electronics

MDPI

Article

Sandpiper Optimization with a Deep Learning Enabled Fault Diagnosis Model for Complex Industrial Systems

Mesfer Al Duhayyim [1,*], Heba G. Mohamed [2], Jaber S. Alzahrani [3], Rana Alabdan [4], Amira Sayed A. Aziz [5], Abu Sarwar Zamani [6], Ishfaq Yaseen [6] and Mohamed Ibrahim Alsaid [6]

1 Department of Computer Science, College of Sciences and Humanities-Aflaj, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
2 Department of Electrical Engineering, College of Engineering, Princess Nourah Bint Abdulrahman University, Riyadh 11671, Saudi Arabia
3 Department of Industrial Engineering, College of Engineering at Alqunfudah, Umm Al-Qura University, Mecca 21961, Saudi Arabia
4 Department of Information Systems, College of Computer and Information Science, Majmaah University, Al-Majma'ah 11952, Saudi Arabia
5 Department of Digital Media, Faculty of Computers and Information Technology, Future University in Egypt, New Cairo 11835, Egypt
6 Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
* Correspondence: m.alduhayyim@psau.edu.sa

Abstract: Recently, artificial intelligence (AI)-enabled technologies have been widely employed for complex industrial applications. AI technologies can be utilized to improve efficiency and reduce human labor in industrial applications. At the same time, fault diagnosis (FD) and detection in rotating machinery (RM) becomes a hot research field to assure safety and product quality. Numerous studies based on statistical, machine learning (ML), and mathematical models have been available in the literature for automated fault diagnosis. From this perspective, this study presents a novel sandpiper optimization with an artificial-intelligence-enabled fault diagnosis (SPOAI-FD) technique for intelligent industrial applications. The aim is to detect the existence of faults in machineries. The proposed model involves the design of a continuous wavelet transform (CWT)-based pre-processing approach, which transforms the raw vibration signal into a useful format. In addition, a bidirectional long short-term memory (BLSTM) model is applied as a classifier, and the Faster SqueezeNet model is applied as a feature extractor. In order to modify the hyperparameter values of the BLSTM model, the sandpiper optimization algorithm (SPOA) can be utilized, showing the novelty of the work. A wide range of simulation analyses were conducted on benchmark datasets, and the results highlighted the supremacy of the SPOAI-FD algorithm over recent approaches.

Keywords: industrial applications; intelligent systems; artificial intelligence; fault diagnosis; rotating machines

check for updates

Citation: Al Duhayyim, M.; G. Mohamed, H.; S. Alzahrani, J.; Alabdan, R.; Aziz, A.S.A.; Zamani, A.S.; Yaseen, I.; Alsaid, M.I. Sandpiper Optimization with a Deep Learning Enabled Fault Diagnosis Model for Complex Industrial Systems. *Electronics* **2022**, *11*, 4190. https://doi.org/10.3390/electronics11244190

Academic Editor: Mohamed Benbouzid

Received: 5 October 2022
Accepted: 24 November 2022
Published: 15 December 2022

1. Introduction

The deployment of artificial intelligence (AI) is critical for success in the complex industrial sector. In particular, AI solutions have become increasingly important as they assist in developing effective smart services, optimizing production process, and forecasting machinery failure [1]. Using this information, industrial professionals could make more informed decisions with improved productivity, efficiency, and safety. Consequently, industries are automated, and people have become increasingly linked, more than ever before. Because of the extensive benefits of smart industry, several fields have started to use it. Fields such as agriculture, energy, automobiles, oil, gas, and so on are some of the typical examples. However, advances in the technology of smart industrial applications have become critical for meeting the requirements of industry 4.0 [2]. The possibilities of using AI in smart industry are relatively diverse and broad. According to the application requirement, the usage of AI provides trusted recommendations, assists with anticipated needs, and manages tasks. Adapting AI-driven technology will provide a competitive benefit through several smart industrial applications [3]. This is due to the fact that irrespective of industry type and size, AI provides potential solutions to all sectors.

Rotating machinery has become a crucial equipment in industries [4]. Over recent years, efficient RM has been deployed in the production of accurate machine tool spindles, the latest supersonic vector aircraft engine, efficient marine propulsion motors, massive generator sets, among others—all of which are designed to achieve unmanned, automated, and maximal speeds. For approving the scalability and security of RMs, it is necessary to design smart and proficient health monitoring and FD systems. Incipient fault diagnosis provides a minimum of consequences for the consistency of rotating machines, while it is very easy and simple to handle. However, the features of incipient fault are not highly reliable, and predicting the micro-fault is more difficult than a typical fault [5].

Several fault diagnoses models were introduced, and they are categorized into three classes: the data-driven method, quantitative model-based method, and qualitative model-based method [6]. As the difficulty of the current process increases, it becomes increasingly difficult to construct mathematical models that efficiently capture system dynamic behavior [7]. Consequently, the data-driven method, which only relies on the data derived from the process, is receiving considerable interest. The primary stage in the data-driven method is feature extraction, in which the processed information is converted into a lower dimension, with more informative data. The artificial neural network (ANN)-based method is an alternative way that has gained considerable attention over the last few years [8]. An artificial neural network is a network of neurons that learn complicated functions over a sequence of non-linear conversions, and, with the emergence of deep learning (DL) methods, it is effectively employed for complicated classification tasks, including speech recognition and image recognition. However, many of the studies used shallow neural networks or neural networks with hierarchical structures. Therefore, the wider possibility of deep neural networks being used to address fault diagnoses has yet to be explored [9,10].

This study introduces a novel sandpiper optimization with an artificial intelligence-enabled fault diagnosis (SPOAI-FD) model for intelligent industrial applications. The proposed SPOAI-FD technique involves the design of a continuous wavelet transform (CWT)-based pre-processing approach, which converts the raw vibration signal into a useful format. Moreover, a bidirectional long short-term memory (BLSTM) model is applied as a classifier, and the Faster SqueezeNet model is employed as a feature extractor. For effectively adjusting the hyperparameter values of the BLSTM, the sandpiper optimization algorithm (SPOA) can be applied. In order to highlight the better performance of the presented model, a comprehensive investigation was conducted, comparing the results against benchmark datasets. The major contributions of the study are as follows.

- An intelligent SPOAI-FD technique comprising pre-processing, Faster SqueezeNet feature extraction, BLSTM classification, and SPOA-based parameter tuning for fault diagnosis is presented. To the best of our knowledge, the SPOAI-FD technique has never been presented in the literature.
- Employ the Faster SqueezeNet model for feature extraction and the BLSTM model for classification.
- Hyperparameter optimization of the BLSTM model using SPOA algorithm using cross-validation helps to boost the predictive outcome of the proposed model for unseen data.

2. Literature Review

Wu et al. [11] designed a CNN for direct learning of the features in the novel vibration signal and Fault Diagnosis (FD). In the study conducted earlier [12], an ensemble transfer CNN, determined by multi-channel signals, was presented. Primarily, a sequence of the source CNN was changed with stochastic pooling, whereas the Leaky ReLU (LReLU) was pre-trained to utilize the multichannel signals. Secondarily, the learned parameter data of all the individuals' source CNN was transmitted to initialize the equivalent target CNN after fine-tuning with some of the target-trained instances. At last, a novel decision fusion approach was planned for flexible fusion of all the individuals' target CNN to obtain the detailed outcome.

In the literature [13], a new FD approach was proposed based on Max-Relevance Min-Redundancy (mRMR) and Improved Multiscale Dispersion Entropy (IMDE). The mRMR technique was employed for automatic selection of the sensitive features from the candidate multi-scale features without any prior data. At last, the sensitive feature vector was set after which the normalized treatment was recorded. The ELM technique was used to train the intelligent analysis method which produced FD outcomes. In the study conducted earlier [14], an FD technique was presented based on DCNN and SVM techniques. Being a data-driven DL approach, the DCNN technique was executed in this study to extract the fault feature automatically. The fault-feature data was removed adaptively based on the minute variances from the local fault signal.

Chen et al. [15] examined a data-driven intelligent FD technique for RM based on a novel Continuous Wavelet Transform-Local Binary CNN (CWT-LBCNN) technique. The presented technique created an end-to-end analysis process without any need for manual extraction of the features. Using the feed and the input vibration signals, the features were taken adaptably, and fault states of the RM were analyzed automatically. Dibaj et al. [16] presented a novel end-to-end FD technique using the fine-tuned VMD and CNN mechanisms. An essential proposal is that CNN can be trained only using healthy and single fault data sets, whereas compound faults data from the training phase cannot be utilized. During the testing phase of CNN technique, the intelligent technique alarmed an untrained compound faults' state, when the developed probability of the CNN outcomes fulfills a group of probabilistic states. In the study conducted earlier [17], a novel technique was proposed based on RNN to identify the fault types from the RM. In this study, 1D time-series vibration signal was initially converted into 2D images. Next, the GRU was established to exploit the temporal data of time-series data and learn the representative features of the created images. Last, the MLP was utilized to execute the fault detection.

Although DL-based fault diagnosis methods exist abundantly in the literature, there is still a need to design an automated fault diagnosis model with an enhanced detection rate. As the increasing number of DL models can result in model overfitting, optimal hyperparameter selection becomes essential. Since the trial-and-error method for hyperparameter tuning is a tedious and erroneous process, metaheuristic algorithms are applied. Therefore, in this work, SPOA algorithm is deployed for the parameter selection of the BLSTM model.

