Research Article

Condition Monitoring of Chain Sprocket Drive System Based on IoT Device and Convolutional Neural Network

Sang Kwon Lee ¹, **Jiseon Back**¹, **Kanghyun An**¹, **Sunwon Kim**¹, **Changho Lee**²,
and **Pungil Kim**²

¹Department of Mechanical Engineering, Inha University, Incheon 22201, Republic of Korea

²Korea Conveyor Inc. Co. Ltd., 627-3, Gojan-dong, Namdong-gu, Incheon 21633, Republic of Korea

Correspondence should be addressed to Sang Kwon Lee; sangkwon@inha.ac.kr

Received 20 May 2020; Revised 30 June 2020; Accepted 6 July 2020; Published 25 July 2020

Academic Editor: Francisco Beltran-Carbajal

Copyright © 2020 Sang Kwon Lee et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a condition monitoring method for the early defect detection in a chain sprocket drive (CSD) system and classification of fault types before a catastrophic failure occurs. In the operation of a CSD system, early defect detection is very useful in preventing system failure. In this work, eight fault types associated with the CSD system components, such as the gear tooth, bearings, and drive motor shaft, were arbitrarily damaged and incorporated into the CSD system. To detect the fault signals during the CSD system operation, the vibration was measured using an Internet of Things (IoT) device, which features a wireless MEMS accelerometer, Bluetooth function, Wi-Fi function, and battery. The IoT device was mounted on the gearbox housing. The measured one-dimensional vibration time-series was transformed into time-scale images using continuous wavelet transform (CWT). A convolution neural network (CNN) was employed to extract deep features embedded in the images, which are closely related to fault types. To update the learning parameters of the CNN, the RMSprop learning algorithm was applied, and the CNN was trained using 500 image samples. Multiple-classification performance of the trained network was tested using 100 image samples. Feature maps for different fault types were obtained from the final CNN convolution layer. For the visualization of fault types, t -stochastic neighbor embedding was employed and applied to the feature maps to convert high-dimensional data into two-dimensional data. Two-dimensional features enabled excellent classification of the eight fault types and one normal type.

1. Introduction

In a mechanical system, the gear train, belt pulley drive, and chain sprocket drive (CSD) are used to transmit power using rotating shafts. When the drive shaft and driven shaft are near each other, a gear drive unit is used. For a location where the distance between the drive shaft and driven shaft is relatively short, a belt pulley drive unit or CSD unit is often used. A CSD system has the advantage of transmitting power with greater force than a belt pulley drive system and does so by transmitting power without slip during the transmission process. Thus, CSDs have been applied in nearly all mechanical industries, such as machine tools, marine and aerospace drives, motorcycles, and the timing system for automotive engines. A CSD system failure while in operation can cause catastrophic human and economic losses.

Therefore, it is critical to identify defects in a CSD system before it breaks down. A CSD system generally consists of a chain sprocket unit and its driving system. Damage to the chain sprocket unit is mainly due to roller chain fatigue [1]. The driving system of a chain sprocket unit consists of the electric motor, gears, bearings, and rotating shaft. Damage to a driving system is caused by a defect or failure [2, 3] of the bearing supporting the rotating shaft and gear tooth, which transmits the driving force to drive the chain sprocket unit. Several studies have been conducted to evaluate methods for rotating machinery. For the feature extraction of fault signals, traditional fault diagnosis methods such as time-average method [4], cepstral analysis [5], pseudo-Wigner-Ville distribution (PWVD) [6], discrete wavelet transform (DWT) [7], higher-order method [8], adaptive line enhancement [9], empirical mode decomposition (EMD) [10], and

cyclostationary analysis [10, 11] have been used. CWT is nonorthogonal wavelet transform, and DWT is orthogonal transform. Fourier transform of mother wavelets for CWT and that for DWT are different from each other. The smoothing effect of CWT is better than that of DWT in the frequency domain. DWT is a useful method for the compression or recovery of signal, but CWT has advantage for the time-frequency analysis of signal to extract the fault information. In PWVD (pseudo-Wigner-Ville distribution) if we determine wrong kernel function, there is a cross-term problem. EMD is theoretically incomplete tool and not useful for the analysis in case of multiple faults. Thus, CWT has been widely used for time-frequency analysis of a fault signal. To classify fault patterns using fault features, k-nearest neighbor algorithms [12], Bayesian classifiers [13], support vector machines (SVMs) [14], and artificial neural networks (ANNs) [15] have been used. Most recently, owing to the development of computational performance and learning algorithms, deep learning approaches that include both feature extraction and classification of fault patterns have been applied in the field of fault diagnosis [16, 17] and have become the most popular diagnosis methods [18–23]. Most papers on the fault diagnosis of rotating machinery using deep learning techniques have focused on the classification of a specific single component defect, such as the bearing wear pattern [16, 18–20], gear tooth failure pattern [17, 21, 22], and unbalanced rotating shaft [23]. These previous studies analyzed and classified the fault pattern of single components such as a tooth, bearing, and rotating shaft of the mechanical system. This paper presents the diagnosis of a CSD system and multiple classifications of fault patterns using a deep learning technique. In this work, the CSD system components such as the electric motor shaft, bearings, and gears were arbitrarily damaged and assembled. A total of eight fault-states were created. The signals used for fault diagnosis were the vibrational acceleration measured in the CSD system. In this work, Internet of Things (IoT) sensor has been developed for the wireless measurement of fault signal and the real time transmission of measured data. The developed IoT sensor was firstly applied to the laboratory test for the validation of the new method. It is being used for the detection of chain convey system. An IoT device measured the acceleration. The IoT device is newly developed and consists of a wireless microelectromechanical system (MEMS) accelerometer, Bluetooth function, Wi-Fi function, and battery. The wireless MEMS accelerometer is useful in IoT applications because its small size and battery-powered operation are the typical requirements for IoT sensors [24, 25]. For the fault stage diagnosis or normal state using deep learning, image data are essential. Therefore, one-dimensional time data were converted into time-scale image data by continuous wavelet transform (CWT). The image data include time-frequency information related to the fault types. A convolutional neural network (CNN) was employed for multiple classifications of eight fault-states and one normal state, and the image data were used as the CNN input. The combination of CWT and CNN was introduced in field of fault detection [26–28]. The innovative difference for these algorithms is the difference in the CNN structure

