Fault diagnosis of gas turbine based on matrix capsules with EM routing

Yunji Zhao, Menglin Zhou, Nannan Zhang, Xiaozhuo Xu and Xinliang Zhang

School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo, People’s Republic of China

ABSTRACT
The fault detection and diagnosis of a gas turbine is of great significance for guaranteeing the complicated dynamic systems working normally and safely. Most of the existing fault diagnosis methods, based on convolutional neural networks (CNN), have certain limitations in extracting correlations of multi-channel data features. The accuracy of fault diagnosis still needs to be improved. In this paper, an approach of fault diagnosis, based on matrix capsules with EM routing, is presented. First of all, three channels data, which respectively represent acceleration, pressure and pulse, are integrated into one image to feed into the network. Secondly, network models based on the matrix capsules start to be trained by using input dataset which contains fault image and normal image. Finally, the pre-trained capsules model is used to diagnose the state of testing data. Besides, to verify the superiority of the algorithm used in this paper, a comparative experiment is implemented between matrix capsule networks and CNN. The results demonstrate that the testing accuracy is 99.995%.

1. Introduction
A gas turbine engine is an air-dependent, thermal jet propulsion device that uses exhaust-gas-driven turbine wheels to drive a compressor, making continuous operation of the engine possible. The gas turbine can use a variety of fuels such as gasoline, diesel, kerosene, peanut oil and almost any flammable liquid or even mixed liquids. There is no cold start problem for gas turbine, so it is not necessary to add antifreeze liquid in a cold weather area (Petrenya, 2019). Due to the above advantages, the gas turbine is widely used in military ships and electrical systems. With the wide use of gas turbine, it is important to accurately diagnose whether the machine is running normally because it will incur huge cost if it breaks down (Alblawi, 2020).

The traditional methods of fault diagnosis always consist of three parts: estimation techniques based on mathematical models and parameters (Kuestenmacher & Plöger, 2016; Wang et al., 2019), analysis methods based on data such as mathematical statistics (Li et al., 2019; Liu et al., 2020) and expertise approaches based on expert experience and knowledge. However, it is extremely difficult to accurately accomplish the fault diagnosis by building a mathematical model. It is also difficult to analyse whether the machine is running normally based on expertise knowledge in complex working environments and massive working data.

In recent years, artificial intelligence has been widely used in many fields especially in industrial processes with its rapid development. It greatly accelerates the development of intelligent fault diagnosis methods. For example, Li et al propose a method based on empirical wavelet transform (EWT) to determine the number and boundaries of modes (Li et al., 2020). The binary tree SVM, based on the analytic hierarchy process proposed by Miao D et al, aims at achieving classification accuracy by structure design (Miao, 2019). Zhang et al propose an unsupervised method based on fast intrinsic component filtering and pseudo-normalization to get consistent features and achieve robust results (Zhang et al. 2020). And there is a development of fault diagnosis method from the traditional artificial neural network, such as BP neural network, (Bi & Liu, 2016; Huo et al., 2018; Liu & Zhang, 2019; Yan et al., 2019) to deep learning, such as convolution neural network (CNN) (Bi & Liu, 2016; Kou et al., 2020; Li et al., 2020; Shao et al., 2019; Zhao et al., 2020; Zhou et al., 2020). Or the combination of multiple methods that Liang et al improve testing accuracy even in complex working conditions by proposing an intelligent failure detection method via wavelet transform, generative adversarial nets and convolutional neural network (Liang et al., 2020). It is undeniable that the above methods have achieved some success in fault diagnosis to a certain extent. However, there still is a lot of room for improvement.
improving the accuracy and stability of fault diagnosis for the drawbacks of above-mentioned methods:

(1) The learning speed of BP neural network is slow. Because simple problems have to be learned hundreds of times and easy to fall into local minima.

(2) CNN has a limitation in extracting the correlation between multichannel data. It needs to use the same knowledge at all locations in an image and there are limitations in using the information of space such as the relationship between part and whole (Hinton et al., 2018). Besides, massive useful information will be lost after pooling which results in the output is almost constant when the input changes slightly.

(3) Massive amount of data needs to be collected. The collected data is input into the network to train a neural network model so that an accurate fault diagnosis can be achieved. However, there are not that much failure data when the gas turbine is in actual operation.

