연구논문

Journal of Drive and Control, Vol.20 No.4 pp.54-63 Dec. 2023
ISSN 2671-7972(print) ISSN 2671-7980(online)
http://dx.doi.org/10.7839/ksfc.2023.20.4.054

Reliable Fault Diagnosis Method Based on An Optimized Deep Belief Network for Gearbox

Oybek Eraliev[1], Ozodbek Xakimov[2] and Chul-Hee Lee[3*]

Received: 18 Oct. 2023, Accepted: 16 Nov. 2023

Key Words: Fault diagnosis, Gearbox, Deep Belief Network, Particle Swarm Optimization, Classification

Abstract: High and intermittent loading cycles induce fatigue damage to transmission components, resulting in premature gearbox failure. To identify gearbox defects, numerous vibration-based diagnostics techniques, using several artificial intelligence (AI) algorithms, have recently been presented. In this paper, an optimized deep belief network (DBN) model for gearbox problem diagnosis was designed based on time-frequency visual pattern identification. To optimize the hyperparameters of the model, a particle swarm optimization (PSO) approach was integrated into the DBN. The proposed model was tested on two gearbox datasets: a wind turbine gearbox and an experimental gearbox. The optimized DBN model demonstrated strong and robust performance in classification accuracy. In addition, the accuracy of the generated datasets was compared using traditional ML and DL algorithms. Furthermore, the proposed model was evaluated on different partitions of the dataset. The results showed that, even with a small amount of sample data, the optimized DBN model achieved high accuracy in diagnosis.

Nomenclature

AI: Artificial intelligence
BP: Backpropagation
CD: Contrastive divergence
CNN: Convolutional neural network
DBN: Deep belief network
DE: Differential evolution
DL: Deep learning
DNN: Deep neural network
DT: Decision tree
EMD: Empirical mode decomposition
FFT: Fast Fourier transform
GA: Genetic algorithm
KNN: k-nearest neighbour
ML: Machine learning
NN: Neural network
PSO: Particle swarm optimization
RBM: Restricted Boltzmann machines
SSAE: Stacked sparse autoencoder
STFT: Short-time Fourier transform
SVM: Support vector machine
WT: Wavelet transform

* Corresponding author: gdhong@hankook.ac.kr
1 Department of Control & Mechanical Engineering, the Graduate School, Hankook University, Busan, 48547 Korea
2 Department of Mechanical Engineering, Jeju National University, Jeju 63243, Korea

1. Introduction

The three main terms used to describe the maintenance strategy for machinery parts are system health management, health monitoring, and fault identification and diagnosis. Mechanical transmission systems account for 30% of a machine's overall maintenance cost and the bulk of breakdowns in rotating machinery[1)]. The gearbox plays a key role in the transmission system that consists of gears, bearings and driving shafts[2-3)]. Any gearbox problems might result in unwelcome downtime, costly repairs, and possibly human casualties. Hence, it is critical to spot and diagnose problems at the early stage.

Different techniques, including vibration signal analysis, noise signal analysis, lubricant character analysis, and temperature monitoring, can be utilized to identify the problem with a machine. Vibration, acoustic, oil-based, electrical, and thermal signals can all be used to reflect gearbox conditions[4-9)]. Diagnostic by vibration signal is the most common method because it is assumed that every machine has a normal spectrum until something goes wrong, at which point the spectrum change [10-11)].

Normally, signal acquisition, signal preprocessing, extraction of features from signal, feature reduction/selection, and defect detection are the five basic processes in traditional fault detection systems based on machine learning (ML) models[12)]. The vibration signal-based approach has been shown to be useful in gathering signal data from machines operating at high speeds. However, extraneous signals from unrelated components frequently mask the vibration signal[13)]. Therefore,

filtering these undesirable signals necessitates signal preprocessing techniques. Deep learning (DL) models have simplified fault detection to three main processes in recent years: signal acquiring, signal processing, and failure identification[14]. The feature extraction and selection processes in DL models are automated because of the various hidden layers. Many applications of DL have been demonstrated, including medicine[15], speech recognition, natural language processing, and computer vision[16]. According to Zhao et al.[17], systems for defect diagnosis have frequently used DL models including deep belief networks (DBNs), stacked sparse autoencoders (SSAEs), and convolutional neural networks (CNNs). The specifics of the DL model variations have been covered by Guo et al.[18].

The performance of DL models improves as the quantity of the training dataset grows, which is a significant advantage over ML models. Once ML models reach a particular level of diagnosis performance, increasing the number of datasets has no effect on their performance. However, the proper setting up of DL models for specific applications is still a work in progress. Many hyperparameters in DL models must be manually specified, which takes time. The general hyperparameters of DL models are the number of epochs, learning rate, hidden node's number, number of hidden layers, and activation function. Hyperparameters are unique to each type of DL model. For instance, the pooling function, filter's number, padding, filter size, and filter stride are all features of CNN models[19]. Sadoughi and Hu[20] go into great length about the CNN hyperparameters. On the other hand, a DBN has a momentum factor and a learning rate that must be determined before beginning the training procedure[21]. Because of its capacity to extract significant deep features, the DBN has performed exceptionally well in multivariate data classification and prediction tasks among DL models. The DBN model can be used in a wide range of applications, from complex forecasting tasks to sensitive medical data prediction, thanks to restricted Boltzmann machines (RBMs) that can extract representative features from data. Furthermore, the DBN architecture contains less hyperparameter variables than the CNN architecture. In comparison to CNN, optimizing the DBN's hyperparameters is much simpler. For example, researchers[22-25]have conducted several studies on gearbox fault diagnosis using DBN. But the DBN's hyperparameters were manually chosen for the majority of the investigation to ensure precise fault diagnosis.

In this study, the suggested model is built on a DBN, which has various advantages such as strong representation of features and extraction capabilities, unsupervised pretraining, and simple architecture. The key contributions of this article are summarized as follows:

The proper hyperparameter selection has been performed on the DBN model by the use of the PSO algorithm. The learning rate and momentum factor are the important hyperparameters of the DBN model. The proper hyperparameter configuration depends on the application, and the grid search or manual hand tuning approach are typically used to choose hyperparameters.

The suggested DBN model has a strong capacity to extract and represent features. Because the vibration signal dataset is not always linear, the DBN is employed to find complex relationships between them.