3. The Proposed Model

In this study, a new SPOAI-FD technique was developed to detect the faults in the rotating machinery. The proposed method includes a few sub-processes, such as the CWT-based pre-processing, Faster SqueezeNet feature extraction, BLSTM classifier, and the SPOA-based hyperparameter optimization. The SPOA method can be applied to modify the hyperparameter values of the BLSTM, thus producing the improved classification performance.

3.1. Data Pre-Processing

Rotating Machinery is a procedure that involves different types of rotating loads and speeds. In order to implement the fault detection process in some operational states and to train the model, the vibration signal of the machine from the entire load and its speed range are essential [18]. At the beginning, the vibration signal was collected from the rotating speed dataset. Particularly, the rotating speed from the trained instance was assumed to be constant, since it is collected when the machinery is at a constant functioning procedure. The CWT retains and generates the localized design of STFT. The CWT of the signal $x(t)$ is defined as the convolutional of signal $x(t)$ that makes use of the wavelet function, $\Psi_{a,b}(t)$.

During this method, the CWT was applied to decompose the data in the scale range of 1 to l in which l generally refers to a higher value or equal to $2q$

$$C_a(k) = \int x(t) \cdot \overline{\Psi}_{a,b}(t)dt \tag{1}$$

where $C_a(a = 1, 2, 3, \ldots, l)$ refers to the wavelet co-efficient of $x(t)$ in *ath* scale and $\overline{\Psi}_{a,b}(t)$ implies the complex conjugate. The CWT produces the co-efficient on several portions of the signals from the scaling factor. When employing the wavelet co-efficient, a signal from the time frequency field is directly projected by the 2D images. The graph of the wavelet co-efficient generates the CWTS.

After obtaining all the wavelet coefficients from a matrix $P = [C_1, C_2, \ldots, C_l]$, it can be changed for gray matrix, i.e., P_{new} as given below.

$$P_{new}(i,j) = \left[\frac{P(i,j) - p_{\min}}{p_{\max} - p_{\min}} \times 255 + \frac{1}{2}\right] \tag{2}$$

where $p_{\min}$ and $p_{\max}$ denote the lower and the higher elements of P, correspondingly. The value of the components from P_{new} refers to the gray value in the order of 0 to 255. Therefore, P_{new} defines the CWTS of the novel signals.

3.2. Feature Extraction: Faster SqueezeNet Model

In this study, Fast SqueezeNet was presented to improve the performance and accuracy of electronic element classification. To avoid over-fitting, the authors added residual and BatchNorm frameworks. Simultaneously, concat was also utilized like DenseNet to connect distinct layers for the purpose of improving the expressiveness of the initial layer. Fast SqueezeNet comprises three block layers, such as a BatchNorm layer, a global average pooling layer, and four convolution layers. Fast SqueezeNet is primarily enhanced as given below.

(1) The current study imitated the DenseNet architecture and presents a distinct connection mode for additional improvement of the data flow among the layers [19,20]. This comprises a fire module and a pooling layer. At last, the two concat layers were also interconnected to the following convolutional layer.

The existing layer obtains each feature map of the preceding layer and employs $x_0, \ldots, x_{l-1}$ as the input; next x_l is demonstrated as given below.

$$x_l = H_l\left([x_0, x_1, \ldots, x_{l-1}]\right), \tag{3}$$

where $[x_0, x_1, \ldots, x_{l-1}]$ represents the connection of the feature graph created in the layer, $0, 1, \ldots, l-1$ and H_l denotes the concatenate of several inputs. Further, x_0 denotes the max pooling layer, x_1 signifies the Fire layer and x_l represents the concat layer.

In the absence of extreme improvement in the number of network parameters, the efficiency can be optimized in its earlier stage. Simultaneously, any two-layer networks can transfer the data directly.

(2) In order to ensure a good network convergence, the ResNet architecture was thoroughly learnt, and distinct components were presented with a fire module and a pooling layer. At last, two layers were added and interconnected to the following convolution layer. Generally, the fundamental mapping is represented as $H(x)$. Consider the stacked non-linear layer to fit other mappings of $F(x) := H(x) - x$. The original mapping is reorganized into $F(x) + x$. $F(x) + x$ is realized as a structure named as shortcut connection. It utilizes the residual architecture of ResNet to resolve issues, such as gradient degradation and disappearing without improving the amount of network variables.

3.3. Fault Detection and Classification: BLSTM Model

BLSTM is an integration of LSTM and Bidirectional RNN (BRNN) techniques. RNN is a significant development of ANN and is used for processing the sequence and time-series data. It has a huge benefit in terms of encrypting the dependency amongst the input values. However, in the case of long data sequence, RNN creates vanishing and exploding states besides their gradient [21]. Next, LSTM was generated to overcome the long-term problem of RNN. LSTM comprises output, input, and forget gates. However, LSTM and RNN can obtain data from the preceding context, thus it achieved heavy improvement with the help of BRNN. It can manage two datasets from the front and back. Figure 1 shows the BLSTM structure.

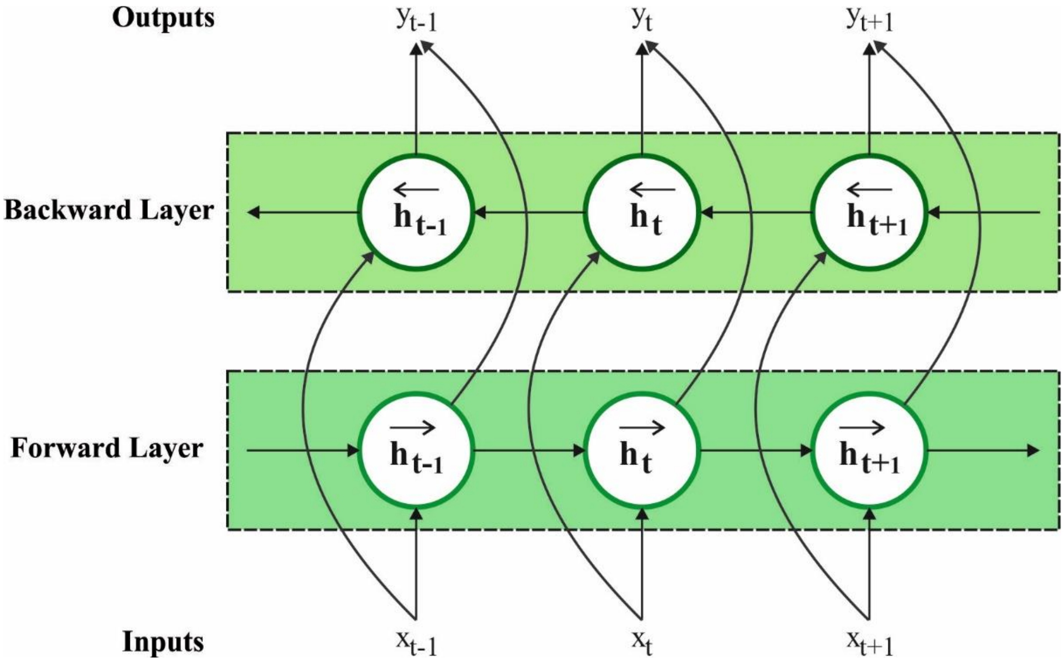

Figure 1. BLSTM structure.

The integration of BRNN and LSTM results in the formation of BLSTM. Hence, the benefits of both LSTM in terms of cell memory storage and BRNN in terms of accessing the data in context before and after making the BLSTM can be exploited. BLSTM has the benefit of LSTM with feedback to the following layer. However, BLSTM can manage the information alternatively with a dependency on longer range. The forwarded function of the BLSTM with an input of L unit and H as the amount of hidden states is estimated as follows.

$$a_h^t = \sum_{l=1}^{L} x_l^t w_{lh} +, \sum_{h'=1,t>0}^{H} b_{h'}^{t-1} w_{h'h} \tag{4}$$

$$b_h^t = \boldsymbol{\Theta_h}\left(a_h^t\right) \tag{5}$$

where $\{x^{\mathrm{t}}\}$ represents a sequence input, a_h^{t} denotes the input of the network to LSTM of the unit h at t time. Θ_h denotes the activation function of the h hidden unit. w_{lh} shows the weight of the input l towards h. The activation function of h, at time t, is represented by a b_h^t. $w_{h,h}$ indicates the weight of h hidden units toward the h' hidden units. The backward estimation of the BLSTM is determined as follows.

$$\frac{\delta O}{\delta w_{hk}} = \sum_{t=1}^{T} \frac{\delta O}{\delta a_h^t} b_h^t$$

$$\frac{\delta O}{\delta a_h^t} = \Theta_h' \left(a_h^t \sum_{k=1}^{K} \frac{\delta O}{\delta a_h^t} w_{hk} + \sum_{h'=1,t>0}^{H} \frac{\delta O}{\delta a_{h'}^{t+1}} w_{hh'} \right)$$

where O represents the objective function with unit K output.

