Research on fault diagnosis of time-domain vibration signal based on convolutional neural networks

Mingyong Li , Qingmin Wei , Hongya Wang and Xuekang Zhang

ABSTRACT
In order to maintain the safe operation of various types of equipment, the health status of main components should be monitored in real-time, and the demand for intelligent fault diagnosis algorithm has increased sharply. However, the traditional intelligent diagnosis algorithm is based on the manual method for signal feature extraction, which has high requirements for expert experience and poor generality. A convolutional neural network, with big data as its engine, is the most effective pattern classification algorithm at present. In this paper, the convolutional neural network is applied to time-domain vibration signal fault diagnosis, taking the bearing as an example, and an intelligent diagnosis method of bearing based on convolution neural network is proposed. The proposed method does not need manual feature extraction, and can automatically complete feature extraction and automatic fault recognition. The convolutional neural network has three convolutional layers. We use data enhancement techniques for the input raw signal and convert the one-dimensional original time-domain vibration signal into a two-dimensional signal. The model shows good results on CWRU dataset and the recognition accuracy of the algorithm in the CWRU bearing database is more than 96%.

1. Introduction
With more and more modern electromechanical equipment, monitoring sensors on electromechanical equipment will obtain massive data, with the increase of the number and sampling frequency of equipment sensors, mechanical health monitoring has entered the era of big data. In the era of electromechanical big data, we need to use new technologies and new methods to automatically mine features from massive data, gradually replace experts to extract features, carry out real-time detection on important parts such as bearings, and ensure the accuracy and efficiency of fault diagnosis and prediction.

Bearing fault diagnosis is a hot research direction of mechanical fault diagnosis. The core of the algorithm lies in signal feature extraction and pattern classification. In the field of bearing fault diagnosis, common feature extraction algorithms include Fast Fourier Transform (Rai & Mohanty, 2007), Wavelet Transform (Lou & Loparo, 2004), Empirical Mode Decomposition (Yu, YuDejie, & Junsheng, 2006), statistical characteristics of signals (Samanta & Al-Balushi, 2003) and so on. Common pattern classification algorithms include Support Vector Machine (Konar & Chattopadhyay, 2011), BP neural network (Li, Chow, Tipsuwan, & Hung, 2000), Bayesian Classifier (Muralidharan & Sugumaran, 2012), Nearest Neighbour Classifier (Pandya, Upadhyay, & Harsha, 2013) and so on.

The current research hotspot of bearing fault diagnosis can be summarized into three categories: seeking for better feature representations, searching for the most suitable feature representations and combination of classifiers, and inventing new sensors. In 2015, Jinane from the 11th French University proposed a new feature extraction method, Global Spectrum Analysis (Harmouche, Delpha, & Diallo, 2015), to analyse bearing vibration signals. In 2016, Christelle from Aachen University of Technology in Germany proposed to use the current signal to diagnose the bearing fault (Mbo'o & Hameyer, 2016). In the same year, Georgoulas proposed a vibration signal feature fusion method for bearing fault diagnosis (Georgoulas & Nikolakopoulos, 2016).

In recent years, convolutional neural networks have achieved a great success in the field of pattern recognition (LeCun, Bengio, & Hinton, 2015). It can automatically mine features from signals and images, replacing the cumbersome feature engineering of traditional
algorithms. This makes the convolutional neural network gradually stand out in the era of big data. In 2015, Xueqian Wang from Tsinghua University constructed a deep neural network using a stacked sparse self-encoder to diagnose the faults of rolling bearings (Junbo et al., 2015). In 2016, Janssen from the University of Ghent in Belgium used convolutional neural network for the first time to diagnose the faults of bearings and gears in the gear box (Janssens et al., 2016). In the same year, Ruqiang Yan from southeast university used a deep neural network based on Denoising Sparse auto-encoders to diagnose faults in bearings and motors (Sun et al., 2016). Chang-an Zhu of China university of science and technology, based on deep belief networks and diagnose the faults of bearing using HDN (Hierarchical Diagnosis Network) (Gan, Wang, & Zhu, 2016). Yaguo Lei from Xi ‘an Jiaotong University proposed a four-layer deep neural network for bearing fault diagnosis (Jia, Lei, Lin, Zhou, & Lu, 2016).

