

Xueyi LI
Jialin LI
David HE
Yongzhi QU

GEAR PITTING FAULT DIAGNOSIS USING RAW ACOUSTIC EMISSION SIGNAL BASED ON DEEP LEARNING

DIAGNOSTYKA PITTINGU KÓŁ ZĘBATYCH NA PODSTAWIE SUROWEGO SYGNAŁU EMISJI AKUSTYCZNEJ W OPARCIU O GŁĘBOKIE UCZENIE MASZYNOWE

Gear pitting fault is one of the most common faults in mechanical transmission. Acoustic emission (AE) signals have been effective for gear fault detection because they are less affected by ambient noise than traditional vibration signals. To overcome the problem of low gear pitting fault recognition rate using AE signals and convolutional neural networks, this paper proposes a new method named augmented convolution sparse autoencoder (ACSAE) for gear pitting fault diagnosis using raw AE signals. First, the proposed method combines sparse autoencoder and one-dimensional convolutional neural networks for unsupervised learning and then uses the reinforcement theory to enhance the adaptability and robustness of the network. The ACSAE method can automatically extract fault features directly from the original AE signals without time and frequency domain conversion of the AE signals. AE signals collected from gear test experiments are used to validate the ACSAE method. The analysis result of the gear pitting fault test shows that the proposed method can effectively performing recognition of the gear pitting faults, and the recognition rate reaches above 98%. The comparative analysis shows that in comparison with fully-connected neural networks, convolutional neural networks, and recurrent neural networks, the ACSAE method has achieved a better diagnostic accuracy for gear fitting faults.

Keywords: gear pitting fault diagnosis, autoencoder, one-dimensional convolutional neural network, acoustic emission signal.

Pitting kół zębatych stanowi jedno z najczęstszych uszkodzeń przekładni mechanicznych. Do wykrywania takich uszkodzeń stosuje się sygnały emisji akustycznej (AE), które, ze względu na niższą wrażliwość na hałas otoczenia, stanowią skuteczniejsze narzędzie diagnostyczne niż tradycyjne sygnały wibracyjne. Wykrywalność zużycia guzłkowatego (pittingu) kół zębatych przy użyciu sygnałów AE i splotowych sieci neuronowych jest jednak niska. Aby rozwiązać ten problem, w niniejszym artykule zaproponowano nową metodę diagnozowania uszkodzeń kół zębatych za pomocą surowych sygnałów AE, którą nazwano augmented convolution sparse autoencoder (konwoluacją rozszerzoną z wykorzystaniem autoenkodera rzadkiego, ACSAE). Jest to metoda samouczenia jednowymiarowych splotowych sieci neuronowych realizowanego za pomocą autoenkodera rzadkiego. Metoda ta wykorzystuje teorię wzmocnienia do zwiększania adaptacyjności i odporności sieci. Metoda ACSAE pozwala na automatyczne wyodrębnianie cech degradacji bezpośrednio z oryginalnych sygnałów AE bez konieczności ich konwersji do domeny czasu i częstotliwości. Walidację metody przeprowadzono na podstawie sygnałów AE otrzymanych w badaniach kół zębatych. Analiza wyników badań pittingu kół zębatych wskazuje, że proponowana metoda pozwala na skuteczną detekcję tego typu uszkodzeń, przy wskaźniku wykrywalności powyżej 98%. Analiza porównawcza pokazuje, że metoda ACSAE cechuje się większą trafnością diagnostyczną w wykrywaniu błędów montażowych kół zębatych w porównaniu z sieciami neuronowymi w pełni połączonymi, splotowymi i rekurencyjnymi.

Słowa kluczowe: diagnostyka pittingu kół zębatych, autoenkoder, jednowymiarowa splotowa sieć neuronowa, sygnał emisji akustycznej.

1. Introduction

With the rapid development of the modern industries, the detection and identification of gear pitting faults in mechanical transmission systems have become one of the critical issues. Establishing a reliable health detection system, especially for gear fitting faults, is the key to ensuring smooth operation of industrial equipment.

Traditionally, vibrational signals are used as monitoring signals for gear fault diagnosis as reported in the literature. For example, Praveenkumar *et al.* [18] extracted statistical inquiries from the acquired gearbox vibration signals. The extracted features were given as input to the support vector machine for fault identification. Zuber *et al.* [35] used a complete set of vibration features as input to the self-organized feature maps and discussed the implementation of fea-

ture-based artificial neural networks and vibration analysis to achieve automatic gearbox fault identification. Of course, these methods have achieved good results. However, due to the influence of the external environment, the vibration signals contain a large amount of environmental noise. Therefore, acoustic emission (AE) technology has gradually been introduced into mechanical fault diagnosis. Crivelli *et al.* [7] proposed that AE is a potentially suitable technique for detecting early fatigue cracks because it is sensitive to high frequencies generated by crack propagation and is not affected by low-frequency noise. The method proposed by He *et al.* [13] uses a short-time Fourier transform to pre-process the AE sensor fault signal of the bearing. The verification results show an accurate classification. Qu *et al.* [20] proposed an AE signal processing method based on an improved time

synchronous average. This provides us with an inspiration to develop a gear pitting fault diagnosis using AE signals.

For gear fault detection, Sharma *et al.* [23] used the Hertzian contact approach to propose a theoretical model for establishing the relationship between the magnitude of the fault and the AE energy produced in the gear. Aouabdi *et al.* [2] proposed a tool for anomaly detection based on the multi-scale entropy algorithm, which used the phase current measured by the induction motor driving the gearbox to identify local gear tooth defects. Ratni *et al.* [21] based on a combination of maximum correlation kurtosis deconvolution and spectral kurtosis to fault diagnosis. Yao *et al.* [29] used the local mean decomposition based on the adaptive decomposition of the signal itself to perform fault diagnosis of the gearbox. Song *et al.* [25] considered the random noise and the ensemble empirical mode decomposition (EEMD). A fault diagnosis method based on singular value decomposition and improved EEMD was proposed. Feng *et al.* [10] proposed a time-frequency analysis method for the diagnosis of planetary gearboxes. And Feng *et al.* [9] also proposed a simple and effective method for diagnosing planetary gearbox faults based on the amplitude and frequency demodulation. These contributions have greatly enriched the literature on gearbox testing.

