

# **Classification of Bearing Fault Location and Severity Using Cascade ANNs with Statistical and Spectral Features**

**Punyisa Kuendee and Udom Janjarassuk**

Department of Industrial Engineering, Faculty of Engineering  
King Mongkut's Institute of Technology Ladkrabang  
Bangkok, Thailand  
[poonpoon99@hotmail.com](mailto:poonpoon99@hotmail.com), [kjudom@gmail.com](mailto:kjudom@gmail.com)

## **Abstract**

Advance in machine learning techniques for machine condition monitoring is now an active field of interest in modern industry for predicting the future health of machines. In this paper, Artificial Neural Networks (ANNs) technique is investigated by using the existing datasets of ball bearing fault to classify the fault location and severity. To enhance the accuracy, the utilization of statistical features and spectral features are used in pre-processing with the proposed method of cascade ANNs. Motor vibration data is collected by using accelerometers attached to the housing in the drive end and fan end. To diagnose the faults in the rolling bearing, the vibration data is pre-processing by using Fast Fourier Transform (FFT) and other statistical features such as standard deviation, skewness and kurtosis etc. Then the feature data is fed into cascade ANN\_1 and ANN\_2 to classify the fault location and severity. The objective of this research is to solve the problem of poor features such as slippage frequency component in FFT response and inappropriate statistical feature datasets to help increase in accuracy and efficiency. The results show that the proposed method gives more accuracy when the poor features are existed and guarantee the accuracy near to 100%.

## **Keywords**

Machine learning, Ball bearing, Vibration data, Condition monitoring and Fault classification.

## **1. Introduction**

Machine fault detection and diagnosis play an important role in real world applications for predicting or monitoring the health condition of machines in modern industry. Generally, machine faults can be occurred in anytime of operation. The chance of faults in bearings is the most frequently occurrence by approximately 51% of all failures (Kang et al. 2015). The bearing faults detection and diagnosis is also a challenging topic for many researchers to study the pattern of the fault signals. Many researches use vibration signal and acoustic emission (AE) collected from accelerometer sensors and microphone sensors respectively attached in the housing of the machine. Most researches use vibration signal rather than acoustic emission. However, when the machine run in low speed the vibration signal seems to be ineffective because the small vibration fault generates signal in small magnitude and hard to classify by conventional methods (Widodo et al. 2009). Thus, the use of AE as input to the next classification process for machine faults running at low speed is mentioned in (Islam et al., 2017).

Feature extraction process are classified into time domain (TD) method and frequency domain (FD) method. Features extraction in TD as in (M. M. Tahir et al. 2017) used 10 statistical features to reduce undesired impact of fluctuations on the vibration signal. In addition, the use of Laplace wavelet analysis for TD vibration signal for feature extraction is introduced in (Al-Raheem and Abdul-Karem 2010). A study on three types of ANNs namely, multilayer perceptron (MLP), radial basis function (RBF) network, and probabilistic neural network (PNN) along with the selection of input features, are optimized using genetic algorithms (GA) for the TD vibration signal is presented in (B. Samanta et al. 2004). Furthermore, simple statistical features such as standard deviation, skewness, kurtosis etc. of the TD vibration signal segments along with peaks of the signal are used as features to the input of ANN and Support Vector Machine (SVM) with Discrete Wavelet Transform (DWT) prior to feature extraction process (Commander Sunil Tyagi 2008) and (O.R. Seryasat et al. 2014).

For features extraction in FD, the FD vibration signal is pre-processed by FFT. Milind Natu (2013) uses FFT and wavelet analysis to detect bearing faults and the author mentioned that this wavelet analysis or advance methods can be used for depth analysis of the signal for the bearing fault problems. In another field of research, Anuradha Abewardana and Upul Sonnadara (2012) learned the bird call (or sound) types of 10 species of Sri Lankan resident birds by using FFT and feed-forward neural network to categorize a selected set of Sri Lankan bird. However, the use of FFT may have a problem because the rolling elements experience some slippage as they enter and leave the bearing load zone. Therefore, the occurrence of the impacts never reproduces exactly at the same position from one cycle to another (Al-Raheem and Abdul-Karem 2010).

