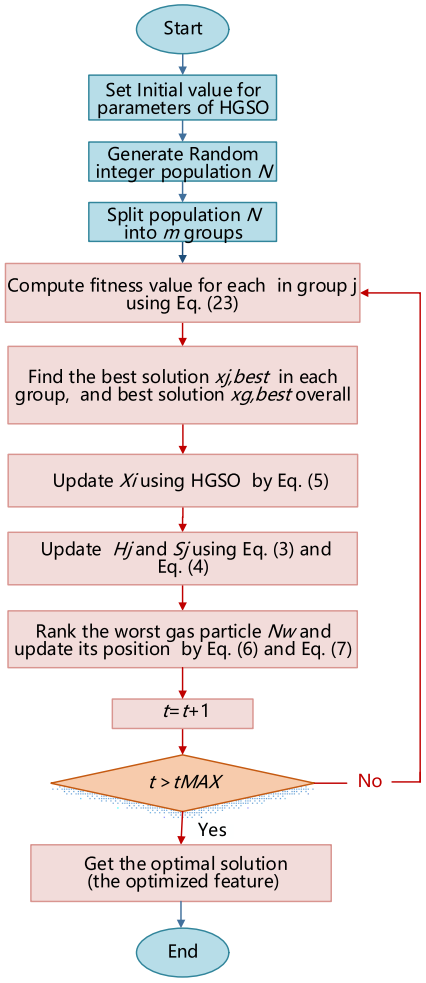

FIGURE 5. Flowchart of the HGSO feature selection.

B. MULTIDOMAIN FEATURE EXTRACTION

A signal can be conveyed by multidomain information, such as waveform, statistical facts, spectrum, etc. Compared with the traditional single domain features, multidomain features are used in this study. The signal is observed from multiple perspectives to obtain different forms of expressions and achieve more comprehensive fault information. Table 1 lists 25 features (time and frequency domain) selected in this paper.

C. PROCESS OF THE HGSO FEATURE SELECTION

The selection of appropriate information for exploitation and classification is crucial. A good feature selection(FS) algorithm can efficiently extract essential information from a dataset while removing redundant information and irrelevant features. In this context, the primary objective of feature selection is to improve the classification accuracy under a specific evaluation criterion or reduce the number of feature dimensions without compromising classification accuracy. It can select a subset of essential features from the original high-dimensional features and then use the selected subset of features together with some practical

TABLE 2. Pseudocode of HGSO feature selection.

1.Input:t_{MAX}: maximum number of iterations， and n: number of solution
2.Initialization HGSO: Generate the initial solutions N using Eq. (1)
3.Distribute the initial solutions N into j groups with the same Henry's coefficient H_j and constant value C_j, while H_j and C_j are different in every cluster.
4.while $t<t_{MAX}$
5. for each agent j do
6. Compute the fitness value for each N_i in group j using Eq.(23).
7. Determine the best solution $x_{j,best}$ in each group, and the best solution $x_{g,best}$ overall N.
8. Update each N using operators of HGSO as defined in Eq. (5)
9. Update coefficients: while t<iteration number
10. Update Henry's coefficient H_j and solubility $S_{i,j}$ using Eq.(3)-(4)
11. Update worst agents
12. Rank the worst gas particle N_w and update its position X_w using Eq. (6)-(7)
13. End for
14.End while
15.Return the best particle and the best solution
16.Output: the best solution

algorithms to accomplish the task of data clustering and classification.

Since the extensively large feature size in fault datasets of aeroengine sensors leading a rather large solution space, there is a high probability that traditional feature selection methods will encounter local optimality problems. This shortcoming may be overcome by metaheuristic algorithms, which evolve with a stochastic element allowing them to escape from local optimal solutions and thus obtain globally optimal solutions. The Henry Gas Solubility Optimization (HGSO) algorithm is a novel metaheuristic algorithm and an efficient search method. Its dynamic search behavior and global search capability enable it to solve complex FS problems efficiently.

This section will discuss the HGSO algorithm and its implementation steps for the FS problem, as shown in Figure 5. The proposed algorithm is described in the following subsections.

1) INITIAL POPULATION

An initial population containing N candidate solutions is first generated, where each individual represents a subset of the features to be selected. Each candidate solution's upper and lower bound are in the range of [0,1]. To facilitate the selection of feature subsets, the solution x_i^0 is converted into Boolean solution x_i^{bin}, as shown in Eq.(22) above.

$$x_i^{bin} = \begin{cases} 1 & if\ x_i^0 > 0.5 \\ 0 & otherwise \end{cases} \tag{22}$$

where 1 indicates the selected feature and 0 represents the unselected feature.

2) EVALUATING SOLUTIONS

In this step, the fitness function Fit_i of each solution x_i^{bin} is calculated. The Fit_i can serve as an evaluation index for feature quality, and is defined as

$$Fit_i = w_1 \times ERR_i + w_2 \times \frac{d_i}{D} \tag{23}$$

where ERR_i refers to the diagnostic error obtained according to the features selected from the i^{th} solution. d_i indicates the number of selected features, and D is the number of features in the original data set. w_1 and w_2 are weighting factors, which can be set as $w_1 = 0.99$, $w_2 = 1 - w_1$.

3) UPDATING SOLUTIONS

The optimal solutions $x_{j,best}$ and $x_{g,best}$ with optimal fitness are first determined. Then Eq.(3)-(7) is applied to update some solutions and coefficients based on the classical HGSO strategy.

4) STOPPING CONDITION

Keep repeating steps (2) (evaluating solutions) and Step (3) (updating solutions) until the maximum number of iterations is reached. Finally, the optimal solution is obtained, and the optimized multidomain feature subset is obtained.

The pseudocode of HGSO feature selection is given in Table 2.

D. FAULT IDENTIFICATION MODEL BASED ON FEATURE FUSION AND ADBN

1) FEATURE FUSION

The proposed multidomain feature extraction and feature selection methods provide the most effective feature combinations for fault identification. Subsequently, the multi-domain feature subsets are fused with the operating conditions of the aeroengine to obtain a complete feature set containing information about the operating environment. The multi-domain features and operating conditions represent different engine information. The details are described as follows.

Multi-domain characteristics are the cross-sectional parameters measured by the sensor that represent engine performance, such as temperature, pressure, etc. Sensor fault

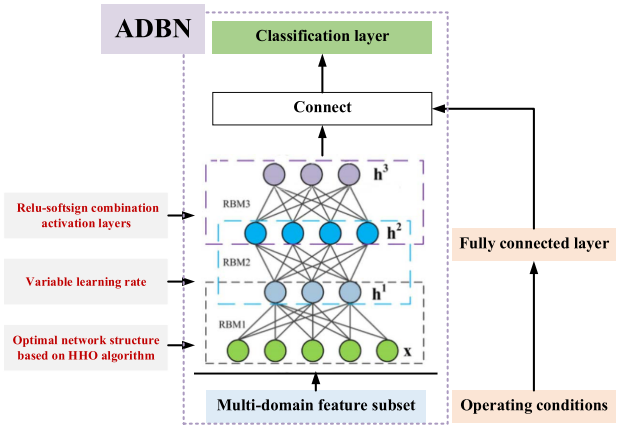

FIGURE 6. Fault identification model based on feature fusion and ADBN.