according to application area. The CNN structure such as filter size (kernel size), number of layers, and number of inputs and outputs should be determined optimally for the successful application in different systems. In this study, the combination of CWT and CNN was also employed for the condition monitoring of the CSD system. The new optimal structure of CNN was presented for diagnosis of the multiple faults in the CSD system. Throughout the CNN, the patterns of eight fault types and a normal type were successfully classified, and their feature maps were well extracted.

2. Theory

2.1. Theory of CWT. The CWT [29] is based upon a family of functions:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad a > b \in \mathfrak{R}, \quad (1)$$

where ψ is a fixed function called the “mother wavelet,” which is localized in terms of both time and frequency. The function $\psi_{a,b}(t)$ is obtained by applying the operations of shifting (b -translation) in the time domain and scaling in the frequency domain (a -dilation) to the mother wavelet. The mother wavelet used throughout this study is the Morlet wavelet [30],

$$\psi(t) = \frac{1}{\sqrt{\pi B}} e^{j\omega_0 t - (t^2/B)}, \quad (2)$$

where ω_0 is the center frequency of the “mother wavelet” when the mother wavelet is transformed to the frequency domain. B is the bandwidth defined as the variances of the Fourier transform $\psi(f)$ of the Morlet wavelet,

$$B = \int_{-\infty}^{+\infty} f^2 \psi^*(f) df, \quad (3)$$

where f indicates the frequency and $*$ denotes the complex conjugate.

The CWT of a signal $x(t)$ is defined by

$$W_x^\psi(a, b) = \int_{-\infty}^{+\infty} x(t) \psi_{a,b}^*(t) dt = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi^*\left(\frac{t-b}{a}\right) dt, \quad (4)$$

where $\psi^*(\cdot)$ is the complex conjugate of $\psi(\cdot)$ and the function $x(t)$ satisfies the condition

$$\|x\|^2 = \int_{-\infty}^{+\infty} |x(t)|^2 dt < +\infty. \quad (5)$$

Here, $\psi_{a,b}(t)$ plays an analogous role to the $e^{j\omega t}$ in the definition of the Fourier transform. If the mother wavelet $\psi(t)$ satisfies the admissibility condition:

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < +\infty, \quad (6)$$

then the inverse wavelet transform can be obtained by

$$x(t) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W_x^\psi(a, b) \psi_{a,b}(t) \frac{db da}{a^2}. \quad (7)$$

2.2. Theory of CNN. Artificial neural networks have been widely used for the prediction and classification of sound and vibration signals. Before deep neural networks (DNNs) were introduced, ANNs with shallow neural network (SNN) structures were used [31–33]. An SNN uses a supervised training process with a feature vector. However, it is difficult to extract the system fault features if the dynamic system characteristics are unknown. Therefore, the DNN structure was developed for feature extraction [34]. A CNN is one of the DNN structures [35], as shown in Figure 1. A CNN uses an unsupervised training process and feature maps related to fault information, which are extracted from the stages of several convolutional and pooling layers. The neurons in a CNN are arranged in the form of feature maps. The input to a convolutional layer is the image x of size $m \times n$ in the CNN. The convolutional layer contains f filters (kernels) of size $r \times s$, which have smaller dimensions than the input image. The output of the convolutional layer is a set of f feature maps of size $(m - s + 1) \times (n - s + 1)$ by striding over one pixel. The filter, realized by assigning a weight f_{ij} to each pixel in the input image and calculated as a weighted sum, extracts certain features contained in the image. The weighted sums are then added by an additive bias and passed through a nonlinear function to obtain pixels in the convolutional map. Traditionally, sigmoid and hyperbolic tangent functions were used; recently, rectified linear units [35] have become popular. The activation output y_j^l of a particular feature map j in the convolutional layer l is given as

$$y_j^l = \phi \left(b_j^l + \sum_{i \in Z_j^l} y_i^{l-1} \otimes f_{ij}^l \right), \quad (8)$$

where ϕ is the nonlinear activation function; b_j^l is the scalar bias for the l^{th} layer; Z_j^l is the selected feature map i in the $(l-1)^{\text{th}}$ layer, which is summed up by the feature map j in the l^{th} layer; $\otimes$ denotes the convolutional operator that convolutes the activation y_i^{l-1} of the preceding layer; and f_{ij}^l is the filter used to perform the convolutional operation.

