IEC
Institute of Electronics and Computer

Journal of the Institute of Electronics and Computer, 2020, 2, 72-92
https://iecscience.org/journals/JIEC
ISSN Online: 2643-8240

The Fault Diagnosis based on Deep Long Short-Term Memory Model from the Vibration Signals in the Computer Numerical Control Machines

Kemal Polat

Department of Electrical and Electronics Engineering, Faculty of Engineering, Bolu Abant Izzet Baysal University, 14280, Bolu, Turkey
E-mail: kpolat@ibu.edu.tr

How to cite this paper: Kemal Polat (2020). The Fault Diagnosis based on Deep Long Short-Term Memory Model from the Vibration Signals in the Computer Numerical Control Machines. Journal of the Institute of Electronics and Computer, 2, 72-92. https://doi.org/10.33969/JIEC.2020.21006.

Received: February 24, 2020
Accepted: April 6, 2020
Published: April 10, 2020

Abstract

Rotating machines have become indispensable in every field of life. Because of people's commitment to machines, the machine's faults have devastating consequences. Fault analysis of devices is common in the industry. Studies in the literature mostly work on simple models. These simple models become inefficient as data complexity increases. In this paper, three types of faults are detected in the drill bit of the CNC (Computer Numerical Control) machine. Deep LSTM (Long Short-Term Memory)-based classification model has been proposed for fault diagnosis of CNC machines. Different structures can be combined to obtain higher accuracy rates. The proposed method consists of two stages: (I) the feature extraction including time, frequency domain, and Morlet wavelet coefficients information from vibration signals from CNC machines, (II) the classification of fault types using deep LSTM model based on the extracted features. LSTM structures having two layers, three layers, four layers, and five layers have been proposed and then compared with each other with respect to the classification accuracy of fault diagnosis. The highest accuracy rate obtained in this study was 99.53%. It can be seen that deep LSTM gives outstanding results in the structures with sequential data input and the proposed system gives promising results in the field of fault analysis of machines.

Keywords

LSTM (Long Short-Term Memory) classification method, Deep learning, Machine based Fault Diagnosis, Drill Bit Fault

1. Introduction

With the advances in technology and science, the mechanical systems used in the modern industry have become more functional and more complex than ever. Rotating machines have become an indispensable part of the industry. Failure of the

rotating machinery, the decrease of the efficiency of the machine, or the stop of the machine even for a short period can cause significant damage to the workplace. Therefore, fault analysis in rotating machines is one of the most important issues in system design and maintenance. Fault analysis of rotating machines has basically three types of tasks: (1) to diagnose whether the system operates in a stable manner, (2) to diagnose the fault when it is in its initial phase or to estimate it before a fault occurs, (3) to predict the direction in which the fault will proceed before the faulty condition deteriorates [1]. Thus, fault analysis is shown as the identification of problems in the working conditions of the machines.

CNC machines can produce industrial parts that are free from human errors. They are often used to produce prototypes in the R&D field as well as to produce products in mass production. With the use of CNC machines, the workforce needed in the industry and the time required to carry out the process are considerably shortened. Due to these advantages, CNC machines have become indispensable today. It is essential to diagnose faults in such a widely used machine.

Measurement of parameters such as current, voltage, and vibration is often used to perform fault diagnosis in machines [2, 3, 4]. Because of the complexity of the signals obtained from the machines, it is impossible to make a fault diagnosis directly from the raw signal. Thus, a two-step method is commonly used for machine fault diagnosis, first in the feature extraction, and in the second stage in diagnosing the fault [1, 5]. The extracted features are used as input data of machine learning methods. There are a variety of machine learning methods, such as predictive methods, mathematical optimization, statistical learning, and classification. The most commonly used methods for fault diagnosis of rotating machines are classification and statistical learning. The most commonly used classification and statistical learning algorithms are the k nearest neighbor (k-NN) [6], the Bayesian Classifier [7], the Support Vector Machine (SVM) [8] and the artificial neural network (ANN) [9]. In recent years, deep learning algorithms have been used in the field of fault diagnosis.

Saravanan et al. (2009), using vibration signals, fuzzy logic, and decision tree structures, have performed error analysis work [27]. Sakthivel et al. (2010), C4.5 decision tree algorithm with the help of the analysis of centrifugal pump errors have done [28]. Pandya et al. (2012), for the diagnosis of bearing failure, feature extraction process with wavelet packet transformation, multi-layer ANN have performed the classification process [29]. Seshadrinath et al. (2013) used complex wavelets related to various errors occurring in induction motor drives. Discrete wavelet transform, DVM, and k nearest neighbor algorithms are also used for comparison [30]. Jafari et al. (2014) used the Acoustic Emission and ANN methods

to diagnose the valve failure of the internal combustion engine [31]. Ali et al. (2015), after performing the feature extraction process from vibration signals, performed the training of ANN structure with the obtained properties and conducted a study to predict bearing failure [32]. Janssens et al. (2016) conducted convective Neural Networks (ESA) based error analysis on spinning machines. In this study, the vibration signals are given by the servo motor during bearing failure, data preprocessing, feature extraction, feature selection, and ESA model creation stages were passed through the 93.61 and 87.25 accuracy rates were obtained [33]. Liu et al. (2018) conducted a comprehensive study with a k-nearest neighbor, naive Bayes classifier, DVM, ANN, and Deep Learning applications to perform error analysis on rotating machines [34]. Jing et al. (2017) obtained a high accuracy rate of 99.33 with the multi-layered ESA structure on the dataset formed by the PHM community with vibration signals received from the gearbox (gearbox) in the planetary arrangement. In this study, Random Forest (RO), DVM, and fully connected ANN operations were applied and the results were compared [35]. Kiranyaz et al. (2018) used a 1-dimensional ESA structure to diagnose the failure of the Modular Multilevel Transducer (MDSD) device. 2D ESA is mainly developed for image processing applications, so it is not efficient to use in vector-shaped data sets. Therefore, they preferred to use 1-dimensional ESA instead of 2-dimensional ESA in this study [36].

In this paper, deep LSTM classification has been formed for the diagnosis of the faults occurring at the drill bit of the CNC machine. First, feature extraction was performed from the vibration data. There are eight features in the time domain, eight features in the frequency domain, and five features with Morlet Wavelet Transform (MWT). After the feature extraction, the number of vibration data in the dataset was increased gradually to provide better results. The number of data was increased to 2, 4, 8, and 16 times, respectively. After increasing the number of data, the normalization process was applied to eliminate the significant differences between the data. Min-max and z-score normalization methods were preferred for the normalization process. After the normalization of the data, the deep Long-Short Time Memory (LSTM) was proposed for the classification process. A total of 4 deep LSTM networks were obtained by adding various layers to the LSTM method. In addition to these steps, the effects on the results have been observed by increasing the number of the LSTM layer outputs and changing the activation functions in the layers.