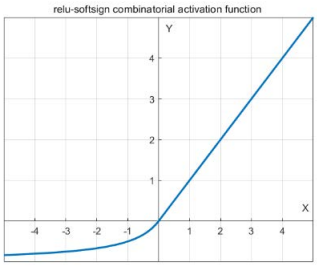

FIGURE 7. Curve of relu-softsign combinatorial activation function.

will be reflected in the variation of these measurements and is the main feature in this paper.

Operating conditions refer to the measured parameters of the aeroengine running environment (altitude and Mach number), which are used as auxiliary features in this investigation. This was done to account for the fact that when an aeroengine operates under varying operating conditions, its sensor measurements are affected by the operating conditions to some extent. Previous studies have not considered, which can hurt the fault diagnosis accuracy.

The feature fusion step proposed in this paper combines operating conditions and the features subset extracted from the cross-sectional parameters measured by the sensors in the previous subsection to provide more information for the fault diagnosis system.

The specific connection process is as follows: multi-domain features are input into the ADBN, and the first output is obtained after three RBM. Simultaneously, operational conditions serve as auxiliary input features that are fed into a fully connected layer. The two outputs are concatenated to form a complete feature set. Lastly, the classification layer of ADBN is used for fault diagnosis. The general layout of this step is displayed in Figure 6.

2) THE PROPOSED ADBN MODEL FOR FAULT IDENTIFICATION

Deep learning algorithms that combine brain-like mechanisms possess superior data mining capabilities and logical expression. As a result, it is better to use deep learning as a fault identification model than other methods. Among them, DBN, as one of the classical deep learning models, has become a buzzword in the field of intelligent fault diagnosis. Driven by the same spirit, DBN is selected as the fault identification model in this study.

As mentioned before, the performance limitations of traditional DBN negatively impact diagnostic accuracy. For example, constant learning rate and saturated activation layers (such as sigmoid, tanh) lead to low diagnostic accuracy and slow convergence [27]. Again, for example, the structure of deep learning networks applicable for a specific problem often comes from very time-consuming trial and error in the context of experience.

This paper proposes a series of improvements to the DBN model, and the adaptive DBN(ADBN) model is proposed.

TABLE 3. Pseudocode of optimization search process.

1.Input: t_{MAX}: maximum number of iterations， and N: Size of population.
2: Generate random N solutions (X).
3: Set the initial value for t= 1.
4: While t<t_{MAX} do
5: Compute the fitness value (F_i) for each X_i.
6: Set $X_{rabbitas}$ as the location of prey (best location).
7: For each X_i do
8: If $|E| \geqslant 1$ then
9: Use Eq. (14) to enhance X_i.
10: End if
11: If $|E| < 1$ then
12: If $|E| \geq 0.5$ and $r \geq 0.5$ then
13: Use Eq. (16) to enhance X_i.
14: Else if $|E| < 0.5$ and $r \geq 0.5$ then
15: Use Eq. (17) to enhance X_i.
16: Else if $|E| \geq 0.5$ and $r < 0.5$ then
17: Use Eq. (18) to enhance X_i.
18: Else if $|E| < 0.5$ and $r < 0.5$ then
19: Use Eq. (19) to enhance X_i.
20: End if
21: End if
22: End for
23: End while
24: Output: the best solution.

TABLE 4. Types and labels of sensor faults.

Fault type	Cause of fault	Label
Short-circuit	Pollution caused by the bridge road corrosion line short connection	0
Open-circuit	Signal line is broken, chip pin is not connected	1
Spike	Random disturbance in power supply and ground wire, surge,	2
Bias	Bias current or bias voltage	3
Drift	Temperature drift	4
Normal	------	5
Periodic disturbance	50Hz interference from the power supply	6

The upgrades adopt a new activation function and use an adaptive learning rate. Further, the optimal structure of the DBN determined by the advanced HHO algorithm allows the model to be adaptively applied to specific problems. The above process is described in detail as follows.

a: THE RELU-SOFTSIGN COMBINATORIAL ACTIVATION FUNCTION

Activation-functions play an crucial role in DBN to enhance the expressiveness of the network. However, the more widely used activation function such as Sigmoid and tanh functions are prone to the problem of gradient dispersion.

To address this problem, an improved activation function is developed and applied to the DBN model. The relu-softsign combinatorial activation function makes the x-negative semi-axis derivative no longer constant to zero, effectively alleviating the irreversible necrosis of neurons. It speeds up the convergence of the model and improves the learning

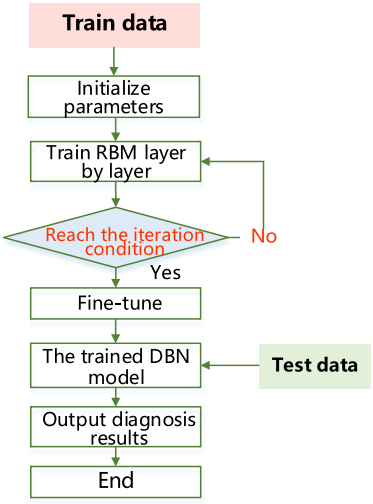

FIGURE 8. The training process of ADBN.

and generalization ability of the model. Its mathematical formula is:

$$f(x) = \begin{cases} \dfrac{x}{1+|x|}, & x \leqslant 0 \\ x, & x > 0 \end{cases} \tag{24}$$

The relu-softsign function is shown in Figure 7.

b: ADAPTIVE LEARNING RATE

A fixed learning rate usually leads to slow training convergence. Considering the drawback of a fixed learning rate, it is replaced by an adaptive learning rate. First, a small learning rate is initialized, and then the learning rate is adaptively adapted according to the reconstruction error. The calculated equation is described by:

$$\begin{aligned} &\Delta E = E(k) - E(k-1); \quad E_R = \left| \frac{\Delta E}{E(k)} \right| \\ &\gamma^k = f_d \gamma^{k-1}; \quad \Delta E \geqslant 0 \\ &\gamma^k = f_i \gamma^{k-1}; \quad \Delta E < 0 \& E_R < \sigma \\ &\gamma^k = \gamma^{k-1}; \quad \Delta E < 0 \& E_R \geq \sigma \end{aligned} \tag{25}$$

c: ADAPTIVE NETWORK STRUCTURE OPTIMIZED BY HHO

The number of nodes in the hidden layer has a significant impact on classification performance. Theoretically, as the number of nodes increases, the network's learning ability also increases. However, more nodes bring higher complexity and higher computational cost, which may leading to overfitting. An evolutionary strategy of HHO with an adaptive training procedure is used to obtain the optimal structure of the DBN. The decision variable is set to be the number of hidden layers of the DBN, and the fitness function is the diagnostic error *ERR*.

The pseudocode of the optimization search process is shown in Table 3.