For time-frequency domain (T-FD) or other hybrid method, Islam et al. (2017) use hybrid features extracted from acoustic emissions and a Bayesian inference-based one-against-all support vector machine (Bayesian OAASVM) for multi-class classification. The wavelet packet decomposition method (D.H. Pandya et al. 2012) is used for time-frequency feature extraction from the vibration signal. V.Hariharan and PSS. Srinivasan (2009) presents a new approach using time-frequency vibration signal for feature extraction and use RBFN and PNN to classify the bearing defects. Yuan Xie and Tao Zhang (2017) present the feature extraction algorithm based on empirical mode decomposition (EMD), FFT, statistical feature and convolutional neural network (CNN) techniques and use SVM/Softmax in training to classify faults. Quang Hung Do (2017) proposes an ANN model along with Gravitational Search Algorithm (GSA) and a Back-Propagation (BP) Algorithm - Gradient Descent with Momentum (GDM). The hybrid algorithm of improved GSA and GDM is then utilized to optimize the weights between layers and biases of the neural network. Ran Zhang et al. (2017) proposes a fault diagnosis model based on Deep Neural Networks (DNN). The model can directly recognize raw time series sensor data without feature selection and signal processing. It also takes advantage of the temporal coherence of the data. Muhammad Sohaib et al. (2017) presents a two-layered bearing fault diagnosis scheme for the identification of fault pattern and crack size for a given fault type. A hybrid feature pool is used in combination with sparse stacked autoencoder (SAE)-based deep neural networks (DNNs) to perform effective diagnosis of bearing faults of multiple severities.

As reviews mentioned above, we see that the researches play attention to features extraction and learning algorithms for classification or even optimization of parameters in processing signal in TD, FD and T-FD. The pros and cons of those methods are considered when we use them with big data which have noise, low level of signal and slippage frequency. Hence, this let to propose the use of statistical and spectral analysis for features extraction to the input of ANN and we present the use of cascade neural networks in classification process. This may increase accuracy in classification and reduce computational time with compromised easy-hard methodology. The example of normal vibration signal and fault vibration signal of ball bearing is shown in Figure 1. From the naked eyes, we can see the normal condition in Figure 1 a) and found that the normal vibration amplitudes are not exceeded value of 1.0. For fault condition, Figure 1 b) shows the peak amplitudes are exceeded the value of 1.0.

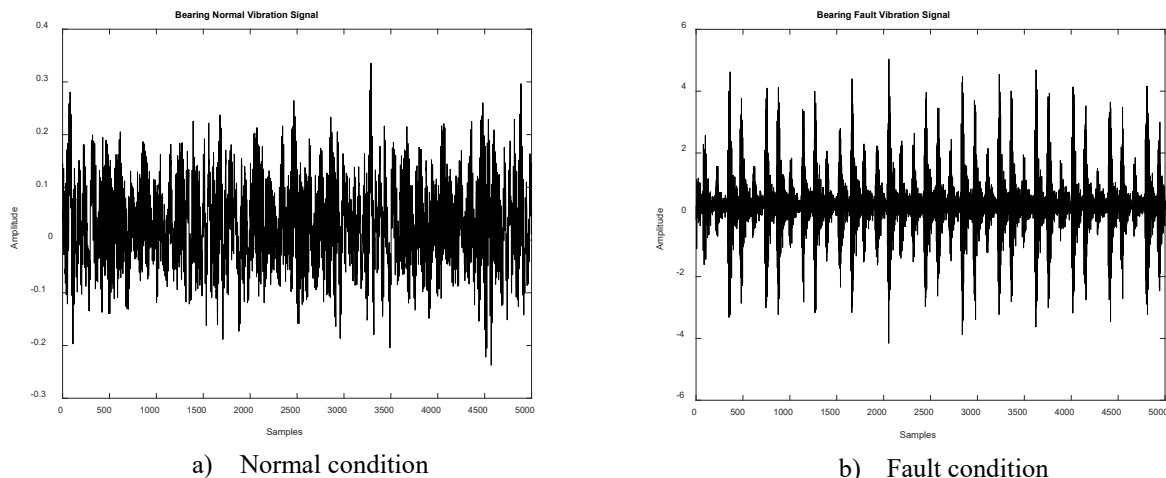


Figure 1. TD Vibration signal of ball bearing

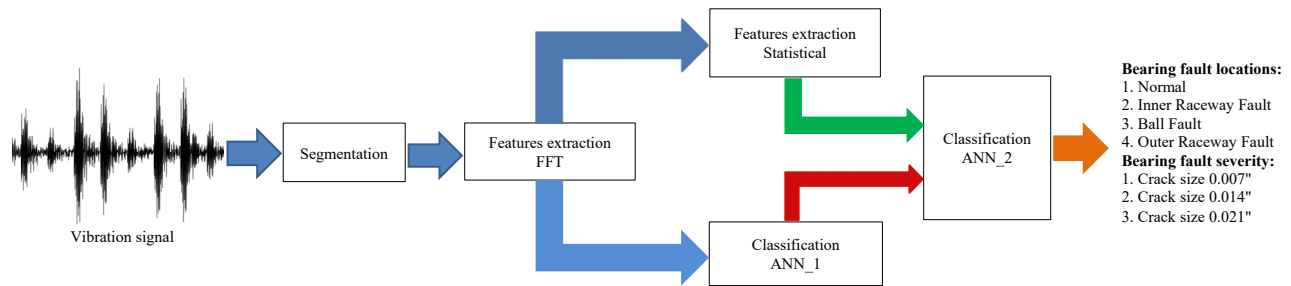


Figure 2. The proposed cascade neural networks scheme for bearing fault detection