Many gear fault processing methods extract features based on prior knowledge and then use machine learning to classify the degree of pitting, such as frequency domain analysis, wavelet transform, mathematical morphology, support vector machine, etc. Yuan *et al.* [30] proposed empirical mode decomposition (EMD) of integrated overall noise reconstruction to overcome the main problems of user-defined parameters and the incompatibility of high Signal to Noise Ratio (SNR) conditions. Liu *et al.* [15] proposed an integrated method of EMD and Wigner-Ville distribution for fault feature extraction. Krishnakumari *et al.* [14] proposed that under feature extraction, statistical characteristics such as standard deviation, kurtosis, etc. are considered as characteristics of the signal. Zhang *et al.* [32] used the EEMD to decompose various fault signals to achieve diagnose results. Chen *et al.* [6] proposed a method for identifying planetary gear faults based on double-tree complex wavelet transform denoising and Laplace feature map. Widodo *et al.* [27] used a short-time Fourier transform method to monitor faults in the gearbox. Li *et al.* [16] supposed relative wavelet energy can identify the trend from the normal state to crack failure before the occurrence of broken tooth failure. Shao *et al.* [22] used principal component analysis and kernel principal component analysis to extract features from the fault features of selected data features and to reduce the effect of dimensional analysis. Bafroui *et al.* [3] used Monet wavelets to process non-stationary vibration signals. Elforjani *et al.* [8] used vector machine regression, multi-layer artificial neural network models and Gaussian process regression to correlate AE features with corresponding natural pitting. These methods play an important role in the development of gear fault diagnosis. However, with the increase in the number of equipment detection points and the sampling frequency, mechanical health monitoring has entered the era of "big data". The traditional intelligent diagnosis algorithm based on signal processing for feature extraction and classifier has high requirements for expert experience and cannot guarantee universality. It cannot meet the big data requirements of gear fault detection.

Fortunately, researchers have tried to use deep learning in the troubleshooting of gears. Zhao *et al.* [33] was inspired by the success of the deep learning method and redefines the representation learning of raw data. They proposed a local feature-based gated recurrent unit network for fault diagnosis. Zhang *et al.* [31] proposed a gear fault diagnosis method based on singular value decomposition and Radial Basis Function neural network for the problem that the weak gear fault signal is difficult to detect. Bangalore *et al.* [4] proposed an artificial neural network (ANN) based condition monitoring method using data from monitoring. Ali *et al.* [1] used ANN computational

modeling to correlate data feedforward and recursive Eman neural network algorithms from AE sensors for the development of ANN models. Sreepadha *et al.* [26] used ANN to perform prediction and classification based on heuristic models based on spur gears. Recently, in the field of mechanical gear fault diagnosis, Qu *et al.* [19] proposed a new method of unsupervised detection of gear pitting failure based on autoencoder theory. The proposed method was developed based on a deep sparse autoencoder. This method integrates the dictionary learning to an autoencoder network in the sparse coding for analysis of fault diagnosis. Cao *et al.* [5] proposed a transfer learning method using the vibration signal to use the convolutional neural network (CNN) for gear fault detection. These methods of deep learning do not require manual extraction of fault features and also achieve better fault detection results. Moreover, CNN greatly reduces the number of network parameters through local weight sharing and can avoid the over-fitting of the network when the number of samples is insufficient. However, these methods still use vibration signals to achieve gear fault detection. As mentioned above, the vibration signal contains a large amount of environmental noise due to the influence of the external environment. It is not the ideal data for deep learning fault extraction.

Based on the above considerations, this paper takes the gear as the research object and improves the method of gear pitting fault diagnosis for the standard one-dimensional CNN model, which enhances the generalization performance of the model. A set of intelligent diagnosis algorithm named augmented convolution sparse autoencoder (ACSAE) being proposed. ACSAE combines sparse autoencoder into one-dimensional CNN, and the algorithm automatically performs feature extraction and fault identification using original AE data with low environmental interference. ACSAE does not require any test set information and it does not require any noise reduction pre-processing and time domain and frequency domain conversion. First of all, the model is trained using an automatic encoder and a one-dimensional CNN. Then, through unsupervised learning, the decoded data that is very close to the original data is obtained, and the training set is added to improve the robustness and adaptability of the network. Next, the model structure and model parameters retain unchanged to ensure that the encoder part retains the characteristics of the original data. Finally, through the fusion of one-dimensional CNN automatic learning features on the softmax classifier, a number of gearbox AE data is selected for fine-tuning the network, training classification model, to achieve accurate identification of gear pitting fault. The experimental results show that the proposed method not only improves the accuracy of gear fault identification, but also improves the correct rate of gear faults from 90.5% of the one-dimensional CNN alone to 97.9% of the ACSAE. The universality of the classification algorithm is significantly improved, which can be used as an auxiliary basis for engineers and technicians to judge the degree of gear pitting fault.

The remainder of the paper is organized as follows. Section 2 describes the ACSAE method presented in this paper. This section mainly introduces the principle of autoencoder and the one-dimensional CNN algorithm and the proposed ACSAE method. In Section 3, five experiments were designed to verify whether the ACSAE algorithm can effectively classify gear pitting faults and give specific experimental parameters. Section 4 analyzes and discusses the experimental results. The current research conclusions and future research directions are summarized in Section 5.

2. The methodology

AE technology is a new type of dynamic non-destructive testing technology, which uses the internal particles of the material to release the strain energy in the form of elastic stress waves due to the relative motion to characterize the internal structure of the object [34]. Since the AE signal is emitted by the fault source itself, the AE technology

can quickly detect and judge the gear pitting fault defect under the operating condition of the equipment.