The classifier of ADBN is a Softmax regression model. The Softmax classification layer determines the output of the ADBN, and the category corresponding to the maximum probability is the type to which the fault belongs. In this study, the aeroengine control system sensor has a total of seven health conditions: six fault types and a normal state. That is, the fault identification model has seven possible outputs: the sensor is normal, or the sensor has a fault labeled 1, or the sensor has a fault labeled 2, and so on. It should be added that since supervised training is essential in this paper, labels should be set up based on the health of the sensors. A detailed description of the faults of the sensors and the labeling process is depicted in Table 4.

The loss function is a cross-entropy loss function. Combining the two employs an interclass competition mechanism to learn inter-class information effectively. The cross-entropy loss function is calculated as:

$$F_c = -\frac{1}{n} \sum_{k=1}^{N_S} [y_c \times ln y_r + (1 - y_c) \times ln(1 - y_r)] \tag{26}$$

where, N_S denotes the dimensionality of the samples in the dataset, y_c and y_r are the output labels and reference labels of the classifier, respectively.

The training process of DBN consists of two phases: pre-training and fine-tuning, as illustrated in Figure 8. In pretraining, the log-likelihood function of the RBM is maximized using stochastic gradient ascent approach and the joint distribution defined by the RBM model is obtained using K times Gibbs sampling, called the contrastive divergence(CD) approach. And the RBMs are trained in an unsupervised manner using the hierarchical greedy technique. The lower layer serves as input to the upper layer until the last RBM is trained. After that, the model parameters are fine-tuned using a back-propagation algorithm based on the labels of known fault types. The purpose of fine-tuning is to optimize the model training results to achieve the desired performance.

E. EVALUATION INDICATOR

To better evaluate the performance of the proposed fault diagnosis method, the following evaluation indicators are selected in this paper.

1) AVERAGE ACCURACY

Accuracy refers to the probability that a fault is correctly classified. The average accuracy is the average of the accuracies obtained after running the fault diagnosis method several times and is calculated by:

$$ACC_{mean} = \frac{1}{M} \frac{1}{N_S} \sum_{k=1}^{M} \sum_{r=1}^{N_S} (y_c == y_r) \tag{27}$$

where M is the number of runs, $sum(y_c^j == y_r^j)$ denotes the summation of the number of samples with the same model output and labeled results, i.e., the total number of correctly classified faults of each type.

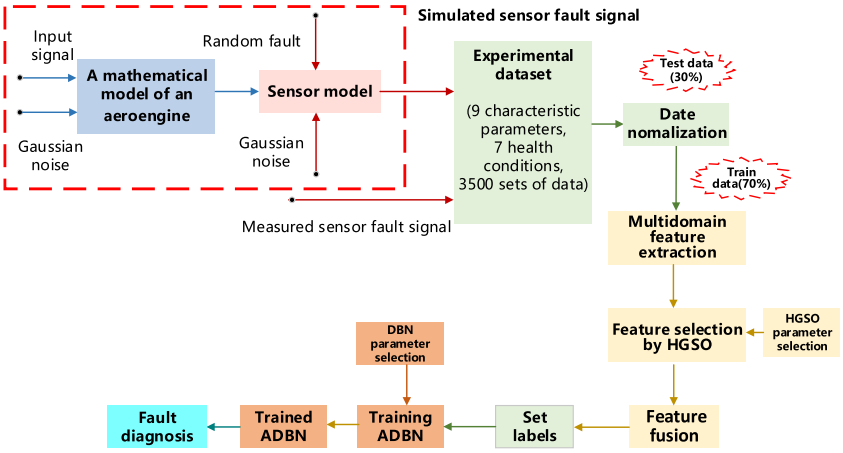

FIGURE 9. Flowchart of experiment.

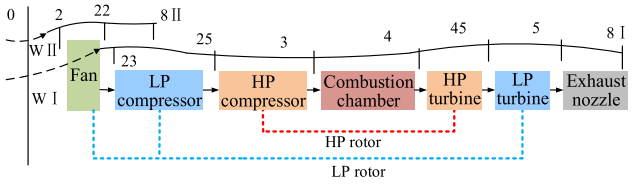

FIGURE 10. Structural sketch of the aeroengine model.

TABLE 5. Fault simulation methods

Fault	Simulation method
Short-circuit	Signal is close to 0.1
Open-circuit	Signal approaches maximum
Spike	Add a pulse signal to the original signal
Bias	Add a small constant or random signal to the original signal
Drift	Signal is offset by a certain rate
Normal	No change
Periodic disturbance	Signal of a certain frequency superimposed on the original signal

2) AVERAGE COMPUTATION TIME

This index emphasizes the rate of operation. Its mathematical expression is as follows.

$$T_{mean} = \frac{1}{M}\sum_{k=1}^{M}(T_k) \tag{28}$$

where T_k is the time spent for the i^{th} operation.

IV. EXPERIMENT AND DISCUSSION

The effectiveness of the proposed method is validated experimentally in the MATLAB 2018b environment. Inspired by ablation experiments, existing methods are compared to demonstrate the performance of the proposed method, The experiment's flowchart is visualized in Figure 9.

A. DATA PREPARATION

The dataset used in the following experimentation is provided from two scenarios. One part is from some sensor signals collected from the running records of a certain type of aeroengine, and the other part is from MATLAB simulation. There are about 300 pieces of real data and 3200 pieces of simulated data.

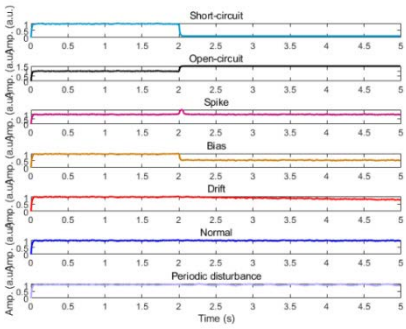

FIGURE 11. Temporal waveform of sensor signals for 7 health conditions.

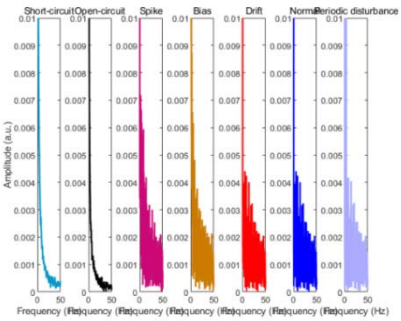

FIGURE 12. Frequency-domain waveform of sensor signals for 7 health conditions.

The simulation model of the sensor needs to be determined first. Referring to the previous literature, the second-order inertia plan [4] was used to build the simulation model of the sensor. Its transfer function is:

$$1) \quad G(s) = \frac{w_n^2}{s^2 + 2\xi w_n s + w_n^2} \cdot e^{-\tau s} \tag{29}$$

where $\xi = 1.25$, $w_n = 9$, $\tau = 1.2$.

TABLE 6. Description of the dataset.

Fault type	Collected signals	Label
Short-circuit	500	0
Open-circuit	500	1
Spike	500	2
Bias	500	3
Drift	500	4
Normal	500	5
Periodic disturbance	500	6

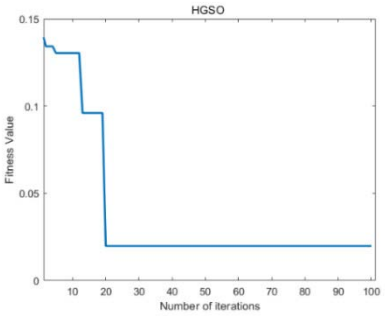

FIGURE 13. The graph of fitness changes of HGSO with iteration.

