DOI 10.5162/sensoren2019/5.4.2

Feature-based analysis of reproducible bearing damages based on a neural network

Andreas Beering, Jakob Döring, Karl-Ludwig Krieger
Institute of Electrodynamics and *Microelectronics (ITEM.ae), University of Bremen, Germany*
Email: {beering, doering, krieger}@item.uni-bremen.de

Abstract

In the interest of detecting damage to tapered roller bearings at an early stage and avoiding further consequential damage in future, an investigation of reproducible damage to bearing inner races, outer races and rolling elements is carried out. The vibration signals generated by the contact of the damaged surfaces during bearing runtime are recorded via a piezoelectric vibration transducer. Different scenarios with regard to the rotational speed and the size of the damage are investigated. Based on a calculation of features, a feed-forward neural network is trained and used to classify the damage. To assess the quality of the neural network, the receiver operator characteristic (ROC) and the area under curve (AUC) are compared for the neural network as well as for other popular classification algorithms such as support vector machine (SVM), decision tree and k-Nearest-Neighbor (KNN). Based on the AUC values, this approach showed that the neural network used has the best performance for the classification of bearing damage with an AUC of 0.994 and an overall classification accuracy of 93.5 %.

Keywords: Bearing faults, feed-forward neural network, piezoelectric vibration transducer, classification performance, condition monitoring

I. Introduction

In the field of condition monitoring, a large number of machine components are already monitored in order to optimize processes or detect damage. However, many areas have not yet been sufficiently investigated, so that avoidable machine failures still occur. This includes undetected bearing or gear damage in agricultural machines, which can result in fatal consequential damage and, due to the repair and downtime of the machine, cause considerable costs. A solution to avoid such costs in future could be the detection of damage by an analysis of vibration data.

Previous approaches to the detection of faults on rotation machine elements have largely been based on envelope analysis of vibration signals. For this analysis, the theoretical damage frequencies of the bearing, which result from the geometry and the rotation speed, are calculated first [1]. Subsequently, the vibration signal generated by the damaged bearing is recorded and the envelope curve of the signal is calculated. From the frequency bands occurring in the amplitude spectrum and the theoretically defined damage frequencies, the damage to the bearing or the gear is concluded [2, 3]. Another approach to the investigation of bearing damage is based on empirical mode decomposition (EMD) to obtain a sum of intrinsic mode functions (IMF) followed by feature extraction and analysis [4]. A further approach to the detection of bearing damage uses an envelope analysis, followed by a classification using a decision tree [5].

This publication presents a study of reproducible artificial damage to tapered roller bearings. The damage is caused to the rolling elements, the inner races and the outer races of the bearings in defined sizes. The resulting vibration signals of the differently damaged bearings are measured on a laboratory test bench. For this purpose, the bearing is inserted into a bearing socket, to which a piezoelectric vibration transducer is attached by a screw connection, and driven at a defined speed via a three-phase motor and a frequency inverter. Feature extraction is performed for the acquired vibration signals. Not only common features from the classification of rotating machines [6], but other acoustic features such as the spectral crest factor or the mel-frequency cepstrum coefficients (MFCCs) are used. Once the features have been calculated, a feed-forward neural network is trained based on a training data set with labelled data. Afterwards, the feed-forward neural network is evaluated and compared with other classifiers using the ROC and AUC.

The remainder of this paper is organized as follows. In Section II the feature extraction is outlined and the theoretical fundamentals of the used feed-forward neural networks are described. Furthermore, the receiver operator characteristic and the area under curve are

described, which are regarded as measures for the performance of a classifier. Section III outlines the structure of the used tapered roller bearing with the reproducible induced damage. Afterwards, the monitoring setup is presented. In Section IV the experimental results are discussed and the used feed-forward neural network is compared to other classification algorithms. Finally, a conclusion and an outlook on future work is drawn in Section V.