The filter weights are trained to detect specific features. Hence, effective feature selection in successive stages that can distinguish between different categories is necessary for the accurate classification of new input images. This is followed by a pooling layer. Each feature map is subjected to region-wise pooling, such as the maximum or average of nonoverlapping pixels. The output of the pooling layer leads to a dimensional reduction depending on the chosen stride. The activation output p_d^l after downsizing the feature map d in a layer l is given as

$$p_d^l = \chi(d_j^l, N^l), \quad (9)$$

where χ is the downsizing function, such as the average or maximum function downsized by a factor of N_j^l , and d_j^l is the convoluted feature map to be downsized.

As the original input passes through successive convolution and pooling processes, the network learns to efficiently represent all of the images. The last neural network layer is a fully connected layer whose output y_p is given by

$$y_p = \psi(\mathbf{b}_p + \mathbf{W}\mathbf{f}), \quad (10)$$

where $\mathbf{b}_p$ is the bias for the output layer, $\mathbf{W}$ is the weight matrix between the input and output layers of the fully connected layer, $\mathbf{f}$ denotes the feature maps of the fully connected input layer, and ψ is the softmax function [35]. The parameters b_j^l , f_{ij}^l , b_p , and W were learned during training. The training takes place via stochastic gradient descent (SGD) with the objective of minimizing the error between the actual and desired output. The gradient was computed using the backpropagation method [36]. All the filter weights and biases were updated according to the objective function for each input sample until an optimal representation is obtained for the training samples. For the backpropagation method, the cost function J is defined as

$$J = \frac{1}{2} \sum_p^P (t_p - y_p)^2, \quad (11)$$

where P is the number of output neurons, t_p is the p^{th} element of the target output for the p^{th} fault, and y_p is the actual output of the network for the p^{th} fault. A common problem encountered in training CNNs is overfitting, which results in poor performance in a set of holdout tests after the network is trained on a small or even large training set; this affects the ability of the model to generalize unseen data. The SGD algorithm has been used as a learning algorithm [34] in backpropagation. To overcome the overfitting problem, adaptive moment estimation (ADAM) [37, 38] was proposed, which is another method that computes the adaptive learning rates for each parameter. Recently, Hinton [39] proposed the RMSprop algorithm. RMSprop is an unpublished, adaptive learning rate method; it lies within the realm of adaptive learning rate methods that have seen growing popularity in recent years. In this study, the RMSprop algorithm was employed. For the RMSprop algorithm, the update rule is mathematically given by

$$E[(\nabla J)^2]_k = \beta E[(\nabla J)^2]_{k-1} + (1 - \beta) \left(\frac{\delta J}{\delta w} \right)^2, \quad (12)$$

$$w_k = w_{k-1} - \frac{\eta}{\sqrt{E[(\nabla J)^2]_k}} \left(\frac{\delta J}{\delta w} \right),$$

where $E[(\nabla J)^2]$ is the moving average of squared gradients, $\delta J / \delta w$ is the gradient of the cost function with respect to the weight, η is the learning rate, and β is the moving average parameter. The default value for the moving average parameter that can be used in projects is 0.9, which works very well for most applications.

3. Experiment

3.1. CSD System and Synthetic Fault Patterns for Test. A visualization of the CSD system used for the experiment is shown in Figure 2. The CSD system used for testing is the one on the right in Figure 2(a). The technical specifications are summarized in Table 1.

FIGURE 1: Flowchart of signal processing for fault detection and classification.

FIGURE 2: Visualization of the CSD system and sensors used for experimentation: (a) total view; (b) sensor positions; (c) motor, bearings, and gearbox.

In the CSD system used for testing, there are four bearings, four gears, one chain, two sprockets, and one electric motor. The input shaft rotational speed is 1800 rpm (30 Hz), and that of the output shaft is 59.4 rpm

(0.99 Hz). The speed reduction ratio is 30. If the chain and sprocket are damaged, the CSD system cannot work; therefore, artificial faults were created only in the CSD system drive portion, such as the motor shaft, bearings,

TABLE 1: Technical specifications of the CSD system.

