ISSN: 1231-4005
e-ISSN: 2354-0133
DOI: 10.2478/kones-2019-0060

ROTATING SHAFT FAULT PREDICTION USING CONVOLUTIONAL NEURAL NETWORK: A PRELIMINARY STUDY

Davor Kolar, Dragutin Lisjak

University of Zagreb
Faculty of Mechanical Engineering and Naval Architecture
Ivana Lučića Street 5, 10002 Zagreb, Croatia
tel.: +385 01 6168 308, +385 01 6168 377
e-mail: dkolar@fsb.hr, dragutin.lisjak@fsb.hr

Michał Pająk

University of Technology and Humanities in Radom
Department of Thermal Technology
Stasieckiego Street 54, 26-600 Radom, Poland
tel.: +48 48 3617149
e-mail: m.pajak@uthrad.pl

Abstract

One of the most important subsystems of the vehicles and machines operating currently in industry and transportation are the rotating subsystems. During the operation, due to the forcing factors influence, the technical state of them is changing and the failure can occur. Fault diagnosis is maintenance task considered as an essential in such subsystems, since possibility of an early detection and diagnosis of the faulty condition can save both time and money. To do this the analysis of the subsystems vibrations is performed. The identified technical state should be considered in a context of the ability and different inability states. Therefore, the first step of the diagnostic procedure is the ability and different inability states identification.

Traditional data-driven techniques of fault diagnosis require signal processing for feature extraction, as they are unable to work with raw signal data, consequently leading to need for both expert knowledge and human work. The emergence of deep learning architectures in condition-based maintenance promises to ensure high performance fault diagnosis while lowering necessity for expert knowledge and human work. This article presents authors initial research in deep learning-based data-driven fault diagnosis of rotating subsystems. The proposed technique input raw three-axis accelerometer signal as high-definition image into deep learning layers, which automatically extract signal features, enabling high classification accuracy.

Keywords: *condition-based maintenance, rotating systems, fault diagnosis, convolutional neural networks*

1. Introduction

Rotating machines in general consist of three major parts: a rotor, rolling or journal bearings (fluid or anti-friction bearings) and a foundation. Rotating subsystems are one of the most important elements used in different kind of vehicles and machines in order to transform the energy and transmit the power. Since rotary machinery usually operates under a tough working environment, it makes it more vulnerable to various types of faults and increases the complexity of fault diagnosis. Both studies and experience show that faults develop and occur in rotating machines during normal operation results in not only the loss of productivity but also in the delayed delivery of goods and services to customers and may even lead to safety, economic and environmental problems.

Vibrations produced by rotary machinery elements occur irrespective of type of rotary machinery or transport equipment. In general, it can be concluded that vibrations are produced by shafts, axes, fans, pumps and turbines depending on the type of subsystems. Vibration in any rotating machinery is caused by faults like imbalance, misalignment, crack, etc.

Vibration monitoring is considered leading technique for rotary equipment condition detection and diagnostics [1]. In past few years, many techniques for signal processing and extraction of information in fault diagnosis was titled in research, primarily focusing in improving the currently available (traditional) or developing new techniques. Additionally, the development of techniques supported by artificial intelligence and application in the field of maintenance by the state have demonstrated their better performance compared to analytical models with classic approaches [2].

Due to the recent increase for the amount of data collected, more and more effort is being invested in the development of techniques for computing the condition of rotation equipment. In recent years, deep learning techniques have achieved huge success in image [3, 4] and speech [5], [6] recognition.

Deep learning stands for class of machine learning techniques specific by its many layers of information processing stages in deep architectures that are exploited for pattern classification and other tasks [7].

Authors focused their efforts on vibration signals time-domain analysis. Firstly, using the experimental simulation stand, the experiment simulations described in chapter 2 were performed in order to record vibrations signals of rotational subsystem operating in ability and different inability states.

For the purposes of research, a convolutional neural network with the ability to capture collected data without further pre-processing or features calculation has been developed. Design of the neural network is described in chapter 3. Chapter 4 presents training process and results of the experiment performed. The article is summed up by some conclusions formulated in chapter 6.

2. Performed experiment simulations

In the study, the vibration signals acquired from a machine fault simulator were used. A SpectraQuest variable speed Machinery Fault Simulator (*MFS*) was used to generate both normal operation (*NS*) and faulty condition data. The simulation stand (Fig. 1) comprised 1 *HP* variable speed motor driving a shaft-rotor component via coupling supported with two sets of ball bearings. The *MFS* is outfitted with three-axis accelerometer and a tachometer that were connected to a National Instruments *DAQ* System.

