Research Article

Journal of Robotics and Automation Research

Gearbox Fault Feature Extraction and Diagnosis

Chuanhao Wang[1], Yongjian Sun[2*], Xiaohong Wang[3]

[1]*School of Electrical Engineering, University of Jinan, Jinan, Shandong, China*

[2]*School of Electrical Engineering, University of Jinan, Jinan, Shandong, China*

[3]*School of Electrical Engineering, University of Jinan, Jinan, Shandong, China*

[*]Corresponding author
Yongjian Sun, School of Electrical Engineering, University of Jinan, Jinan, Shandong, China.

Submitted: 19 July 2021; **Accepted**: 27 July 2021; **Published**: 10 Aug 2021

Citation: *Chuanhao Wang, Yongjian Sun, Xiaohong Wang. (2021). Gearbox Fault Feature Extraction and Diagnosis. J Robot Auto Res, 2(2), 14-18.*

Abstract
Gearboxes, as an essential connection in mechanical equipment as well as a universal component for transmitting power, have a wide range of applications in modern industrial development. Gear systems can be damaged to different degrees during operation, such as gear breakage, pitting, wear, etc., all of which generate periodic impulse impact force and produce the modulation phenomenon of vibration signal, and extract modulation information from the signal and analyze its intensity and frequency to determine the degree and parts of the parts damage. This paper mainly discusses the analysis and processing of the vibration data of gears by MATLAB, based on the snowflake diagram obtained by analyzing the SDP (Symmetrized Dot Pattern) method, the images are filtered and binarized separately, and the images obtained are compared with the similarity with the standard images, thus greatly leveling the fault identification purpose.

Keywords: Gearbox, Feature Extraction, Fault Diagnosis, Image Processing

Introduction
In the construction of the national economy, with the accelerated development of modern industry, modern industry is developing in the direction of automation and large- scale, modern equipment also tends to be more and more automated, high-speed and lighter development, which has a great impact on the market, leading to higher and higher indicators in all aspects of the production system, while make in large machinery and equipment more and more perfect, slowly gears have become the current stage of industrial production The necessary parts. With the continuous development of the national economy, the rapid development of the global economy, increasingly at the expense of the destruction of the ecological environment of the Earth in exchange for the economic development of the countries of the world [1]. The development of new energy sources has become a top priority. New energy sources, mainly solar, wind and tidal energy, have become the main clean energy sources, and wind energy as a clean energy source is a force to be reckoned with. Wind turbine gearbox is the most important core component in wind power generation, and the failure of gearbox is the most important failure in the whole process of wind power generation [2]. According to the statistics of the relevant departments, the proportion of functional failure of the gear itself accounts for 60% of the gearbox failure parts in the whole automation industry, which shows that the gear failure is the most important cause of mechanical failure leading to industrial shutdown. Not only the fan gear box, in the country's industrial development gear has always played an extremely important identity, the unique design of the gear is destined to be an indispensable part of industrial production, but also because of its design led to the gear is the most prone to failure of mechanical equipment parts. A small transmission gear untimely failure, or will probably cause the entire automated machine downtime, and even then lead to the entire assembly line shutdown, or even the entire automated work shop stop production, serious will bring unnecessary casualties, to the factory, the country caused unnecessary economic losses, and even bring unnecessary casualties. In general, the research of fault diagnosis technology of gearbox has been widely concerned and studied all over the world.

In order to meet the needs of the continuous development of China's machinery and equipment in the direction of automation and high reliability, as well as the transformation to the direction of high performance and high efficiency, it is imperative to improve the accuracy and rapidity of the gear- box fault diagnosis and diagnosis process. With the continuous development of science and technology in the new era, the popularity of various kinds of machinery and equipment and the emergence of various new technologies, various fault diagnosis technologies have achieved rapid development, and the fault diagnosis technology for gearboxes has also made great progress [3]. In the construction of China's national economy, modern industrial development is gradually developing in the direction of automation, mass production and high-speed production, which has a very positive effect on reducing production costs and improving production efficiency of factories. Relative to mechanical equipment, a gear box gear equipment can directly determine the stability of the equipment, affecting the safety of the equipment, because the gear box as the main components of large mechanical equipment, as long as once the gear box failure caused by the mechanical equipment to stop running, will cause great economic losses [4]. Therefore, early detection of faults is imperative, rather than waiting for faults to occur and then dealing with them, finding out the cause of faults and solving the mistake factor in improving the operational efficiency of the system.