II. Classification algorithm and performance

A. *Feature extraction*

Feature extraction describes the representation of characteristic pattern properties for the purpose of differentiating pattern classes. This also ensures that irrelevant information is not represented and the effective amount of data is reduced. The majority of machine learning classification algorithms uses features as input parameters to assign patterns to specific classes [7]. In this approach known features from the analysis of rotating machines such as kurtosis, skewness or variance are calculated [6]. In addition, features from the field of speech processing such as the mel-frequency cepstrum coefficients, the spectral crest factor or various features for the evaluation of the signal power are calculated. Overall, a total of 71 features are calculated from the vibration data of the bearing and used as inputs for the feed-forward neural network.

B. *Feed-forward neural network*

Neural networks can be described as a series of functional transformations, which combine input signals with weights and biases for the classification of several classes. Feed-forward neural networks represent a special case of neural networks. They allow the signals to propagate exclusively from input to output without feedback. The two-layer feed-forward network used in this paper is shown in Fig. 1 and consists of 71 inputs, 10 neurons in the hidden layer, 4 neurons in the output layer and 4 outputs. For the inputs of the neural network, a total of 71 features in the time and frequency domain have been calculated.

The basic neural network model to describe the first layer of the network is given as

$$a_j = \sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)} \qquad (1)$$

where D describes the number of input signals x_i of the network, j the number of neurons in the first layer, $w_{ji}^{(1)}$ the weights and $w_{j0}^{(1)}$ the biases. The resulting values a_j are known as activations.

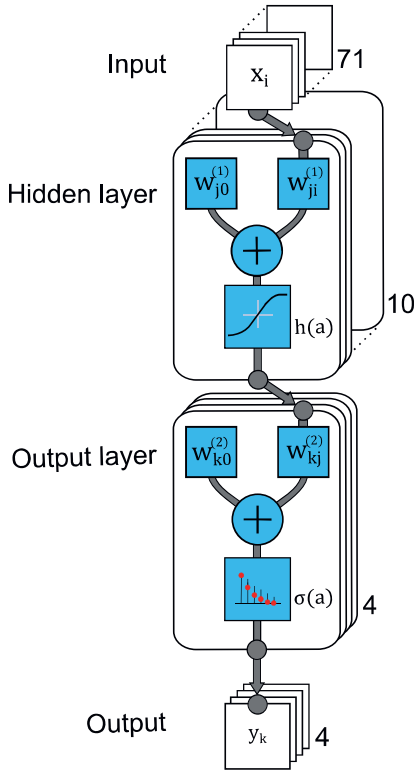

Fig. 1: Feed-forward neural network structure with 71 inputs, a hidden layer with 10 neurons and an output layer with 4 neurons for the classification of 4 classes.

Each activation is transformed using a nonlinear activation function h to form

$$z_j = h(a_j). \qquad (2)$$

In this approach a hyperbolic tangent sigmoid transfer function is used, which is represented by the following term

$$h(a) = \frac{2}{1 + e^{-2a}} - 1. \qquad (3)$$

The second layer of the network can be described by an analogous consideration to form

$$a_k = \sum_{j=1}^{M} w_{kj}^{(2)} z_j + w_{k0}^{(2)} \qquad (4)$$

where M describes the number of inputs z_j from the first layer, k the number of neurons of the second layer, $w_{kj}^{(2)}$ the weights and $w_{k0}^{(2)}$ the biases. To give a set of network outputs y_k, the activations are transformed using an appropriate activation function to form

$$y_k = \sigma(a_k). \qquad (5)$$

The activation function for the second layer used in this approach is the softmax activation function which can be described as

$$\sigma(a) = \frac{2}{1 + e^{-a}}. \tag{6}$$

The whole feed-forward neural network can be described as a combination of the previous stages and takes the form the following equation

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} h\left(\sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right). \tag{7}$$

The output y_k of the network is given as a non-linear function of a set of input signals $\mathbf{x}$ controlled by a vector of adjustable weights $\mathbf{w}$ [7].