The sections of this paper are as follows; the dataset is presented in section 2. The classification procedures used for error analysis are shown in section 3. The experimental results and the case studies in this area are listed in section 4. Finally,

conclusions are given in section 5.

2. Material

The dataset used in this paper was obtained by Verma et al. to diagnose various wears that occurred at the drill bit of the CNC machine. In the study, EMCO Concept Mill 105 3-axis computerized CNC machine shown in Figure 1 is used [5].

Figure 1. EMCO Concept Mill 105 [11].

HSS Twist 9mm diameter drill bit was used to drill the metal plate. In order to capture the vibration data, a single axis accelerometer named PCB 63001 is used. The sampling frequency is 32.768 Hz. In order to convey the vibration data to the computer, the data acquisition system used the NI 9234 signal processing unit and the NI 9172 computer connection [5].

When measuring the vibration data, the rotational speeds of 160, 170, 180, 190, and 200 rpm are used in addition to the feed rates of 4mm, 8mm, and 12mm per minute. Thus, a total of 15 different variations are obtained. For each variation, 8-second vibration data is recorded both when drilling the metal plate and after drilling the metal plate. This procedure is repeated for 3 faults and 1 normal condition [5].

Chisel wear: At this stage, the recording is started as soon as the drill bit blinding occurs. The blin is resulting from high temperature, and tip deformation is shown in Figure 2(a) [12].

Flank wear: The wear occurs due to the friction force between the metal plate and the drill bit. This wear, which shown in Figure 2(b), increases as the rotational speed of the drill bit increases [12].

Outer corner wear: Due to the high frictional force that occurs between the drill bit and the inner side of the hole, the outer corner of the drill bit is worn out, and the piece can break off. This wear, as shown in Figure 2(c), may occur at one corner or both sides of the drill bit [12].

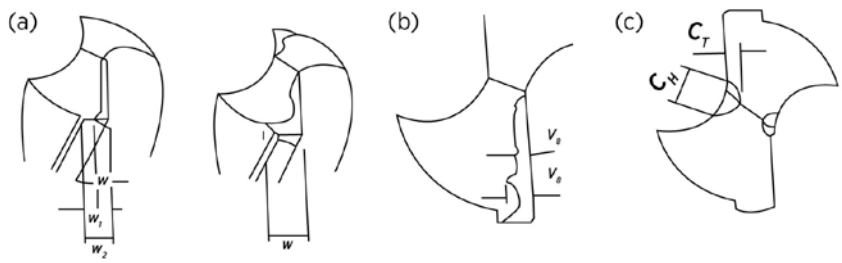
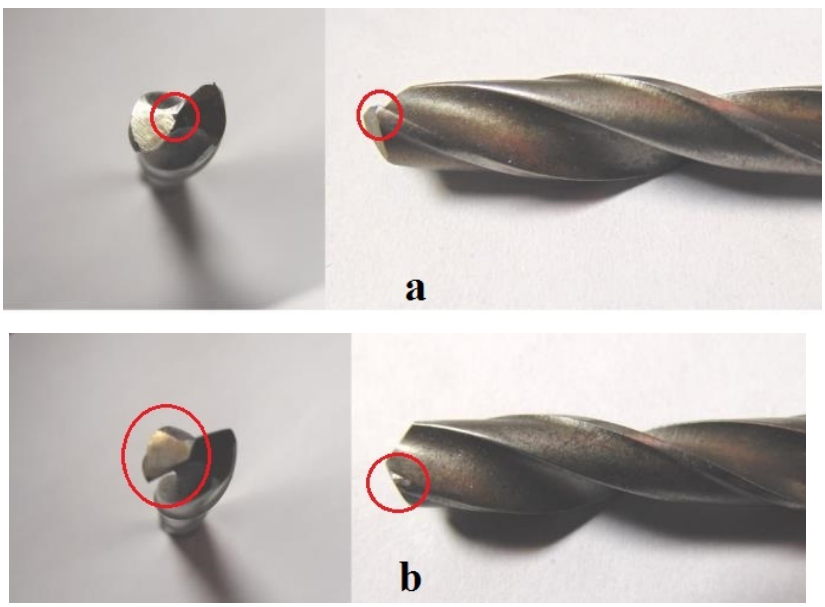


Figure 2. Drill Bit Wears (a) chisel wear (b) flank wear (c) outer corner wear [12].

The actual situation pictures of the faulty cases are given in Figure 3.



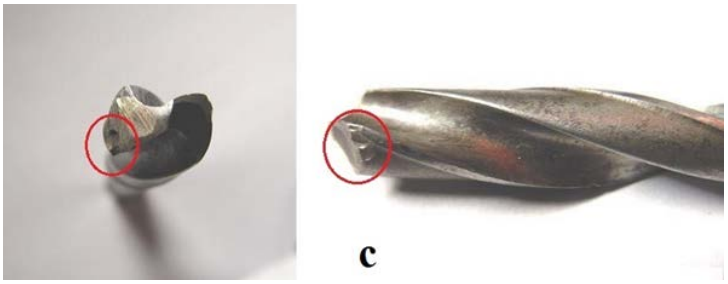


Figure 3. Actual Pictures of Faulty Situations (a) chisel wear (b) flank wear (c) outer corner wear [13].

In this dataset, the signals belonging to each fault and normal conditions are given in the below figures. Figure 4 shows the vibration signal of the drill blanking error in the dataset. Figure 5 gives the vibration signal of wear error of side of the drill bit. Figure 6 denotes the Vibration signal of wear error of the outer corner of the drill bit. Figure 7 demonstrates the vibration signal of the normal state in the dataset.

