machines

MDPI

Article

A Multi-Strategy Optimized Framework for Health Status Assessment of Air Compressors

Dali Hou and Xiaoran Wang *

School of Mechanical Engineering, Shanghai Institute of Technology, Shanghai 201400, China; houdali@sit.edu.cn
* Correspondence: 226091140@mail.sit.edu.cn

Abstract: Air compressors play a crucial role in industrial production, and accurately assessing their health status is vital for ensuring stable operation. The field of health status assessment has made significant progress; however, challenges such as dataset class imbalance, feature selection, and accuracy improvement remain and require further refinement. To address these issues, this paper proposes a novel algorithm based on multi-strategy optimization, using air compressors as the research subject. During data preprocessing, the Synthetic Minority Over-sampling Technique (SMOTE) is introduced to effectively balance class distribution. By integrating the Squeeze-and-Excitation (SE) mechanism with Convolutional Neural Networks (CNNs), key features within the dataset are extracted and emphasized, reducing the impact of irrelevant features on model efficiency. Finally, Bidirectional Long Short-Term Memory (BiLSTM) networks are employed for health status assessment and classification of the air compressor. The Ivy algorithm (IVYA) is introduced to optimize the BiLSTM's hyperparameters to improve classification accuracy and avoid local optima. Through comparative and ablation experiments, the effectiveness of the proposed SMOTE-IVY-SE-CNN-BiLSTM model is validated, demonstrating its ability to significantly enhance the accuracy of air compressor health status assessment.

Keywords: air compressors; health status; class imbalance; ivy algorithm; feature extraction

check for updates

Academic Editors: Jie Liu and Wentao Mao

Received: 20 February 2025
Revised: 13 March 2025
Accepted: 18 March 2025
Published: 20 March 2025

Citation: Hou, D.; Wang, X. A Multi-Strategy Optimized Framework for Health Status Assessment of Air Compressors. *Machines* **2025**, *13*, 248. https://doi.org/10.3390/machines13030248

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

1. Introduction

With the advent of the Industry 4.0 era, industrial production is undergoing an unprecedented intelligent transformation [1]. In this context, mechanical equipment is becoming increasingly complex, raising higher requirements for real-time monitoring and health management of its operating status. In the natural gas transmission sector, air compressors, as core equipment in the compressed air system [2], directly impact the safety and operational efficiency of the entire oil and gas transmission system [3]. Therefore, real-time and accurate monitoring and evaluation of an air compressor's health have become key technological challenges to ensure production safety and improve system reliability [4].

The methods for assessing the health status of air compressors have transitioned from conventional manual inspections to advanced intelligent monitoring systems utilizing sensors [5]. Conventional approaches depend largely on routine manual examinations and predetermined maintenance schedules, which are often inefficient and prone to human error. Therefore, the development of intelligent health status assessment algorithms has become a key research direction.

In recent years, neural networks have been more popular in health status assessment. They have shown good performance in equipment health diagnostics, and related research has been more comprehensive [6–8]. For example, a method based on KS density

and Bidirectional Long Short-Term Memory (BiLSTM) networks was proposed for health prediction of rotating motor bearings [9]. Nambiar et al. [10] have employed predictive machine learning models and feature fusion techniques for fault diagnosis of single-acting, single-stage reciprocating air compressors. Shar et al. [11] developed a diagnostic framework centered on energy performance, integrating time series analysis with isolation forest methods to accurately identify faults and evaluate the operational health of centrifugal gas compressors. Tran et al. [12] introduced a hybrid deep belief network (HDBN) approach, leveraging vibration and acoustic emission data to achieve high-precision fault detection in reciprocating compressor valves.

Although the neural networks have shown good results in classification, challenges remain, particularly in hyperparameter optimization and model configuration. Traditional gradient-based optimization methods are prone to getting stuck in local optima, whereas heuristic optimization algorithms with global search capabilities, such as Genetic Algorithms [13], Particle Swarm Optimization [14], and Ant Colony Optimization [15], offer some solutions to this problem. In 2024, Mojtaba Ghasemi et al. introduced the Ivy algorithm (IVYA) [16], offering an innovative methodology [16]. The Ivy algorithm simulates the growth pattern of ivy plants and achieves global search by coordinating population growth, diffusion, and evolution, thereby improving optimization efficiency and accuracy. The IVYA has been applied to optimize the Random Forest (RF) model to enhance the accuracy of ship diesel generator systems in fault diagnosis, and the superior performance of the IVYA in this application was verified [17].

However, air compressor data often contain intricate nonlinear patterns, which continues to pose challenges for health status assessment. Therefore, it is necessary to employ more effective techniques to optimize the feature extraction process. This approach allows the model to better interpret data and identify critical features with greater precision, ultimately boosting its evaluation and generalization performance. Xu et al. [18] demonstrated an innovative multi-label feature extraction technique by combining two least-squares formulations within Principal Component Analysis (PCA) and Multi-label Dimensionality Reduction via Dependence Maximization (MDDM) in a linear framework. Additionally, Jogin et al. [19] employed Convolutional Neural Networks (CNNs) and Deep Learning methodologies for effective feature extraction. CNNs have been proven to be effective in extracting local features and capturing periodic and trend patterns in the data [20]. CNNs automatically extract key information through convolution operations, while pooling layers reduce the dimensionality of features, reduce computational demands, and improve the model's generalization capabilities [21].

We built an innovative optimized model to solve the above issues. The primary design principle of this model is to enhance the accuracy of air compressor health assessments through a more effective method, following recent trends of multi-strategy optimization and advanced algorithm integration. Therefore, an efficient optimization algorithm should be considered, one that follows the latest trend of adaptive adjustment of model parameters, aiming to dynamically optimize hyperparameters and weight configurations to enhance the model's stability. Guided by this design philosophy, the model's modules and their functions, as well as the interrelationships between the modules, have been thoroughly considered and optimized. We strive to ensure that each module is custom-designed to address specific issues in the health status assessment of air compressors, maximizing the model's performance. The structure of our proposed model can be summarized as follows: First, the Synthetic Minority Over-sampling Technique (SMOTE) [22] technique is used to augment the minority class samples. Next, CNNs are used to automatically extract key features from the dataset. Subsequently, the Squeeze-and-Excitation (SE) mechanism is integrated into our model to dynamically recalibrate feature weights, thereby enhancing

the feature extraction capabilities of the CNN. Finally, the IVY-optimized BiLSTM module is employed to refine the model's performance by optimizing its hyperparameters, leading to improved accuracy in health status classification.

To rigorously evaluate the performance of our model, we not only conducted comparative experiments with other existing algorithms but also performed ablation experiments to thoroughly analyze the contribution of each module to the overall performance. In terms of health status classification accuracy, the proposed model achieved the better results. The primary contributions of this work are summarized as follows:

(1) We have proposed a novel multi-strategy optimized health status classification model for air compressors. The model integrates three modules: SMOTE, SE-CNN, and IVY-BiLSTM. First, the SMOTE module effectively addresses the class imbalance issue. Second, the SE-CNN module extracts the data features, highlighting important features while suppressing redundant or irrelevant ones. Finally, by integrating the IVY-optimized BiLSTM network, the model is able to more accurately learn the behavior patterns of the compressor in different states, leading to improved health status classification.
(2) Our goal is to find an excellent optimization algorithm that can fully leverage the potential of the model and achieve optimal classification performance when handling complex classification problems.
(3) Our goal is to effectively integrate data preprocessing and feature extraction into the optimized algorithm, thereby improving the efficiency of health assessment and increasing the precision of the evaluation results.
(4) Our model demonstrates its relatively superior performance through comparative and ablation experiments. The model's generalization ability is also validated by applying it to different datasets.

2. Materials and Methods

The framework of the SMOTE-IVY-SE-CNN-BiLSTM is shown in Figure 1. This model integrates the SMOTE, Ivy algorithm, CNN, SE mechanism, and BiLSTM. The process can be illustrated as follows.

Step 1: The SMOTE is used to over-sample the minority class and generate synthetic samples, effectively balancing the proportion of minority and majority class samples in the air compressor status data. This approach helps mitigate misclassification issues caused by the insufficient number of minority class samples in the original data.

Step 2: After data preprocessing, the CNN is employed to extract key local features, capturing periodic, trending, or abrupt changes in the air compressor's data. The SE mechanism is introduced to enhance the performance of the CNN. The SE mechanism adaptively adjusts the feature weights, enabling the CNN to capture features more effectively while suppressing less relevant ones.

Step 3: To achieve more precise health status recognition, the IVYA is utilized to optimize the hyperparameters of the BiLSTM, including the number of neurons in the hidden layer, the initial learning rate, and the L2 regularization coefficient.

Step 4: The information output from step 2 is transmitted to the optimized BiLSTM layer to obtain the ultimate state classification outcome.

