



Article Hybrid Multimodal Feature Fusion with Multi-Sensor for Bearing Fault Diagnosis

Zhenzhong Xu^{1,†}, Xu Chen^{2,†}, Yilin Li³ and Jiangtao Xu^{1,*}

- ¹ College of Aerospace and Civil Engineering, Harbin Engineering University, Harbin 150001, China
- ² School of Reliability and Systems Engineering, Beihang University, Beijing 100191, China
- ³ School of Foreign Languages, Beijing Institute of Technology, Beijing 102488, China
- * Correspondence: hit_xjt@163.com
- ⁺ These authors contributed equally to this work.

Abstract: Aiming at the traditional single sensor vibration signal cannot fully express the bearing running state, and in the high noise background, the traditional algorithm is insufficient for fault feature extraction. This paper proposes a fault diagnosis algorithm based on multi-sensor and hybrid multimodal feature fusion to achieve high-precision fault diagnosis by leveraging the operating state information of bearings in a high-noise environment to the fullest extent possible. First, the horizontal and vertical vibration signals from two sensors are fused using principal component analysis, aiming to provide a more comprehensive description of the bearing's operating condition, followed by data set segmentation. Following fusion, time-frequency feature maps are generated using a continuous wavelet transform for global time-frequency feature extraction. A first diagnostic model is then developed utilizing a residual neural network. Meanwhile, the feature data is normalized, and 28 time-frequency feature indexes are extracted. Subsequently, a second diagnostic model is constructed using a support vector machine. Lastly, the two diagnosis models are integrated to derive the final model through an ensemble learning algorithm fused at the decision level and complemented by a genetic algorithm solution to improve the diagnosis accuracy. Experimental results demonstrate the effectiveness of the proposed algorithm in achieving superior diagnostic performance with a 97.54% accuracy rate.

Keywords: multimodal feature fusion; multi-sensor; PCA; ResNet; SVM; ensemble learning

1. Introduction

In contemporary industrial production environments, rotating machinery stands as one of the most prevalent pieces of mechanical equipment [1], with bearings serving as crucial components within such machinery. Faults in bearings can result in a decrease in technical indicators, substantial economic losses, and even significant safety incidents [2–5]. The majority of faults in rotating machinery stem from bearing issues [6]. Hence, timely and accurate identification of bearing faults holds immense importance in mitigating the occurrence of accidents [7]. The operational status of a bearing cannot be adequately captured by a single sensor due to variations in sensor monitoring locations and environmental interferences. Moreover, the operational environment of bearings frequently experiences disruptions from shock, wear, and corrosion, leading to changes in the bearing surface geometry and therefore to non-stationary and non-linear impulse responses on damaged contact surfaces [8]. Extracting and identifying fault characteristics pose significant challenges. Thus, it is imperative to address the non-stationary and non-linear attributes of vibration signals captured from multiple sensors to extract fault characteristics effectively and develop a high-precision fault diagnosis model based on these features.

In a practical working environment, it is difficult to diagnose bearing faults directly. Sensors can capture digital signals, such as vibration signals, reflecting the condition of



Citation: Xu, Z.; Chen, X.; Li, Y.; Xu, J. Hybrid Multimodal Feature Fusion with Multi-Sensor for Bearing Fault Diagnosis. *Sensors* 2024, 24, 1792. https://doi.org/10.3390/ s24061792

Academic Editor: Jongmyon Kim

Received: 8 February 2024 Revised: 3 March 2024 Accepted: 6 March 2024 Published: 11 March 2024



Copyright: © 2024 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/). the bearing. Signal processing and analysis enable the determination of the bearing's condition [9,10]. During operation, the bearing generates vibration signals characterized by non-stationarity and non-linearity. Traditionally, signal features are extracted using methods such as empirical modal decomposition (EMD) [11] and wavelet transform (WT). EMD is primarily an empirical method. However, challenges such as modal aliasing arise during the signal decomposition process, limiting the applicability of EMD [12]. WT divides the signal's frequency band into multiple scales, facilitating feature extraction from both the time and frequency domains. This approach has demonstrated effectiveness in practical applications [13]. However, the above methods inevitably lose fault information during the process; consequently, this paper proposes a multimodal feature fusion method based on multi-sensors in order to maximize the retention and extraction of fault information.

In real industrial environments, multiple sensors are commonly deployed. The fusion of signals or extracted features from these sensors constitutes a pivotal step preceding fault diagnosis. Lupea et al. [14] extracted signal features from both time and frequency domains and used them in the classification process to solve the proposed multi-fault detection problem. Ye et al. introduced a multi-level feature fusion network for integrating vibration signals, yielding promising outcomes [15]. Chen et al. employed sparse autoencoders to fuse signals from multiple sensors and utilized deep belief networks for fault diagnosis [16]. Zhu et al. employed wavelet packet transform and multi-weight singular value decomposition for extracting time-frequency features. They employed a support vector machine (SVM) classifier [17]. Shao et al. utilized denoising autoencoders and contractive autoencoders to learn features. Subsequently, they employed locality-preserving projection for feature fusion, enhanced feature quality, and finally applied softmax for fault diagnosis [18]. Wang et al. introduced a three-stage multimodal feature fusion approach for integrating vibration and torque signals. They coupled this with an attention-based multidimensional concatenated convolutional neural network for fault diagnosis [19]. Buchaiah et al. employed a random forest for feature selection. They utilized 14 feature extraction techniques to extract 2D features and employed a SVM classifier for fault diagnosis [20]. Wang et al. applied principal component analysis (PCA) to fuse multidimensional features, characterizing the operating condition of rolling bearings [21].