The convolutional neural network has made major breakthroughs in the fields of two-dimensional image recognition and diagnosis. The convolutional neural network is an ‘end-to-end’ network structure, which can complete the whole process of feature extraction, feature dimension reduction and classifier classification through a neural network. This feature of convolutional neural network undoubtedly makes up the shortcomings of the current fault diagnosis method, and can provide new ideas for the current fault diagnosis algorithm and improve the diagnostic performance of the current fault diagnosis algorithm.

As a powerful feature extractor and classifier, CNN (convolutional neural network) can obtain the most suitable features for classification tasks through training. In this paper, bearing, the most important part in machinery, is taken as the research object, and the convolutional neural network is applied in bearing fault diagnosis. The model can automatically complete feature extraction and fault identification. The main work of this paper is as follows:

1. The proposed method in this paper can directly use the raw time-domain vibration signal as the input of CNN, which means that this method can be used to solve intelligent diagnosis problems of other machine systems.

2. In the fault diagnosis algorithm, feature extraction of signal is necessary. The fault diagnosis method proposed in this paper is based on convolutional neural network, which is an end-to-end diagnosis method without the need for hand-crafted feature extraction, and can automatically feature extraction and automatic fault recognition.

3. The experimental results show that the proposed method can extract the appropriate fault features adaptively from the raw bearing vibration data and classify the faults with high accuracy and stability. The recognition accuracy of the proposed method in the CWRU bearing database is more than 96%.

2. Convolutional neural network structure

Convolution neural network is a multilayer neural network, which includes convolutional layers, pooling layers, activation layers and full connected layers. The filter stage is designed to extract features from the inputs, which contains two kinds of layers, the convolutional layer and the pooling layer. The full connection layer classifies the learning features. The function of each type of layer will be described below.

2.1. Convolution layer

Convolution layer uses Convolutional Kernels to perform convolution operation on the local region of the input signal, then generates the output features by the activation unit. Each filter uses the same kernel which is also known as weight sharing, to extract the local feature of the input local region. A filter outputs a frame on the next layer, and the number of frames is called the depth of this layer. We use $K_i^l$ to denote the weights of the $i$-th filter kernel in layer $l$, and use $X^{(i)}_l$ to denote the $j$-th local region in convolutional layer $l$. Therefore, the convolution process is described as follows:

$$y^{(i)}_l = K_j^l \ast X^{(i)}_l = \sum_{j=0}^{W-1} \left( K_j^l \ast X^{(i)}_l \right),$$

where $K_j^l$ is the $j$-th weight of the $i$-th convolution kernel in layer $l$, $X^{(i)}_l$ is the $j$-th local region convolved in layer $l$, and $W$ is the width of the convolution kernel.

2.2. Pooling layer

It is common to add a pooling layer after a convolutional layer in the CNN network. It functions as a downsampling operation which reduces the spatial size of the features and the parameters of the network, at the same time, it has the effect of anti-displacement and anti-deformation. The commonly used Pooling functions are average-pooling and max-pooling, the most commonly used pooling layer is a max-pooling layer, which takes the maximum value in the perception domain as the output, to reduce the parameters and obtain location-invariant...
features. The max-pooling transformation is described as follows:

$$p_l^{ij} = \max_{j=1}^{W_l+1} a_l^{ij(1)},$$

(2)

where $a_l^{ij}$ is the activation value of the $i$-th neuron in the $l$-th frame of the layer $l$, $W_l$ is the width of the pooling region, and $p_l^{ij}$ denotes the corresponding value of the neuron in layer $l$ of the pooling operation.

### 2.3. Activation layer

After the convolution operation, then the activation function will be used. The purpose of the activation function is to acquire a nonlinear expression of the input signal to enhance the representation ability and make the learned features more divisible. Activation functions commonly used in neural networks include Sigmoid functions, Tanh and ReLU (Rectified Linear Unit). In recent years, ReLU was widely used as an activation unit to accelerate the convergence of the CNNs. ReLU makes the weights in the shallow layer more trainable when using back-propagation learning method to adjust the parameters. The formula of ReLU is described as follows:

$$a_l^{ij} = f(y_l^{ij}) = \max(0, y_l^{ij}),$$

(3)

where $a_l^{ij}$ is the activation value of the output $y_l^{ij}$ of the convolution layer.