The input signal for the sensor simulation model comes from the C-MAPSS aeroengine simulation model. It can be clearly seen in Figure 9 that the sensor model is connected in series with the engine model. This engine model is built based on the component characteristics and the input-output relationships between the components. The aeroengine model includes the fan, compressor, combustion chamber, turbine, and exhaust nozzle. Its structure sketch is illustrated in Figure 10.

Next, random faults are injected into the output signals of the sensor model the sensor signals. That is, the sensor signals are processed according to the method in Table 5 for fault simulation. In this study, a large amount of simulated fault data is generated by randomly varying the fault set time, mode, level, period, and other parameters. Ambient noise (Gaussian white noise) is then randomly injected into the input and output signals to make the experimental data as realistic as possible and to realistically reflect the robustness of the proposed method.

During the experiment, the data acquisition time interval is 0.01s. In practice, sensor faults might be subjected to different types. The dataset used in this paper contains sensor signals for seven health conditions. That is, it encompasses six faults types and the normal state of the sensor. The time domain and frequency domain curves of sensor signals are presented in Figures 11 and 12.

It must be mentioned that although the simulated fault data are hardly identical to the recorded operating data, this study still attempts to simulate the fault to a large extent and provide the most considerable discrimination. It makes meaningful sense since deep learning algorithms require a lot of data

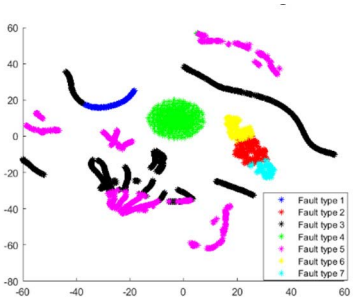

FIGURE 14. Raw sensor signals.

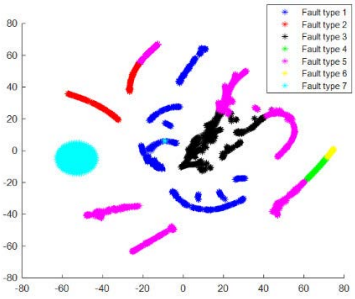

FIGURE 15. Raw multi-domain features without feature selection.

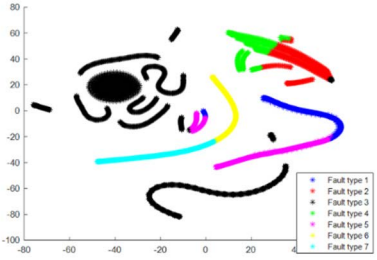

FIGURE 16. The optimized multi-domain features by HGSO.

for training. Due to limited conditions, there is not enough running data recorded to support the good performance of the deep learning algorithm. Furthermore, because some faults are naturally rare and difficult to obtain, an effective method is needed to utilize this information, just like the suggested study.

The data set contains 7 types of sensor health conditions. There are 500 samples for each sensor fault type, and the total number of samples is 3500. The dataset was randomly divided into two parts: 70% for the training set and 30% for the test set. Regarding the generalizability of the model, the idea of K-fold cross-validation was adopted. The original dataset was divided into 10 equal parts. Randomly, three of these parts were used as the test set and the rest were used as the training set to train the model and calculate the accuracy of the model on the test set. Each time, different data is considered as the test set and repeated 30 times(M=30), and the average accuracy is finally taken as the final model accuracy.

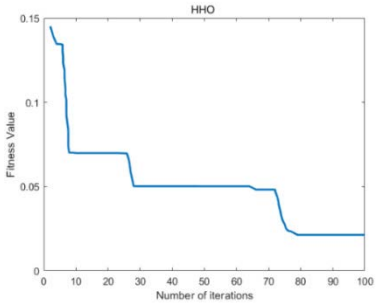

FIGURE 17. The graph of fitness changes of HHO with iteration.

TABLE 7. Network parameters

Name	Parameter of DBN	Parameter of ADBN
Input layer	Input data(1*10)	Input data(1*10)
Hidden layer 1	10	10
Hidden layer 2	30	30
Hidden layer 3	30	30
Hidden layer 4	30	30
Output layer	Output data(7*1)	Output data(7*1)
Learning rate for weights	0.1	Adaptive
The max epoch	100	100

B. RESULTS AND DISCUSSION

1) INFLUENCE OF OPTIMIZED FEATURES

To capture the diagnostic information, there are 18 time-domain features and 7 frequency-domain features extracted from the sensor signal, as listed in Table 1, 25 in total. Then, to obtain a sufficiently reduced feature subset, feature optimization is performed by the HGSO algorithm. The maximum number of iterations is 100, and the remaining of the parameters that need to be set for HGSO are mentioned in the previous description of the algorithm. It should be noted that most of the parameters are set with reference to the literature [29]. The feature selection process was repeated 30 times. The HGSO algorithm automatically selects ten primary features from a pool of twenty-five features after multiple iterations. The variation of HGSO fitness value during the iterative process is shown in Figure 13. The final feature selection's results were [6], [8], [11], [15], [16], [17], [19], [21], [23], [24], representing peak-peak value, variance, skewness, pulse factor, square root amplitude, margin factor, gravity frequency, RMS frequency, frequency standard deviation, and spectral entropy, respectively. These ten features are the most fault-sensitive features that were selected. They form the best subset of features for providing the fault recognition model with minimal but high-quality fault information, thereby reducing the computational burden and enhancing the classification accuracy.

To verify the effect of feature optimization, the idea of an ablation experiment is adopted to compare the optimized multi-domain features with the raw ones. The metric for comparison is the data classifiability effect by the T-distributed stochastic neighbor embedding (T-SNE) technology. TSNE technology is a nonlinear method for visualizing dimensionality reduction [23]. It allows evaluating the separability of data and visualizing the data structure based on clustering performance. T-SNE is used to evaluate the effectiveness of optimized multi-domain features. Data of the same fault type are clustered in T-SNE. The higher the similarity between data points, the better the clustering performance of the data. Additionally, the visual graphs of the raw sensor signal is drawn. As can be seen from Figure 14, the characteristics of sensor data for different health conditions are overlapped and crossed together, making it difficult to distinguish the fault type. This again demonstrates the need for feature extraction.

Figure 15 and Figure 16 display the graphical representations of the raw multi-domain features and the simplified features, respectively. It is intuitive to conclude that the data separability in Figure 16 is better compared to Figure 15. That is, the optimized multi-domain features are more effective than the raw ones. It can be seen that using the raw features have a negative impact on the performance of fault diagnosis. It is also obvious that the raw features have a higher dimension which increases the structural complexity of the network. The optimization of the features permits the retention of salient features and the discarding of redundant information. The most appropriate architecture subset results in efficient fault detection performance when accompanied by thoroughly chosen by HGSO. The performance of the proposed method is outstanding, even in the case of minimal knowledge.