In summary, it is challenging to combine signals from multiple sensors and perform feature fusion to extract information that characterizes the operational status of the bearing. Furthermore, integrating the extracted features and leveraging diverse classifiers for a final diagnosis represents a crucial and challenging phase. In this paper, a fault diagnosis algorithm based on multi-sensor and hybrid multimodal feature fusion is proposed. Ultimately, the algorithm's effectiveness is validated through comprehensive experimental results. The main contributions of this paper include the following:

- (1) The fusion of horizontal vibration signals (HVS) and vertical vibration signals (VVS) from a multi-sensor in a feature layer yields dual-channel data. This approach maximizes the integration of feature information from both sensors, thereby enhancing the robustness and generalization capabilities of the algorithm.
- (2) Initially, the dual-channel data is consolidated into a single-channel dataset. Continuous wavelet transform (CWT) is employed to extract global time-frequency feature information, generating time-frequency feature maps for training the residual neural network (ResNet). Simultaneously, time-frequency feature indexes are extracted postnormalization to obtain key indexes, facilitating training of the SVM. Utilizing global time-frequency features and key indexes for model training enhances the algorithm's learning capability.
- (3) Ensemble learning is employed to achieve decision-level fusion. The genetic algorithm (GA) is combined to address the multi-objective optimization model for obtaining weight parameters for the ResNet and SVM models. This integration harnesses the strengths of both base models, resulting in a diagnostic model with superior classification accuracy.

The remainder of the paper is organized as follows: The theoretical background, including multimodal feature fusion, time-frequency feature extraction, residual neural networks, support vector machines, and genetic algorithms, is briefly introduced in Section 2. Section 3 provides a description of the proposed diagnosis method in detail. In Section 4, the experimental results are analyzed and discussed. Finally, the conclusions are given in Section 5.

2. Theoretical Background

2.1. Multimodal Feature Fusion

PCA is a widely utilized algorithm for data dimensionality reduction. It recombines the original correlated variables into new variables that are uncorrelated with each other, aiming to retain as much information from the original variables as possible [22]. The information in the data is primarily captured by the variance, where a higher variance indicates more information content. The specific calculation steps are outlined below:

- (1) Data preprocessing;
- (2) Compute the matrix of correlation coefficients between variables, denoted as $R = (r_{st}), s, t = 1, 2, ..., n$;
- (3) Determine the eigenvalues λ_i and corresponding eigenvectors $e_i = (e_{i1}, e_{i2}, \dots, e_{ip})$, $||e_i|| = 1$ for R;
- (4) Compute the variance contribution *Vcr* and the cumulative variance contribution ratio *Cvcr* for the first *l* principal components as follows:

$$Vcr = \frac{\lambda_i}{\sum_{i=1}^n \lambda_i} \tag{1}$$

$$Cvcr = \frac{\sum_{i=1}^{l} \lambda_i}{\sum_{i=1}^{n} \lambda_i}$$
(2)

(5) Select the first *l* principal components $F_1, F_2, ..., F_l$ based on the following cumulative variance contribution ratio:

$$F_k = e_{k1}X_1 + e_{k2}X_2 + \dots + e_{kn}X_n, \ k = 1, 2, \dots, l$$
(3)

where F_i is uncorrelated with F_i ($i \neq j$; i, j = 1, 2, ..., l) and $ar(F_i) > var(F_j)$, i < j.

2.2. Time-Frequency Feature Extraction

2.2.1. Continuous Wavelet Transform

WT, introduced by Gilles [23], is a signal processing technique designed for analyzing non-smooth, non-linear vibration signals. Similar to the Fourier transform, it employs a family of wavelet functions to represent the signal. Let us assume the function $\psi \in L^2(R) \cap L^1(R)$, $\hat{\psi}(0) = 0$. ψ is translated and scaled to generate the following set of functions:

$$\psi_{a,b}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right), a, b \in \mathbb{R}, a \neq 0$$

$$\tag{4}$$

where $\psi_{a,b}$ represents an analytic or continuous wavelet; ψ is a fundamental wavelet; a is a scaling factor; and b is a translation factor.

CWT for an arbitrary function $f(t) \in L^2(R)$ is as follows:

$$W_f(a,b) = \langle f, \psi_{a,b} \rangle = |a|^{-1/2} \int f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt$$
(5)

where $\psi(t)$ represents the complex conjugate of $\psi(t)$; < f, $\psi_{a,b} >$ denotes the inner product; $W_f(a, b)$ denotes the coefficients of the wavelet function for a given scale factor *a* and

translation factor *b*; *a* and *b* are continuous variables, and when the WT is performed in this manner, it is referred to as CWT.

In this study, CWT is employed to convert the vibration signals into time-frequency feature maps. The process is as follows:

- (1) f_s represents the sampling frequency, F_c is the wavelet center frequency, and the actual center frequency corresponding to *a* is $F_a = F_c \times f_s/a$;
- (2) Let *totalscal* denote the length of the scale sequence during WT of the signal and *c* be a constant. The scale sequence takes the form ^c/_{totalscal}, ..., ^c/_{totalscal-1}, ^c/₄, ^c/₂, *c*;
- (3) According to step (1), *c*/*toalscal* corresponds to the actual frequency of $f_s/2$, thus $c = 2 \times F_c \times totalscal$. The scale sequence *t* can be calculated based on step (2);
- (4) After determining the wavelet base and scale, the wavelet coefficient $W_f(a, b)$ is computed according to the principle. Following step (1), the scale sequence is converted into a frequency sequence f, which is then combined with the time sequence t to obtain time-frequency feature maps.

2.2.2. Time-Frequency Indexes Extraction

When a fault occurs in the bearing, changes are observed in both time- and frequencydomain signals. Time-domain signals primarily undergo alterations in amplitude and distribution, while frequency-domain signals exhibit new frequency components and changes in spectral convergence. In the work of this paper, a total of 28 time-frequency indexes are selected, comprising 15 time-domain indexes and 13 frequency-domain indexes. These time-frequency characteristic indexes can comprehensively reflect the statistical characteristics of the bearing at the time of fault, thus helping to establish a fault diagnosis model.