Our architecture adopts a modular decoupled design where sub-models achieve both functional autonomy and collaborative coordination. The SMOTE (data processing module), SE-CNN (feature engineering module), and IVY-BiLSTM (classification module) are designed as self-contained units interconnected through standardized interfaces. This configuration allows each module to operate independently while maintaining seamless interaction and coordination. This architecture also allows isolated hyperparameter op-

timization for the BiLSTM, avoiding the combinatorial complexity of end-to-end tuning. Such modularity not only improves the transparency of workflow but also enables targeted optimization of individual components without disrupting system integrity, aligning with industrial requirements for flexibility and maintainability. Brief core algorithm can be seen at Appendix A.

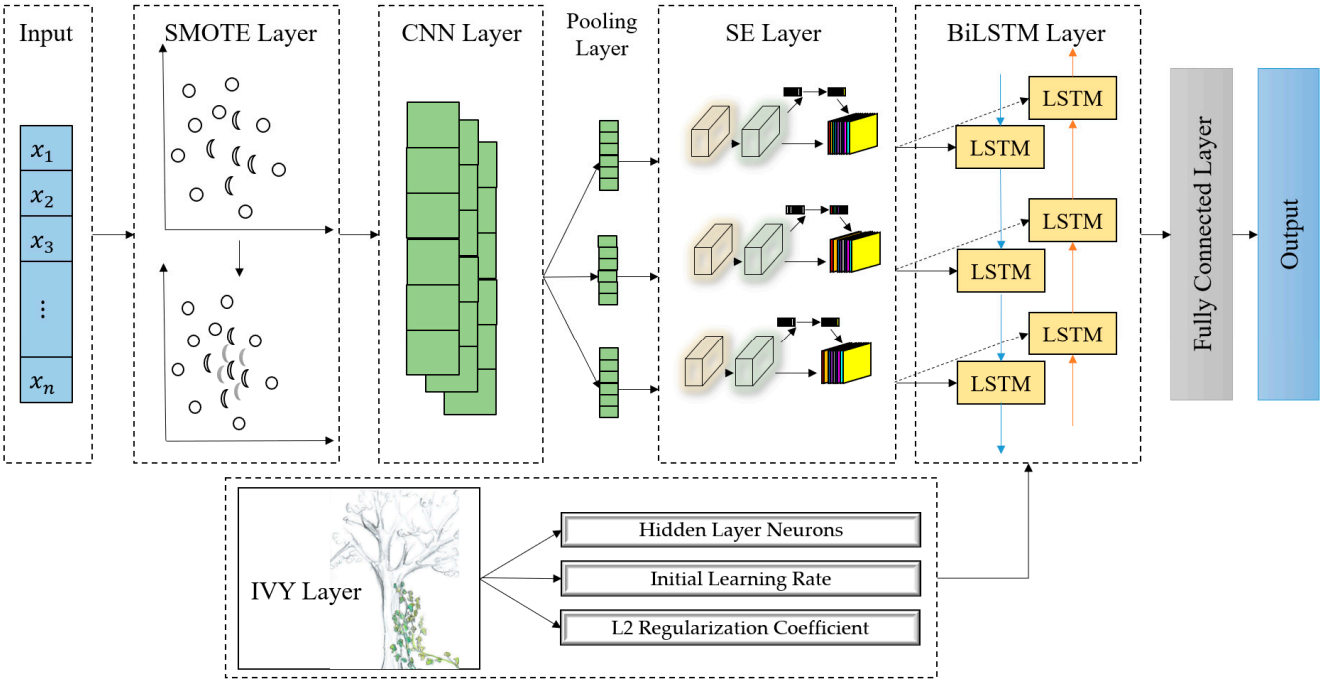

Figure 1. The framework of the SMOTE-IVY-SE-CNN-BiLSTM.

2.1. Synthetic Minority Over-Sampling Technique

The SMOTE, introduced by Chawla et al. [22], has become widely used in machine learning and data mining, particularly for addressing class imbalance issues. The core idea of the SMOTE is as follows: First, a minority class sample is randomly selected, and its k nearest neighbors are identified within its neighborhood. Then, based on a certain probability, one of these k neighbors is randomly chosen, and a new synthetic sample is created through "interpolation".

The generation of new synthetic samples can be illustrated as follows:

$$Y = Z_i + rand \times (Z_1 - Z_2) \tag{1}$$

where Z_1 represents a majority class sample, Z_2 is one of the k nearest samples to Z_i, *rand* is a random number between 0 and 1, and Y represents newly generated minority class sample.

The explanation of the SMOTE over-sampling method is shown in Figure 2.

We use the modular and loosely coupled design of the SMOTE algorithm, and its plug-and-play capability enables selective deactivation of the preprocessing module in scenarios with balanced data distributions, thereby achieving dynamic optimization of model redundancy and targeted reduction in computational overhead.

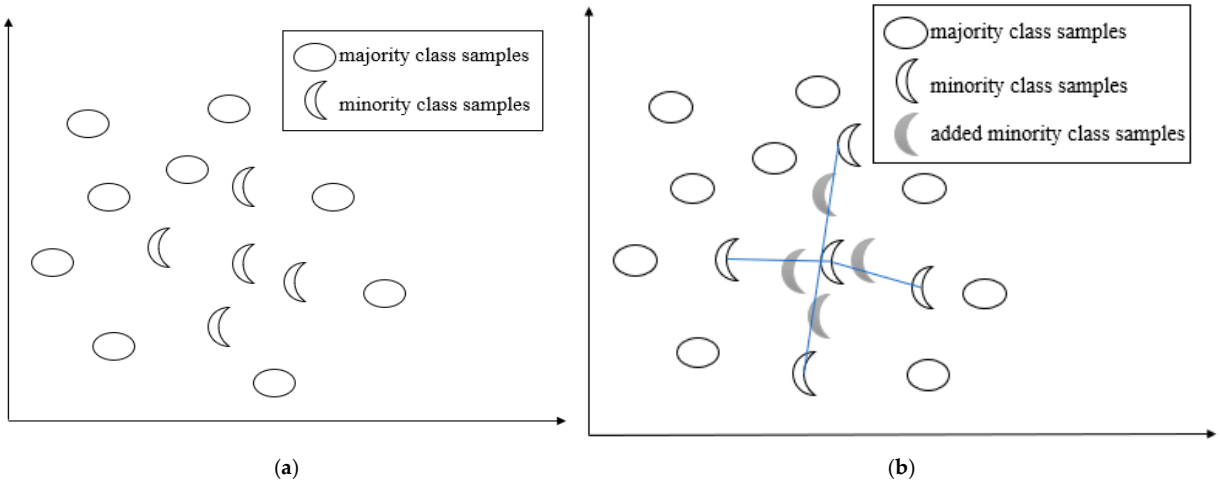

Figure 2. Explanation of the SMOTE over-sampling method: (**a**) unprocessed sample distribution; (**b**) processed sample distribution, the blue line represents the "interpolation" process of SMOTE.

2.2. Establishment of the SE-CNN Model

2.2.1. Convolutional Neural Network

The CNN can be used for effective feature extraction of air compressors, helping the model to make better health status assessments.

A standard CNN comprises the input layer, convolutional layers, pooling layers, fully connected layers, and the output layer [23]. The architecture of the model is depicted in Figure 3.

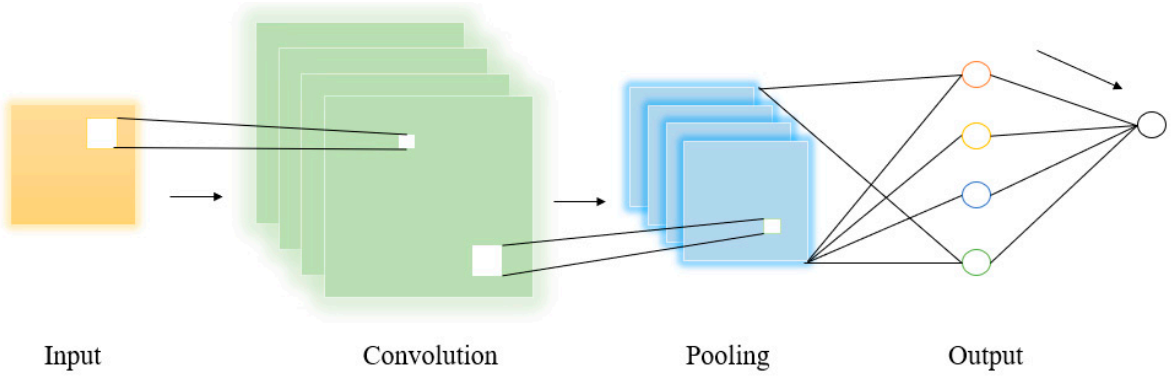

Figure 3. The basic structure of the CNN.