### 2.4. Fully connected layer

The fully connected (FC) layer classifies the features extracted from the convolution layer. Specifically, the output of the last pooling layer is firstly developed into a one-dimensional feature vector as the input of the FC layer. Then the input and output are fully connected. The activation function used by the hidden layer is ReLU, and the activation function used by the output layer is Softmax. The purpose of Softmax function is to convert the input neurons into the probability distribution with a sum of 1, which is conducive to the establishment of the subsequent multi-classification objective function. The forward propagation formula of the FC layer is shown as follows:

$$Z_l^{i+1} = \sum_{i=1}^{n} W_l^{ij} a_l^{i} + b_l^j,$$

(4)

where $W_l^{ij}$ is the weight between the $i$-th neuron in layer $l$ and the $j$-th neuron in layer $l + 1$. $Z_l^{i+1}$ is the logits value of the $j$-th output neuron in layer $l + 1$, $b_l^j$ is the offset value of all neurons in the layer $l$ to the $j$-th neuron in layer $l + 1$.

### 2.5. Objective function

In the convolutional neural network, the output results obtained by forwarding propagation needs to be compared with the actual label results to calculate the gap between them. The gap size is defined by the objective function, also known as Loss Function. The common objective functions are Squared Loss Function and Cross-entropy Loss Function. In machine learning, Cross-entropy Loss Function is often regarded as the negative log-likelihood of softmax distribution. Therefore, Cross-entropy Loss Function is adopted as the objective function in this paper. The Cross-entropy Loss Function is the objective function, as shown in Equation (5):

$$L = \frac{1}{m} \sum_{k=1}^{m} \sum_{j} p_k^j \log q_k^j,$$

(5)

where $m$ is the size of the input mini-batch.

### 2.6. CNN back propagation

Loss back propagation is the key step of weight optimization for the neural network. The main approach is to first solve the derivative of the objective function with respect to the last layer of neurons, and then calculate the derivative of the objective function with respect to the ownership value layer by layer from back to front through the chain rule. It mainly includes fully connected (FC) layer back propagation and convolution layer back propagation.

The FC layer reverse derivative. First, calculate the derivative of the objective function $L$ with respect to the last logits value $z_l^{i+1}$ as shown in Equation (6):

$$\frac{\partial L}{\partial z_l^{i+1}} = \sum_{k=1}^{m} p_k^j q_k^j - p_k^j.$$

(6)

Then calculate the derivative of the FC layer objective function $L$ with respect to the weight $W_l^{ij}$ and partial $b_l^j$ of the FC layer, as shown in Equations (7) and (8):

$$\frac{\partial L}{\partial W_l^{ij}} = \frac{\partial L}{\partial z_l^{i+1}} \frac{\partial z_l^{i+1}}{\partial W_l^{ij}} = \frac{\partial L}{\partial z_l^{i+1}} \cdot a_l^{i},$$

(7)

$$\frac{\partial L}{\partial b_l^j} = \frac{\partial L}{\partial z_l^{i+1}} \frac{\partial z_l^{i+1}}{\partial b_l^j} = \frac{\partial L}{\partial z_l^{i+1}}.$$

(8)

Finally, calculate the derivatives of the objective function $L$ with respect to the activation value $a_l^{i}$ and logits value $z_l^{i}$ of the FC hidden layer with the activation
function ReLU, as shown in Equations (9) and (10):
\[
\frac{\partial L}{\partial a^{l}(i)} = \sum_{j} \frac{\partial L}{\partial z_{l}^{j}(i)} \cdot \frac{\partial z_{l}^{j}(i)}{\partial a_{l}^{j}(i)} = \sum_{j} \frac{\partial L}{\partial z_{l}^{j}(i)} \cdot W_{l}^{j}_{ij}, (9)
\]
\[
\frac{\partial L}{\partial z_{l}^{j}(i)} = \frac{\partial L}{\partial a_{l}^{j}(i)} \cdot \frac{\partial a_{l}^{j}(i)}{\partial z_{l}^{j}(i)} = \begin{cases} 
0 & \text{if } z_{l}^{j}(i) \leq 0 \\
\frac{\partial L}{\partial a_{l}^{j}(i)} & \text{if } z_{l}^{j}(i) > 0
\end{cases} (10)
\]

After the value of \(\frac{\partial L}{\partial z_{l}^{j}(i)}\) is calculated by the Equation (10), the derivatives of the objective function \(L\) with respect to the weight \(W_{l}^{j-1}\) and the bias \(b_{l}^{j-1}\) of the FC hidden layer can be solved according to Equations (7) and (8). The inverse derivative of the convolution layer is similar to the FC layer, so they will not be described here.

3. The bearing fault diagnosis model based CNN

There are three steps in our CNN diagnosis model: Data preparation, Model Training, and Model testing, as shown in Figure 1.