The ACSAE method presented in this paper combines the one-dimensional CNN with the sparse autoencoder to enhance and optimize the network model and applies to original AE data for gear pitting fault diagnosis. The unsupervised layer-by-layer learning strategy can make the network parameters more precise. The problem of non-convex optimization makes the detection effect of gear pitting fault more stable and reliable. The CNN greatly reduces the number of network parameters through local weight sharing and can avoid the over-fitting of the network when the number of samples is insufficient [11]. However, the traditional CNN method of random initialization still has the risk of falling into a local optimum. In response to this situation, this paper presents a method to update traditional CNN using sparse autoencoder (SAE). Because SAE can compress the key features of the input signals, it is more effective to use these key data for the gear pitting diagnosis. As shown in Fig. 1, first, the original AE signals are encoded and compressed through convolutional networks. Then the convolutional features are deconvolved back to generate the reconstructed signals, and the weights are optimized to minimize the error between the reconstructed signals and the original signals. After that, the decoded features are added to the input data, encoded in the same way, and sequentially cycled. Finally, a fully connected network and a classifier are added to the network, and the network is fine-tuned with a small number of tagged samples to provide enhanced gear pitting fault identification.

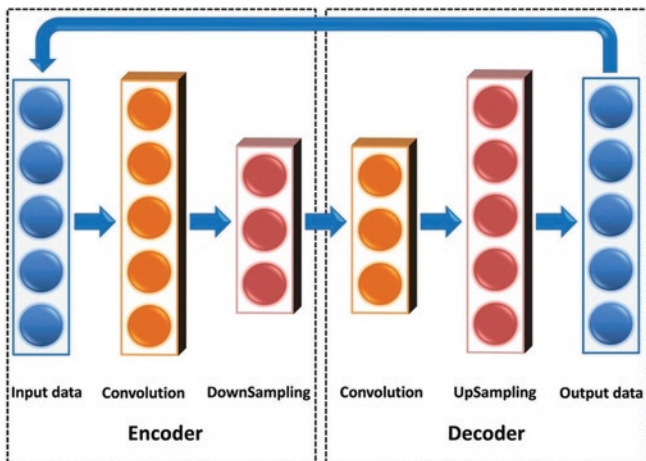


Fig. 1. The general procedure of the proposed ACSAE method

2.1. Sparse autoencoder

An autoencoder is an unsupervised learning method in deep learning, which uses a back propagation algorithm and has a good ability to learn the characteristics of the dataset [17]. Given a training sample set $\mathbf{X} = \{X_1, X_2, X_3, X_4, X_5, X_6\}$ for the autoencoder network structure as shown in Fig. 2, the autoencoder network tries to pass the original data \mathbf{X} , through the process of encoding and decoding, restore the data \mathbf{X}' so that $\mathbf{X}' \approx \mathbf{X}$. Different hidden layer units can have different activation levels for the input data. In general, a SAE network achieves the goal of sparseness by sparsely constraining the hidden layer of the autoencoder network, that is, only a small number of hidden layer units are activated [28]. SAE can usually learn a low-dimensional representation of input data that is very similar to principal component analysis results.

Fig. 2 shows an example of a neural network with three hidden layers. The entire neural network has a layer depth of 5, including an input layer, three hidden layers, and an output layer. Express the

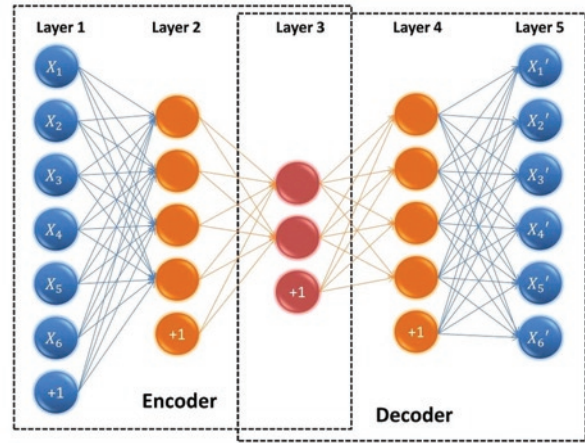


Fig. 2. The structure of autoencoder

equation in vector form, defining z_{j+1} in Eq. (1) and a_{j+1} in Eq. (2). Where W is the weight, and the weight is continuously updated by the gradient descent method to reduce the function loss. z_j is expressed as the sum of the multiplication of the input and weight values of the j^{th} layer and the bias b_j . f is the activation function:

$$z_{j+1} = W_j a_j + b_j \tag{1}$$

$$a_{j+1} = f(z_{j+1}) \tag{2}$$

Assume that a_j indicates the output value of the hidden layer neuron j with input X_i . Then the average activate degrees of the hidden layer neurons j — $\hat{\rho}_j$ can be computed in Eq. (3). ρ is a sparse parameter, which is close to zero:

$$\hat{\rho}_j = \frac{1}{m} \sum_{i=1}^m [a_j(X_i)] \tag{3}$$

To limit the average activate degrees where $\hat{\rho}_j$ is different from ρ by adding a penalty term to the objective function. This allows the average activate degrees of the hidden layer neurons to be distributed in a small range. s_2 represents the number of hidden neurons in the hidden layer, and j represents the j^{th} neuron in the hidden layer.

When $\hat{\rho}_j = \rho$, $\sum_{j=1}^{s_2} K_L(\rho \parallel \hat{\rho}_j) = 0$. If the difference between them increases, the result increases. Based on relative entropy, the sum of the selected penalty factors (4) as follow:

$$\sum_{j=1}^{s_2} K_L(\rho \parallel \hat{\rho}_j) = \sum_{j=1}^{s_2} \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j} \tag{4}$$

The total cost function of the sparse autoencoder system can be expressed as Eq (5). In Eq (5), s_2 represents the number of neurons in the hidden layer, β represents the weight of the penalty term, W is the coding matrix, and b is the coding deviation. As the gap between the two sparse parameter larger, the value of the penalty factor rises sharply:

$$J_{sparse}(W, b) = J(W, b) + \lambda \sum_{j=1}^{s_2} K_L(p \| p'_j) \quad (5)$$

2.2. One-dimensional CNN

A CNN is a feedforward neural network. As shown in Fig. 3, it is usually composed of alternating convolutional layers and pooling layers. The convolutional layer can capture the regional connection features in the input data, and the principle of weight sharing is applied. The principle of weight sharing of convolution kernels is due to the fact that there can be multiple convolution kernels in a convolutional layer, but each convolution kernel corresponds to a unique filtering feature. All features of the same sample are from identical convolution kernels. The number of parameters to train is greatly reduced in CNN. The pooling layer combines adjacent nodes into one to merge similar features, further reducing the amount of data trained. These characteristics make the CNN have a certain degree of translation, scaling, and distortion invariance. The fewer parameters make the training faster. The backpropagation algorithm is used to update the weight in the training process. The convolution kernel of a one-dimensional CNN can be seen as a sliding on a time series, extracting short-term features between sequences. These features are further aggregated by the pooling layer, which preserves the local features of the time series well, and then directly connects back and forth through the fully connected layer.