Three-axis accelerometer was mounted on the bearing housing on the shaft side opposite of the motor position. The sampling frequency was set to 6.4 kHz, while revolving speed during the experiment was 1500 rpm. Vibration signals in three directions (X, Y, Z) were acquired when the system operated under normal condition (*NS*) and faulty conditions. There were two faulty shaft conditions simulated: eccentric rotor fault (*ERF*) and unbalanced rotor fault (*IMRF*). Operation under normal conditions was interpreted as operating of the rotation subsystem remaining in ability state while operation under any of faulty conditions was interpreted as operating of the rotation subsystem remaining in inability state.

Each acquired sample of 6400 data points is stored as dataset representing state. Vibration signals under three different working conditions are used in this study. They are divided into training and testing datasets separately, which are randomized before being used in training and testing the model. The descriptions of them are listed in Tab. 1.

Tab. 1. Simulated fault conditions

No.	Condition	Description
1	Normal state	Machine is running without simulated fault
2	Unbalanced rotor	Machine is running with simulated fault of imbalance on main shaft
3	Eccentric rotor	Fault is simulated by adding eccentric rotor on main shaft

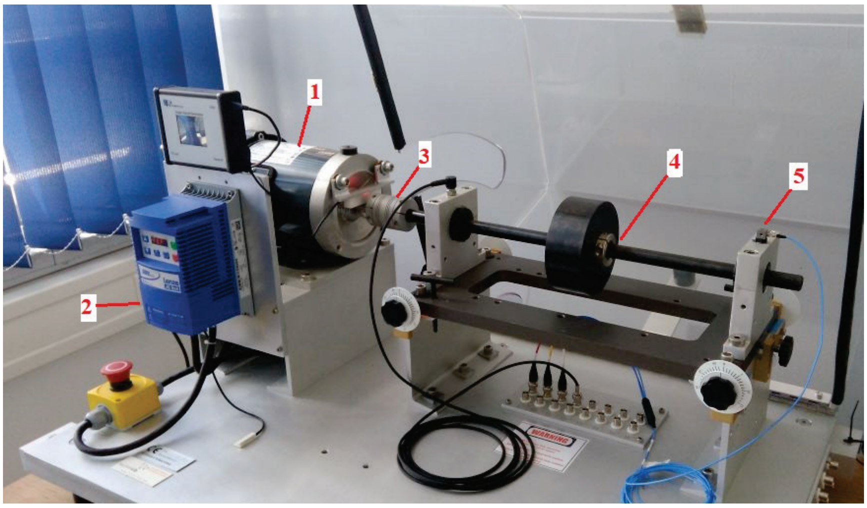

Fig. 1. Fault simulator: 1 – Three-phase induction motor, 2 – Variable speed motor drive 3 – clutch, 4 – main shaft with load, 5 – 3-axis accelerometer

Each vibration sample comprising 6400 values was stored as a separate vibration signal. The 3150 datasets have been collected to train the convolutional neural network data-driven model for failure classification and separately 1350 measurement has been made to collect test data. Tab. 2 illustrates the data composition of collected samples. From all the samples, 70% of the data is used for training and validation during training while rest of 30% is used for testing the model. The 10% of training data is used for validation during training. The samples for training, testing, and validation during the experiment were selected randomly.

Tab. 2. Composition of collected samples for fault classification

Machinery state	Training samples	Test samples	Sum
Normal working condition	1050 samples	Group 1 (150 samples)	4500 datasets collected (28 800 000 data points)
		Group 2 (150 samples)	
		Group 3 (150 samples)	
Unbalanced rotor	1050 samples	Group 1 (150 samples)	
		Group 2 (150 samples)	
		Group 3 (150 samples)	
Eccentric rotor	1050 samples	Group 1 (150 samples)	
		Group 2 (150 samples)	
		Group 3 (150 samples)	

Convolutional neural network training is done on GPU of our machine learning platform that consist of Intel i7-7700 CPU, 32GB of RAM and CUDA capable GeForce RTX 2070 graphics card with 2304 Cuda Cores and 1620 MHz base clock.

3. Convolutional Neural Network Design

Convolution Neural Networks (CNN) are a type of artificial neural network that are adapted to relatively fast and efficient resolution of the problems of high-dimensional inputs or inputs that possess a multitude of features with its internal structure. Its main application CNN has gained in processing, classifying and recognizing objects in the images, but recently examples of the application of such architecture in the maintenance field can be found [8-10].

