Air Compressor Fault Diagnosis Algorithm Using Voiceprint Feature Fusion

Guoliang Feng, Fumin Wang, Tianming Yu, and Ce Xu

Abstract—**As an important piece of equipment in industrial production, the health of an air compressor directly impacts the success of production. Therefore, researching fault diagnosis methods for air compressors is of significant importance in improving the continuity, reliability, and safety of production. Traditional fault diagnosis methods, however, struggle to obtain accurate fault features. The measurement of feature distribution differences between various working conditions lacks sufficient domain adaptability, making it challenging to achieve high recognition accuracy. Additionally, the operation of air compressors generates background noise, which can introduce interference and impact the accuracy in fault detection. In order to overcome these limitations, a fault diagnosis method for air compressors based on feature fusion is proposed. Firstly, the Mel-frequency cepstral coefficients (MFCC) features and wavelet transform features of the air compressor are extracted separately. Then, late fusion is applied at the decision level to combine confidence scores and predicted bounding boxes. The best network model is determined based on evaluation metrics to complete the classification. Based on the experimental results analysis, the feature fusion method demonstrated superior recognition performance.**

Index Terms—**Feature fusion, Voiceprint recognition, Fault recognition, Feature extraction.**

I. INTRODUCTION

THE operational status of an air compressor is directly related to the safety of the entire workflow. Therefore, it is necessary to diagnose and monitor the air compressor for faults [1], [2]. Recently, deep learning (DL) methods have been widely applied, among which Convolutional Neural Network (CNN), leveraging their powerful non-linear mapping capability, are extensively used to extract features and classify raw vibration signals. They are employed to address mechanical fault diagnosis problems traditionally reliant on signal processing and expert knowledge [3], [4], [5], [6], [7]. He et al. [8] proposed a novel deep learning framework called ResNet-ELM (RNELM) for classifying mechanical faults. The time-domain analysis of raw fault signals allows for extracting comprehensive and significant fault features encompassing both time-scale and frequency-scale information [9]. Additionally, there exists a correlation between the time scale and frequency scale [10]. Thus, the method begins by transforming one-dimensional time data from different fault categories into two-dimensional time-frequency images using Continuous Wavelet Transform (CWT). Following this, a deep ResNet structure is employed to extract comprehensive data features from these time-frequency images, facilitating precise fault diagnosis through advanced feature extraction techniques [11]. Prosvirin et al. [12] have developed a new method for air compressor fault diagnosis, utilizing an end-to-end approach. They introduced a deep neural architecture combining an Autoencoder (AE) and CNN, aimed at enhancing diagnostic intelligence and accuracy. In current CNN-based methods, the performance of CNN is influenced by the parameters of the convolutional kernel function, which in turn affects the diagnostic accuracy. However, the set parameters cannot adaptively reflect the differences in feature distribution under various operating conditions, thus reducing the diagnostic accuracy [13]. Nowadays, industrial equipment can generate a large amount of time series and diverse data, leading to a growing interest in data-driven fault diagnosis methods in recent years. Wang et al. [14] Detecting faults in air compressors through an integrated approach using CWT, dual-channel convolutional neural network (DCCNN), and long short-term memory network (LSTM). Hamzeh et al. [15] used an Adaptive Neuro-Fuzzy Inference System (ANFIS) to conduct analyses related to the reliability and stability of the production process. Soltanali et al. [16] prioritizing Risk Priority Numbers (RPN) for fault diagnosis using Support Vector Machines (SVM). Li et al. [17] pointed out that although SVM and Particle Swarm Optimization (PSO) can achieve fault type classification, the inherent limitations of traditional machine learning, such as sensitivity to feature selection dimensions, the need for independent feature extraction and classification, and shallow learning structures, prevent them from meeting the requirements. [18]. The CNN, as a mainstream deep learning approach, can effectively address these issues. [19], [20]. Wang et al. [21] proposed a DCNN that uses time-frequency images as inputs to address the classification of planetary gear bearing faults. Xiao et al. [22] modified a parallel ResNet network to extract time-frequency image features. Zhang et al. [23] employed transfer learning to train the model. Tian et al. [24] utilized residual networks to extract feature information, employing the ResNet50 network for feature extraction and augmentation. They fine-tuned the transfer learning using a

Manuscript received January 5, 2024; revised August 16, 2024.

This work was supported by the Natural Science Foundation of China under Grant No. 52376096.

Guoliang Feng is a professor of the School of Automation Engineering, Northeast Electric Power University, Jilin, Jilin 132012 China. (e-mail: fengguoliang@neepu.edu.cn).

Fumin Wang is a postgraduate student of the School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin, Jilin 100190 China. (e-mail: wfm521622@foxmail.com).

Tianming Yu is an associate professor of the School of Automation Engineering, Northeast Electric Power University, Jilin, Jilin 132012 China. (e-mail: 20192928@neepu.edu.cn).

Ce Xu is a senior engineer of the Department of China Petroleum Liaohe Equipment Company, Panjin, Liaoning 124000 China. (e-mail: xuce2001@163.com).

pre-trained dataset of CNN [25]. Liang et al. [26] integrated expanded convolution with residual networks to enhance fault diagnosis capability in noisy environments. While CNNs have shown significant improvements, the existing issues in fault diagnosis cannot be overlooked.

The current methods solely focus on minimizing the overall feature distribution gap without taking into account the variations in class-specific features between the source and target domains. Not only does it blur the decision boundaries between classes, but it also decreases accuracy. In CNN architectures, simply concatenating or adding two convolutional layers does not effectively enhance information complementarity. On the other hand, although richer convolutional features are advantageous for enhancing fault classification accuracy, they can also lead to fault diagnosis model networks with deep layers. This, in turn, can cause issues such as gradient vanishing/exploding and challenges in parameter optimization. In practical industrial production environments, significant machine noises are common, making it challenging to extract features for noise identification and reducing the accuracy of fault determination.

To effectively address various existing issues, this paper proposes a feature fusion-based model for air compressor fault diagnosis. The contributions of this paper are outlined as follows:

1) We developed a neural network architecture search algorithm utilizing transfer learning and Bayesian optimization. Changing the network structure to expand artificially designed networks into hypernetworks, we conducted a structural search using network morphism. Subnetworks of the supernet were obtained as candidate networks for training and evaluation. The best model structure for classification was ultimately determined based on the evaluation metrics.

2) In the data preprocessing stage, we introduce noise with a random signal-to-noise ratio to simulate real working environments. We construct an ensemble classifier to train and extract features from the wavelet scattering transform, and a neural network to train the MFCC features. By combining the input of wavelet scattering transform and MFCC features, our approach encompasses both types of features. Furthermore, considering the issue of underfitting in single-model classification, we employ late fusion to combine the output of the wavelet classifier and CNN. The combination generates a vector that displays the relative confidence level of decisions. which is then multiplied to create a late fusion system. This approach aims to enhance classification accuracy and generalization performance. This method can clarify the decision boundaries of different categories in the target domain, thereby improving fault diagnosis accuracy.

II. Theoretical Background

A. Mel Frequency Cepstral Coefficients

MFCC features are classical auditory perceptual features based on the human ear's perception of sound. Due to the nonlinear relationship between perceived pitch and frequency in the human auditory system, it is necessary to transform the signal into a frequency power spectrum and apply cepstral analysis to obtain features suitable for auditory perception. Therefore, extracting the optimal parameter representation of noise signals can enhance fault detection performance. [27].

The relationship between sound frequency and Mel frequency transformation is as follows:

$$Mel(f) = 2595\lg(1 + f / 700) \tag{1}$$

Given an input sound frequency f, the computation of MFCC involves six steps: framing, windowing, fast Fourier transform, filter bank processing, logarithmic operation, and discrete cosine transform (DCT). The process of MFCC feature extraction is illustrated in Figure 1.

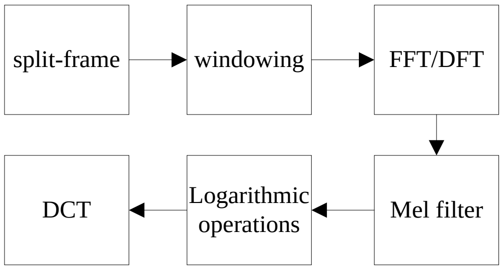


Fig. 1. The process of MFCC feature extraction

The sound signal and frames undergo changes; for instance, the slope of resonance peaks varies during transition periods. Therefore, it is essential to incorporate features related to the temporal changes of cepstral characteristics. In total, 13 delta or velocity features (12 cepstral features plus energy) and 39 double delta or acceleration features are added [28]. The energy in the frame of the signal x in the window from time sample t_1 to time sample t_2 is represented by the following equation:

$$Energy = \sum X^2[t] \tag{2}$$

The frame-to-frame variations of cepstral or energy features in Equation 2 correspond to each of the 13 delta features, while the delta features correspond to each of the 39 double delta features.