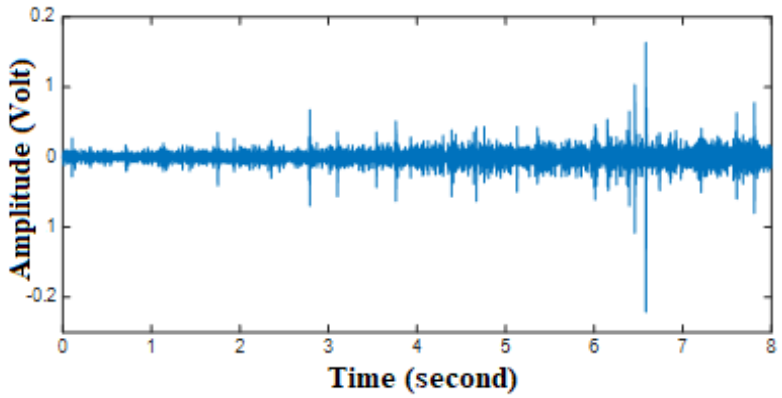


Figure 4. Vibration signal of drill blanking error in the dataset

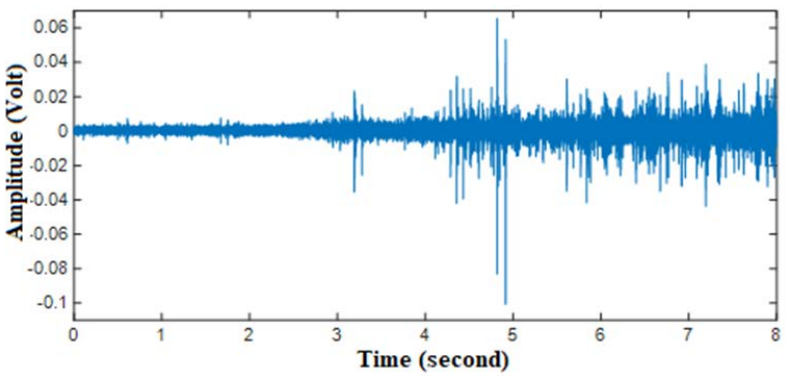


Figure 5. Vibration signal of wear error of side of the drill bit in the dataset

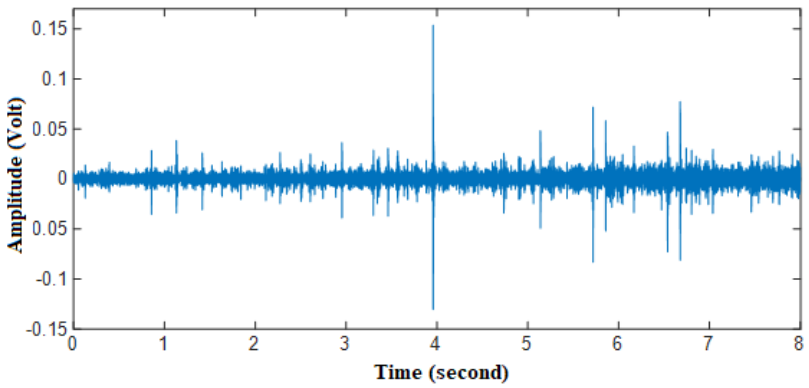


Figure 6. Vibration signal of wear error of outer corner of the drill bit in the dataset

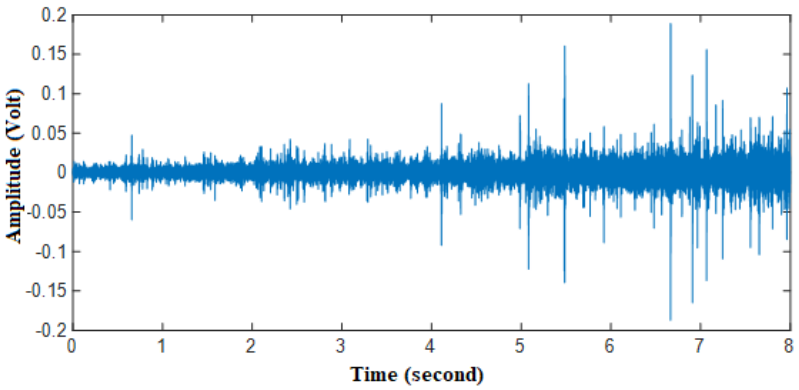


Figure 7. Vibration signal of normal state in the dataset

3. The Proposed Method

In this paper, to carry out the classification of fault types in the CNC machines, a three stages hybrid model has been proposed. In the first stage, we have extracted the features including time domain, frequency domain, and Morlet wavelet coefficients from the vibration signals. In the second stage, to standardize the dataset, two different normalization methods comprising the z-score and min-max. normalization has been used. In the third stage, the normalized data have been classified by deep LSTM networks. In Fig. 8, the methods used for diagnosing the fault are shown in steps.

3.1. Feature Extraction Stage

In this study, 21 features from the vibration signals for each case in the dataset have been extracted from the vibration signals. Among these 21 features, eight features are extracted in the time domain. These features; absolute mean, maximum peak,

RMS value, variance, kurtosis, peak factor, shape factor, and skewness. Eight features are extracted through Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT). Five features are extracted through Morlet Wavelet Transform (MWT), which are standard deviations, wavelet entropy, kurtosis, skewness, and variance [17, 38, 39, 40].

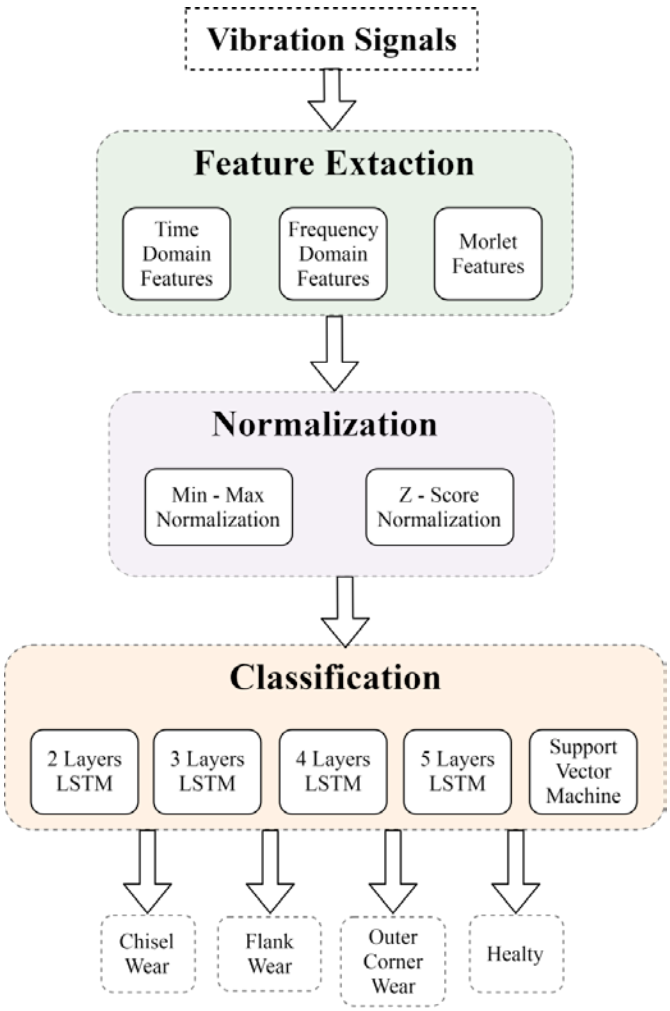


Figure 8. The block diagram of the proposed method

3.2. Normalization Stage

We have used two different normalization methods including z-score and minimum-maximum normalization for the dataset. Z-score normalization is a strategy of normalizing data that estop this outlier problem. The equation for z-score

normalization is shown in Eq. (1).

$$z - score = \frac{Value - \mu}{\sigma} \quad (1)$$

Here, μ is the mean value of the feature and σ is the standard deviation of the feature.