2) INFLUENCE OF FUSED FEATURES AND FAULT CLASSIFIER

Once a subset of multidomain features has been obtained, the ADBN model is applied to fault diagnosis. The proposed ADBN model as a fault classifier is one of the ways to improve the accuracy of engine sensor fault diagnosis, which is the subject of this study. In addition, the fusion feature considering H and Ma is also one of the advantages of the proposed method. To further demonstrate the superiority of the proposed method in this paper, the proposed method (referred to as Algorithm 3), the other two fault diagnosis methods. In the same spirit as the ablation experiment, one comparative fault diagnosis method did not consider the fusion features (referred to as Algorithm 1), and the other utilizes the traditional DBN model as the fault classifier (referred to as Algorithm 2).

In this research, the ADBN model consists of three layers of RBM superimposed. The number of nodes in the input layer is set according to the dimensionality of the input data. This suggests that the input of ADBN is the optimized feature with the size of 1*10. The output layer nodes are set to 1*7. This means that the output of ADBN is the health condition of the sensor. Since the number of nodes in the hidden layer has a great impact on the classification performance. In this study, the HHO algorithm is employed to determine the optimal structure of the ADBN, which improves diagnostic accuracy and reduces training time. The evolutionary trajectory of

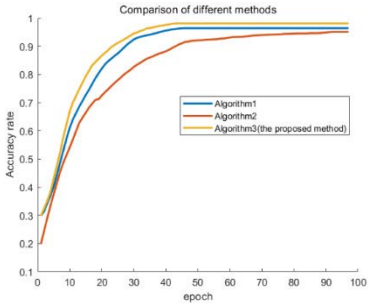

FIGURE 18. The comparison of different methods.

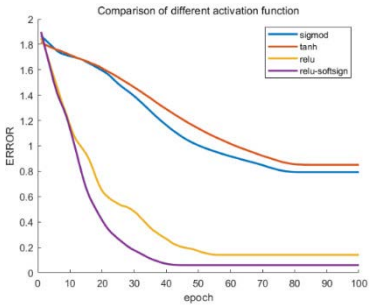

FIGURE 19. The comparison of different activation function.

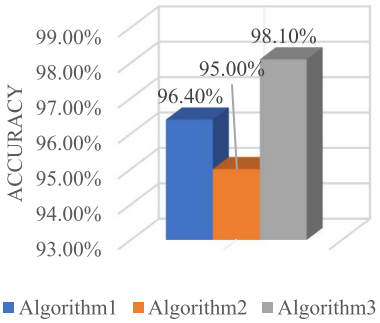

FIGURE 20. The Intuitive comparison of accuracy.

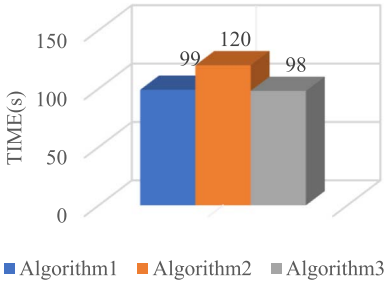

FIGURE 21. The Intuitive comparison of computation time.

the optimization process is given in Figure 17. Finally, the number of nodes in the hidden layer of the ADBN provided by HHO is [30, 30, 30]. The parameters to be set are listed in Table 7. Among them, the maximum number of epochs for training the model is a relatively important parameter. Therefore, the loss of the model on the test set after each epoch is dynamically observed during the training process. The selection criterion was a compromise between accuracy and computational burden. The parameter settings in DBN used are given in Table 7, and the number of nodes in the hidden layer of DBN is obtained using the traditional try and error method.

The obtained classification results are shown in Figure 18, and the specific calculation results are in Table 8. It is mentioned that the proposed method in the research offers considerable performance. From Figure 18 and Table 8, it is observed that the proposed method achieves an eminent diagnostic accuracy of 98.1%. It indicated that the proposed method is prone to improving the discrimination of faults. It is worth mentioning that the computation time of the proposed method is also ideal, which is 98s. It shows the merit of the proposed method in the data processing. Further, the ADBN model is considered an excellent fault recognition model due to its powerful feature learning and data expression capabilities. When facing complex tasks, the deep structure of ADBN is a propulsive factor. This advantage becomes more evident in the fault diagnosis of aeroengine sensors because there are numerous variables and a highly non-linear relationship.

According to the experimental findings, short-circuit fault, and open-circuit fault have the highest diagnostic accuracy. Short circuit fault, open circuit fault, pulse fault, and bias

TABLE 8. Comprehensive comparison of the performance of each method.

Number	Name	Algorithm1	Algorithm 2	Algorithm 3 (the proposed method)
1	Short-circuit	0.993	0.973	1.00
2	Open-circuit	0.993	0.960	0.993
3	Spike	0.967	0.953	0.987
4	Bias	0.947	0.940	0.980
5	Drift	0.933	0.920	0.960
6	Normal	0.940	0.940	0.960
7	Periodic disturbance	0.973	0.960	0.987
Total accuracy		0.964	0.950	0.981
Train time(s)		99	120	98

fault belong to sudden hard faults that are easier to identify and have a higher accuracy. Once an open/short circuit fault occurs, the measured value of the sensor quickly rushes to its maximum/minimum value, which is the easiest fault to identify. High accuracy rates of them were achieved in all three algorithms. In particular, in Algorithm 3, the accuracy of short-circuit faults and open-circuit faults reached 100% and 99.3%. Pulse fault, bias fault, and periodic fault were the next highest. The Pulse fault and bias fault are also sudden hard faults that are easier to identify. Because the duration of pulse fault is relatively short, the accuracy rate is slightly lower compared to faults with longer fault durations, such as open-circuit fault or short-circuit fault, which reached 98.7% in Algorithm 3. However, despite the small magnitude of individual bias faults, the overall accuracy of this type of fault is lower than that of open-circuit or short-circuit faults, which have larger fault magnitudes. In Algorithm 3, the bias faults reached 98.0%. Periodic faults are not hard

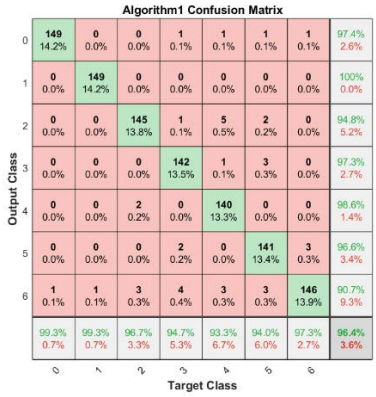

FIGURE 22. The confusion matrix of Algorithm 1.

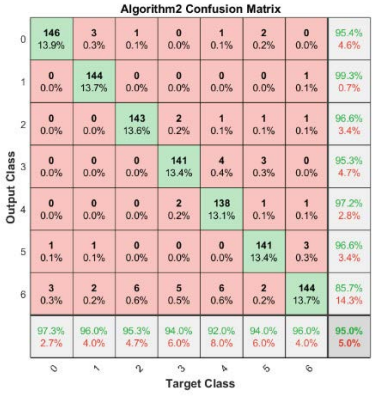

FIGURE 23. The confusion matrix of Algorithm 2.

