

Article

Diagnosing Faults of Pneumatic Soft Actuators Based on Multimodal Spatiotemporal Features and Ensemble Learning

Tao Duan ^{1,2}, Yi Lv ¹, Liyuan Wang ², Haifan Li ¹, Teng Yi ¹, Yigang He ² and Zhongming Lv ^{2,*} 

¹ Hubei Technology Innovation Center for Smart Hydropower, Wuhan 430000, China; duantao@whu.edu.cn (T.D.); lv_yi@ctg.com.cn (Y.L.); li_haifan@ctg.com.cn (H.L.); yi_teng@ctg.com.cn (T.Y.)

² School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China; wangly1027@163.com (L.W.); yghe1221@whu.edu.cn (Y.H.)

* Correspondence: lvzhongming1989@126.com

Abstract

Soft robots demonstrate significant advantages in applications within complex environments due to their unique material properties and structural designs. However, they also face challenges in fault diagnosis, such as nonlinearity, time variability, and the difficulty of precise modeling. To address these issues, this paper proposes a fault diagnosis method based on multimodal spatiotemporal features and ensemble learning. First, a sliding-window Kalman filter is utilized to eliminate noise interference from multi-source signals, constructing separate temporal and spatial representation spaces. Subsequently, an adaptive weight strategy for feature fusion is applied to train a heterogeneous decision tree model, followed by a dynamic weighted voting mechanism based on confidence levels to obtain diagnostic results. This method optimizes the feature extraction and fusion process in stages, combined with a dynamic ensemble strategy. Experimental results indicate a significant improvement in diagnostic accuracy and model robustness, achieving precise identification of faults in soft robots.

Keywords: soft robotics; spatiotemporal features; ensemble learning; fault diagnosis; multimodal features



Academic Editor: Ahmed Abu-Siada

Received: 8 July 2025

Revised: 19 August 2025

Accepted: 20 August 2025

Published: 21 August 2025

Citation: Duan, T.; Lv, Y.; Wang, L.; Li, H.; Yi, T.; He, Y.; Lv, Z. Diagnosing Faults of Pneumatic Soft Actuators Based on Multimodal Spatiotemporal Features and Ensemble Learning. *Machines* **2025**, *13*, 749. <https://doi.org/10.3390/machines13080749>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As an emerging branch in the field of robotics, soft robotics is gradually reshaping the application landscape of traditional rigid robots through its unique material properties and structural designs. Unlike traditional rigid robots that rely on metal links and joints, soft robots are fabricated using highly compliant materials such as polydimethylsiloxane (PDMS), polyurethane (PU), and thermoplastic polyurethane (TPU). These materials combine elasticity, ease of processing, and low cost, endowing the robots with enhanced flexibility [1]. This inherent structural compliance grants soft robots infinite degrees of freedom, enabling them to undergo passive deformation in response to environmental constraints, thereby performing tasks effectively in complex and unstructured environments, which compensates for the limitations of traditional rigid robots [2]. Soft robots can be driven in various ways, including pneumatic pressure, cable-driven tendons, electroactive polymers (EAPs), and pneumatic actuation, among which pneumatic soft actuators generate corresponding mechanical movements through fluid pressure [3].

Soft robots have demonstrated broad application prospects in constrained environments such as medical intervention, industrial inspection, and power system operation and maintenance [2]. Compared with rigid robots, soft robots lack rigid structures and exhibit

limited stiffness, making their postures under external loads difficult to predict accurately, which further increases the complexity of fault diagnosis [4]. In terms of actuation, both McKibben pneumatic artificial muscles and fluid-driven actuators face challenges such as sealing failure and performance degradation caused by material fatigue, while the highly nonlinear buckling characteristics of single elastic actuators significantly aggravate the difficulty of fault diagnosis [5]. In addition, by integrating multimodal sensing systems, such as ultrasonic sensors and flexible triboelectric sensors, soft robots are capable of autonomous object localization and multidimensional information acquisition [6]. However, the intrinsic body structures composed of elastic materials render them vulnerable to diverse fault types, which typically exhibit nonlinear, time-varying, and hard-to-model characteristics [7]. From the perspective of physical damage, although the self-sealing ability of elastomer materials provides resilience by enabling continuous operation after puncture, it simultaneously makes damage features difficult to capture through conventional detection methods [8]. Furthermore, the unique design requirements of soft grippers for deformation forces not only extend their ability to manipulate fragile objects but also fundamentally differentiate their gripping fault mechanisms from those of traditional rigid structures [9]. These compounded challenges, arising from the constitutive properties of materials, flexible structural configurations, and bio-inspired actuation mechanisms, fundamentally restrict the applicability of existing diagnostic methods in soft robotics. Moreover, vibration signals acquired from soft robots are often non-stationary and nonlinear, and are subject to the influence of complex operating conditions and electromagnetic environments [10,11]. Therefore, there is an urgent need to develop novel diagnostic approaches that integrate multiphysics coupling analysis with intelligent feature extraction.

At present, robotic fault diagnosis methods can be broadly categorized into model-based approaches and data-driven approaches. Model-based methods rely on precise mathematical descriptions of the system's dynamic behavior and typically implement fault detection by constructing state observers or residual generators. In [12], a multi-model-based fault detection and diagnosis (FDD) method for internal sensors of mobile robots is proposed. Each filter is associated with a model tailored to a specific sensor fault mode, and fault decisions are made by comparing condition estimates of sensor gains across different models. Reference [13] presents a fault detection method based on power consumption monitoring in industrial robots. This method employs an accurate mathematical model derived from reference power patterns to monitor system performance. The proposed energy-based diagnostic technique can be easily integrated into existing industrial robot power systems and allows for remote monitoring. These approaches have demonstrated significant success in rigid robots with deterministic models. However, when applied to soft robots, challenges arise due to the highly nonlinear constitutive behavior of hyperelastic materials and the complex coupling of multi-chamber deformations. These characteristics hinder the formulation of accurate mathematical models, leading to model mismatches and increased false-alarm rates. In contrast, data-driven methods extract fault features directly from monitoring data through machine learning techniques, thereby avoiding complex modeling processes. Supported by advancements in support vector machines [14], fuzzy systems [15,16], and artificial neural networks [17,18], data-driven approaches have emerged as a dominant trend in intelligent diagnosis research. To improve the practicality of deep-learning-based fault diagnosis methods for proton exchange membrane fuel cells (PEMFCs), reference [19] proposed a deep-learning-based PEMFC fault diagnosis framework combined with a frequency selection method to reduce the influence of non-fault operations on diagnosis results. In reference [20], neural networks are employed to identify unknown nonlinear characteristics in the system, construct an observer model and a general fault model to estimate unavailable states, and describe faults. These methods are

particularly well-suited for soft robotic systems, which are characterized by nonlinearity, disturbances, and time-varying behaviors, and they exhibit strong potential for enhancing diagnostic accuracy and system robustness.