Therefore, it is still a vital and meaningful task to diagnose whether the machine is in normal operation by extracting relevant characteristics and features of limited input data accurately. Aiming at solving the problems mentioned above and achieve high accuracy for fault diagnosis, this paper proposed a method based on matrix capsules with EM routing whose advantages are listed as follows:

(1) The raw dataset for training is easy to get and there is no need to pre-process data but to combine three channels data that represent acceleration, pressure and pulse together.

(2) The accuracy of fault diagnosis is extremely high which can achieve 99.995% for the advantage of matrix capsules network that it can extract global features and the correlations among multi-channel data.

(3) The timeliness of fault detection is guaranteed. With the network model pre-trained, the detection speed is so fast that 10,000 groups of data just need 60 s.

This paper consists of four parts: materials and methods, experiment, results and discussion. In section 2, theory and principle of the methods used in this paper are elaborated in detail. Section 3 introduces how the experiment data have been gotten and the process of training a network. Section 4 shows results and section 5 is the overall discussion.

2. Materials and methods

2.1. Capsule network

The convolutional neural networks (CNN) understand and acquire the local features and information of the input through the convolutional layers. When low-level features and information are acquired, groups are combined into more complex and abstract features. Then pooling layers are needed to reduce the size of the output tensor or feature map, which will lose other information such as location. Based on such limitations of CNN, Hinton et al. proposed the concept of capsule network in Dynamic Routing Between Capsules (Sabour et al., 2017), using vectors instead of scalars as input and output, and understanding the characteristic properties of input such as rotation or equal scale transformation.

Capsules are essentially a set of neurons whose input/output vectors represent the instantiation parameters of a particular entity type (that is, the probability of the existing of a particular object and some properties). The length of the input/output vectors represents the probability of entity existence, which is compressed into the interval of 0–1, as shown in Formula (1). The direction of the vector represents the instantiation parameter. The instantiation parameters of a capsule i of the same level are predicted by the transformation matrix for a higher-level capsule j. For all i ∈ ΩL and j ∈ ΩL+1. When most of the vote prediction is consistent, the capsules are activated.

\[ V_j = \frac{||s_j||}{1 + ||s_j||^2} \]

\[ s_j = \sum_i c_i \hat{u}_{ij} \quad u_{ij} = W_{ij} u_i \]

\[ c_i = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \]  

where \( u_i \) is the output vector of the capsule i and Wij is learned through back propagation and loss function which represents vectors instead of elements. \( \hat{u}_{ij} \) is a linear combination of \( u_i \), which can also be understood as the strength of the capsule i in the lower layer connecting to the capsule j in the higher layer. bj is learned by back propagation and loss function. \( c_i \) denotes coupling coefficient, which is calculated iteratively through dynamic routing. At this point, the higher-level capsules become active when the multiple predictions are consistent. Flowchart of the dynamic routing capsule network is shown in Figure 1.

Capsule network was first used for handwritten numeral recognition. After regular convolution of the
convolutional layer, the input number images are converted to 20 * 20 characteristics which are 256-d. After the operation of PrimaryCaps, 32 capsules of 6x6 are obtained and each capsule contains eight components. It can be understood that PrimaryCaps is equivalent to a depth of 32 general convolution layers, each layer of previously only scalar value into a vector of length 8. Finally, 10 standard Capsule units with an output vector of 16 elements each are processed through DigitCaps network. Dynamic routing occurs between PrimaryCaps and DigitCaps. And the specific parameter settings are shown in Table 1 where k_s represents the size of the convolution kernels and s denotes the convolution stride.

### 2.2. Fault diagnosis based on matrix capsule with EM routing

Because of its own structural composition and characteristics, the capsules can capture the possibility of features and their variants. Therefore, the capsules can not only detect different features but also learn and detect the variants through training. After the dynamic routing between capsules, Hinton et al. proposed the matrix capsules with EM routing. The similarity with the dynamic routing capsule network lies in that the capsule j feature of the upper layer is detected through the voting agreement of capsule i of the lower layer. The difference is that each capsule layer of matrix capsule with EM routing contains two parts: pose matrix M and activation probability a. Both M and a are obtained by the EM routing algorithm using step E and step M alternately. The network of Matrix capsule with EM routing is shown in Figure 2.