B. Wavelet Scattering Transform

Usually, when an air compressor is disturbed or malfunctions, it may exhibit abnormal vibrations and produce unusual noises, which align with elastic theory and wave theory. [29], we can find that:

$$\sigma = E\varepsilon \tag{3}$$

$$\varepsilon = \frac{v}{c} \tag{4}$$

where σ is the stress generated by the vibrations of the air compressor, E is elastic modulus, ε is strain and v, c is particle vibration velocity, propagation speed of vibration wave, respectively Substitution of EQ. (3) into EQ. (4) gives:

$$\sigma = \frac{Ev}{c} \tag{5}$$

Traditional wavelet transforms ignore high-frequency signals when decomposing the low-frequency part of the signal, leading to the loss of high-frequency signal details. However, the wavelet scattering transform can recover lost high-frequency signals by utilizing the wavelet transform modulus after extracting low-frequency features. This paper utilizes wavelet scattering transform based on wavelet transform theory to extract more complex features. The wavelet scattering transform can reduce sample data size, minimize intra-class differences, and retain distinguishability among different classes.

The wavelet scattering transform consists of three cascaded stages. In the first stage, the signal x is decomposed and convolved with the mother wavelet ψ centered at frequency λ, yielding $x*\psi^{\lambda}$. In the second stage, signals often undergo convolutional non-linear operations, which typically increase their frequency and can compensate for information loss caused by downsampling [30]. Finally, applying a time-averaged low-pass filter in the form of a scaling function ϕ yields the absolute convolution signal:

The zeroth-order scattering coefficient describes the signal's local translational invariance: s_0

$$s_0 = x*\phi \tag{6}$$

In each level, the loss of high-frequency components in the convolution signal is due to the averaging operation. Recovering these components requires convolving the signal with the wavelet of the next level.

Therefore, the first-order scattering coefficient S_1 is defined as the average absolute amplitude of any wavelet coefficient at scale $1 \le j \le J$ over a half-overlapping time window of size 2^j :

$$S_1 = \left| x*\psi_{\lambda 1} \right| \tag{7}$$

III. Suggested methods

A. Presentation of THE Overall Structure of The Methods

During the operation of air compressors, due to the high noise often present in the working environment, fault diagnosis is considered a challenging task. To address this issue, this paper proposes a fault diagnosis method based on feature fusion. The method aims to integrate wavelet features with MFCC features to enhance recognition accuracy, while also leveraging the classification capability of ensemble classifiers and the stable global dynamic search capability of Bayesian optimization. The overall method is illustrated in Figure 2.

1) Preprocess the noisy data from the air compressor and extract MFCC features and wavelet scattering transform features separately.

2) Train the MFCC features using CNN and train the wavelet scattering transform features using an ensemble classifier.

3) Utilize the Bayesian optimization algorithm to optimize the hyperparameters of the neural network to enhance the model accuracy.

4) Predict the probabilities of the target noise belonging to each class separately for the two classification models.

5) Perform decision-level fusion of the classification results from the two different feature recognition models to obtain the fused decision outcome.

B. Bayesian Optimization

Let $f(x)$ denote the mapping from the hyperparameter vector x to the model's generalization performance, where $x \in X$, $X \subseteq R$, d and X represents the dimensions of a hyperparameter space of size d [31]. The objective of hyperparameter optimization is to search within this d -dimensional hyperparameter space for the optimal hyperparameter x^* that maximizes the model's generalization performance. The expression is defined as follows:

$$x^* = \arg\max f(x) \tag{8}$$

As the function f evaluates the model's generalization metrics concerning its hyperparameters, such as generalization accuracy, it is a high-cost black-box objective function. This is because assessing f demands substantial computational resources and time for each training and evaluation of a set of hyperparameters.

We propose using the Bayesian optimization algorithm to optimize the following network hyperparameters: initial learning rate, stochastic gradient descent momentum, L2 regularization strength, maximum number of epochs, validation frequency, and minimum batch size. The input comprises training and validation data used to create the objective function for the Bayesian optimizer. This function trains a convolutional neural network and subsequently returns the classification error rate on the validation set. The algorithm flow is illustrated in Algorithm 1. Since Bayesian optimization is used to choose the best model based on the validation set error rate, The final outcome may suffer from overfitting, so it's necessary to validate on an independent test set, assess the error, and then select the best model. Bayesian optimization is conducted to further enhance the neural network by minimizing the classification error on the validation set as much as possible. The Bayesian optimization process is illustrated in Algorithm 1.

C. Comprehensive Classifier Based on Subspace Method

Using ensemble classifier based on subspace method, we can extract wavelet transform features from the noise generated by an air compressor and conduct classification. The subspace method involves selecting a subset of features from the original data to create new data [32]. As shown in Figure 3, the extracted wavelet transform features are used to create a new dataset by selecting various wavelet features from each sample. The new dataset is then used to train a classifier. The remaining features from each sample can be used to create a second new dataset, which can be utilized to train a second classifier, Continue this iterative process to train the classifier ensemble [33].

The wavelet transform features extracted from the dataset are inputted into various classifiers for parallel training. Each classifier includes the target prediction output, and the predicted output probabilities from different models are weighted and averaged to create the final dataset. The ensemble classifier is then trained using this dataset to generate the final classification result of the dataset [34]. Figure 4 illustrates the training process of the ensemble classifier.