Min-max. Normalization is one of the most common ways in order to normalize data. For every feature, the minimum value of that feature gets transformed into a 0, the maximum value gets transformed into a 1, and every other value gets transformed into a decimal between 0 and 1. The equation for Min.-max. Normalization is shown in Eq. (2).

$$Normalization = \frac{value - min.}{max. - min.} \quad (2)$$

3.3. Deep LSTM Classification Model

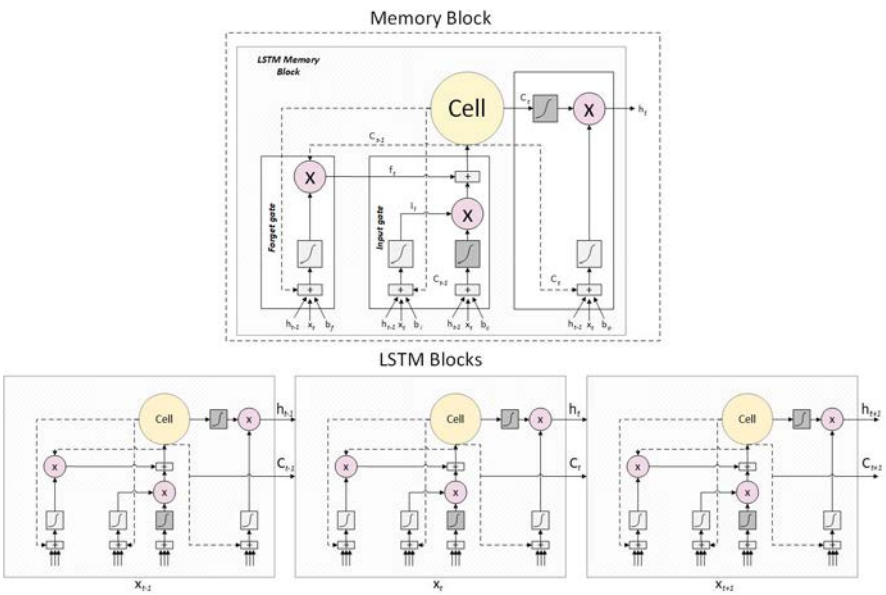


Figure 9. The proposed deep LSTM model for fault diagnosis

In this study, a multilayer classification algorithm based on deep LSTM is proposed to automatically recognize error types after the feature is extracted from vibration signals on CNC machines. Deep learning has recently been used frequently in problems such as biomedical, measurement systems, pattern recognition, control, communication, and electronic system design. In this study, 2-layer, 3-layer, 3-layer, and 4-layer LSTM models have been proposed and compared in terms of system performance in order to automatically identify four different error types. The block

diagram of the proposed system is given in Figure 9.

While creating multi-layer network structures, several layers were added to the LSTM method, and results were observed. New structures were formed by improving the layers on the observed results. A total of 4 structures are shown below. Table 1 shows the parameters of the Layers in LSTM modeling for our dataset.

The multilayered structures of the proposed deep LSTM are given below;

- LSTM layer + output layer (2 Layers),
- LSTM layer + pooling layer + fully-connected layer + output layer (4 Layers),
- LSTM layer + fully-connected layer + output layer (3 Layers),
- LSTM layer + 3 fully-connected layers + output layer (5 Layers)

Table 1. Parameters of the Layers in LSTM modelling for our dataset

Layers	Variables and Dimensions	Activation Functions
LSTM Layer	1-10 outputs 100 outputs	ReLU function Tanh function
Fully Connected Layer	30 nodes	ReLU function Tanh function
Pooling Layer	S=2	-
Output Layer	6 outputs	Softmax

4. Experimental Results

4.1. Performance Criteria

The performance criterion for the results is shown in Figure 10. Features were obtained in 4 areas. Then, a normalization layer containing three methods was applied for each feature area. Afterward, the results were obtained by 4 data segmentation methods for each normalization layer

These procedures have been repeated for the 3 steps shown below.

- LSTM outputs number 10 and activation function is ReLU
- LSTM outputs number 10, and the activation function is TANH
- LSTM outputs number 100, and the activation function is TANH

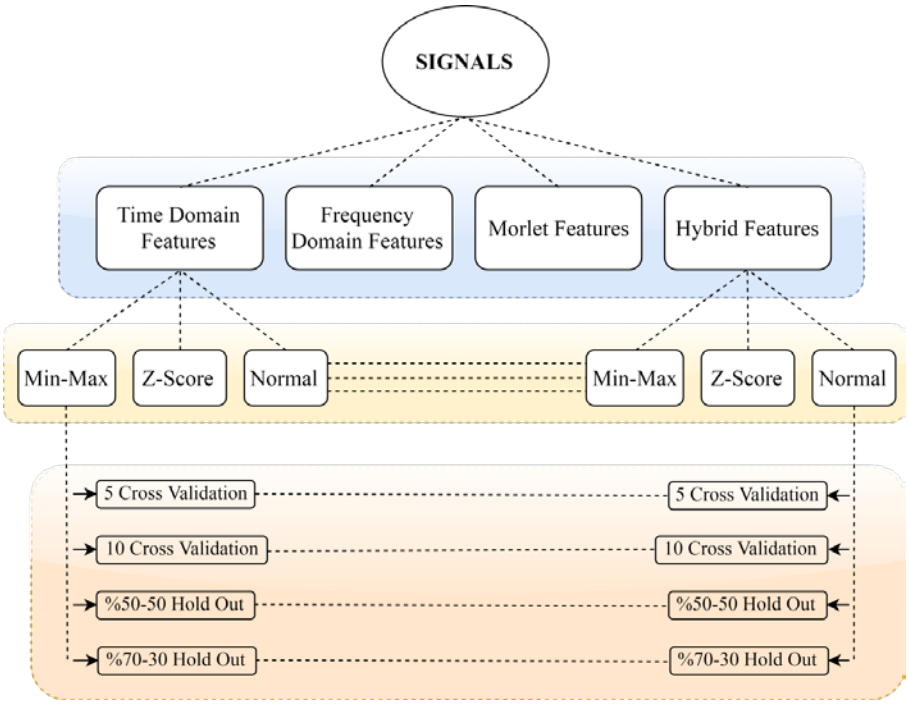


Figure 10. The used performance criteria for our study

4.2. Results

4.2.1. Time Domain Features Results

The best results obtained using only time-domain features are shown in this section. The results of the situation at which the output number of the LSTM layer is 10 and the activation function is the RELU is shown in Table 2. The best result for this situation is obtained with a 5-layer structure.