In the network structure, A denotes channels of convolutional layers which can be 32 or 64. B, C and D represent capsule types of each capsule layer, respectively. B, C and D are 32, 32, 32, respectively if A = 32. Otherwise, B, C and D are 8, 16, 16, respectively. The voting V_ij from capsule i to parent capsule j can be calculated by multiplying the pose matrix M_i of capsule i by an angle invariant transformation matrix W_ij:

\[
    v_{ij} = M_i W_{ij} 
\]

W_ij is obtained through back propagation and loss function. Meanwhile, pose matrix M_i needs to be calculated by EM algorithm iteration. In the EM route, the pose matrix of the parent capsule j is modelled by the Gaussian model. Since the pose matrix is a 4 x 4 matrix with 16 elements, the attitude matrix is modelled by the Gaussian model with 16 \( \mu \) and 16 \( \sigma \). Each represents one element of the attitude matrix. The Gaussian probability density function
is as follows:

\[ P(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]  

(5)

Therefore, the probability that \( v_{ij} \) belongs to the Gaussian model of capsule \( j \):

\[
\begin{align*}
P_{ij} &= \frac{2}{\sqrt{2\pi}h^2} \exp \left( -\frac{(v_{ij} - \mu_j)^2}{2\sigma_j^2} \right) \\
\ln (P_{ij}) &= -\frac{(v_{ij} - \mu_j)^2}{2\sigma_j^2} - \frac{\ln (2\pi)}{2} - \ln (\sigma_j^2)
\end{align*}
\]  

(6)

where \( h \) denotes the \( h \)-th element in the matrix.

The cost of activated parent capsules is obtained by the following formula:

\[
\begin{align*}
cos t_{ij} &= \sum -r_{ij}\ln (P_{ij}) \\
&= \sum r_{ij} \left( \ln (\sigma_j) + \frac{\ln (2\pi)}{2} + \frac{1}{2} \right)
\end{align*}
\]  

(7)

\( r_{ij} \) denotes probability distribution. Due to the different degree of association between the lower capsule \( i \) and the upper capsule \( j \), the allocation probability \( r_{ij} \) is initially set as \( 1/\Omega (L + 1) \) and iteratively calculated by EM routing.

\( a_j \) is used to determine whether the upper capsule \( j \) is activated:

\[
a_j = \text{logistic} \left( \lambda \left( \beta_a - \sum \cos t_{ij} - \beta_\mu \sum r_{ij} \right) \right)
\]  

(8)

where \( \beta_a \) and \( \beta_\mu \) are discriminatively learned by back propagation and loss functions. \( \lambda \) is set as a fixed schedule as an inverse temperature parameter.

At this point, \( \mu \) and \( \sigma \) in the Gaussian as well as the activation value of \( a \) is calculated by M-step of EM routing based on the initial \( r_{ij} \). Then E-step of EM routing will recalculate \( r_{ij} \) based on current \( r_{ij}, a_j \) and \( V_{ij} \). The loop is repeated three times until the algorithm is completed.

3. Experiment

3.1. Data set acquisition

The working medium of Gas turbine comes from the atmosphere. First of all, the compressor continuously draws in air from the outside and compresses it. Then the compressed air is mixed with fuel in the combustion chamber to start burning. Gas is produced at the same time. The gas flows into the turbine and expands, producing acceleration, as the starter carries the gas turbine from static state. At this point, the turbine blade rotates with the compressor impeller until acceleration to the point where the gas turbine functions independently. Finally, the starter is detached and the gas turbine starts working.

Due to the gas turbine working principle, there are three kinds of states work data of gas turbine used for training fault diagnosis models based on Matrix Capsule with EM routing. They are 2000rpm_no_air (no load, air injection is not started), 2000rpm (air jet starts working, it contains normal data), 2000rpm_fault_simulation (air jet starts, it contains fault data), respectively. The composition of the three data types is shown in Table 2, the data of each channel are 655, 360 x 1.

Since Channel_1, the time dimension data have no influence on gas turbine fault detection; this channel data can be discarded. Channel_2, Channel_3 and Channel_4 data that represent acceleration, pressure and pulse, respectively, can be considered.