The convolutional layer serves as a fundamental element of a CNN, composed of a collection of parallel feature maps. The CNN performs convolution operations on the input raw air compressor data through the convolutional layer. Convolution operations fundamentally act as a method for detecting local features, utilizing a sliding window to extract localized patterns from the data. By sliding multiple convolutional kernels over the data, the CNN can automatically identify important local features, which are essential for assessing the health status of the air compressor. The convolution operation is computed as shown in Equation (2).

$$x_j^r = f\left(\sum x_i^{r-1} \times k_{i,j}^r + b_j^r\right) \tag{2}$$

where x_j^r corresponds to the j-th feature map produced by the r-th convolution operation, while b_j^r represents the bias term associated with the j-th convolutional kernel in the r-th layer, the term $k_{i,j}^r$ refers to the j-th kernel applied to the i-th input feature map, and f denotes the nonlinear activation function used in the process.

The pooling layer is commonly positioned following the convolutional layer, serving the key purpose of down-sampling the extracted convolutional features. Pooling operations aim to reduce feature dimensionality while preserving essential information. Through max pooling or average pooling, the CNN can select the most representative features and effectively suppress noise. This module uses global average pooling for calculation. The global average pooling layer computes the mean value of all pixels within each feature map, producing a condensed feature vector. The calculation is expressed as follows:

$$y_{k,i,j} = \frac{1}{|R_{i,j}|} \sum_{(m,n)\in R_{i,j}} x_{k,m,n} \tag{3}$$

where $y_{k,i,j}$ represents the average pooling value associated for the k-th feature map within the rectangular area $R_{i,j}$, $x_{k,m,n}$ denotes the element located at position (m, n) within $R_{i,j}$, and $|R_{i,j}|$ represents the number of elements in the region $R_{i,j}$.

The fully connected layer performs a fully connected operation on all the neurons from the last pooling layer in the CNN, and its equation is represented as follows:

$$O(x) = f(wx + b) \tag{4}$$

where x is the input to the fully connected layer, w represents the weight matrix, and b is the bias vector.

2.2.2. Squeeze-and-Excitation Mechanism

When using CNNs to extract the operating features of the air compressor, there is often the issue of irrelevant information being extracted by the network. Irrelevant information often spreads across the channel dimensions of feature maps, distracting the CNN's focus.

To address this issue, this paper introduced the SE mechanism. The SE mechanism comprises two primary operations: Squeeze and Excitation [24]. The Squeeze operation F_{sq} compresses the spatial features of each channel into a global descriptor using global average pooling, capturing essential information from the entire channel. The Excitation operation F_{ex} then learns the weight coefficients for each channel based on its global feature, adjusting the output of different channels accordingly. This way, the SE mechanism effectively strengthens the features critical to the health status assessment of the air compressor while suppressing the interference of irrelevant or redundant information, thus enhancing the model's discriminative capability.

The structure of the SE model is shown in Figure 4.

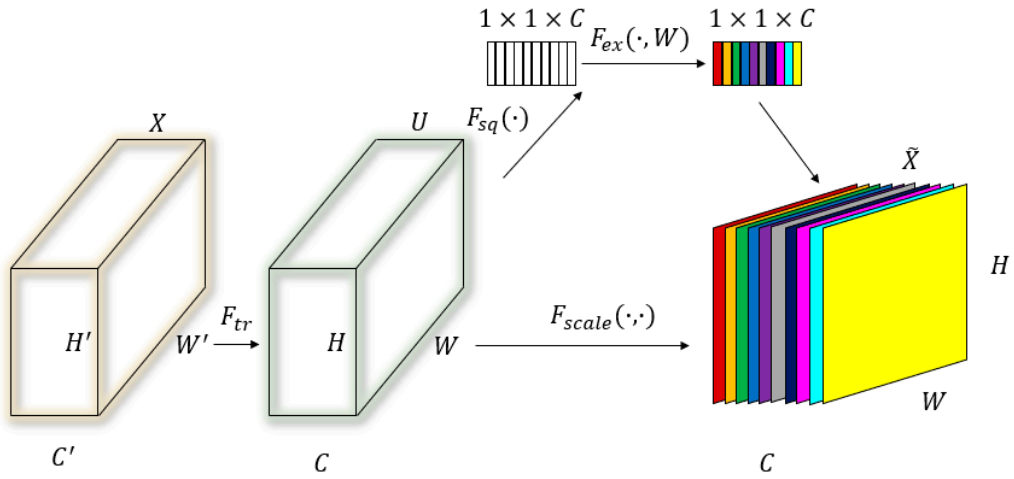

Figure 4. Structure of the SE model.

In this figure, C indicates the number of channels in the feature map, W and H represent the width and height dimensions of the feature map, X represents the input feature map, $\tilde{X}$ signifies the output feature map after the scaling operation, U stands for the transformed feature map, and F_{tr} corresponds to the transformation process from X to U.

2.2.3. Construction of the SE-CNN Model

The SE-CNN model can be illustrated as Figure 5, which includes two convolutional layers, two pooling layers, and one SE module.

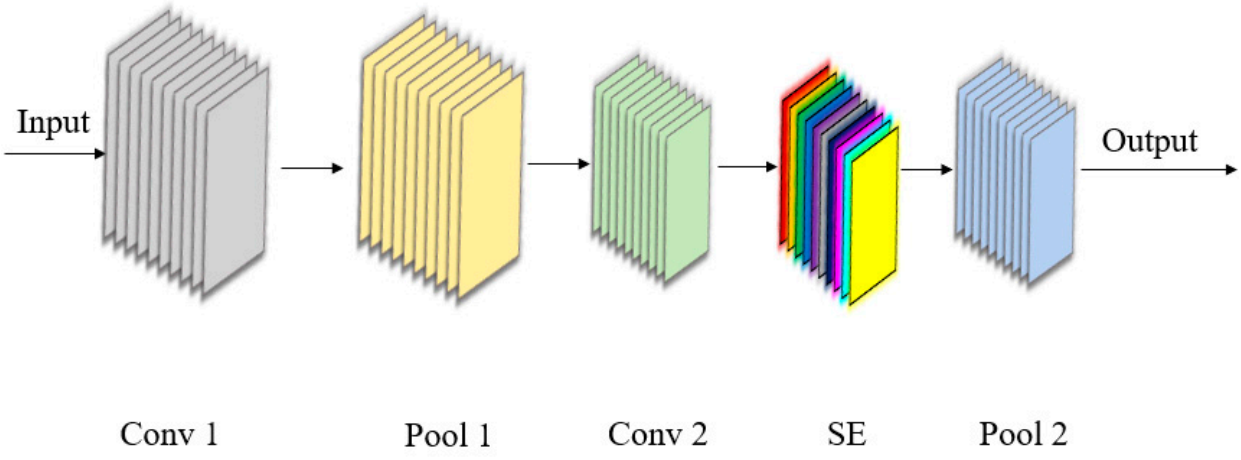

Figure 5. The model of the SE-CNN. Different colors represent different layers of the SE-CNN.

The feature data collected from the air compressor are used as the input, which passes through the first convolutional layer to extract local features and generate feature maps. The pooling layer then down-samples these feature maps, reducing their spatial dimensions while retaining critical information The second convolutional layer processes the pooled feature maps, further extracting more abstract features. Then, the SE module introduces a global feature re-scaling mechanism, assigning adaptive weights to each channel of the feature map to emphasize or de-emphasize their significance. Finally, the pooling layer down-samples the weighted feature maps from the SE module, further compressing their size and simplifying the feature representation.

2.3. Ivy Algorithm Optimization for Bidirectional Long Short-Term Memory Network

2.3.1. Ivy Algorithm

The Ivy algorithm (IVYA) is an intelligent optimization algorithm proposed by Mojtaba Ghasemi et al. in 2024 [16]. This algorithm simulates the growth pattern of ivy plants and aims to achieve optimization objectives through the coordinated and orderly growth of populations, as well as the spread and evolution of the ivy plants. The algorithm leverages data from neighboring ivy plants to identify the optimal growth direction and enhances its performance by focusing on the closest and most significant neighboring plants.

Step 1: Initialization.

The algorithm begins by randomly assigning the initial positions of the IVY population within the defined search space. This is achieved using the following equation:

$$I_i = I_{min} + rand(1, Z) \odot (I_{max} - I_{min}) \tag{5}$$

Here, $rand(1,\ Z)$ is a vector of size Z containing uniformly distributed random numbers within the range [0, 1]. I_{max} and I_{min} define the upper and lower boundaries of the search space, respectively, and the Hadamard product is represented by the operation "$\odot$".

Step 2: Coordinated and ordered population growth.

The growth rate Gr of each ivy plant is modeled using the following equation:

$$\triangle Gr(t+1) = rand^2 \cdot (N(1,Z) \odot \triangle Gr(t)) \tag{6}$$

where $\triangle Gr$ denotes the growth rates in a discrete-time system, $rand^2$ is a random number generated from a probability density function $\frac{1}{2\sqrt{x}}$, and $N(1,Z)$ is a random vector of dimension Z, where each component is drawn from a standard Gaussian distribution.