However, current studies predominantly focus on mining single-modal data, which limits the comprehensive characterization of the multidimensional information evolution during the operation of soft robots. Typically, soft robots are equipped with various flexible sensors, such as pressure, temperature, and acceleration sensors. The data generated by these sensors exhibit significant modality differences and are often contaminated by noise. Conventional fusion strategies—such as feature concatenation or simple averaging—fail to dynamically adjust fusion weights according to the importance of each modality, which may lead to information redundancy or the suppression of critical features [21,22]. Moreover, at the model level, most diagnostic approaches rely on a single classifier for decision making. This makes it difficult to balance model expressiveness and generalization capability. In non-ideal scenarios such as working condition variations or changes in fault types, these models often suffer from reduced recognition accuracy and lack of stability [23]. In addition, the data generated during soft robot operation present evident spatiotemporal coupling characteristics. Fault evolution is not only influenced by the current system state but also constrained by historical trajectories and spatial distributions. Therefore, it is essential to develop diagnostic frameworks that possess both temporal modeling capabilities and spatial perception mechanisms in order to enable in-depth analysis of state evolution patterns and accurate identification of fault trends.

With the continuous advancement of cutting-edge technologies such as multimodal data fusion, intelligent perception, and spatiotemporal modeling, significant progress has been made in the state monitoring and fault diagnosis of soft robots under complex working conditions. Against this backdrop, this study addresses the common diagnostic challenges encountered during the operation of soft robots, including strong nonlinear responses, multimodal coupling, and dynamic uncertainties. A novel fault diagnosis method is proposed, which integrates spatiotemporal feature modeling and dynamic ensemble discrimination. First, a sliding-window Kalman filtering approach is introduced to fundamentally suppress interference and eliminate noise. Then, by incorporating both temporal and spatial characteristics, a spatiotemporal-aware diagnostic method is developed. This method establishes a multi-stage-optimized intelligent diagnosis framework, which focuses on preprocessing and noise reduction of multi-source signals, modeling intermodal correlations, decoupling and fusing critical spatiotemporal features, and implementing ensemble classification strategies for multi-condition scenarios. The objective is to enhance the accuracy of state perception and the robustness of fault recognition in soft robots operating in complex environments. The proposed multi-source heterogeneous signal fusion method enables deep mining of critical fault information by decoupling spatiotemporal features. Meanwhile, the dynamic weighting mechanism in ensemble learning demonstrates excellent robustness and adaptability when dealing with time-varying system states and increased uncertainties, thereby providing both a theoretical foundation and a practical pathway for the integrated development of intelligent perception and adaptive diagnosis technologies.

2. Overview of Soft Robot Fault Diagnosis

Diagnostic System Framework

This paper proposes a general architecture for fault diagnosis methods incorporating temporal-spatial and local-global feature fusion, as shown in Figure 1. To address the issue of reduced classification accuracy caused by severe noise interference in fault data that drowns out fault features, a Kalman filtering model with a sliding window is employed to filter the collected data. The vibration data of the soft robot are considered as spatial

feature local information, while information such as torque and current signals is treated as temporal feature local information. In the spatial feature extraction process, variational mode decomposition (VMD) is applied to obtain several intrinsic mode function (IMF) components with single frequencies, and multiscale permutation entropy (MPE) is utilized to quantify the nonlinear dynamic characteristics of each IMF component, thereby capturing multiscale information. In the temporal feature extraction process, weights are dynamically allocated based on an attention mechanism, followed by the application of a reverse cloud model to calculate the expectation, entropy, and hyper-entropy, quantifying the randomness and fuzziness of multimodal data. Through adaptive weight strategy-based feature fusion, a heterogeneous decision tree model is trained, and a dynamic weighted voting mechanism based on confidence is employed to obtain the diagnostic results.

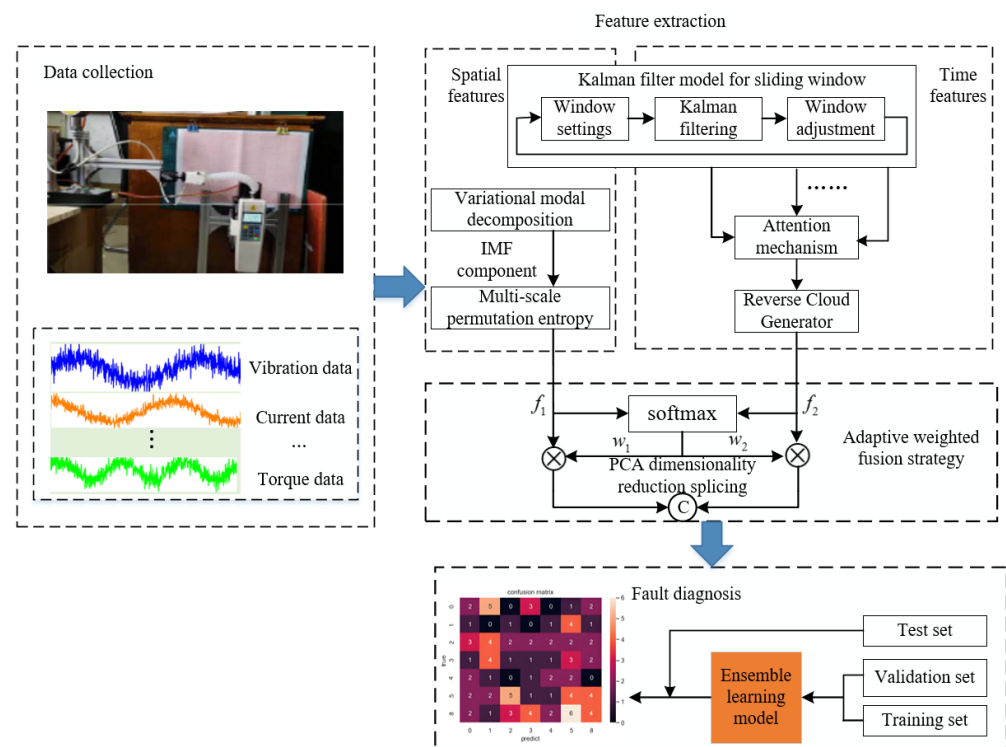


Figure 1. Overall architecture of the fault diagnosis method for pneumatic soft actuators.

3. Methodology

Robot faults mainly occur in sensors and actuators; however, the lack of complete and publicly available fault datasets limits the accuracy of deep-learning-based fault diagnosis. Traditional data augmentation techniques are primarily designed for image data and are not directly applicable to one-dimensional sensor signals [22]. To address the issue of limited fault samples, this study proposes a sliding-sampling-based dataset augmentation method to expand the fault data for soft robot sensors and actuators, thereby improving the accuracy of fault classification.