When using matrix capsule and EM routing algorithm for fault diagnosis, it is not necessary to determine the working state of the machine, but only to make it contain the three types of data mentioned above except the time dimension. Therefore, the acceleration, pressure and pulse of the three working states are combined to form three new channel matrices: channel_a, channel_pr and channel_pu which, respectively, represent the total acceleration, pressure and pulse. At this time, the data size of each channel is 1,966,080 x 1(655, 360 x 3 = 1,966,080), and the data in columns 1 - 1, 310, 720(655, 360 x 2 = 1, 310, 720) are the acceleration, pressure and pulse values during normal operation, while the data in columns 1, 310, 721 - 1, 966, 080 are fault data.

The original matrix capsule and EM routing network are applied to handwritten digit recognition, with input data of 28 x 28 images. Therefore, in order to successfully apply the input gas turbine work data to matrix capsule and EM routing network, 261, 261 and 262 data can be collected from channel_a, channel_pr and channel_pu, respectively to form a 1 x 784 matrix and recombine it into 28 x 28 images. The specific operation is to desperately collect 35,000 sets of normal data from 1 to 1,310,720 in three channels with 35 steps and 261, 261 and 262 as the total number. With 10 steps and 261, 261 and 262 as the total number, 35,000 sets of fault data are separately

| State               | Time   | Acceleration | Pressure | Pulse |
|---------------------|--------|--------------|----------|-------|
| 2000 rpm_no_air     | Channel_1 | Channel_2    | Channel_3 | Channel_4 |
| 2000 rpm            | Channel_1 | Channel_2    | Channel_3 | Channel_4 |
| 2000 rpm_fault_simulation | Channel_1 | Channel_2    | Channel_3 | Channel_4 |
collected from 1,310,721–1,966,080 in the three channels. The three features of fault-free data and normal data are, respectively, combined to form a data matrix containing 70,000 sets of 784-dimensional samples. Among them, the positive and negative samples account for half, 60,000 groups are used for training and 10,000 groups are used for testing. The flowchart of the data preprocess is shown in Figure 3.

3.2. Experimental data and settings

The simulation experiment used Python as the programming language and Pytorch as the deep learning framework. It ran at the Ubuntu system, Intel(R) Core(TM) i7-9700 K processor, 32.0GB of memory and NVIDIA GeForce GTX 1080. After collecting and sorting out 70,000 sets of 784-dimensional gas turbine working data, the matrix capsule and EM routing network were first converted into 28*28-dimensional image information.

During the experiment, the 4 × 4 trainable transformation matrix \( W_{ij} \) was learned by back propagation and loss function. \( r_{ij}, \mu, \sigma \) and \( q_{ij} \) were iteratively calculated through EM routing. The batch size was 64 and the batch number was 935, and the dynamic learning rate was used for network training. The training and testing were repeated for five times, and the average value was taken as the accuracy of the final result. Parameter Settings in matrix capsule and EM routing algorithm network are shown in Table 3 where \( k_s \) is the convolutional kernel size and \( s \) denotes the convolution stride.

### Table 3. Parameters in matrix capsule and EM routing.

| Layer          | Using       | Parameters                  | Output          |
|----------------|-------------|-----------------------------|-----------------|
| Input          | Picture     |                             | (28*28)         |
| ReLU Conv      | \( k_s = 5*5, s = 2 \) | \( 5*5*1*32 \) \( \rightarrow \) \( 14*14*32 \) |                 |
| PrimaryCaps    | \( k_s = 1*1, s = 1 \) | \( 32*32*4+1 \) \( \rightarrow \) \( 14*14*32*4 \) \( \rightarrow \) \( Activation(14*14*32*1) \) |                 |
| ConvCaps1      | \( k_s = 3*3, s = 2 \) | \( 3*3*32*4 \) \( \rightarrow \) \( 6*6*32*4 \) \( \rightarrow \) \( Pose(6*6*32*4) \) \( \rightarrow \) \( Pose(6*6*32*4) \) |                 |
| ConvCaps2      | \( k_s = 3*3, s = 1 \) | \( 3*3*32*4 \) \( \rightarrow \) \( 4*4*32*4 \) \( \rightarrow \) \( Pose(4*4*32*4) \) \( \rightarrow \) \( Pose(4*4*32*4) \) \( \rightarrow \) \( Activation(4*4*32*1) \) \( \rightarrow \) \( Activation(4*4*32*1) \) |                 |
| Class Capsules | \( k_s = 1*1 \) | \( 32*2*4*4 \) \( \rightarrow \) \( 2*4*4 \) \( \rightarrow \) \( Pose(2*4*4) \) \( \rightarrow \) \( Activation(2*1) \) |                 |