CNNs have been widely used in the fields of image and audio learning because of their good properties [24]. However, traditional CNNs use supervised learning algorithms, which require a large amount of supervised data in the learning process, which is costly. The key feature of CNNs is a local connection and weight sharing. For ease of understanding, both the input variable X and the model parameter W_i are represented by a matrix, \times represents a convolution operation, b_i is a bias, and only one bias is introduced for each feature, otherwise the degree of freedom of learning is too large. σ is the activation function. The encoding process for CNN can be defined as the following equation:

$$Y_i = \sigma(X \times W_i + b_i) \quad (6)$$

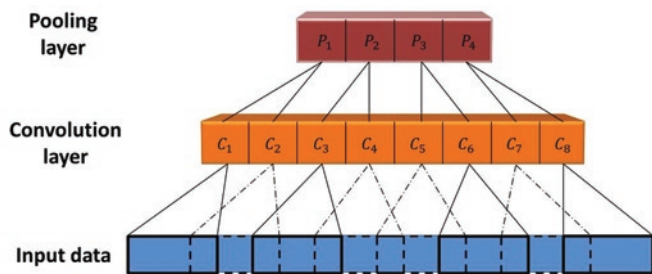


Fig. 3. The principle of one-dimensional CNN.

In this paper, we need to deal with the time series of AE signals. Therefore, one-dimensional convolution is used as the convolution layer to construct a one-dimensional CNN suitable for feature extraction of AE signals. Given an input signal sequence $X_i, i = 1, \dots, n$, and filters $W_j, j = 1, \dots, m$. The filter sequentially performs a local

convolution operation on the input features. In general, the length m of the filter is much smaller than the length n of the signal sequence. The output of the convolution is defined as:

$$y_i = \sum_{k=1}^m W_j \times X_{i-j+1} \quad (7)$$

In the convolutional layer, each neuron in the L layer is connected to neurons in a part of the $L-1$ layer to form a local connection network. The convolutional layer requires an activation function $f(x)$ for nonlinear feature mapping, and W^L is an m -dimensional filter that is the same for all neurons. $a_{(i+m-1)i}^L = [a_{(i+m-1)i}^L, \dots, a_{(i+m-1)i}^L]$, b_i is partial set the function, $i = 1, 2, \dots, n$. Then the output of the neuron of the L layer is defined as:

$$a_i^L = f\left(\sum_{j=1}^m W_j^L \times a_{i-j+m}^{L-1} + b^L\right) = f\left(W^L \times a_{(i+m-1)i}^{L-1} + b_i\right) \quad (8)$$

Pooling is a self-sampling process that greatly reduces the number of features, avoids overfitting, and allows the next layer of neurons to remain invariant to small morphological changes, providing strong robustness. The operation of the pooling layer is also a feature obtained by some way from a region. The common pooling method is to take the maximum or average value of all neurons in the region. For a feature map a_i obtained by the convolutional layer, it is divided into a plurality of regions $R_k, k = 1, \dots, n$. Taking the maximum value of each region as:

$$pool_{max}(R_k) = \max_{i \in R_k} a_i \quad (9)$$

The base of the convolution operation in Eq. (6) is added to the decoding portion of the autoencoder. During the decoding process, multiple feature values are obtained according to the hidden layer. Specifically, each deconvolution operation is performed by using a convolution kernel, and c is used as a bias. After the results obtained by the convolution kernels, an activation function is used to obtain an activation value of the output variable X' . The decoding process of reconstructing the input variable X' to the convolutional autoencoder network is defined as:

$$X' = \sigma\left(\sum_{j \in Y} Y_j * X * W'_j + c_j\right) \quad (10)$$

In addition, to ensure that the decoding process can restore the data to its original size, the CNN uses full convolution in both the encoding and decoding processes. In the ACSAE network, the pooling layer and the de-pooling layer are added after the corresponding convolution layer and deconvolution layer. Due to the information loss in the pooling operation, the convolution autoencoder network has a reconstruction error. An approximate representation of the input data is obtained by decoding. However, due to the fact that the reconstruction of the AE signal accuracy is not important for the general target task, the compressed and decompressed AE data has no distortion.

3. Gear test experimental setup and data processing

In order to verify whether the presented method is effective or not for gear pitting fault diagnosis, an experiment was designed and conducted on a gearbox test rig in the laboratory. The schematic of

the gearbox test rig is shown in Fig. 4. It consists of two 45 kW Siemens servo motors, one of which is the drive motor and the other the load motor. The AE sensor was located on the gear housing close to the faulty gear. The main parameters of the gearbox are shown in Table 1.

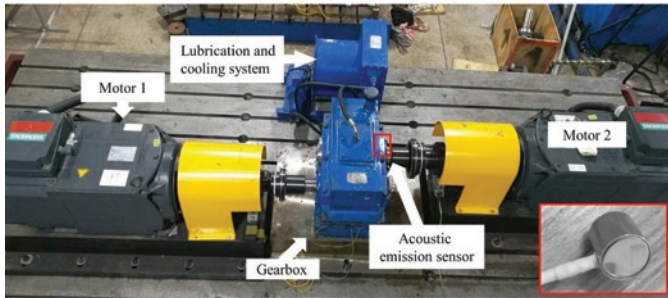


Fig. 4. Schematic of the gearbox test rig

Table 1. The major parameters of the gearbox

Gear Parameter	Driving Gear	Driving Gear
Tooth number	72	40
Module (mm)	3	3
Pitch diameter (mm)	120	120
Base circle diameter (mm)	202.974	112.763
Pressure angle (°)	20	20
Tooth width (mm)	85	85