The fault diagnosis technology of gearboxes has great significance, and its existence is not only to reduce the failure rate of machine equipment, but also to reduce the maintenance cost of machine equipment, and the maintenance time is also greatly reduced, and it is greatly to improve the operation time of machinery and equipment. According to Japanese statistics, after the adoption of fault diagnosis technology, the accident rate of machinery and equipment has been reduced by 75%, and the maintenance cost has been reduced by 25%-50% [5]. Therefore, from the perspective of the development of the new era of the 21st world, whether from the perspective of people's industrial production safety, or from the economic interests of domestic and foreign enterprises, or from the international development trend, the 21st century for us to establish a perfect gearbox fault detection system and the establishment of a complete diagnostic system has an extremely far-reaching significance.

The main content of this paper is the study of gearbox fault feature extraction and diagnosis method, which is mainly, aimed at the analysis and study of the fault states of gears in gearboxes under various operating conditions [6]. This paper focuses on the study of the characteristics of the gearbox by the basic structural basis of the gearbox and the study of the common faults that exist in the gearbox [7]. The method chosen for this project is the SD method, which is to transform the abstract vibration signal of the Gearin to a visualized snowflake-like image, and to convert the signal problem, into an image problem for analysis, so as to achieve the purpose of fault signal detection.

Scheme Design

Basic Structure of Gearbox

The gearbox is one of the most important components of wind turbines. The double-fed wind turbine gearbox contains the most important shaft and gears, as well as the cage that fixes the entire gearbox state. Through the coordination between these components, the wind energy in nature drives the rotation of the fan blades, and the rotation of the fan blades drives the rotation of the motor through the gear structure, which can convert the wind energy into mechanical energy, and then the motor generates electricity to pro- duce electricity, so as to do the need of power generation. The shaft and gear in the gearbox are the two most important parts of the fan gearbox. The inner ring of the gearbox is connected to the whole shell, and the inner ring is the center of the whole planetary wheel and the core of the whole gearbox, and various external loads need to be borne by it, and what I am going to carry out in this design is the study of the gear failure in the gearbox [8]. Due to the different design process, the structure of the fan gear box will have different deviations, but their design principles are the same, are mechanical energy from the blade to the first level of planetary gear, and then to the second level of fixed shaft wheel system, and finally to the generator. As shown in figure 1 that is the structure of the gear box, its main working principle is: the natural wind energy to bring the blade a certain speed, the blade will be this mechanical energy transfer to the input shaft, the input shaft received mechanical energy impact will produce a certain torque continue to backward transmission, and then through the planetary frame drive power transmission to the big gear, then further by the big gear to the gear shaft, and then with the big gear through the output axis The torque is output by the large gear through the generator, thus converting the mechanical energy into electrical energy through the generator [9].

The basic Principle of SDP Method

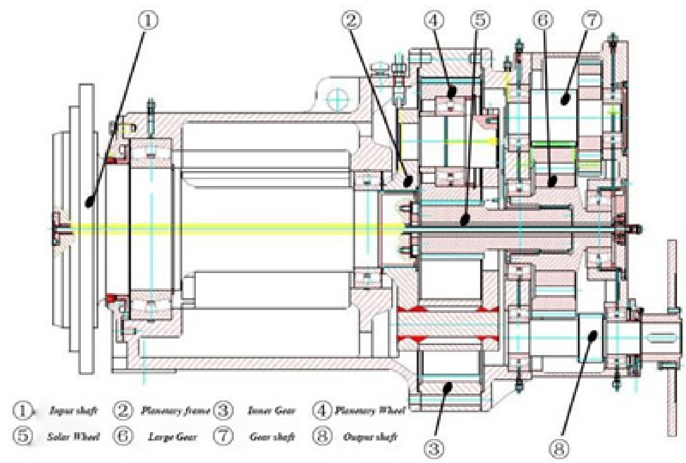

Figure 1: Wind turbine gearbox structure diagram

In the discrete data of vibration signal, the vibration amplitude of moment is xn and the vibration amplitude of moment n+l is xn. By transforming these data through the SDP formula, asymmetric image in the polar coordinate space $p[\gamma(n),\theta(n),\varphi(n)]$ is obtained. The schematic diagram of SDP (Symmetrized Dot Pattern) formula is shown in figure 2. Unlike the time domain frequency domain analysis described above, the SDP can visualize the vibration signal as a mirror symmetric image, which is more obvious, more intuitive and simpler. The SDP equation is as follows.