There are two types of gearbox dataset and a variety of train–test partitions applied. It can be a chance to not only compare proposed optimized DBN model to other works such as ML algorithms, deep neural network (DNN), and CNN, but also learn how the model responds to training data of various sizes.

The following is how the rest of the article is structured. The background of the suggested technique is discussed in Section 2. The method of experimentation is described in Section 3. The outcome of the suggested model is talked about in Section 4. The conclusion is presented in Section 5.

2. Method

The diagnosis system used in this study is based on a combination of DBN, PSO, and spectrogram of short-time Fourier transform (STFT) images, as illustrated in Fig. 1. STFT method, which is described in section 2.3, is utilized in the gearbox datasets. The analysis uses a low-resolution image size of a 28×28 grayscale image as input data for fault diagnosis analysis. According to[26], a 28×28 image size is sufficient for defect identification analysis. The hyperparameters, including momentum factor and learning rate, for each RBM and neural network (NN) are optimized by the PSO algorithm. The fitness function is the classification error of the DBN model as expressed in Eq. (1).

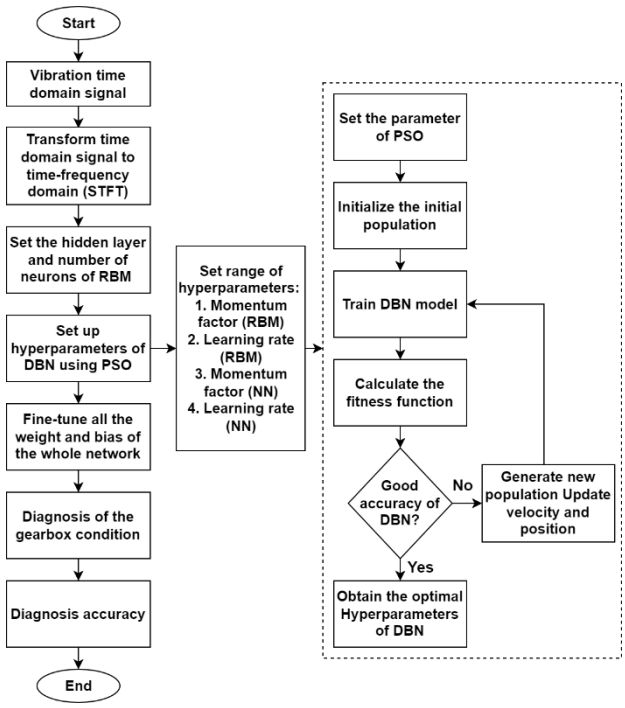

Fig. 1 Overall flowchart of the diagnosis system.

$$Test\ error = \frac{true_{value} - predicted_{value}}{total_{predictions}} \quad (1)$$

where, $true\_value$ denotes actual class number or ground truth, $predicted\_value$ denotes predicted class number by the model and $total\_predictions$ is the number of predictions performed by the model.

2.1 Deep Belief Network

One variety of deep neural networks is the DBN that is made up of numerous limited Boltzmann machines stacked on top of each other. The RBM is a two-layer generative neural network with a visible layer $v = (v_1, v_2, v_3, \dots, v_m)$ and a hidden layer $h = (h_1, h_2, h_3, \dots, h_n)$ both of which are built on probabilistic binary units. The basic construction of an RBM is shown in Fig. 2, with n and m units in the hidden and visible layers, respectively.

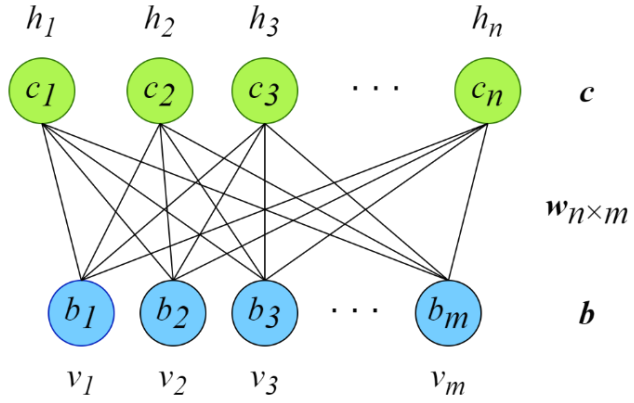

Fig. 2 Basic structure of a RBM with m visible nodes and n hidden nodes.

An RBM is an energy-based model in which each layer's unit has an energy value assigned to it. As shown in Eq. (2), the energy function is defined as follows:

$$E(\boldsymbol{v}, \boldsymbol{h}) = -\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij} v_i h_j - \sum_{i=1}^{m} b_i v_i - \sum_{i=1}^{n} c_j h_j \quad (2)$$

where w_{ij} is the symmetric weight, v_i, h_j are the binary states, and b_i, c_j are their biases, respectively.

Pre-training and fine-tuning phases make up the RBM's training technique. Unsupervised learning is used to extract a deep hierarchical representation of training data in the pre-training phase by choosing the proper weights and biases, namely w, b and c. The pre-training phase's major goal is to find the best parameter set w, b and c, so that the energy function in Eq. (2) minimizes to a convergent state. It is more likely that the association $(\boldsymbol{v}, \boldsymbol{h})$ with lesser energy exists. As a result, each correlation's likelihood is inversely proportional to its energy function:

$$P(\boldsymbol{v}, \boldsymbol{h}) = \frac{e^{-E(\boldsymbol{v},\boldsymbol{h})}}{\sum_{\boldsymbol{v},\boldsymbol{h}} e^{-E(\boldsymbol{v},\boldsymbol{h})}} \quad (3)$$

where $\sum_{\boldsymbol{v},\boldsymbol{h}} e^{-E(\boldsymbol{v},\boldsymbol{h})}$ represents a normalization factor. Hinton[27] has suggested contrastive divergence (CD-k) as a faster training method for the DBN because it is computationally costly. According to the algorithm, $P(\boldsymbol{h}|\boldsymbol{v})$ is used to forecast hidden values h for a given input vector v. $P(\boldsymbol{v}|\boldsymbol{h})$ is then used to anticipate new visible units, v, based on the hidden values obtained. The conditional probabilities are computed using Eqs. (4) and (5), which are also known as alternating Gibbs sampling:

$$P(\boldsymbol{v} = \boldsymbol{1}|\boldsymbol{h}) = Sigmoid\left(\sum_{j=1}^{n} w_{ij}^T h_j + b_i\right) \quad (4)$$