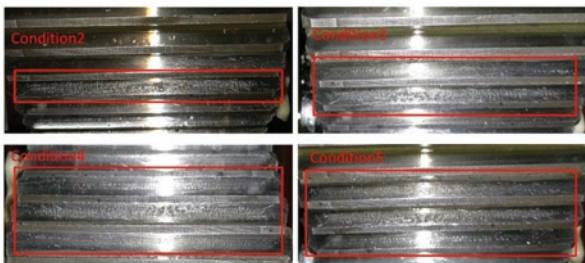


Fig. 5. Pitting degree of driven gears

In this experiment, the speed of the gear was set to 3600 RPM and 100 Nm. Table 2 shows the five gear pitting conditions. Condition 1 is a normal gear. In condition 2, the middle teeth were approximately

Table 2. The approximate percentage of pitting area under 5 gear conditions

Gear Condition	Upper tooth	Middle tooth	Lower tooth
Condition1	Normal	Normal	Normal
Condition2	Normal	50%	Normal
Condition3	10%	50%	Normal
Condition4	10%	50%	10%
Condition5	30%	50%	10%

50% worn and the other two teeth were normal. Condition 3 had 50% pitting on the intermediate gear, about 10% pitting on the upper teeth, and the other tooth was normal. Under condition 4, approximately 50% of the intermediate gears were worn out, and both the upper tooth and lower tooth had approximately 10% of the teeth been worn. Under condition 5, the intermediate teeth were approximately 50% worn, the upper teeth were 30% worn, and the lower teeth were 10% worn. The pitting conditions of the gear is shown in Fig. 5.

For AE data acquisition, a true differential wideband sensor with high sensitivity and bandwidth was used. It has a good frequency response over the range of 100–900 kHz. Differential sensors offer a lower noise output from a pre-amplifier. The original AE signals of 500 and 10,000 data points are shown in Figs 6(a) and (b) respectively. As can be seen from Fig. 6, the AE data in various states are very close, and it is almost impossible for the naked eye to judge which one of the sections should be.

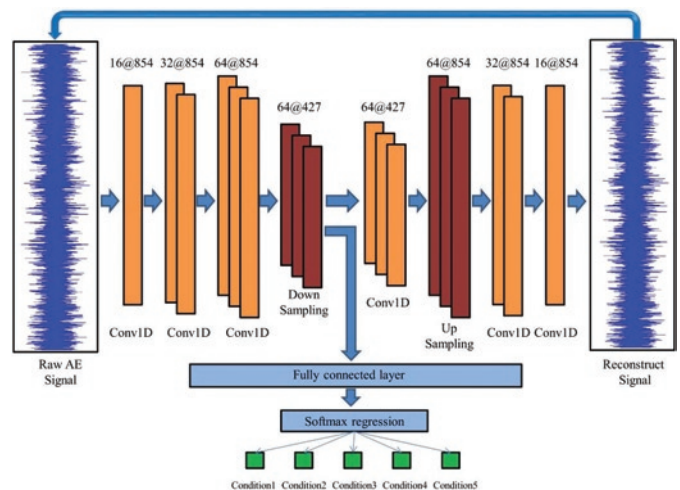
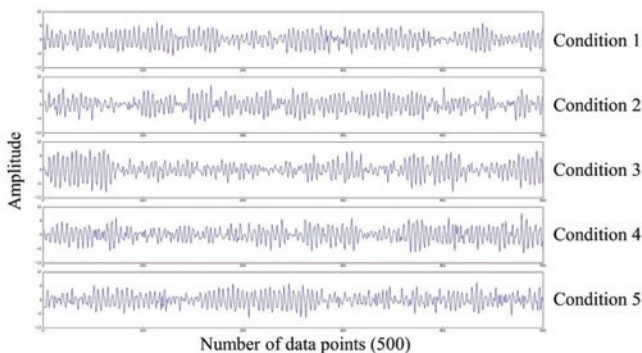
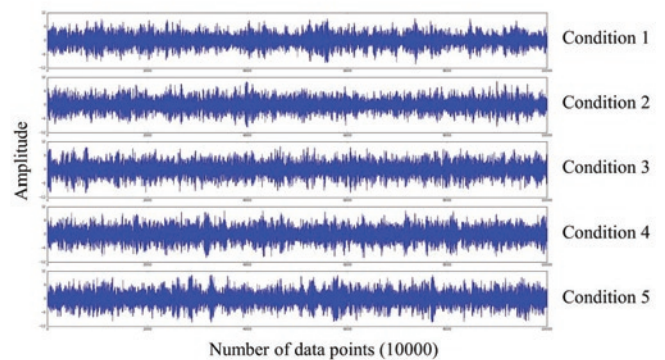


Fig. 7. The general procedure of the data processing using ACSAE method



a)



b)

Fig. 6. Raw AE signals of gear pitting fault conditions

The original AE signals were processed according to the process shown in Fig. 7. The test was conducted in five sessions with sampling features of 51,200. First, the training sample set and the test sample set were determined. 80% of the data were used for training and the rest of the data used for validation and testing. 854 data points were set as the length for each sample. After selecting the training samples and test samples, the specific classification steps are as follows:

- Step 1. Perform he-normal initialization [12] on the weight matrix of the network, and perform standard deviation random initialization on parameters such as offset.
- Step 2. Scale the data to make it sensitive to the activation function relu.
- Step 3. Use ACSAE to conduct layer-by-layer unsupervised training on the network.
 - a. Calculate the error between the actual output vector and the output vector.
 - b. The error obtained in the previous step is back-propagated layer by layer, and then the gradient of the error cost function is obtained by the stochastic gradient descent method, and then the weight parameters are updated.
 - c. Perform multiple iterations to improve the accuracy of the network.
 - d. The iteration is stopped when the specified iteration termination condition is reached.
 - e. Re-tune the network with the expanded data.
- Step 4. Retain the weight of the encoder part and fine-tune the network by two fully connected layers and a softmax classification layer.
- Step 5. Enter the test sample into the trained neural network to obtain the classification accuracy.