3.4. Hyperparameter Optimization

To select the hyperparameter values of the BLSTM, the SPOA model was utilized. Sandpipers are seabirds that exist all over the globe. They commonly live in colonies [22] and possess the highest level of intelligence in finding and attacking prey. The important characteristics of the sandpiper are its migratory power and attacking behavior. Such characteristics were utilized to derive the SPOA. The mathematical formulation of these behaviors is offered herewith.

3.4.1. Exploration Process

The SPOA investigated the collection of the sandpipers that move through different positions at the time of migration. Here, the sandpipers need to fulfill the three criteria given below.

Collision evading: An extra parameter C_A is applied to compute the newly attained searching agents and to avoid the collisions among the nearby sandpipers.

$$\vec{C_{sp}} = C_A \times \vec{P_{sp}}(z) \tag{6}$$

where $\vec{C_{sp}}$ designates the location of the searching agents that do not collide with other searching agents, $\vec{P_{sp}}$ denotes the existing location of the searching agent, z specifies the existing round, and C_A describes the motion of the searching agents in the searching region.

$$C_A = C_f - \left(z \times \left(C_f / \text{Max}_{iterations}\right)\right)$$

where

$$z = 0, 1, 2, \ldots, \text{Max}_{iterations} \tag{7}$$

where C_f implies a control frequency for C_A parameter adjustment that is linearly decreased from C_f to 0. The fitness function can be represented as a process that determines the population and derives a score. In addition to avoiding the collisions, the searching agent converges in the direction of its optimal neighbors.

$$\vec{M_{sp}} = C_B \times \left(\vec{P_{bst}}(z) - \vec{P_{sp}}(z)\right) \tag{8}$$

where $\vec{M_{sp}}$ stipulates the location of the searching agent and $\vec{P_{sp}}$ indicates the optimal searching agent, $\vec{P_{bst}}$. C_B implies an arbitrary parameter that is important for an effective exploration which can be determined using Equation (9).

$$C_B = 0.5 \times R_{and} \tag{9}$$

where R_{and} represents an arbitrary integer that exists in the interval of $[0,1]$. At last, the sandpipers or searching agents upgrade their location with respect to the optimum searching agents.

$$\vec{D_{sp}} = \vec{C_{sp}} + \vec{M_{sp}} \tag{10}$$

where $\vec{D_{sp}}$ states the gap between the searching agents and the optimally fit searching agents.

3.4.2. Exploitation Process

At the time of migration, the sandpiper modifies its speed and the angle of attacks in a seamless fashion. It utilizes its wings to increase its height. The sandpiper follows a spiral formation at the time of prey attack. These characteristics in the 3D plane are defined as follows.

$$x' = R_{adius} \times sin(i) \tag{11}$$

$$y' = R_{adius} \times cos(i) \tag{12}$$

$$z' = R_{adius} \times i \tag{13}$$

$$r = u \times e^{kv} \tag{14}$$

where R_{adius} indicates the radius of every turn of the spiral, i implies a variable that exists in the interval of $[0 \leq k \leq 2\pi]$. u and v denote the constant values that describe the spiral shapes, and e symbolizes the base of the natural logarithm. The values of u and v are denoted as 1. If the value exceeds 1, then the spiral shape becomes difficult. Consequently, the location of the search agent is upgraded as follows.

$$\overrightarrow{P_{sp}}(z) = \left(\overrightarrow{D_{sp}} \times \left(x' + y' + z'\right)\right) \times \overrightarrow{P_{bst}}(z) \tag{15}$$

where $\overrightarrow{P_{sp}}(z)$ upgrades the position of another searching agent and the store's best solution.

The SPOA approach develops an FF to accomplish a better classification accuracy. It describes a positive integer to demonstrate the optimal performance of the candidate solution. In such cases, the minimized classification error rate is considered to be the FF as given below. The optimum solution is a lesser error rate, and the worst solution reaches a maximum error rate.

$$\begin{aligned} fitness(x_i) &= Classifier\ Error\ Rate(x_i) \\ &= \frac{number\ of\ misclassified\ instances}{Total\ number\ of\ instances} * 100 \end{aligned} \tag{16}$$

4. Experimental Validation

The proposed SPOAI-FD technique was experimentally validated by means of automotive gearbox and bearing fault datasets [23,24]. The former dataset comprises seven classes whereas the latter dataset includes a total of 10 classes. The first dataset holds seven types of health statuses, such as outer race bearing fault, minor-chipped gear fault, missed tooth gear fault, and three types of compound faults (Normal, Minor-chipped tooth, Missing tooth (0.2 mm), and the Missing tooth (2 mm)). The second dataset has both normal as well as fault data. The bearing fault has a few types, such as the Inner race (IF), Outer race (OF), and Ball faults (BF). Therefore, 10 kinds of bearing health status under varying loads were studied. The details of the dataset are shown in Table 1.

Table 2 and Figure 2 show the accuracy examination results achieved by the proposed SPOAI-FD model on gearbox dataset under distinct classes. The results exhibit that the proposed SPOAI-FD method attained better accuracy values for every run. For example, with Class 1, the proposed SPOAI-FD method obtained the accuracy values 0.9941, 0.9936, 0.9937, 0.9945, and 0.9940 correspondingly. Similarly, with Class 2, the presented SPOAI-FD technique attained the accuracy values 0.9915, 0.9927, 0.9935, 0.9914, and 0.9906, correspondingly. Likewise, with Class 3, the SPOAI-FD approach produced the following accuracy values, 0.9920, 0.9932, 0.9926, 0.99145, and 0.9921, correspondingly. Simultaneously, with Class 7, the proposed SPOAI-FD method obtained the accuracy values 0.9946, 0.9903, 0.9924, 0.9945, and 0.9921, correspondingly.

Table 1. Dataset details.

Dataset	Class Number	Class Label
Dataset-I	Class 1	Outer Race Bearing Fault
	Class 2	Minor Chipped Gear Fault
	Class 3	Missed Tooth Gear Fault
	Class 4	Normal
	Class 5	Minor chipped tooth
	Class 6	Missing tooth (0.2 mm)
	Class 7	Missing tooth (2 mm)
Dataset-II	Class 1	Outer Race Bearing Fault
	Class 2	Minor Chipped Gear Fault
	Class 3	Missed Tooth Gear Fault
	Class 4	Normal
	Class 5	Minor chipped tooth
	Class 6	Missing tooth (0.2 mm)
	Class 7	Missing tooth (2 mm)
	Class 8	Inner race (IF)
	Class 9	Outer race (OF)
	Class 10	Ball faults (BF)

Table 2. Analytical results of the SPOAI-FD technique under Gearbox Dataset.

Class Label	Run-1	Run-2	Run-3	Run-4	Run-5	Average
Class-1	0.9941	0.9936	0.9937	0.9945	0.9940	0.9940
Class-2	0.9915	0.9927	0.9935	0.9914	0.9906	0.9919
Class-3	0.9920	0.9932	0.9926	0.9914	0.9921	0.9923
Class-4	0.9945	0.9909	0.9934	0.9946	0.9927	0.9932
Class-5	0.9941	0.9921	0.9925	0.9910	0.9925	0.9924
Class-6	0.9947	0.9908	0.9928	0.9944	0.9910	0.9927
Class-7	0.9946	0.9903	0.9924	0.9945	0.9921	0.9928

Table 3 and Figure 3 show the comparative accuracy analysis results attained by the proposed SPOAI-FD and other recent approaches [25,26] on gearbox datasets. The result demonstrate that the proposed SPOAI-FD method achieved better accuracy values than the rest of the methods under all the classes. For example, with Class 1, the SPOAI-FD model accomplished a high accuracy of 0.9940 whereas the FFTKNN, FFTSVM, FFTDBN, FFTSAE, CNN, CNN2 and the IIFD-SOIR technique obtained the least accuracy values, such as 0.8364, 0.9886, 0.9746, 0.9855, 0.9885, 0.9881, and 0.9876, correspondingly. Simultaneously, with Class 2, the proposed SPOAI-FD model gained an increased accuracy of 0.9919 although the existing models, such as FFTKNN, FFTSVM, FFTDBN, FFTSAE, CNN, CNN2, and IIFD-SOIR techniques resulted in low accuracy values, such as 0.9195, 0.9801, 0.9693, 0.9836, 0.9821, 0.9797, and 0.9852, correspondingly. Concurrently, with Class 3, the presented SPOAI-FD model accomplished a high accuracy of 0.9923 whereas the other models, such as FFTKNN, FFTSVM, FFTDBN, FFTSAE, CNN, CNN2, and IIFD-SOIR techniques, obtained the least accuracy values, such as 0.9811, 0.9837, 0.9777, 0.9802, 0.9684, 0.9676, and 0.9811, correspondingly.