Component	Property	Value
Motor	Rotational speed (no slip)	1800 rpm
	Rotational speed (slip)	1735 rpm
	No. of poles	4
	No. of rotors	8
	No. of wings	7
Gearbox	No. of teeth in drive gear 1	15
	No. of teeth in driven gear 1	82
	No. of teeth in drive gear 2	16
	No. of teeth in driven gear 2	88
	Rotational speed	0.99 Hz
Sprocket	No. of teeth in drive sprocket	19
	No. of teeth in driven sprocket	47
	Pitch drive sprocket	192.9 mm
	Pitch driven sprocket	475.4 mm
	Rotational speed of chain	0.99 Hz
Chain	Pitch of chain	31.75 mm
	Length of chain	3048 mm
	No. of links	96
	Pitch driven sprocket	

	Property	Rotation frequency (Hz)	No. of balls (EA)	Spec. unit Diameter of ball (mm)	Pitch circle (mm)	Contact angle (°)
Bearing	Bearing ①	30	9	7.93	38.95	0
	Bearing ②	5.49	7	7.93	29.47	0
	Bearing ③	5.49	7	9.52	35.47	0
	Bearing ④	0.99	8	13.48	58.33	0

and gear teeth. Faults were not introduced in the sprocket and chain.

The faults and conditions created are as follows:

- (1) Cracked tooth in Driven Gear 2 (Fault 1)
- (2) Broken tooth in Driven Gear 2 (Fault 2)
- (3) Unbalanced shaft in electric motor (Fault 3)
- (4) Outer-raceway fault in Bearing ① (Fault 4)
- (5) Inner-raceway fault in Bearing ① by inserting matter (Fault 5)
- (6) Unbalanced shaft in electric motor + broken tooth in Driven Gear 2 (Fault 6)
- (7) Unbalanced shaft in electric motor + broken tooth in Driven Gear 2 + outer-raceway fault in Bearing ① (Fault 7)
- (8) Unbalanced shaft in electric motor + broken tooth in Driven Gear 2 + inner-raceway fault in Bearing ① (Fault 8)
- (9) Normal condition (normal)

Some images of synthetic fault conditions are presented in Figure 3. Figures 3(a)–3(e) show the cracked gear tooth, broken gear tooth, unbalanced shaft of the electric motor, outer-raceway fault by making a hole, and inner-raceway fault by inserting matter through the hole, respectively.

3.2. Data Acquisition. The experiments were performed under the conditions shown in Figure 2(c). The helical gearbox consists of two gears mounted on independent shafts. The input shaft is connected to the electric motor, which transforms electrical energy into rotational movement for transmission to the mechanical system. The output shaft is linked to the chain sprocket unit, which has a chain connected to the final output shaft and transforms electrical energy into mechanical force, as opposed to the rotational movement of the final shaft. The final output shaft can be used to drive the conveyor system. The most important details regarding these mechanical components are listed in Table 1. Vibration data were obtained using a vibration sensor unit mounted on the gearbox housing to determine the acceleration in the vertical direction. Sound data were obtained using a microphone placed at a distance of 1 m from the CSD system to measure radiated noise, as shown in Figure 2(b). A vibration sensor unit, called “Happy-Go,” was built with the following components:

STM32F030C8T6 (MCU) (CPU)
 Intsain BPT2640TXSRF (BLE) (Bluetooth)
 Intsain WIS3200 (WLAN) (IoT development platform)
 Analog devices ADXL337 (MEMS accelerometer)
 Battery (3.7 V, 2600 mAh)

FIGURE 3: Examples of different bearings and gear conditions and unbalanced shaft of the electric motor: (a) cracked gear tooth, (b) broken gear tooth, (c) unbalanced motor shaft, (d) outer-raceway fault by making hole, and (e) inner-raceway fault by inserting matter through the hole.

FIGURE 4: “Happy-Go” IoT device for the measurement of vibration: (a) MEMS accelerometer, (b) CPU, (c) IoT development platform, and (d) top-view of IoT device.

“Happy-Go” has its own built-in low-pass filter set to 1/2 the sampling rate; the device is shown in Figure 4.

The accelerometer also has a built-in high-pass filter, but it was deactivated. The test rotor was sampled at 2000 Hz, indicating that the cut-off frequency of the low-pass filter was 1000 Hz. The measurement range, sensitivity, noise density, and frequency bandwidth of the sensor are ± 3 g, 270 mV/ga, 175 μ g, and 2000 Hz, respectively. The vibration data were transmitted to a computer via Wi-Fi. Bluetooth is installed for the measured data transmission to the smart phone in the future but was not used in this study. A 1/2-inch free-field microphone (B&K 4192, Denmark) was used to measure sound data, which were transmitted to the computer through a data acquisition system (NI 9233, USA). With this setup, a dataset was created incorporating the healthy and faulty conditions presented in Section 3.1. A test was performed for each fault condition, resulting in a total of 700 test runs. Each test had a runtime of 5 min, from which the last 30 s of vibration data were captured using the accelerometers.

4. Signal Processing for the Measured Data

4.1. Data Analysis Based on Vibration Theory. The vibration data measured in the time domain were analyzed in the frequency and time-frequency domains using the MATLAB (MathWorks, USA) signal processing toolbox. Figures 5 shows the time history of eight faults and one

normal signal measured by the accelerometer during one test.