Step 3: The growth of ivy plants.

When an individual l_i identifies its most significant and closest neighbor l_{ii} during the global search, its movement is determined by the fitness function as follows:

$$l_i^{new} = l_i + |N(1,Z)| \odot (l_{ii} - l_i) + N(1,Z) \odot \triangle Gr_i \tag{7}$$

$$\triangle Gr_i = \begin{cases} l_i \odot (l_{max} - l_{min}), & Iter = 1 \\ rand^2 \odot (N(1,Z) \odot \triangle Gr_i). & Iter > 1 \end{cases} \tag{8}$$

where l_i^{new} denotes the updated position of the i-th individual, and $|N(1,Z)|$ denotes the absolute value of the random vector $N(1,Z)$.

Step 4: Spreading and evolution of ivy plants.

During the final stages of the search process, each individual l_i will progressively move towards the most optimal individual l_{best}, mimicking the strategy of focusing on the best solution in the vicinity, as represented in Equations (9) and (10).

$$l_i^{new} = l_{best} \odot rand(1,Z) + N(1,Z) \odot \triangle Gr_i \tag{9}$$

$$\triangle Gr_i^{new} = l_i^{new} \oslash (l_{max} - l_{min}) \tag{10}$$

where $\triangle Gr_i^{new}$ represents the updated growth rate.

2.3.2. Bidirectional Long Short-Term Memory Network

The advantage of LSTM neural networks is that they retain the structure of traditional multilayer perceptrons while incorporating self-regulating feedback connections. However, the LSTM still employs unidirectional information flow, limiting its ability to capture all the information. The BiLSTM structure consists of two LSTMs [25–27], enabling a more comprehensive data analysis. The structure is illustrated in Figure 6.

The computation process of the BiLSTM is shown in the following equation.

$$\begin{cases} f_i = s(W_j x_j + U_j h_{t-1} + b_j) \\ \widetilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \\ i_i = s(W_i x_t + U_i h_{t-1} + b_i) \\ o_i = s(W_o x_t + U_o h_{t-1} + b_o) \\ c_t = f_i * \widetilde{c}_{t-1} + i_i * \widetilde{c}_t \\ h_t = o_i * \tanh(c_t) \end{cases} \tag{11}$$

where f_i, i_i, and o_i are the output parameters of the forget gate, update gate, and output gate, respectively; s is the nonlinear activation function; W_j, U_j, W_c, U_c, W_i, U_i, W_o, and U_o are the weights for each part; b_j, b_c, b_i, and b_o are the bias matrices; the input to the current LSTM unit is x_t; h_{t-1} represents the hidden state output from the previous LSTM unit, which also serves as part of the current output; and $\widetilde{c}_t$ and c_t refer to the candidate cell state and the final cell state, respectively, which are crucial for maintaining long-term dependencies in the network.

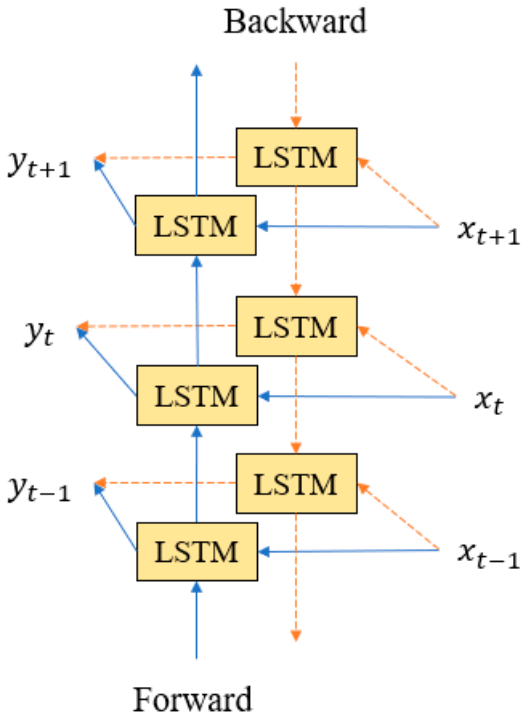

Figure 6. The structure of the BiLSTM. The blue arrows represent the forward process of the BiLSTM, while the orange arrows represent the backward process of the BiLSTM.

2.3.3. Optimization Process

The Ivy algorithm (IVY) optimizes the hyperparameters of the BiLSTM model by combining both local and global search strategies to successfully avoid the model becoming trapped in local optima. During the climbing phase, the IVY algorithm fine-tunes the hyperparameters through detailed local search, gradually improving the performance of the BiLSTM and enhancing its fitness, preventing premature convergence. In the expansion phase, the IVY performs a global search, exploring a broader solution space to ensure the discovery of a global optimum.

The optimization process is outlined below.

Step 1: Initialization.

First, initialize a population consisting of multiple ivy plant individuals, where each individual represents a set of BiLSTM hyperparameters. These hyperparameters include learning rate, L2 regularization coefficient, and number of hidden layer neurons.

Step 2: Climbing phase: local search.

During the climbing phase, ivy plants select the closest and most suitable neighbors and move in that direction to promote growth. This process simulates the optimization of the IVY algorithm within the local solution space. In the BiLSTM model, the climbing phase optimizes the current hyperparameter combination by fine-tuning it, gradually improving the model's performance while avoiding local optima.

Step 3: Expansion phase: global search.

In the expansion phase, ivy plants explore new areas through extensive expansion to find more space for growth. This mechanism emulates the global search capability of the IVY algorithm. During the optimization of BiLSTM hyperparameters, the expansion phase explores the global solution space by introducing larger variations, thereby preventing convergence to local optima and facilitating the identification of potential global optima.

Step 4: Fitness evaluation and update.

After each iteration, the fitness value of the BiLSTM model is calculated. The population is ranked based on the fitness value, and the best individuals are retained. These

individuals' hyperparameters are used to update the others. An elitism strategy is typically employed to ensure that the best solutions are preserved after each iteration.

Step 5: Termination condition and output.

The IVY algorithm terminates under two conditions: either the specified maximum number of iterations is completed, or the performance of the BiLSTM model reaches a stable state, signifying that further optimization yields negligible improvements. The hyperparameters corresponding to the best individual are returned as the final result. The BiLSTM model is then trained with the optimized hyperparameters.

The process of the IVY-BiLSTM is shown in Figure 7.

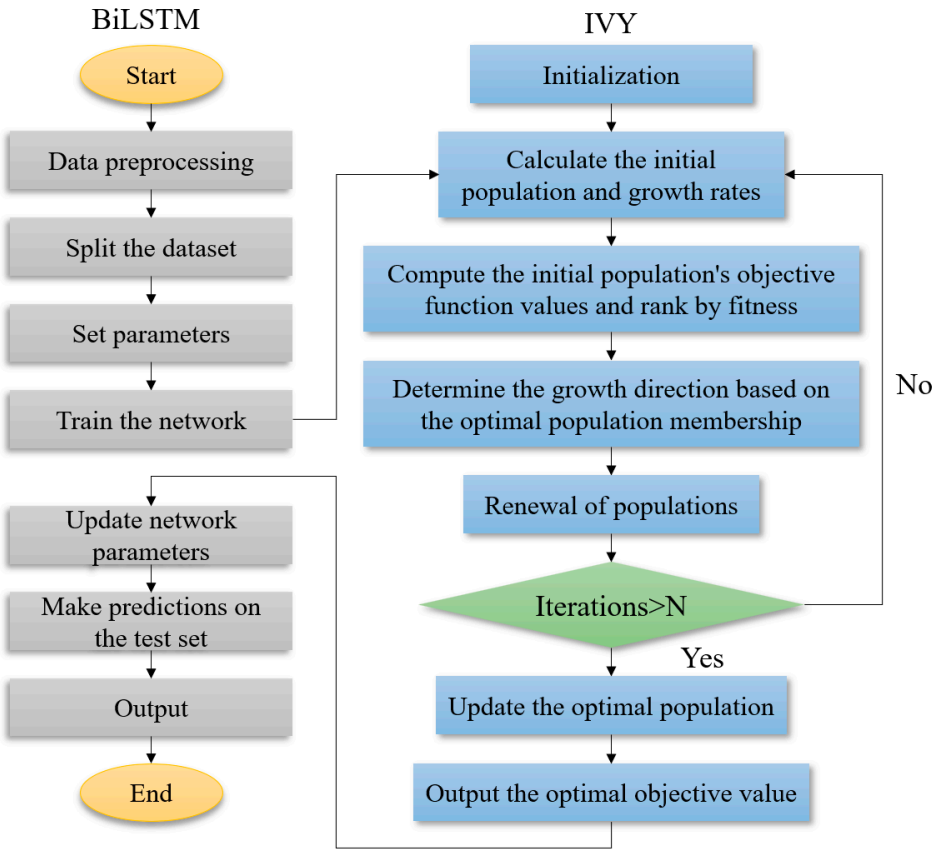

Figure 7. The process of the IVY-BiLSTM.