C. Cross-validation

Cross-validation is a technique used to evaluate the performance of a model in machine learning. With new data sets, which were not used in the training phase, the performance of the prediction is tested. This is done by partitioning a data set into subsets for training and testing the algorithm. Since cross-validation does not use all data when developing a model, it is a commonly used method to prevent over-fitting during training. One of the most common cross-validation techniques is k-fold cross-validation. The data set is divided into k subsets of approximately the same size. Each subset is used to validate the model, while the remaining subset is used for training. In total, this process is repeated k times to use each subset once for validation. The k validation runs allow the mean performance of the algorithm to be determined [9].

D. Receiver operating characteristic

A receiver operating characteristic (ROC) curve illustrates the relative trade-off between benefits (true positives) and costs (false positives) of a classification model at all classification thresholds. A ROC diagram represents the true positive rate over the false positive rate, as shown in Fig. 2.

The true positive rate TPR is given by

$$TPR = \frac{TP}{TP + FN} \tag{8}$$

where TP are the true positive classified samples and FN the false negatives. The false positive rate FPR is given by

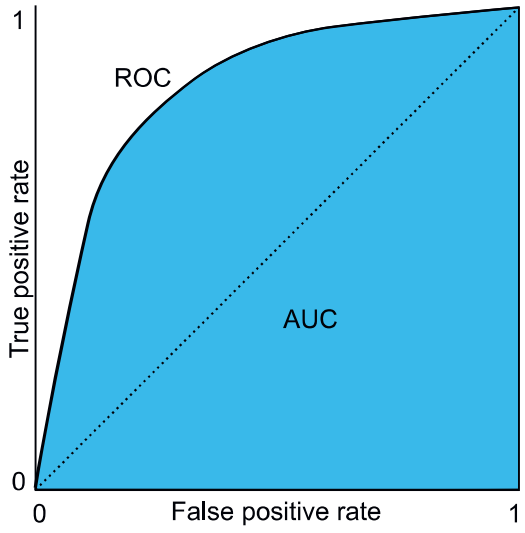

Fig. 2: Receiver operating characteristic (ROC) curve and the area under curve (AUC)

$$FPR = \frac{FP}{FP + TN} \tag{9}$$

where FP are the false positive classified samples and TN the true negatives [10]. The dotted line in Fig. 2 represents the strategy of randomly guessing a class. Classifiers whose ROC curves lay in the lower right triangle are understood as misinterpretation of the data. Whereas classifiers whose ROC curves lay in the upper left triangle are considered good classifications, depending on the size of the area under the ROC curve (AUC) is. Since the AUC is a portion of the unit square, its value will be between 0 and 1. For the strategy of random guessing, the AUC value is 0.5. The height of the AUC value thus describes the quality of the classifier, whereby a value of 1 represents the ideal classifier [11].

III. Experimental Methodology

A. Tapered roller bearing structure

Tapered roller bearings are a special type of roller bearings that can support both axial and radial forces. Due to the rollers being situated perpendicular to the rolling axis, these bearings can transmit higher forces than deep groove ball bearings. The essential components of such a bearing consist of the outer race, the rolling elements, the inner race and the cage. The cage holds the rolling elements on the running surface of the inner race. The outer race is detached from the remaining part of the bearing. During the rotation of the bearing, the greatest forces are transmitted between the running surfaces of the inner race, the outer race and the

rolling elements. The schematic structure of the investigated bearing is shown in Fig. 3.

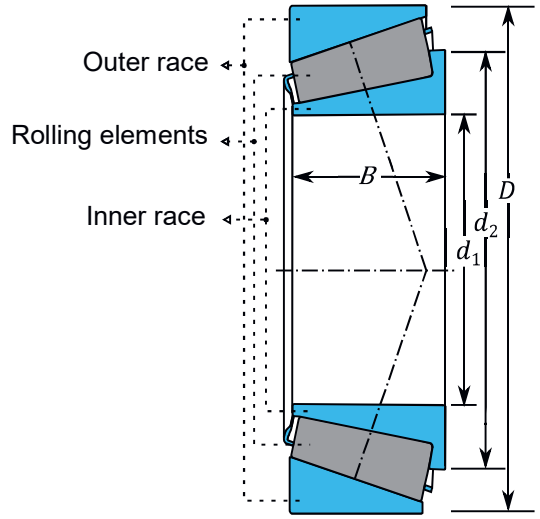

Fig. 3: Schematic structure of the investigated tapered roller bearing.