As shown in Tables 1 and 2, parameters p_1 , $p_{12}-p_{15}$ primarily reflect changes in signal amplitude in the time domain, while parameters p_2 , p_6-p_{11} mainly describe the distribution of the signal in the time domain. Parameters p_3-p_5 mainly indicate the distribution of the energy in the time domain, while parameter p_{16} primarily represents vibration energy in the frequency domain. Parameters $p_{17}-p_{19}$, p_{21} , $p_{25}-p_{28}$ primarily describe changes in spectral power convergence, while parameters p_{20} , $p_{22}-p_{24}$ indicate positional changes in frequency [24].

The formulas for calculating the time domain characteristic indexes are shown in Table 1:

Table 1. Parameters of time domain characteristic indexes.

$p_1 = \frac{\sum_{n=1}^N x(n)}{N}$	$p_2 = \sqrt{rac{\sum_{n=1}^{N} (x(n) - p_1)^2}{(N-1)}}$
$p_3 = \left(\frac{\sum_{n=1}^{N} \sqrt{ x(n) }}{N}\right)^2$	$p_4 = \sqrt{\frac{\sum_{n=1}^{N} (x(n))^2}{N}}$
$p_5 = \max x(n) $	$p_6 = \frac{\sum_{n=1}^{N} (x(n) - p_1)^3}{(N-1)p_2^3}$
$p_7 = \frac{\sum_{n=1}^{N} (x(n) - p_1)^4}{(N-1)p_2^2}$	$p_8 = \frac{p_5}{p_3}$
$p_9 = \frac{p_5}{p_3}$	$p_{10} = \frac{p_4}{\frac{1}{N}\sum_{n=1}^{N} x(n) }$
$p_{11} = rac{p_5}{rac{1}{N}\sum_{n=1}^{N} x(n) }$	$p_{12} = \max x(n)$
$p_{13} = \min x(n)$	$p_{14} = p_{12} + p_{13} $
$p_{15} = p_1 $	

 $\overline{x(n)}$ is a time series, n = 1, 2, ..., N. *N* is the number of data points.

The formulas for calculating the frequency domain characteristic indexes are shown in Table 2:

Table 2. Parameters of frequency domain characteristic indexes.

$p_{16} = \frac{\sum_{k=1}^{K} s(k)}{K} \\ p_{18} = \frac{\sum_{k=1}^{K} (s(k) - p_{16})^3}{K (\sqrt{p_{17}})^3}$	$p_{17} = \frac{\sum_{k=1}^{K} (s(k) - p_{16})^2}{K - 1}$ $p_{19} = \frac{\sum_{k=1}^{K} (s(k) - p_{16})^4}{K p_{17}^2}$
$p_{20} = \frac{\sum_{k=1}^{K} f_k s(k)}{\sum_{k=1}^{K} s(k)}$	$p_{21} = \sqrt{\frac{\sum_{k=1}^{K} (f_k - p_{20})^2 s(k)}{K}}$

Table 2. Cont.

$$p_{22} = \sqrt{\frac{\sum_{k=1}^{K} f_k^{\ 2} s(k)}{\sum_{k=1}^{K} s(k)}}}{p_{24} = \sqrt{\frac{\sum_{k=1}^{K} f_k^{\ 2} s(k)}{\sum_{k=1}^{K} f_k^{\ 2} s(k)}}} \qquad p_{23} = \sqrt{\frac{\sum_{k=1}^{K} f_k^{\ 4} s(k)}{\sum_{k=1}^{K} f_k^{\ 2} s(k)}}}{p_{25} = \frac{p_{21}}{p_{20}}}$$

$$p_{26} = \frac{\sum_{k=1}^{K} (f_k - p_{20})^3 s(k)}{Kp_{21}^3} \qquad p_{27} = \frac{\sum_{k=1}^{K} (f_k - p_{20})^4 s(k)}{Kp_{21}^4}}{p_{28}}$$

s(k) is the spectrum, k = 1, 2, ..., K. *K* is the number of spectral lines, and f_k is the frequency value of the *K*th spectral line.

2.3. Residual Neural Network

To address the issue of gradient disappearance encountered in traditional convolutional neural networks (CNN) as depth increases, He [25] introduced the ResNet. The key innovation of this network lies in the incorporation of residual blocks. The structure of the residual block is shown in Figure 1, which directly connects the residual block to the forward neural network. The forward neural network contains a convolutional layer, a batch normalization layer, and a Relu activation function layer. The input of the residual block is x, the output is H_x , and the forward neural network layer is F_x , $F_x = H_x - x$. For instance, in the case of a forward neural network layer with two weight layers, F_x and x are elementwise summed up during direct concatenation, ensuring equal dimension mapping.

$$\begin{cases} y = F(x, \{W_i\}) + x \\ F = W_2 \sigma(W_1 x) \end{cases}$$
(6)

where $F(x, \{W_i\})$ is the residual mapping; *F* is the residual function; *W* is the weight parameter; and σ is the activation function.



Figure 1. Residual block.

When the dimensions of the mappings within the residual block are dissimilar, a linear mapping is performed on *x* to ensure dimensional alignment. Subsequently, the outputs

of the linear mapping and the neural network layers are added together while preserving the dimensions.

$$\begin{cases} y = F(x, \{W_i\}) + W_s x\\ F = W_2 \sigma(W_1 x) \end{cases}$$
(7)

where W_s is the mapping weight of the input *x*.

ResNet offers enhanced optimization and improved accuracy compared to traditional convolutional neural networks, particularly as network depth increases. Different types and sizes of data can benefit from various configurations of ResNet, such as ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152.