The structure of the ADCAE model designed for the number of channels in each convolution network was: 16-32-64-64-32-16. In the convolutional layer, the kernel size of all convolutional layers were 7; stride was 1; padding selects 'same'. In the pooling layer, all pool sizes were 2. padding selects 'valid'. The optimizer used stochastic gradient descent algorithm, and loss used categorical cross-entropy. The initial value of learning rate was set as 0.001, decay as 1e-6, momentum as 0.9, and activation set as relu.

4. Results and discussions

Fig. 8 shows a comparison of accuracy in the diagnosis of gear pitting faults between the ADCAE algorithm and other algorithms. As shown in Fig. 8, the accuracy of the ACSAE algorithm is close to 98% for gear pitting fault diagnosis. Its fluctuations are small and more stable. The accuracy of the CNN algorithm is approximately 90%. The fully-connected network (FCN) performed well, reaching 93%. The two methods of the recurrent neural network, gated recurrent unit (GRU) and long short term memory (LSTM), did not reach 85%. In summary, the ADCAE algorithm is superior to other algorithms for gear pitting fault diagnosis.

Table 3. The accuracy comparison between the ACSAE algorithm and the other algorithm's for acoustic emission data sets

Accuracy	Training set	Validation set	Testing set
ACSAE	0.9787	0.9774	0.9753
FCN	0.9340	0.9328	0.9311
CNN	0.9014	0.8947	0.8836
GRU	0.8179	0.8141	0.8119
LSTM	0.6793	0.6782	0.6567

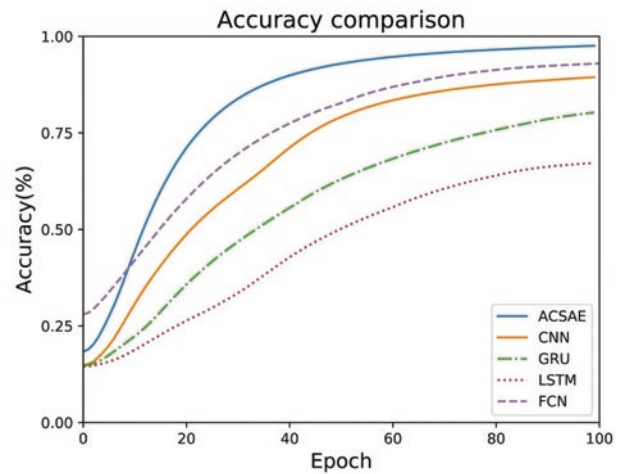


Fig. 8. The comparison of accuracy between the ACSAE algorithm and the other algorithm for gear pitting fault diagnosis

Table 3 provides a comparison of the accuracy of the ACSAE algorithm and other algorithms, including the training set, the validation set, and the test set. As shown in Table 3, the accuracy of the ACSAE for the training can be as high as 97.87%. The accuracy of the validation and test set is close to the accuracy of the training. The results show that the ACSAE algorithm can achieve good gear pitting fault diagnosis results without observed overfitting problem. In contrast, the accuracy and training accuracy of CNN algorithms and other algorithms used for validation and testing are not as good as the ACSAE algorithm. As a result, ACSAE's gear pitting fault diagnosis results are more accurate using the original AE signals.

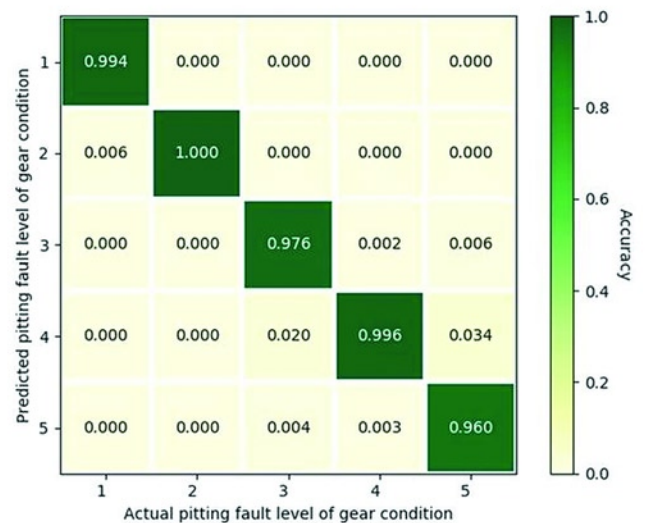


Fig. 9. The confusion matrix by the ACSAE algorithm for the testing set

Each type of data was tested in 500 groups. The test results are shown in Fig. 9. It can be seen from the confusion matrix of Fig. 9 that the ACSAE algorithm presented in this paper has a better gear pitting fault diagnosis. The pitting fault diagnosis accuracy of condition 1 has over 99%. The recognition rate of condition 2 is 100%. For the condition 3, 4, and 5, because the working conditions are very similar, there has little misjudgment, but the lowest accuracy is still above 96%.

For each gear condition, 50 sets of test data were used for visualization of the dimensionality reduction. Fig. 10 is a three-

dimensional view of the features. It can be seen from Fig. 10 that the ACSAE algorithm presented in this paper can clearly distinguish between various gear conditions. Clustering formed in Fig. 10 can also verify the previous conclusions, the overall recognition rate is higher, and some points that are slightly closer to 3, 4 and 5 are likely to cause misjudgment of the neural network.

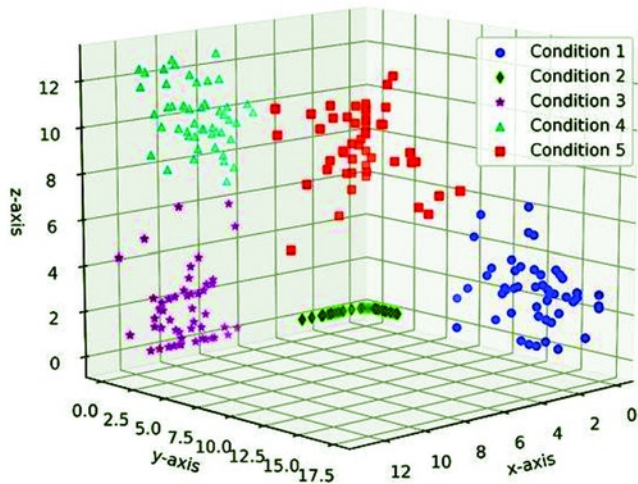


Fig. 10. The visualization of three-dimensional features of gear pitting conditions

The compressed three-dimensional features are visualized in two dimensions using t-SNE as shown in Fig. 11. As can be seen from Fig. 11, five different gear pitting faults can be clearly clustered. It can also be seen from the figure that the point adjacent in the high-dimensional data space are similar to the low-dimensional projections, which proves that the ACSAE is effective for the diagnosis of gear pitting faults.