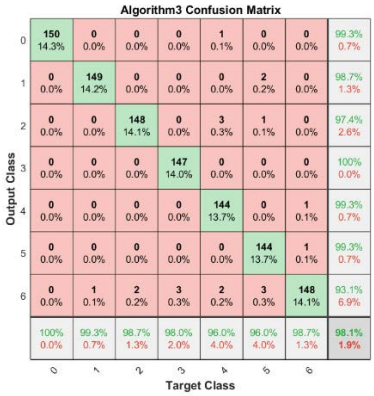

FIGURE 24. The confusion matrix of Algorithm 3.

faults, but the fault accuracy is also higher due to their fixed frequency. In Algorithm 3, the accuracy of periodic fault reaches 98.7%. The drift fault, on the other hand, is a time-dependent soft fault, and its fault severity increases gradually as time increases. It has weak characteristics in the early stage of the fault. It is easily drowned in the noise, making it more difficult to distinguish from the normal condition. Therefore, the diagnosis accuracy of the drift fault and the normal state is the lowest. In Algorithm 3, their accuracy reaches both 96%.

To further verify that the fusion features proposed in this paper have a positive impact on fault diagnosis results, Algorithm 1 is compared with Algorithm 3. Both methods use the same ADBN and optimized multi-domain features; the only difference is that Algorithm 2 does not consider feature fusion. The experimental result is represented graphically, as shown in Figures 18, 20, and 21. It demonstrates that the average accuracy of Algorithm 1 is only 96.7%, which is lower than Algorithm 3. And as can also be seen in Table 8, the computation time for Algorithm 1 is 99s. The main difference between Algorithm 1 and Algorithm 3 is the improvement of fault accuracy rather than time. This is because the inclusion of additional features makes fault diagnosis results less influenced by operating conditions. Due to aerodynamic relationships, the measured values of the engine's sensors are affected by H and Ma. Fusing the operating conditions with the features learned by the ADBN enables the classifier to understand their relationship and thus can provide better diagnostic results.

Comparing Algorithm 3 with Algorithm 2, it is clear that Algorithm 3 has the best diagnostic accuracy. The experimental results in Table 8 show that even though both use the same features as input, the accuracy of Algorithm 3 is higher than that of Algorithm 2 because of the improvements made to the DBN model. Specifically, the accuracy of Algorithm 3 was as high as 98.1%, but the accuracy of Algorithm 2 was only 95%. It indicates that the proposed ADBN classification is more accurate compared to the traditional DBN model. More importantly, the results show that the proposed ADBN has a fast training speed. As can be seen in Figure 18, ADBN converges at about 40 rounds, while DBN converges at about 50 rounds, which is slower than the ADBN's convergence rate. The results validate the improved effectiveness of the proposed ADBN to a great extent, involving adaptive learning rates and adaptive structures.

To further verify the positive effect of relu-softsign on ADBN, we added comparative experiments with different ADBN structures. Each structure of the ADBN model uses a different activation function, Sigmoid, tanh, relu, relu-softsign. The results in Figure 20 show the error function of the training process after each training period. It is clear that the training process of the proposed relu-softsign is fast and smooth with the fastest convergence rate at about round 40.

In effort to more explicitly describe the diagnostic capability of the three methods, confusion matrices are calculated for each. The confusion matrix is a visual sketch of the classification effect, which can depict the relationship between the real class attributes of the sample data and the recognition outcomes. It is frequently employed to assess classifier performance. Figures 22, 23, and 24 illustrate the confusion matrix for each method. For instance, there are 150 samples of fault type 1, of which 149 samples were correctly classified by Algorithm 1 with an accuracy rate of 93.3%. Algorithm 2 correctly classified 146 samples with an accuracy of 97.3%. Algorithm 3 proposed in this paper correctly classified all 150 samples with an accuracy rate

TABLE 9. Comprehensive comparison of the performance of each method.

Rank	Number	Weight	Rank	Number	Weight
1	6	1	14	18	0.592
2	11	0.942	15	3	0.590
3	19	0.939	16	7	0.589
4	21	0.916	17	13	0.532
5	24	0.905	18	14	0.528
6	23	0.877	19	17	0.522
7	20	0.861	20	25	0.522
8	8	0.839	21	10	0.339
9	11	0.809	22	9	0.233
10	16	0.762	23	1	0.231
11	15	0.678	24	2	0.086
12	12	0.671	25	5	0
13	4	0.627			

TABLE 10. Comprehensive comparison of the performance of each method.

Number	Name	Algorithm 3 the proposed method	Algorithm 4	Algorithm 5
1	Short-circuit	1.00	0.980	0.960
2	Open-circuit	0.993	0.973	0.953
3	Spike	0.987	0.947	0.933
4	Bias	0.980	0.933	0.913
5	Drift	0.960	0.927	0.900
6	Normal	0.960	0.940	0.920
7	Periodic disturbance	0.987	0.953	0.940
Total accuracy		0.981	0.950	0.931

of 100%. These three confusion matrices demonstrate the significant advantages of the proposed method for aeroengine sensor fault diagnosis. The main reason is the use of feature selection, feature fusion, and ADBN. The above influencing factors lead to better results for the fault diagnosis proposed in this paper.

3) COMPARISON WITH OTHER METHODS

The quality of the developed algorithm in this paper is compared with some popular methods that has been applied with success to solve the fault diagnosis problem.

The excellent performance of the proposed method in this paper is due to good features and an excellent fault identifier. In order to demonstrate the superiority of the used meta-heuristic intelligent algorithm approach(HGSO), the ReliefF algorithm was selected as a comparison algorithm for feature selection (referred to as Algorithm 4). Also, during the fault diagnosis process, they use the same ADBN model as fault classifier and auxiliary features for a fair comparison between these two methods.

The ReliefF algorithm is considered as one of the most successful preprocessing algorithms due to its advantages such as high efficiency and no restriction on data types, and is also widely used in the field of fault diagnosis[40], [41]. The algorithm assigns different weights to features according to the relevance of each feature and category, and features with weights less than a certain threshold will be removed. The advantage of feature ranking method is that it is independent of classifier used and the features are selected based on their ranking. Tables 9 shows the feature ranking of the calculated features, and Table 10 shows the comparison results of different methods.

Features are ranked in order of weight from largest to smallest, and then the weights are normalized. The normalized weights less than 0.6 are eliminated, and the features are selected according to the order of their weights to form the feature set. It is observed from Table 9 that 13 characteristics are eligible. The final feature selection's results were [4], [6], [8], [11], [12], [15], [19], [20], [21], [23], [24]. Similar to the method used in this paper, a subset of features and auxiliary features are fed into the ADBN model to automatically obtain fault diagnosis results. Since by comparing the behavior of ReliefF with HGSO, it is noticed the superiority of HGSO over ReliefF. In Table 10, the accuracy of Algorithm 3 was 98.1%, and the accuracy of Algorithm 4 was 94.6%. for one thing, the ReliefF actually ranks the original feature variables in terms of merit, and does not determine the number of extracted feature dimensions. It requires manual setting of thresholds to filter the desired features, and cannot achieve adaptive feature extraction. In contrast, the HGSO method can adaptively determine the optimal combination of feature subsets, which can produce higher classification accuracy. For another, since the evaluation criteria of ReliefF are independent of the specific learning algorithm, while HGSO belongs to the wrapper type method, which takes the fault classification accuracy as the evaluation criteria of the feature subset, the feature set selected by ReliefF is lower than HGSO in the classification accuracy method.