3. Experiments

3.1. Datasets and Preprocessing

This paper collects monitoring data of air compressors under different health conditions, including healthy, sub-healthy, dangerous, and faulty states, with about 1000 samples in the dataset. The dataset originates from an air compressor remote monitoring and fault diagnosis system under the Industrial Internet of Things (IIoT) framework of a natural gas transmission company. The system employs Remote Terminal Units (RTUs) [28] deployed on air compressor units to achieve real-time acquisition of multi-source sensor data, utilizing the Modbus TCP/IP protocol for remote transmission and aggregation of monitoring signals. The collected data at the system terminal underwent processing and filtering, followed by time-synchronized alignment of alarm logs with operational parameters, ultimately constructing an air compressor operational condition dataset annotated with health status labels.

The compressor parameters are shown in Table 1.

Table 1. Operating parameters of the air compressor.

Operating Parameter Name	Unit	Data Type
Air Pressure	Bar	1/L
Pressure Difference	Bar	1/L
Intermediate Pressure	Bar	1/L
Oil Pressure	Bar	1/L
Lubricating Oil Temperature	°C	1/L
1st Stage Rotor Exhaust Temperature	°C	1/L
2nd Stage Rotor Inlet Temperature	°C	1/L
2nd Stage Rotor Exhaust Temperature	°C	1/L
Inlet Temperature	°C	1/L
Exhaust Temperature	°C	1/L

To train and evaluate the model, the dataset is divided into three subsets: 60% for training, 10% for validation, and 30% for testing. Based on practical experience and the specific situation, this ratio usually provides a good balance.

Analysis of the air compressor dataset reveals that the "healthy" state accounts for the majority of all states, while the "dangerous" and "faulty" states are relatively less represented. Thus, the model may not be able to fully learn the characteristics of "faulty" states, leading to an increased error rate when identifying this class, thereby affecting classification accuracy. To address this issue, this paper employs the SMOTE to create synthetic samples for the underrepresented classes in the training dataset, thereby augmenting their representation and achieving a more balanced class distribution.

The data distribution before and after applying the SMOTE is shown in Table 2.

Table 2. Training data distribution before and after applying the SMOTE.

State Type	Quantity (Before)	Quantity (After)	Label
healthy	488	488	1
sub-healthy	308	417	2
dangerous	146	360	3
faulty	58	312	4

3.2. Evaluation Metrics

To evaluate the performance of this model, we utilize Cohen's Kappa, recall, F1 score, and accuracy as the evaluation metrics.

The Cohen's Kappa spans from 0 to 1, with values approaching 1 indicating stronger agreement between the model's predictions and the actual outcomes, thus reflecting a more reliable and accurate evaluation.

$$k = \frac{P_o - P_i}{1 - P_i} \tag{12}$$

where P_o represents the observed accuracy, and P_i represents the expected accuracy.

Recall refers to the proportion of actual samples belonging to a certain category that are correctly predicted as that category by the model. There is a corresponding recall rate for each category.

The F1 score is derived as the harmonic mean of precision P and recall R.

$$F1 = 2 \times \frac{P \times R}{P + R} \tag{13}$$

3.3. SE-CNN Model Parameter Configuration

Traditional air compressor health state assessment methods indicate a strong correlation between the operational state of the compressor and its state parameters. However, with the increasing volume of data, there may be issues of information redundancy. These redundant features not only increase the computational burden on the model but may also affect its performance. Therefore, reducing data dimensionality and model simplification have become necessary optimization strategies. To address such issue, this study proposes the SE-CNN model, designed to refine and optimize the feature extraction process for air compressor data.

The configuration parameters for the SE-CNN model are detailed in Table 3. These parameters were chosen based on insights from the existing literature and practical experience [29].

Table 3. Parameters of the SE-CNN model.

Parameter Name	Parameter Value
Convolution Layer	2
Global Average Pooling Layer	2
SE Module	1
Convolution Kernel Size	3×1

3.4. IVY-BiLSTM Model Parameter Configuration

To determine the optimal hyperparameters for the BiLSTM, the IVY algorithm is used for optimization, forming the IVY-BiLSTM model. The first step is to determine the initial parameters. Based on experience and references from the literature, start with some common initial values, and dynamically adjust these parameters through multiple trials to ultimately determine the more suitable initial parameters [30]. Table 4 displays the initial parameter configurations.

Table 4. Initial parameters of the IVY-BiLSTM model.

Parameter Name	Initial Parameter Value
Initial Population Size	12
Learning Rate	$10^{-4} - 10^{-2}$
Hidden Layer Neurons	10–30
Regularization Coefficient	$10^{-4} - 10^{-1}$
Maximum Number of Iterations	10
Maximum Number of Training Iterations	150

Figure 8 illustrates the trend of accuracy as a function of the number of epochs. As the training process advances, the accuracy consistently rises, demonstrating that the model is progressively enhancing its predictive capabilities. In the early stages, the accuracy rises rapidly, then enters a stabilization phase, where it remains at a high level, suggesting that the model's fitting on the training set gradually improves.

Figure 9 presents the evolution of the loss function over the number of epochs during the training phase. Initially, the loss exhibits a sharp decline, indicating rapid convergence and effective learning. As training continues, the loss curve flattens and stabilizes at a minimal value, signifying that the model has successfully reduced prediction errors and achieved a robust and efficient training outcome. This behavior underscores the model's ability to optimize performance and maintain stability throughout the learning process.

After determining the initial parameters, we can obtain the optimal hyperparameters. The IVY-SE-CNN-BiLSTM model is used to classify the data and output the final health status.

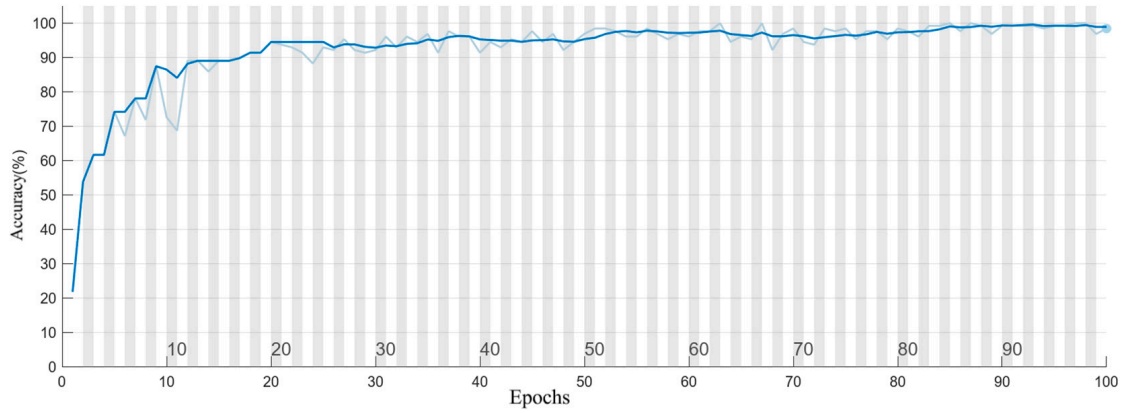

Figure 8. Accuracy trend with respect to epochs. The blue line represents the actual trend, while the dark blue line represents the predicted trend by the system.

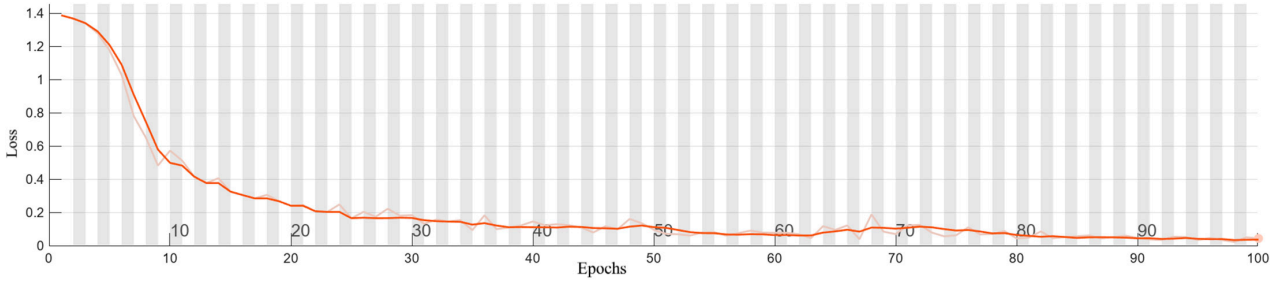

Figure 9. Loss trend with respect to epochs. The orange line represents the actual trend, while the dark orange line represents the predicted trend by the system.