$$P(\boldsymbol{h} = \boldsymbol{1}|\boldsymbol{v}) = Sigmoid\left(\sum_{i=1}^{m} w_{ij}^T v_i + c_j\right) \quad (5)$$

where the sigmoid function is expressed as follows:

$$f_S(x) = Sigmoid(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

Samples $(\boldsymbol{v}^{(k)}, \boldsymbol{h}^{(k)})$ are guaranteed to be practical samples of $P(v = 1|h)$ when the value of k is high enough. One-step sampling (CD-1) is used in this investigation since it is good enough to offer a realistic estimate. Computing the differences between data and reconstruction values yields the reconstruction error (ΔW) and the weight matrix update, as stated in Eqs. (7) and (8):

$$\Delta W = \frac{1}{m}\sum_{i,j,m}\left(v_{mi}h_{mj}^{(0)} - v_{mi}h_{mj}^{(1)}\right) \quad (7)$$

$$W \leftarrow \gamma_p W + \alpha_p \Delta W \quad (8)$$

where m represents number of samples, γ_p is a momentum factor and α_p is the learning rate. Correspondingly, biases b and c can be updated by convergently running one of these chains, as expressed in Eqs. (9) and (10):

$$b = \gamma_p b + \alpha_p \frac{1}{m}\sum_{m=1}^{m}\left(v_{mi}^{(0)} - v_{mi}^{(1)}\right) \quad (9)$$

$$c = \gamma_p c + \alpha_p \frac{1}{m}\sum_{m=1}^{m}\left(v_{mj}^{(0)} - v_{mj}^{(1)}\right) \quad (10)$$

Following the pre-training phase can be used to fine-tune the parameter sets W, b and c using the algorithm of backpropagation (BP). The training is done in this step in the same way that a NN is trained. Weights that are bidirectional can be transformed to unidirectional weights. The model training is substantially faster because the

parameters have already been initialized by unsupervised learning. Equations (11) and (12) show the parameter modifications in the NN:

$$W_{NN} \leftarrow \gamma_f W_{NN} - \alpha_f \frac{\partial Err}{\partial W_{NN}} \tag{11}$$

$$a \leftarrow \gamma_f a - \alpha_f \frac{\partial Err}{\partial W_{NN}} \tag{12}$$

where W_{NN} presents the weight of the NN, γ_p and α_p present the momentum factor and the learning rate, respectively, and a is a bias for the fine-tuning phase.

The overall structure of the DBN used in this study is illustrated in Fig. 3.

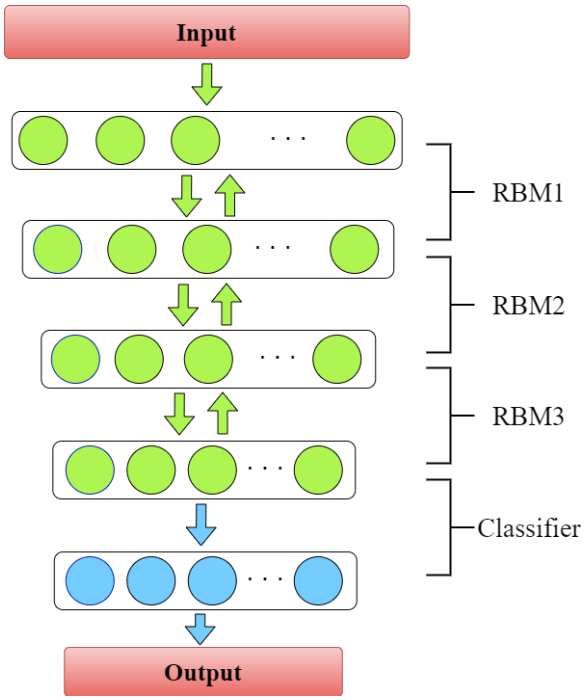

Fig. 3 Structure of a deep belief network

2.2 Particle Swarm Optimization

With the advancement of machine intelligence, approaches such as random search, PSO, grid search, differential evolution (DE) and genetic algorithm (GA) can now be used to optimize hyperparameters of deep learning models. M. M. A. Wahab and et. al.[28]have investigated optimization strategies in depth, and the analysis revealed that PSO approach is alike to the DE approach.

The PSO method is made up of three key parameters: an overall experience (Gbest), a personal experience (Pbest), and the current movement of the particles in the search area to identify their next places.

Based on the data, each particle adjusts its velocity and position as follow:

$$v_{id}^{k+1} = w \cdot v_{id}^{k} + c_1 \cdot rand_1 \cdot \left(pbest_{id}^{k} - p_{id}^{k}\right) + c_2 \cdot rand_2 \cdot \left(gbest_{id}^{k} - p_{id}^{k}\right) \quad and \tag{14}$$

$$p_{id}^{k+1} = p_{id}^{k} + v_{id}^{k+1}$$

where v represents the velocity of the ith particle and d expresses the a particle's size in the kth iteration $(1 \le d \le n)$, $rand_1$ and $rand_2$ are random numbers in the range of [0, 1], c_1 and c_2 indicate personal and social learning factors, which are positive numbers, respectively and w represents the weight of inertia. It is possible to balance the capacities of global and local exploration by using equation (15).

$$w_k = w_{max} - \frac{w_{max} - w_{min}}{t_{max}} \cdot t \tag{15}$$

where w_{max} indicates the weight of maximum inertia, w_{min} indicates the weight of maximum inertia and t is the present iteration. The swarm size has been set to 5 and iterations have been set to 30 during the analysis.