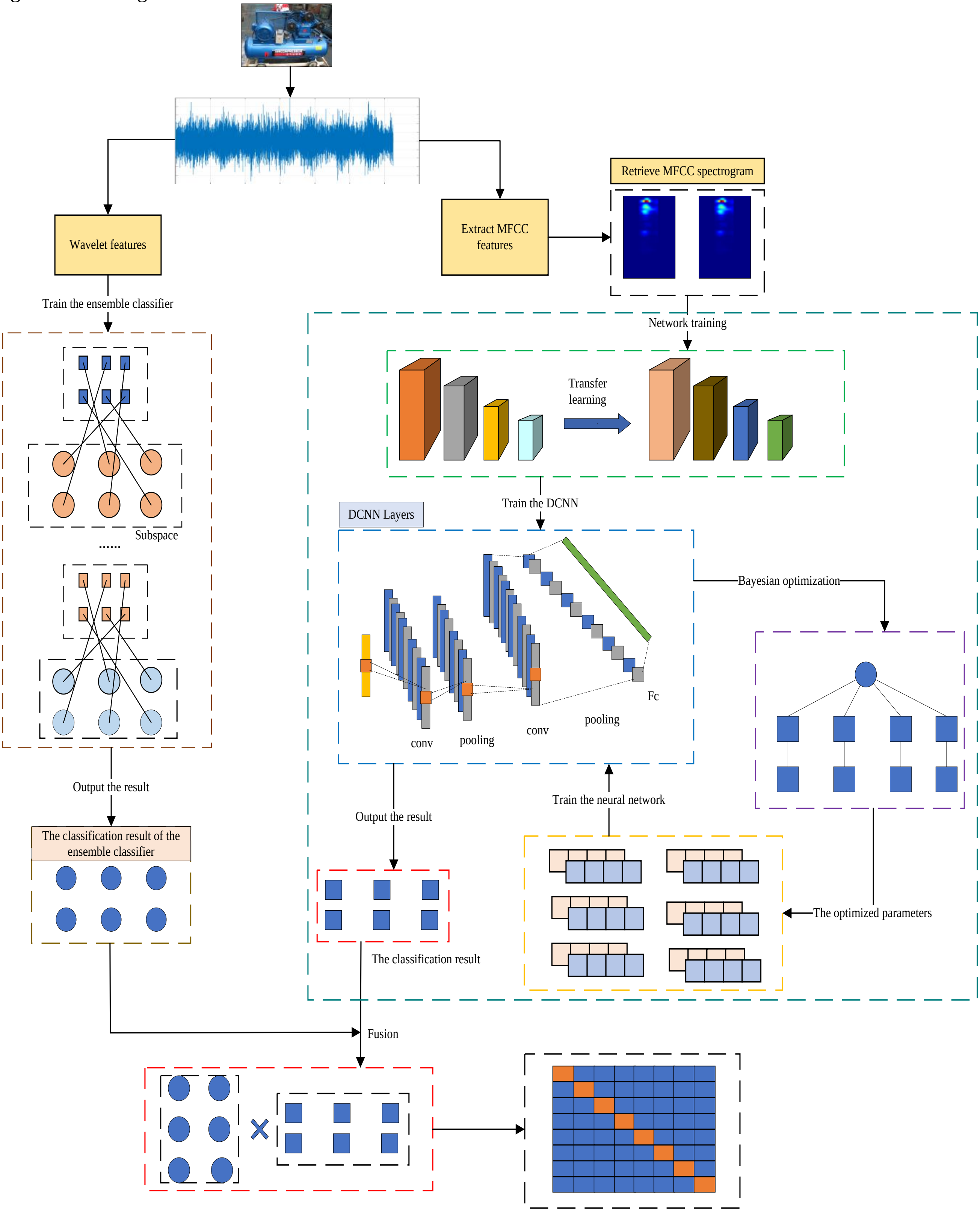


Fig. 2. Overall Flow

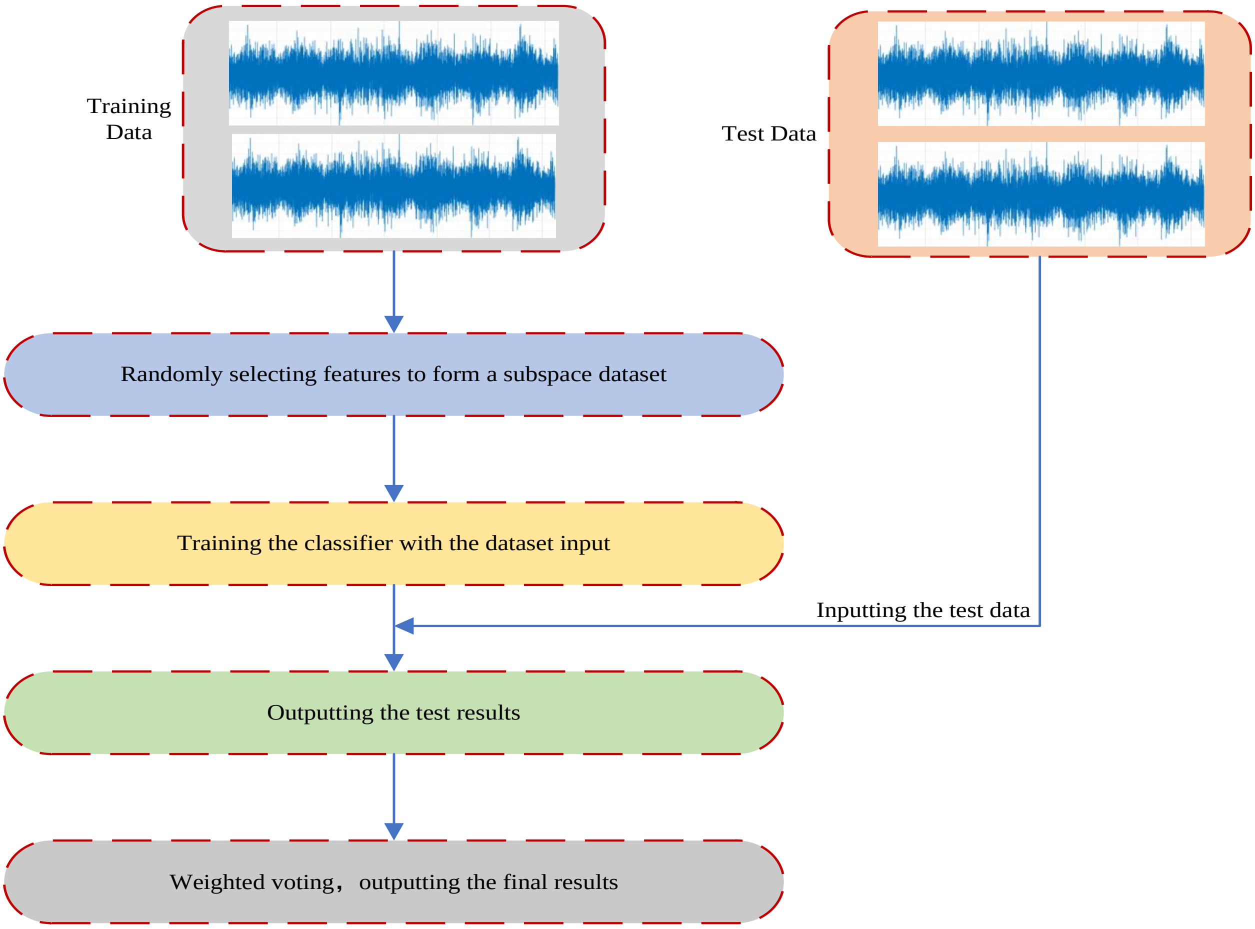


Fig. 3. Subspace method

Algorithm 1 Bayesian Optimization

Input:

Model f , Collection function α , Existing samples D .

Output:

Hyperparameter vector x^* .

1: **for** t =1, 2, ... , T **do**

2: Maximize the acquisition function to get the next evaluation point $x_t = \arg\max_{x \in X^\alpha} (x \mid D_{1:t-1})$.

3: the value of the objective function $y_t = f(x_t) + \epsilon_t$.

4: Integration of data $D_t = D_{t-1} U x_t$, y_t , And update the model.

5: end for

Algorithm 2 Integrated classifier based on subspace method

Input:

Set of wavelet samples X_i , Set of wavelet labels Y_i , Set of old classifiers E .

Output:

Integrated Classifier E_{new} .

1: Update the weights for each classifier.

2: Update the weights for each classifier .

3: For each classifier $\psi \in E$.

4: For Y_i each unique label in y .

5: Find every sample $X^y \subset X_i$ label y .

6: **if** y is a new label for classifier ψ **then**

7: X^y to update ψ to make it acceptable to the new category.

8: else

9: Use a multi-objective algorithm to select a subset of the samples of X^y .

10: the Update ψ with the selected samples.

11: end if

12: Generating new subspaces through existing classifiers.

13: The weight of ψ_{new} is set to 1.

14: Train the classifier ψ_{new} with the new subspace and Data (X_i , Y_i) and add the set E_{new} .

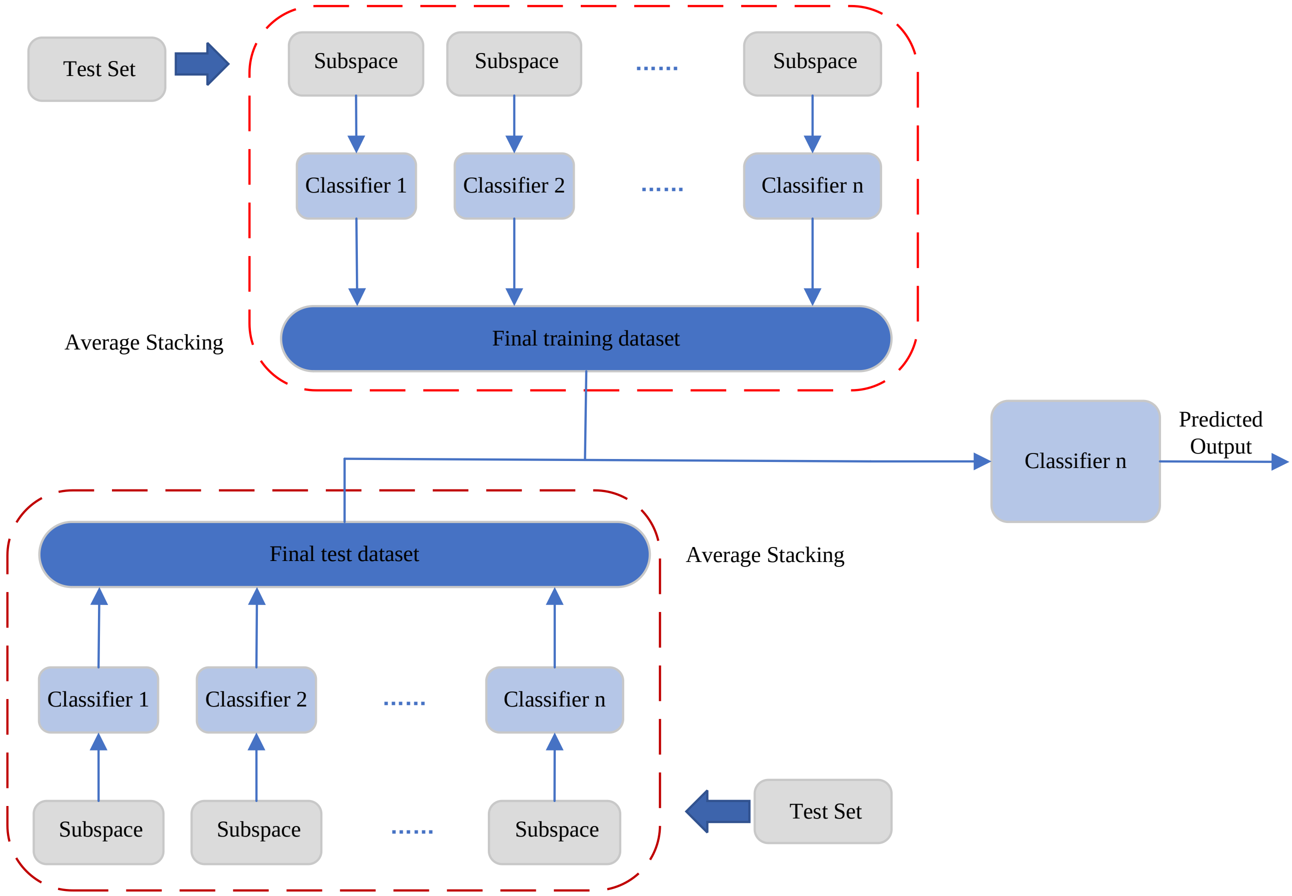


Fig. 4. Flow of integrated classifier training

D. Feature Fusion

The focus is on integrating the classification results of different features under varying conditions. The classification results of MFCC features and wavelet transform features are fused at the decision level to combine the recognition results from different features and classifiers.

