Research on Engine Failure Diagnosis Algorithm Based on Sound Signal

Ziyang Hu
Army Academy of Artillery & Air Defense, Hefei 230031, China
hzyyqgmz1991@126.com

Abstract. Effective diagnosis of engine failure is an important factor affecting vehicle safety. Based on infrared, ultrasonic, thermal, vibration and other methods, there are drawbacks such as the need to modify the engine, high deployment costs, etc. In this thesis, a failure diagnosis model of RNN denoising and BP neural network identification based on fbank feature of sound signal is proposed. It is found through experiments that it can effectively diagnose engine failures.

1. Introduction
Failure of the engine will seriously affect the normal operation of the vehicle. To effectively control and reduce the impact of engine failure, it is necessary to strengthen the study of engine failure diagnosis: Before the occurrence of the failure, the engine status is monitored, the failure signs and hidden dangers are found in time; after the engine failure occurs, the failure cause is analyzed in time, and corresponding solutions are formulated to prevent failures from recurring or even accidents.

In recent years, many researchers have devoted themselves to the research of engine failure diagnosis, and have proposed failure diagnosis methods such as infrared, ultrasonic, thermal, vibration and other methods. However, the above method has obvious drawbacks, that is, the engine needs to be modified, and the deployment cost, time cost and decision cost are relatively high. The solution based on acoustic diagnostic technology provides a non-contact technical path for engine failure diagnosis. By the sound signal, this thesis proposes a failure diagnosis model based on fbank feature for RNN noise reduction and BP neural network recognition.

2. Failure diagnosis model
2.1. Fbank feature extraction
In the short-term analysis, the sound signal is divided into one frame and one frame (the length is generally 10 to 30 ms, and in order to make a smooth transition between the frames, the adjacent frames generally overlap 50% or more of the frame length) to analyze [1], so that each frame feature parameter together constitutes a feature sequence of the sound signal. Sound signal analysis mainly includes time domain analysis, frequency domain analysis and cepstrum domain analysis [2]; frequency domain analysis is more important than time domain analysis. This thesis will mainly adopt the Mel filter bank features. (Fbank).
2.2. RNN-based noise reduction model

The essential feature of RNN is that there are both internal feedback connections and feedforward connections between its processing units. Its internal feedback connection can preserve the state of the hidden layer nodes and provide memory for the network. The RNN will remember the previous information and use the previous information to influence the subsequent output. [2] Because the RNN above will have gradient disappearance or gradient explosion when dealing with long-term dependency problems [3], it is not used much in practical applications. This problem has also led to the emergence of many variant RNNs, the most commonly used are Long Short-Term Memory (LSTM) and Gated Recurrent Unit–RNN (GRU). The GRU is mainly introduced here.

The Gate Mechanism is added to better capture the dependence of the time step distance in the time series. It controls the flow of information through the gate. The gate mechanism controls input, memory, and makes predictions at the current time step. [4] Gated Recurrent Unit has two gates, a reset gate and an update gate, and its structure is shown in Figure 1.

![Figure 1. GRU gate control loop mechanism](image)

The traditional deep learning-based noise reduction process usually uses the neural network model to directly process the FFT spectrum, and outputs the noise-reduced spectrum. However, due to the high dimension of the FFT spectrum, it leads to the noise reduction model requires many parameters and a large amount of calculation. Therefore, this thesis simplifies the traditional deep learning-based noise reduction process to reduce the number of parameters and the amount of calculation required of the noise reduction model. The dimension of the fbank feature is much lower than the dimension of the FFT spectrum. Therefore, the neural network noise reduction model directly processes the fbank feature to reduce parameters and the amount of calculation.

