2019 IEEE 6th International Conference on Industrial Engineering and Applications

Motor Fault Diagnosis Using CNN Based Deep Learning Algorithm Considering Motor Rotating Speed

Ji-Hoon Han, Dong-Jin Choi, Sun-Ki Hong
Department of Digital Control Engineering
Hoseo University
Asan, Korea
e-mail: sunkih4@naver.com

Hyeong-Seob Kim
Department of Information Control Engineering
Hoseo University
Asan, Korea
e-mail: uiophh@nate.com

Abstract—**Motor fault is a major problem in unmanned oriented systems such as smart factories. Recently, some studies have been actively carried out to make such a fault diagnosis unattended by a deep learning algorithm. However, these studies do not take into account the speed at which motors are driven, and therefore they are not appropriate for the actual system. In this paper, the experiment was performed to develop a deep learning algorithm that considered the motor speed. The method is that regard the vibration signal as an image and uses the algorithm, CNN, which is suitable for this processing. Adopting the suitable model reduced over-fitting and increased accuracy by reducing the model complexity. As a result, it is shown that fault diagnosis can be performed considering the motor driving speed by using the deep learning algorithm.**

Keywords-deep learning; motor fault diagnosis; CNN; BPNN; data acquisition; motor speed; gear fault

I. INTRODUCTION

With the advent of the fourth industrial revolution, the studies on the 'Internet of Things' and 'Deep Learning' are actively underway. Among them, the research on the smart factory to which information and communication technologies to the production process is applied has received considerable attention. However, the studies on the smart factory are focused on automation. Therefore, it is still not enough to cope with the fault situation of factory facilities. It is the electric motor that occupies most of the factory facilities. The signals that indicate mechanical fault such as damaged bearing or gear in the motor are noise and vibration. In practice, most faults at industrial sites are detected through the noise and vibration felt by workers. In addition, the classical fault diagnosis is performed by experts who have knowledge of the motor signal. And this work is done in the industrial site. The above two factors are not suitable for the smart factory which is aiming to be unattended.

For these reasons, studies of unattended fault diagnosis using deep learning are being actively carried out [1]. Applying deep learning to the fault diagnosis will result in a cost advantage because the system operates unattended. In addition, since fault diagnosis is always carried out, it helps maintenance of factory facilities. But the biggest advantage of this technology is that it can overcome the problem of diagnosing classic motor faults. The problem is that the motor fault signals are difficult to analyze even by experts. This is because the fault signals vary depending on the size and specifications of the motor and how it is operated. The fault signal is also very susceptible to disturbance.

These problems can be solved by using data-based learning and classification technique called as deep learning. However, previous studies that apply deep learning to fault diagnosis of the motor have generally assumed various states of the fault condition, but have experimented with the fixed motor driving speed [2~4]. However, the motors used in actual factory facilities have different driving speeds for that purpose. In this paper, the study was performed to develop a deep learning algorithm that can diagnose faults at unfixed speeds in order to overcome the fixing motor speeds problem of previous studies. In order to cope with the data problem whose characteristics change according to the situation, a study on the developing of the deep learning algorithm which is relatively insensitive to over-fitting has been performed.

The composition of the paper is as follows. Section II describes the differences between this paper and previous studies, and Section III describes the experimental environment and hyper-parameters. In Section IV, the results of the experiment for evaluating the performance of the algorithm considering the motor speed are explained. Finally, the conclusion describes the efficiency and applicability of fault diagnosis that using deep learning at the unfixed motor speed.

II. THE DIFFERENCE FROM PREVIOUS STUDIES

This section explains the differences between the deep learning model proposed in this paper and previous studies [2~4]. The difference is represented by the data population and the simplified model.

A. *Differences in Data Population*

First, the difference between previous studies and this paper is the data population. The used training data is the motor vibration signal. As described in the previous section, the operating conditions of the motor are so diverse depending on the needs of the user. Therefore, in this paper, the vibration signal measured by changing the motor speed, which is one of the operating conditions of the motor, was used as training data.