Table 2. The obtained classification accuracy results of LSTM output:10, ReLU (5 Layers) in the classification of fault diagnosis from the vibration signals based on only time-domain features

	5 Cross-Validation			**10 Cross-Validation**			**50-50% Hold Out**			**70-30 % Hold Out**		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	53,33	21,67	38,33	52,50	19,17	40,00	36,67	23,33	23,33	36,11	19,44	27,78
2x Data	65,42	30,42	32,08	66,25	29,58	30,42	45,00	16,67	30,83	59,72	27,78	31,94
4x Data	71,67	50,62	37,92	73,54	49,79	36,46	54,58	34,58	34,58	63,89	52,08	37,50
8x Data	87,60	61,14	39,27	88,54	60,42	38,54	82,29	38,96	34,37	80,90	49,65	29,86
16x Data	**91,15**	83,96	40,16	90,00	85,21	41,77	88,33	79,06	36,56	90,10	80,90	40,62

The results of the situation at which the output number of the LSTM layer is 10 and

the activation function is the TANH is shown in Table 3. The best result for this situation is obtained with a 5-layer structure.

Table 3. The obtained classification accuracy results of LSTM output:10, TANH (5 Layers) in the classification of fault diagnosis from the vibration signals based on only time-domain features.

	5 Cross-Validation			10 Cross-Validation			50-50% Hold Out			70-30 % Hold Out		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	70,00	18,33	36,67	70,00	20,83	34,17	58,33	21,67	33,33	69,44	27,78	33,33
2x Data	78,33	36,25	32,92	76,25	36,25	33,33	69,17	23,33	30,83	84,72	26,38	31,94
4x Data	81,46	53,12	37,92	85,21	55,20	40,62	59,17	37,92	35,00	78,47	52,08	33,33
8x Data	88,54	69,58	44,06	88,96	74,89	43,44	85,62	55,00	41,46	85,07	56,59	38,89
16x Data	**94,27**	88,59	44,63	93,80	88,12	46,20	85,83	87,39	45,52	92,18	83,68	45,66

The results of the situation at which the output number of the LSTM layer is 100 and the activation function is the TANH is shown in Table 4. The best result for this situation is obtained with a 5-layer structure.

Table 4. The classification accuracies of LSTM output:100, TANH (5 Layers) in the classification of fault diagnosis from the vibration signals based on only time-domain features.

	5 Cross-Validation			10 Cross-Validation			50-50% Hold Out			70-30 % Hold Out		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	72,50	16,67	35,00	71,67	20,83	34,17	60,00	23,33	30,00	69,44	33,33	36,11
2x Data	77,92	34,17	33,75	78,75	32,92	35,83	64,17	23,33	29,17	76,39	26,39	31,94
4x Data	84,79	56,25	38,54	85,62	59,17	41,87	62,92	35,83	33,75	84,72	54,17	33,33
8x Data	87,71	72,50	42,29	89,58	80,94	46,77	88,96	52,29	37,71	86,11	60,76	37,50
16x Data	**92,66**	89,27	43,02	92,60	89,69	48,85	89,48	88,33	47,50	91,32	84,20	45,83

4.2.2. Frequency Domain Features Results

The best results obtained using only frequency domain features are shown in this section. The results of the situation at which the output number of the LSTM layer is 10 and the activation function is the RELU is shown in Table 5. The best result for this situation is obtained with a 5-layer structure.

Table 5. The obtained classification accuracy results of LSTM output:10, ReLU (5 Layers) in the classification of fault diagnosis from the vibration signals based on only frequency domain features

	5 Cross-Validation			10 Cross-Validation			50-50% Hold Out			70-30 % Hold Out		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	40,83	27,50	66,67	43,33	20,00	65,00	30,00	25,00	60,00	22,22	16,67	69,44
2x Data	62,50	32,50	75,83	62,92	37,08	72,08	56,67	19,17	74,17	59,72	23,61	76,38
4x Data	72,08	42,29	78,75	71,04	44,58	75,83	60,83	33,75	78,75	68,75	33,33	81,25
8x Data	77,70	56,35	80,00	77,50	58,23	81,25	71,87	44,17	76,46	81,25	52,43	78,82
16x Data	86,61	65,73	84,74	**88,28**	67,86	84,74	84,69	64,06	83,75	86,28	65,10	84,03

The results of the situation at which the output number of the LSTM layer is 10 and the activation function is the TANH is shown in Table 6. The best result for this situation is obtained with a 5-layer structure.

Table 6. The obtained classification accuracy results of LSTM output:10, TANH (5 Layers) in the classification of fault diagnosis from the vibration signals based on only frequency domain features

	5 Cross-Validation			10 Cross-Validation			50-50% Hold Out			70-30 % Hold Out		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	49,17	30,00	71,67	48,33	30,00	71,67	36,67	18,33	68,33	33,33	13,89	77,78
2x Data	66,67	40,83	80,00	67,50	38,75	80,42	55,00	20,83	75,83	77,78	31,94	76,39
4x Data	77,92	46,87	80,00	76,25	47,29	78,96	67,08	38,75	78,33	75,69	43,05	82,64
8x Data	77,71	61,35	79,79	79,79	61,04	81,77	73,75	50,00	76,25	81,25	57,29	81,25
16x Data	88,91	70,42	83,75	**90,10**	77,03	84,27	87,92	65,42	83,33	87,50	69,96	83,33

The results of the situation at which the output number of the LSTM layer is 100 and the activation function is the TANH is shown in Table 7.