However, this time domain signal may have a problem in classification due to time shift if we split the signal in wrong segmentation period. The vibration data used in this study are taken from Case Western Reserve University Bearing Data Center Website and granted permission from Professor Dr. Kenneth A. Loparo, Faculty director, ISSACS: Institute for Smart, Secure and Connected Systems, Co-Academic Director, IoT Collaborative, Department of Electrical Engineering and Computer Science, Case Western Reserve University.

## 2. Methodology

The scheme of the proposed approach is shown in Figure 2. We use a hybrid approach between statistical and spectral analysis. First of all, the vibration signals from Case Western Reserve University Bearing Data Center Website (2018) are used and then split it or do segmentation. The segmentation process can reduce the size of input datasets to help in protection on computer memory shortage. The dataset of any cases has sampling frequency of 12,000 samples/second or 48,000 samples/second and the dataset has 120,000 – 480,000 samples/file.

For features extraction, after segmentation, the split signal is pre-processed by FFT for spectral features. FD features are put into 2 neural networks ANN\_1 and ANN\_2. FD statistical features are calculated from split signals with 5 statistical formula (see section 2.3) and feed to be the input candidates of ANN\_2. The output of ANN\_1 which are trained from FD input features are fed into the input of ANN\_2 also. Finally, first hybrid features (5 statistical + 4 spectral) are trained and classified by cascade ANN\_2 with the target data of bearing fault location and second hybrid features (5 statistical + 3 spectral) are mixed to be the input of ANN\_2 to classify the severity.

### 2.1 Data segmentation

Segmentation of vibration signal is the method to partition signal into several small pieces according to the significant data which can cover the overall signal or not. This was achieved by manual segmentation to the original vibration signal. The manual selection of samples is the important thing because the small signal must represent the big signal characteristics. In this research, we segment the signal to 12,000 samples/piece from original 120,000 samples because this segmented period can cover all frequency response and represents the original signal. Then, we do normalize the segmented signal because this make all vibration signals in percentage of maximum amplitude which help simplify the calculation, especially in training process of neural networks.

### 2.2 Spectral Features

In this research, we use spectral analysis by using FFT to the segmented signal for features extraction. FFT is an algorithm of computation the N-point Discrete Fourier Transform (DFT) with a computational complexity of  $O(N \log N)$ . An example of FFT to the vibration signal is shown in Figure 3. As showing in Figure 3 a), we notice that the normalized frequencies of the spectrums are mostly less than 0.4 in normal condition, however in Figure 3 b), the normalized frequencies of the spectrums are mostly greater than 0.4 in fault condition.

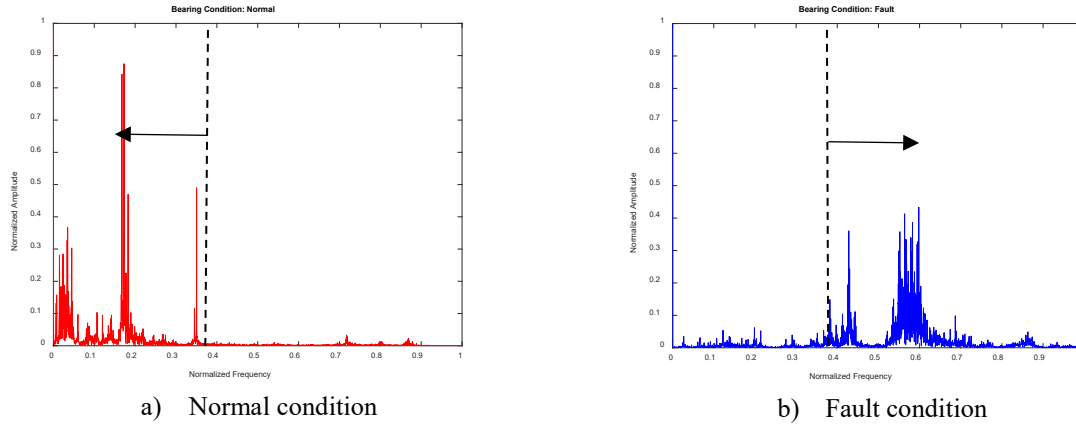


Figure 3. FFT of Vibration signal

## 2.3 Statistical Features

The statistical features of the vibration signal used in this paper are taken from (Muhammad Sohaib et al. 2017). The authors use 11 features to be included in the hybrid feature pool. These features are the root mean square value (RMS), kurtosis value (KV), square root of the magnitude (SRM), peak-to-peak value (PPV), standard deviation (SD), skewness value (SV), margin factor (MF), crest factor (CF), impulse factor (IF), kurtosis factor (KF), and mean value (MV).