These results have shown that the ACSAE algorithm is effective for using the original AE signals for gear pitting fault diagnosis. It is worth noting that this study only examines the case where the pitting of the intermediate gear, and the different degree of pitting of the adjacent gears. It is not clear whether the expected results will be achieved

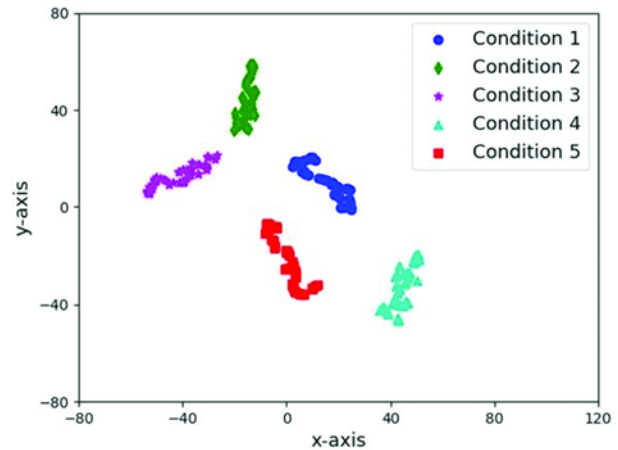


Fig. 11. The visualization of two-dimensional features of gear pitting conditions

when the type of failure is more complicated. More in-depth research will be conducted in the future.

5. Conclusions

In this paper, a gear pitting fault diagnosis method based on the integration of one-dimensional CNN and sparse autoencoder was presented. The presented method was validated and tested with AE data collected from a gear test rig in the laboratory. The validation and test results have shown that the presented method has achieved more stable and reliable performance, and the accuracy of fault diagnosis has reached 97.9%. The comparative results have shown that the diagnosis accuracy using the ACSAE algorithm is better than other methods.

Acknowledgement

This research was funded by National Natural Science Foundation of China (Grant No. 51675089).

References

1. Ali Y H, Abd Rahman R, Hamzah R I R. Artificial Neural Network Model for Monitoring Oil Film Regime in Spur Gear Based on Acoustic Emission Data. *Shock and Vibration*, 2015, <https://doi.org/10.1155/2015/106945>.
2. Aouabdi S, Taibi M, Bouras S, Boutasseta N. Using multi-scale entropy and principal component analysis to monitor gears degradation via the motor current signature analysis. *Mechanical Systems and Signal Processing* 2017; 90: 298-316, <https://doi.org/10.1016/j.ymssp.2016.12.027>.
3. Bafroui H H, Ohadi A. Application of wavelet energy and Shannon entropy for feature extraction in gearbox fault detection under varying speed conditions. *Neurocomputing* 2014; 133: 437-445, <https://doi.org/10.1016/j.neucom.2013.12.018>.
4. Bangalore P, Tjernberg L B. An Artificial Neural Network Approach for Early Fault Detection of Gearbox Bearings. *Ieee Transactions on Smart Grid* 2015; 6: 980-987, <https://doi.org/10.1109/TSG.2014.2386305>.
5. Cao P, Zhang S L, Tang J. Preprocessing-Free Gear Fault Diagnosis Using Small Datasets With Deep Convolutional Neural Network-Based Transfer Learning. *Ieee Access* 2018; 6: 26241-26253, <https://doi.org/10.1109/ACCESS.2018.2837621>.
6. Chen X H, Cheng G, Li H Y, Li Y. Fault identification method for planetary gear based on DT-CWT threshold denoising and LE. *Journal of Mechanical Science and Technology* 2017; 31: 1035-1047, <https://doi.org/10.1007/s12206-017-0202-5>.
7. Crivelli D, McCrory J, Miccoli S, Pullin R, Clarke A. Gear tooth root fatigue test monitoring with continuous acoustic emission: Advanced signal processing techniques for detection of incipient failure. *Structural Health Monitoring-an International Journal* 2018; 17: 423-433, <https://doi.org/10.1177/1475921717700567>.
8. Elforjani M, Shanbr S. Prognosis of Bearing Acoustic Emission Signals Using Supervised Machine Learning. *Ieee Transactions on Industrial Electronics* 2018; 65: 5864-5871, <https://doi.org/10.1109/TIE.2017.2767551>.
9. Feng Z P, Liang P M. Fault diagnosis of wind turbine planetary gearbox under nonstationary conditions via adaptive optimal kernel time frequency analysis. *Renewable Energy* 2014; 66: 468-477, <https://doi.org/10.1016/j.renene.2013.12.047>.