2.4. Support Vector Machine

SVM is a machine learning algorithm grounded in statistical theory, primarily employed for classification tasks [26]. Fault diagnosis based on time-frequency feature indexes inherently presents a challenge as a small-sample and nonlinear classification problem. SVM typically exhibits superior generalization to small-sample data compared to Back Propagation (BP), Extreme Learning Machines (ELM), Long Short-Term Memory (LSTM), and Deep Belief Network (DBN), which may suffer from overfitting. While SVM leverages kernel methods to address nonlinear classification issues, BP, LSTM, and DBN necessitate intricate architectures and parameter tuning to handle nonlinear relationships. SVM offers enhanced interpretability relative to BP and ELM, with the added benefit of avoiding local minima due to their convex optimization nature. LSTM excels at processing temporal information, yet in this paper, time-frequency feature indexes are extracted from temporal signals, which diverges from temporal information. In summary, for the specific problem addressed in this paper, SVM demonstrates a pronounced advantage. Its fundamental concept revolves around identifying the optimal classification hyperplane by maximizing the margin between the hyperplane and the classification data, as depicted in Figure 2:



Figure 2. Optimal classification hyperplane.

Let the input data be $x_i = (i = 1, 2, ..., M)$, with M representing the number of samples. For the binary classification problem, the classification label $y \in \{-1, +1\}$. For linearly separable data, the corresponding hyperplane is defined as follows:

$$f_x = w^T x + b = 0 \tag{8}$$

where w represents the M-dimensional weight vector and b is the bias scalar. These two parameters determine the distance between the hyperplane and the origin. Maximization of the margin between different classes is achieved by solving the following optimization problem:

n

$$uax \frac{2b}{||w||^2} \tag{9}$$

For cases where linear separability is not possible, a soft-spaced support vector machine is implemented by introducing the slack variable $\varepsilon_i \ge 0$ and the penalty factor *C* as follows:

$$min\frac{1}{2}||w||^2 + C\sum_{i=1}^n \varepsilon_i \tag{10}$$

For multi-classification problems involving *N* classes of samples, N(N-1)/2 standard SVM classifiers are trained, and the final result is based on the predictions of each classifier.

In SVM, the kernel function is employed to compute the similarity or inner product between input vectors, expressing the relationship between data in the feature space. The Radial Basis Function (RBF) is chosen as the kernel function due to its suitability for highdimensional data and nonlinear mapping. The expression of the RBF kernel function is given by the following equation:

$$k(x_i, x_j) = exp(-\gamma ||x_i - x_j||)$$
(11)

where x_i, x_j represent the input vector and center vector, respectively; $||x_i - x_j||$ denotes the Euclidean distance; and γ is a monotonic function that takes a small value when x_i is far from x_j .

2.5. Genetic Algorithm

GA is an adaptive global optimization search technique inspired by nature's genetic mechanisms and biological evolution, employing binary genetic coding [27]. Assuming that allele $\Gamma = \{0, 1\}$ and individual space $H_L = \{0, 1\}^L$. Reproduction comprises crossover and mutation. The optimization steps are shown below:

- (1) Determine the population size N, crossover probability P_c , mutation probability P_m , and termination criterion. Randomly generate N individuals as the initial population X(0) and set an algebraic counter $t \leftarrow 0$.
- (2) Calculate the fitness of individuals in X(t).
- (3) Select M/2 pairs of matrices from X(t) using a selection operator, where $M \ge N$.
- (4) Perform crossover among the selected pairs to create M intermediate individuals according to P_c .
- (5) Apply mutations to the *M* intermediate individuals according to P_m to obtain *M* candidate individuals.
- (6) Select *N* individuals from the *M* candidates based on fitness to form the new generation population X(t + 1).
- (7) If the termination criterion has been satisfied, output the individual with maximum fitness in X(t + 1) as the optimal solution; otherwise, $t \leftarrow t + 1$ and return to (5).

3. Proposed Method

The proposed fault diagnosis algorithm leverages multi-sensor and hybrid multimodal feature fusion, as depicted in Figure 3. The model framework comprises the following three primary components: multimodal signal feature fusion, model training optimization, and multimodal decision fusion.

Initially, HVS and VVS from the bearing are captured by two sensors and normalized via min-max scaling. Subsequently, the vibration signals in both directions undergo fusion using min-max scaling and PCA. In the next stage, the fused feature signals (FVS) from the two sensors are partitioned into datasets. Time-frequency maps are generated using CWT, while a ResNet model is trained to establish a diagnostic model. Simultaneously, 28 time-frequency feature index parameters are extracted, normalized using min-max scaling, and used to train SVM for another diagnostic model. Finally, leveraging the base models obtained, GA optimizes the weight hyperparameters. The final diagnostic model is constructed using ensemble learning techniques, thereby enhancing the overall accuracy of the diagnostic system.



Figure 3. Fault diagnosis algorithm based on multi-sensor and hybrid multimodal feature fusion.

3.1. Multimodal Signal Feature Fusion

Two sensors, such as acceleration meters, are deployed at distinct locations on the test rig. Each sensor captures vibration signals across various modes in two directions: HVS and VVS. These signals, influenced by the sensor deployment positions and vibration directions, exhibit distinct fault characteristics. Initially, these signals are normalized using

min-max scaling. Employing PCA facilitates data dimensionality reduction and enables the fusion of diverse modal signals from each sensor to yield FVS, as depicted in Figure 4.



Figure 4. Single-sensor data fusion.

3.2. Model Training Optimization

3.2.1. ResNet Diagnostic Model Based on Time-Frequency Feature Maps

After acquiring the FVS through PCA, the signal undergoes segmentation to generate a dataset based on the sampling frequency and interval. Subsequently, CWT is applied to analyze and process feature information across time and frequency domains, resulting in the generation of time-frequency maps from three channels. Finally, the ResNet-34 model is trained to establish the initial base classification model. The model architecture and steps are as depicted in Figure 5:



Figure 5. ResNet-34 diagnostic model flowchart.