The CWRU Data preprocessing process is illustrated with the following data as an example: 48k Drive End Bearing Fault Data (48k_Drive_End_B007_2). The specific process is shown in Figure 2. The converted two-dimensional Data is suitable as the input sample of the convolutional neural network.

In the field of computer vision, the dataset can be enhanced by image mirroring, rotation, panning, cropping, and contrast adjustment. However, in the field of fault diagnosis, there is no specific dataset enhancement technology, and the training sample size of some diagnosis algorithms is very small, which is easy to cause overfitting. For the one-dimensional fault diagnosis signal, due to its unique timing and periodicity, the image dataset enhancement technology is not completely applicable. The data enhancement method proposed in this paper is overlapping sampling.

For a given raw accelerometer signal, we use a sliding window of size 512 with shift step size 200 to scan the signal and generate the raw data samples. Thus, for any two consecutive samples, there will be an overlap of 300 data points.

4. The CNN structure for bearing fault diagnosis

At present, the commonly used convolutional neural networks, such as VGGnet (Simonyan & Zisserman, 2014), ResNet, and Google’s Inception V4 (Szegedy et al., 2016), all contain stacked \(3 \times 3\) convolution kernel. This can not only deepen the depth of the network, but also achieve a larger field with fewer parameters, thus suppressing overfitting. The network structure of the visual field is not suitable for bearing fault diagnosis.

The structure of the convolutional neural network proposed in this paper is shown in Figure 3. The convolutional network consists of three convolutional layers (conv1, conv2, conv3), two pooling layers (pool1, pool3), two fully connected hidden layers (hidden4, hidden5), and one Softmax layer. The diagnostic signal is transformed into a set of Feature Maps through the first convolution layer and the ReLU activation layer, and then the max-pooling is used for down-sampling. Repeat the above operation to connect the feature map of the last
The CWRU dataset is one-dimensional time-domain data, which is converted into $30 \times 40$ two-dimensional data for easy processing in CNN. The structural parameters of the convolutional neural network are shown in Table 1. The model has three layers of convolution and pooling, the size of the first layer of convolution kernel is $3 \times 3$, the size of the second and third layer of convolution kernel are $2 \times 2$, and the size of the two pooling layers are $2 \times 2$. The number of neurons in the two full connective layers are 500. Softmax layer has 10 outputs for 10 bearing states.

Adam optimization algorithm: For the shallow neural network, the SGD, which used in BP neural networks can converge to the global optimum. However, for the deep convolutional neural network proposed in this paper, due to the large number of parameters and hyperparameters, if the selection of hyperparameters is not good, the SGD training often falls into the local optimum. Therefore, Adam (adaptive moments) algorithm is adopted in this paper (Kingma & Ba, 2014). Adam is a learning rate-adaptive optimization algorithm that dynamically adjusts the learning rate of each parameter by using the first-moment estimation and the second-moment estimation of the gradient.

Dropout algorithm: Dropout was proposed by Srivastava of the University of Toronto (Srivastava et al., 2014), which could suppress overfitting of the neural network and improve its generalization performance. The Dropout algorithm is to adjust the value of neurons in the network to zero in probability $p$, so as to prevent co-adaptation between neurons in the same layer. Therefore, Dropout makes the feature expression of neurons more independent and can improve the expression ability of network. In general, Dropout is used for full connection layer, because
it takes the largest proportion of parameters in the whole convolutional neural network.

5. Bearing fault diagnosis experiment

5.1. Experiment settings

The experimental data in this paper are from the Rolling Bearing Data Center of Case Western Reserve University (CWRU) (Csegroups). The CWRU dataset is the bearing fault diagnostic standard dataset, which is recognized worldwide. The bearing data acquisition system is shown in Figure 4. The experimental platform consists of a 2 HP motor (left), a torque sensor (middle), a power meter (right) and an electronic control unit. A single point of failure was arranged on the bearing by using EDM technology. There were three defect positions of the bearing diagnosed, namely rolling element damage, outer ring damage and inner ring damage, and the fault diameters are 0.007, 0.014, and 0.021 in., respectively. The bearings use SKF bearings. Vibration signals of active end bearings are collected by acceleration in experiments. The sampling frequency of the digital signal is 12,000 Hz, and the fault data of the drive end bearing are also collected at the sampling rate of 48,000 Hz.

As shown in Figure 5, the drive end signal of frequency 48 KHZ is used to display the rolling element, inner ring, outer ring fault signal and the normal signal from top to bottom.