4. Results and Discussion

4.1. Comparative Analysis of Sample Balancing Methods

To verify the model's capability in managing imbalanced datasets and explore the importance of data balancing in the process of the health status assessment, this paper designs a comparative experiment from a data processing perspective. The experiment users three different datasets: the first dataset is the raw data without any preprocessing, the second dataset is balanced using random over-sampling, and the third dataset is balanced using the SMOTE. These three datasets, serving as the sole variable, are fed into the model for both training and testing purposes. Table 5 presents the detailed results of the health status assessment for the three different datasets.

Table 5. Performance of the three datasets.

Experimental Method	Accuracy	Kappa
SMOTE	**0.9689**	0.9370
Random Over-sampling	0.9365	0.9107
Without Sampling	0.8601	0.8305

The best results are indicated in bold font.

The results clearly show that using the SMOTE sampling improved the model's accuracy. The accuracy of the air compressor health assessment using the SMOTE sampling

for the test set reaches 96.89%, while the accuracy with random over-sampling is 93.65%, and the accuracy using the imbalanced dataset is only 86.01%.

When the original data are normalized and directly divided for input into the health status assessment model, the model struggles to generalize effectively to the "faulty" category because of the insufficient number of samples for that category in the original dataset, leading to the worse evaluation performance. This is because, with an imbalanced dataset, the model tends to prioritize learning the more abundant classes (such as the "healthy" category) while neglecting the less represented classes (such as the "faulty" category).

Random over-sampling balances the class distribution by simply duplicating minority class samples, enabling the model to learn and capture relevant features from the minority class. Compared to the original dataset, the performance of the model improved by 8.88% when tested on the over-sampled dataset. However, simply replicating minority class samples may cause the model to repeatedly encounter the same samples, preventing it from learning diverse features. Additionally, random over-sampling can lead to biased sample distributions by excessively increasing the number of minority class samples, which not only fails to improve the model's learning of the minority class but may also degrade the model's performance and stability in real-world applications.

In contrast, the SMOTE offers distinct advantages. The SMOTE does not simply replicate existing minority class samples; instead, it creates new synthetic samples by interpolating between neighboring minority class samples in the feature space. Unlike simple replication methods, this approach enhances the diversity of the minority class data, providing more samples that enable the model to more effectively learn the distinctive features of the minority class. After applying the SMOTE, the accuracy of the air compressor health status assessment improved by 12.65%.

4.2. Analysis of the Results of the Air Compressor Health Assessment Model Based on the SMOTE-IVY-SE-CNN-BiLSTM

This section presents an in-depth analysis of the performance of the air compressor health assessment model based on the SMOTE-IVY-SE-CNN-BiLSTM. After training the model, we evaluate its performance by applying the test set features to multiple models. Along with the approaches introduced in this study, seven other algorithms, namely, the Backpropagation Neural Network (BP) [31], Particle Swarm Optimization–Bidirectional Long Short-Term Memory network (PSO-BiLSTM) [32], Random Forest (RF) [33], Support Vector Machine (SVM) [34], BiLSTM, LSTM, and IVY-BiLSTM are selected as comparison models.

The BP adjusts the weights and biases using the backpropagation algorithm to optimize classification performance. The RF enhances classification stability and accuracy by combining multiple decision trees and employing a voting mechanism. The SVM, on the other hand, identifies the optimal hyperplane to maximize the margin between classes and can address nonlinear classification challenges through the use of kernel functions. The PSO-BiLSTM classification model combines the PSO algorithm with the BiLSTM network to improve the performance of classification tasks by optimizing the hyperparameters of the BiLSTM. The PSO-BiLSTM model utilizes the max iteration of 10 and the particle swarm population size of 20. The range of the learning rate, hidden layer node, and L2 regularization coefficient are $10^{-4} - 10^{-1}$, 10–30, and $10^{-4} - 10^{-1}$, respectively. The parameters of the IVY-BiLSTM model are same as the SMOTE-IVY-SE-CNN-BiLSTM.

The performance of these models are compared in Table 6.

In Table 6, the evaluation metrics for all models include Kappa value, F1 score, and accuracy, which reflect the models' performance in the health status assessment task of air compressors.

Table 6. Performance of the different models.

	Method	Training			Test		
		Kappa	**F1**	**Accuracy**	**Kappa**	**F1**	**Accuracy**
1	BP	0.7039	0.7382	0.7139	0.6779	0.7132	0.6758
2	SVM	0.8531	0.8702	0.8591	0.8003	0.8327	0.8237
3	RF	0.8613	0.8791	0.8627	0.8009	0.8562	0.8423
4	LSTM	0.9301	0.9277	0.9190	0.8774	0.8801	0.8791
5	BiLSTM	0.9347	0.9401	0.9337	0.9005	0.9117	0.9091
6	PSO-BiLSTM	0.9668	0.9703	0.9681	0.9341	0.9517	0.9428
7	IVY-BiLSTM	0.9713	0.9886	0.9758	0.9581	0.9633	0.9511
8	IVY-SE-CNN-BiLSTM	0.9548	0.9803	**0.9779**	0.9370	0.9725	**0.9689**

The best results are indicated in bold font.

In the comparison between Model 1 and Model 5, the BiLSTM performs the best. The BiLSTM, with its bidirectional structure, can capture and utilize the contextual relationships in the data more effectively, allowing for a deeper understanding of the multi-dimensional features of the input, thereby improving classification accuracy. In contrast, the BP, SVM, and RF, due to their simpler architectures, lack deep feature extraction capabilities and are more prone to overfitting. As a result, they perform worse than the BiLSTM in handling complex health status classification tasks. The LSTM, with its unidirectional structure, performs slightly worse than the BiLSTM. In the comparison between the PSO-BiLSTM (Model 6) and the IVY-BiLSTM (Model 7), the IVY optimization algorithm demonstrates its advantages. The IVY optimization improves the performance of the BiLSTM model by making more extensive adjustments globally and performing detailed optimization at the local level. This is reflected in the fact that the IVY-BiLSTM outperforms the PSO-BiLSTM in terms of Kappa value, F1 score, and accuracy.

Compared to these, it is evident that the SMOTE-IVY-SE-CNN-BiLSTM model exhibits advantages in both during the training and testing phases. On the test set, the Kappa value of the SMOTE-IVY-SE-CNN-BiLSTM is 0.9799, F1 score is 0.9773, and accuracy is 0.9722. The higher Kappa value indicates stronger consistency between the model's classifications and the actual labels, suggesting greater stability and trustworthiness for the model in real-world applications. The higher F1 score means the model can not only maintain high precision but also identify potentially overlooked anomalies. The advantage of the accuracy demonstrates that the SMOTE-IVY-SE-CNN-BiLSTM has a stronger discriminative ability in classification tasks, enabling more accurate health status assessments of air compressors and reducing misclassifications.

Following the foundational comparative analysis of the air compressor health assessment model, this study systematically designed noise interference contrast experiments by injecting Gaussian white noise (standard deviation: 5% of the baseline signal amplitude) into raw multi-source sensor time-series data. This controlled noise injection effectively replicates potential interference scenarios in real-world industrial environments. The experiments not only quantify the impact of noise on diagnostic outcomes but also demonstrate the model's strong anti-interference capabilities, thereby providing a foundation for future optimizations to enhance the performance of health state classification models.

Table 7 compares the performance of various models on the test set with 5% added noise. As shown in the table, under 5% noise conditions, our model demonstrates strong anti-interference capability, achieving accuracy, Kappa coefficient, and F1 score of 92.13%, 91.19%, and 95.06%, respectively. These metrics surpass those of other algorithms, including the BP, SVM, RF, LSTM, BiLSTM, and PSO-BiLSTM. Compared to its performance under noise-free conditions, the degradation magnitude is relatively small. This indicates that

in the presence of noise interference, our model can effectively suppress noise-induced disturbances to classification decision boundaries through three key mechanisms: the SMOTE-enhanced sample balance, the SE-CNN feature selection mechanism for improving signal-to-noise ratio, and the dynamic feature perception capability of the IVY-BiLSTM. These findings provide reliability assurance for online diagnostic systems in industrial environments with complex noise conditions, while establishing an extendable technical pathway for future anti-interference optimization.

Table 7. Performance of the different models with added noise.

Method	Kappa	F1	Accuracy
BP	0.4729	0.4733	0.4962
SVM	0.6029	0.6059	0.6217
RF	0.6073	0.6401	0.6229
LSTM	0.7126	0.7352	0.7189
BiLSTM	0.7430	0.7905	0.7576
PSO-BiLSTM	0.8395	0.8876	0.8519
IVY-SE-CNN-BiLSTM	0.9119	0.9506	**0.9213**

The best results are indicated in bold font.