In this research, we use only 5 statistical features because we need to know or study the effect of poor features that is the main objective of this research. The equations of these features are shown in equation (1) – (5). The benefit of using statistical features is that it can reduce the effect of slippage frequency problem in FD. This can be corrected by statistical mean or accumulative mean.

$$\text{Mean value (MV)} \quad \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (1)$$

$$\text{Standard deviation (SD)} \quad \sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (2)$$

$$\text{Root mean square (RMS)} \quad RMS = \sqrt{\left( \frac{1}{N} \sum_{i=1}^N x_i^2 \right)} \quad (3)$$

$$\text{Kurtosis value (KV)} \quad KV = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \bar{x}}{\sigma} \right)^4 \quad (4)$$

$$\text{Skewness value (SV)} \quad SV = \frac{1}{N} \sum_{i=1}^N \left( \frac{x_i - \bar{x}}{\sigma} \right)^3 \quad (5)$$

## 2.4 Data Normalization

Input features of the neural networks need to be normalized because the variation range in magnitude can cause the training results to be divergent. Normalization  $X$  is the ratio between input feature  $x$  to the maximum value of all features,  $\max(x)$  that is  $X = x / \max(x)$  and normalized range between  $[0.00 - 1.00]$ . All input features are essential to do in the same range to the ANNs model. This guarantee for optimization of weight and bias values to be convergent because normalization will eliminate the influence of one factor over another. In addition, the gradient descent with momentum GDM algorithm which used for backpropagation converges faster with normalized data than with un-normalized data.

## 2.5 Artificial Neural Network

Artificial Neural Network (ANN) is a feedforward neural network for function approximation and classification. ANN consists of input layer, hidden layers, and output layer. Node  $i$ , also called a neuron, in an ANN is shown in Figure 4. It includes a summer and a nonlinear activation function  $g$ .

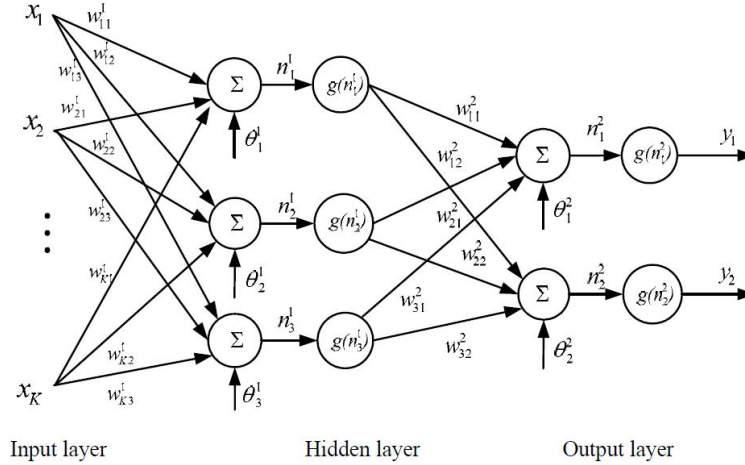


Figure 4. Artificial neural network

From Figure 4, the output  $y_i$ ,  $i=1,2$ , of the neural network becomes

$$y_i = g_2 \left( \sum_{j=1}^3 w_{ji}^2 g_1(n_j^1) + \theta_i^2 \right) = g_2 \left( \sum_{j=1}^3 w_{ji}^2 g_1 \left( \sum_{k=1}^K w_{kj}^1 x_k + \theta_j^1 \right) + \theta_i^2 \right) \quad (6)$$

From (6) the set of input features  $x_k$  in the input layer pass through neurons or nodes and then multiplied by the weights  $w_k$ . Then, the bias  $\theta_j$  is added to the result and activate the neurons  $n_j$  by activation function  $g_1$  in hidden layer and  $g_2$  in output layer. In this paper, we chose  $g_1$  as *sigmoid* and  $g_2$  as *softmax* (Mark Hudson Beale et al. 2017), and use scaled conjugate gradient backpropagation for learning and estimating the weights and biases of the final outputs  $y_1$  and  $y_2$ .

## 3. Experimental Data

The experimental data or seeded fault data is taken from Case Western Reserve University. The experiment setup is shown in Figure 5 and the main equipment consists of 2 hp electric motor (left), a torque transducer/encoder (center), a dynamometer (right), and control electronics (not shown).