$$\gamma(n) = \frac{x_n - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

$$\theta(n) = \theta + \frac{x_{n+l} - x_{\min}}{x_{\max} - x_{\min}} g \quad (2)$$

$$\varphi(n) = \theta - \frac{x_{n+l} - x_{\min}}{x_{\max} - x_{\min}} g \quad (3)$$

In the above three equations, $\gamma(n)$ represents the radius in the polar coordinate system; $\theta(n)$ represents the angle of rotation of the polar coordinates counter clock wise along the initial line; $\varphi(n)$ while is the angle of rotation of the polar coordinates clockwise along the initial line. The specific calculation formula is shown above. In its calculation, xminineq.1is the minimum value in the discrete sequence of vibration signal; xmax is the maximum value in the discrete sequence of vibration signal; the vibration amplitude of moment nine q.2 and eq.3 is xn, the vibration amplitude of moment $n+l$ is x_n, g is the angle amplification factor, θ is the rotation angle of the mirror plane of symmetry, and l is the time interval of parameter.

Cosine similarity

In a multidimensional space, two vectors A and B, we can regard these two vectors A and B as the extensions of two different directions past the origin in a multi-dimensional space, so an angle will be formed between the two vectors, and the size of this angle determines the similarity of vectors A and B. When there is no angle between vectors A and B, it means that these two vectors are similar. When there is no angle between vector A and vector B, it means that these two vectors are coincident, and we can also say that these two vectors are completely similar; when the angle between vector A and vector B is 90 degrees, it means that these two vectors are two vectors completely dissimilar; when the angle between two vectors A and B is 180 degrees, the two vectors are in opposite directions. Therefore, we can clarify the degree of similarity between two vectors by judging the size of the angle between two vectors A and B.

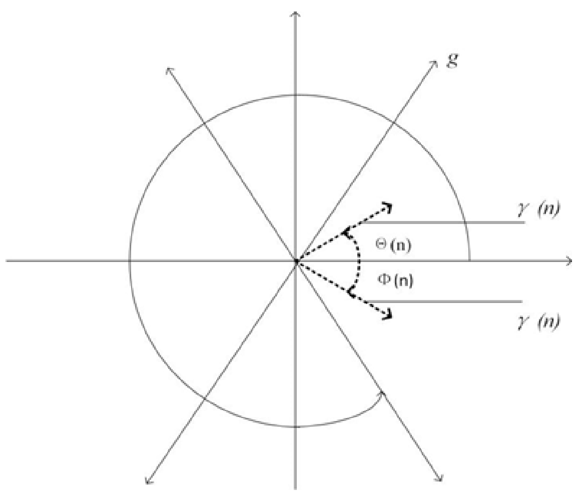

Figure 2: SDP schematic diagram

The smaller the angle between two vectors, the higher the degree of similarity. First, we obtain the gray levels corresponding to each position of the test and standard images respectively, count the number of appearances of 256 gray levels and represent them in a histogram, and use the histogram to represent the number of 256 gray levels, then we divide the 256 gray levels in the represented histogram into 64 zones on average, which means that each zones 4 consecutive gray levels, and perform the summation operation for each zone of 4 The summation operation of the 4 gray levels of each zone can obtain one element, one element of the vector. Finally, the cosine similarity can be calculated by vectors A and B to determine the degree of similarity between the two images. The formula is in equation 4.

$$\cos\theta = \frac{\sum_{i=1}^{n}\left(A_i \times B_i\right)}{\sqrt{\sum_{i=1}^{n}\left(A_i\right)^2} \times \sqrt{\sum_{i=1}^{n}\left(B_i\right)^2}} \quad 4$$

Simulation Experiment

Image Processing

Image processing is an important prerequisite for image feature extraction and a key part of image feature extraction. The good or bad image processing is directly related to the accuracy of subsequent image fault feature extraction and the correct rate of classification. In this paper, the obtained image is grayscale processed then median filtered, and the image is binarized and averaged to obtain the standard image for image feature extraction.