The bearing used in this paper has an inner diameter of d_1 = 35 mm, a diameter from the center to the rolling elements of d_2 = 51.95 mm, an outer diameter of D = 72 mm, and a width of B = 17 mm. Due to the force applied to the contact surfaces, these surfaces are most susceptible to damage [12]. Within the scope of this paper, damage to the outer race, inner race and rolling element surfaces is specifically investigated.

B. Reproducible bearing damage

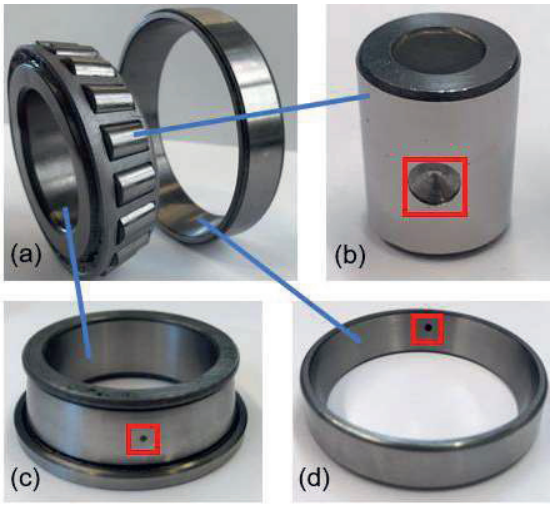

Fig. 4. Entire tapered roller bearing (a) and its individual parts with the artificially induced damage to the rolling element (b), the inner race (c) and the outer race (d).

A total of twelve damaged bearings are being investigated. Four different bearings have been examined for an outer race damage, four for an inner race damage and four for a rolling element damage. The damage is introduced into the inner races, outer races and one rolling elements via erosion. Each class of damage is examined individually. For each damage class a damage of 1.5 mm, 2 mm, 2.5 mm and 3 mm diameter is examined. Fig. 4 (a) shows the tapered roller bearing type examined and the positions of the respective damage for the individual test objects. Fig. 4 (b) shows a rolling element damage, Fig. 4 (c) an inner race damage and Fig. 4 (d) an outer race damage.

C. Experimental setup

In order to investigate the artificial damage, a laboratory test bench has been set up, which is shown in Fig. 5. A bearing socket has been manufactured which can hold different bearing sizes. Piezoelectric vibration transducers can be attached to this by means of screw connections with a defined prestressing force to ensure a good signal transmission [13]. A defined radial force can be applied to the bearing's outer ring by means of a screw attached to the bearing block. The bearing, which is set in the socket, is connected with a 450 W three-phase motor via a low-vibration coupling. A frequency inverter is used to set the speed of the three-phase motor to an adjustable value.

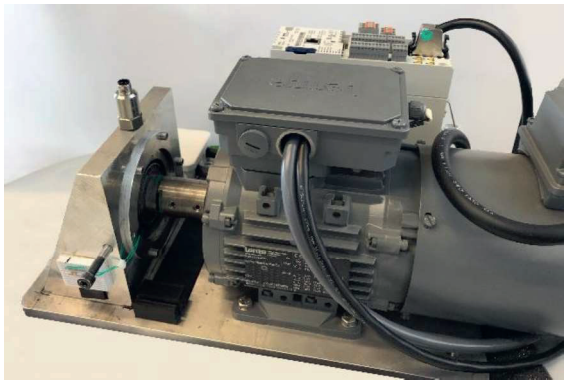

Fig. 5: Test bench for the investigation of bearing damages consisting of a three-phase motor, a frequency inverter, a bearing socket and a piezoelectric vibration transducer.