Figure 6 shows a comparison between the power spectrum of the normal vibration signal and eight fault vibration signals, respectively. The spectrum shapes of the eight fault signals were different from that of the normal signal. According to vibration theory of rotating machinery [40–43], under normal conditions, in rotating machinery like the CSD system, there are major vibration sources such as impact vibration between the sprocket and chain, gear meshing vibration, bearing rolling vibration, and electric motor shaft rotor vibration. The frequencies of these sources in the CSD system were calculated using the system specifications listed in Table 1:

- (i) Gear meshing frequency in Gear 1: 450 Hz (rotating frequency $\times$ number of teeth: 30 Hz $\times$ 15)
- (ii) Gear meshing frequency in Gear 2: 88 Hz (rotating frequency $\times$ number of teeth: 30 Hz $\times$ 15/82 $\times$ 160)
- (iii) Bearing 1 rolling frequency: 270 Hz (rotating frequency $\times$ number of balls: 30 Hz $\times$ 9)
- (iv) Bearing 2 and 3 rolling frequency: 39 Hz (rotating frequency $\times$ number of balls: 5.49 Hz $\times$ 7)
- (v) Bearing 4 rolling frequency: 8 Hz (rotating frequency $\times$ number of teeth: 0.99 Hz $\times$ 8)
- (vi) Chain and sprocket impact frequency at drive shaft: 18.8 Hz (rotating frequency $\times$ number of sprockets: 0.99 Hz $\times$ 19)

FIGURE 5: Time history of vibration data recorded by the MEMS accelerometer mounted on the CSD system. (a) Fault no. 1. (b) Fault no. 2. (c) Fault no. 3. (d) Fault no. 4. (e) Fault no. 5. (f) Fault no. 6. (g) Fault no. 7. (h) Fault no. 8. (i) Normal.

(vii) Structural dynamic resonance of the CSD system: 57 Hz

The frequencies have several vibration peaks corresponding to these sources, as shown in Figure 6. The frequency region of these vibration peaks is under 500 Hz. The variations in vibration caused by the eight faults primarily occur in the frequency region above 500 Hz, as shown in

Figure 6. The peaks under 500 Hz are related to the rotating frequency of the shaft, the teeth meshing frequency, the contact frequency of balls in bearings, and the contact frequency between chain and sprocket. The high frequency peaks are related to mechanical faults such as broken teeth, bearing wear, and rotating shaft imbalance. The mutual spectrum shapes of the eight fault signals were also different.

FIGURE 6: Comparison between spectrum for faults and normal vibration data. (a) Fault no. 1. (b) Fault no. 2. (c) Fault no. 3. (d) Fault no. 4. (e) Fault no. 5. (f) Fault no. 6. (g) Fault no. 7. (h) Fault no. 8.

According to the spectrum shape differences, the eight fault types can be distinguished and classified using vibration signals. For a more meaningful analysis of fault classification, CWT was applied to the vibration signals, and the time-frequency information for vibration signals was obtained. Figure 7 shows the CWT analysis results applied to the eight vibration signals. Detailed explanation of the major frequency components in the CWT analysis for the normal vibration signal is given in Figure 8 and listed in Table 2. The vibrations of these major frequencies are related to the vibration sources and listed in Table 2.

4.2. Feature Extraction. The traditional fault classification method in the field of machine learning uses a feature vector. Feature vectors extracted from raw signals have been used as the input of classifiers such as SVM and SNN. Therefore, numerous feature extraction methods have been studied for many years [4–11]. The major features of rotating machinery can be summarized [17, 44] as follows:

- (i) Peak to peak
- (ii) Root mean square time and frequency
- (iii) Standard deviation time and frequency
- (iv) Shape factor
- (v) Frequency center
- (vi) Impulse factor
- (vii) Crest factor
- (viii) Mean time and frequency (first moment)
- (ix) Variation time and frequency (second moment)
- (x) Skewness (third moment)
- (xi) Kurtosis (fourth moment)

To reduce feature dimensionality and improve classification accuracy, feature selection is critical to subsequent classifications. Several researchers have proposed effective methods of feature selection [45]. The classification machine using feature vectors has adopted the supervised training method [34]. However, it is difficult to extract the feature vectors for multiple-fault classification. Therefore, it is necessary to extract the features automatically based on the unsupervised training method.

5. CNN for Multiple-Fault Classification

5.1. Input for CNN. Recently, machine classification using an unsupervised training method based on DNN was proposed. The CNN is a classification machine using an unsupervised training method. The supervised training method demands a feature vector correlated to fault characteristics as the input of a classification machine. The major feature vectors of rotating machinery were mentioned in Section 4.2. These feature vectors are useful for the fault pattern classification of one mechanical element such as the tooth itself, bearing itself, or shaft itself. However, when multiple faults are taking place during CSD system operation, it is difficult to find which feature vectors are correlated to the multiple faults of tooth, bearings, shaft imbalance, and their

combination. Therefore, the feature vectors for the multiple classifications should be self-extracted using a DNN such as CNN. The CNN uses the raw signal instead of using features and generates a feature map at each CNN convolution layer stage. It is an unsupervised training method because the raw signal is used as the input of machine learning based on CNN. In general, the CNN uses images as input data. In this study, the CWT was applied to the signal recorded by a wireless MEMS accelerometer. The image data obtained by the CWT were used as the CNN input image, as shown in Figure 7. Vibration data recorded every 30 s were classified by the multiclass classifier as Fault 1, Fault 2, . . . , Fault 8 or normal. Among the images recorded every 30 s ($126 \times 60,000$), a reduced image (126×400) recorded for 0.2 s was used as the CNN input.