Drawing Images by SDP Method

Firstly, the SDP method is used to transform the image of 10 groups of 600 data representing 6 working conditions respectively, and through the SDP formula, the snowflake image representing 6 working conditions of the gear is derived, as shown in figure 3, which can be very intuitive to see the difference of the snowflake image in 6, and the six working conditions of the gear reflected by the snowflake image are also completely different.

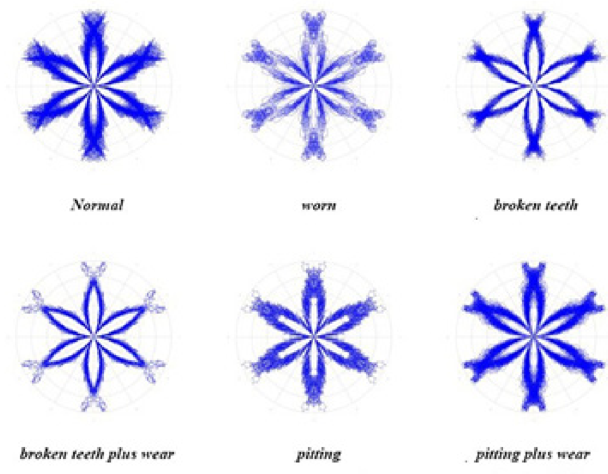

Figure 3: Six types of working conditions polar graph

Binarization Process

Image binarization is also one of the most common and the most convenient means of image processing in image preprocessing. In this paper, the image is processed by selecting a suitable threshold value so that only two gray values of 0 or 1 exist in the snowflake image, and in this way the snowflake image is made to appear only black or white. This can make the data in the snow flake image to achieve the maximum degree of reduction, is the whole snowflake image becomes simpler, and then make the image features become more obvious, more conducive to the subsequent work in the image recognition. Through this series of operations, we successfully grayed out the denoised snow flake (4) image, and then binarized it to generate a binarized image. As shown in figure 4, the binarized snowflake image is shown for the normal gear condition and other fault conditions.

Median Filter Denoising

The snowflake images generated with the original sequence data have coarser images, which can lead to difficulties in later recognition, and in order to reduce the interference of noise on later snowflake image recognition and get high-quality snowflake images, we denoise the obtained original snowflake images. Compared with mean filter de- noising, median filter denoising can more effectively protect the six corners of the snowflake image, which are the sharper edge parts of the snowflake image. Therefore, in the exposition of this paper, median filter denoising is selected for snowflake images. Further, by selecting and cropping the binary matrix of the snowflake image, the white border around the image is removed to retain the most valuable area of the lower snowflake image as much as possible, and a size standard suitable for various working conditions is selected to unify the size of each snowflake image, which is more conducive to the recognition of each image at a later stage. As in figure 5.

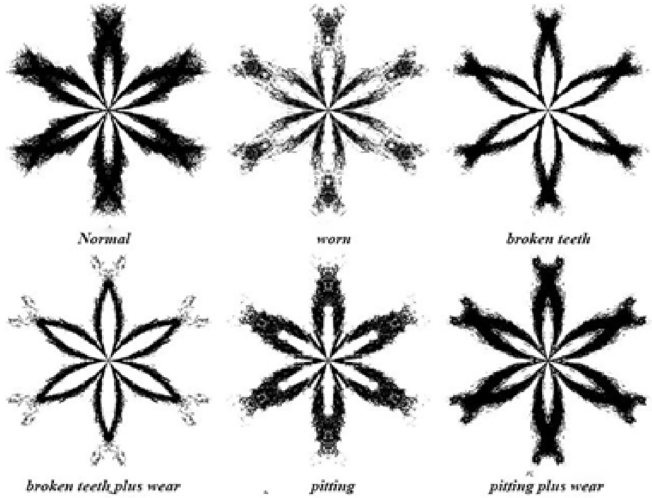

Figure 5: Median filter denoising

Fault Type Identification

After the standard snowflake image is output, 6000 data are called, 10 groups of 600 data are generated from the called data, the polar plot is established by using the SDP method, the images are filtered, binarized and localized in turn, and the 10 polar images are averaged immediately afterwards to obtain the test image. The test images generated by the program are compared with the standard snowflake images of the six working conditions respectively for similarity, and the column with the highest similarity is output for the statistical accuracy later.