A total of 4 intact tapered roller bearings and 12 with artificial damage have been examined. Damage in the form of circular surfaces with diameters of 1.5 mm, 2 mm, 2.5 mm and 3 mm have been investigated for each of the three damage classes (outer race, inner race, rolling elements). All tested bearings have been investigated on the laboratory test bench at speeds of 100 min^{-1}, 200 min^{-1}, 300 min^{-1}, 400 min^{-1}

and 500 min^{-1}. The piezoelectric vibration transducer (IDS Innomic, ICS80) mounted on the bearing block via a screw connection has a linear frequency range up to approximately 10 kHz. In order to cover the entire linear range of the sensor, a sampling rate of 20 kHz is used.

IV. Experimental results

With respect to the classification algorithms, the four bearing states are combined into classes according to Tab. 1.

Tab. 1: List of the damage assigned to the classes for the classification

	Class description
Class 1	Intact bearing
Class 2	Inner race damage
Class 3	Outer race damage
Class 4	Rolling element damage

For each class, a data set of 3440 time windows with a length of 2048 sample points and a sampling rate of 20 kHz is examined. Each measurement series for a class contains measurements of five different rotational speeds for four bearings each. For the investigation of the classification algorithms a 10-folded cross-validation has been used.

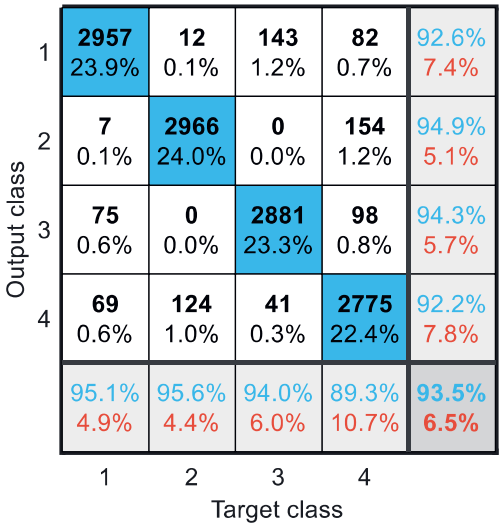

Fig. 6: Confusion matrix of the test data set classified via the feedforward neural network.

Fig. 6 shows the average confusion matrix for the classification of the test data set via the feed-forward neural network after the 10-folded cross-validation. The rows correspond to the predicted class of the neural network and the columns to the true class. Accordingly, the diagonal fields show the correctly classified time windows and the non-diagonal fields show the incorrectly classified time windows. The column on the right shows the percentages of all time windows belonging to each class that are correctly or incorrectly classified. The bottom row shows the percentages of all time windows belonging to each class that are correctly and incorrectly classified. The field at the bottom right shows the total accuracy.

The classification of the neural network is very accurate for all classes with an accuracy for each class of at least 89.3 %. The overall accuracy of the network is 93.5 %, which can be interpreted as a very good prediction of the damage investigated via the network.

In order to compare the classification of the feed-forward neural network with other classification algorithms, a linear support vector machine, a fine decision tree and a k-nearest neighbor (k = 10) classification have been evaluated. For each of these classification algorithms the same data set and features have been used.

For the comparison of the classifier performance the receiver operating characteristic (ROC) curve for each classification algorithm has been determined.