2.3 Time-Frequency Transformation

Time and frequency representations of the machinery signal's energy can be revealed using time-frequency analysis. For machinery defect diagnostics, various time-frequency transformations have been utilized, including empirical mode decomposition (EMD), wavelet transform (WT), STFT, and spectral kurtosis diagram (kurtogram). Among them, the STFT is frequently employed in time-frequency analysis of a row vibration signal[29-31]. The STFT method transforms a time-domain signal to a time-frequency-domain signal. It divides the large time-series signal into several the same size of pieces using windowing function and calculates fast Fourier transform coefficients (FFTs) for each piece. A matrix is used to store the calculated values that correspond to the segments. One of the benefits of STFT is that its results can be used immediately in the training of ML models. The STFT is expressed mathematically as follow:

$$STFT(\tau, w) = \int_{-\infty}^{\infty} s(t) w(t-\tau) e^{-j\omega t} dt \tag{16}$$

where $s(t)$ represents the time domain vibration signal, $w(t)$ represents a windowing function, t represents time, and τ represents time index. STFT supports a variety of windowing functions, including triangular, rectangular, Hamm, Blackman, Kaiser, Gaussian, and Hann. When compared to the impacts of other windows, Hann windowing is often a nice select and is widely used with random data since it has a negligible effect on the frequency resolution and amplitude correctness of the resulting frequency spectrum[32]. As a result, in this investigation, the Hann windowing function is applied, which can be written as follows:

$$h(t) = \begin{cases} 0.5\left[1 - \cos\left(\frac{2\pi n}{M-1}\right)\right], 0 \le n \le M-1 \\ 0, otherwise \end{cases} \tag{17}$$

where n is the time index and M is the number of samples.

The librosa open-source package in Python is used to conduct STFT on the vibration signal. The output is an $n_{n\_fft} \times m_s$ matrix with complex numbers, and the output's absolute values are used. The number of FFT coefficients for each window (n_fft = 1024) is given by n_fft. The following formula is used to determine n_{n_fft}:

$$n_{n_{fft}} = \frac{n_{fft}}{2} + 1 \tag{18}$$

The hop size is set to 512. Following that, the spectrogram of the data is converted to 28x28 grayscale images. Each experiment has been repeated 20 times, with the average results have been used in the study.

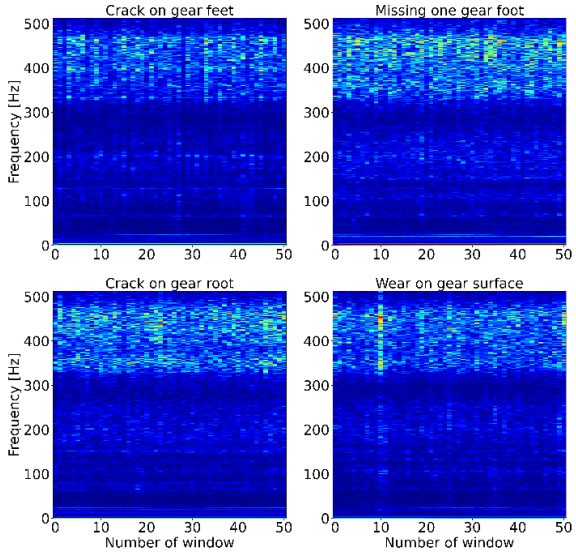

Fig. 4 Spectrogram of the vibration data (20 Hz).

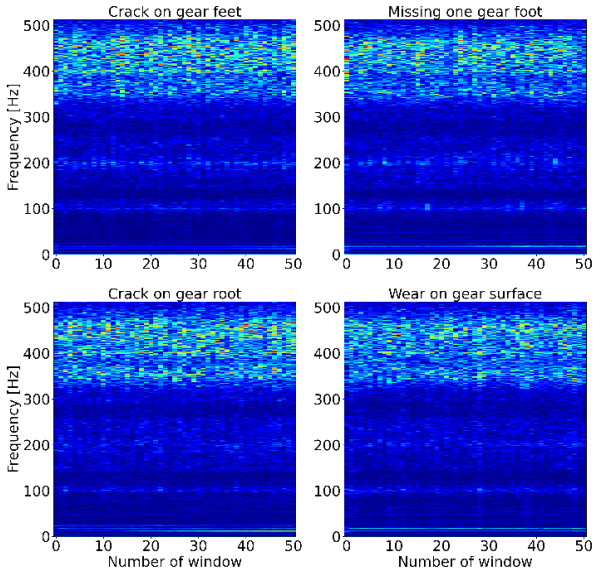

Fig. 5 Spectrogram of the vibration data (30 Hz).

By using STFT analysis and sampling frequency data, the gearbox's time-domain signal can be converted into a time-frequency signal. The spectrograms for both the 20 Hz and 30 Hz datasets are shown in Fig. 4 and Fig. 5, respectively.

2.4 ML and DL Models Used for Comparative Study

Support Vector Machine (SVM): SVM is a powerful supervised learning algorithm used for classification and regression tasks. It operates by finding the hyperplane that best separates data points belonging to different classes. SVM aims to maximize the margin between classes, defined as the distance between the hyperplane and the nearest data point of each class. SVM can handle high-dimensional data effectively and is particularly useful when the data is not linearly separable by transforming it into a higher-dimensional space using the kernel trick.

Decision Tree (DT): DTs are versatile models used for classification and regression. They represent decisions in a tree-like structure, where each node represents a decision based on a feature. DTs make decisions by recursively splitting the dataset based on features, creating a tree structure. The splits are chosen to maximize information gain or minimize impurity. One of the key advantages of DTs is their interpretability, allowing users to easily understand and visualize the decision-making process.

k-Nearest Neighbors (KNN): KNN is a straightforward and intuitive algorithm for classification and regression. It classifies a data point based on the majority class of its k nearest neighbors in the feature space. KNN is a non-parametric and instance-based algorithm, making predictions directly from the training dataset. It is effective for tasks with complex decision boundaries. The choice of the parameter k influences the algorithm's performance, with smaller values making the model more sensitive to local fluctuations and larger values leading to smoother decision boundaries.

DNN: DNN employed in this study was meticulously configured to capture intricate patterns within the dataset. The model was trained over 100 epochs, utilizing the Adam optimizer with a learning rate set at 0.001. The architecture comprises two hidden layers, each housing 100 neurons. This configuration was chosen to strike a balance between model complexity and efficiency. The training process involved the iterative refinement of internal parameters through the backpropagation algorithm, enabling the DNN to learn hierarchical representations of features within the data. The use of two hidden layers with a considerable number of neurons facilitated the model's capacity to discern nuanced relationships present in the vibration signal dataset.