During the fault data acquisition phase, vibration and inductance signals are collected from each joint of the soft robot. In the data preprocessing stage, the sliding window method segments the data along the time dimension. The sliding window method is a commonly used data augmentation technique in fault diagnosis and can significantly increase the number of training samples, as illustrated in Figure 2. A time window of length w is applied to segment the data, with a step size of S , and the total number of data points

is T . The formula for calculating the number of training samples N is given in Equation (1), with the result rounded to the nearest integer.

$$N = \frac{T - W}{S} + 1 \quad (1)$$

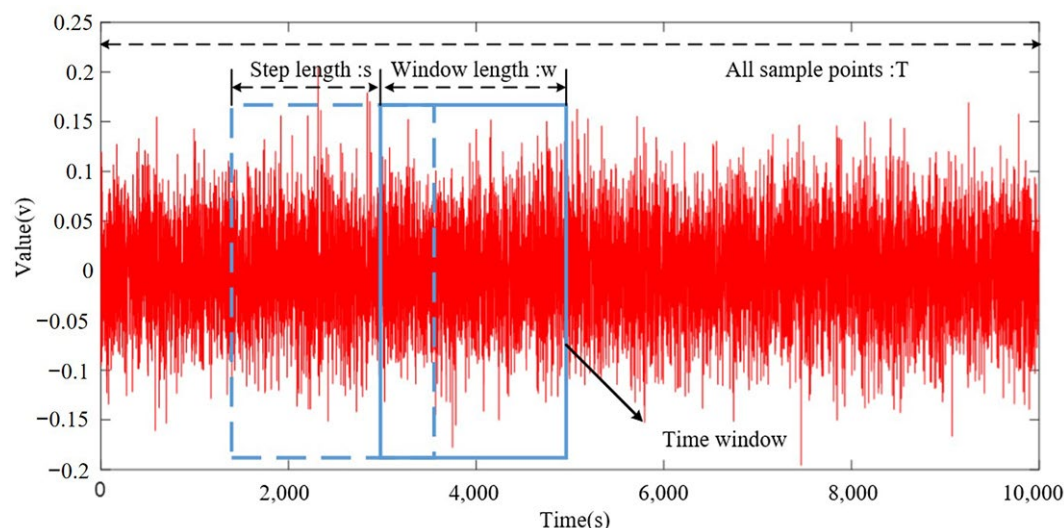


Figure 2. The method of sliding the window.

3.1. The VMD Method

VMD is an adaptive signal processing method based on variational principles, aiming to decompose a non-stationary signal $f(t)$ into several modal components $u_k(t)$ with limited bandwidth and tunable center frequencies, such that these modes are additive in the time domain and exhibit good localization in the frequency domain [24]. VMD achieves this decomposition by solving the following variational problem:

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[u_k(t) \cdot e^{-j\omega_k t} \right] \right\|_2^2 \right\} \text{ s.t. } \sum_k u_k(t) = f(t) \quad (2)$$

Here, $u_k(t) \cdot e^{-j\omega_k t}$ denotes the complex analytic signal of the k -th mode demodulated to the baseband, and the second norm of its derivative, $\| \cdot \|_2$ is used to measure its frequency bandwidth. This optimization objective indicates that minimizing the frequency bandwidth of each mode under the constraint that the sum of all modes equals the original signal. To facilitate the solution, Lagrange multipliers $\lambda(t)$ and penalty factors α are introduced, transforming the constrained optimization problem into an unconstrained augmented Lagrangian function:

$$\mathcal{L}(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[u_k(t) \cdot e^{-j\omega_k t} \right] \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \quad (3)$$

where VMD is solved using the Alternating Direction Method of Multipliers (ADMM), where, in each iteration, the mode u_k , its corresponding center frequency ω_k , and the Lagrange multiplier λ are updated alternately until the convergence condition is satisfied.

3.2. Multiscale Permutation Entropy Algorithm

The MPE algorithm is an optimized multiscale permutation entropy algorithm [25]. The principle of the multiscale permutation entropy algorithm is as follows:

- (1) Coarse-grain the time series $\{X(1), X(2), \dots, X(n)\}$ to obtain a new vector:

$$Y_n^\tau = \frac{1}{\tau} \sum_{n=(j-1)\pi}^{j\pi} X(n), \quad 1 \leq j \leq \frac{N}{\tau} \quad (4)$$

where N and τ are integers; τ represents the scale factor and the time-series scale.

- (2) Reconstruct the subsequences into m -dimensional time series:

$$Z(t) = [Y(t), \dots, Y(\tau + (m-1)t)] \quad (5)$$

in the equation, t represents the delay time, and m is the embedding dimension.

- (3) Arrange $Z(t)$ and calculate the probability:

$$P(i) = \frac{K}{N/\tau - m + 1} \quad (6)$$

where K represents the sum of the frequencies of occurrence of each distinct time series.

- (4) Therefore, the MPE calculation formula for time series signals is

$$H_\tau(P) = - \sum_{i=1}^{m!} P_\tau(i) \ln(P_\tau(i)) \quad (7)$$

3.3. Cloud Model Theory

The cloud model reveals the uncertainty of any event through two transformation models: a forward cloud generator that obtains quantitative values from qualitative concepts and a backward cloud generator that transforms quantitative values into qualitative concept descriptions [26]. Generally, the cloud model is defined as follows: Let the set U be the universe of discourse of variable X , C be a qualitative concept on U , and sample x be a random realization of the qualitative concept C . Then, the certainty degree of x with respect to C is a random number with a stable tendency, which can also be represented by Equation (8).

$$\mu : U \rightarrow [0, 1], \forall x \in U, x \rightarrow \mu(x) \quad (8)$$

The cloud model generates three numerical characteristics—expected value (Ex), entropy (En), and hyper-entropy (He)—to represent the overall properties of the cloud model concept, effectively integrating and characterizing the randomness and fuzziness of qualitative concepts [27]. Ex measures the certainty of a qualitative concept and represents the ideal value of the qualitative concept. En measures the uncertainty of a qualitative concept, reflecting the dispersion of cloud droplets and determining the certainty degree of acceptable cloud droplets. He quantifies the uncertainty of En ; a smaller He indicates a higher degree of acceptance of the concept, whereas a larger He suggests greater difficulty in reaching a consensus on the concept.

The forward normal cloud generation algorithm utilizes the certainty expression $\mu_A(x) = \exp\left\{-\frac{(x-Ex)^2}{2(En')^2}\right\}$ based on the three numerical characteristics of qualitative concepts, generating N x_i values with certainty μ_i , which constitute N cloud drops of the cloud.