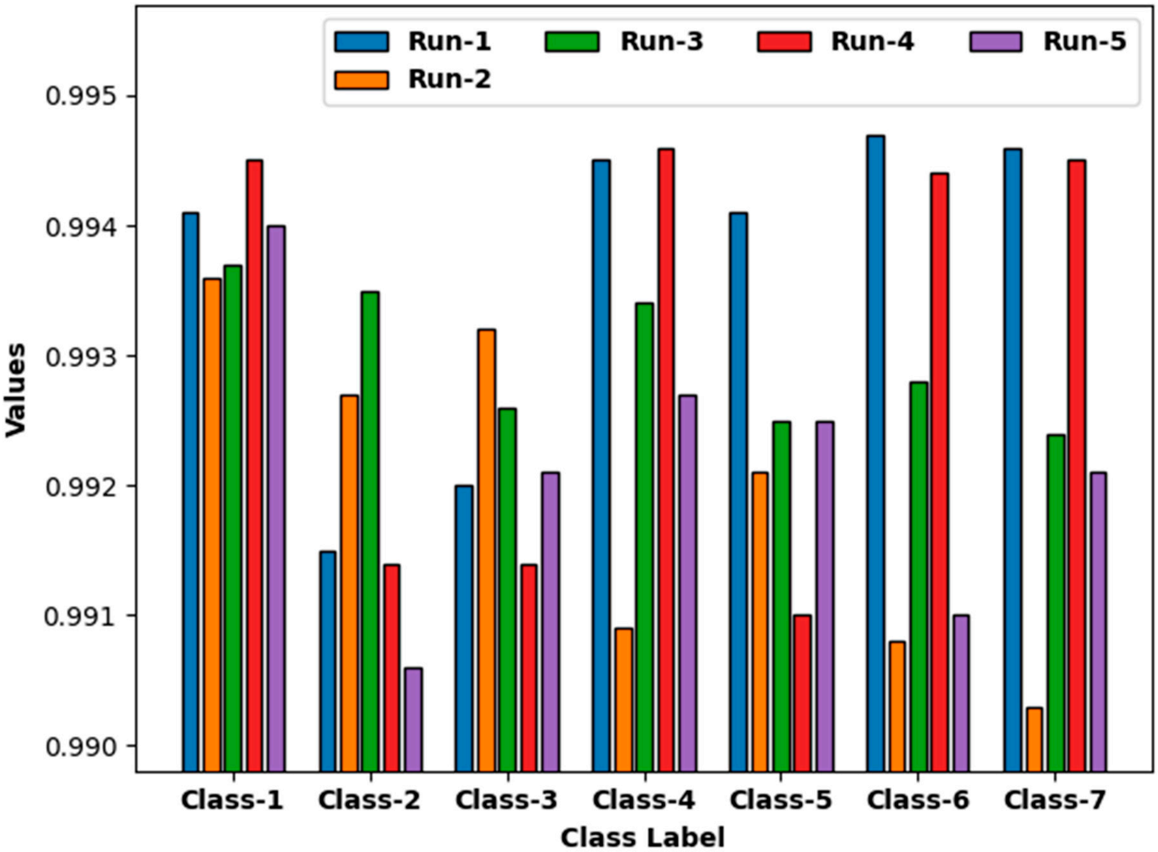

Figure 2. Analytical results of the SPOAI-FD technique under Gearbox dataset.

Table 3. Analytical results of Each Fault Class of different methods on gearbox dataset.

Methods	Gearbox Dataset							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**Average**
FFTKNN	0.8364	0.9195	0.9811	0.9869	0.8678	0.6768	0.6697	0.8483
FFTSVM	0.9886	0.9801	0.9837	0.9850	0.9801	0.9609	0.8685	0.9638
FFTDBN	0.9746	0.9693	0.9777	0.9755	0.9792	0.9468	0.9385	0.9659
FFTSAE	0.9855	0.9836	0.9802	0.9877	0.9854	0.9677	0.9558	0.9780
CNN	0.9885	0.9821	0.9684	0.9827	0.9876	0.9815	0.8857	0.9681
CNN2	0.9881	0.9797	0.9676	0.9861	0.9832	0.9577	0.9063	0.9670
IIFD-SOIR	0.9876	0.9852	0.9811	0.9855	0.9862	0.9823	0.9771	0.9836
SPOAI-FD	0.9940	0.9919	0.9923	0.9932	0.9924	0.9927	0.9928	0.9940

The accuracy investigation outcomes, obtained by the proposed SPOAI-FD approach, under gearbox dataset, are shown in Figure 4. The result demonstrates that the proposed SPOAI-FD technique gained an increment in its validation accuracy compared to the training accuracy. Furthermore, it is obvious that the accuracy value becomes saturated based on the count of epochs.

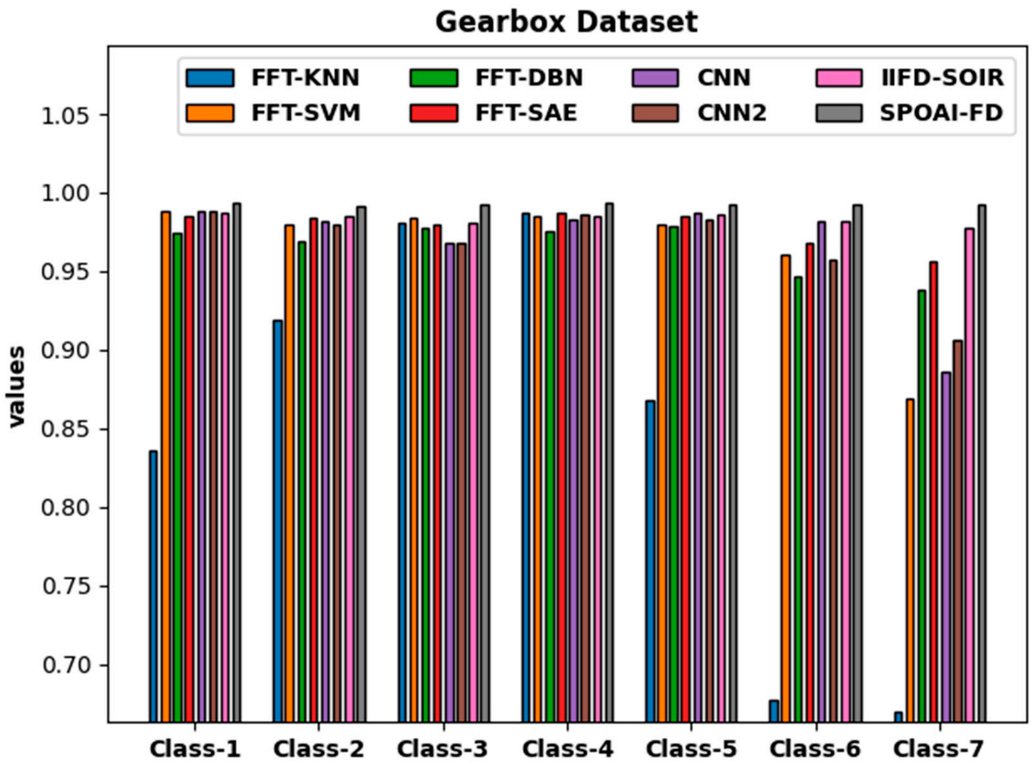

Figure 3. Accuracy analysis results of SPOAI-FD technique under Gearbox dataset.

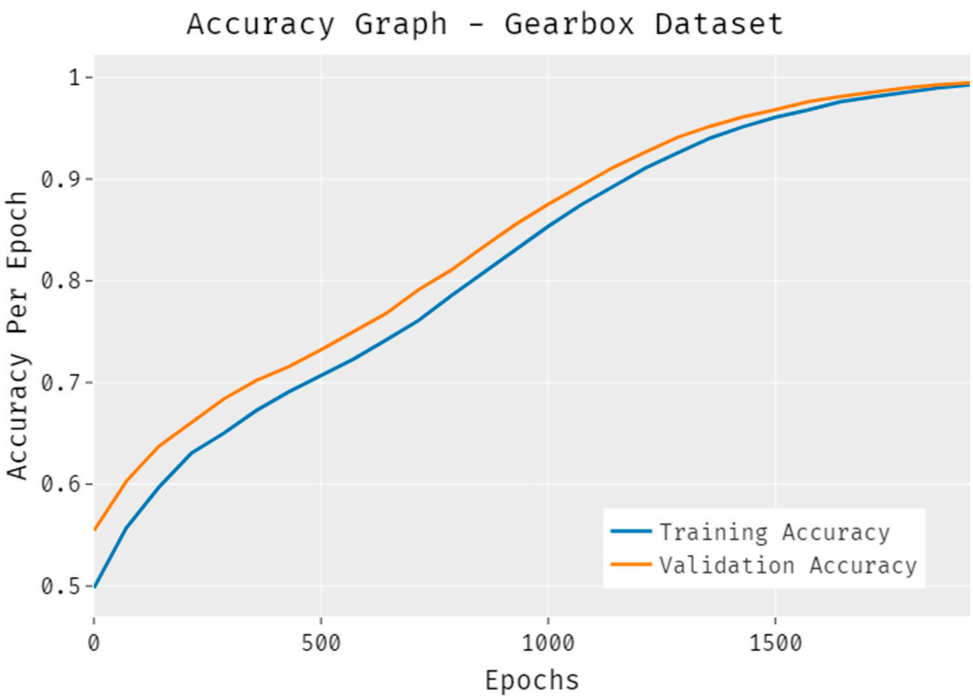

Figure 4. Accuracy analysis graph of the SPOAI-FD methodology under Gearbox dataset.

The loss investigation outcomes of the SPOAI-FD system, under gearbox dataset, are illustrated in Figure 5. The figure reveals that the proposed SPOAI-FD method significantly reduced the validation loss over training loss. Additionally, it is noted that the loss value becomes saturated with the count of epochs.