CNN: CNN employed in our research was intricately designed to extract spatial hierarchies from grid-structured data, particularly suitable for our vibration signal dataset. The model underwent training for 100 epochs, employing the Adam optimizer with a learning rate of 0.001. The architectural design of the CNN is characterized by two convolutional layers, comprising 64

and 32 filters, respectively. Subsequently, two fully connected layers, each with 50 neurons, were incorporated to capture higher-level abstractions. This architectural choice was made to ensure the effective extraction of hierarchical features, promoting adaptability to the diverse patterns inherent in gearbox fault signals. Throughout the training process, the CNN iteratively learned relevant filters and weights through backpropagation, enhancing its capability to recognize complex patterns within the input data. The combination of convolutional and fully connected layers facilitated the model's ability to discern spatial hierarchies and relationships, contributing to its robust performance in fault diagnosis tasks.

3. Experimental Procedure

Two acceleration datasets, including an experimental gearbox and a wind turbine gearbox, are used in this study. Fig. 6 presents the experimental setup, and it is detailed in[33].

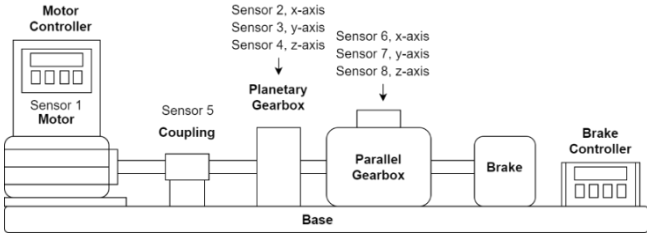

Fig. 6 Scheme of an experimental setup of the gearbox

The analysis in this paper mainly focused on gear issues. Five separate gear's row vibration signals are gathered during the data collecting process: healthy, crack on feet of gear, missing one foot of gear, wear on surface of gear and crack on gear root. For the fault diagnosis analysis, each gear situation is assigned a class. The test setup has eight sensors. The gearbox runs at 20 and 30 Hz. Sensors 1is for motor, while sensor 5 is torque sensors mounted on coupling part. And acceleration sensors 2, 3, and 4 are mounted on the planetary gearbox, while the rest of acceleration sensors are mounted on the parallel gearbox. Therefore, only sensors 1 and 5 are not employed to evaluate the suggested diagnosis approach. 2 000 Hz is used as the sampling frequency.

4. Results and Discussion

Instead of using a system that combines many sensors, this study examines the model that is suggested on a single sensor. Because not all sensors in practical applications offer a reliable signal for analysis, a multi-sensor fusion method may result in an appropriate diagnosis. As a result, the primary goal is to perform accurate problem identification with a single sensor. The impact of the DBN's hidden layer size on diagnosis performance is examined in this article, as the number of RBM is one of the elements that can lead the network to overfitting.

Fig. 7 and Fig. 8 depict the proposed DBN model's diagnosis performance at 20 and 30 Hz, respectively. From the graphs, the DBN with 2 RBMs has a slightly lower performance compared to the rest of the DBNs. There is no discernible difference in the performance of DBNs with 3 RBMs and 4 RBMs. Therefore, the DBN with 3 RBMs is chosen for diagnosis because it takes less time to train for parameter and hyperparameter tuning. Overall accuracy of the optimized DBN is over 99.9% on the 20Hz dataset as shown in Fig. 7.

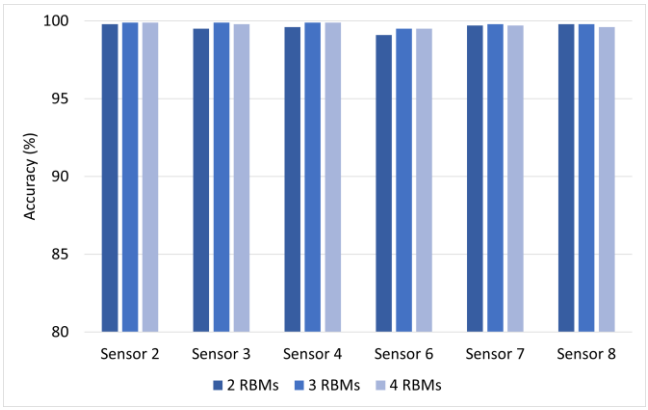

Fig. 4 Diagnostic accuracy of the optimized DBN with different layer size (20 Hz).

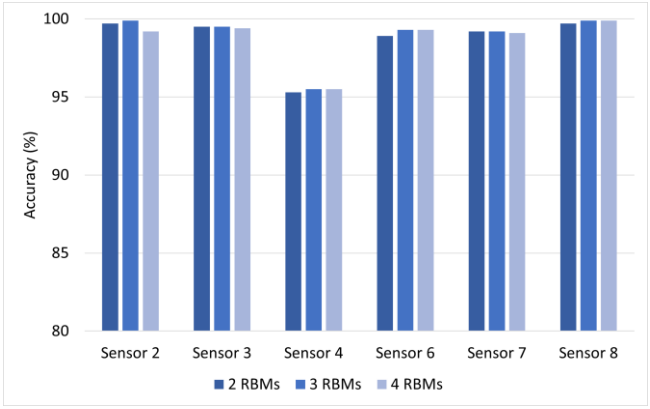

Fig. 5 Diagnostic accuracy of the optimized DBN with different layer size (30 Hz).

The optimized DBN is also tested on the 30 Hz dataset. Fig. 8 illustrates the results of the test. The proposed model attains maximum level of performance accuracy of 99.9% for sensors 2 and 8, while the lowest performance accuracy of 95% is observed in sensor 4. The overall performance is over at 95%, which is slightly lower than that of the 20 Hz dataset.

4.1 Comparative Study of Traditional ML And DL Models

Traditional fault diagnosis has been performed for the datasets. Each equation listed in Table 1 is considered as one feature. So, overall, thirteen statistical features are obtained from each 1024 data points of the row vibration

signal. These feature vectors are used as an input to traditional ML and DL models.

Table 1 Features from the row vibration signal.