The reverse normal cloud generation algorithm operates inversely to the forward normal generation algorithm, transforming N given cloud drops (quantitative dataset x_i) into digital characteristics (Ex, En, He) representing qualitative concepts. Using only the

quantitative values of the cloud drops x_i to recover the three parameters of the cloud, the reverse normal cloud generation algorithm is expressed as follows:

$$Ex = \frac{1}{N} \sum_{i=1}^N x_i \quad (9)$$

$$En = \sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |x_i - Ex| \quad (10)$$

$$He = \sqrt{\frac{1}{N-1} \sum_{i=1}^n (x_i - Ex)^2 - \left(\sqrt{\frac{\pi}{2}} \times \frac{1}{N} \sum_{i=1}^N |x_i - Ex| \right)^2} \quad (11)$$

3.4. Ensemble Learning Model

In the field of machine learning, ensemble learning is a technique that improves overall performance by combining the prediction results of multiple models. Its core idea is to construct a strong learner by integrating multiple weak learners, thereby enhancing both the accuracy and robustness of predictions [28]. Bagging is a commonly used ensemble learning method that generates multiple training sets through bootstrap sampling, trains several models independently, and then combines their outputs via voting or averaging to obtain the final result. Random forest is a classical implementation of bagging. It constructs multiple decision trees, each trained with randomly selected features, which helps to reduce the risk of overfitting [29]. Moreover, random forest demonstrates good generalization performance with small-scale training datasets, shows tolerance to missing data, and exhibits strong noise resistance. The algorithmic flow of random forest (RF) is illustrated in Figure 3.

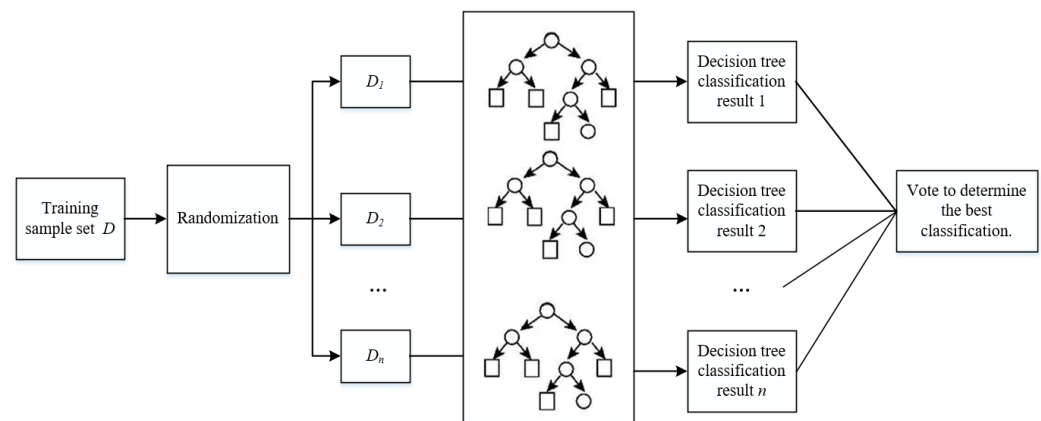


Figure 3. Flowchart of the random forest algorithm.

3.5. Bayesian-Optimization-Based Random Forest Model

Ensemble algorithms typically involve a large number of hyperparameters, and their diagnostic performance is highly dependent on the selection of an appropriate parameter set. Bayesian optimization (BO) offers an efficient solution to this challenge. It primarily consists of two components: a probabilistic surrogate model and an acquisition function (AC). The core idea is to fully leverage prior knowledge by fitting a probabilistic model to the observed values of the black-box objective function, thereby obtaining the posterior distribution of the target function [30]. In terms of parameter optimization strategy, Bayesian optimization employs a Gaussian process (GP)-based intelligent update mechanism. During each iteration, it integrates empirical knowledge from previously evaluated parameter configurations. Bayesian optimization is shown in Algorithm 1. The parameter tuning pro-

cess of the random forest (RF) model using Bayesian optimization is illustrated in Figure 4. The major advantage of this method lies in its relatively low number of iterations and high computational efficiency. This becomes particularly important when the parameter space is large, in which case traditional grid search methods become significantly less efficient and time consuming.

Algorithm 1: Bayesian optimization.

1. For datasets with $n = 1, 2, 3, \dots$
 2. A new function x_{n+1} is obtained by optimizing the acquisition function α .

$$x_{n+1} = \arg \max_x \alpha(x; F_n)$$
 3. The objective function is queried to obtain y_{n+1} .
 4. The dataset is updated as $F_{n+1} = \{F_n, (x_{n+1}, y_{n+1})\}$.
 5. Update the Gaussian probabilistic model.
 6. Iteratively update and select the next point until the stopping criterion is met.
-

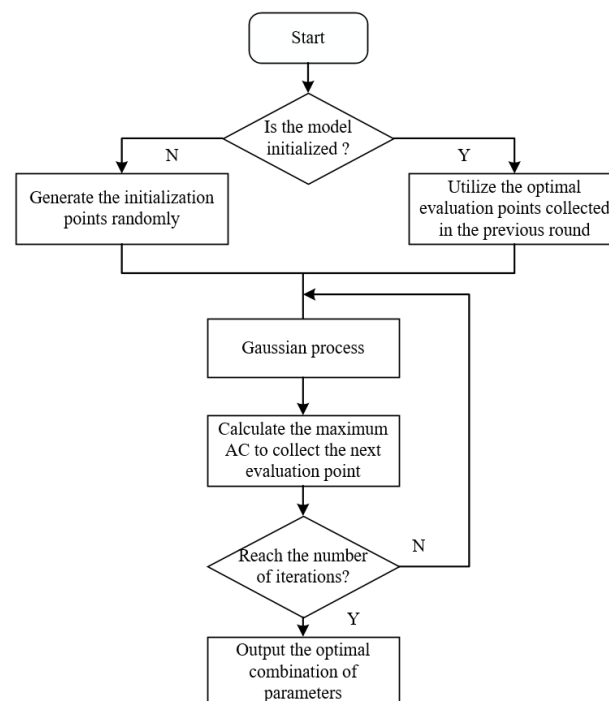


Figure 4. Parameter optimization process of the random forest model.

4. Results and Discussion

4.1. Experimental Platform Construction

The soft actuator in this study adopts a coupled rigid–soft structural form, where metal coils are wound around the outer contour of the soft actuator to form a reinforcing structure (as shown in the enlarged small picture in Figure 5). Additionally, slight variations in the distance and alignment between ring-shaped metal coils with different electromagnetic properties can lead to significant changes in mutual inductance. The soft actuator has an asymmetric corrugated structure in the vertical direction and features multiple joints, enabling the distribution of driving force across the entire actuator and generating a torque-driven bending motion around the center through multi-joint actuation.