Another advantage of the proposed method in this paper is the use of deep learning networks. To verify this, SVM is chosen as the fault classifier for comparison (referred to as Algorithm 5). Both methods take the same features as the indicator vector to characterize the health status. SVM is one of the most popular supervised learning algorithms[42], [43]. SVM can construct a hyper plane or set of hyper planes in a high-dimensional space, which can be used for classification and regression. The parameters of SVM are set with reference to the literature [43]. The experimental results in Table 10 show that the accuracy of method 3 is higher than that of method 5 even if the same features is used. It is seen from Table 9 that the accuracy rate of method 3 is as high as 98.1%, but the accuracy rate of method 5 is only 92.8%. Compared with SVM, the ADBN model has powerful feature learning and data representation capabilities, which can be well adapted to complex fault identification tasks. ADBN extracts more detailed and obvious feature differences of fault types through multi-dimensional multi-layer mapping, and achieves the optimization of diagnostic model parameters through backward tuning. The deep structure of ADBN is its advantage, and this advantage becomes more obvious in fault diagnosis of aero-engine sensors, where highly nonlinear relationships and the influence of disturbances are included in the process. The experimental results demonstrate ADBN has improved diagnostic accuracy compared to SVM.

V. CONCLUSION

Fault diagnosis of aeroengine control system sensors has become a vital issue that requires prompt solutions. To meet the need for high accuracy and low computational effort, this paper presents a new intelligent fault diagnosis method for aeroengine sensor with better features and more effective fault classifiers. Some conclusions are summarized as follows:

(1) The selected multidomain feature by HGSO reduces the input dimension of the fault identification model and redundant information in the signal. This improves the classified performance compared with the unprocessed multidomain feature.

(2) The solution considers the influence of operating conditions on the measured values of the aeroengine sensors. The operating conditions and features learned by ADBN are fused to improve diagnostic performance.

(3) The proposed ADBN is verified to be potential for efficient sensor fault classification. The optimal structure of DBN is determined by HHO. A Relu-softsign activation layer and variable learning rate are developed to speed up the training process. The ADBN has improved convergence speed and detection performance. It also can be a generic solution that can be applied to different aeroengines.

(4) The framework's performance is validated experimentally and is proved to be promising for accommodating a compromise between the requirements for high diagnostic accuracy and low computational burden. In short, this study provides a efficient and adaptable way to aeroengine sensor fault diagnosis.

Future work will focus on building a more efficient and robust framework for applications of sensor fault diagnosis when unknown faults exist, as well as applying transfer learning to improve the generalizability of the models.

REFERENCES

[1] F. Lu, Y. Huang, J. Huang, and X. Qiu, "A hybrid Kalman filtering approach based on federated framework for gas turbine engine health monitoring," *IEEE Access*, vol. 6, pp. 9841–9853, 2018.

[2] H. Yu *et al.*, "Sensor fault diagnosis of gas turbine engines using an integrated scheme based on improved least squares support vector regression," *Proc. Inst. Mech. Eng., G, J. Aerosp. Eng.*, vol. 234, no. 3, pp. 607–623, 2019.

[3] H. Sun, Y. Guo, and W. Zhao, "Fault detection for aircraft turbofan engine using a modified moving window KPCA," *IEEE Access*, vol. 8, pp. 166541–166552, 2020.

[4] B. Pourbabaee, N. Meskin, and K. Khorasani, "Sensor fault detection, isolation, and identification using multiple-model-based hybrid Kalman filter for gas turbine engines," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 4, pp. 1184–1200, Jul. 2016.

[5] C.-W. Fei, H. Li, H.-T. Liu, C. Lu, L.-Q. An, L. Han, and Y.-J. Zhao, "Enhanced network learning model with intelligent operator for the motion reliability evaluation of flexible mechanism," *Aerosp. Sci. Technol.*, vol. 107, Dec. 2020, Art. no. 106342.

[6] C.-W. Fei, H. Li, C. Lu, L. Han, B. Keshtegar, and O. Taylan, "Vectorial surrogate modeling method for multi-objective reliability design," *Appl. Math. Model.*, vol. 109, pp. 1–20, Sep. 2022.

[7] H. Li, L. Gou, H. Zheng, and H. Li, "Intelligent fault diagnosis of aeroengine sensors using improved pattern gradient spectrum entropy," *Int. J. Aerosp. Eng.*, vol. 2021, pp. 1–20, Jan. 2021.

[8] T. Khaoula, C. Nizar, V. Sylvain, and T. Teodor, "Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges," *Ann. Rev. Control*, vol. 42, pp. 63–81, Dec. 2016.

[9] Y. Yuan, S. Ding, X. Liu, and Q. Pan, "Hybrid diagnosis system for aeroengine sensor and actuator faults," *J. Aerosp. Eng.*, vol. 33, no. 1, Jan. 2020, Art. no. 04019108.

[10] T. Kobayashi and D. L. Simon, "Evaluation of an enhanced bank of Kalman filters for in-flight aircraft engine sensor fault diagnostics," *J. Eng. Gas Turbines Power-Trans.*, vol. 127, no. 3, pp. 497–504, 2005.

[11] E. Kiyak, A. Kahvecioglu, and F. Caliskan, "Aircraft sensor and actuator fault detection, isolation, and accommodation," *J. Aerosp. Eng.*, vol. 24, no. 1, pp. 46–58, 2011.

[12] L. Gou, Y. Shen, H. Zheng, and X. Zeng, "Multi-fault diagnosis of an aeroengine control system using joint sliding mode observers," *IEEE Access*, vol. 8, pp. 10186–10197, 2020.

[13] R. J. Patton and J. Chen, "Review of parity space approaches to fault diagnosis for aerospace systems," *J. Guid., Control, Dyn.*, vol. 17, no. 2, pp. 278–285, 1994.

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

[15] C. L. Wen, F. Y. Lv, Z. J. Bao, and M. Q. Liu, "A review of data driven-based incipient fault diagnosis," *J. Automatica Sinica*, vol. 42, no. 9, pp. 1285–1299, Sep. 2016.

[16] S. L. Brunton, J. N. Kutz, K. Manohar, A. Y. Aravkin, K. Morgansen, J. Klemisch, N. Goebel, J. Buttrick, J. Poskin, A. W. Blom-Schieber, T. Hogan, and D. McDonald, "Data-driven aerospace engineering: Reframing the industry with machine learning," *AIAA J.*, vol. 59, no. 8, pp. 2820–2847, 2021.

[17] J. Tian, G.-W. Yi, C.-W. Fei, J. Zhou, Y.-T. Ai, and F.-L. Zhang, "Quantum entropy-based hierarchical strategy for inter-shaft bearing fault detection," *Struct. Control Health Monit.*, vol. 28, no. 12, p. e2839, 2021.