Fault diagnosis of a motor considering the driving speed needs to be interpreted differently than fault diagnosis with many fault states. The linear regression training model

978-1-7281-0851-3/19/$31.00 ©2019 IEEE

output in this paper is the value calculated by the soft-max function as shown in (1).

$$y_k = \frac{\exp(a_k + C')}{\sum_{i=1}^{n} \exp(a_i + C')} \tag{1}$$

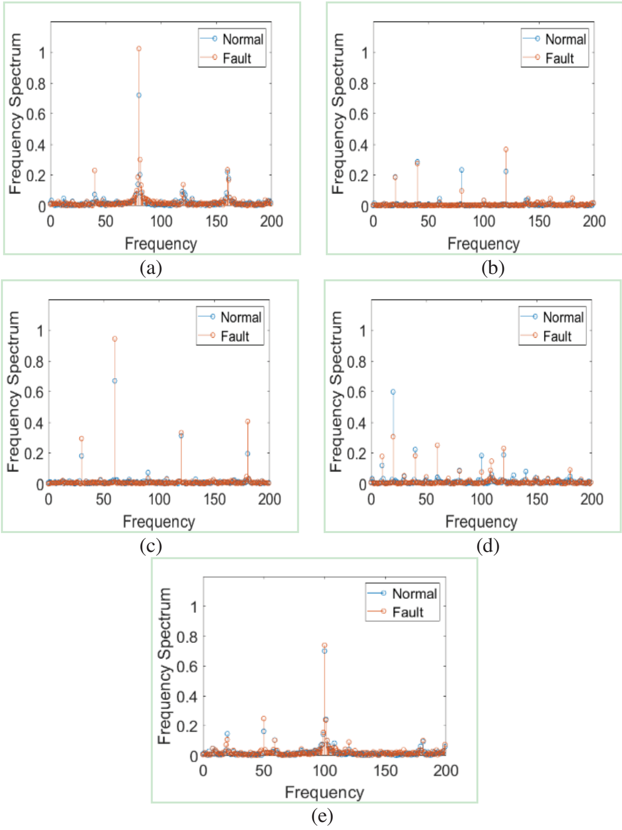

Figure 1. Normal/Fault FFT waveform for each driving frequency: (a) 10Hz, (b) 20Hz, (c) 30Hz, (d) 40Hz and (e) 50Hz.

where a_k is the output per class, n is the number of output layers, and C' is the value obtained by performing $\log C$ operations at the maximum output value C to prevent overflow in computer operations. Class means the objects to be classified by the algorithm. In this paper, normal and fault are class. The fault diagnosis with many fault states in previous studies become to have many classes. In the case of various conditions for one class means the increment of the data complexity itself with the class fixed. In fact, the data characteristics vary greatly with each driving speed. The FFT results of the vibration signal at each driving frequency are shown in Fig. 1. Fig. 1 shows the normal and fault FFT signals at 10, 20, 30, 40 and 50 Hz. It shows the increased data variance due to speed considerations.

B. Proposed Simplified Model

Secondly, the difference between previous studies and this study is that the model is relatively simple. The reason for simplifying the model is illustrated in Fig. 1, the variance of the data itself is very large even within each class. Error due to the variance of the data itself and the bias and variance of the training model is expressed through (2) [5].

$$\varepsilon = \sigma_D^2 + \beta_M^2 + \sigma_M \tag{2}$$

where ε is the error of the model. σ_D^2 is the existing variance of the data. This value means an error that cannot be reduced. β_M means the error between the approximate function and the actual function of the training model. That is, it means an error that occurs when a wrong assumption is made in the training model. Generally, it is large when the model is simple. σ_M represents how much the approximate function varies in the vicinity of the average. Generally, when this value is large, an over-fitting problem arises [5]. Fig. 2 shows the relationship shown in (2).

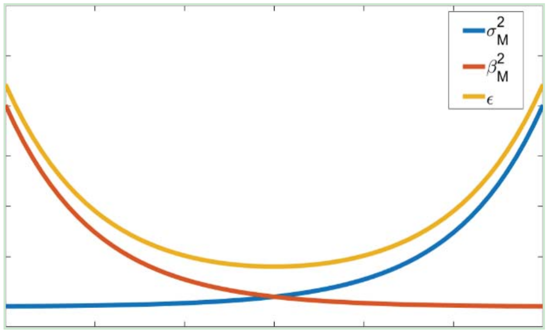

Figure 2. Bias-variance trade-off.