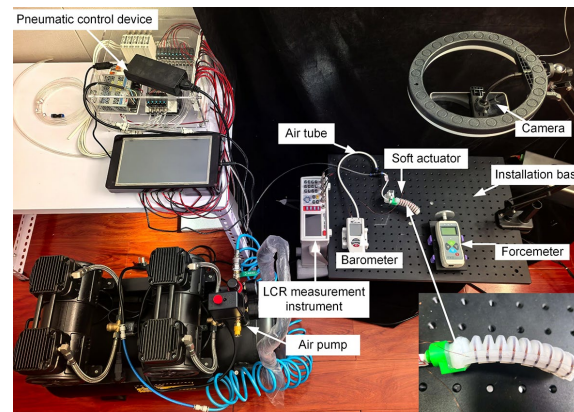


Figure 5. Experimental platform.

The experiments conducted on the soft actuator primarily involved the inductance signal of the actuator itself and the output vibration signal. Both experiments required measuring the deformation (inductance signal) and output force data (vibration signal) of the soft actuator under varying levels of positive pressure. In this study, the deformation data were measured using an inductive self-sensing sensor. The inductance of the sensor was calibrated using an LCR meter (Victor 4090A, Shenzhen Yisheng Victory Technology Co., Ltd., Shenzhen, China). This device operates by observing the impedance between low and high frequencies. The acquisition of vibration signals corresponds to the vibration output force signal of the soft actuator; the output force data were obtained by measuring the vibration signal using a force sensor (HP-500, Yueqing Handpi Instruments Co., Ltd., Wenzhou, China). Therefore, the pneumatic system must have the capability to regulate and maintain positive pressure. During the positive pressure actuation experiments, the electro-pneumatic proportional valve adjusted the output air pressure according to the magnitude of the analog voltage signal. To eliminate the influence of gravity on the movement of the soft actuator, the middle axial plane of the actuator was set to remain horizontal with the ground, and its motion was parallel to the horizontal plane. The control system adjusted the pressure applied by the proportional valve in increments of 1 kPa. The final assembled testing platform is shown in Figure 5, where the mounting base could be adjusted according to different testing requirements. In this experiment, three types of single faults in the soft robot were simulated through manual fault injection: blockage at the front-end structure (F1), leakage at the root-bonding area (F2), and rupture-induced leakage in the middle bellows structure (F3). Additionally, the normal mode (F0) was included, achieving a total simulation of four different data types, as shown in Figure 6.

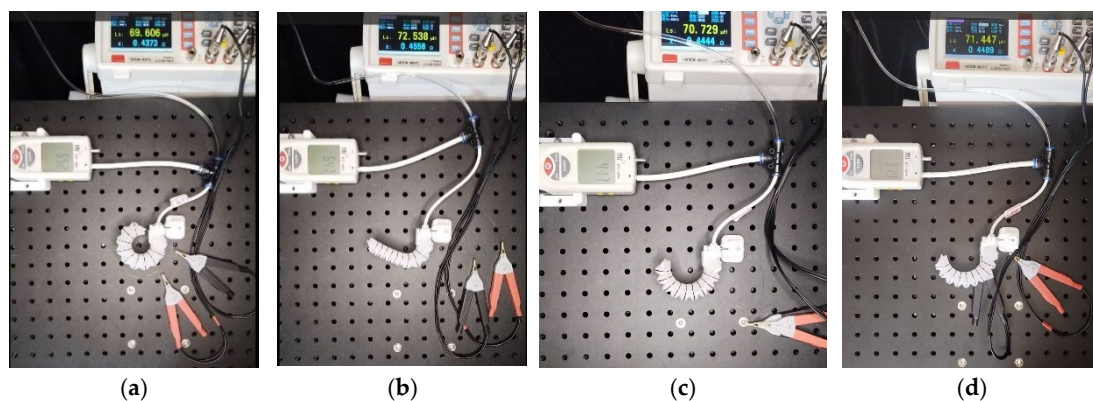


Figure 6. Fault mode. (a) Normal group, (b) blockage at the front-end structure, (c) leakage at the root-bonding area, (d) rupture-induced leakage in the middle bellows structure.

4.2. Experimental Results

Vibration signals from each joint were obtained through vibration sensors, while inductance and vibration signals were acquired through the robot controller. The sampling frequency was set to 512 Hz, and a sliding window approach was employed with a window length of 500 and a step size of 0.8 s. As a result, a total of 1041 samples were obtained per mode, with each sample containing 2500 data points. Each sample included signals from two channels: vibration and inductance values under bending angles. Finally, all data were divided into three parts: 70% for the training dataset, 20% for the validation dataset, and 10% for the testing dataset.

Since the differences between vibration output signals corresponding to various faults in soft robots are not obvious, the IMF components generated by VMD decomposition can amplify these signal differences. However, when extracting features from fault signals, careful consideration must be given to the selection of feature parameters, and the MPE values of the IMF components are calculated as a basis for distinguishing different fault types.

Two important parameters of the VMD algorithm are set as $\alpha = 2000$ and $K = 8$. In the process of constructing the feature vector, computing the MPE values of all IMF components would significantly affect the training speed; therefore, only the MPE values of the first three IMF components are considered as feature values. The scale factor τ and embedding dimension m have a significant impact on the MPE values. When the scale factor τ was 5 and 10, the MPE values of the F0 state signal were calculated and are plotted in Figure 5. In this paper, the scale factor is set to five. Figure 7 shows the permutation entropy values of normal samples with different scale factors $\tau = 5$. Experimental results show that the best diagnostic performance is achieved when the embedding dimension is three. For each IMF component, a 5×1 entropy value vector can be obtained, and part of the fault feature vector data is shown in Table 1.

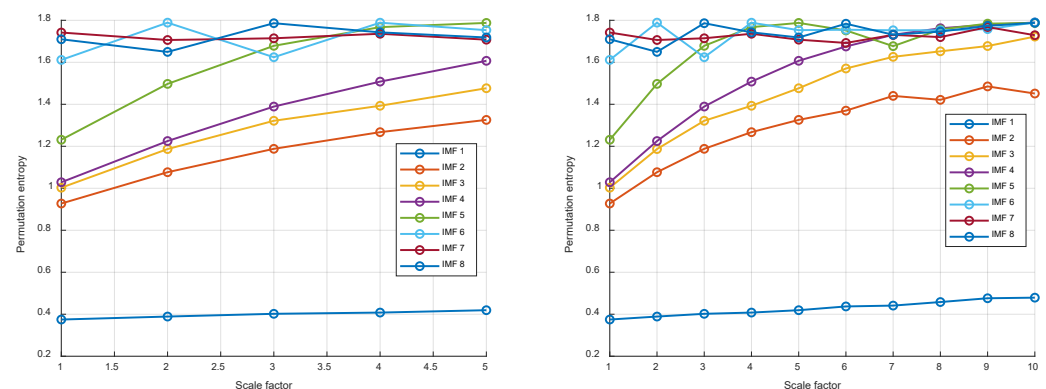


Figure 7. MPE curves for normal samples corresponding to different scale factors.