5.2. Experimental implementation

The test objects of this test are the drive end bearing. The model of the bearing to be diagnosed is the deep groove ball bearing SKF6205, and the sampling frequency is 48 kHz. There are three defect positions in the bearing to be diagnosed, namely ball, inner race and outer race. The fault diameters are 0.007, 0.014, and 0.021 in., respectively. There are nine damage states in total. On experiment, 2000 data points were used for diagnosis for each sample.

A total of four datasets were prepared for the experiment, as shown in Table 2. Dataset 1, Dataset 2, and Dataset 3 are loaded with 1, 2, and 3 hp, respectively. Each dataset contains 2400 training samples and 480 test samples, and there is no overlap between the training samples and the test samples. Dataset 4 is the union of the Dataset 1, Dataset 2, and Dataset 3, which includes 3 load states, a total of 7200 training samples and 1440 test samples.
### Table 2. Test dataset description.

| Damage location | Normal | Rolling body | Inner ring | Outer ring | Load |
|-----------------|--------|--------------|------------|------------|------|
| Label           | 0      | 1            | 2          | 3          | 4    |
| Damage diameter | 0.007  | 0.014        | 0.021      | 0.007      | 0.014|
| Datas Train     | 240    | 240          | 240        | 240        | 240  |
| Et1 valid       | 48     | 48           | 48         | 48         | 48   |
| Datas Train     | 240    | 240          | 240        | 240        | 240  |
| Et2 valid       | 48     | 48           | 48         | 48         | 48   |
| Datas Train     | 240    | 240          | 240        | 240        | 240  |
| Et3 valid       | 48     | 48           | 48         | 48         | 48   |
| Datas Train     | 720    | 720          | 720        | 720        | 720  |
| Et4 valid       | 144    | 144          | 144        | 144        | 144  |

**Figure 6.** Test accuracy of convolution network on four datasets.

### 5.3. Experimental results

The deep-learning framework used in the experiment was Lasagne. Lasagne is a lightweight library based on Theano to build and train neural networks. It supports feedforward networks such as CNN, and RNN includes Long-term and Short-term Time Memory (LSTM), and any combination of them, to supports the CPU and GPU to run. The configuration of the computer used is CPU: Intel core i7-4770, with 8GB of memory capacity and 1TB SeaGate hard disk. The system environment is Ubuntu 64-bit version 16.04, Python 2.7, which builds a convolutional neural network model in the Lasagne environment. After the training data is trained by the model, the value of the objective function is calculated, and then the weight in the model is updated through the training module. Finally, the trained model is used to diagnose the input signal in the test stage.

The Dataset 1, Dataset 2, Dataset 3, and Dataset 4 are trained using the built network. The initial parameters of the Adam algorithm during training are as follows:

- update = adam,
- update_learning_rate = 0.0001,
- update_beta1 = 0.9,
- update_beta2 = 0.999,
- update_epsilon = 1e-8.

Dropout suppression overfitting is used on the third and fourth layers of the network, respectively. The specific parameters are set as follows:

- dropout3_p = 0.3,
- dropout4_p = 0.5.

For Dataset 1, Dataset 2, Dataset 3 and Dataset 4, in order to verify the reliability of the results of each training, 10 times were trained for each Dataset. The
test results are shown in Figure 6. It can be seen that the recognition rate of the convolutional neural network on each dataset reaches 98%, with a minimum of 96%. The experimental results show that the convolutional neural network is feasible and can achieve high recognition rate when applied to the vibration signal in time-domain.

6. Conclusion

Under the background of intelligent manufacturing and machine learning, this paper applies the convolutional neural network to the bearing fault diagnosis, proposes the convolutional neural network model applied to the time-domain vibration signal, and carries out the bearing fault diagnosis. The proposed method in this paper can directly use the raw time-domain vibration signal as the input of CNN, which means that this method can be used to solve intelligent diagnosis problems of other machine systems. The proposed method is an end-to-end diagnosis method without the need for hand-crafted feature extraction, and can automatically feature extraction and automatic fault recognition. The experimental results show that the proposed method can extract the appropriate fault features adaptively from the raw bearing vibration data and classify the faults with high accuracy and stability. The recognition accuracy of the proposed method in the CWRU bearing database is more than 96%. In this paper, bearing is taken as an example, in the future work, the proposed method can be applied to fault diagnosis of rotating parts such as gears, or other neural networks other than CNN, such as Long Short-Term Memory network (LSTM).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Natural Science Foundation of China [grant number 61370205]; Science and Technology Project of Chongqing Education Commission of China [grant number KJ1600332]; Ministry of Education Humanities and Social Science Project [grant number 18YJA880061]; Humanities and Social Science Project of Chongqing Education Commission of China [grant number 19SKGH035]; and Chongqing Education Scientific Planning Project [grant number 2017-GX-116].