The data of 6 working conditions are called, and each working condition generates 10 groups of 600 data volumes, and the polar coordinate map is established by using SDP method, and the resulting polar coordinate map is subjected to median filtering and denoising, binarization of polar coordinate map, and localization of images in turn, and 60 images are obtained for image averaging respectively, and the standard images of 6 working conditions are finally obtained. As shown in figure 6.

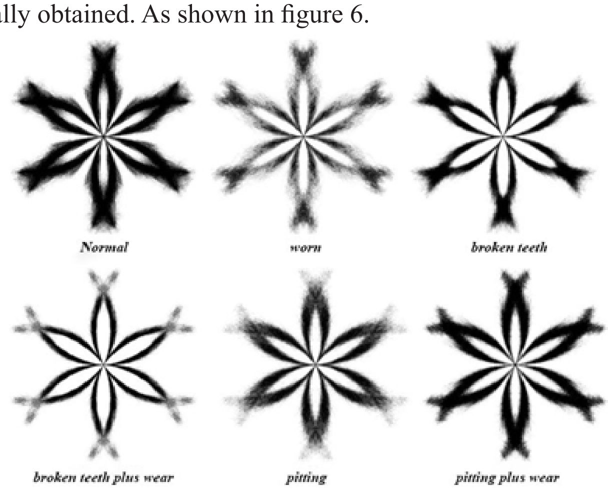

Figure 6: Standard images of six working conditions

After the test criteria are obtained, the overall program design is carried out. The data of the six working conditions are put into the six columns of the table in order, and one working condition is selected randomly, and the number of columns corresponding to the selected working condition is recorded and stored in the cell array s1. The 6,000 data are called, and the called data are generated into 10 groups of 600 data, and the SDP method is used to create a polar plot to convert the signal data into a snowflake image with symmetry about the mirror image. Next, the polar coordinate snowflake image s are sequentially median filtered to prepare the foundation for the subsequent is image analysis. The binarization selected a suitable threshold value, so that the snowflake image exists only two gray values of 0 or 1, by this way the snowflake image only presents two effects of black or white. Then the localization process is carried out, by selecting and cropping the binary matrix of the snowflake image, removing the white border around the image, retaining the most valuable area of the snowflake image as much as possible, and selecting the size standard suitable for various working conditions, unifying the size of each snowflake image, which is more conducive to the recognition of each image later. The final 10 polar coordinate images obtained will be averaged to obtain the test image, and the six standard images will be use d as the standard for subsequent data processing. Then the test images are compared with the standard images of the six working conditions in turn, and the similarity with the standard images of the six working conditions is compared, and the number of columns of the working conditions corresponding to the maximum similarity is recorded, and the corresponding columns are stored in the cell array s2, and then this is cycled 500 times, and finally the array s2obtained from the test is subtracted from the randomly selected array s1, and the number of zeroes is the number of correct matches. The number of zeroes is the number of correct matches, and the accuracy rate is calculated. The accuracy rate is calculated. Fault test discriminating process is as shown in figure 7.

- The data of 6 working conditions are stored in order in 6 lists.
- Randomly select a working condition and store the data in the array S1.
- 6000 data were selected and divided equally into 10 groups, and polar plots were created using the sdp method.
- Generated images for image processing (gray scale processing, denoising, binarization).
- Compare the test image with the standard image for similarity.
- Record the column number of working conditions corresponding to the maximum similarity and save it to S2, then loop 500 times
- Subtract the data S1 from S2, count the number of zeros and calculate the correctrate