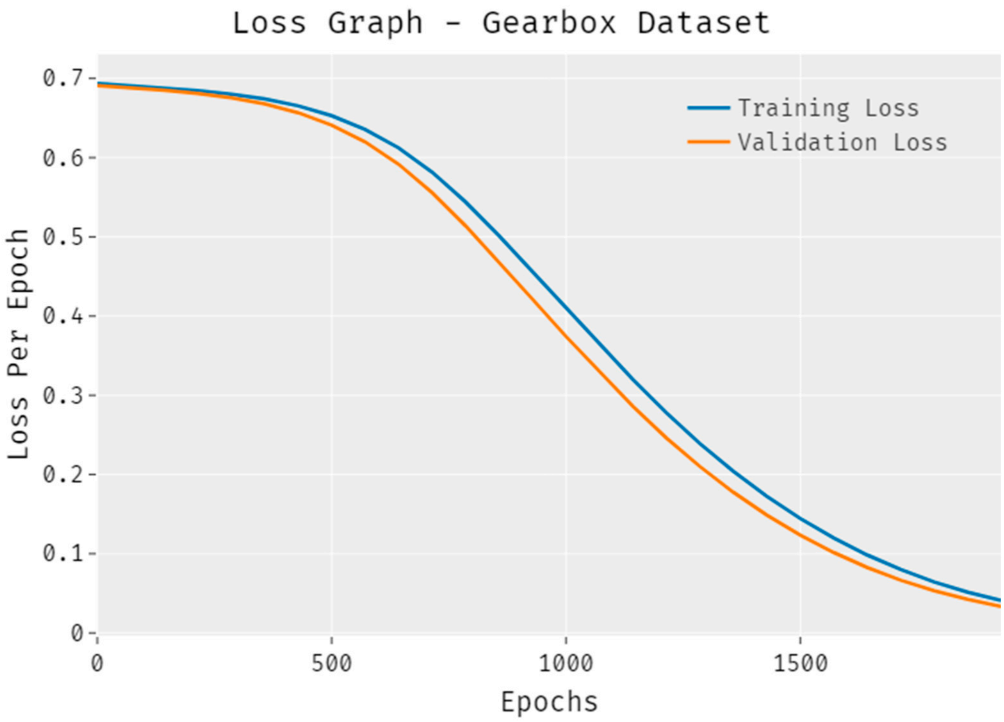

Figure 5. Loss analysis results of SPOAI-FD technique under Gearbox dataset.

Table 4 and Figure 6 depict the accuracy investigation outcomes attained by the proposed SPOAI-FD approach on bearing dataset under diverse classes. The experimental values demonstrate that the proposed SPOAI-FD approach achieved improved accuracy values under all the runs. For example, with Class 1, the SPOAI-FD methodology gained the accuracy values 0.9922, 0.9939, 0.9945, 0.9906, and 0.9906, correspondingly. Likewise, with Class 2, the SPOAI-FD algorithm yielded the following accuracy values, 0.9946, 0.9940, 0.9920, 0.9910, and 0.9922, correspondingly. Moreover, with Class 3, the SPOAI-FD approach accomplished the accuracy values 0.9911, 0.9927, 0.9941, 0.9943, and 0.9937, correspondingly. At last, with class 7, the proposed SPOAI-FD technique reached the accuracy values 0.9923, 0.9927, 0.9933, 0.9930, and 0.9901, correspondingly.

Table 4. Analytical results of SPOAI-FD technique under Bearing dataset.

Class Label	Run-1	Run-2	Run-3	Run-4	Run-5	Average
Class-1	0.9922	0.9939	0.9945	0.9906	0.9906	0.9924
Class-2	0.9946	0.9940	0.9920	0.9910	0.9922	0.9928
Class-3	0.9911	0.9927	0.9941	0.9943	0.9937	0.9932
Class-4	0.9932	0.9945	0.9920	0.9917	0.9903	0.9923
Class-5	0.9926	0.9936	0.9944	0.9940	0.9906	0.9930
Class-6	0.9902	0.9949	0.9934	0.9935	0.9946	0.9933
Class-7	0.9923	0.9927	0.9933	0.9930	0.9901	0.9923
Class-8	0.9946	0.9917	0.9934	0.9946	0.9950	0.9939
Class-9	0.9920	0.9925	0.9924	0.9912	0.9914	0.9919
Class-10	0.9920	0.9908	0.9933	0.9905	0.9911	0.9915

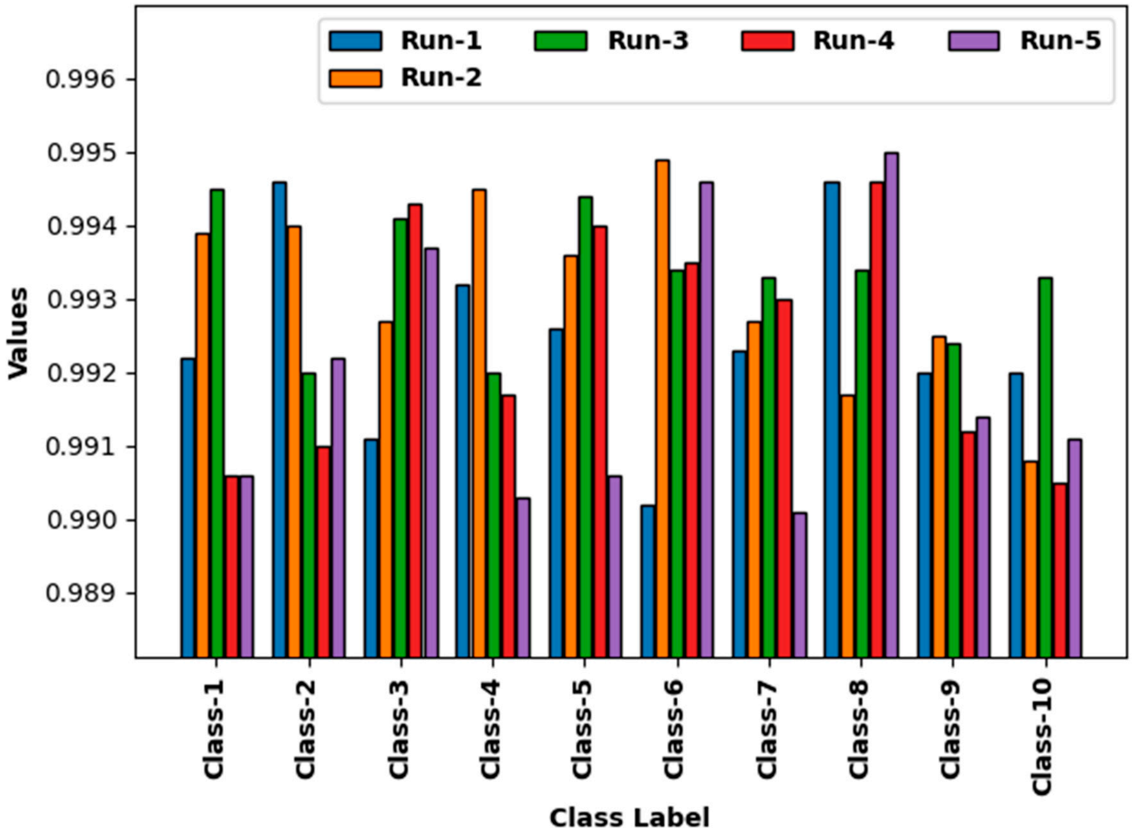

Figure 6. Accuracy analysis of SPOAI-FD technique under Bearing dataset.

Table 5 and Figure 7 portray the brief comparison study outcomes accomplished by the proposed SPOAI-FD and other recent approaches on bearing dataset. The simulation values depict that the SPOAI-FD system gained enhanced accuracy values over the rest of the methods under all the classes. For example, with Class 1, the SPOAI-FD model gained increased accuracy values, such as 0.9924, while FFTKNN, FFTSVM, FFTDBN, FFTSAE, CNN, CNN2, and IIFD-SOIR techniques achieved the least accuracy values, such as 0.9706, 0.9892, 0.9815, 0.9860, 0.9791, 0.9833, and 0.9870, correspondingly. Simultaneously, with Class 2, the proposed SPOAI-FD method yielded an enhanced accuracy of 0.9928. However, the other models, such as FFTKNN, FFTSVM, FFTDBN, FFTSAE, CNN, CNN2, and IIFD-SOIR techniques, achieved minimal accuracy values, such as 0.9582, 0.9446, 0.9741, 0.9691, 0.9382, 0.9127, and 0.9771, correspondingly. Concurrently, with Class 3, the SPOAI-FD model accomplished a maximum accuracy of 0.9932. However, the FFTKNN, FFTSVM, FFTDBN, FFTSAE, CNN, CNN2, and the IIFD-SOIR techniques obtained the least accuracy values, such as 0.9721, 0.9862, 0.9836, 0.9864, 0.9792, 0.9934, and 0.9867, correspondingly.

The accuracy investigation outcomes of the SPOAI-FD approach under bearing dataset are depicted in Figure 8. The result exhibits that the proposed SPOAI-FD methodology gained a better validation accuracy compared to the training accuracy. Additionally, it is noticeable that the accuracy value becomes saturated with the count of epochs.