The loss function adopts the cross-entropy loss function, which is defined as shown in the following equation:

$$L = \frac{1}{N} \sum_{i} L_{i} = -\frac{1}{N} \sum_{i} \sum_{c=1}^{M} y_{ic} log(p_{ic})$$
(12)

where *M* is the number of categories; y_{ic} denotes the sign function (0 or 1), taking 1 if the true category of sample *i* is equal to *c* and 0 otherwise; and p_{ic} is the predicted probability of observing that sample *i* belongs to category *c*. The optimization algorithm used in this paper is stochastic gradient descent.

3.2.2. SVM Diagnostic Model Based on Time-Frequency Indexes

On the alternate pathway, employing the time-frequency index formulas, 28 key indexes are computed for each data segment. Following normalization via min-max scaling, SVM is trained. The One-vs.-Rest strategy is adopted to enable multi-class classification,



thereby yielding an additional base model. The model architecture and steps are outlined as depicted in Figure 6:

Figure 6. SVM diagnostic model flowchart.

The dataset utilized in this study comprises high-dimensional, linearly indivisible data, prompting the adoption of the RBF as the kernel function due to its ability to handle such data structures effectively. The penalty factor *C* and the RBF parameter γ play pivotal roles in shaping the accuracy of the classification model.

To optimize the SVM classification model, this paper employed the grid search method, a technique widely recognized for its effectiveness in determining optimal parameter values. Through systematic exploration within predefined ranges, the grid search method facilitates the identification of the most suitable values for *C* and γ , ensuring the creation of an accurate classification model.

3.3. Multimodal Decision Fusion

After training to obtain two base models, ResNet and SVM, set a weight parameter for each base model to combine their diagnostic results for the purpose of decision fusion. Define two weight coefficients a, b as shown in the following equation:

$$0 \le a \le 1 \tag{13}$$

$$a+b=1\tag{14}$$

$$P(x) = argmax(\frac{a \times P_{Resnet}(x) + b \times P_{SVM}(x)}{2})$$
(15)

where P(x) is the predicted category; $P_{Resnet}(x)$ is the probability of the category predicted by the ResNet model; $P_{SVM}(x)$ is the probability of the category predicted by the SVM model; *a* is the weight coefficient assigned to the ResNet model; and *b* is the weight coefficient assigned to the SVM model.

Table 3 shows the confusion matrix for the binary classification, which presents the complete evaluation of the classification results.

A

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$
(16)

$$Precision = \frac{TP}{TP + FP} \tag{17}$$

$$Recall = \frac{TP}{TP + FN}$$
(18)

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$
(19)

Table 3. Confusion matrix for the binary classification.

Total Population	Condition Positive	Condition Negative	
Predicted condition positive	True positive (<i>TP</i>)	False positive (<i>FP</i>)	
Predicted condition negative	False negative (<i>FN</i>)	True negative (<i>TN</i>)	

GA is employed to optimize the weight hyperparameters a, b in the context of maximizing both accuracy and F1 score. This optimization task presents a multi-objective problem, where the objective function aims to balance the trade-off between accuracy and F1 score as follows:

$$max(Accuracy, F1) \tag{20}$$

4. Case Study

In order to evaluate the effectiveness of the proposed algorithm, experimental tests were conducted using the test rig depicted in Figure 7. The components of the test rig include an electric motor, an inertia wheel for applying radial load, a coupling, a conveyor belt drive mechanism, a conveyor belt, a crank-connecting rod mechanism, a gearbox, and a reciprocating mechanism with a spring. The bearing under investigation is mounted close to the motor side and is identified as a contact deep groove ball bearing (MB ER-10K). Three distinct types of faults, namely inner race fault (IF), outer race fault (OF), and rolling element fault (RF), were intentionally implanted in the bearing.



Figure 7. A schematic of the bearing test rig.

For fault implantation, localized cracks with a width and depth of 0.2 mm were introduced in the outer race groove, inner race groove, and ball surface of the bearing. Two acceleration meters (AM), specifically PCB Model 608A11, were strategically mounted at locations indicated in Figure 7. These meters, with a sensitivity of $10.2 \,(\text{mV/ms}^{-2})/100 \,(\text{mv/g})$, played a crucial role in capturing vibration data essential for fault diagnosis. The rotational speed is 900 revolutions per minute. The acquisition card used in the experiment is the NI 9234, operating at a sampling frequency of 25.6 kHz. $\frac{AM}{I} \frac{Fault Type}{II} \frac{HVS}{I} \frac{VVS}{I}$ $\frac{IF}{I} OF \frac{1}{1} OF \frac{1$

Table 4. Vibration signal time domain.

Table 4:

RF

IF refers to inner race fault, OF to outer race fault, and RF to rolling element fault; AM refers to acceleration meter, HVS to horizontal vibration signals, and VVS to vertical vibration signals.

The raw signals of the three types of faults measured by the multi-sensor are shown in

From Table 4, it is evident that the amplitude of time-domain signals varies significantly across different fault types. Typically, the signals exhibit distinct amplitudes: the RF signal displays the highest amplitude, followed by the IF signal, with the OF signal registering the smallest amplitude. Additionally, it is observed that the amplitude difference among measured signals in various directions for the same fault remains relatively minor.

4.1. Feature Layer Fusion

HVS and VVS from each of the multi-sensors are fused using PCA. The contribution of the two principal components after fusion is shown in Table 5:

Table	5.	PCA	results.
-------	----	-----	----------

AM	Fault Type	Principal Component 1	Principal Component 2
1	IF	0.83593733	0.16406267
	OF	0.63439911	0.36560089
	RF	0.81361577	0.18638423
2	IF	0.76050419	0.23949581
	OF	0.64756928	0.35243072
	RF	0.71948348	0.28051652

IF refers to inner race fault, OF to outer race fault, and RF to rolling element fault; AM refers to acceleration meter, HVS to horizontal vibration signals, and VVS to vertical vibration signals.