Table 7. The obtained classification accuracy results of LSTM output:100, TANH (5 Layers) in the classification of fault diagnosis from the vibration signals based on only frequency domain features

	5 Cross-Validation			10 Cross-Validation			50-50% Hold Out			70-30 % Hold Out		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	59,17	23,33	71,67	58,33	25,83	71,67	40,00	30,00	76,67	50,00	13,89	72,22
2x Data	68,33	37,08	77,92	68,75	34,58	79,58	70,83	23,33	76,67	75,00	38,89	73,61
4x Data	78,75	50,00	79,17	80,62	53,33	81,04	70,00	40,42	77,08	81,94	51,39	83,33
8x Data	78,96	68,33	80,83	80,42	72,60	81,98	72,71	59,17	74,79	79,86	66,32	81,94
16x Data	91,04	81,35	84,27	**92,34**	83,75	84,69	89,79	77,92	83,96	92,01	81,60	82,46

The best result for this situation is obtained with a 5-layer structure. According to

the results obtained from frequency domain features, the best result is obtained when the LSTM output number is 100, the activation function is TANH, min-max normalization is applied, and the data number is increased to 16 times. At this stage, the best result is obtained through the 5-layer LSTM structure and 10-fold cross-validation method. When the above tables are examined, it is possible to obtain a maximum accuracy of 92,34% using only frequency domain features. Min-max normalization should be preferred for the best result, and high results are also obtained with non-normalized data, but the results do not exceed 90% in this wise. In this stage, the 10-fold cross-validation method is better than other methods.

4.2.3. Morlet Wavelet Features Results

The best results obtained using only Morlet Wavelet features are shown in this section. The results of the situation at which the output number of the LSTM layer is 10 and the activation function is the RELU is shown in Table 8. The best result for this situation is obtained with a 5-layer structure.

Table 8. The obtained classification accuracy results of LSTM output:10, ReLU (5 Layers) in the classification of fault diagnosis from the vibration signals based on only Morlet Wavelet Features

	5 Cross-Validation			**10 Cross-Validation**			**50-50% Hold Out**			**70-30 % Hold Out**		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	42,50	31,67	30,00	44,17	33,33	25,00	36,67	23,33	25,00	36,11	30,55	11,11
2x Data	64,58	40,00	39,17	63,33	41,25	34,58	59,17	26,67	26,67	62,50	31,94	30,55
4x Data	86,25	58,75	38,33	86,46	63,33	41,04	83,33	42,50	37,08	88,89	59,03	43,05
8x Data	87,71	64,79	39,37	88,02	64,37	40,52	84,37	59,17	34,58	88,89	63,89	34,37
16x Data	91,87	66,25	41,25	91,09	68,49	42,13	87,92	56,46	40,21	**92,01**	62,85	39,76

Table 9. The obtained classification accuracy results of LSTM output:10, TANH (5 Layers) in the classification of fault diagnosis from the vibration signals based on only Morlet Wavelet Features

	5 Cross-Validation			**10 Cross-Validation**			**50-50% Hold Out**			**70-30 % Hold Out**		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	55,00	35,00	32,50	55,00	30,83	30,83	45,00	30,00	26,67	55,55	25,00	27,78
2x Data	69,58	45,42	37,08	68,33	42,50	35,83	62,50	31,67	30,00	66,67	43,05	27,78
4x Data	86,04	59,37	41,04	89,17	62,70	43,12	87,50	44,58	36,67	90,97	59,72	38,89
8x Data	90,73	59,89	43,65	89,37	64,79	41,77	87,08	57,08	40,42	90,28	59,37	40,97
16x Data	93,18	68,91	41,15	93,65	72,66	43,38	91,87	64,58	41,46	**94,27**	71,70	40,28

The results of the situation at which the output number of the LSTM layer is 10 and the activation function is the TANH is shown in Table 9. The best result for this

situation is obtained with a 5-layer structure.

The results of the situation at which the output number of the LSTM layer is 100 and the activation function is the TANH is shown in Table 10.

Table 10. The obtained classification accuracy results of LSTM output:100, TANH (5 Layers) in the classification of fault diagnosis from the vibration signals based on only Morlet Wavelet Features

	5 Cross-Validation			10 Cross-Validation			50-50% Hold Out			70-30 % Hold Out		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	56,67	25,83	32,50	61,67	28,33	30,00	48,33	31,67	28,33	50,00	30,56	30,56
2x Data	72,92	37,08	40,42	68,75	40,00	35,42	65,83	30,83	33,33	70,83	31,94	41,67
4x Data	87,29	58,75	37,08	87,50	65,21	41,46	87,50	42,50	36,67	91,67	56,25	38,89
8x Data	89,79	65,00	41,77	89,48	69,90	41,87	87,92	64,58	41,46	89,93	67,36	40,28
16x Data	92,40	72,60	41,30	93,02	75,21	42,50	89,69	72,08	40,94	**93,75**	73,96	40,28

The best result for this situation is obtained with a 5-layer structure. According to the results obtained from frequency domain features, the best result is obtained when the LSTM output number is 10, the activation function is TANH, min-max normalization is applied, and the data number is increased to 16 times. At this stage, the best result is obtained through the 5-layer LSTM structure, and 70-30% hold out method. When the above tables are examined, it is possible to obtain a maximum accuracy of 94,27% using only Morlet wavelet domain features. Min-max normalization should be preferred for the best result. In this stage, a 70-30% hold out method is better than other methods.

4.2.3. Hybrid Features Results

The best results obtained using hybrid features (all 21 features combining the time, frequency and Morlet wavelet domain features) are shown in this section. The results of the situation at which the output number of the LSTM layer is 10 and the activation function is the RELU is shown in Table 11. The best result for this situation is obtained with a 3-layer structure.

The results of the situation at which the output number of the LSTM layer is 10 and the activation function is the TANH is shown in Table 12. The best result for this situation is obtained with a 5-layer structure.

The results of the situation at which the output number of the LSTM layer is 100 and the activation function is the TANH is shown in Table 13.

Table 11. The obtained classification accuracy results of LSTM output:10, ReLU (3 Layers) in the classification of fault diagnosis from the vibration signals based on all 21 features combining the time, frequency and Morlet wavelet domain features

	5 Cross-Validation			10 Cross-Validation			50-50% Hold Out			70-30 % Hold Out		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	48,33	24,16	54,16	50,00	23,33	59,16	30,00	26,66	35,00	25,00	16,66	63,88
2x Data	71,66	35,00	76,66	71,25	32,92	77,92	62,50	24,16	49,16	63,88	29,16	65,28
4x Data	96,46	48,54	83,33	96,87	53,75	84,79	93,75	37,50	80,00	97,92	45,14	79,86
8x Data	97,92	57,18	86,25	98,02	59,58	87,39	95,83	48,75	80,83	96,87	61,11	86,46
16x Data	99,06	64,74	89,53	**99,16**	67,03	90,26	98,75	61,04	88,75	98,78	63,02	89,93

Table 12. The obtained classification accuracy results of LSTM output:10, TANH (5 Layers) in the classification of fault diagnosis from the vibration signals based on all 21 features combining the time, frequency and Morlet wavelet domain features