### Table 4. Training and testing accuracy at different epochs.

| Epoch   | Train accuracy | Test accuracy                  |
|---------|----------------|--------------------------------|
| Epoch_0 | 96.665%        | 100%/10,000/10,000            |
| Epoch_1 | 100%           | 100%/10,000/10,000            |
| Epoch_2 | 99.993%        | 100%/10,000/10,000            |
| Epoch_3 | 100%           | 99.99%/9999/10,000            |
| Epoch_4 | 100%           | 99.99%/9999/10,000            |

superiority of the fault diagnosis algorithm used in this paper, a comparison experiment between matrix capsules network and CNN was furthermore set. Obviously, two methods should start with the same dataset so that it would be more intuitive to reflect the contrast effects.

The CNN methods for fault diagnosis, used in this paper, are Siamese network for bearing fault diagnosis. The dataset used was Case Western Reserve University Bearing fault data which has five kinds of fault and can be downloaded from GitHub. The result of the accuracy of Siamese network is 99.47%. The accuracy of fault diagnosis by matrix capsules with EM routing can achieve 99.511% that 2242 groups were diagnosed correctly in 2253 groups of total testing data with the images transformed by Case Western Reserve University Bearing fault data which perform better than Siamese network.

### 3.3. Comparison between CNN and matrix capsules

In the actual running of the machine systems, faced with the complex characteristics of data obtained, diagnostic errors of conventional detection methods still exist in the fault diagnosis filed. In order to clearly verify the
4. Results

The 60,000 groups data are input into the network to start to train a model, then 10,000 groups data are input the trained module for testing. The training and test accuracy can be obtained, respectively. When $A = 64$, $B = 8$, $C = 16$ and $D = 16$, the training and testing accuracy are shown in Table 4. When $A = B = C = D = 32$, there is little difference on testing accuracy. The comparison is shown in Figure 4 where ‘test1’ ‘test2’ dividely on behalf of the former ($A = 64$, $B = 8$, $C = 16$ and $D = 16$) and the later ($A = B = C = D = 64$).

5. Discussion

Matrix capsules with EM routing can extract correlations between multi-channel data. Because of its own characteristics, this paper proposed a method by applying matrix capsules with EM routing for gas turbine fault diagnosis. The lower capsules extract basic whole features and characteristics and then input to higher capsules by voting agreement with distribution coefficients. The extracted features are comprehensive and can feed in different capsules that deal with specific modules. After the training of matrix capsules with EM, the accuracy of gas turbine fault diagnosis can obtain 99.995%. Besides, the training and testing results of the network both perform steadily. The results demonstrate that the gas turbine fault diagnosis based on matrix capsules with EM routing performs better than other traditional methods.

Acknowledgements

We thank the anonymous reviewers for their valuable suggestions and comments on this paper. We also thank the authors of Matrix Capsules with EM Routing for providing source code.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Foundation of Henan Educational Committee grant number 16A413009 and 13B413037; The National Natural Science Foundation of China grant number 61973105 and 61573130; The National Natural Science Foundation of China (NO.U1504506).

References

Alblawi, A. (2020). Fault diagnosis of an industrial gas turbine based on the thermodynamic model coupled with a multi feedforward artificial neural networks. Energy Reports, 6, 1083–1096. https://doi.org/10.1016/j.egyr.2020.04.029

Bi, F. R., & Liu, Y. P. (2016). Fault diagnosis of valve clearance in diesel engine based on BP neural network and support vector machine. Transactions of Tianjin University, 22(6), 536–543. https://doi.org/10.1007/s12209-016-2675-1

Hinton, G. E., Sabour, S., & Frosst, N. (2018). Matrix capsules with EM routing, 6th International Conference on Learning Representations, ICLR, April 30 - May 3, 2018, Vancouver Convention Center, Vancouver, BC, Canada.