The loss investigation results of the SPOAI-FD algorithm under bearing dataset are depicted in Figure 9. The figure reveals that the proposed SPOAI-FD method achieved a reduction in the validation loss than the training loss. Additionally, it is to be noted that the loss value becomes saturated with the count of epochs.

Table 5. Analytical results of Each Fault Class of different methods on bearing dataset.

Methods	Bearing Dataset										
	1	2	3	4	5	6	7	8	9	10	Average
FFTKNN	0.9706	0.9582	0.9721	0.9432	0.9563	0.9655	0.9788	0.9423	0.9793	0.9608	0.9627
FFTSVM	0.9892	0.9446	0.9862	0.9839	0.9772	0.9073	0.9869	0.9407	0.9899	0.8602	0.9566
FFTDBN	0.9815	0.9741	0.9836	0.9746	0.9672	0.9573	0.9759	0.9327	0.9768	0.9434	0.9667
FFTSAE	0.9860	0.9691	0.9864	0.9700	0.9716	0.9587	0.9829	0.9312	0.9876	0.9366	0.9680
CNN	0.9791	0.9382	0.9792	0.9765	0.9836	0.9757	0.9887	0.9635	0.9806	0.9829	0.9748
CNN2	0.9833	0.9127	0.9834	0.9786	0.9598	0.9759	0.9833	0.9302	0.9790	0.9849	0.9671
IIFD-SOIR	0.9870	0.9771	0.9867	0.9807	0.9877	0.9846	0.9824	0.9718	0.9814	0.9860	0.9825
SPOAI-FD	0.9924	0.9928	0.9932	0.9923	0.993	0.9933	0.9923	0.9939	0.9919	0.9915	0.9927

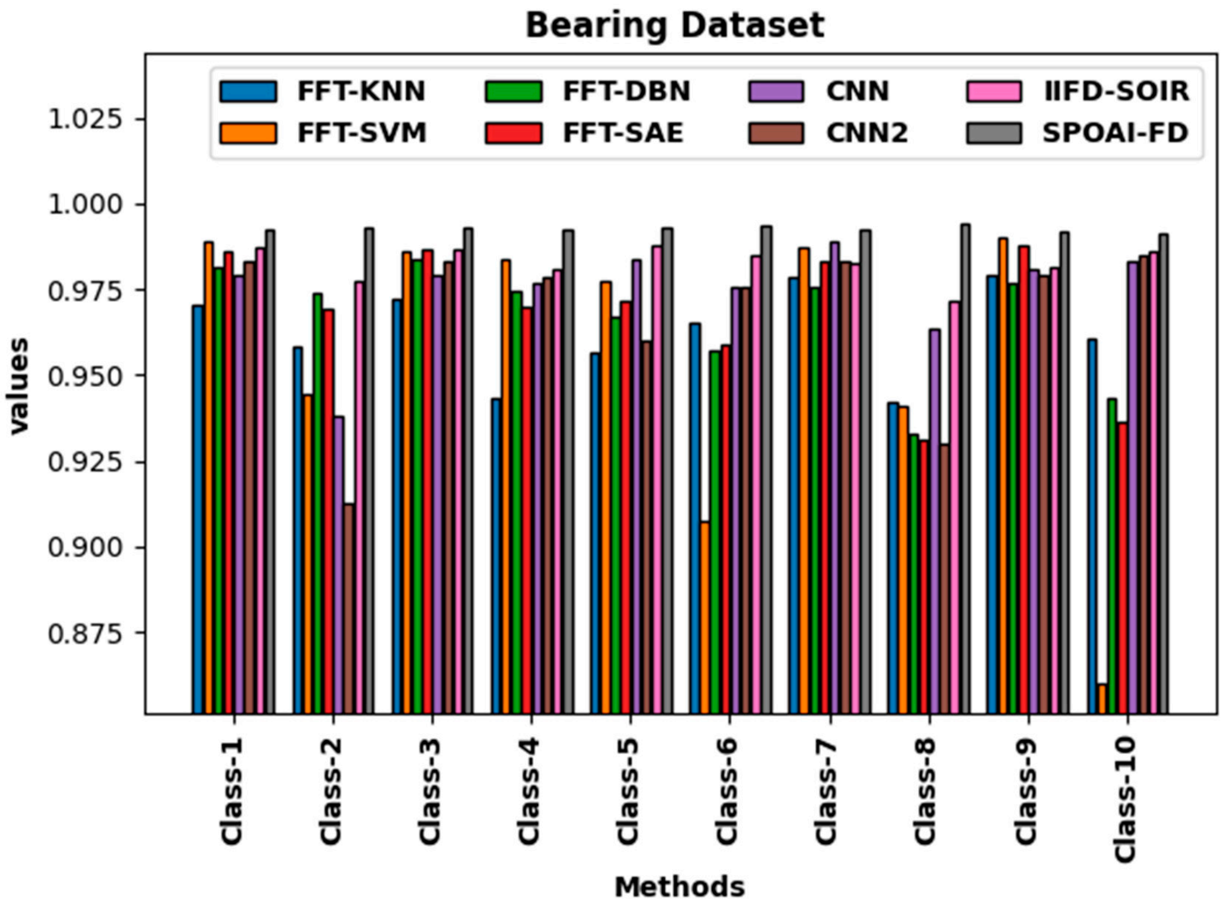

Figure 7. Analytical results of SPOAI-FD technique under Bearing dataset.

Table 6 provides the overall average analysis results of the SPOAI-FD and other recent methodologies. Figure 10 offers the comparative average accuracy analysis outcomes of SPOAI-FD approach and other methods on gearbox dataset. The results show that the SPOAI-FD technique outperformed all other methods with maximum training and testing accuracies.

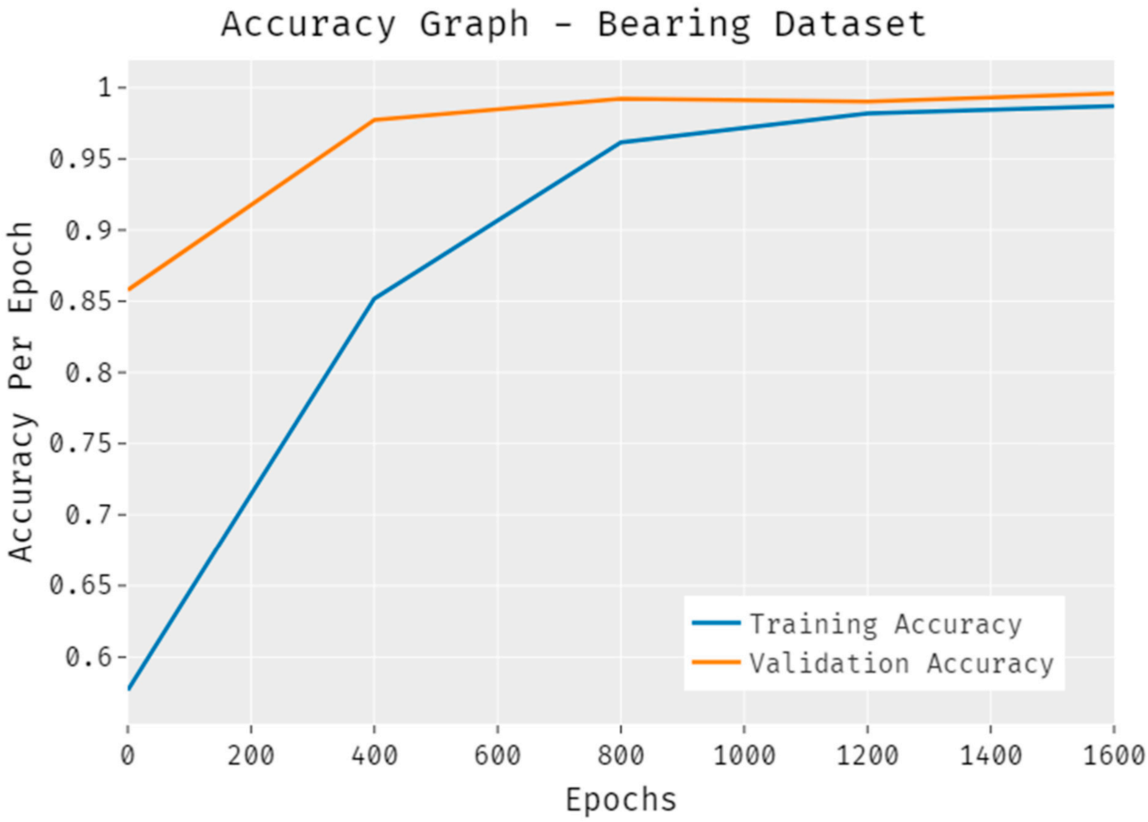

Figure 8. Accuracy analysis outcomes of SPOAI-FD technique under Bearing dataset.