Parameter	Equation
Mean	$T1 = \frac{\sum_{n=1}^{N} s(t)}{N}$
Standard deviation	$T2 = \sqrt{\frac{\sum_{n=1}^{N}(s(t)-T1)^2}{N-1}}$
Variance	$T3 = \left(\frac{\sum_{n=1}^{N}\sqrt{\lvert s(t)\rvert}}{N}\right)^2$
RMS	$T4 = \sqrt{\frac{\sum_{n=1}^{N}(s(t))^2}{N}}$
Absolute maximum	$T5 = \max \lvert s(t)\rvert$
Coefficient of skewness	$T6 = \frac{\sum_{n=1}^{N}(s(t)-T1)^3}{(N-1)(T2)^3}$
Kurtosis	$T7 = \frac{\sum_{n=1}^{N}(s(t)-T1)^4}{(N-1)(T2)^4}$
Crest factor Margin	$T8 = \frac{T5}{T4}$
Margin factor	$T9 = \frac{T5}{T3}$
Shape factor	$T10 = \frac{T4}{\frac{1}{N}\sum_{n=1}^{N}\lvert s(t)\rvert}$
Impulse factor	$T11 = \frac{T5}{\frac{1}{N}\sum_{n=1}^{N}\lvert s(t)\rvert}$
A factor	$T12 = \frac{T5}{T2 \cdot T3}$
B factor	$T13 = \frac{T7 \cdot T8}{T2}$

where $s(t)$ represents the time domain vibration signal.

In this investigation, SVM, DT and KNN traditional ML classifiers are used. Gini diversity index is used for the split criterion in DT model. For the SVM model, a radial basis kernel function is used, while KNN model's hyperparameter is Euclidean distance. Thereafter, two DL models such as DNN and CNN are chosen for comparative analysis. Six different categories of data are obtained from sensors at various positions and are used to test the models. All results of the comparative study for 20 Hz and 30 Hz datasets are listed in Table 2 and Table 3 respectively. It can be noted that from Table 2 and 3 there is no notable differences in the diagnostic accuracy of the traditional ML models. The SVM model, which uses the sensor 2 input for the 20 Hz data set and the sensor 6 input for the 30 Hz data set, achieves the best diagnosis results. The models' accuracy varies from 67% to 85%. The traditional ML models generate inaccurate diagnosis results when compared to the proposed DBN model system.

On the other hand, a slight increase is observed in DL models, especially the CNN model, when compared to traditional ML models. The CNN model achieves the highest performance accuracy in the sensor 2 input for the 20 Hz data set and the sensor 6 input for the 30 Hz data set. Due to its poor performance on sensor 4 (30 Hz), DNN is inadequate for these data sets. Meanwhile, the proposed optimized DBN model achieves 95.5% to 99.9% accuracy, which is higher than that of traditional ML and DL models by 10 to 20%.

Table 2 Comparative diagnosis performance of ML and DL models (20 Hz)

Sensors	DT	SVM	KNN	DNN	CNN	Proposed DBN
Sens. 2	85.2%	85.3%	79.8%	80.3%	87.8%	99.9%
Sens. 3	84.8%	85.1%	80.1%	86.6%	81.5%	99.9%
Sens. 4	84.9%	84.9%	83.1%	84.5%	84.1%	99.9%
Sens. 6	77.3%	84.8%	75.2%	81.1%	77.3%	99.5%
Sens. 7	80.1%	83.7%	80.2%	75.4%	81.6%	96.8%
Sens. 8	79.8%	82.9%	80.0%	79.6%	85.2%	99.8%

Table 3 Comparative diagnosis performance of ML and DL models (30 Hz)

Sensors	DT	SVM	KNN	DNN	CNN	Proposed DBN
Sens. 2	82.1%	82.5%	80.3%	77.5%	80.6%	99.9%
Sens. 3	77.8%	82.2%	78.6%	66.9%	79.7%	99.5%
Sens. 4	76.9%	79.9%	77.3%	80.5%	81.4%	95.5%
Sens. 6	77.2%	83.1%	81.8%	76.2%	82.3%	99.3%
Sens. 7	69.8%	67.5%	69.2%	63.8%	78.7%	99.2%
Sens. 8	77.7%	79.8%	78.2%	65.9%	84.0%	99.9%

To learn how the model responds to training data of various sizes, the datasets are split into different train-test samples, including 80%-20%, 50%-50% and 20%-80% and the optimized DBN model is evaluated. The result of the test is shown in Table 4. According to the table, there is no significant difference between all partitions in diagnosis accuracy of the proposed optimized DBN model. The lowest diagnosis performance accuracy of 91.5% is observed for sensor 4 of the 30 Hz dataset, while the highest trend is seen for sensor 3 of the 30 Hz dataset, when the dataset makes up a partition of 20%-80%. It means that the model can also perform satisfaction diagnosis performance on a small dataset. It has been proven that the optimized DBN model has a strong capacity to extract and represent features.

Table 4 Diagnostic accuracy of the proposed DBN model in different partitions of the dataset

Sensors	80%-20% (20 Hz)	50%-50% (20 Hz)	20%-80% (20 Hz)	80%-20% (30 Hz)	50%-50% (30 Hz)	20%-80% (30 Hz)
Sens. 2	82.1%	82.5%	80.3%	77.5%	80.6%	99.9%
Sens. 3	77.8%	82.2%	78.6%	66.9%	79.7%	99.5%
Sens. 4	76.9%	79.9%	77.3%	80.5%	81.4%	95.5%
Sens. 6	77.2%	83.1%	81.8%	76.2%	82.3%	99.3%
Sens. 7	69.8%	67.5%	69.2%	63.8%	78.7%	99.2%
Sens. 8	77.7%	79.8%	78.2%	65.9%	84.0%	99.9%

4.2 Analysis On Wind Turbine Gearbox

In order to provide strong evidence, the proposed optimized DBN model is tested on the actual gearbox of a wind turbine. There are two different types of row vibration signals in the dataset: healthy and unhealthy

states. During the data collection operation, the wind turbine gearbox is operated at 30 Hz[34]. The results of this study are shown in Fig. 9. The proposed optimized DBN model has achieved the highest diagnosis performance of 99.9%. Although the data acquired from actual structure was generally substantially masked by undesirable noise, making processing challenging, the performance of the proposed DBN model for diagnosis can be found to be good. This demonstrates the robustness of the suggested optimized DBN model.

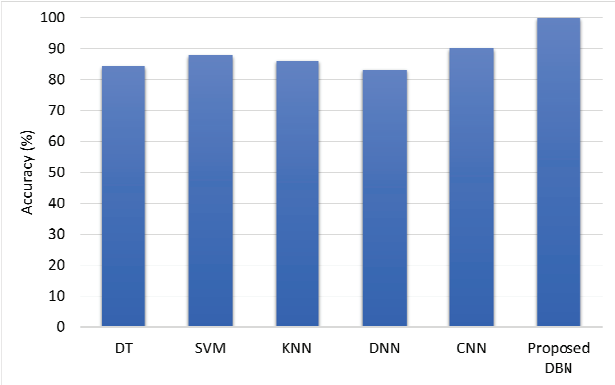

Fig. 6 Performance of the classifiers on wind turbine gearbox dataset.