CNN is a variant of the feed-forward multi-layer neural network and is primarily designed for processing image data in the form of a matrix, taking into account local and global stationary properties [11]. Its structure is similar to a classical multilayer perceptron or a classical feed-forward artificial neural network. CNN is a network composed of layers, in which the output of the previous layer is connected to the input of the next with a set of parameters that can be learned. The main difference with respect to the multilayer perceptron is that each layer is represented as a set of input and output mapping features, calculated by the convolution procedure, with the aim of covering different perspectives of input data.

Generally, the operation of the convolutional network can be divided into several main operations: *Convolution*, *Activation function*, *Pooling*, followed by one or more *fully connected layers* aimed to classify the features learned in convolutional layers, as shown in Fig. 2.

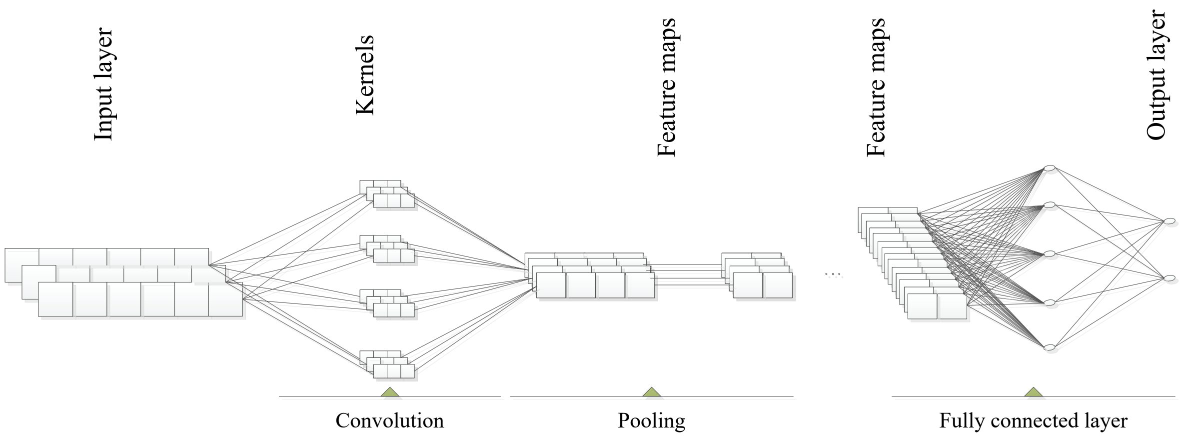

Fig. 2. Structure of convolutional neural networks

Tab. 3. Convolutional neural network layers activations and parameters

Layer	**Description**	**Activations**	**Learnable parameters**
Input layer	Signal input	6400 x 1 x 3	-
Convolutional layer 1	Number of kernels: *k*/2 Kernel size: *k* x 1 x 3 Stride: [1 1]	6400 x 1 x *k*	Weights *k* x 1 x 3 x *k*/2 Biases 1 x 1 x *k*/2
Batch normalization 1	Batch size: 128	6400 x 1 x *k*/2	Offset 1 x 1 x *k*/2 Scale 1 x 1 x *k*/2
Activation Layer 1	ReLU	6400 x 1 x *k*/2	-
Pooling layer 1	Max Pooling [2 1] Stride: [1 1]	6399 x 1 x *k*/2	-
Convolutional layer 2	Number of kernels: 16 Kernel size: *k*/2 x 1 x 4 Stride: [1 1]	6399 x 1 x *k*	Weights *k*/2 x1 x *k*/2 x *k* Biases 1 x 1 x *k*
Batch normalization 2	Batch size: 128	6399 x 1 x *k*	Offset 1 x 1x *k* Scale 1 x 1x *k*
Activation Layer 2	ReLU	6399 x 1 x *k*	-
Pooling layer 2	Max Pooling [2 1] Stride: [1 1]	6398 x 1 x *k*	
Convolutional layer 3	Number of kernels: *k* Kernel size: *k*/2 x 1 x 4 Stride: [1 1]	6398 x 1 x *k*	Weights *k*/2 x 1 x *k* x *k* Biases 1 x 1 x 16
Batch normalization 3	Batch size: 128	6398 x 1 x *k*	Offset 1 x 1x *k* Scale 1 x 1x *k*
Activation Layer 3	ReLU	6398 x 1 x *k*	-
Pooling layer 3	Average Pooling [4 1] Stride: [1 1]	6395 x 1 x *k*	
Fully connected layer	3 fully connected layer	1 x 1 x 3	Weights 3 x (6395 *k*) Biases 3 x 1
SoftMax activation	Softmax function	1 x 1 x 3	-
Output layer	Classification output	-	-

The CNN structure in this study contains three alternating convolutional and pooling layers with one fully connected layer followed by SoftMax activation function and classification output layer. By using such a combination of layers, all 6400 univariate time series points spread across 3 channels from each of the samples are used for feature learning.