5.2. Network Setup and Training. After the vibration signals were processed using CWT and converted into two-dimensional images, they were used as CNN inputs. The DNN toolbox for MATLAB (MathWorks, USA) was utilized. The network architecture is shown in Figure 9. The first layer following the input layer is a convolutional layer with eight feature maps of filter size 3×3 . This is followed by a mean-pooling layer of size 2×2 . The next layer is a convolutional layer with 16 feature maps of filter size 3×3 , followed by a 2×2 mean-pooling layer. The output layer contains nine neurons corresponding to the eight different fault conditions and one normal condition. All the layers are fully connected. The softmax function was used as the classification function. The RMSprop method [39] was used to train the network with an initial learning rate of 0.001. The batch size was taken as 128. Training was carried out for 1000 iterations (50 epochs). The change in accuracy during the learning iterations is shown in Figure 10. An optimal value of weight is achieved in 390 iterations with minimum error. For this training, out of the 700 samples, 500 samples were used for training, 100 samples for validation, and 100 samples for testing, and 10 different networks were produced.

5.3. Results and Discussion. The classification accuracy ratio is the ratio of the number of correctly classified test samples to the total number of test samples. In this case, among the 100 test samples, 97 samples were correctly classified. As shown in Figure 10, the accuracy rate arrives at the maximum value and becomes stable at 390 iterations. The filter size in the convolution layer and the number of filters were optimized to ensure that the accuracy rate converges to the maximum value and is stable. The accuracy rate is defined as follows [20]:

$$\text{accuracy rate} = \frac{\text{the corrected classification of sample}}{\text{the total number of samples}}. \quad (13)$$

The network performed with a classification accuracy rate of 97%, misclassifying just three samples, as shown in Figure 11.

The feature maps for the eight fault-states and one normal state were extracted from the final convolutional layer of one network among the 10 trained networks, as shown in Figure 12. The figure clearly shows different feature maps for the eight fault-states and one normal state. According to these results, the feature map of each fault shows a different time-frequency characteristic.

The features must be visualized to verify the image recognition accuracy. If the data for each pixel are used for features as shown in Figure 12, then the features are high-dimensional data and show the time-frequency characteristics related to each fault. Therefore, *t*-stochastic neighbor embedding (*t*-SNE) was used for visualization by converting high-dimensional data into low two-dimensional data, as shown in Figure 13 [46]. “High-dimensional” implies that

FIGURE 7: Time-frequency analysis for the vibration by applying CWT to raw vibration signals. (a) Fault no. 1. (b) Fault no. 2. (c) Fault no. 3. (d) Fault no. 4. (e) Fault no. 5. (f) Fault no. 6. (g) Fault no. 7. (h) Fault no. 8. (i) Normal.

FIGURE 8: Detailed explanation for major frequency components based on CWT analysis for normal vibration signal.

TABLE 2: Signal analysis for vibration data measured by a sensor.

Mechanical source	Frequency (Hz)
Bearing 1 rolling	270
Gear 2 meshing harmonic	174
Gear 2 meshing	87
Vibration resonance of system	57
Bearing 2 rolling	37
Rotor vibration in electric motor	30
Impact at driver sprocket	18.27

the number of features is high and “two-dimensional” indicates that the number of features is two. Figure 13(a) shows the reduced two-dimensional features for 20 input images of the CNN; Figure 13(b) shows the reduced two-dimensional features for 20 iterated feature maps extracted from the final network convolution layer trained by the CNN. Finally, for the 10 feature maps obtained from the final convolution layer of 10 different networks trained by 10

FIGURE 9: Convolutional neural network architecture with two stages.

FIGURE 10: Learning error and accuracy rates during training process.

FIGURE 11: Classification results and accuracy.

FIGURE 12: Feature maps for eight fault-states and one normal state extracted from the final convolutional layer of the trained CNN network. (a) Fault no. 1. (b) Fault no. 2. (c) Fault no. 3. (d) Fault no. 4. (e) Fault no. 5. (f) Fault no. 6. (g) Fault no. 7. (h) Fault no. 8. (i) Normal.

FIGURE 13: Two-dimensional features reduced by t-SNE using (a) 180 raw input images, (b) feature map obtained by 20 iterations of one trained network, and (c) feature maps from 10 different networks.

different input datasets, the t-SNE was used for visualization by converting these features into low two-dimensional data. The results were plotted as shown in Figure 13(c). The features of the raw images are scattered randomly, making it difficult to classify the fault and normal states. However, the features obtained by the trained network are grouped according to fault or normal conditions. Therefore, it is possible to classify the fault and normal states.