As shown in Fig. 1, if the speed is different, the σ_D^2 will increase and this will cause an increase of ε. The way to reduce ε is to reduce β_M and σ_M. However, the trade-off relationship is generally established between β_M and σ_M. a large error due to the large σ_D^2. That is, only one of β_M or σ_M must be reduced. Assuming that the σ_M is increased to reduce the β_M, the training model would have to be more complex to increase the σ_M. However, if the σ_M increases, the possibility of overfitting is increased as mentioned above, and in the field where the difference between the experimental training data set and the actual data set is likely to be large, such as fault diagnosis, the over-fitting may be a fatal disadvantage of the training model. Previous studies [2~4] may cause problems that are over-fitting because the model is relatively complex. Conversely, if β_M is increased to reduce the σ_M, that is, if the model is to be simplified, there would be a problem in classifying the data with large σ_D^2 as shown in Fig. 1. For the above reasons, a simple model is proposed in this paper to classify data which have a large σ_D^2. This proposed model consists of a combination of a single-layer CNN and a single-layer DNN (linear regression model). The reason why CNN is combined with DNN is that CNN is a structure that is mainly used to process images [6]. If the concept of CNN structure that can extract even very fine image features by using the filter is brought to the motor fault diagnosis, it can be seen as a simple image rather than a signal. That is, if the waveform of the FFT signal is regarded as an image in units of pixels, it is not difficult to extract features using a simple CNN structure. Fig. 3 shows the overall structure of the neural network used in this paper.

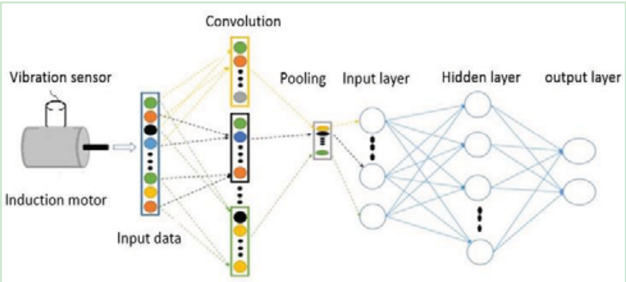

Figure 3. The structure of simplified CNN model.

III. Experimental Setup

A. Experimental Environment

In this paper, the experiment is performed to confirm the validity of the deep learning algorithm for motor fault diagnosis considering motor speed. The motor used in the experiment is a 4 pole induction motor. The motor rated electrical capacity is 200 watt. The 10:1 reduction gear is attached to this motor. The internal view of the gear is shown in Fig. 4.

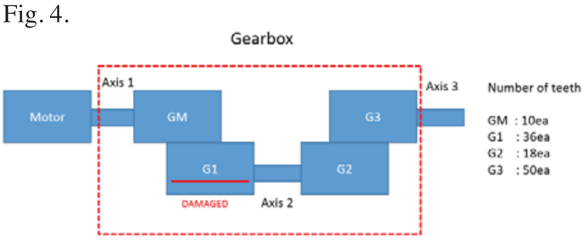

Figure 4. Structure of the Gearbox.

The data used for training is the motor x-axis vibration signal. To measure this data, a single axis acceleration vibration sensor is used. This signal is collected through the PXI. The sampling frequency is 20kHz. The experimental environment used in this paper is shown in Fig. 5.

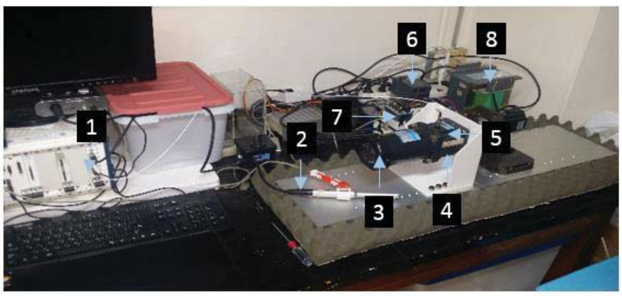

Figure 5. Experimental setup : (1) Data Acquisition Board, (2) Noise Measurement Equipment, (3) Induction Motor, (4) Gearbox, (5) Vibration Sensor, (6) Inverter, (7) Voltage and Current Measurement Equipment and (8) Transformer.

The code for training and verification of the training model is written through MATLAB, and the CPU of the computer is i7. The fault condition is one tooth broken in the center axis gear G2 as shown in Fig. 4. The damaged gear is shown in Fig. 6.