Overall, the SMOTE-IVY-SE-CNN-BiLSTM model demonstrates better performance in air compressor health status assessment, with outstanding results in the three key metrics: Kappa value, F1 score, and accuracy. We believe that this achievement is mainly attributed to the effective execution of each module in the model. The SMOTE module, as the foundation of the model, effectively increases the minority class samples, promoting data class balance and thereby improving the model's capability to classify minority classes. The SE-CNN module is essential for extracting both local and global features. It not only effectively expands the model's perceptual range to better facilitate global feature modeling but also significantly boosts the model's capacity to capture intricate features. This is particularly important when handling high-dimensional and complex structured input data. Without this module, redundant information could interfere with the model's accurate analysis of features, which would reduce its performance. The introduction of the IVY optimization algorithm further optimizes the hyperparameters of the BiLSTM model, enabling it to learn features from the data more efficiently. Compared to traditional hyperparameter selection methods, the IVY optimization algorithm explores the hyperparameter space more comprehensively, selecting the most suitable hyperparameter configuration for the current task.

Through the collaborative effort of these modules, our model not only performs excellently in data preprocessing, feature extraction, and model optimization but also achieves outstanding performance in the final health status evaluation task. This underscores the critical role each module plays in the model, collectively contributing to the overall enhancement of performance.

To substantiate this claim, we will present the ablation study in the following section.

4.3. Ablation Study

To validate the effectiveness of the proposed improvements in the air compressor health status assessment task, particularly in improving accuracy, an ablation study is conducted. To assess the contribution of each component to the overall performance, we conduct an ablation study by selectively removing key modules. The five models—IVY-SE-CNN-BiLSTM, SE-CNN-BiLSTM, CNN-BiLSTM, and BiLSTM—allowed for a detailed evaluation of the impact of each module on the model's effectiveness. Specifically, the

study tests the impact of the feature extraction module CNN, the Ivy algorithm, and the SE attention mechanism on model performance.

As shown in Table 8, by comparing the training and testing results of different models, the impact and enhancement of each module on model performance can be clearly observed.

Table 8. Performance of the different models in the ablation study.

Method	Training			Test		
	Kappa	**F1**	**Accuracy**	**Kappa**	**F1**	**Accuracy**
IVY-SE-CNN-BiLSTM	0.9548	0.9803	**0.9779**	0.9370	0.9725	**0.9689**
SE-CNN-BiLSTM	0.9464	0.9711	0.9673	0.9119	0.9622	0.9407
CNN-BiLSTM	0.9413	0.9501	0.9458	0.9047	0.9361	0.9219
BiLSTM	0.9347	0.9401	0.9337	0.9005	0.9117	0.9091

The best results are indicated in bold font.

The IVY-SE-CNN-BiLSTM model demonstrates strong performance in the air compressor health status assessment task.

However, after removing the IVY optimization, the model's performance significantly deteriorates. Although the BiLSTM can still effectively process data, the lack of precise hyperparameter tuning provided by the IVY algorithm reduces the model's adaptability and generalization ability when facing complex tasks, resulting in greater fluctuations in the stability of the test results. The accuracy on the test set decreased by 3.35% after removing the IVY optimization, highlighting the significant impact of hyperparameter tuning on improving model performance.

In subsequent experiments, removing the SE module further degraded performance. The SE module plays a critical role in dynamically recalibrating the significance of feature weights; without it, the model lost its ability to prioritize the most relevant features.

Finally, when the CNN module was removed, the model lost its powerful feature extraction capability, which made it weaker when handling complex patterns in the data. Without the CNN module, the model struggled to capture critical local features from the input signals, causing the accuracy on the test set to decrease by 1.41%. This result further emphasizes the crucial role of feature extraction in enhancing the model's recognition capability and handling complex data.

In conclusion, the ablation experiments clearly highlight the key role of each module in the model. Although the introduction of the SE module contributed to an improvement in overall performance, its effect on performance enhancement was relatively minor within the SE-CNN module combination. Therefore, future research could explore the incorporation of more efficient attention mechanisms to further improve the model's ability in feature selection and representation. This direction provides important guidance for future work.

4.4. Model Performance Across Diverse Datasets

To assess the model's generalization ability and robustness, this section conducts experiments using three distinct datasets: (1) hydraulic system condition assessment dataset, (2) motor condition assessment dataset, and (3) Case Western Reserve University bearing fault dataset. After applying the model to these varied datasets, the results consistently exhibit good classification performance, as shown in Table 9, further reinforcing the model's generalization capacity. Regardless of the complexity of the datasets or the variations in data distributions and label categories, the model exhibits a relatively strong capacity to learn from the data and identify key features, enabling it to achieve precise and reliable classification outcomes. This suggests that the model exhibits remarkable flexibility and

applicability, thereby demonstrating that its ability and reliability in specific industrial scenarios have been validated.

Table 9. Model performance across diverse datasets.

Dataset	Training			Test		
	Kappa	**F1**	**Accuracy**	**Kappa**	**F1**	**Accuracy**
1	0.9378	0.9631	0.9527	0.9007	0.9423	0.9401
2	0.9215	0.9411	0.9470	0.8869	0.9211	0.9307
3	0.9071	0.9025	0.9433	0.8634	0.8859	0.9225
Our dataset	0.9548	0.9803	**0.9779**	0.9370	0.9725	**0.9689**

The best results are indicated in bold font.

5. Conclusions

This study presents an air compressor health status assessment model based on the SMOTE-IVY-SE-CNN-BiLSTM, which integrates the data handling capabilities of the SMOTE, the feature extraction power of the SE-CNN, and the local–global optimization strategy of the IVY to enhance the classification accuracy of compressor health status. In our model, every module is meticulously crafted and follows a set of essential design principles to maximize its effectiveness in enhancing accuracy. The SMOTE-IVY-SE-CNN-BiLSTM model achieved an accuracy of 98.58% on the training set and 97.22% on the test set. Ablation experiments show that our model effectively addresses issues in health status evaluation, with each module fully leveraging its strengths. Compared to other classification models, our approach outperforms in terms of performance.

Although the SMOTE alleviates the class imbalance problem, the improvement in performance is limited. Future research could focus on adopting more refined sampling methods that incorporate data distribution information to optimize the sample generation process. Additionally, due to the integration of multiple modules, the model's computational complexity is relatively high. In the future, we plan to adopt a more lightweight architecture: (1) Model compression: we will achieve it by implementing a customized model compression mechanism to reduce the number of model parameters and computational complexity, while preserving critical feature extraction and classification capabilities. (2) Simplification of SE attention mechanism: we will use the appropriate and effective lightweight attention to maintain competitive model performance and organically integrate it into the feature extraction algorithm to reduce the complexity of our model. (3) If there is a higher demand for simplification, knowledge distillation may be considered to use to compress the framework without compromising functionality.

Finally, while this study demonstrates superior robustness and anti-interference capability in current-stage air compressor health diagnosis tasks, two critical challenges remain in real-time operational environments: (1) complex multi-source interference and diverse health state characterizations and (2) generalized recognition requirements for non-standard health patterns under non-steady-state operating conditions. To further enhance the model's robustness and generalization adaptability in dynamic industrial environments, subsequent work will focus on the following technical pathways. We will embed the model into real-time diagnostic systems and deploy an online active learning framework based on incremental learning. By dynamically integrating real-time operational condition data streams with the SMOTE-augmented datasets, we will develop adaptive parameter updating strategies based on data distribution shift detection, progressively reduce the proportion of synthetic samples, and ultimately achieve lossless decommissioning of the SMOTE module through completeness validation of real fault data. The model's outputs will be integrated with professional experience-based judgments and field operational

parameters, correlating historical maintenance records with real-time diagnostic results. Reverse optimization of feature selection thresholds and dynamic refinement and expansion of health state classification granularity will enable the distinction of finer health state categories to improve diagnostic accuracy and application flexibility.

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

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. Due to privacy and ethical restrictions, the data are not publicly accessible. The datasets used in the Section 4.4 can be accessed from the link: (1). https://archive.ics.uci.edu/dataset/447/condition+monitoring+of+hydraulic+systems (accessed on 11 March 2025), (2). https://pan.baidu.com/s/19OddVvs6CYPh-giRRLJfwg Extract code: cgvn (accessed on 11 March 2025), (3). https://www.kaggle.com/datasets/brjapon/cwru-bearing-datasets (accessed on 11 March 2025).

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

Abbreviations

The following abbreviations are used in this manuscript:

SMOTE	Synthetic Minority Over-sampling Technique
IVYA	Ivy algorithm
SE	Squeeze-and-Excitation
CNNs	Convolutional Neural Networks
BiLSTM	Bidirectional Long Short-Term Memory network
PCA	Principal Component Analysis
MDDM	Multi-label Dimensionality reduction via Dependence Maximization
BP	Neural Network
PSO-LSTM	Particle Swarm Optimization–Bidirectional Long Short-Term Memory network
RF	Random Forest
SVM	Support Vector Machine

Appendix A

(1) The brief core algorithm implementation of the IVY-BiLSTM can be seen at https://pan.baidu.com/s/1mbWG_aCR8QPhfd0PL1lqUQ Extract code: 3s3s (accessed on 9 March 2025).