From Table 5, it can be seen that for the three faults, the principal component 1 of fused vibration signals from the two sensors can cover 63–84% of the original information. This table can prove that the signals after PCA fusion well extract the fault feature information. In this paper, principal component 1 is chosen as FVS. The fused feature signal is shown in Figure 8:



Figure 8. FVS. (a) AM 1 IF; (b) AM 1 OF; (c) AM 1 RF; (d) AM 2 IF; (e) AM 2 OF; (f) AM 2 RF.

In Figure 8, it is evident that following data normalization, the amplitudes predominantly fall within the range of -0.5 to 0.5. Notably, the amplitude ranges vary noticeably across the three different faults examined. Specifically, the amplitude of the IF signal exhibits a broader distribution, while that of the RF signal appears more concentrated. This distinction suggests varying fault characteristics within the analyzed signals. Furthermore, despite normalization, the overall waveform remains largely consistent with the original vibration signal, indicating the preservation of fundamental signal characteristics post-normalization.

4.2. Base Model Training Optimization

The dataset was obtained with a sampling frequency of 25.6 kHz. Each segment of FVS is sliced into intervals of 0.05 s, containing 1280 sample points. In total, 2240 segments are generated from the dataset. The dataset is divided into training and test sets according to an 8:2 ratio, resulting in 1792 segments allocated for training and 448 segments for testing.

4.2.1. ResNet-34 Diagnostic Model

CWT is employed to process the datasets and generate time-frequency feature maps. Given that the fault frequencies predominantly lie within the low-frequency band, the vertical coordinate display range is constrained to 0–1000 Hz to capture the primary global feature information. This study utilizes the Gaussian wavelet as the wavelet function for the CWT, with a scale factor set to 256 to ensure comprehensive feature extraction across the signal spectrum.

For illustration, consider a subset of three signals from each fault type. Their respective time-frequency feature maps are presented as depicted in Figure 9:

The ResNet model, comprising ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152 architectures, is employed for training on the extracted time-frequency feature maps. These models are evaluated to select the most effective diagnostic model for the task at hand. Table 6 provides a summary of the model parameters for each ResNet architecture.



Figure 9. CWT time-frequency feature maps. (a) IF; (b) OF; (c) RF.

Tab	le 6	5. I	Resl	Net	mod	lel	parameters.	
-----	------	------	------	-----	-----	-----	-------------	--

Layer Name	Output Size	18-Layer 34-Layer 50-Layer		101-Layer	152-Layer		
Input layer		(200,200,3)					
Conv1	$\begin{array}{c} 112 \times 112 \\ 56 \times 56 \end{array}$		7 × 7, 64, stride 2 bn, 3 × 3 max pool, stride 2				
Conv2_x	56 × 56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	
Conv3_x	28 ×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 8$	
Conv4_x	14×14	$\begin{bmatrix} 3 \times 3,256 \\ 3 \times 3,256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 36$	
Conv5_x	7×7	$ \begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2 $	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	
Output layer	1×1	Average pool, 1000-d fc, softmax					

The input image size is set to 200×200 pixels with three channels (RGB). Conv1 serves as a preprocessing convolutional layer, featuring a 7 × 7 convolutional layer with a stride of 2, followed by a batch normalization layer, and a 3 × 3 maximum pooling layer with a stride of 2. Subsequently, Conv2_x, Conv3_x, Conv4_x, and Conv5_x represent the residual structures, each comprising a variable number of residual blocks. The output of the final residual layer undergoes average pooling before entering the fully connected layer (FC) for classification. Finally, the softmax classifier outputs the probability values corresponding to the categories, yielding the recognition results.

In this paper, the training iteration is 40 epochs, and the model effect is tested at five epoch intervals. The training and testing results are shown in Figure 10.



Figure 10. Train and test results (a) train accuracy; (b) test accuracy; (c) train loss; (d) test loss.

From Figure 10a,c, it is observed that the ResNet-34 model demonstrates faster convergence in terms of accuracy and loss on the training set, outperforming other models. From Figure 10b,d, the ResNet-34 model achieves an accuracy of 89.06% with a corresponding loss of 0.396 after 40 epochs of training, indicating superior performance compared to alternative models.

The classification results of the faults are shown in Figure 11, which reveal that the ResNet-34 model achieves a precision of 89.22%, a recall of 89.47%, and an F1 score of 0.8928. These metrics collectively demonstrate the model's effectiveness in fault classification, with ResNet-34 showcasing superior performance in comparison to other models. Furthermore, the AUC values for all three fault types are observed to be close to one, indicating excellent classification performance across the board. Consequently, the ResNet-34 model emerges as a robust base model for subsequent ensemble learning diagnostic models.



Figure 11. ResNet-34 diagnosis results: (a) confusion matrix; (b) ROC and AUC.

Utilizing the 28 time-frequency index calculation formulas, 28 time-frequency feature indexes are obtained for each piece of data. In order to visualize the expression of these time-frequency feature indicators for the features, the t-SNE method is used to visualize the high-dimensional data.

Figure 12a demonstrates that FVS on a two-dimensional plane makes it challenging to distinguish between various types of data points effectively. However, Figure 12b reveals that after extracting 28 time-frequency indexes, clear differentiation among data points is observed, particularly with RF being notably distinct from IF and OF, while differentiation between IF and OF is less pronounced. In this study, min-max normalization is employed to normalize the time-frequency indexes. Figure 12c illustrates that after normalization, the three types of fault data points become highly distinguishable, effectively expressing the characteristic information of each fault type while minimizing the impact of noise. The enhanced distinguishability among fault types is crucial for accurate fault diagnosis and analysis.