10. Feng Z P, Liang P M, Zhang Y, Hou S M. Fault diagnosis for wind turbine planetary gearboxes via demodulation analysis based on ensemble empirical mode decomposition and energy separation. *Renewable Energy* 2012; 47: 112-126, <https://doi.org/10.1016/j.renene.2012.04.019>.
11. Guo S, Yang T, Gao W, Zhang C. A Novel Fault Diagnosis Method for Rotating Machinery Based on a Convolutional Neural Network. *Sensors* 2018; 18(5): 1429-1444, <https://doi.org/10.3390/s18051429>.
12. He K M, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international conference* 2015; 1026-1034, <https://doi.org/10.1109/ICCV.2015.123>.
13. He M, He D. Deep Learning Based Approach for Bearing Fault Diagnosis. *Ieee Transactions on Industry Applications* 2017; 53: 3057-3065, <https://doi.org/10.1109/TIA.2017.2661250>.
14. Krishnakumari A, Elayaperumal A, Saravanan M, Arvindan C. Fault diagnostics of spur gear using decision tree and fuzzy classifier. *International Journal of Advanced Manufacturing Technology* 2017; 89: 3487-3494, <https://doi.org/10.1007/s00170-016-9307-8>.
15. Liu X, Jia Y X, He Z W, Zhou J. Application of EMD-WVD and particle filter for gearbox fault feature extraction and remaining useful life prediction. *Journal of Vibroengineering* 2017; 19: 1793-1808, <https://doi.org/10.21595/jve.2017.17680>.
16. Li Z, Ma Z Y, Liu Y B, Teng W, Jiang R. Crack Fault Detection for a Gearbox Using Discrete Wavelet Transform and an Adaptive Resonance Theory Neural Network. *Strojniski Vestnik-Journal of Mechanical Engineering* 2015; 61: 63-73, <https://doi.org/10.5545/sv-jme.2014.1769>.
17. Lu C, Wang Z Y, Qin W L, Ma J. Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification. *Signal Process* 2017; 130: 377-388, <https://doi.org/10.1016/j.sigpro.2016.07.028>.
18. Praveenkumar T, Saimurugan M, Krishnakumar P, Ramachandran K I. Fault diagnosis of automobile gearbox based on machine learning techniques. *12th Global Congress on Manufacturing and Management* 2014; 97: 2092-2098, <https://doi.org/10.1016/j.proeng.2014.12.452>.
19. Qu Y Z, He M, Deutsch J, He D. Detection of Pitting in Gears Using a Deep Sparse Autoencoder. *Applied Sciences-Basel* 2017; 7(5): 515-529, <https://doi.org/10.3390/app7050515>.
20. Qu Y Z, Zhu J D, He D, Qiu B, Bechhoefer E. Development of a New Acoustic Emission Based Fault Diagnosis Tool for Gearbox. *2013 Ieee International Conference on Prognostics and Health Management*, 2013; 1-9, <https://doi.org/10.1109/ICPHM.2013.6621418>.
21. Ratni A, Rahmoune C, Benazzouz D. A new method to enhance of fault detection and diagnosis in gearbox systems. *Journal of Vibroengineering* 2017; 19: 176-188, <https://doi.org/10.21595/jve.2016.17214>.
22. Shao R P, Hu W T, Wang Y Y, Qi X K. The fault feature extraction and classification of gear using principal component analysis and kernel principal component analysis based on the wavelet packet transform. *Measurement* 2014; 54: 118-132, <https://doi.org/10.1016/j.measurement.2014.04.016>.
23. Sharma R B, Parey A. Modelling of acoustic emission generated due to pitting on spur gear. *Engineering Failure Analysis* 2018; 86: 1-20, <https://doi.org/10.1016/j.engfailanal.2017.12.016>.
24. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
25. Song M M, Xiao S G. A Fault Diagnosis Method of Gear Based on SVD and Improved EEMD. *Intelligent Computing, Networked Control, and Their Engineering Applications* 2017; 762: 65-74, https://doi.org/10.1007/978-981-10-6373-2_7.
26. Sreepadha C, Kumari A K, Perumal A E, Panda R C, Harshabardhan K, Aribalagan M. Neural network model for condition monitoring of wear and film thickness in a gearbox. *Neural Computing & Applications* 2014; 24: 1943-1952, <https://doi.org/10.1007/s00521-013-1427-6>.
27. Widodo A, Satrijo D, Prahasto T, Haryanto I. Fault Detection of Gearbox Using Time-Frequency Method. *7th International Conference on Mechanical and Manufacturing Engineering* 2017; 1831, <https://doi.org/10.1063/1.4981194>.
28. Yang Z, Jing H. A deep learning method based on hybrid auto-encoder model. *Technology, Networking, Electronic and Automation Control Conference* 2017; 1100-1104, <https://doi.org/10.1109/ITNEC.2017.8284911>.
29. Yao Z T, Hu Y C. Gearbox fault diagnosis based on LMD and Cyclostationary Demodulation. *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence* 2016; 984-989.
30. Yuan J, Ji F, Gao Y, Zhu J, Wei C J, Zhou Y. Integrated ensemble noise-reconstructed empirical mode decomposition for mechanical fault detection. *Mechanical Systems and Signal Processing* 2018; 104: 323-346, <https://doi.org/10.1016/j.ymssp.2017.11.004>.
31. Zhang Q, Zhao W, Xiao S G. Fault Diagnosis of Gear Based on Singular Value Decomposition and RBF Neural Network. *2017 2nd International Conference on Frontiers of Sensors Technologies* 2017; 470-474.
32. Zhang Q, Zhao W, Xiao S G, Song M M. Method of Gear Fault Diagnosis Based on EEMD and Improved Elman Neural Network. *Materials Science, Energy Technology, and Power Engineering I* 2017; 1839(1): 020111, <https://doi.org/10.1063/1.4982476>.
33. Zhao R, Wang D Z, Yan R Q, Mao K Z, Shen F, Wang J J. Machine Health Monitoring Using Local Feature-Based Gated Recurrent Unit Networks. *Ieee Transactions on Industrial Electronics* 2018; 65: 1539-1548, <https://doi.org/10.1109/TIE.2017.2733438>.
34. Zhou Y, Lin L, Wang D, He M, He D. A new method to classify railway vehicle axle fatigue crack AE signal. *Applied Acoustics* 2018; 131: 174-185, <https://doi.org/10.1016/j.apacoust.2017.10.025>.
35. Zuber N, Bajric R, Sostakov R. Gearbox Faults Identification Using Vibration Signal Analysis and Artificial Intelligence Methods. *Eksploracija I Niezawodnosc-Maintenance and Reliability* 2014; 16: 61-65.

Xueyi LI**Jialin LI**

School of Mechanical Engineering and Automation
Northeastern University
3-11 Wenhua Road, Heping District
Shenyang, 110819, China

David HE

Department of Mechanical and Industrial Engineering
University of Illinois at Chicago
842 West Taylor Street, Chicago, IL 60607, USA

Yongzhi QU

School of Mechanical and Electronic Engineering
Wuhan University of Technology
122 Luoshi Road, Wuhan, 430070, China

Emails: lixueyineu@gmail.com, jialinli_neu@163.com,
davidhe@uic.edu, quwong@whut.edu.cn