Electro-discharge machine is used to seed faults with diameters of 0.007, 0.014, and 0.021 inches on the inner raceway (IR), outer raceway (OR), and Ball (BA) at the drive end bearings. Then, the vibration signals were collected by accelerometer attached to the housing of the drive end, fan end and ball at 12 o'clock with sampling rate of 12,000 Hz. The motor is taken mechanical loads from 0 to 3 hp causing the variation of motor speed ranging from 1722 to 1797 revolutions per minute (rpm).

In this study, we choose the datasets for location faults in crack size diameter of 0.014 inches only on the drive end with 3 fault locations of IR, OR, BA and 1 Normal with mechanical load of 0 hp. For fault severities, we choose 3 fault crack sizes of 0.007, 0.014, and 0.021 inches on IR only with mechanical load of 0 hp.

The vibration data have 120,000 samples with sampling rate 12,000 Hz, so we split it to 10 pieces per set. Thus, each feature has 12,000 vibration element values. Thus, the location fault target matrix is a  $4 \times 10$  matrix and the severities target matrix is a  $3 \times 10$  matrix. The details of the bearing dataset are given in Table 1.

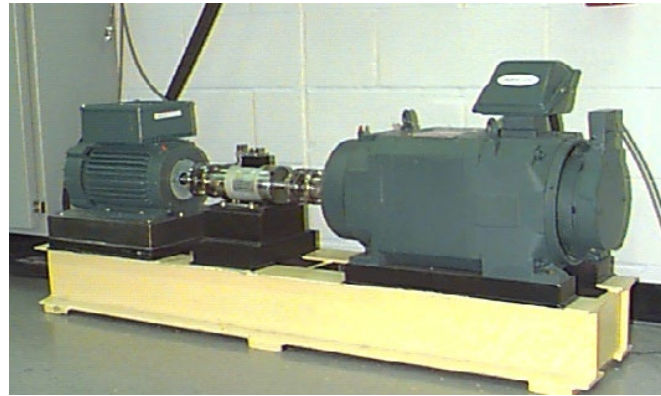


Figure 5. Case Western Reserve University's seeded fault bearing testbed

Table 1. Details of the bearings and dataset

Fault Type	Fault Location	Crack size (Inches)	All Samples	Training Samples	Test & Validation Samples	Sample Length	Accelerometer Position	Load (hp)
Normal	None	0	10	70%	15% (Test) +15% (Validation)	12,000	Drive End	0
Inner raceway	IR	0.007	10					
	IR	0.014	10					
	IR	0.021	10					
Outer raceway	OR	0.014	10					
Ball	BA	0.014	10					

## 4. Results and Discussion

The proposed cascade neural networks can be categorized into three research schemes namely pure statistical scheme, pure spectral scheme and hybrid-cascade ANNs scheme. The testing results of these schemes are compared in the percentage of output accuracy and the performance of network convergence. For this study, the neural network performance is presented in cross-entropy versus epochs as shown in Figure 6. The worst neural network performance is pure statistical scheme as shown in Figure 6 a). The better performance is pure spectral scheme as shown in Figure 6 b) and the best performance is hybrid-cascade ANNs scheme as shown in Figure 6 c).

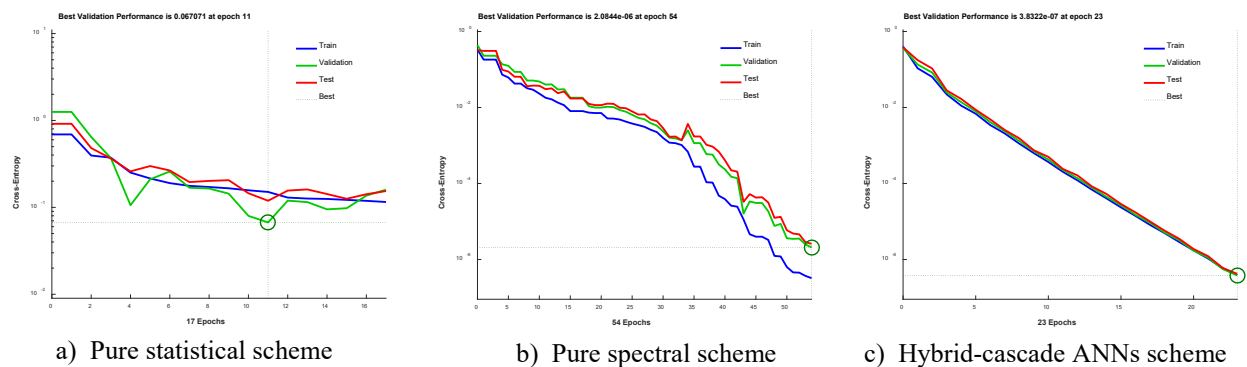


Figure 6. Neural network performance of three type features

#### 4.1 Fault location classification

The contribution of this proposed cascade ANNs is that it can help in dominance the poor features with the strong features and the data may not use many in training for the ANN step 2 (ANN\_2). However, the data used for training in step 1 (ANN\_1) may have more data but in appropriate amounts up to the neural network performance. The classification results of fault locations by these schemes are shown in Figure 7.