6. Conclusions

In this paper, a novel condition monitoring approach for the CSD system is proposed, based on the integration of CWT and CNN. In the CSD system, eight fault-states and one normal state were artificially manufactured. An IoT device for vibration measurement was also developed. The device features a wireless MEMS accelerometer, Bluetooth function, and Wi-Fi function. The vibration data were measured using the wireless MEMS accelerometer mounted on the CSD system. One-dimensional vibration signals in the time domain were transformed into time-scale images via CWT. These images were then classified by a CNN, which can extract the underlying, deep features embedded in images that are closely related to fault types. As critical factors affecting the network classification accuracy, the filter size and number of filters were optimized in the convolutional and pooling layers of the CNN structures. Input images obtained by the CWT and feature maps extracted by the CNN are high-dimensional data. Thus, *t*-SNE was used for visualization by converting the high-dimensional data into low two-dimensional data. Two-dimensional features enabled clear classification of the eight fault-states and one normal state. The results showed that the condition monitoring approach for the CSD system based on the integration of CWT and CNN is an excellent classification method.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Inha University Research Grant.

References

- [1] S. Papadopolou, I. Pressas, A. Vazdirvanidis, and G. Pantazopoulos, "Fatigue failure analysis of roll steel pins from a chain assembly: fracture mechanism and numerical modeling," *Engineering Failure Analysis*, vol. 101, pp. 320–328, 2019.
- [2] M. Amarnath and S.-K. Lee, "Assessment of surface contact fatigue failure in a spur geared system based on the tribological and vibration parameter analysis," *Measurement*, vol. 76, no. 6, pp. 32–44, 2015.
- [3] N. K. Arakere, "Gigacycle rolling contact fatigue of bearing steels: a review," *International Journal of Fatigue*, vol. 93, pp. 238–249, 2016.
- [4] P. D. McFadden, "Examination of a technique for the early detection of failure in gears by signal processing of the time domain average of the meshing vibration," *Mechanical Systems and Signal Processing*, vol. 1, no. 2, pp. 177–183, 1987.
- [5] R. Randall, "Cepstrum analysis and gearbox fault detection," Technical Report, pp. 13–150, B & K Application, Copenhagen, Denmark, 1982.
- [6] W. J. Wang and P. D. McFadden, "Early detection of gear failure by vibration analysis i. calculation of the time-