Figure 12. Two-dimensional visualization by t-SNE. (**a**) FVS; (**b**) time-frequency indexes; (**c**) time-frequency indexes after normalization.

In this study, RBF is employed as the kernel function for SVM to perform classification tasks. The classification effectiveness of the SVM model is primarily influenced by the following two key parameters: the penalty factor *C* and the parameter γ of the RBF kernel function. The values of the penalty factor *C* and the parameter γ of the RBF kernel function are shown in Table 7.

Table 7. Parameter val	lues
------------------------	------

(·	γ
0.10	0.50	0.001	0.005
0.46	2.32	0.010	0.050
2.15	10.77	0.100	0.500
10.00	50.00	10.000	50.000

In this paper, the grid search method is used to solve the penalty factor C and the parameter γ . The classification accuracy of the SVM model is shown in Figure 13.

The model classification accuracy is 95.76% when C is 50 and γ is 0.5, which is the best performance among the parameter combinations.

The results of fault diagnosis using the SVM model are presented in Figure 14a. The model demonstrates high precision (96.04%), recall (95.97%), and F1 score (0.9581), indicating its effectiveness in accurately classifying the three types of faults. Additionally, Figure 14b illustrates that the AUC values for the three fault types are close to one, underscoring the SVM's robust classification performance across fault categories. These findings affirm the SVM's superior classification efficacy in fault diagnosis, warranting its selection as another base model for subsequent ensemble learning diagnostic models.



Figure 13. SVM classification accuracy.



Figure 14. SVM diagnosis results: (a) confusion matrix; (b) ROC and AUC.

4.3. Diagnostic Model Based on Ensemble Learning

In this study, ensemble learning techniques are employed to combine the diagnostic results of two base models through decision-level fusion, aiming to enhance the accuracy of fault diagnosis. The synthesis of diagnostic results is achieved using GA, which iteratively optimizes the weight coefficient values *a*,*b* assigned to the base models.

GA aims to maximize the accuracy and F1 score, treating them as fitness functions guiding the optimization process. A population size of 50 individuals is initialized, with a maximum population size capped at 100 to ensure computational efficiency. A mutation probability of 0.2 and a crossover probability of 0.7 are defined to introduce variability and exploration within the population.

The optimization process iterates over 1–15 generations, seeking to converge towards individuals with the highest fitness values, which represent the optimal weight coefficient values for combining the base models. By leveraging GA, the study seeks to identify the

most effective combination of base models that maximizes diagnostic accuracy and *F*1 score, ultimately improving fault diagnosis performance.

According to the GA solution results in Figure 15, the accuracy and F1 of the model reached their maximum value when the population was iterated to the ninth generation and a was found to be 0.4859 and b was 0.5141. Therefore, these two values were used as weights for the base model ResNet-34 and the SVM diagnostic model.



Figure 15. GA results.

Using this integrated learning model for the test set, the diagnostic results are obtained as shown in Figure 16.



Figure 16. Ensemble model diagnosis results: (a) confusion matrix; (b) ROC and AUC.

The diagnostic results of the ensemble learning model for three types of faults are presented in Figure 16a. The model achieves an impressive accuracy rate of 97.54%, a precision rate of 97.63%, a recall rate of 97.68%, and an F1 score of 0.9757. Additionally, Figure 16b illustrates that the AUC values for the three fault types are very close to one, indicating excellent classification performance across fault categories.

19 of 21

As depicted in Figure 11, Figure 14, and Figure 16, distinguishing between IF and OF poses a challenge primarily due to the mechanical similarity between the inner and outer races. The operational faults occurring in the inner and outer races of the bearing result in a partial overlap of their spectral components. Additionally, factors such as sampling frequency, sensor placement, and noise interference further complicate the distinction between IF and OF.

As shown in Table 8, these evaluation metrics collectively demonstrate the superior performance of the ensemble learning model compared to the individual base models. The ultimate diagnostic model demonstrated an 8.48% and 1.78% enhancement in accuracy compared to the two foundational models, accompanied by improvements in the other three metrics to varying degrees. By combining the strengths of the ResNet-34 and SVM diagnostic models through ensemble learning, the model achieves a better diagnostic effect, effectively enhancing fault diagnosis accuracy and reliability.

Table 8. Classification evaluation metrics.

Algorithm	Accuracy	Precision	Recall	F1
ResNet-34	0.8906	0.8922	0.8947	0.8928
SVM	0.9576	0.9604	0.9597	0.9581
The proposed ensemble model	0.9754	0.9763	0.9768	0.9757

5. Conclusions

The proposed fault diagnosis algorithm based on multi-sensor and hybrid multimodal feature fusion shows promise in improving diagnostic accuracy and credibility. The fusion of hybrid multimodal features is delineated into two stages. In the initial phase, vibration signals from distinct sensors in varying directions are amalgamated at the feature layer. Subsequently, in the second stage, diagnostic outcomes from diverse modal data sources are integrated at the decision layer employing ensemble learning techniques. By fusing features from the HVS and VVS using PCA, the algorithm extracts global time-frequency information through CWT and key index information using 28 time-frequency indexes. This approach addresses the non-stationary and non-linear characteristics of vibration signals and enhances the accuracy and credibility of diagnostic results.

The utilization of ResNet-34 and SVM models for fault diagnosis, followed by ensemble learning for decision layer fusion, further improves diagnostic performance. Despite the overall effectiveness of the proposed algorithm, misclassifications between IF and OF remain a challenge. Subsequent research should optimize the feature extraction technique, enhance the classifier structure, and enrich the model training with additional data. These steps are essential for effectively discerning between bearing IF and OF.