2.3. Recognition Model Based on BP Neural Network

BP neural network (BPNN) has powerful nonlinear mapping, self-learning and self-adaptation, generalization and fault tolerance capabilities. It mainly consists of input layer, hidden layer and output layer. The BPNN algorithm structure is shown in Figure 2, where X(1), X(2),...,X(n) are the input values of the network, Y(1), Y(2),...,Y(n) is the predicted value of the network. The input value of the network propagates from the input layer to the output layer through the hidden layer, and the neurons in the output layer obtain the response of the network to the input. Then, according to the direction of reducing the error between the target output and the actual output, the connection weights are corrected layer by layer from the output layer through each middle layer and finally returned to the input layer. [5] Within a certain range, as the error back propagation progresses, the correct rate of feedback in the network test also increases. Through the two processes of forward propagation and back propagation, the advancement of the algorithm and the error correction are realized, and the output value is approximated to the ideal value. [6]
3. Experimental analysis

3.1. Experimental data
The sound data used in this experiment is the engine sound signal of the QE1118GA model. Considering the influence of the engine idle speed and the acceleration speed on its sound signal, the sound signals of idle speed and acceleration to 2000rpm are collected. For each type of speed, the positions of 1m in front left, 1m in front, and 1m in front of the right were selected; the length of each segment was 30s; the sampling rate was 48 KHz. The experiment records the failure sample by manually setting the engine failure on the normal vehicle. The specific method is to loosen the piston ring of two cylinders of the engine, so that the engine is running lack of cylinders, and the sound signal is collected in the same way. The engine sound data of each speed, position and state are randomly assigned to the training set, the verification set and the test set at the ratio of 3:1:1. The noise data is the field noise of the yard which is recorded several times during the acquisition process, and the acquisition time length and sampling rate of each segment are consistent with the foregoing.

The engine sound data and the noise data are mixed with a random signal-to-noise ratio (SNR) to form new sound signal, that is, each training set and verification set data will be mixed with random SNR when the noise reduction model is trained.

3.2. Noise reduction experiment
The noise reduction model of this experiment takes the fbank feature of the engine sound mixing noise as the input. The input data is processed through several layers of GRUs, then processed through a fully connected layer to obtain a template, which is then multiplied by the input to output the denoised fbank feature. The corresponding engine sound signal fbank feature is used as the learning target, and training is performed using the Minimum Mean Squared Error (MMSE) criterion to minimize the mean square error (MSE) between the noise-reduced fbank feature and the fbank feature of the original engine sound signal. The same indicator was used during the test.

In order to find the best results, this paper experiments multiple sets of model parameters: the number of GRU layers is 2, 3, the number of GRU output neurons is 64, 80, 96, 128, 160, and 192, respectively. The number of neurons in the TDense layer is 40 which is consistent with the input fbank feature dimension. Using the MMSE criterion as the optimization criterion, using the Nadam algorithm as the optimization algorithm, the learning rate is 0.002 and the number of batches is 64. The entire training set is cycled 30 times. In addition, the loss on the verification set is monitored, and if the verification set loss does not decrease for three consecutive times, the learning rate is multiplied by 0.2. Take the parameters when the loss of the verification set reaches the minimum as the final parameter of the noise reduction model. Finally, comparing the MSE of the test set in the noise reduction models with each set of parameters, it was found that the lowest MSE was achieved using the noise reduction model with 2-layer GRU and the output neurons number of 128.
3.3. Failure identification experiment

The sound data in front of the engine at idle speed in the training set is used to train. 5 segment sound data of normal status and 5 segment sound data of failure status are selected. The fbank feature of noise-reduced sound signal is taken as input. The guide output value is 100 for failure samples; the guide output value is 0 for normal samples. After training the BPNN with this guide value, the closer the output value is to 100, the greater the possibility that the input data comes from the sound data of failure and the closer to 0, the opposite. The number of repetitions in this training is 2000; the convergence error is 1e-7; the learning rate is 0.01. Execute the script, and the training situation and the output result of BPNN are shown in Figure 3.

It can be seen from Figure 3 that the experimental results did not reach the expected value. When Epoch<50, the MSE value has partially converged and cannot reach the set value of 1e-7. There are three possible reasons for the experimental results did not reach the expected value: First, the learning rate setting of BPNN is not suitable, which leads to partial convergence in the training process. Second, the number of BPNN middle layers is too small to complete the abstraction of implicit commonality in different data. Third, the correlation between the sound data in front of the engine at idle speed and the failure is low, that is, the failure does not cause a significant change in the sound data in front of the engine at idle speed.