As see from Figure 7 a) the statistical approach is suited for normal condition which has average accuracy more than 80%. For spectral and hybrid features in Figure 7 b) and 7 c), the spectral have the average accuracy more than 99% and the cascade ANNs has all accuracy 100% in all cases (Normal, IR, BA and OR). As presented results, we may conclude that the FFT for spectral feature extraction is an effective method for classification the pattern of bearing fault in any locations. Especially, the cascade approach can help increasing the accuracy of the overall system.

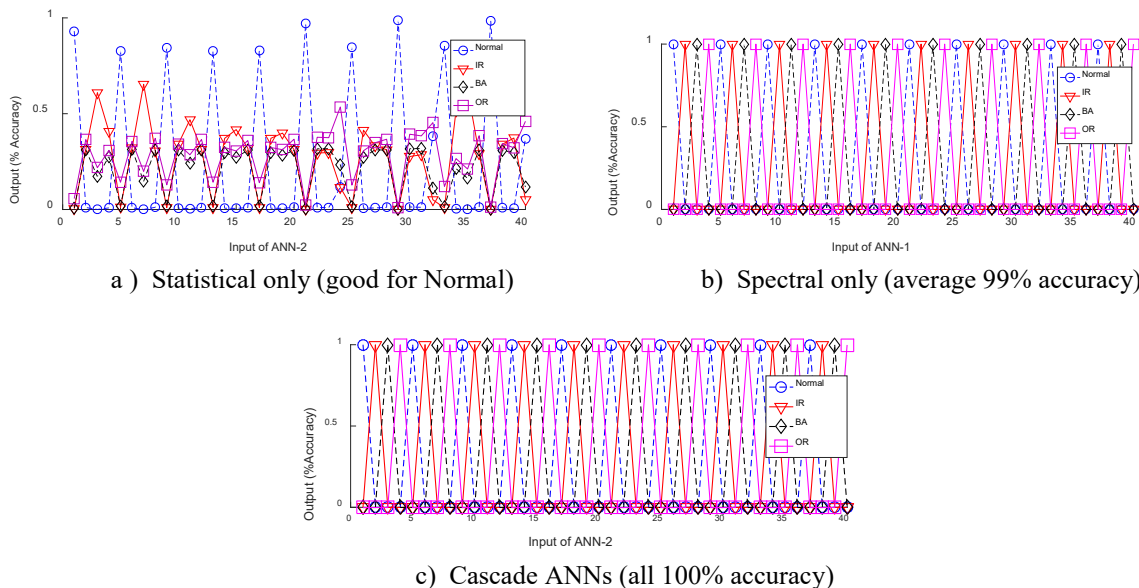


Figure 7. Percentage of output classification accuracy (Locations)