Huo, L., Zhang, X. Y., & Li, H. D. (2018). Bearing fault diagnosis based on BP neural network. IOP Conference Series: Earth and Environmental Science, 208, article 012092. https://doi.org/10.1088/1755-1315/208/1/012092
Kou, L. L., Qin, Y., Zhao, X. J., & Chan, X. A. (2020). A multi-dimension end-to-end CNN model for rotating devices fault diagnosis on high-speed train bogie. *IEEE Transactions on Vehicular Technology, 69*(3), 2513–2524. https://doi.org/10.1109/TVT.2019.2955221

Kuestenmacher, A., & Plöger, P. G. (2016). Model-based fault diagnosis techniques for mobile robots. *IFAC-PapersOnLine, 49*(15), 50–56. https://doi.org/10.1016/j.ifacol.2016.07.613

Li, Y. B., Du, X. Q., Wan, F. Y., Wang, X. Z., & Yu, H. C. (2020). Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging. *Chinese Journal of Aeronautics, 33*(2), 427–438. https://doi.org/10.1016/j.cja.2019.08.014

Li, J. M., Wang, H., Wang, X. D., & Zhang, Y. G. (2020). Rolling bearing fault diagnosis based on improved adaptive parameterless empirical wavelet transform and sparse denoising. *Measurement, 152*, article 107392. https://doi.org/10.1016/j.measurement.2019.107392

Li, J., Ying, Y., & Ji, C. (2019). Study on gas turbine gas-path fault diagnosis method based on quadratic entropy feature extraction. *IEEE Access, 7*, 89118–89127. https://doi.org/10.1109/ACCESS.2019.2927306

Liang, P. F., Deng, C., Wu, J., & Yang, Z. X. (2020). Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network. *Measurement, 159*, article 107768. https://doi.org/10.1016/j.measurement.2020.107768

Liu, W., Chen, W., & Zhang, Z. (2020). A novel fault diagnosis approach for rolling bearing based on high-order synchrosqueezing transform and detrended fluctuation analysis. *IEEE Access, 8*, 12533–12541. https://doi.org/10.1109/ACCESS.2020.2965744

Liu, P., & Zhang, W. (2019). A fault diagnosis intelligent algorithm based on improved BP neural network. *International Journal of Pattern Recognition & Artificial Intelligence, 33*(9), 50–56. https://doi.org/10.1142/S0218001419590286

Miao, D. (2019). Research on fault diagnosis of high-voltage circuit breaker based on support vector machine. *International Journal of Pattern Recognition & Artificial Intelligence, 33*(6), 603–608. https://doi.org/10.1142/S0218001419590195

Petrenya, Y. K. (2019). Development of gas turbine energy technologies in Russia. *Herald of the Russian Academy of Sciences, 89*(2), 101–104. https://doi.org/10.1134/S1019331619020163

Sabour, S., Frosst, N., & Hinton, G. E. (2017). 31st Conference on Neural Information Processing Systems. Dec 4, 2017 – Dec 9, 2017. Dynamic routing between capsules. Neural Information Processing Systems: NIPS Proceeding, Long Beach, CA, USA (pp. 3856–3866), arXiv:1710.09829v2.

Shao, S. Y., Yan, R. Q., Lu, Y. D., & Gao, R. X. (2019). DCNN-based multi-signal induction motor fault diagnosis. *IEEE Transactions on Instrumentation and Measurement, 69*(6), 2658–2669. https://doi.org/10.1109/TIM.2019.2925247

Wang, X. G., Liu, X. Y., & Li, Y. (2019). An incremental model transfer method for complex process fault diagnosis. *IEEE/CAA Journal of Automatica Sinica, 6*(5), 1268–1280. https://doi.org/10.1109/JAS.2019.1911618

Yan, C., Li, M., & Liu, W. (2019). Transformer fault diagnosis based on BP-Adaboost and PNN series connection. *Mathematical Problems in Engineering, 2019*(10), 1–10. https://doi.org/10.1155/2019/1019845

Zhang, Z., Li, S., Lu, J., Wang, J., & Jiang, X. (2020). A novel intelligent fault diagnosis method based on fast intrinsic component filtering and pseudo-normalization. *Mechanical Systems and Signal Processing, 145*, article 106923. https://doi.org/10.1016/j.ymssp.2020.106923

Zhao, B., Zhang, X. M., Hai, L., & Yang, Z. B. (2020). Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions. *Knowledge-Based Systems, 199*, article 105971. https://doi.org/10.1016/j.knosys.2020.105971

Zhou, J. D., Yao, Q. B., Wu, H., Ma, S. X., & Zhang, H. S. (2020). Fault diagnosis of gas turbine based on partly interpretable convolutional neural networks. *Energy, 200*, article 117467. https://doi.org/10.1016/j.energy.2020.117467