Table 1. Partial data of fault feature vectors based on VMD-MPE.

State	IMF1			IMF2			IMF3		
	MPE1	MPE2	MPE3	MPE1	MPE2	MPE3	MPE1	MPE2	MPE3
F0	0.3748	0.3890	0.4020	0.9273	1.0765	1.1878	1.0027	1.1867	1.3214
F1	0.4078	0.4222	0.4377	0.8933	1.0299	1.1336	1.0883	1.3340	1.504
F2	0.5721	0.5870	0.6009	0.8283	0.9142	0.9722	0.9510	1.1138	1.2422
F3	0.3748	0.3874	0.4020	0.8764	0.9723	1.0487	0.9883	1.1164	1.2424

The deformation (inductance signal) data of the soft actuator are input into the inverse cloud generator to obtain the digital characteristics of the cloud under each fault mode. Subsequently, the forward cloud generator is utilized to generate cloud droplets. In Figure 8,

the cloud models and sample distributions of 10,000 cloud droplets produced by the normal mode F0 and the fault mode F1 are presented. Through comparison, it can be observed that the degree of distinguishability from the cloud model analysis, from high to low, is expectation, certainty, and randomness. The range of accepted cloud droplets in the domain space and the degree of condensation of cloud droplets vary significantly for each fault mode, and clear differences exist between the normal state and the fault states.

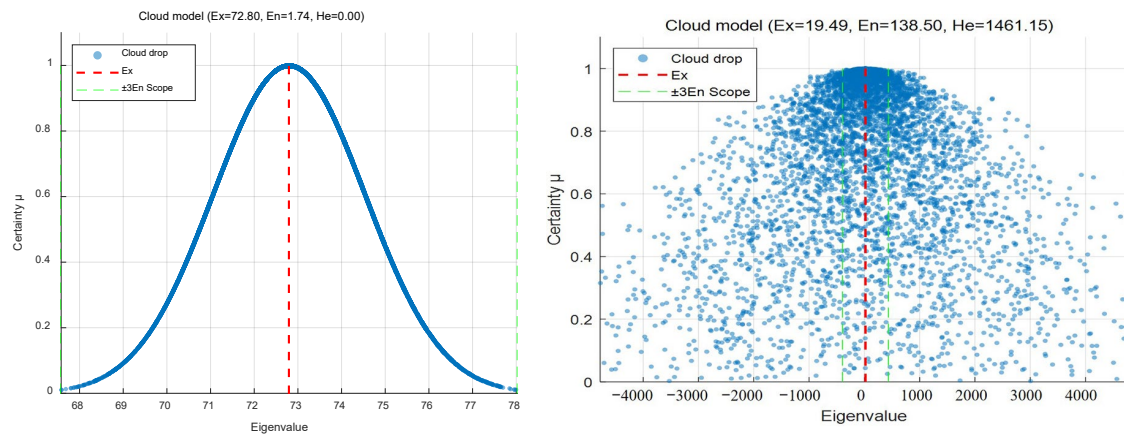


Figure 8. Cloud model diagrams under normal mode F0 and fault mode F1.

The random forest (RF) model involves four primary hyperparameters, whose Bayesian optimization (BO) search ranges are provided in Table 1. Among these parameters, the number of decision trees (N_{est}) represents the model's learning capacity. A small N_{est} value may lead to underfitting due to insufficient learning, while an excessively large value significantly increases computational cost and only marginally improves performance beyond a certain threshold, which may result in overfitting. The remaining three parameters—maximum tree depth (D_{tre}), minimum samples required to split a node (T_{spit}), and minimum samples required at a leaf node (T_{leaf})—govern the strategy for utilizing training samples. For large-scale datasets, the maximum tree depth D_{tre} should be limited to prevent overly complex trees. A node will stop splitting when the number of samples falls below T_{spit} , and pruning is performed when a leaf node contains fewer samples than T_{leaf} . The search space for RF hyperparameters optimized via Bayesian optimization is detailed in Table 2.

Table 2. Search space of RF hyperparameters for Bayesian optimization.

Description	Hyperparameter	Search Range
Number of decision trees	N_{est}	[10, 100]
Maximum depth of each tree	D_{tre}	[10, 200]
Minimum number of samples to split a node	T_{spit}	[2, 50]
Minimum number of samples at a leaf node	T_{leaf}	[1, 20]

Principal component analysis (PCA) was employed to align the dimensionality of the extracted spatial and temporal features, which were then concatenated to form a unified feature set. The fused features were randomly split into training and testing sets according to a fixed proportion. A Bayesian optimization (BO) strategy was adopted to jointly tune four key hyperparameters of the random forest (RF) model, with five-fold cross-validation error used as the objective function to be minimized. The optimization was performed over 30 iterations and took approximately 43.64 s. The optimal hyperparameter configuration obtained through the BO process was $N_{est} = 62$, $D_{tre} = 54$, $T_{spit} = 32$, and $T_{leaf} = 8$. The RF model was retrained using these optimized parameters, achieving 100% classification

accuracy on both the training and testing sets. The training results of the BO-RF model are illustrated in Figure 9, demonstrating both excellent fitting ability and strong generalization performance. Specifically, Figure 9a presents a 3D visualization of the objective function model during the BO process, clearly illustrating the joint influence of hyperparameter combinations on model performance. Figure 9b shows the feature importance bar chart, indicating the relative contribution of each feature to model prediction. Figure 9c,d depict the classification results for the training and testing sets, respectively. In this classification scheme, label 0 denotes no fault, label 1 indicates front-end leakage, label 2 refers to front-end blockage, and label 3 represents root leakage. To further evaluate the performance of the proposed approach, a comparative analysis was conducted with k-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Fully Connected Neural Network (FCNN) models. The average diagnostic accuracy and Kappa coefficients of all methods are summarized in Table 3. The results demonstrate that the BO-RF model achieves superior performance in terms of both accuracy and consistency.