[18] Y. Ai, J.-Y. Guan, C.-W. Fei, J. Tian, and F. -L. Zhang, "Fusion information entropy method of rolling bearing fault diagnosis based on N-dimensional characteristic parameter distance," *Mech. Syst. Signal Process.*, vol. 88, pp. 123–136, May 2017.

[19] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1693–1702, Jul. 2017.

[20] A. Abid, M. T. Khan, and M. S. Khan, "Multidomain features-based GA optimized artificial immune system for bearing fault detection," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 1, pp. 348–359, Jan. 2020.

[21] Z. N. S. Vanini, N. Meskin, and K. Khorasani, "Multiple-model sensor and components fault diagnosis in gas turbine engines using autoassociative neural networks," *J. Eng. Gas Turbines Power*, vol. 136, no. 9, Sep. 2014, Art. no. 091603.

[22] K. Cai, S. Xie, and W. Yang, "Fault diagnosis and adaptive reconfiguration control for sensors in aeroengine based on improved least squares support vector," *J. Aerosp. Power*, vol. 23, no. 6, pp. 1118–1126, 2008.

[23] A. Fentaye, S. I. U.-H. Gilani, and A. T. Baheta, "Performance-based fault diagnosis of a gas turbine engine using an integrated support vector machine and artificial neural network method," *Proc. Inst. Mech. Eng.*, vol. 233, no. 6, pp. 786–802, 2019.

[24] O. Fink, Q. Wang, M. Svensen, P. Dersin, and W. J. Lee, "Potential, challenges and future directions for deep learning in prognostics and health management applications," *Eng. Appl. Artif. Intell.*, vol. 92, Jun. 2020, Art. no. 103678.

[25] D. Ezzat, A. E. Hassanien, A. Darwish, M. Yahia, A. Ahmed, and S. Abdelghafar, "Multi-objective hybrid artificial intelligence approach for fault diagnosis of aerospace systems," *IEEE Access*, vol. 9, pp. 41717–41730, 2021.

[26] L. Gou, H. Li, H. Zheng, H. Li, and X. Pei, "Aeroengine control system sensor fault diagnosis based on CWT and CNN," *Math. Problems Eng.*, vol. 2020, pp. 1–12, Jan. 2020.

[27] M. Lv, S. Liu, X. Su, and C. Chen, "Deep transfer network with multi-kernel dynamic distribution adaptation for cross-machine fault diagnosis," *IEEE Access*, vol. 9, pp. 16392–16409, 2021.

[28] K. Zhang *et al.*, "Fault diagnosis of planetary gearbox using a novel semi-supervised method of multiple association layers networks," *Mech. Syst. Signal Process.*, vol. 131, pp. 243–260, Sep. 2019.

[29] P. Tamilselvan and P. Wang, "Failure diagnosis using deep belief learning based health state classification," *Rel. Eng. Syst. Saf.*, vol. 17, no. 12, pp. 1287–1304, 2016.

[30] Y. D. Liu, X. X. Xie, and X. F. Zheng, "Fault diagnosis of aero-engine sensor based on deep learning," *Transducer Microsyst. Technol.*, vol. 36, no. 9, pp. 147–150, 2017.
[31] D.-L. Feng, M.-Q. Xiao, Y.-X. Liu, H.-F. Song, Z. Yang, and Z.-W. Hu, "Finite-sensor fault-diagnosis simulation study of gas turbine engine using information entropy and deep belief networks," *Frontiers Inf. Technol. Electron. Eng.*, vol. 17, no. 12, pp. 1287–1304, Dec. 2016.
[32] Y. Hu, Z. Sun, L. Cao, Y. Zhang, and P. Pan, "Optimization configuration of gas path sensors using a hybrid method based on Tabu search artificial bee colony and improved genetic algorithm in turbofan engine," *Aerosp. Sci. Technol.*, vol. 112, May 2021, Art. no. 106642.
[33] F. A. Hashim, E. H. Houssein, M. S. Mabrouk, W. Al-Atabany, and S. Mirjalili, "Henry gas solubility optimization: A novel physics-based algorithm," *Future Gener. Comput. Syst.*, vol. 101, pp. 646–667, Dec. 2019.
[34] N. Neggaz, E. H. Houssein, and K. Hussain, "An efficient Henry gas solubility optimization for feature selection," *Exp. Syst. Appl.*, vol. 152, Aug. 2020, Art. no. 113364.
[35] B. S. Yildiz, N. Pholdee, N. Panagant, S. Bureerat, A. R. Yildiz, and S. M. Sait, "A novel chaotic Henry gas solubility optimization algorithm for solving real-world engineering problems," *Eng. Comput.*, vol. 38, no. S2, pp. 871–883, Jun. 2022.
[36] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.
[37] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Gener. Comput. Syst.*, vol. 97, pp. 849–872, Aug. 2019.
[38] C. A. Hao, A. A. Heidari, H. Chen, M. Wang, Z. Pan, and A. H. Gandomi, "Multi-population differential evolution-assisted Harris hawks optimization: Framework and case studies—ScienceDirect," *Future Gener. Comput. Syst.*, vol. 111, pp. 175–198, Oct. 2020.
[39] H. M. Alabool, D. Alarabiat, L. Abualigah, and A. A. Heidari, "Harris hawks optimization: A comprehensive review of recent variants and applications," *Neural Comput. Appl.*, vol. 33, no. 15, pp. 8939–8980, Aug. 2021.
[40] Z. G. Wang *et al.*, "Gas path fault mode identification of turboshaft engine based on ReliefF-LMBP algorithm," *J. Propuls. Technol.*, vol. 42, no. 1, pp. 220–229, 2021.
[41] X. Wen and Z. Xu, "Wind turbine fault diagnosis based on ReliefF-PCA and DNN," *Exp. Syst. Appl.*, vol. 178, Sep. 2021, Art. no. 115016.
[42] Z. F. Wang, J. L. Zarader, and S. Argentieri, "A novel aircraft engine fault diagnostic and prognostic system based on SVM," in *Proc. IEEE Int. Conf. Condition Monit. Diagnosis*, Sep. 2012, pp. 723–728.
[43] Q. Xu and S. H. I. Jun, "Fault diagnosis for aero-engine applying a new multi-class support vector algorithm," *Chin. J. Aeronaut.*, vol. 19, no. 3, pp. 175–182, 2006.

HUIHUI LI received the B.S. degree from the School of Power and Energy, Northwestern Polytechnical University, Xi'an, China, where she is currently pursuing the Ph.D. degree. Her research interests include aeroengine control and fault diagnosis.

LINFENG GOU received the Ph.D. degree from Northwestern Polytechnical University, in 2010. He is currently a Professor of power control engineering and the Vice President of the Power and Energy College, Northwestern Polytechnical University. He has participated in many national key projects. His main research interests include fault diagnosis and fault tolerant control of aeroengine control systems.

YINGXUE CHEN received the Ph.D. degree from the Queen Mary University of London, in 2019. She is currently an Associate Professor of power control engineering. Her main research interests include wind and solar energy systems and engineering.

HUACONG LI received the Ph.D. degree from Northwestern Polytechnical University, in 2008. He is currently a Professor of power control engineering with Northwestern Polytechnical University. His main research interest includes aircraft power control.