The superior performance of the optimized DBN model compared to traditional ML models, as well as DNN and CNN architectures, can be attributed to several key factors.

Hierarchical feature learning: DBN excels in automatically learning hierarchical representations of features from the input data. The model's ability to capture intricate patterns and dependencies in time-frequency visual patterns of gearbox vibrations is crucial for effective fault diagnosis. Traditional ML models may struggle to automatically extract such complex hierarchical features, and even DNN and CNN architectures might not inherently capture them without specialized design considerations.

Unsupervised pre-training: DBN employs an unsupervised pre-training phase, allowing the model to learn latent representations of the input data before fine-tuning on the specific diagnostic task. This pre-training strategy can enhance the model's ability to generalize well to new and unseen data, a feature that may contribute to its improved performance compared to models that lack this pre-training phase.

Efficient representation of non-linear relationships: The inherent architecture of DBN, with its stacked layers of stochastic latent variables, enables efficient representation of non-linear relationships within the data. This is particularly advantageous for capturing the intricate and non-linear nature of gearbox fault patterns, which may be challenging for traditional ML models that rely on linear assumptions.

PSO for hyperparameter tuning: The integration of PSO for hyperparameter tuning in the optimization process further refines the DBN model's performance. This adaptive optimization technique allows the model to navigate the complex hyperparameter space effectively, resulting in a more finely tuned and optimized network compared to conventional ML models and simpler deep learning architectures.

Versatility and adaptability: DBN's versatility in handling diverse datasets, such as those from wind turbine and experimental gearboxes, highlights its adaptability to different operational conditions. Traditional ML models and even some DNN and CNN architectures may struggle to exhibit such adaptability without extensive customization.

5. Conclusion

In conclusion, this study has successfully demonstrated the effectiveness of an optimized DBN model in diagnosing gearbox problems induced by high and intermittent loading cycles. The incorporation of time-frequency visual pattern identification, coupled with PSO for hyperparameter tuning, has yielded a robust and high-performing diagnostic tool. The application of the proposed model to wind turbine and experimental gearbox datasets showcased its versatility and reliability in real-world scenarios.

The academic significance of this research lies in the advancement of gearbox diagnostics through the fusion of vibration-based techniques and artificial intelligence algorithms. The optimized DBN model has shown superior classification accuracy when compared to traditional ML and DL algorithms, underscoring its efficacy in addressing the challenges posed by fatigue damage in transmission components.

However, it is essential to acknowledge certain limitations in this study. The evaluation of the proposed model was conducted on a specific set of gearbox datasets, namely the wind turbine gearbox and the experimental gearbox. While these datasets provide valuable insights, the generalizability of the model to a broader range of gearbox types and operational conditions should be explored in future research. Additionally, the study's focus on high accuracy with a small amount of sample data raises questions about the model's performance scalability to larger and more diverse datasets.

Future research endeavors could enhance the applicability of the optimized DBN model by expanding the dataset variety and incorporating more complex real-world conditions. Moreover, investigating the model's performance under varying levels of noise, uncertainties, and fault severity would contribute to a more comprehensive understanding of its robustness. Exploring the integration of additional sensor modalities and considering the practical implementation of the proposed diagnostic tool in industrial settings would further validate its efficacy.

Acknowledgement

This research was supported by a grant(2023-MO IS35-005) of Policy-linked Technology Development Program on Natural Disaster Prevention and Mitigati on funded by Ministry of Interior and Safety (MOIS, Korea).

Conflicts of Interest

The authors declare that there is no conflict of interest.

References

1) F. Elasha, M. Greaves, and D. Mba, “Planetary bearing defect detection in a commercial helicopter main gearbox with vibration and acoustic emission,” Struct. Heal. Monit., vol. 17, no. 5, pp. 1192–1212, 2018, doi: 10.1177/1475921717738713.2) S. Butzmann and J. Melbert, "Sensorless Control of Electromagnetic Actuators for Variable Train," SAE Paper 2000-01-1225, pp. 325-423, 2000.
2) Y. Wang, S. Yang, and R. V. Sanchez, “Gearbox fault diagnosis based on a novel hybrid feature reduction method,” IEEE Access, vol. 6, pp. 75813–75823, 2018, doi: 10.1109/ACCESS.2018.2882801.
3) Y. B. Lee, G. C. Lee, J. J. Lee, S. Y. Lim, W. J. Kim, and K. M. Kim, “ A Study on the Acceleration Durability Test of In-Wheel Drive Gearbox for Military Special Vehicles,” vol. 19, no. 3, pp. 32–38, 2022.
4)Z. Chen, C. Li, and R. Sanchez, “Gearbox Fault Identification and Classification with Convolutional Neural Networks,” Shock Vib., vol. 2015, pp. 1–10, 2015, doi: 10.1155/2015/390134.
5) C. Li and M. Liang, “Extraction of oil debris signature using integral enhanced empirical mode decomposition and correlated reconstruction,” Meas. Sci. Technol., vol. 22, no. 8, 2011, doi: 10.1088/0957-0233/22/8/085701.
6) A. M. D. Younus and B. S. Yang, “Intelligent fault diagnosis of rotating machinery using infrared thermal image,” Expert Syst. Appl., vol. 39, no. 2, pp. 2082–2091, 2012, doi: 10.1016/j.eswa.2011.08.004.
7) T. Toutountzakis, C. K. Tan, and D. Mba, “Application of acoustic emission to seeded gear fault detection,” NDT E Int., vol. 38, no. 1, pp. 27–36, 2005, doi: 10.1016/j.ndteint.2004.06.008.
8) J. R. Ottewill and M. Orkisz, “Condition monitoring of gearboxes using synchronously averaged electric motor signals,” Mech. Syst. Signal Process., vol. 38, no. 2, pp. 482–498, 2013, doi: 10.1016/j.ymssp.2013.01.008.
9) J. W. Park, S. Han, H. S. Lee, and S. Yun, “A Study on the Hydraulic Vibration Characteristics of the Prefill Check Valve-monitoring,” vol. 18, no. 3, pp. 8–15, 2021.
10) H. S. Baek, J. H. Shin, and S. J. Kim, "Development of AI-Based Condition Monitoring System for Failure Diagnosis of Excavator ’ s Travel Device,” vol. 18, no. 1, pp. 24–30, 2021.
11) Y. K. Kang and J. S. Jang, "Feasibility Study on the Vibration Reduction for Hydraulic Breaker by the Dynamic Vibration Absorber,” vol. 18, no. 4, pp. 65–71, 2021.
12) S. Khan and T. Yairi, “A review on the application of deep learning in system health management,” Mech. Syst. Signal Process., vol. 107, pp. 241–265, 2018, doi: 10.1016/j.ymssp.2017.11.024.
13) T. Wang, Q. Han, F. Chu, and Z. Feng, “Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review,” Mech. Syst. Signal Process., vol. 126, pp. 662–685, 2019, doi: 10.1016/j.ymssp.2019.02.051.
14) S. R. Saufi, Z. A. Bin Ahmad, M. S. Leong, and M. H. Lim, “Low-Speed Bearing Fault Diagnosis Based on ArSSAE Model Using Acoustic Emission and Vibration Signals,” IEEE Access, vol. 7, pp. 46885–46897, 2019, doi: 10.1109/ACCESS.2019.2909756.
15) J. Xu, L. Xiang, R. Hang, and J. Wu, “Stacked Sparse Autoencoder (SSAE) based framework for nuclei patch classification on breast cancer histopathology,” 2014 IEEE 11th Int. Symp. Biomed. Imaging, ISBI 2014, no. 61273259, pp. 999–1002, 2014, doi: 10.1109/isbi.2014.6868041.
16) L. Deng, G. Hinton, and B. Kingsbury, “New types of deep neural network learning for speech recognition and related applications: An overview,” ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc., pp. 8599–8603, 2013, doi: 10.1109/ICASSP.2013.6639344.
17) R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, “Deep learning and its applications to machine health monitoring,” Mech. Syst. Signal Process., vol. 115, pp. 213–237, 2019, doi: 10.1016/j.ymssp.2018.05.050.
18) Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, “Deep learning for visual understanding: A review,” Neurocomputing, vol. 187, pp. 27–48, 2016, doi: 10.1016/j.neucom.2015.09.116.
19) W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, “A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load,” Mech. Syst. Signal Process., vol. 100, pp. 439–453, 2018, doi: 10.1016/j.ymssp.2017.06.022.
20) M. Sadoughi and C. Hu, “Physics-Based Convolutional Neural Network for Fault Diagnosis of Rolling Element Bearings,” IEEE Sens. J., vol.