The specific steps are as follows:

1) Utilize a CNN model to extract MFCC features from the noise of the air compressor and perform classification. Use an ensemble classifier to extract wavelet features and conduct classification. Predict the probability of the fault category for each of the two classification models.

2) Fuse the classification results of the low-recognition-rate model with those of the high-recognition-rate model to enhance the recognition performance and decision-making ability of the model. Fuse the classification results obtained from several CNN models and the ensemble classifier separately, then select the optimal fusion model through comparison.

3) Perform decision-level fusion of the classification results from different features in the deep learning models. Utilize post-fusion theory to combine the optimal CNN classification results and the ensemble classifier at the decision level, resulting in the final classification model.

4) Invoke the model to plot a confusion matrix to visualize the fused classification accuracy and obtain the final decision results.

IV. Experimental Studies

Analyze the classification recognition results of the CNN model and the decision fusion classification results of the four neural networks and ensemble classifiers respectively, validating the effectiveness of the proposed deep learning and machine learning decision fusion algorithms for air compressor noise fault detection.

A. Data Sources

This article utilizes a dataset created from acoustic recordings sampled at 16 kHz in a single-stage reciprocating air compressor. The specifications of the air compressor are as follows:

- Air pressure range: 0-500 lb/m^2, 0-35kg/cm^2.
- Induction Motor: 5HP, 415V, 5Am, 50 Hz, 1440rpm.
- Pressure Switch: Model PR-15, Range 100-213 PSI.

Dataset 1 consists of 1800 noise data samples, including normal state and 7 types of fault states. The seven fault states are: Inlet Valve Leakage (LIV) fault, Outlet Valve Leakage (LOV) fault, Non-Return Valve (NRV) fault, Piston Ring fault, Flywheel fault, Rider Belt fault, and Bearing fault. Dataset 2 consists of 600 noise data samples, including normal and fault states. Table 1 presents the fundamental information of air compressor noise dataset 1 for this experiment, while Table 2 displays the basic information of air compressor noise dataset 2 for this experiment. Add white noise with a random signal-to-noise ratio (SNR) ranging from -20 to 20 to create datasets with varying SNR conditions. For dataset 1, divide it into a training set, validation set, and test set in a ratio of 0.7: 0.2 : 0.1, while for dataset 2, divide it into a training set, validation set, and test set in a ratio of 0.6: 0.2: 0.2. Apply data augmentation to both dataset 1 and dataset 2 by flipping along the X and Y axes, shifting between -30 and 30, and randomly scaling between 0.8 and 1.2.

TABLE I
AIR COMPRESSOR NOISE DATASET 1 BASIC INFORMATION

operational state	serial number	quantities
Leakage at inlet valve fault	LIV	225
Leakage at outlet valve fault	LOV	225
Non-return valve fault	NRV	225
Pistonring fault	Piston	225
Flywheel fault	Flywheel	225
Riderbelt fault	Rider	225
Bearing fault	Bearing	225
Healthy	Healthy	225

TABLE II
AIR COMPRESSOR NOISE DATASET 2 BASIC INFORMATION

operational state	serial number	quantities
normal	normal	300
abnormal	abnormal	300

B. Data Processing

The waveforms of eight types of faults and normal states in Dataset 1 are shown in Figure 5, and the waveforms of faults and normal states in Dataset 2 are shown in Figure 6. The analysis results of wavelet scattering transformation are obtained by averaging the scattering coefficients of 10-second audio segments. Parallel feature extraction is performed using the tall array and cellfun. Wavelet feature vectors are extracted, and the results of MFCC analysis are presented in the form of MFCC spectrograms.

As shown in Figure 7, through comparative analysis of the Mel spectrogram between faults and normal states, differences in the distribution of sample data can be observed. In the Mel spectrogram, each audio segment can be represented by a corresponding Mel-spectrogram, with each frame corresponding to a certain frequency band as a feature. The color of each point represents the magnitude of the inverse frequency coefficient. Under normal circumstances, the colors in the high-frequency region are darker. When the air compressor is in a fault state, the colors in the low-frequency region are darker, which differs significantly from the normal state. Therefore, using the Mel spectrogram obtained from MFCC analysis of air compressor vibration signals can more accurately detect faults.

C. Comparative Analysis

The performance of the fault diagnosis model for air compressors based on feature fusion is validated through three aspects: comparison between single and multiple features, selection of feature fusion neural networks, and choice of feature fusion methods. The models requiring feature fusion include DCNN, GoogleNet, SqueezeNet, ResNet-50, and ensemble classifiers. These five models are evaluated based on accuracy, recall, and F1 score, with each model trained for 30 epochs.

1) Comparative analysis of single feature and feature fusion recognition results

By comparing six different network models including DCNN, GoogleNet, SqueezeNet, ResNet-50, ensemble classifier, and the fusion model, we trained DCNN, GoogleNet, SqueezeNet, and ResNet-50 based on MFCC features, while the ensemble classifier was trained based on wavelet features. These two sets of features were fused at the decision level to obtain the final fusion model. Table 3 presents the training statistics of the models. Through comparison, we can observe that on both datasets, the training performance of the fusion model is superior to that of individual models.

As shown in Figure 9a, b, the multi-class confusion matrices for the feature fusion models on Dataset 1 and Dataset 2 are presented. The multi-class confusion matrices provide a detailed display of the classification results of this method, with the horizontal and vertical axes representing the predicted labels by the model and the true data labels, respectively. It is evident that the proposed method effectively learns the fault feature information, as different fault states exhibit a distinct separation without overlapping. There is no overlap between categories, demonstrating the outstanding defect diagnostic capability of the proposed method.

2) Comparison of neural network models

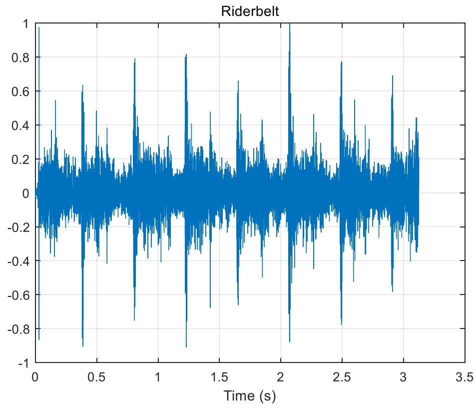
After assessing the performance of various feature fusion models and considering factors such as MFCC feature extraction, training time, and the accuracy of fusion results, a decision fusion training was conducted using different deep learning neural networks and ensemble classifiers. The goal was to select the optimal feature fusion model by combining the classification results of these models.

The experimental results, as shown in Table 4, involved comparing four selected neural networks to determine the most suitable one for fusion with ensemble classifiers. The accuracy of these four models is depicted in Figure 8a, b, c, d. By combining the information from Tables 3 and 4, it was observed that ResNet-50 performed the best on Dataset 1 but had the longest training time. Therefore, while ensuring accuracy, priority should be given to training time. On Dataset 2, DCNN exhibited the best performance with the shortest training time. This is because complex network models may overfit small sample datasets, whereas DCNN, with its shallower depth, performed well in recognizing small sample data. Thus, after comparing DCNN and ResNet-50, DCNN was chosen as the preferred model. Referring to Table 3, the precision of DCNN was 0.83 and 0.99 respectively, while ResNet-50's precision was 0.88 and 0.55 respectively. Considering all factors, DCNN showed strong applicability, shorter training time, and better precision. Therefore, DCNN was selected for late fusion with ensemble classifiers.