For the first reason, the learning rate is changed and the training test is carried out. The learning rate was changed to 0.1, 0.2, 0.3, 0.02, 0.05, 0.001, 0.002, 0.005, and the experiment was repeated several times. In all cases, the BPNN training showed a partial convergence state. As shown in Figure 4. The first reason can be ruled out.
Next, measures were taken for the second reason, the number of middle layers is changed to 2. The training situation and test results are as shown in Figure 5. In addition, the number of middle layers has also been tried to be modified to 3,4,5,6. However, the test results are still not satisfactory. The second reason can be ruled out.

For the third reason, all kinds of sound data in the training set are analyzed in turn to test the output of BPNN, including the sound data in right front of the engine at idle speed, the sound data in left front
of the engine at idle speed, the sound data in front of the engine at acceleration speed, the sound data in right front of the engine at acceleration speed and the sound data in left front of the engine at acceleration speed. When the sound data in front of the engine at acceleration speed is tested, the output of BPNN is in line with expectations. It can be seen that in this training, the MSE value decreases steadily as the number of training increases, and the expected convergence error of 1e-7 is reached at Epoch=1009. The test result is as shown in Figure 6. The sound data in right front of the engine at acceleration speed and the sound data in left front of the engine at acceleration speed are also tried, and the BPNN output is also in line with expectations.

![Image of Neural Network](image1.jpg)

**Figure 6.** Using the sound data in front of the engine at acceleration speed

In order to verify the failure identification capability of the BPNN, the sound data of the engine at acceleration speed is selected in the test set, and it is mixed with the noise data at different signal-to-noise ratios to form new sound signal. The fbank features are extracted from 100 sets of new sound signal (the first 50 groups are failure sound data, and the last 50 groups are normal sound data) and noise reduced, and then input to the BPNN, and finally the output result is tested. As shown in Figure 7, 98 sets of output values in the 100 sets of tests meet the expected output. It is verified that the BPNN has good failure recognition capability.

![Image of Validation Result](image2.jpg)

**Figure 7.** Validation result
4. Conclusion

In order to diagnose the engine failure by sound signal, this thesis proposes a failure diagnosis model of RNN denoising and BP neural network identification based on fbank feature of sound signal. Firstly, the noise reduction on the fbank feature is proposed to reduce the number of parameters and the calculation amount required for the neural network noise reduction model. Then, the RNN-based noise reduction model is used to denoise the fbank features. Finally, the BP neural network recognition model is trained and verified by using the fbank feature after noise reduction. The results show that when the sound data of the engine at acceleration speed is used, the output of BPNN is as expected, which proves that the failure diagnosis model has the ability to use the engine sound signal for failure diagnosis. There is a small amount of missed detection in this algorithm. For this reason, it is necessary to further study the feature extraction and recognition algorithms in the later stage to reduce the missed detection rate.

References

[1] Lin Huibin. Anti-noise Performance of Discrete Spectrum Correction Theory and Their Application on Engineering [D]. South China University of Technology: South China University of Technology, 2010.

[2] Han Yan, Chen Lihua. Design and implementation of a speaker recognition system [J]. Journal of Zhejiang University(Engineering Science), 2001 (2): 165-169.

[3] Zhang Jinafu, Song Yu, Li Gang, et al. A method of fault diagnosis for rolling bearing of wind turbine based on long short-time memory neural network [J]. Computer Measurement & Control, 2017, 25 (1): 16-19.

[4] Qiao Dalei, DaiLikun, Zou Yujuan. Research and Design of Predictive Maintenance System for Unmanned Surface Vehicles Engine Cabin [J], Machinery Design & Manufacture, 2018, 6, 270-272.

[5] Long Quan, Liu Yongqian, Yang Yongping. Fault diagnosis method for wind turbine gearbox based on particle swarm optimization BP neural network [J]. Acta Energiae Solaris Sinica, 2012, 33 (1): 120-124.

[6] Yang Jie, WU Zhongru, GU Chongshi. Dam Deformation Monitoring Model and Forecast Based on BP Algorithm of Artificial Neural Networks [J]. Journal of Xi'an University of Technology, 2001, 17 (1), 25-29.