	5 Cross-Validation			10 Cross-Validation			50-50% Hold Out			70-30 % Hold Out		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	62,50	30,00	70,00	74,17	30,83	72,50	50,00	23,33	48,33	58,33	27,78	72,22
2x Data	82,08	44,17	81,25	84,17	39,17	82,92	65,83	22,50	76,67	80,55	31,94	80,55
4x Data	96,67	63,33	86,04	96,87	62,08	86,46	96,67	46,25	82,92	97,92	58,33	88,89
8x Data	98,02	70,83	87,81	98,26	73,54	88,44	96,67	60,21	85,42	96,53	68,40	88,19
16x Data	98,85	81,41	91,09	**99,27**	82,19	91,82	98,54	77,08	89,48	97,74	76,04	90,80

Table 13. The obtained classification accuracy results of LSTM output:100, TANH (3 Layers) in the classification of fault diagnosis from the vibration signals based on all 21 features combining the time, frequency and Morlet wavelet domain features

	5 Cross-Validation			10 Cross-Validation			50-50% Hold Out			70-30 % Hold Out		
Input Size	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard	minmax	z-score	standard
Raw Data	74.17	28.33	76,67	79,17	32,50	79,17	68,33	31,67	73,33	77,78	25,00	80,55
2x Data	86.67	47.92	85,42	90,42	46,25	85,00	85,83	35,00	80,83	90,28	43,05	81,94
4x Data	98.33	70.42	87,92	98,54	75,00	88,33	97,92	55,00	87,92	97,92	68,75	87,50
8x Data	98.44	80.31	89,68	98,96	83,96	90,73	97,92	74,58	85,62	96,87	79,51	89,58
16x Data	99.43	89.58	92,34	**99,53**	92,45	93,02	98,96	88,75	90,94	99,30	91,15	92,88

The best result for this situation is obtained with a 3-layer structure. According to the results obtained from hybrid features, the best result is obtained when the LSTM output number is 100, the activation function is TANH, min-max normalization is applied, and the data number is increased to 16 times. At this stage, the best result is obtained through the 3-layer LSTM structure and 10-fold cross-validation method.

When the above tables are examined, it is possible to obtain a maximum accuracy

of 99.53% using all features. Min-max normalization should be preferred for the best result. In this stage, the 10-fold cross-validation method is better than other methods.

4.3. Discussion

Compared to previous studies in this area, the LSTM method offers both easier feasibility and higher accuracy ratios. However, it is seen that only the LSTM layer is not sufficient in complex data. Only one study using the drill bit monitoring data set was found in the literature search. This study and our method's results are shown in Table 14. The other studies and their accuracy ratios in the literature are shown in Table 15.

Table 14. The conducted studies using same dataset in literature

Authors	Year	Method	Features	Classification Accuracy (%)
Thirukovalluru at al. [18]	2016	SVM	FFT	99.3
			WPT	96.46
Our Study	**2019**	**Deep LSTM**	**Time-Frequency-Wavelet**	**99.53**

Table 15. The comparison of our method with other classifier algorithms in the classification of fault diagnosis using different datasets in the literature

Authors	Year	Method	Accuracy Ratio (%)
Zarekar, Khajavi & Payganeh [19]	2019	EEMD & MLP	**78.3**
Nerella, Ratnam & Rao [20]	2018	ANN	**93.10**
Yang, Han & Hwang [21]	2005	SVM (All features)	**97.50**
Konar & Chattopadhyay [22]	2011	ANN	**96.67**
Li at al. [23]	2000	ANN	**96.25**
Janssens at al. [24]	2016	CNN	**93.61**
Sathish & Karthick [25]	2018	HAIWF	**67**
Li, Huang & Ji [26]	2019	CNNEPDNN	**98.10**
		CNN	**97.79**
		DNN	**89.89**

As it is done in this paper, LSTM provides high-level results when supported by

different structures. As can be seen in Table 15, the accuracy ratios increase to 90% and above when the number of data is doubled.

5. Conclusions

When the studies in the field of fault diagnosis in rotating machines are examined, it is seen that LSTM and its variations are rarely used. In this study, the LSTM method is successful in applications with sequential input data. It was found that generating different ANN structures without being dependent on the single-layer structure is efficient. Also, the effect of the output number of the LSTM layer and the activation functions used in the layers were observed. As a result of the transactions carried out, the fault diagnosis was carried out successfully with 99.53% accuracy rate. With this study, it has been shown that outstanding results can be obtained in more complex data in this area. When using LSTM in the field of fault diagnosis in rotating machines, it is seen that only the LSTM layer is not sufficient. Also, it is seen that there is no need for algorithms with too many layers in order to get the best results, and if the right choices are made, the best results can be obtained even in 3 layers structures. While the K Fold Cross Validation system is used for classifications, the accuracy rate increases as k number increases. However, the doubling of the number of k increases the processing time by approximately 100%. This is a drawback for large datasets. Also, the number of outputs of the LSTM layer significantly affects the processing time. The effect of normalizing the input data on the result is also observed. The best result is obtained with min-max normalization. With non-normalized data, rarely more than 90% of results can be obtained. As a result of the observations, if the output number of the LSTM layer is set to be more than 100, the result is not changed. It was also found that the pooling layer did not change the result but increased the processing time. This shows that; the algorithms developed for image processing applications, such as CNN, need to be modified to give results in sequential data. Otherwise, no results are obtained or the yield is very low.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] Liu, R., Yang, B., Zio, E., & Chen, X. (2018). Artificial intelligence for fault diagnosis of rotating machinery: A review. Mechanical Systems and Signal Processing, 108, 33-47.

[2] Peter, W. T., Yang, W. X., & Tam, H. Y. (2004). Machine fault diagnosis

through an effective exact wavelet analysis. Journal of sound and vibration, 277(4-5), 1005-1024.

[3] Benbouzid, M. E. H. (2000). A review of induction motors signature analysis as a medium for faults detection. IEEE transactions on industrial electronics, 47(5), 984-993.

[4] Tavner, P. J. (2008). Review of condition monitoring of rotating electrical machines. IET Electric Power Applications, 2(4), 215-247.

[5] Verma, N. K., Sevakula, R. K., Dixit, S., & Salour, A. (2015). Data-driven approach for drill bit monitoring. IEEE Reliab. Mag, 19-26.

[6] Wang, D. (2016). K-nearest neighbors based methods for identification of different gear crack levels under different motor speeds and loads: Revisited. Mechanical Systems and Signal Processing, 70, 201-208.

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

[8] Widodo, A., & Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. Mechanical systems and signal processing, 21(6), 2560-2574.

[9] Haykin, S., & Network, N. (2004). A comprehensive foundation. Neural networks, 2(2004), 41.