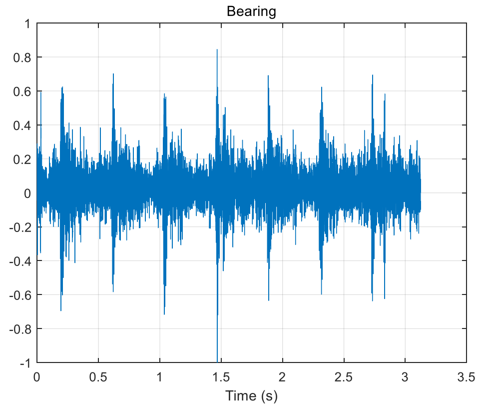
As illustrated in Figures 9a, b, the multi-class confusion matrices for Dataset 1 and Dataset 2 based on the DCNN and ensemble classifier feature fusion model are provided. These matrices detailedly display the classification results of the method, the horizontal axis represents predicted labels, and the vertical axis represents true data labels. Clearly, the proposed method effectively learned fault feature information, as different fault states exhibited distinct separation without overlap. The absence of overlap between categories demonstrates the excellent fault diagnostic capability of this method.

3) Comparing with other feature fusion methods

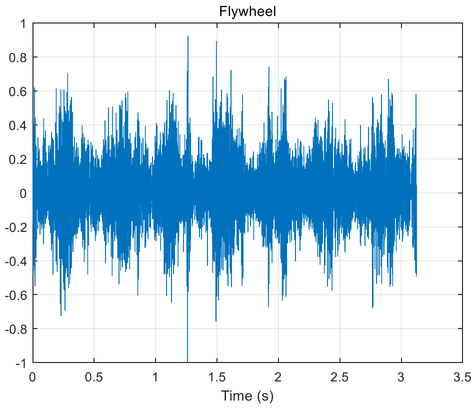
Based on the comparison between model-based feature fusion and late feature fusion, we explored their differences in handling complex tasks. Model fusion integrates the outputs of multiple models, while late fusion applies additional fusion strategies after model output. The comparative results, as shown in the Figure 10, indicate that late fusion performs better in detecting faults in air compressors.



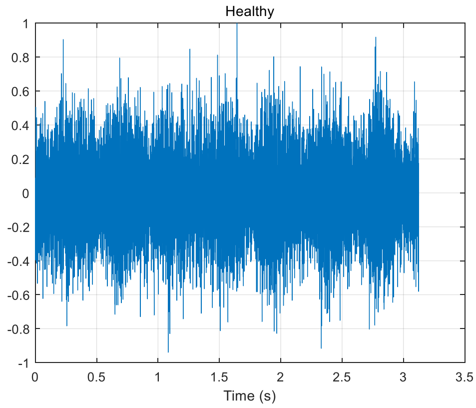
(a) Riderbelt



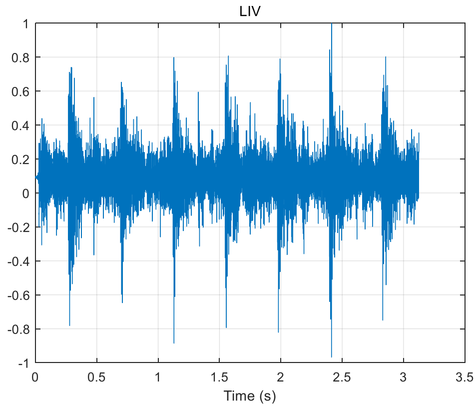
(b) Bearing



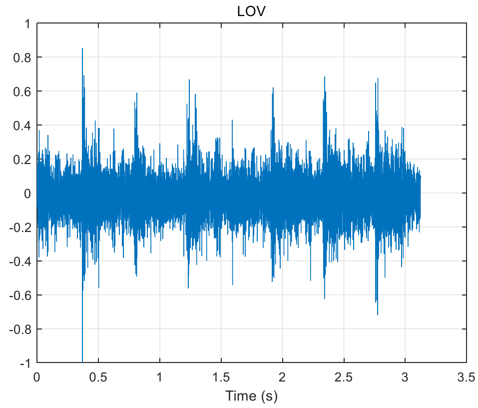
(c) Flywheel



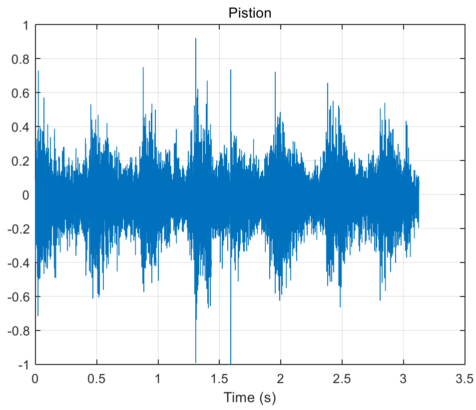
(d) Healthy



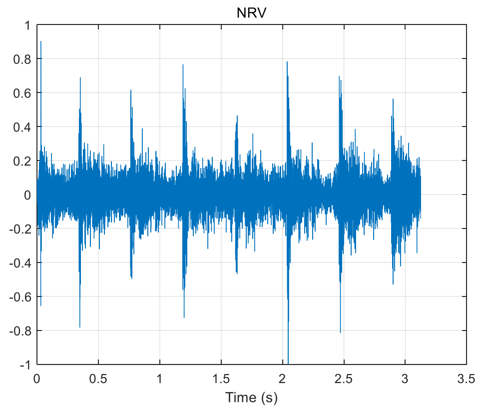
(e) LIV



(f) LOV

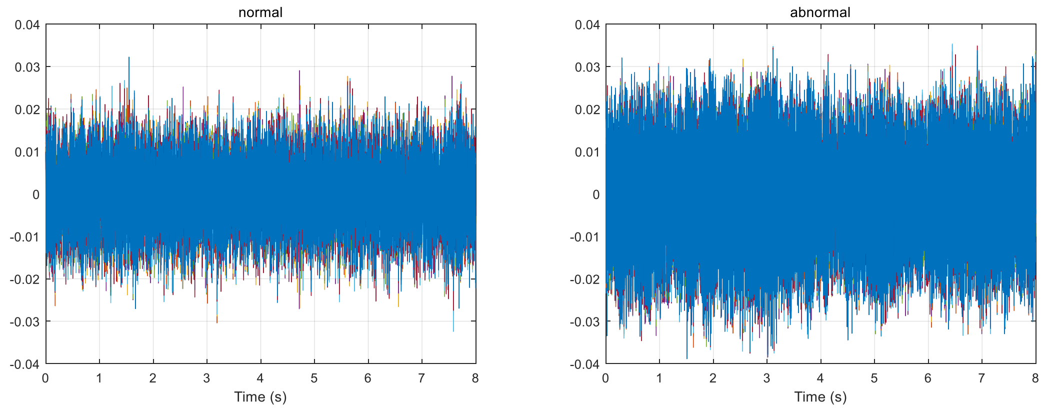


(g) Pistion



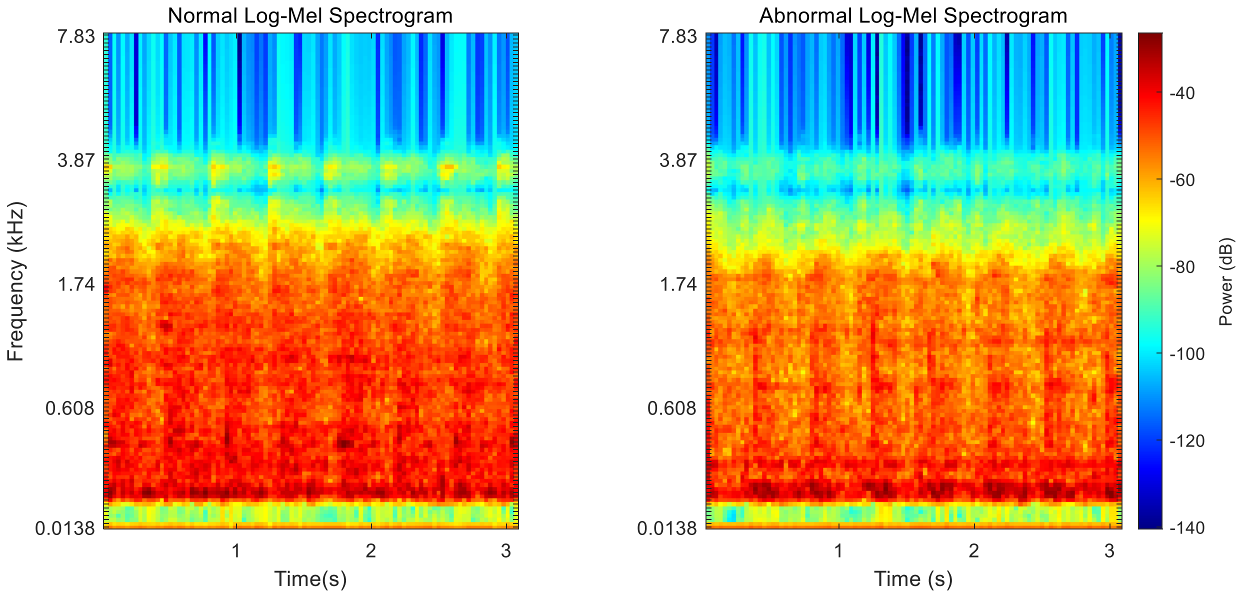
(h) NRV

Fig. 5. The waveforms of dataset 1



(a) normal (b) abnormal

Fig. 6. The waveforms of dataset 2



(a) Normal MFCC Characterization Spectra (b) Abnormal MFCC Characterization Spectra