Best combination of hyper parameters was sought by testing the grid of numbers k = [4 8 12 16], respectively. For this research, hyper parameters named number of kernels and kernel size were considered for grid search. First convolutional layer output consists of $k/2$ feature maps calculated using $k/2$ number of kernels with size k x 1 x 3, that are translated into second layer inputs. Further on, second and third convolutional layer consist of k number of kernels where k feature maps with kernel size of $k/2$ x 1 x 3 layer are calculated. The activations and learnable parameters of each layer along with description of each layer are presented in Tab. 4.

4. Training and results

Neural network learning is an iterative procedure for determining the weights and biases of neurons in the network. The developed CNN adjusts learnable parameters by minimizing previously defined loss function (1).

$$E = -\sum_t \sum_k y_k^*(t) \log(y_k(t)), \tag{1}$$

where: $y_k^*(t)$ and $y_k(t)$ are the targets and predicted values of the t-th training example of the k-th class, respectively. Widely used backpropagation algorithm is used to minimize the loss function by calculating stochastic gradient descent that allows network to update parameters during training. For the purpose of research constant learning rate of 0.005 and momentum of 0.95 were used, respectively.

As defined previously, 4 different kernel size and number of kernels were used to obtain better network structure. Network with different values of k were developed and trained for 10 consecutive times. After each training iteration, trained network was evaluated by calculating accuracy on every test group and all calculated values were stored.

Tab. 4. Hyper parameters grid search – accuracies

k	Test 1	Test 2	Test 3	Average	Worst	Best	StDev
4	93.23%	92.90%	92.66%	92.93%	78.17%	98.12%	0.0786
8	99.83%	99.90%	99.73%	99.82%	99.33%	100%	0.0038
12	99.93%	99.93%	99.96%	99.94%	99.83%	100%	0.001
16	99.87%	99.83%	99.90%	99.87%	99.33%	100%	0.0025

Table 4 presents calculated averages of accuracies for each test group, followed by average, worst and best values and standard deviation obtained for each k value. It can be seen that CNN with k = 4, i.e. k smaller than 8 produces considerably poorer than networks with larger number k. It can be concluded, that network with k = 4, which has only 2 feature maps in first convolutional layer, is not capable of adjusting learnable parameters to efficiently classify test set data. On the other hand, networks with larger k were able to estimate class with larger accuracy. Best evaluation results for each k value larger than 4 were 100% on each of the test group samples. Further on, networks trained with k = 12 yields lowest standard deviation altogether with greatest average among all test samples. Although training performed with k = 16 resulted second best score in both average accuracy and standard deviation, higher k value means longer training and higher potential of overfitting due to quantity of learnable parameters. CNN-s are widely presented as black-box solutions and they are somewhat hard to understand inner operating mechanisms. To better understand how features are being generated, visualization of the CNN layers activations for network with k = 12 and 100% accuracy on all three test sets has been done.

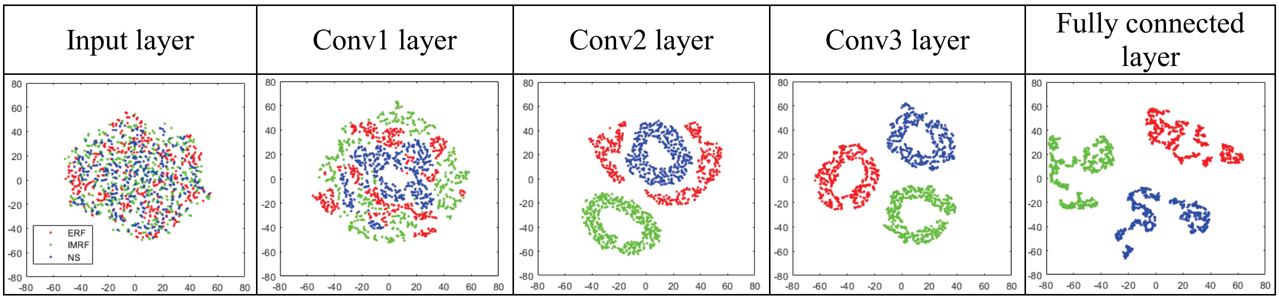

Fig. 3. t-SNE feature representations

For the purpose of visualizing this high dimensional data, t-SNE [12] algorithm with is implemented. The t-SNE is used for dimensionality of high dimensional points represent layer activations point for all test data samples combined. Nearby points in the high-dimensional space correspond to nearby embedded low-dimensional points, and distant points in high-dimensional space correspond to distant embedded low-dimensional points. Learned feature representation is presented in Fig. 3.