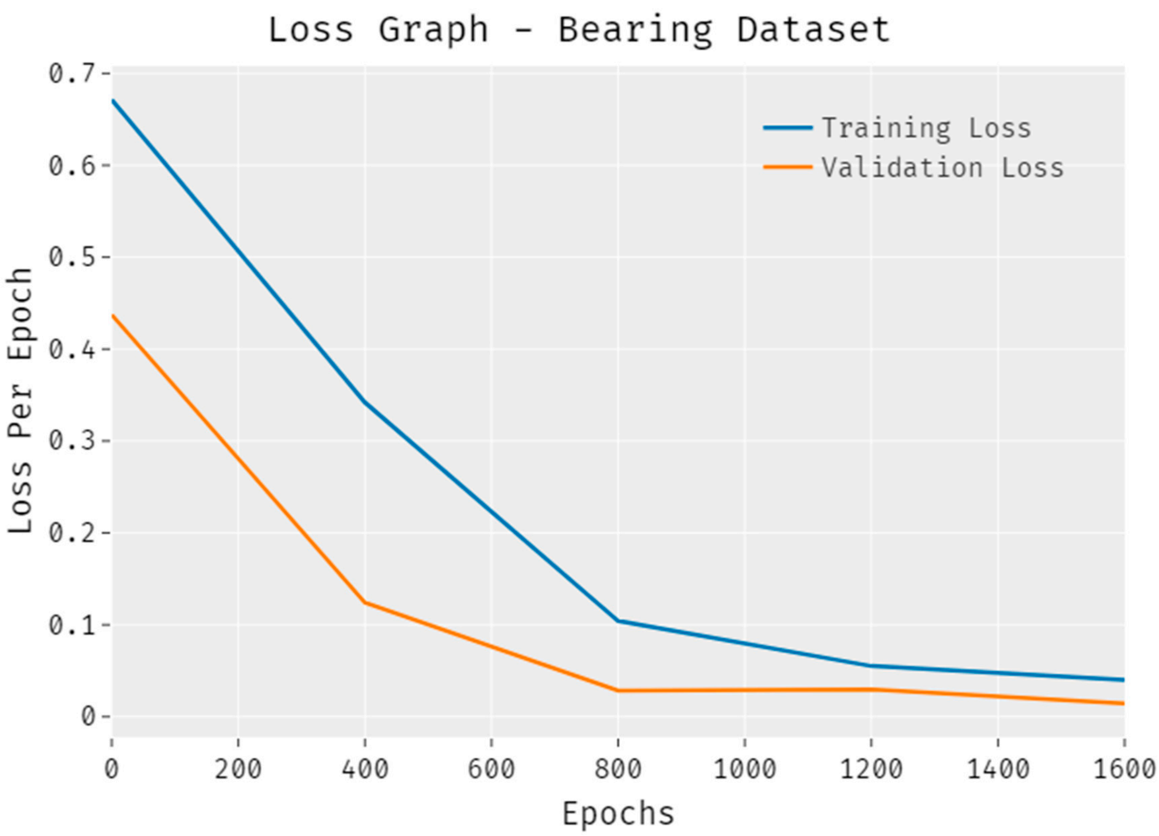

Figure 9. Loss analysis outcomes of SPOAI-FD technique under Bearing dataset.

Table 6. Training and testing accuracies of the proposed SPOAI-FD approach and other recent methods.

Method	Gearbox Dataset		Bearing Dataset	
	Training	Testing	Training	Testing
FFTKNN	0.8567	0.8483	0.9754	0.9627
FFTSVM	0.9753	0.9638	0.9622	0.9566
FFTDBN	0.9711	0.9659	0.9814	0.9667
FFTSAE	0.9864	0.9780	0.9740	0.9680
CNN	0.9764	0.9681	0.9789	0.9748
CNN2	0.9726	0.9670	0.9768	0.9671
IIFD-SOIR	0.9899	0.9836	0.9890	0.9825
SPOAI-FD	0.9960	0.9940	0.9951	0.9927

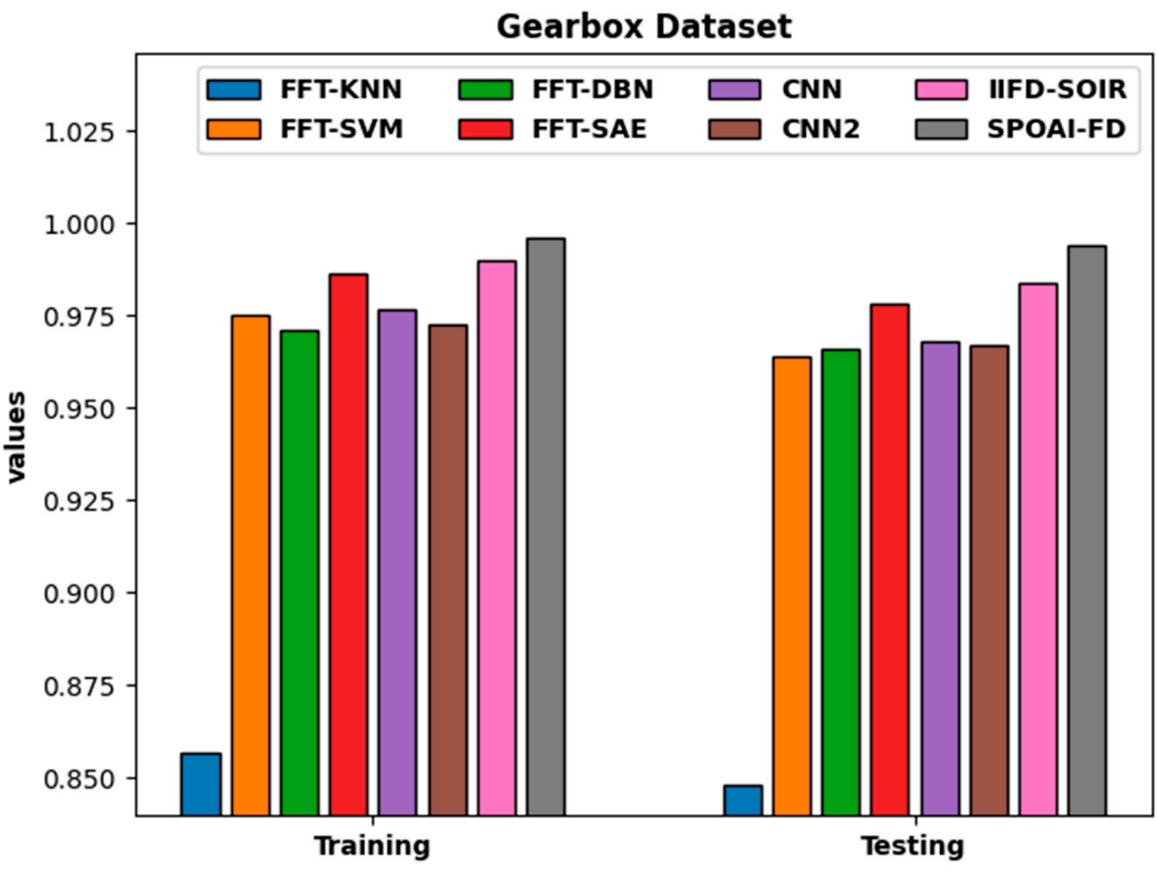

Figure 10. Comparative analysis outcomes of the SPOAI-FD technique under Gearbox dataset.

For example, with respect to training accuracy, the SPOAI-FD approach reached a maximum training accuracy of 0.9960. However, other methods such as the FFTKNN, FFTSVM, FFTDBN, FFTSAE, CNN, CNN-2, and IIFD-SOIR algorithms gained lesser training accuracies, such as 0.8567, 0.9753, 0.9711, 0.9864, 0.9764, 0.9726, and 0.9899, correspondingly.

Figure 11 illustrates the detailed average accuracy analysis outcomes achieved by the proposed SPOAI-FD algorithm and other techniques on bearing dataset. The results obtained showcase that the SPOAI-FD method surpassed all other existing techniques with maximum training and testing accuracies. For instance, the proposed SPOAI-FD system reached an increased training accuracy of 0.99510, whereas the other methods such as FFTKNN, FFTSVM, FFTDBN, FFTSAE, CNN, CNN-2, and IIFD-SOIR techniques produced the least training accuracy values, such as 0.9754, 0.9622, 0.9814, 0.9740, 0.9789, 0.9768, and 0.9890, correspondingly. By observing the abovementioned outcomes, it can be inferred that the proposed SPOAI-FD system has an enhanced fault diagnosis efficiency over other methods.