Fig. 7. Comparative analysis of Mel spectrograms

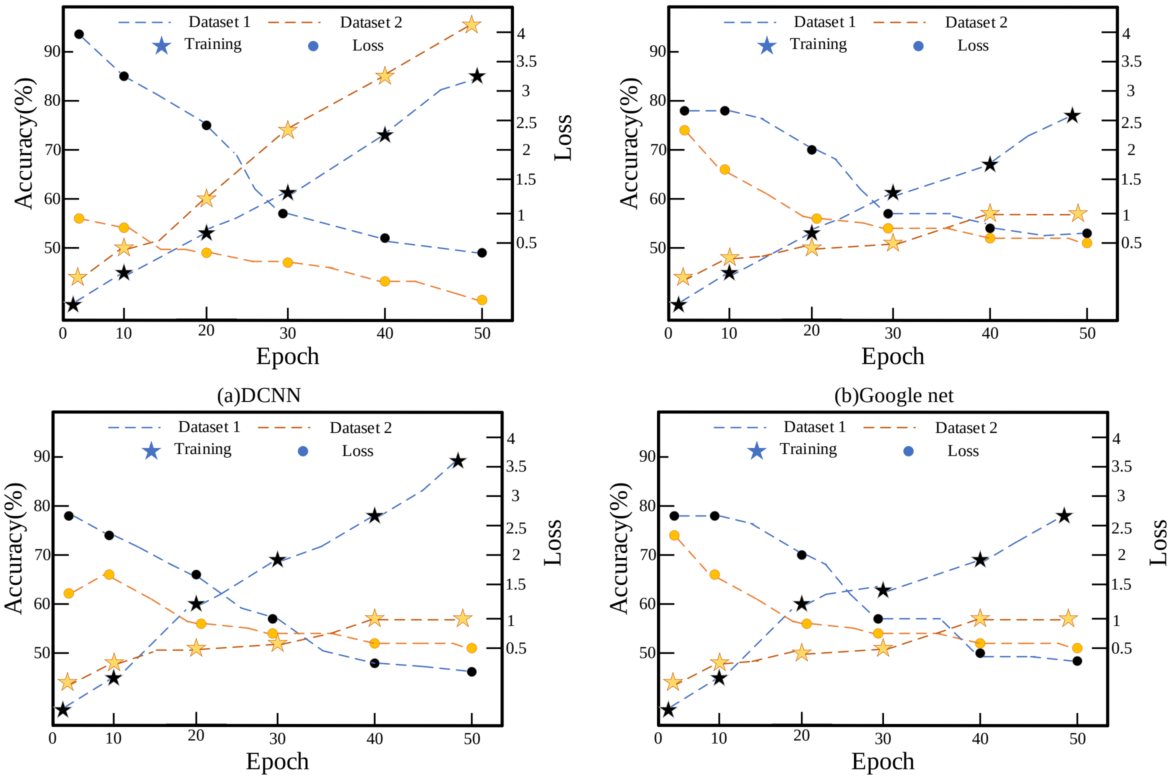
TABLE III
COMPARISON BETWEEN SINGLE-FEATURE NETWORK MODEL AND FUSION MODEL

Neural network model	Data 1 Accuracy	Recall	F1-score	Data 2 Accuracy	Recall	F1-score
DCNN	0.83	0.80	0.81	0.99	0.99	0.99
Squeeze net	0.74	0.72	0.72	0.52	0.48	0.50
Google net	0.73	0.70	0.69	0.52	0.51	0.51
Resnet-50	0.88	0.81	0.81	0.55	0.50	0.50
Integrated Classifier	0.99	0.98	0.99	0.98	0.97	0.98
Fusion	1.00	1.00	1.00	0.99	1.00	0.99

TABLE IV
COMPARISON OF THE NETWORK MODELS

Neural network model	Data 1 Accuracy	Time (min)	Data 2 Accuracy	Time (min)
DCNN-Integrated Classifier	1.00	15	0.99	13
Squeeze net-Integrated Classifier	1.00	18	0.99	18
Google net-Integrated Classifier	1.00	19	0.99	19
Resnet-50-Integrated Classifie	1.00	22	0.99	20

Late fusion allows for the application of various fusion strategies after model output, thus offering greater flexibility. This means that the most suitable fusion method can be chosen based on task requirements and data characteristics. Late fusion does not require training multiple models simultaneously but performs fusion processing independently after each model is trained. This reduces the overall training cost. Both deep learning and machine learning models can utilize late fusion techniques, thereby demonstrating strong compatibility. This enhances its general applicability in practical scenarios.



(a)DCNN (b)Google net

(c)Resnet-50 (d)Squeeze net

Fig. 8. Model Accuracy Chart

TABLE V
COMPARISON BETWEEN THE TRANSFER LEARNING-BASED NETWORK AND THE FUSION MODEL

Neural network model	Data 1 Accuracy	Recall	F1-score	Data 2 Accuracy	Recall	F1-score
DCNN	0.85	0.82	0.83	0.99	0.99	0.99
Squeeze net	0.78	0.75	0.74	0.53	0.50	0.51
Google net	0.80	0.80	0.79	0.55	0.52	0.51
Resnet-50	0.90	0.90	0.89	0.57	0.55	0.55
Fusion	1.00	1.00	1.00	0.99	1.00	0.99

TABLE VI
COMPARISON BETWEEN THE NETWORK BASED ON BAYESIAN OPTIMIZATION AND THE FUSION MODEL

Hyperparameters	Range
Initialized learning rate	[1e-2,1]
Stochastic gradient descent momentum	[0.8,0.98]
L2 regularization strength	[1e-10,1e-2]
Maximum number of rounds	[20,40]
Validation frequency	[5,40]
Minimum lot size	[20,50]

D. *Parametric Analysis*

To address the complexity of deep learning neural network structures and further enhance model performance, it is essential to conduct parameter analysis following comparative analysis. Parameter settings are crucial as they directly impact the final experimental results and model performance. To further improve the overall classification accuracy of neural networks, we suggest employing two optimization methods: transfer learning and Bayesian optimization. Additionally, we also consider the influence of signal-to-noise ratio on the model.

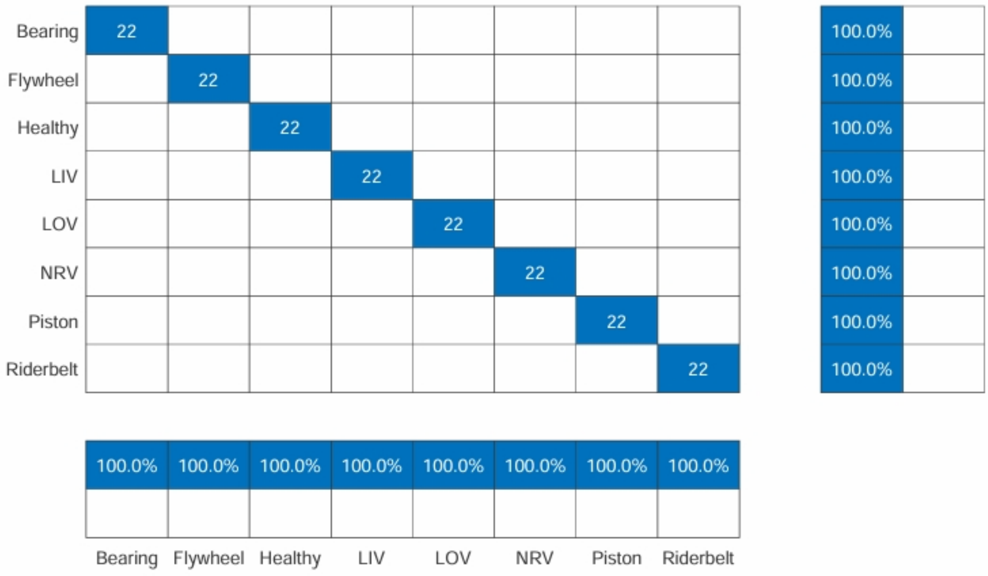
1) Comparison of Different Signal-to-Noise Ratios

By conducting experiments under different signal-to-noise ratio (SNR) conditions, we can observe the robustness and performance variations of the system under different levels of noise. A higher SNR typically indicates less noise interference, while a lower SNR may lead to confusion between signal and noise, thereby affecting the system's performance.