In conclusion, this paper presents an effective solution for leveraging multi-sensor vibration signals of different modes for fault diagnosis in high-noise environments. The proposed approach employs a two-stage fusion process involving the feature layer and the decision layer, resulting in an efficient fault diagnosis algorithm. The algorithm shows promise as a robust classifier, opening avenues for further refinement and optimization in future research endeavors.

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

Funding: This research was funded by the National Natural Science Foundation of China under Grant No. 11372080.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data contained within the article.

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

References

- 1. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* 2018, 108, 33–47. [CrossRef]
- 2. Li, J.; Wang, Y.; Zi, Y.; Jiang, S. A local weighted multi-instance multilabel network for fault diagnosis of rolling bearings using encoder signal. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 8580–8589.
- 3. Li, Y.; Yang, Y.; Wang, X.; Liu, B.; Liang, X. Early fault diagnosis of rolling bearings based on hierarchical symbol dynamic entropy and binary tree support vector machine. *J. Sound Vib.* **2018**, *428*, 72–86. [CrossRef]
- Mikic, D.A.; Epifanic-Pajic, V.L.; Desnica, E.; Asonja, A.L.; Stojanovic, B.L. Reliability analysis of ball bearing on the crankshaft of piston compressors. J. Balk. Tribol. Assoc. 2016, 22, 5060–5070.
- Hadi, R.H.; Hady, H.N.; Hasan, A.M.; Al-Jodah, A.; Humaidi, A.J. Improved Fault Classification for Predictive Maintenance in Industrial IoT Based on AutoML: A Case Study of Ball-Bearing Faults. *Processes* 2023, 11, 1507. [CrossRef]
- 6. Zhang, B.; Zhang, L.; Xu, J. Degradation feature selection for remaining useful life prediction of rolling element bearings. *Qual. Reliab. Eng. Int.* **2016**, *32*, 547–554. [CrossRef]
- Wei, Y.; Li, Y.; Xu, M.; Huang, W. A review of early fault diagnosis approaches and their applications in rotating machinery. *Entropy* 2019, 21, 409. [CrossRef]
- Mikić, D.; Desnica, E.; Kiss, I.; Mikić, V. Reliability Analysis of Rolling Ball Bearings Considering the Bearing Radial Clearance and Operating Temperature. *Adv. Eng. Lett.* 2022, 1, 16–22. [CrossRef]
- 9. Wang, Z.; Xiao, F. An improved multisensor data fusion method and its application in fault diagnosis. *IEEE Access* 2018, 7, 3928–3937. [CrossRef]
- 10. Liu, Q.C.; Wang, H.P.B. A case study on multisensor data fusion for imbalance diagnosis of rotating machinery. *Ai Edam* **2001**, *15*, 203–210. [CrossRef]
- 11. Ambrożkiewicz, B.; Syta, A.; Georgiadis, A.; Gassner, A.; Litak, G.; Meier, N. Intelligent Diagnostics of Radial Internal Clearance in Ball Bearings with Machine Learning Methods. *Sensors* **2023**, *23*, 5875. [CrossRef] [PubMed]
- 12. Wang, J.; Du, G.; Zhu, Z.; Shen, C.; He, Q. Fault diagnosis of rotating machines based on the EMD manifold. *Mech. Syst. Signal Process.* 2020, 135, 106443. [CrossRef]
- 13. Feng, Z.; Liang, M.; Chu, F. Recent advances in time-frequency analysis methods for machinery fault diagnosis: A review with application examples. *Mech. Syst. Signal Process.* **2013**, *38*, 165–205. [CrossRef]
- 14. Lupea, I.; Lupea, M. Machine Learning Techniques for Multi-Fault Analysis and Detection on a Rotating Test Rig Using Vibration Signal. *Symmetry* **2023**, *15*, 86. [CrossRef]
- 15. Ye, Z.; Yu, J. Multi-level features fusion network-based feature learning for machinery fault diagnosis. *Appl. Soft Comput.* **2022**, 122, 108900. [CrossRef]
- 16. Chen, Z.; Li, W. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 1693–1702. [CrossRef]
- 17. Zhu, H.; He, Z.; Wei, J.; Wang, J.; Zhou, H. Bearing fault feature extraction and fault diagnosis method based on feature fusion. *Sensors* **2021**, *21*, 2524. [CrossRef] [PubMed]
- Shao, H.; Jiang, H.; Wang, F.; Zhao, H. An enhancement deep feature fusion method for rotating machinery fault diagnosis. *Knowl.-Based Syst.* 2017, 119, 200–220. [CrossRef]
- 19. Wang, D.; Li, Y.; Jia, L.; Song, Y.; Liu, Y. Novel three-stage feature fusion method of multimodal data for bearing fault diagnosis. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–10. [CrossRef]
- 20. Buchaiah, S.; Shakya, P. Bearing fault diagnosis and prognosis using data fusion based feature extraction and feature selection. *Measurement* **2022**, *188*, 110506. [CrossRef]
- 21. Wang, H.; Ni, G.; Chen, J.; Qu, J. Research on rolling bearing state health monitoring and life prediction based on PCA and Internet of things with multi-sensor. *Measurement* **2020**, *157*, 107657. [CrossRef]
- 22. Maćkiewicz, A.; Ratajczak, W. Principal components analysis (PCA). Comput. Geosci. 1993, 19, 303–342. [CrossRef]
- 23. Gilles, J. Empirical wavelet transform. *IEEE Trans. Signal Process.* 2013, *61*, 3999–4010. [CrossRef]
- 24. Lei, Y.; He, Z.; Zi, Y.; Hu, Q. Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs. *Mech. Syst. Signal Process.* **2007**, *21*, 2280–2294. [CrossRef]
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

- 26. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28. [CrossRef]
- 27. Mirjalili, S. Genetic algorithm. In *Evolutionary Algorithms and Neural Networks: Theory and Applications;* Springer: Cham, Switzerland, 2019; pp. 43–55.

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.