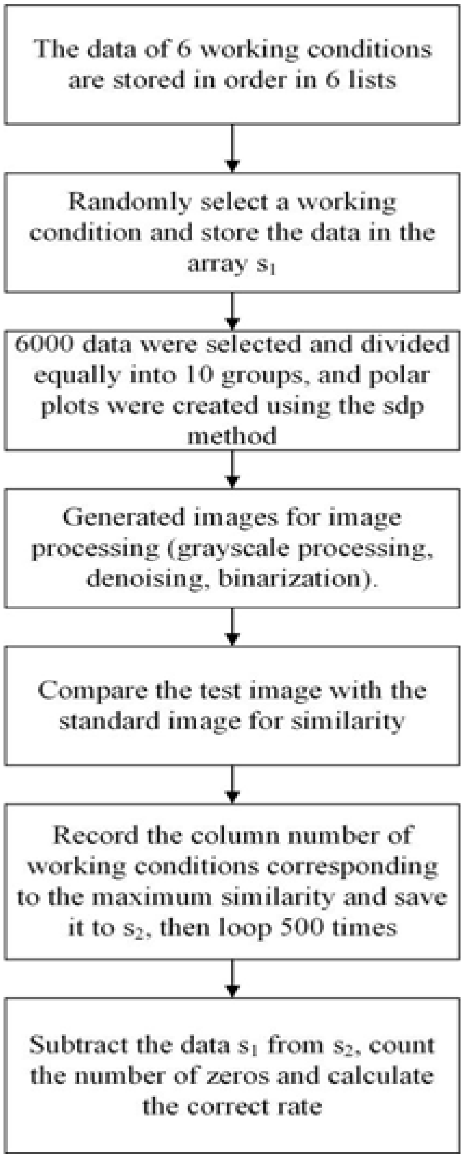

Figure 7: Fault test discriminating process

Test Results

After their own multiple measurements, only one group of ten sets of data accuracy did not reach 100%, but the test results in the presence of object error accuracy also reached 99.8%, while the remaining nine groups of test results are correct 100%, therefore, this test can be seen as the identification of randomly selected data and the selected data where the working conditions data are exactly the same, the correct rate reached 100%, test The results are shown in table 1.

Conclusion

The main purpose of the analysis of gear faults is to classify and identify these fault patterns by analyzing the characteristic signals of gear faults, comparing the signal characteristics in the absence of faults, and finally testing the effectiveness of the algorithm masa whole using the data tested. The above table we can easily see that the correct rate of image fault classification reached 100%, this correct rate fully reflects the feasibility of the method.

Table 1: Test results

Group	Noun	Correct Rate
group 1	500	100%
group 2	500	100%
group 3	500	99.8%
group 4	500	100%
group 5	500	100%
group 6	500	100%
group 7	500	100%
group 8	500	100%
group 9	500	100%
group 10	500	100%
mean	500	100%

References

1. Kim, D. K., Rho, K. H., Na, Y., & Kim, M. (2021). Evaluation of energy storage technologies for efficient usage of wind power in the far-eastern region: A techno-economic analysis. Journal of Energy Storage, 39, 102595.
2. Dameshghi, A., & Refan, M. H. (2019). Wind turbine gearbox condition monitoring and fault diagnosis based on multi-sensor information fusion of SCADA and DSER-PSO-WRVM method. International Journal of Modelling and Simulation, 39(1), 48-72.
3. Wang, P., Song, L., Hao, Y., Wang, H., Li, S., & Cui, L. (2021). A light intelligent diagnosis model based on improved Online Dictionary Learning sample-making and simplified convolutional neural network. Measurement, 109813.
4. Liu, R., Zuo, H., Sun, J., & Wang, L. (2017). Electrostatic monitoring of wind turbine gearbox on oil-lubricated system. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 231(19), 3649-3664.
5. Zhang, Q., Wang, Y., Lin, W., Luo, Y., & Wu, X. (2020). Contact Mechanics Analysis and Optimization of Shape Modification of Electric Vehicle Gearbox. Jordan Journal of Mechanical & Industrial Engineering, 14(1).
6. Durbhaka, G. K., Selvaraj, B., Mittal, M., Saba, T., Rehman, A., & Goyal, L. M. (2021). Swarm-LSTM: Condition monitoring of gearbox fault diagnosis based on hybrid LSTM deep neural network optimized by swarm intelligence algorithms. CMC-Comput. Mater. Continua, 66(2), 2041-2059.
7. Schmidt, S., Heyns, P. S., & Gryllias, K. C. (2021). An informative frequency band identification framework for gearbox fault diagnosis under time-varying operating conditions. Mechanical Systems and Signal Processing, 158, 107771.
8. Zhu, X., Zhao, J., Hou, D., & Han, Z. (2019). An SDP characteristic information fusion-based CNN vibration fault diagnosis method. Shock and Vibration, 2019.
9. Zhu, X., Hou, D., Zhou, P., Han, Z., Yuan, Y., Zhou, W., & Yin, Q. (2019). Rotor fault diagnosis using a convolutional neural network with symmetrized dot pattern images. Measurement, 138, 526-535.