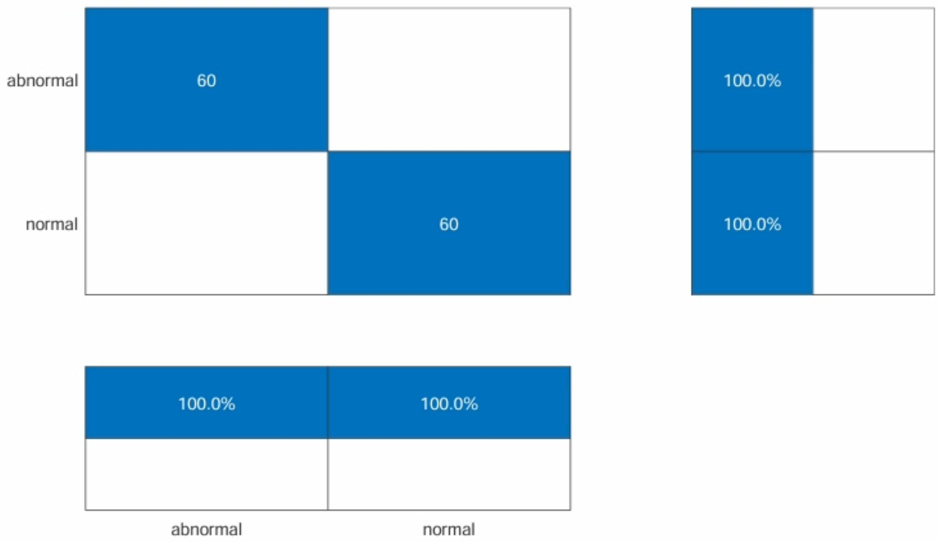
In our experiments, we simulated different SNR conditions by adjusting the relative strengths of the signal and noise. We collected data under each SNR condition and evaluated and compared the system's performance. By comparing the experimental results under different SNR conditions, as shown in Figure 11, we can observe that the feature fusion model exhibits better noise robustness in different environments.

2) Transfer learning

A comparative test was conducted on the fault diagnosis performance of neural networks using transfer learning. The effectiveness of the method was analyzed by replacing the final layer and freezing the initial layers for each network. As shown in Table 5, it can be observed that the accuracy of different neural networks improved in both sets of data. In dataset 1, the recognition accuracy of the ResNet-50 network model showed a significant improvement, attributed to its efficient resolution of the vanishing gradient problem through residual structures. However, the improvement in model performance was less pronounced in dataset 2. Nonetheless, the fused model exhibited the highest accuracy, recall rate, and F1 score.



(a) Data 1 Late Fusion Confusion Matrix Map



(b) Data 2 Late Fusion Confusion Matrix Map

Fig. 9. Late Fusion Confusion Matrix Map

TABLE VII
COMPARISON OF MODELS AFTER BAYESIAN OPTIMIZATION

Neural network model	Data 1 Accuracy	Recall	F1-score	Data 2 Accuracy	Recall	F1-score
DCNN	0.86	0.83	0.83	0.99	0.99	0.99
Squeeze net	0.79	0.77	0.78	0.54	0.53	0.52
Google net	0.82	0.80	0.80	0.55	0.52	0.51
Resnet-50	0.93	0.92	0.90	0.58	0.56	0.55
Fusion	1.00	1.00	1.00	0.99	1.00	0.99

3) Bayesian optimization

Based on transfer learning, to achieve higher model accuracy, this paper opts to use Bayesian optimization. In this optimization process, we use training data and validation data as inputs to optimize specific variables, as shown in Table 6. We create an objective function for the Bayesian optimizer, which train the CNN and returns the classification error on the validation set. Due to the setup based on Bayesian optimization, overfitting may occur, necessitating additional independent testing. Through maximizing the reduction of classification error on the validation set during Bayesian optimization, we found that all networks in dataset 1 showed some degree of improvement in multiple tests, as shown in Table 7, there is an overall improvement in accuracy, and the fusion model continues to perform the best.

E. Experimental Results

Evaluate the effectiveness of the method proposed in this paper, we fused various neural networks with ensemble classifiers and optimized different model parameters through a series of experiments to select the optimal fusion model. Based on the performance of different models, we ultimately decided to adopt decision fusion between DCNN and the ensemble classifier. The resulting model not only demonstrates good accuracy in fault detection but also requires less training time and exhibits better adaptability to different datasets.

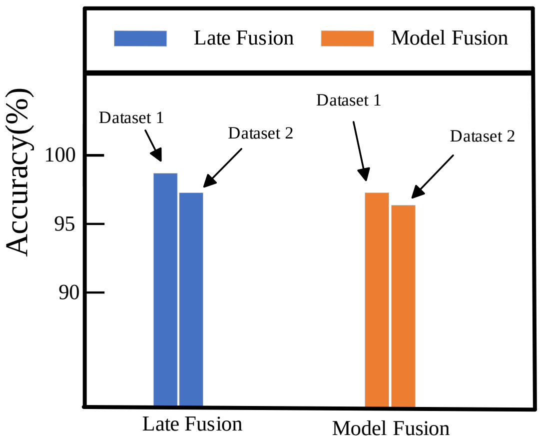
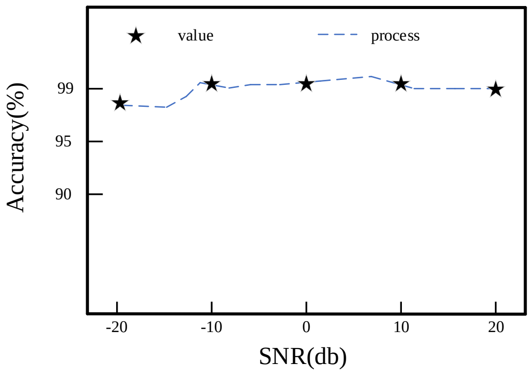
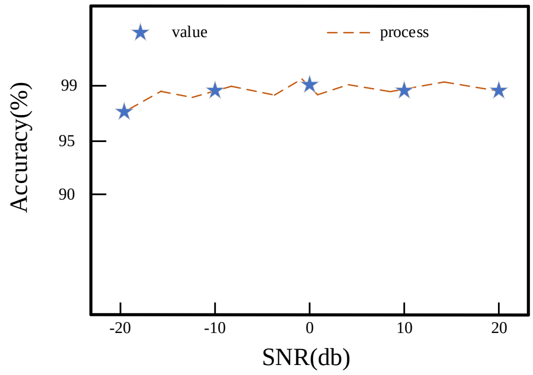


Fig. 10. Comparison of different feature fusion methods



(a) Dataset 1



(b) Dataset 2

Fig. 11. Comparison of Feature Fusion Models at Different Signal-to-Noise Ratios

V. CONCLUSIONS

Through the use of transfer learning, Bayesian optimization, and late fusion methods, it has been demonstrated through comparative analysis that feature fusion is more effective than transfer learning and Bayesian optimization. This paper presents a fault diagnosis model for air compressors based on feature fusion, which extracts wavelet scattering features for input integration with a classifier and MFCC features for input into a convolutional neural network for training and testing. After training, the classification results are fused to obtain the final fault diagnosis result.

The following conclusions can be drawn: The wavelet scattering transform, after extracting low-frequency features, utilizes wavelet transform modulation to recover lost high-frequency signals and extract more complex features, thereby minimizing intra-class differences to the greatest extent. Meanwhile, MFCC features preserve the discriminability between different classes. The late fusion method consistently achieves good results in fault recognition accuracy after multiple tests. In comparison to other deep learning networks that rely on single-feature extraction, the late fusion-based approach proposed in this paper significantly enhances recognition accuracy and consistently outperforms other literature-based detection methods in fault recognition accuracy. The air compressor fault diagnosis method proposed in this paper can enhance fault identification accuracy, decrease training time for deep learning networks, and provide valuable insights for production. The late fusion-based air compressor fault diagnosis model still has room for improvement in feature extraction. It can extract additional sound features to further enhance the accuracy of air compressor fault recognition.