(a) (b)

Figure 6. Damaged gear : (a) Top View of Damaged Gearbox and (b) G1 Gear of Fig.4

B. Training Data Set

In this paper, the training data are the motor x-axis vibration signal. These are used by FFT as shown in Fig. 1. The reason for this is that the configuration of training data is set to vibration signal at all frequencies from 10 Hz to 50 Hz for fault diagnosis. By FFT that can clearly identify the data characteristics can be confirmed in narrower frequencies band. This affects the verification time and the accuracy of the training model. The complexity of the raw signals at each driving frequency is shown in Fig. 7.

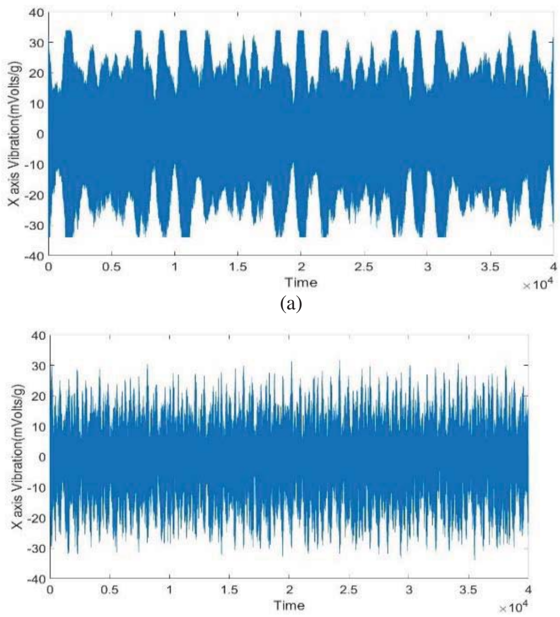

Figure 7. Raw signal in normal state : (a) driving frequency 10Hz and (b) driving frequency 50Hz

Compared to Fig. 1 the characteristic differences in Fig. 7 are relatively small. In addition, the length of the data required to identify the characteristics is also shorter when FFT is performed. For the above reasons, FFT signal for fault diagnosis at the unfixed driving frequency is selected as the training data.

C. Hyper-Parameter Setup

The training model used in this paper is based on supervised learning, where learning is the process of finding the optimum values of weights and filter parameters. However, the learning rate and the size of the batch, which are not the weights and filter values that change as the learning progresses, must be set by the human. These values are called hyper-parameters. These hyper-parameters have a profound effect on learning. Optimizing this should not be based solely on training data. If the performance of the training model is verified by immediately applying the hyper-parameter to the training data, the performance of the hyper-parameter should be estimated by making separate verification data because the hyper-parameters cause over-fitting the training data. Hyper-parameter optimization methods include experimental performance analysis and Bayesian optimization [7] based on mathematical theory. The hyper-parameters of the training model in this paper are the size of the batch, the learning rate, the number of epochs, and the momentum coefficient. The momentum coefficient is the parameter set when adding the concept of speed to the weight update. In this paper, experimental optimization is used for hyper-parameter optimization. The effects of hyper-parameters on the training model can be seen in Fig. 8.

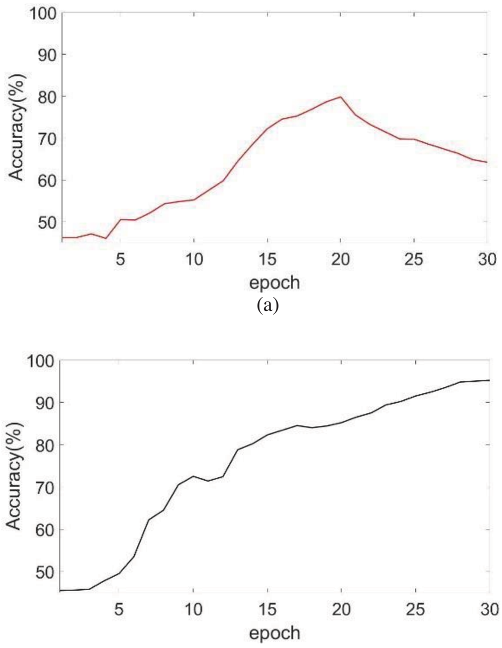

Figure 8. The accuracy of Epoch according to learning rate: (a) Inappropriate hyper parameters and (b) appropriate hyper parameters.