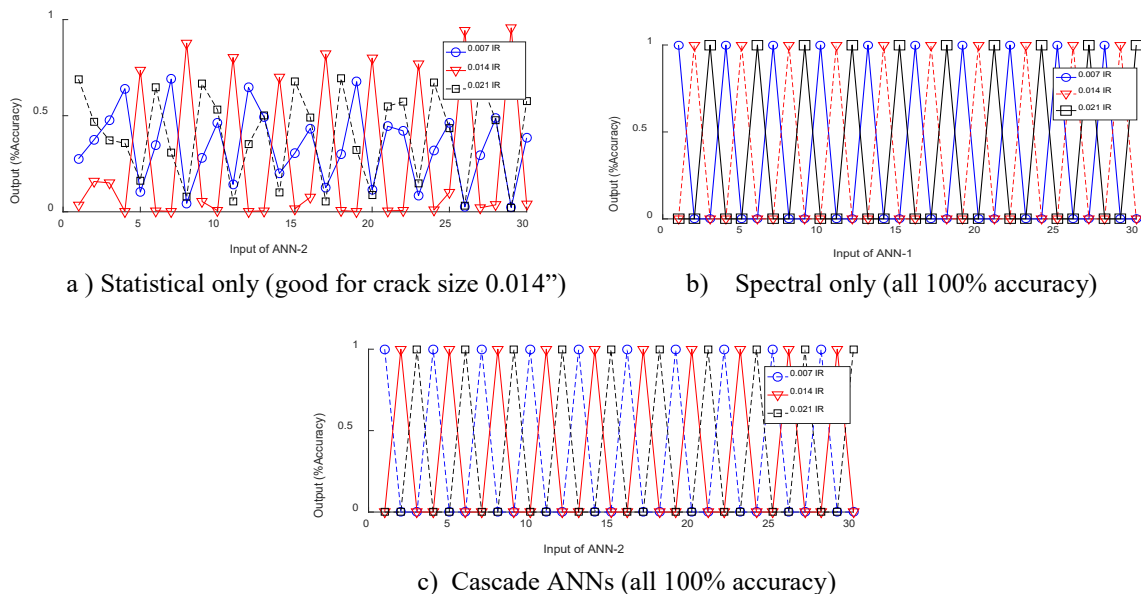


Figure 8. Percentage of output classification accuracy (Severities)



## 4.2 Fault severity classification

The vibration data of seeded fault element is brought to classify in different crack size diameters namely 0.007, 0.014 and 0.021 inches. We perform the same experiment as in fault location cases mentioned above. Figure 8 a), 8 b) and 8 c) shows the results of severity of three different scheme.

For severity test, Figure 8 a) shows that the statistical features can be used for classification the crack size 0.014 inches which its accuracy more than 70% but the accuracy results are unclear in crack size 0.007 and 0.021 inches. Better results, from the spectral features and hybrid features are shown in Figure 8 b) and 8 c) respectively with 100% accuracy. Notice that the statistical features cannot give a good result for localization and severity classification. This is because the spectrum frequency of the location cases: IR, BA and OR has the normalized amplitudes around 0.5 at the normalized frequency as shown in Figure 9 a), which is the same as severity cases shown in Figure 9 b). The statistical features cannot be classified effectively in this situation because the amplitude is nearly the same value.

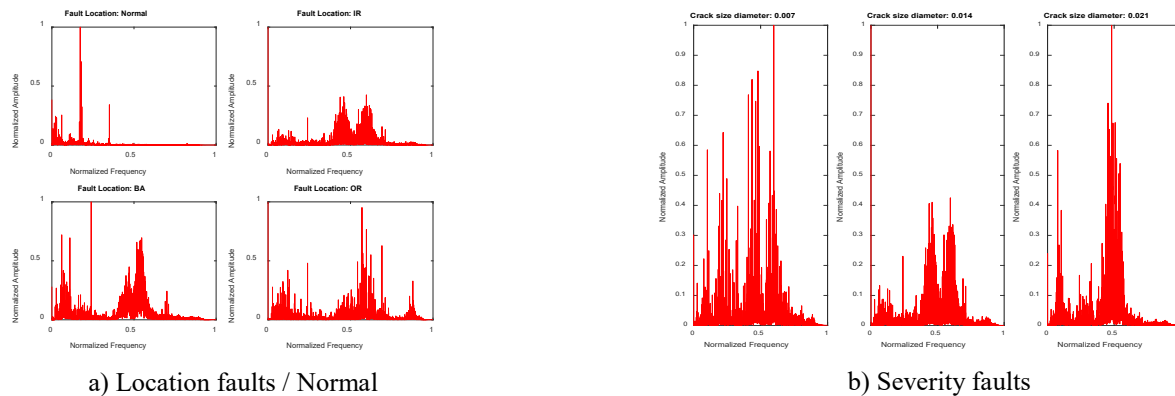


Figure 9. Comparative of spectral features

## 5. Conclusion

The proposed cascade neural networks for bearing faults detection and diagnosis can be used to classify the location and severity of faults effectively. The use of statistical features such as mean, standard deviation, root mean square, skewness and kurtosis to the FFT signal cannot give a good result in both location and severity classification because the FFT amplitudes are almost equal around normalized frequency at 0.5. The proposed cascade ANNs method has its contribution in the case where the systems have poor features and these features cannot be recognized effectively in the next classification process (ANN\_1). In such case, we use the cascade neural network 2 (ANN\_2) to help dominate with the good features. Especially, in the case where we do not know when the poor features exist. This proposed method guarantees the solution output of the neural networks to be accurate near 100%. In addition, it is a simplified method for general machine fault detection and diagnosis.

## 6. Future Works

In future works of this research, we may pay attention to the optimization algorithms of learning process, new statistical features or new feature extraction process such as parallel neural networks, convolution neural networks, etc. Furthermore, we may use sound signal (acoustic emission) in the future research instead of vibration signal as it is rarely studied in this fields.