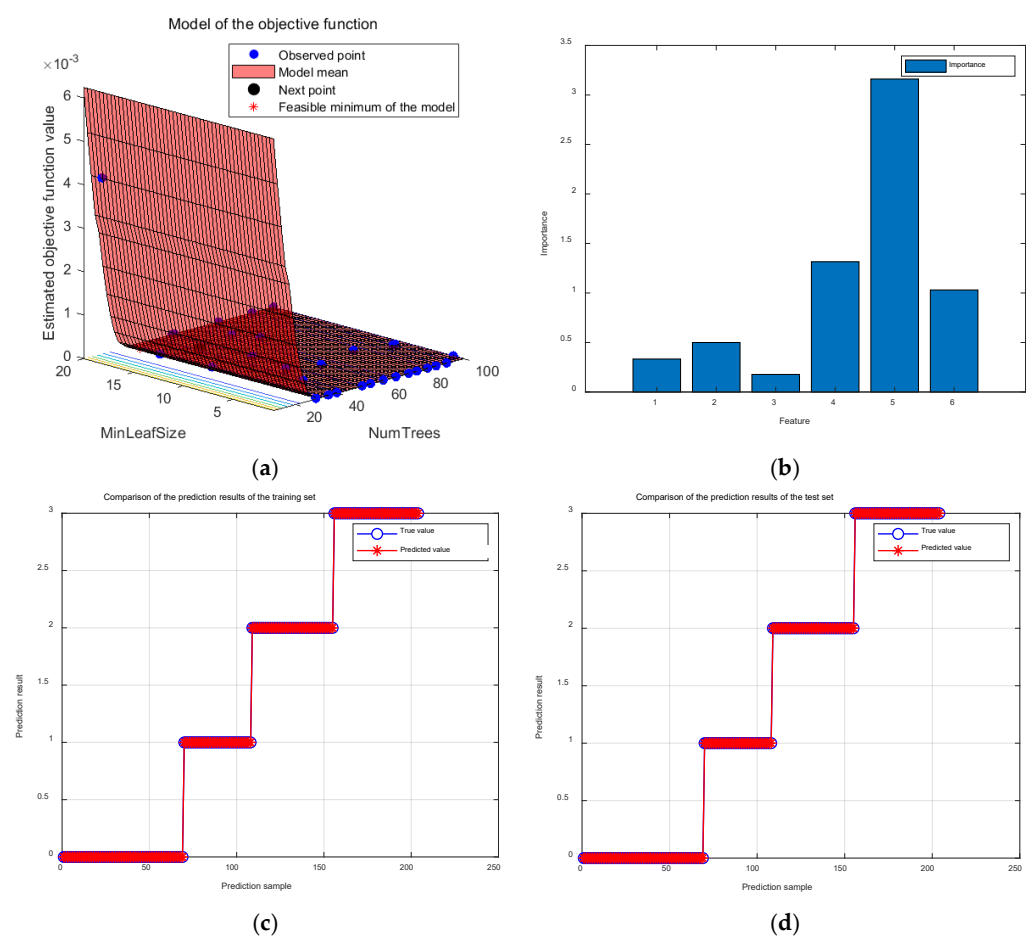


Figure 9. Training results of the BO-RF model: (a) 3D visualization of the objective function model; (b) bar chart of feature importance; (c) prediction results on the training set; (d) prediction results on the testing set.

Table 3. Comparative analysis of diagnostic model results.

Diagnostic Algorithm	Test Accuracy (%)	Kappa Coefficient (Test Set)
KNN	73.00	0.637
SVM	97.30	0.964
Neural Network	93.90	0.919
This work	99.10	0.986

5. Conclusions

This study investigates the fault classification problem of soft robots and conducts experimental validation on a soft actuator platform. A novel fault classification method incorporating spatiotemporal feature information is proposed. First, to address the issue of noisy fault data that often mask critical fault information and result in low classification accuracy, a sliding-window-based filtering model is developed. Then, considering the sequential nature of soft robot data and the semantic differences among sensor modalities, a fault classification framework that separates temporal and spatial features is designed. Multiscale permutation entropy and numerical features derived from the cloud model are extracted, dimensionally reduced, fused, and then fed into a Bayesian optimized random forest (BO-RF) algorithm for classification. The proposed method achieves an average diagnostic accuracy of 99.10% on the test set. However, the Bayesian optimization process typically requires numerous iterations, which may lead to computational and time bottlenecks with large-scale datasets or high-dimensional feature spaces. Additionally, although random forests are relatively interpretable, the integration of Bayesian optimization increases model complexity and further compromises interpretability. Future work will expand the range of working conditions and explore lightweight feature selection strategies and efficient optimization mechanisms in multi-task and multi-fault scenarios. These efforts aim to reduce the model's dependency on computational resources, improve the real-time performance and deployment efficiency of the diagnostic system, and ultimately provide more reliable and comprehensive technical support for the intelligent diagnosis of soft robots in real-world engineering applications.

The fault diagnosis method proposed in this study, based on multimodal spatiotemporal features and ensemble learning, demonstrates promising prospects for engineering applications and cross-industry promotion. The core technical framework of this method can be widely applied in equipment health monitoring in fields such as intelligent manufacturing, aerospace, and automotive industries, enabling early fault warning and precise diagnosis through real-time integration of multidimensional sensor data. With the continued advancement of Industry 4.0 and intelligent manufacturing, this technology is expected to significantly improve equipment maintenance efficiency, reduce maintenance costs, and provide crucial technical support for building more intelligent and reliable industrial systems, thereby holding substantial theoretical significance and practical value.

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

Funding: This work was supported by the Open Research Fund of Hubei Technology Innovation Center for Smart Hydropower (Grant No. 1523020038), and supported by the China Postdoctoral Science Foundation (Grant No. 2024M762476).

Data Availability Statement: The data can be provided by the corresponding author.

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

References

1. Ahn, S.; Jung, W.; Ko, K.; Lee, Y.; Lee, C.; Hwang, Y. Thermopneumatic soft micro bellows actuator for standalone operation. *Micromachines* **2021**, *12*, 46. [[CrossRef](#)]
2. Chen, Y.; Sun, Q.; Guo, Q.; Gong, Y. Dynamic modeling and experimental validation of a water hydraulic soft manipulator based on an improved newton—Euler iterative method. *Micromachines* **2022**, *13*, 130. [[CrossRef](#)]