19, no. 11, pp. 4181–4192, 2019, doi: 10.1109/JSEN.2019.2898634.
21) B. Ibrokhimov, C. Hur, H. Kim, and S. Kang, "A-DBNF: adaptive deep belief network framework for regression and classification tasks," Appl. Intell., vol. 51, no. 7, pp. 4199–4213, 2021, doi: 10.1007/s10489-020-02050-2.
22) H. Chen, J. Wang, B. Tang, K. Xiao, and J. Li, "An integrated approach to planetary gearbox fault diagnosis using deep belief networks," Meas. Sci. Technol., vol. 28, no. 2, 2017, doi: 10.1088/1361-6501/aa50e7.
23) H. Zheng and Y. Dai, "Fault Prediction of Fan Gearbox Based on Deep Belief Network," J. Phys. Conf. Ser., vol. 1449, no. 1, 2020, doi: 10.1088/1742-6596/1449/1/012050.
24) J. Yu and G. Liu, "Knowledge extraction and insertion to deep belief network for gearbox fault diagnosis," Knowledge-Based Syst., vol. 197, p. 105883, 2020, doi: 10.1016/j.knosys.2020.105883.
25) Y. Qin, X. Wang, and J. Zou, "The Optimized Deep Belief Networks with Improved Logistic Sigmoid Units and Their Application in Fault Diagnosis for Planetary Gearboxes of Wind Turbines," IEEE Trans. Ind. Electron., vol. 66, no. 5, pp. 3814–3824, 2019, doi: 10.1109/TIE.2018.2856205.
26) D. Verstraete, A. Ferrada, E. L. Droguett, V. Meruane, and M. Modarres, "Deep Learning Enabled Fault Diagnosis Using Time-Frequency Image Analysis of Rolling Element Bearings," Shock Vib., vol. 2017, pp. 1–17, 2017, doi: 10.1155/2017/5067651.
27) G. H.-N. computation and undefined 2002, "Training products of experts by minimizing contrastive divergence," ieeexplore.ieee.org, Accessed: Jun. 06, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6789337/
28) M. N. Ab Wahab, S. Nefti-Meziani, and A. Atyabi, "A comprehensive review of swarm optimization algorithms," PLoS One, vol. 10, no. 5, pp. 1–36, 2015, doi: 10.1371/journal.pone.0122827.
29) J. Y. Lee, "Variable short-time Fourier transform for vibration signals with transients," JVC/Journal Vib. Control, vol. 21, no. 7, pp. 1383–1397, 2015, doi: 10.1177/1077546313499389.
30) O. Eraliev, K.-H. Lee, and C.-H. Lee, "Vibration-Based Loosening Detection of a Multi-Bolt Structure Using Machine Learning Algorithms," Sensors, vol. 22, no. 3, p. 1210, 2022, doi: 10.3390/s22031210.
31) J. Huang, B. Chen, B. Yao, and W. He, "ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network," IEEE Access, vol. 7, pp. 92871–92880, 2019, doi: 10.1109/ACCESS.2019.2928017.
32) "Encyclopedia of Vibration | ScienceDirect." https://www.sciencedirect.com/referencework/9780122270857/encyclopedia-of-vibration (accessed Jan. 20, 2022).
33) S. Shao, S. McAleer, R. Yan, and P. Baldi, "Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning," IEEE Trans. Ind. Informatics, vol. 15, no. 4, pp. 2446–2455, Apr. 2019, doi: 10.1109/TII.2018.2864759.
34) S. R. Saufi, Z. A. Bin Ahmad, M. S. Leong, and M. H. Lim, "Gearbox Fault Diagnosis Using a Deep Learning Model With Limited Data Sample," IEEE Trans. Ind. Informatics, vol. 16, no. 10, pp. 6263–6271, Oct. 2020, doi: 10.1109/TII.2020.2967822.