## Acknowledgements

The researchers would like to express my very great appreciation to the Department of Industrial Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, Thailand. This work was supported by King Mongkut's Institute of Technology Ladkrabang. The researchers would like to thank Professor Dr. Kenneth A. Loparo in the courtesy of research data. Finally, the researcher wishes to thank my family for their support and encouragement throughout my life.



## References

- Kang, M., Kim, J., Kim, J. M., Tan, A. C. C., Kim, E. Y., and Choi, B. K., Reliable fault diagnosis for low-speed bearings using individually trained support vector machines with kernel discriminative feature analysis, *IEEE Trans. Power Electron*, vol. 30, pp. 2786–2797, 2015.
- Widodo, A., Kim, E. Y., Son, J.-D., Yang, B.-S., Tan, A. C. C., Gu, D.-S., Choi, B.-K., and Mathew, J., Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine, *Expert Syst. Appl.*, vol. 36, pp. 7252–7261, 2009.
- Islam et al., Reliable bearing fault diagnosis using Bayesian inference-based multi-class support vector machines, *The Journal of the Acoustical Society of America*, vol. 141, EL89, 2017.
- M. M. Tahir et al., Enhancing fault classification accuracy of ball bearing, *IEEE Open Access Journal*, February vol. 20, 2017.
- Al-Raheem and Abdul-Karem, Rolling bearing fault diagnostics using artificial neural networks based on Laplace wavelet analysis, *International Journal of Engineering, Science and Technology*, vol. 2, pp. 278-290, 2010.
- B. Samanta et al., Bearing Fault Detection Using Artificial Neural Networks and Genetic Algorithm, *EURASIP Journal on Applied Signal Processing*, vol. 3, pp. 366–377, 2004.
- O.R. Seryasat et al., Fault detection of rolling bearings using discrete wavelet transform and neural network of SVM, *Advances in Environmental Biology*, vol. 8(6) Special, pp. 2175-2183, 2014.
- Milind Natu, Bearing Fault Analysis Using Frequency Analysis and Wavelet Analysis, *International Journal of Innovation, Management and Technology*, vol. 4, no. 1, February, 2013.
- Anuradha Abewardana and Upul Sonnadara, Classification of Birds using FFT and Artificial Neural Networks, *Proceedings of the Technical Sessions*, Institute of Physics – Sri Lanka, vol. 28, pp. 100-105, 2012.
- D.H. Pandya et al., ANN based fault diagnosis of rolling element bearing using time-frequency domain feature, *International Journal of Engineering Science and Technology (IJEST)*, vol. 4, no.6, June, 2012.
- V.Hariharan and PSS. Srinivasan, New approach of classification of rolling element bearing fault using artificial neural network, *Journal of Mechanical Engineering, Transaction of the Mech. Eng. Div., The Institution of Engineers, Bangladesh*, vol. ME 40, no. 2, December 2009.
- Yuan Xie and Tao Zhang, Fault diagnosis for rotating machinery based on convolutional neural network and empirical mode decomposition, *Hindawi Shock and Vibration*, vol. 2017, Article ID 3084197, 12 pages, 2017.
- Quang Hung Do, Predictions of machine vibrations using artificial neural networks trained by gravitational search algorithm and back-propagation, *International Journal of Artificial Intelligence*, vol. 15, pp. 93-111, 2017.
- Ran Zhang et al., Fault diagnosis from raw sensor data using deep neural networks considering temporal coherence, *sensors*, vol. 17, 549, 2017.
- CaseWestern Reserve University. B.D.C. Seeded Fault Test Data. Available online: <http://csegroups.case.edu/bearingdatacenter/home/> (accessed on 29 October 2018).
- Muhammad Sohaib et al., A hybrid feature model and deep-learning-based bearing fault diagnosis, *Sensors* 2017, 17, 2876; doi:10.3390/s17122876, 2017.
- Mark Hudson Beale, Martin T. Hagan, Howard B. Demuth, Neural Network Toolbox™, *User's Guide, The MathWorks, Inc.*; 2017.

## Biographies

**Punyisa Kuendee** is a doctoral student in the Faculty of Industrial Engineering at King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. She received a Bachelor of Science degree in General Science from Kasetsart University, Bangkok, Thailand, and a Master of Engineering degree in Industrial Engineering at Chulalongkorn University, Bangkok, Thailand. Her research interests include manufacturing process, operation research, and quality control.

**Udom Janjarassuk** is an Assistant Professor in Industrial Engineering in the Faculty of Engineering at King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand. He holds a Bachelor of Engineering degree in Electrical Engineering from King Mongkut's Institute of Technology Ladkrabang, a Master of Industrial Engineering and a PhD in Industrial Engineering from Lehigh University, USA. His research interests include automation, simulation, scheduling, and optimization.