[10] Yang, B. S., Han, T., & Hwang, W. W. (2005). Fault diagnosis of rotating machinery based on multi-class support vector machines. Journal of Mechanical Science and Technology, 19(3), 846-859.

[11] https://www.emco-world.com/en/products/industrial-training/machines/milling/cat/26/d/2/p/1000045%2C26/pr/concept-mill-105.html (last accessed: February, 2019)

[12] Kanai, M. (1978). Statistical characteristics of drill wear and drill life for the standardized performance tests. Ann. CIRP, 27(1), 61.

[13] Kumar, A., Ramkumar, J., Verma, N. K., & Dixit, S. (2014, June). Detection and classification for faults in drilling process using vibration analysis. In 2014 International Conference on Prognostics and Health Management (pp. 1-6). IEEE.

[14] https://deeplearning4j.org/docs/latest/deeplearning4j-quickstart (last accessed: February ,2019)

[15] https://deeplearning.cms.waikato.ac.nz/user-guide/getting-started/ (last accessed: February, 2019)

[16] https://deeplearning4j.org/tutorials/10-layers-and-preprocessors (last accessed: February, 2019)

[17] Jing, L., Zhao, M., Li, P., & Xu, X. (2017). A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox. Measurement, 111, 1-10.

[18] Thirukovalluru, R., Dixit, S., Sevakula, R. K., Verma, N. K., & Salour, A. (2016, June). Generating feature sets for fault diagnosis using denoising stacked auto-encoder. In 2016 IEEE International Conference on Prognostics

and Health Management (ICPHM) (pp. 1-7). IEEE.

[19] Zarekar, J., Khajavi, M. N., & Payganeh, G. Roller Bearıng Fault Detectıon Usıng Empırıcal Mode Decomposıtıon And Artıfıcıal Neural Network Methods.

[20] Nerella, M. J., Ratnam, C., & Rao, V. V. (2018). Fault Dıagnosıs Of A Rollıng Element Bearıngs Usıng Acoustıc Condıtıon Monıtorıng And Artıfıcıal Neural Network.

[21] Yang, B. S., Han, T., & Hwang, W. W. (2005). Fault diagnosis of rotating machinery based on multi-class support vector machines. Journal of Mechanical Science and Technology, 19(3), 846-859.

[22] 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.

[23] 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.

[24] Janssens, O., Slavkovikj, V., Vervisch, B., Stockman, K., Loccufier, M., Verstockt, S., ... & Van Hoecke, S. (2016). Convolutional neural network based fault detection for rotating machinery. Journal of Sound and Vibration, 377, 331-345.

[25] Sathish, T., & Karthick, S. (2018). HAIWF-based fault detection and classification for industrial machine condition monitoring. Progress in Industrial Ecology, an International Journal, 12(1-2), 46-58.

[26] Li, H., Huang, J., & Ji, S. (2019). Bearing Fault Diagnosis with a Feature Fusion Method Based on an Ensemble Convolutional Neural Network and Deep Neural Network. Sensors, 19(9), 2034.

[27] Saravanan N, Cholairajan S & Ramachandran KI (2009) “Vibration-Based Fault diagnosis of Spur Bevel Gear Box Using Fuzzy Technique”, Expert systems with applications, 36(2): 3119-3135.

[28] Sakthivel NR, Sugumaran V & Babudevasenapati S (2010) “Vibration Based Fault Diagnosis of Monoblock Centrifugal Pump Using Decision Tree”, Expert Systems with Applications, 37(6): 4040-4049.

[29] Pandya DH, Upadhyay S & Harsha SP (2012) “ANN Based Fault Diagnosis of Rolling Element Bearing Using Time-Frequency Domain Feature”, International Journal of Engineering Science and Technology, 4(6): 2878-2886.

[30] Seshadrinath J, Singh B & Panigrahi BK (2013) “Investigation of Vibration Signatures For Multiple Fault Diagnosis in Variable Frequency Drives Using Complex Wavelets”, IEEE Transactions on Power Electronics, 29(2): 936-945.

[31] Jafari SM, Mehdigholi H & Behzad M (2014) “Valve Fault Diagnosis in Internal Combustion Engines Using Acoustic Emission and Artificial Neural Network”, Shock and Vibration, 2014.

[32] Ali JB, Fnaiech N, Saidi L, Chebel-Morello B & Fnaiech F (2015) “Application of Empirical Mode Decomposition and Artificial Neural Network For Automatic Bearing Fault Diagnosis Based on Vibration Signals”,

Applied Acoustics, 89: 16-27.

[33] Janssens O, Slavkovikj V, Vervisch B, Stockman K, Loccufier M, Verstockt S, deWalle RV & Van Hoecke S (2016) “Convolutional Neural Network Based Fault Detection For Rotating Machinery”, Journal of Sound and Vibration 377: 331-345.

[34] Liu H, Zhou J, Zheng Y, Ji, Zhang Y (2018) “Fault Diagnosis of Rolling Bearings with Recurrent Neural Network-Based Autoencoders”, ISA transactions,77: 167-178.

[35] Jing L, Zhao M, Li P & Xu X (2017) “A Convolutional Neural Network Based Feature Learning and Fault Diagnosis Method For the Condition Monitoring of Gearbox”, Measurement, 111: 1-10.

[36] S. Kiranyaz, A. Gastli, L. Ben-Brahim, N. Al-Emadi and M. Gabbouj, "Real-Time Fault Detection and Identification for MMC Using 1-D Convolutional Neural Networks," in IEEE Transactions on Industrial Electronics, vol. 66, no. 11, pp. 8760-8771, Nov. 2019.

[37] Kemal Polat, Kaan Onur Koc (2020). Detection of Skin Diseases from Dermoscopy Image Using the combination of Convolutional Neural Network and One-versus-All. Journal of Artificial Intelligence and Systems, 2, 80–97. ttps://doi.org/10.33969/AIS.2020.21006.

[38] Murat Arican, Kemal Polat (2020). Binary particle swarm optimization (BPSO) based channel selection in the EEG signals and its application to speller systems. Journal of Artificial Intelligence and Systems, 2, 27–37. https://doi.org/10.33969/AIS.2020.21003.

[39] D. Jude Hemanth (2020). EEG signal based Modified Kohonen Neural Networks for Classification of Human Mental Emotions. Journal of Artificial Intelligence and Systems, 2, 1–13. https://doi.org/10.33969/AIS.2020.21001.

[40] Jatin Arora, Utkarsh Agrawal, and Prerna Sharma (2020). Classification of Maize leaf diseases from healthy leaves using Deep Forest. Journal of Artificial Intelligence and Systems, 2, 14–26. https://doi.org/10.33969/AIS.2020.21002.