3. Zhang, Y.; Wang, T.; He, W.; Zhu, S. Human-powered master controllers for reconfigurable fluidic soft robots. *Soft Robot.* **2023**, *10*, 1126–1136. [\[CrossRef\]](#)
4. Garriga-Casanovas, A.; Treratanakulchai, S.; Franco, E.; Zari, E.; Ferrandy, V.; Viridyawan, V.; y Baena, F.R. Optimised design and performance comparison of soft robotic manipulators. In Proceedings of the 2022 7th International Conference on Mechanical Engineering and Robotics Research (ICMERR), Krakow, Poland, 9–11 December 2022.
5. Lin, Y.; Xu, Y.X.; Juang, J.Y. Single-actuator soft robot for in-pipe crawling. *Soft Robot.* **2023**, *10*, 174–186. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Shi, Q.; Sun, Z.; Le, X.; Xie, J.; Lee, C. Soft robotic perception system with ultrasonic auto-positioning and multimodal sensory intelligence. *ACS Nano* **2023**, *17*, 4985–4998. [\[CrossRef\]](#)
7. Shiva, A.; Sadati, S.H.; Noh, Y.; Fraś, J.; Ataka, A.; Würdemann, H.; Hauser, H.; Walker, I.D.; Nanayakkara, T.; Althoefer, K. Elasticity versus hyperelasticity considerations in quasistatic modeling of a soft finger-like robotic appendage for real-time position and force estimation. *Soft Robot.* **2019**, *6*, 228–249. [\[CrossRef\]](#)
8. Shepherd, R.F.; Stokes, A.A.; Nunes, R.M.D.; Whitesides, G.M. Soft machines that are resistant to puncture and that self seal. *Adv. Mater.* **2013**, *25*, 6709–6713. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Bialek, M.; Rybarczyk, D. Research on the operational properties of the soft gripper pads. *Sci. Rep.* **2024**, *14*, 32133. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Hanada, E.; Takano, K.; Antoku, Y.; Matsumura, K.; Watanabe, Y.; Nose, Y. A practical procedure to prevent electromagnetic interference with electronic medical equipment. *J. Med. Syst.* **2002**, *26*, 61–65. [\[CrossRef\]](#)
11. Liu, A.; Yang, Z.; Li, H.; Wang, C.; Liu, X. Intelligent diagnosis of rolling element bearing based on refined composite multiscale reverse dispersion entropy and random forest. *Sensors* **2022**, *22*, 2046. [\[CrossRef\]](#)
12. Hashimoto, M.; Kawashima, H.; Oba, F. A multi-model based fault detection and diagnosis of internal sensors for mobile robot. In Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003), Las Vegas, NV, USA, 27–31 October 2003; Volume 4, pp. 3787–3792.
13. Nordin, F.H.; Sabry, A.H.; Ab Kadir, M.Z.A. Fault detection and diagnosis of industrial robot based on power consumption modeling. *IEEE Trans. Ind. Electron.* **2019**, *67*, 7929–7940. [\[CrossRef\]](#)
14. Deng, F.; Guo, S.; Zhou, R.; Chen, J. Sensor multifault diagnosis with improved support vector machines. *IEEE Trans. Autom. Sci. Eng.* **2015**, *14*, 1053–1063. [\[CrossRef\]](#)
15. Sneider, H.; Frank, P.M. Observer-based supervision and fault detection in robots using nonlinear and fuzzy logic residual evaluation. *IEEE Trans. Control Syst. Technol.* **2002**, *4*, 274–282. [\[CrossRef\]](#)
16. Sun, X.; Jia, X. A fault diagnosis method of industrial robot rolling bearing based on data driven and random intuitive fuzzy decision. *IEEE Access* **2019**, *7*, 148764–148770. [\[CrossRef\]](#)
17. Pan, J.; Qu, L.; Peng, K. Deep residual neural-network-based robot joint fault diagnosis method. *Sci. Rep.* **2022**, *12*, 17158. [\[CrossRef\]](#)
18. Skoundrianos, E.N.; Tzafestas, S.G. Finding fault-fault diagnosis on the wheels of a mobile robot using local model neural networks. *IEEE Robot. Autom. Mag.* **2004**, *11*, 83–90. [\[CrossRef\]](#)
19. Lv, J.; Yu, Z.; Sun, G.; Liu, J. Deep learning-based fault diagnosis and Electrochemical Impedance Spectroscopy frequency selection method for Proton Exchange Membrane Fuel Cell. *J. Power Sources* **2024**, *591*, 233815. [\[CrossRef\]](#)
20. Wu, C.; Liu, J.; Xiong, Y.; Wu, L. Observer-based adaptive fault-tolerant tracking control of nonlinear nonstrict-feedback systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *29*, 3022–3033. [\[CrossRef\]](#)
21. Wang, Z.; Liang, H.; Chen, C.; Wang, T.; Chen, Z. Graph Fusion and Propagation for Fault Diagnosis in Industrial Robots with Limited Labeled Data. *IEEE Sens. J.* **2025**, *25*, 31742–31753. [\[CrossRef\]](#)
22. Zhang, Y.; Zang, Z.; Zhang, X.; Song, L.; Yu, Z.; Wang, Y. Fault Diagnosis of Industrial Robot Based on Multi-Source Data Fusion and Channel Attention Convolutional Neural Networks. *IEEE Access* **2024**, *12*, 82247–82260. [\[CrossRef\]](#)
23. Miao, Z.; Zhou, F.; Yuan, X.; Xia, Y.; Chen, K. Multi-heterogeneous sensor data fusion method via convolutional neural network for fault diagnosis of wheeled mobile robot. *Appl. Soft Comput.* **2022**, *129*, 109554. [\[CrossRef\]](#)
24. Cui, H.; Guan, Y.; Chen, H. Rolling element fault diagnosis based on VMD and sensitivity MCKD. *IEEE Access* **2021**, *9*, 120297–120308. [\[CrossRef\]](#)
25. Li, Y.; Xu, M.; Wei, Y.; Huang, W. A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree. *Measurement* **2016**, *77*, 80–94. [\[CrossRef\]](#)
26. Zhou, K.; Lu, N.; Jiang, B. Basic probability assignment using intuitive fuzzy cloud model for information fusion and its application in fault diagnosis. *IEEE Trans. Instrum. Meas.* **2023**, *73*, 3504413. [\[CrossRef\]](#)
27. Liu, H.C.; Li, Z.; Song, W.; Su, Q. Failure mode and effect analysis using cloud model theory and PROMETHEE method. *IEEE Trans. Reliab.* **2017**, *66*, 1058–1072. [\[CrossRef\]](#)
28. Le, V.; Yao, X.; Miller, C.; Tsao, B.H. Series DC arc fault detection based on ensemble machine learning. *IEEE Trans. Power Electron.* **2020**, *35*, 7826–7839. [\[CrossRef\]](#)

29. Yu, W.; Zhao, C. Online fault diagnosis for industrial processes with Bayesian network-based probabilistic ensemble learning strategy. *IEEE Trans. Autom. Sci. Eng.* **2019**, *16*, 1922–1932. [[CrossRef](#)]
30. Alfarizi, M.G.; Tajiani, B.; Vatn, J.; Yin, S. Optimized random forest model for remaining useful life prediction of experimental bearings. *IEEE Trans. Ind. Inform.* **2022**, *19*, 7771–7779. [[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.