References

Csegroups. Retrieved from http://csegroups.case.edu/bearing datacenter/pages/welcome-case-western-reserve-university -bearing-data-center-website

Gan, M., Wang, C., & Zhu C. (2016). 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, 72–73, 92–104.

Georgoulas, G., & Nikolakopoulos, G. (2016). Bearing fault detection and diagnosis by fusing vibration data. Industrial Electronics Society, IECON 2016-42nd Annual Conference of the IEEE (pp. 6955–6960). IEEE.

Harmouche, J., Delpha, C., & Diallo, D. (2015). Improved fault diagnosis of ball bearings based on the global spectrum of vibration signals. IEEE Transactions on Energy Conversion, 30(1), 376–383.

Janssens, O., Slavovikj, V., Vervisch, B., Stockman K., Locu-ifier M., Verstoct S., . . . Van Hoecke S. (2016). Convolutional neural network based fault detection for rotating machinery. Journal of Sound and Vibration, 377, 331–345.

Jia, F., Lei, Y., Lin, J., Zhou X., & Lu N. (2016). Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mechanical Systems and Signal Processing, 72–73, 303–315.

Junbo, T., Weining, L., Juneng, A., & Xueqian, W. (2015). Fault diagnosis method study in roller bearing based on wavelet transform and stacked auto-encoder. The 27th Chinese Control and Decision Conference (2015 CCDC) (pp. 4606–4613). IEEE.

Kingma, Diederik P., & Ba, Jimmy. (2014). Adam: A method for stochastic optimization. arXiv preprintarXiv:1412.6980.

Konar, P., & Chattopadhyay, P. (2011). Bearing fault detection of induction motor using wavelet and support vector machines (SVMs). Applied Soft Computing, 11(6), 4203–4211.

Lasagne. Retrieved from http://lasagne.readthedocs.io/en/latest/

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

Li, B., Chow, M. Y., Tipsuwan, Y., & Hung J.C. (2000). Neural-network-based motor rolling bearing fault diagnosis. IEEE Transactions on Industrial Electronics, 47(5), 1060–1069.

Lou, X., & Loparo, K. A. (2004). Bearing fault diagnosis based on wavelet transform and fuzzy inference. Mechanical Systems and Signal Processing, 18(5), 1077–1095.

Mbo'o, C. P., & Hameyer, K. (2016). Fault diagnosis of bearing damage by means of the linear discriminant analysis of stator current features from the frequency selection. IEEE Transactions on Industry Applications, 52(5), 3861–3868.

Muralidharan, V., & Sugumaran, V. (2012). A comparative study of Naive Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. Applied Soft Computing, 12(8), 2023–2029.

Pandya, D. H., Upadhyay, S. H., & Harsha, S. P. (2013). Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN. Expert Systems with Applications, 40(10), 4137–4145.

Rai, V. K., & Mohanty, A. R. (2007). Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform. Mechanical Systems and Signal Processing, 21(6), 2607–2615.

Samanta, B., & Al-Balushi, K. R. (2003). Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. Mechanical Systems and Signal Processing, 17(2), 317–328.

Simonyan, Karen., & Zisserman, Andrew. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
Srivastava, Nitish., Hinton, Geoffrey., Krizhevsky, Alex., & Sutskever, Ilya. (2014). Dropout: A simple way to prevent neural. *The Journal of Machine Learning Research, 15*(1), 1929–1958.

Sun, W., Shao, S., Zhao, R., Yan R., Zhang X., & Chen X. (2016). A sparse auto-encoder-based deep neural network approach for induction motor faults classification. *Measurement, 89*, 171–178.

Szegedy, Christian., Ioffe, Sergey., Vanhoucke, Vincent., & Alemi, Alexander. (2016). Inception-v4, inception-resnet and the impact of residual connections on learning. arXiv preprint arXiv:1602.07261.

Yu, Yang., Dejie, Yu., & Junsheng, Cheng. (2006). A roller bearing fault diagnosis method based on EMD energy entropy and ANN. *Journal of Sound and Vibration, 294*(1), 269–277.