(2) The complete code for the SMOTE can be referenced at https://github.com/kaushalshetty/SMOTE (accessed on 12 March 2025).

(3) The complete code for the SE can be referenced at https://github.com/hujie-frank/SENet/ (accessed on 12 March 2025).

References

1. Md Nor, N.; Che Hassan, C.R.; Hussain, M.A. A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems. *Rev. Chem. Eng.* **2020**, *36*, 513–553. [CrossRef]
2. Ameri, H.; Farzaneh-Gord, M. Effect of compressor emergency shutdown in various compressor station configurations on natural gas transmission pipeline. *Iran. J. Sci. Technol. Trans. Mech. Eng.* **2021**, *45*, 427–440. [CrossRef]
3. Sharp, M. Observations on developing reliability information utilization in a manufacturing environment with case study: Robotic arm manipulators. *Int. J. Adv. Manuf. Technol.* **2019**, *102*, 3243–3264. [CrossRef] [PubMed]

4. Daya, A.A.; Lazakis, I. Developing an advanced reliability analysis framework for marine systems operations and maintenance. *Ocean Eng.* **2023**, *272*, 113766. [CrossRef]
5. Fu, J.; Zhang, G.; Sun, X.; He, T.; Li, H. A combinatorial prediction model for the performance of fuel cell air compressor with air bearings considering accelerated decay characteristics. *Appl. Therm. Eng.* **2024**, *253*, 123784. [CrossRef]
6. De Giorgi, M.G.; Menga, N.; Ficarella, A. Exploring Prognostic and Diagnostic Techniques for Jet Engine Health Monitoring: A Review of Degradation Mechanisms and Advanced Prediction Strategies. *Energies* **2023**, *16*, 2711. [CrossRef]
7. 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]
8. Thakkar, U.; Chaoui, H. Prognostic and health management of an aircraft turbofan engine using machine learning. In Proceedings of the 2023 IEEE Vehicle Power and Propulsion Conference (VPPC), Milan, Italy, 24–27 October 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
9. Habbouche, H.; Benkedjouh, T.; Amirat, Y.; Benbouzid, M. Rotating machine bearing health prognosis using a data driven approach based on KS-density and BiLSTM. *IET Sci. Meas. Technol.* **2025**, *19*, e12215. [CrossRef]
10. Nambiar, A.; Aravinth, S.; Sugumaran, V.; Ramteke, S.M.; Marian, M. Prediction of air compressor faults with feature fusion and machine learning. *Knowl.-Based Syst.* **2024**, *304*, 112519. [CrossRef]
11. Shar, M.A.; Muhammad, M.B.; Mokhtar, A.A.B.; Soomro, M. A Novel Energy Performance-Based Diagnostic Model for Centrifugal Compressor using Hybrid ML Model. *Arab. J. Sci. Eng.* **2024**, *49*, 14835–14853. [CrossRef]
12. Tran, V.T.; AlThobiani, F.; Tinga, T.; Ball, A.; Niu, G. Single and combined fault diagnosis of reciprocating compressor valves using a hybrid deep belief network. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2018**, *232*, 3767–3780. [CrossRef]
13. Babu, R.M.; Satamraju, K.P.; Gangothri, B.N.; Malarkodi, B.; Suresh, C.V. A Hybrid model using genetic algorithm for energy optimization in heterogeneous internet of blockchain things. *Telecommun. Radio Eng.* **2024**, *83*, 1–16. [CrossRef]
14. Sun, X.; Liu, H. Multivariate short-term wind speed prediction based on PSO-VMD-SE-ICEEMDAN two-stage decomposition and Att-S2S. *Energy* **2024**, *305*, 132228. [CrossRef]
15. Ebid, A.M.; Abdel-Kader, M.Y.; Mahdi, I.M.; Abdel-Rasheed, I. Ant Colony Optimization based algorithm to determine the optimum route for overhead power transmission lines. *Ain Shams Eng. J.* **2024**, *15*, 102344. [CrossRef]
16. Ghasemi, M.; Zare, M.; Trojovský, P.; Rao, R.V.; Trojovská, E.; Kandasamy, V. Optimization based on the smart behavior of plants with its engineering applications: Ivy algorithm. *Knowl.-Based Syst.* **2024**, *295*, 111850. [CrossRef]
17. Ouyang, H.; Li, W.; Gao, F.; Huang, K.; Xiao, P. Research on Fault Diagnosis of Ship Diesel Generator System Based on IVY-RF. *Energies* **2024**, *17*, 5799. [CrossRef]
18. Xu, J.; Liu, J.; Yin, J.; Sun, C. A multi-label feature extraction algorithm via maximizing feature variance and feature-label dependence simultaneously. *Knowl.-Based Syst.* **2016**, *98*, 172–184. [CrossRef]
19. Jogin, M.; Madhulika, M.S.; Divya, G.D.; Meghana, R.K.; Apoorva, S. Feature extraction using convolution neural networks (CNN) and deep learning. In Proceedings of the 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Karnataka, India, 18–19 May 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 2319–2323.
20. Luo, S.; Wang, B.; Gao, Q.; Wang, Y.; Pang, X. Stacking integration algorithm based on CNN-BiLSTM-Attention with XGBoost for short-term electricity load forecasting. *Energy Rep.* **2024**, *12*, 2676–2689. [CrossRef]
21. Vatsa, A.; Hati, A.S. Insulation aging condition assessment of transformer in the visual domain based on SE-CNN. *Eng. Appl. Artif. Intell.* **2024**, *128*, 107409. [CrossRef]
22. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]
23. Dao, F.; Zeng, Y.; Qian, J. Fault diagnosis of hydro-turbine via the incorporation of bayesian algorithm optimized CNN-LSTM neural network. *Energy* **2024**, *290*, 130326. [CrossRef]
24. Lv, D.; Zhao, C.; Ye, H.; Fan, Y.; Shu, X. GS-YOLO: A lightweight SAR ship detection model based on enhanced GhostNetV2 and SE attention mechanism. *IEEE Access* **2024**, *12*, 108414–108424. [CrossRef]
25. Hameed, Z.; Garcia-Zapirain, B. Sentiment classification using a single-layered BiLSTM model. *IEEE Access* **2020**, *8*, 73992–74001. [CrossRef]
26. Kowsher, M.; Tahabilder, A.; Sanjid, M.Z.I.; Prottasha, N.J.; Uddin, S.; Hossain, A.; Jilani, A.K. LSTM-ANN & BiLSTM-ANN: Hybrid deep learning models for enhanced classification accuracy. *Procedia Comput. Sci.* **2021**, *193*, 131–140.
27. Pandey, V.; Lilhore, U.K.; Walia, R.; Alroobaea, R.; Alsafyani, M.; Baqasah, A.M.; Algarni, S. Enhancing heart disease classification with M2MASC and CNN-BiLSTM integration for improved accuracy. *Sci. Rep.* **2024**, *14*, 24221. [CrossRef]
28. Bouraiou, A.; Slimani, A.; Neçaibia, A.; Lachtar, S.; Labed, N.; Dahbi, A.; Rouabhia, A.; Rahli, C. A Simple Platform to Establish Supervision, Monitoring, and Control using Arduino/Ethernet Shield and SCADA via Industrial Modbus TCP/IP Communication Protocol. *Alger. J. Renew. Energy Sustain. Dev.* **2024**, *6*, 84–89.
29. Zhang, Z.; Qin, B.; Gao, X.; Ding, T.; Zhang, Y.; Wang, H. SE-CNN based emergency control coordination strategy against voltage instability in multi-infeed hybrid AC/DC systems. *Int. J. Electr. Power Energy Syst.* **2024**, *160*, 110082. [CrossRef]

30. Longjam, T.; Kisku, D.R.; Gupta, P. Writer independent handwritten signature verification on multi-scripted signatures using hybrid CNN-BiLSTM: A novel approach. *Expert Syst. Appl.* **2023**, *214*, 119111. [CrossRef]
31. Li, J.; Cheng, J.; Shi, J.; Huang, F. Brief introduction of back propagation (BP) neural network algorithm and its improvement. In Proceedings of the Advances in Computer Science and Information Engineering: Volume 2, Zhengzhou, China, 19–20 May 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 553–558.
32. Peng, S.; Zhu, J.; Wu, T.; Yuan, C.; Cang, J.; Zhang, K.; Pecht, M. Prediction of wind and PV power by fusing the multi-stage feature extraction and a PSO-BiLSTM model. *Energy* **2024**, *298*, 131345. [CrossRef]
33. Speiser, J.L.; Miller, M.E.; Tooze, J.; Ip, E. A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst. Appl.* **2019**, *134*, 93–101. [CrossRef]
34. Mathur, A.; Foody, G.M. Multiclass and binary SVM classification: Implications for training and classification users. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 241–245. [CrossRef]

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