Fig. 8 shows that the accuracy increases with the increase of the epoch when the hyper-parameters are appropriate. However, if the hyper-parameters are inappropriate, the accuracy increases to a certain level and then decrease. The appropriate hyper-parameters are obtained through experimentation and generally have values between 0 and 1 [8].

IV. EXPERIMENT

For fault diagnosis at unfixed motor rotating speed, the population of training data is set to x-axis vibration signal at drive frequencies from 10 Hz to 50 Hz. Experiments are performed by changing the interval of training data and verification data to 1, 2, 5, and 10 Hz. The reason for adjusting the interval of data is to check the possibility of fault diagnosis for both case where the motor driving speed is limited and the case where the driving speed is not limited. Fig. 9 shows the normal and fault FFT signals at 10, 11, 15, and 20 Hz to show the diversity of the data.

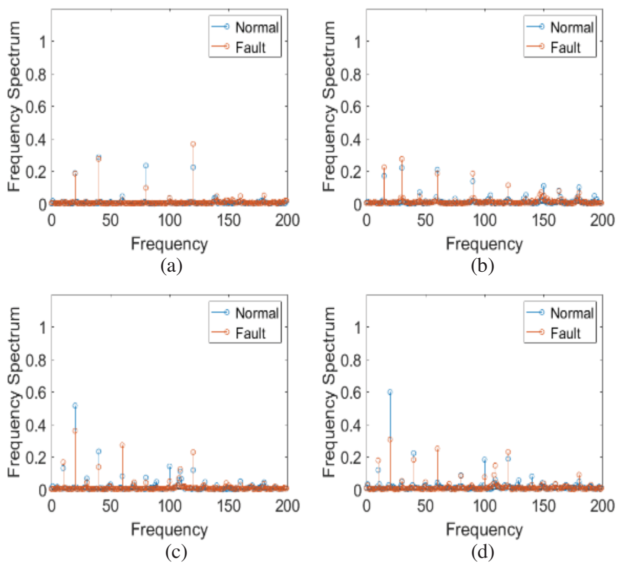

Figure 9. Normal/fault FFT waveform each driving frequency-2 : (a) 10Hz, (b) 11Hz, (c) 15Hz and (d) 20Hz.

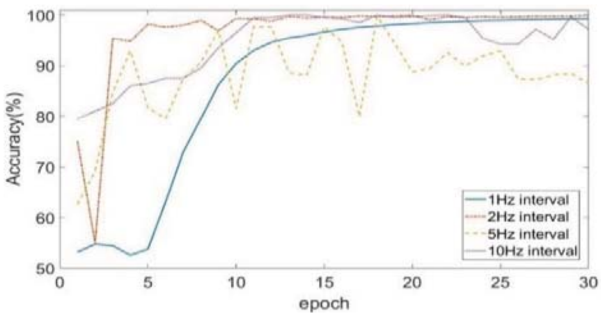

Figure 10. Accuracy per epoch for each population

The indicator of the deep learning algorithms performance is over-fitting. For fault diagnosis, the drop-out algorithm to prevent over-fitting [9] is added to the training process because actual data and the data used for training are highly likely to be different. In this paper, to evaluate the model performance, the number of the epoch to maximize accuracy and maximum accuracy are checked at each interval. The criteria for available accuracy is assumed to be 99% or higher. The reason for adding the epoch to the performance indicator is that if the number of the epoch is excessively large in order to obtain the accuracy even if a satisfactory accuracy (99% or more) is confirmed, the model is difficult to use for the general case. The experiment is performed with the same number of data at all intervals because the number of data affected performance. In addition, the experiment is performed with the same number of data at all frequencies. Fig. 10 and Table. 1 shows the results of the experiment when the number of data at all intervals is equal. Table. 2 show the results of the experiment when the number of data at all frequencies is the same.