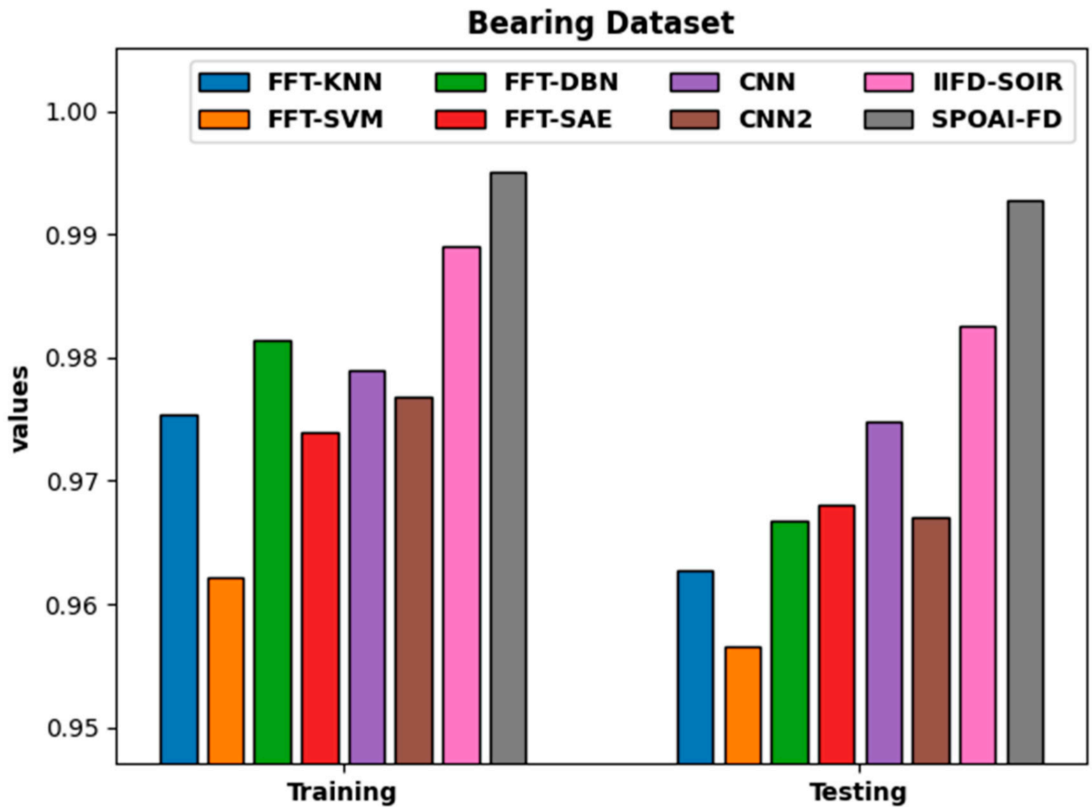

Figure 11. Comparative analysis of the SPOAI-FD technique under Bearing dataset.

5. Conclusions

In the current study, a new SPOAI-FD technique has been designed to detect the faults in RMs. The proposed algorithm includes numerous sub-processes, such as the CWT-based pre-processing, Faster SqueezeNet feature extraction, BLSTM classifier and SPOA-based hyperparameter optimization. The SPOA has been employed to modify the hyperparameter values of the BLSTM model, thus producing better classification performance. In order to highlight the better performance of the presented model, a comprehensive examination was conducted using automotive gearbox and bearing fault datasets. The extensive comparison study outcomes highlighted the supremacy of the proposed SPOAI-FD algorithm over other recent approaches, since the former achieved the maximum accuracy values, such as 0.9960 and 0.9951, on gearbox and bearing datasets, respectively. In the future, hybrid DL mechanisms can be developed to enhance the classifier outcomes on fault detection process.

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

Funding: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R140), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: 22UQU4340237DSR59.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable to this article as no datasets were generated during the current study.

Conflicts of Interest: The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

References

1. Sodhro, A.H.; Pirbhulal, S.; de Albuquerque, V.H.C. Artificial Intelligence-Driven Mechanism for Edge Computing-Based Industrial Applications. *IEEE Trans. Ind. Inform.* **2019**, *15*, 4235–4243. [CrossRef]
2. Lee, J.; Davari, H.; Singh, J.; Pandhare, V. Industrial Artificial Intelligence for industry 4.0-based manufacturing systems. *Manuf. Lett.* **2018**, *18*, 20–23. [CrossRef]
3. Hao, M.; Li, H.; Luo, X.; Xu, G.; Yang, H.; Liu, S. Efficient and Privacy-Enhanced Federated Learning for Industrial Artificial Intelligence. *IEEE Trans. Ind. Inform.* **2019**, *16*, 6532–6542. [CrossRef]
4. Cioffi, R.; Travaglioni, M.; Piscitelli, G.; Petrillo, A.; De Felice, F. Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions. *Sustainability* **2020**, *12*, 492. [CrossRef]
5. Gao, Y.; Chen, M.; Yu, D. Semi-supervised graph convolutional network and its application in intelligent fault diagnosis of rotating machinery. *Measurement* **2021**, *186*, 110084. [CrossRef]
6. Wei, H.; Zhang, Q.; Shang, M.; Gu, Y. Extreme Learning Machine-based Classifier for Fault Diagnosis of Rotating Machinery using a Residual Network and Continuous Wavelet Transform. *Measurement* **2021**, *183*, 109864. [CrossRef]
7. Lu, J.; Qian, W.; Li, S.; Cui, R. Enhanced K-Nearest Neighbor for Intelligent Fault Diagnosis of Rotating Machinery. *Appl. Sci.* **2021**, *11*, 919. [CrossRef]
8. Chen, H.; Liu, Z.; Alippi, C.; Huang, B.; Liu, D. Explainable Intelligent Fault Diagnosis for Nonlinear Dynamic Systems: From Unsupervised to Supervised Learning. *IEEE Trans. Neural Networks Learn. Syst.* **2022**, 1–14. [CrossRef]
9. Chen, H.; Li, L.; Shang, C.; Huang, B. Fault Detection for Nonlinear Dynamic Systems with Consideration of Modeling Errors: A Data-Driven Approach. *IEEE Trans. Cybern.* **2022**, 1–11. [CrossRef]
10. Jin, T.; Yan, C.; Chen, C.; Yang, Z.; Tian, H.; Wang, S. Light neural network with fewer parameters based on CNN for fault diagnosis of rotating machinery. *Measurement* **2021**, *181*, 109639. [CrossRef]
11. Wu, C.; Jiang, P.; Ding, C.; Feng, F.; Chen, T. Intelligent fault diagnosis of rotating machinery based on one-dimensional convolutional neural network. *Comput. Ind.* **2019**, *108*, 53–61. [CrossRef]
12. He, Z.; Shao, H.; Zhong, X.; Zhao, X. Ensemble Transfer CNNs Driven by Multi-Channel Signals for Fault Diagnosis of Rotating Machinery Cross Working Conditions. *Knowl. Based Syst.* **2020**, *207*, 106396. [CrossRef]
13. Yan, X.; Jia, M. Intelligent fault diagnosis of rotating machinery using improved multiscale dispersion entropy and mRMR feature selection. *Knowl. Based Syst.* **2019**, *163*, 450–471. [CrossRef]
14. Xue, Y.; Dou, D.; Yang, J. Multi-fault diagnosis of rotating machinery based on deep convolution neural network and support vector machine. *Measurement* **2020**, *156*, 107571. [CrossRef]
15. Cheng, Y.; Lin, M.; Wu, J.; Zhu, H.; Shao, X. Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network. *Knowl. Based Syst.* **2021**, *216*, 106796. [CrossRef]
16. Dibaj, A.; Ettefagh, M.M.; Hassannejad, R.; Ehghaghi, M.B. A hybrid fine-tuned VMD and CNN scheme for untrained compound fault diagnosis of rotating machinery with unequal-severity faults. *Expert Syst. Appl.* **2021**, *167*, 114094. [CrossRef]
17. Zhang, Y.; Zhou, T.; Huang, X.; Cao, L.; Zhou, Q. Fault diagnosis of rotating machinery based on recurrent neural networks. *Measurement* **2021**, *171*, 108774. [CrossRef]
18. Guo, S.; Yang, T.; Gao, W.; Zhang, C. A Novel Fault Diagnosis Method for Rotating Machinery Based on a Convolutional Neural Network. *Sensors* **2018**, *18*, 1429. [CrossRef] [PubMed]
19. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K. Densely connected convolutional networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
20. Xu, Y.; Yang, G.; Luo, J.; He, J. An Electronic Component Recognition Algorithm Based on Deep Learning with a Faster SqueezeNet. *Math. Probl. Eng.* **2020**, *2020*, 2940286. [CrossRef]
21. Singla, P.; Duhan, M.; Saroha, S. An ensemble method to forecast 24-h ahead solar irradiance using wavelet decomposition and BiLSTM deep learning network. *Earth Sci. Inform.* **2021**, *15*, 291–306. [CrossRef]
22. Kaur, A.; Jain, S.; Goel, S. Sandpiper optimization algorithm: A novel approach for solving real-life engineering problems. *Appl. Intell.* **2020**, *50*, 582–619. [CrossRef]
23. Saravanakumar, R.; Krishnaraj, N.; Venkatraman, S.; Sivakumar, B.; Prasanna, S.; Shankar, K. Hierarchical symbolic analysis and particle swarm optimization based fault diagnosis model for rotating machineries with deep neural networks. *Measurement* **2020**, *171*, 108771. [CrossRef]
24. Smith, W.A.; Randall, R.B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mech. Syst. Signal Process.* **2015**, *64*, 100–131. [CrossRef]

25. Yang, Y.; Zheng, H.; Li, Y.; Xu, M.; Chen, Y. A fault diagnosis scheme for rotating machinery using hierarchical symbolic analysis and convolutional neural network. *ISA Trans.* **2019**, *91*, 235–252. [CrossRef] [PubMed]
26. Surendran, R.; Khalaf, O.I.; Romero, C.A.T. Deep Learning Based Intelligent Industrial Fault Diagnosis Model. *Comput. Mater. Contin.* **2022**, *70*, 6323–6338. [CrossRef]