REFERENCES

[1] U. Hisashi, K. Yuma, S. Shoichiro, N. Akira, and H. Noboru, “Anomaly detection technique in sound to detect faulty equipment,” NTT Technical Review, vol. 15, no.8, pp28–34, 2017.

[2] P. Shi, H. Gao, Y. Yu, X. Xu, and D. Han, “Intelligent fault diagnosis of rolling mills based on dual attention-guided deep learning method under imbalanced data conditions,” Measurement, vol. 204, p111993, 2022.

[3] Xiaoli Duan, Sanwei Liu, Fuyong Huang, Daoyuan Zhang, Yan Zhao, Jianjia Duan, Zeyu Zeng, Ting Yu, Lipeng Zhong, and Bin Dai, "Cleaning of Transient Fault Data in Distribution Network Based on

Clustering by Fast Search and Find of Density Peaks," Engineering Letters, vol. 31, no.4, pp1348-1358, 2023.
[4] Guang Yang, and Xinrong Liu, "Study on Method of Bearing Fault Detection Based on Vibration Signal Analysis," Engineering Letters, vol. 31, no.3, pp1009-1015, 2023.
[5] J. S. Abbasi, F. Bashir, K. N. Qureshi, M. N. ul Islam, and G. Jeon, “Deep learning-based feature extraction and optimizing pattern match ing for intrusion detection using finite state machine,” Computers & Electrical Engineering, vol. 92, p107094, 2021.
[6] S. K. Pandey and R. R. Janghel, “Classification of electrocardiogram signal using an ensemble of deep learning models,” Data Technologies and Applications, vol. 55, no.3, pp446–460, 2021.
[7] Cha, Young-**, Wooram Choi, and Oral Büyüköztürk. "Deep learning-based crack damage detection using convolutional neural networks." Computer-Aided Civil and Infrastructure Engineering, vol. 32, no.5, pp361– 378, 2017.
[8] H. Wei, Q. Zhang, M. Shang, and Y. Gu, “Extreme learning machine based classifier for fault diagnosis of rotating machinery using a residual network and continuous wavelet transform,” Measurement, vol. 183, p109864, 2021.
[9] P. Liang, C. Deng, J. Wu, Z. Yang, J. Zhu, and Z. Zhang, “Compound fault diagnosis of gearboxes via multi-label convolutional neural network and wavelet transform,” Computers in Industry, vol. 113, p. p103132, 2019.
[10] J. Burriel-Valencia, R. Puche-Panadero, J. Martinez-Roman, A. Sapena-Bano, M. Riera-Guasp, and M. Pineda-Sanchez, “Multi band frequency window for time-frequency fault diagnosis of induction machines,” Energies, vol. 12, no.17, p3361, 2019.
[11] L. Wen, X. Li, and L. Gao, “A transfer convolutional neural network for fault diagnosis based on resnet-50,” Neural Computing and Appli cations, vol. 32, no.10, pp6111–6124, 2020.
[12] A. E. Prosvirin, Z. Ahmad, and J.-M. Kim, “Global and local feature extraction using a convolutional autoencoder and neural networks for diagnosing centrifugal pump mechanical faults,” IEEE Access, vol. 9, pp65838–65854, 2021.
[13] Kuntoğlu, Mustafa, and Hacı Sağlam, “Investigation of signal behaviors for sensor fusion with tool condition monitoring system in turning,” Measurement, vol. 173, p108582, 2021.
[14] Yu Wang, Changfeng Zhu, Qingrong Wang, and Jinhao Fang, "Research on Fault Detection of Rolling Bearing Based on CWT-DCCNN-LSTM," Engineering Letters, vol. 31, no.3, pp987-1000, 2023.
[15] H. Soltanali, A. Rohani, M. Tabasizadeh, M. H. Abbaspour-Fard, and A. Parida, “An improved fuzzy inference system-based risk analysis approach with application to automotive production line,” Neural Computing and Applications, vol. 32, pp10573–10591, 2020.
[16] H. Soltanali, M. Khojastehpour, J. E. D. A. E. Pais, and J. T. Farinha, “Sustainable food production: an intelligent fault diagnosis framework for analyzing the risk of critical processes,” Sustainability, vol. 14, no.3, p1083, 2022.
[17] L. Ma, G. Wang, P. Zhang, and Y. Huo, “Fault diagnosis method of circuit breaker based on ceemdan and pso-gsa-svm,” IEEJ Transactions on Electrical and Electronic Engineering, vol. 17, no.11, pp1598 1605, 2022.
[18] H. Shao, H. Jiang, H. Zhao, and F. Wang, “A novel deep autoen coder feature learning method for rotating machinery fault diagnosis,” Mechanical Systems and Signal Processing, vol. 95, pp187–204, 2017.
[19] A. K. Ozcanli and M. Baysal, “Islanding detection in microgrid using deep learning based on 1d cnn and cnn-lstm networks,” Sustainable Energy, Grids and Networks, vol. 32, p100839, 2022.
[20] S. A. Kawa and M. A. Wani, “Data driven designing of convolutional neural networks architectures for image classification.” Engineering Letters, vol. 31, no.2, pp735–742, 2023.
[21] L. Wang, D. Ping, C. Wang, S. Jiang, J. Shen, and J. Zhang, “Fault diagnosis of rotating machinery bearings based on improved dcnn and woa-delm,” Processes, vol. 11, no.7, p1928, 2023.
[22] G. Li, J. Ao, J. Hu, D. Hu, Y. Liu, and Z. Huang, “Dual-source gramian angular field method and its application on fault diagnosis of drilling pump fluid end,” Expert Systems with Applications, vol. 237, p121521, 2024.
[23] P. Zhang, W. Jiang, Y. Zheng, S. Zhang, S. Zhang, and S. Liu, “Hydraulic-pump fault-diagnosis method based on mean spectrogram bar graph of voiceprint and resnet-50 model transfer,” Journal of Marine Science and Engineering, vol. 11, no.9, pp1678, 2023.
[24] C. Tian, D. Hao, M. Ma, J. Zhuang, Y. Mu, Z. Zhang, X. Zhao, Y. Lu, X. Zuo, and W. Li, “Graded diagnosis of helicobacter pylori infection using hyperspectral images of gastric juice,” Journal of Biophotonics, vol. 17, no.1, pe202300254, 2024.
[25] A. K. Sharma, A. Nandal, A. Dhaka, L. Zhou, A. Alhudhaif, F. Alenezi, and K. Polat, “Brain tumor classification using the mod ified resnet50 model based on transfer learning,” Biomedical Signal Processing and Control, vol. 86, p105299, 2023.
[26] H. Liang and X. Zhao, “Rolling bearing fault diagnosis based on one dimensional dilated convolution network with residual connection,” Ieee Access, vol. 9, pp31078–31091, 2021.
[27] V. Tiwari, “Mfcc and its applications in speaker recognition,” Interna tional journal on emerging technologies, vol. 1, no.1, pp19–22, 2010.
[28] L. Bai, J. Liang, and Y. Guo, “An ensemble clusterer of multiple fuzzy k-means clusterings to recognize arbitrarily shaped clusters,” IEEE Transactions on Fuzzy Systems, vol. 26, no.6, pp3524–3533, 2018.
[29] Yongqing Zeng, Haibo Li, Xiang Xia, Shouchun Deng, Hong Zuo, Haozhen Yue, and Hua Luo, "Research on Time-frequency Characteristics for Blasting Vibration Signal of CO2 Blasting by Frequency Slice Wavelet Transform," Engineering Letters, vol. 28, no.4, pp1047-1057, 2020.
[30] S. Liao, J. Chen, J. Hou, Q. Xiong, and J. Wen, “Deep convolutional neural networks with random subspace learning for short-term traffic f low prediction with incomplete data,” 2018 International Joint Con ference on Neural Networks (IJCNN)), pp1–6, 2018.
[31] Z. Yu, Y. Lu, J. Zhang, J. You, H.-S. Wong, Y. Wang, and G. Han, “Progressive semisupervised learning of multiple classifiers,” IEEE Transactions on Cybernetics, vol. 48, no.2, pp689–702, 2017.
[32] J. J. Rodriguez, L. I. Kuncheva, and C. J. Alonso, “Rotation forest: A new classifier ensemble method,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no.10, pp1619–1630, 2006.
[33] R. Polikar, J. DePasquale, H. S. Mohammed, G. Brown, and L. I. Kuncheva, “Learn++. mf: A random subspace approach for the missing feature problem,” Pattern Recognition, vol. 43, no.11, pp3817 3832, 2010.
[34] E. H¨ ullermeier and S. Vanderlooy, “Combining predictions in pairwise classification: An optimal adaptive voting strategy and its relation to weighted voting,” Pattern Recognition, vol. 43, no.1, pp128–142, 2010