By looking from the input layer through convolutions, it can be more clearly seen features become extracted in the form of layer activations and divided into colour clusters representing classes as we are moving to the fully connected layer.

5. Summary and conclusions

The research presented in the article sums up authors' preliminary study in convolutional neural network application for rotary machinery intelligent fault diagnosis. In order to identify ability state named normal working conditions and specified inability states (i.e. early faults) the experimental tests were accomplished. During the test, vibration signals were recorded. Recorded accelerometer signals in time domain were divided into learning and testing sets without calculating any additional signal feature. Algorithm for automatic feature extraction and rotary machinery state classification based on convolutional neural network is designed and trained. Grid search is used for hyper parameters optimization.

Trained networks were tested on experimental data collected in Laboratory for Maintenance of University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture.

Results shows great potential of the proposed CNN technique in the data-driven fault diagnosis field, especially since vibration signals from three-axis accelerometer enters model without any time-consuming manual feature extraction. Based on the results of the executed tests, where best-trained network performed well on the testing data sets and obtained 100% accuracy, it was stated that the proposed fault diagnosis technique is precise enough to be the object of further industrial research.

Additional testing of proposed technique on different types of failures and on known datasets is essential for performance comparison. Further, on selecting optimal hyper parameters is still a challenge. Finally, training process of developed is time demanding and using GPU hardware is highly advisable. Considering those, future work will be based on additional testing of the technique.

References

[1] Sinha, J. K., Elbhbah, K., *A future possibility of vibration based condition monitoring of rotating machines,* Mechanical Systems and Signal Processing, Vol. 34, No. 1-2, pp. 231-240, 2013.

[2] Yan, J., *Machinery prognostics and prognosis oriented maintenance management*. Singapore: John Wiley & Sons Singapore Pte. Ltd, 2015.

[3] Chan, T.-H., Jia, K., Gao, S., Lu, J., Zeng, Z., Ma, Y., *PCANet: A Simple Deep Learning Baseline for Image Classification?*, IEEE Transactions on Image Processing, Vol. 24, No. 12, pp. 5017-5032, 2015.
[4] He, K., Zhang, X., Ren, S., Sun, J., *Deep Residual Learning for Image Recognition,* in *2016* IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, USA, Las Vegas, NV 2016.
[5] Huang, P.-S., Kim, M., Hasegawa-Johnson, M., Smaragdis, P., *Deep learning for monaural speech separation,* in *2014* IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1562-1566, Florence, Italy 2014.
[6] Noda, K., Yamaguchi, Y., Nakadai, K., Okuno, H. G., Ogata, T., *Audio-visual speech recognition using deep learning,* Applied Intelligence, Vol. 42, No. 4, pp. 722-737, 2015.
[7] Schmidhuber, J., "Deep learning in neural networks: An overview," *Neural Networks*, Vol. 61, pp. 85-117, 2015.
[8] Shaheryar, A., Yin, X.-C., Yousuf, W., *Robust Feature Extraction on Vibration Data under Deep-Learning Framework: An Application for Fault Identification in Rotary Machines,* International Journal of Computer Applications, Vol. 167, No. 4, pp. 37-45, 2017.
[9] Li, S., Liu, G., Tang, X., Lu, J., Hu, J., *An Ensemble Deep Convolutional Neural Network Model with Improved D-S Evidence Fusion for Bearing Fault Diagnosis,* Sensors, Vol. 17, No. 8, p. 1729, 2017.
[10] Shao, S., Sun, W., Wang, P., Gao, R. X., Yan, R., *Learning features from vibration signals for induction motor fault diagnosis*, pp. 71-76, 2016.
[11] Le Cun, Y., Bengio, Y., Hinton, G., *Deep learning*, Nature, Vol. 521, No. 7553, pp. 436-444, 2015.
[12] Van der Maaten, L., Hinton, G., *Visualizing non-metric similarities in multiple maps,* Machine Learning, Vol. 87, No. 1, pp. 33-55, 2012.

Manuscript received 15 July 2019; approved for printing 03 September 2019