TABLE I. EXPERIMENTAL RESULTS WITH THE SAME NUMBER OF DATA AT EACH INTERVAL

Interval	**Maximum Value**		**Number of Data**	
	Accuracy	***The number of Epoch***	***Each frequency***	***Total***
1Hz	99.28%	30	250 Ea	20000 Ea
2Hz	99.75%	20	500 Ea	
5Hz	99.94%	18	1250 Ea	
10Hz	100%	13	2000 Ea	

TABLE II. EXPERIMENTAL RESULTS WITH THE SAME NUMBER OF DATA AT EACH FREQUENCY

Interval	**Maximum Value**		**Number of Data**	
	Accuracy	***The number of Epoch***	***Each frequency***	***Total***
1Hz	99.28%	30	250 Ea	20000 Ea
2Hz	99.32%	29		10000 Ea
5Hz	99.42%	21		4000 Ea
10Hz	99.76%	19		2500 Ea

Analysis of the experiment results shows that the maximum accuracy increases as the data interval increases and the number of the epoch to reach it decreases. This can be explained as a result of an increase in the variance of the data itself when the bias and variance of the model are same through (2). It should be noted that even if the variance in the data has increased, the model accuracy is well enough. This is the result of adopting a suitable model for fault diagnosis considering speed. If fault diagnosis is required at a more detailed speed range, increase the number of data. These results show that if the number of data is sufficient, the fault diagnosis algorithm performs well.

V. CONCLUSIONS

In this paper, the deep learning algorithm considering the motor speed is developed. The previous studies had developed fault diagnosis algorithms at a fixed speed, which may have some limits to apply to industries. Also, when the speed is considered in the fault diagnosis, it is difficult to classify the data because of the data variance increases. To solve this problem, 'Simplified CNN' is proposed as the suitable model. The proposed model reduces the model complexity, thereby reducing the over-fitting and increasing the classification accuracy. To confirm the performance of the proposed model, the accuracy was verified by varying the number of data for each interval. As a result, if the number of data was sufficient the proposed model was confirmed to have sufficient performance of 99% or more accuracy. These results show some advantage in application to the industrial site, since the motor speed that were not considered in previous studies [2~4] has been taken into account.

In this paper, the number of training data was limited to prove the results of the experiment. However, as the experimental results show, the number of data significantly affects the deep learning algorithm performance. For this reason, it is desirable to include 'Internet of Things' in fault diagnosis to store motor signals in the database and ultimately to make big data. A study to apply 'Internet of Things' to fault diagnosis is being performed with this study [10].

Finally, if the training data is sufficient for the motor rotating speed, it is expected that the motor fault diagnosis can be performed with good accuracy using the deep learning algorithm.

ACKNOWLEDGMENT

This research was supported by Korea Electric Power Corporation. [Grant number: R18XA06-23]

REFERENCES

[1] J. W. Sun, R. Wyss, A. Steinecker, and P. Glocker, "Automated fault detection using deep belief networks for the quality inspection of electromotors", Technisches Messen, vol. 81, pp. 255-263, April 2014.

[2] Siyu Shao, Wenjun Sun, Peng Wang, Robert X, Gao, Ruqiang Yan, "Learning features from vibration signals for induction motor fault diagnosis" , 2016 International Symposium on Flexible Automation, pp. 71-76, 2016

[3] W. Li, S. Zhang, and G. He, "Semisupervised distance-preserving selforganizing map for machine-defect detection and classification," IEEE Trans. Instrum. Meas., vol. 62, no. 5, pp. 869–879, May 2013

[4] X. Zhang, Y. Liang, and J. Zhou, "A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM," Measurement, vol. 69, pp. 164–179, June. 2015

[5] Geman, S. and Bienenstock, E. Neural networks and the bias/variance dilemma. Neural Computation, 4:1–58,

[6] K. Simonyan and A. Zisserman. “Very deep convolutional networks for large-scale image recognition” International Conference on Learning Representations (ICLE), 2015.

[7] Jasper Snoek, Hugo Larochelle, Ryan P. Adams, “ Practical Bayesian Optimization of Machine Learning Algorithms “ Advances in Neural Information Processing Systems 25, 2012

[8] James Bergstra and Youshua Bengio, “Random Search for Hyper-Parameter Optimization” , Journal of Machine Learning Research 13, pp. 281-305, February 2012

[9] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout : A simple way to prevent neural networks from overfitting” The Journal of Machine Learning Research, pp. 1929-1958, 2014

[10] D. J. Choi, J. H. Han, H. S. Kim, M. J. Choi, C. W. Kwon, S. K. Hong, “Data Acquisition System for Real Time Motor Fault Diagnosis and Deep Learning Process”, Korean Institute of Electrical Engineers (KIEE) Summer Conference, July 2018.