- frequency distribution," *Mechanical Systems and Signal Processing*, vol. 7, no. 3, pp. 193–203, 1993.
- [7] W. J. Wang and P. D. McFadden, "Application of orthogonal wavelets to early gear damage detection," *Mechanical Systems and Signal Processing*, vol. 9, no. 5, pp. 497–507, 1995.
 - [8] S. K. Lee and P. R. White, "Higher-order time-frequency analysis and its application to fault detection in rotating machinery," *Mechanical Systems and Signal Processing*, vol. 11, no. 4, pp. 637–650, 1997.
 - [9] S. K. Lee and P. R. White, "The enhancement of impulsive noise and vibration signals for fault detection in rotating and reciprocating machinery," *Journal of Sound and Vibration*, vol. 217, no. 3, pp. 485–505, 1998.
 - [10] J.-S. Kim and S.-K. Lee, "Identification of tooth fault in a gearbox based on cyclostationarity and empirical mode decomposition," *Structural Health Monitoring*, vol. 17, no. 3, pp. 494–513, 2018.
 - [11] C. Capdessus, M. Sidahmed, and J. L. Lacoume, "Cyclostationary processes: application in gear faults early diagnosis," *Mechanical Systems and Signal Processing*, vol. 14, no. 3, pp. 371–385, 2000.
 - [12] D. Wang, "K-nearest neighbors based methods for identification of different gear crack levels under different motor speeds and loads: revisited," *Mechanical Systems and Signal Processing*, vol. 70–71, pp. 201–208, 2016.
 - [13] P. Baraldi, L. Podofillini, L. Mkrtychyan, E. Zio, and V. N. Dang, "Comparing the treatment of uncertainty in bayesian networks and fuzzy expert systems used for a human reliability analysis application," *Reliability Engineering & System Safety*, vol. 138, pp. 176–193, 2015.
 - [14] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer Science & Business Media, Berlin, Germany, 2013.
 - [15] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: a review," *Mechanical Systems and Signal Processing*, vol. 108, pp. 33–47, 2018.
 - [16] M. Gan, C. Wang, and C. Zhu, "Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 72–73, pp. 92–104, 2016.
 - [17] O. Janssens, V. Slavkovikj, B. Vervisch et al., "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331–345, 2016.
 - [18] Z. Chen, S. Deng, X. Chen, C. Li, R.-V. Sanchez, and H. Qin, "Deep neural networks-based rolling bearing fault diagnosis," *Microelectronics Reliability*, vol. 75, pp. 327–333, 2017.
 - [19] C. Lu, Z.-Y. Wang, W.-L. Qin, and J. Ma, "Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification," *Signal Processing*, vol. 130, pp. 377–388, 2017.
 - [20] C. Lu, Z. Wang, and B. Zhou, "Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification," *Advanced Engineering Informatics*, vol. 32, pp. 139–151, 2017.
 - [21] D. Yan, A. Wang, S. Wang, B. He, and G. Meng, "Fault diagnosis under variable working conditions based on STFT and transfer deep residual network," *Shock and Vibration*, vol. 2020, pp. 1–18, 2020.
 - [22] D. Cabrera, F. Sancho, C. Li et al., "Automatic feature extraction of time-series applied to fault severity assessment of helical gearbox in stationary and non-stationary speed operation," *Applied Soft Computing*, vol. 58, pp. 53–64, 2017.
 - [23] H. shao, H. Jiang, F. Wang, and H. Zhao, "An enhancement deep feature fusion method for rotating machinery fault diagnosis," *Knowledge-Based Systems*, vol. 119, pp. 200–220, 2017.
 - [24] S. Jiménez, M. O. T. Cole, and P. S. Keogh, "Vibration sensing in smart machine rotors using internal MEMS accelerometers," *Journal of Sound and Vibration*, vol. 377, pp. 58–75, 2016.
 - [25] G. Feng, N. Hu, Z. Mones, F. Gu, and A. D. Ball, "An investigation of the orthogonal outputs from an on-rotor MEMS accelerometer for reciprocating compressor condition monitoring," *Mechanical Systems and Signal Processing*, vol. 76–77, pp. 228–241, 2016.
 - [26] L. F. Guo, H. Li, H. Zheng, H. Li, and X. Pei, "Aeroengine control system sensor fault diagnosis based on CWT and CNN," *Mathematical Problems in Engineering*, vol. 2020, pp. 1–12, 2020.
 - [27] M. F. Guo, X.-D. Zeng, D.-Y. Chen, and N.-C. Yang, "Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant grounding distribution systems," *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1291–1300, 2018.
 - [28] P. Wang, Ananya, R. Yan, and R. X. Gao, "Virtualization and deep recognition for system fault classification," *Journal of Manufacturing Systems*, vol. 44, pp. 310–316, 2017.
 - [29] S.-K. Lee, "An acoustic decay measurement based on time-frequency analysis using wavelet transform," *Journal of Sound and Vibration*, vol. 252, no. 1, pp. 141–153, 2001.
 - [30] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, New York, NY, USA, 1999.
 - [31] S.-K. Lee, T.-G. Kim, and U. Lee, "Sound quality evaluation based on artificial neural network," *Lecture Notes in Computer Science*, Springer, vol. 4221, pp. 545–554, Berlin, Heidelberg, Germany, 2006.
 - [32] E.-Y. Kim, Y.-J. Lee, and S.-K. Lee, "Heath monitoring of a glass transfer robot in the mass production line of liquid crystal display using abnormal operating sounds based on wavelet packet transform and artificial neural network," *Journal of Sound and Vibration*, vol. 331, no. 14, pp. 3412–3427, 2012.
 - [33] B. Samanta, "Gear fault detection using artificial neural networks and support vector machines with genetic algorithms," *Mechanical Systems and Signal Processing*, vol. 18, no. 3, pp. 625–644, 2004.
 - [34] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
 - [35] Y. Lecun, K. Kavukcuoglu, E. Culurciello et al., "Large-scale FPGA-based convolutional networks," *Machine Learning on Very Large Data Sets*, pp. 1–26, Cambridge University Press, Cambridge, UK, 2011.
 - [36] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
 - [37] D. P. Kingma and L. J. Ba, "Adam: a method for stochastic optimization," in *Proceedings of the International Conference on Learning Representations*, pp. 1–13, San Diego, CA, USA, May 2015.
 - [38] A. C. Wilson, R. Roelofs, M. Stern, N. Srebro, and B. Recht, "The marginal value of adaptive gradient methods in machine learning," 2017, <http://arxiv.org/abs/1705.0829v2>.
 - [39] G. Hinton, "Neural networks for machine learning online course," <https://www.coursera.org/learn/neuralnetworks/home/welcome>.
 - [40] X. Liu, W. Wang, W. Sun, T. Wu, J. Liu, and J. Liu, "Design and experimental analyse of low noise double-pitch silent

- chain for conveyor,” *Procedia Engineering*, vol. 29, pp. 2146–2150, 2012.
- [41] L. Lefebvre and F. Laville, “Noise source identification for mechanical systems generating periodic impacts,” *Applied Acoustics*, vol. 69, no. 9, pp. 812–823, 2008.
- [42] H. Zheng, Y. Y. Wang, G. R. Liu et al., “Efficient modelling and prediction of meshing noise from chain drives,” *Journal of Sound and Vibration*, vol. 245, no. 1, pp. 133–150, 2001.
- [43] N. Fuglede and J. J. Thomsen, “Kinematic and dynamic modeling and approximate analysis of a roller chain drive,” *Journal of Sound and Vibration*, vol. 366, pp. 447–470, 2016.
- [44] J. He, S. Yang, and C. Gan, “Unsupervised fault diagnosis of a gear transmission chain using a deep belief network,” *Sensors*, vol. 17, no. 7, p. 1564, 2017.
- [45] Q. Hu, Z. He, Z. Zhang, and Y. Zi, “Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble,” *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 688–705, 2007.
- [46] L. Van der Maaten and G. Hinton, “Visualizing high- dimensional data using t-SNE,” *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 2008.