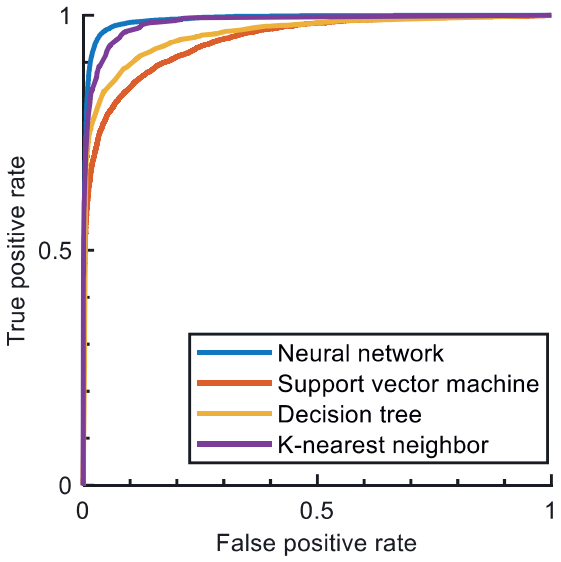

Fig. 7: Receiver operating characteristic (ROC) for the investigated bearing classes determined for four different classifiers.

The averaged results after the 10-folded cross-validation for the ROC curves are shown in Fig. 7. The x-axis shows the false positive rate and the y-axis the true positive rate of the classification. For each individual classification algorithm, the mean value of all four ROC curves for the respective classes is considered. The linear SVM achieves the worst ROC curve in a visual comparison. Followed by the fine decision tree and the k-nearest neighbor algorithm. The

best result is achieved by the feed-forward neural network.
In order to quantify the performance of the ROC curves and thus of the classifiers more precisely, the average AUCs are calculated and listed in Tab. 2.

Tab. 2: Area under curve for the different classification algorithms

Classification algorithm	AUC
Neural network	0.994
Support vector machine	0.968
Decision tree	0.978
K-Nearest neighbors	0.989

Overall, the neural network achieves the best result with 0.994, followed by KNN with 0.989, Decision tree with 0.978 and SVM with 0.968. However due to the high AUC values, all classifiers can therefore be rated as suitable for the detection of bearing damage.

V. Conclusion

This paper presents a study of reproducible damage to inner races, outer races and rolling elements of tapered roller bearings. In the first part of the study, damage caused by erosion in the form of circular surfaces with a diameter of 1.5 mm to 3 mm has been introduced into the various components of the bearings. By varying the speed, several series of measurements have been recorded for each bearing using piezoelectric transducers. A total of 71 features have been calculated for the measurement data and different classification algorithms have been compared using ROC curves and AUC. The best classification result has been achieved by the feed-forward neural network.
In future investigations, the influence of the damage size on the classification accuracy will be investigated. At the same time, the artificial damage will be compared with real damage of bearings from agricultural machines in order to derive a more robust classification. In addition, more damaged bearings will be investigated to build the test and training data set from different bearings. Beyond that wear phenomena, such as pitting, will be investigated in order to predictively determine the remaining useful lifetime of the bearing.

References

[1] G. Brian and K. Starry, “Rolling Element Bearing Analysis,” *Materials Evaluation, Vol. 70, No. 1 pp:78-85,* 2011.

[2] A. Beering, R. Klemm and K.-L. Krieger, “Analyse und Lokalisation von Zahnbrüchen in mehrstufigen Getrieben mit Hilfe von vibroakustischen Signalen,” in *DAGA 2019 - 45. Jahrestagung für Akustik*, Rostock, 2019.

[3] W. Smith and R. Randall, “Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study,” *Mechanical Systems and Signal Processing 64-65 pp: 100–131,* 2015, doi: 10.1016/j.ymssp.2015.04.021.

[4] C. Luo, M. Jia and Y. Wen, “The diagnosis approach for rolling bearing fault based on Kurtosis criterion EMD and Hilbert envelope spectrum,” in *2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference* , Chongqing, 2017, doi: 10.1109/ITOEC.2017.8122439.

[5] H. Van Khang, J. Sri Lal Senanayaka and K. G. Robbersmyr, “Towards Online Bearing fault detection using Envelope Analysis of vibration signal and Decision Tree Classification Algorithm,” in *2017 20th International Conference on Electrical Machines and Systems* , Sydney, 2017, doi: 10.1109/ICEMS.2017.8056146.

[6] Verein Deutscher Ingenieure, Körperschallmessungen zur Zustandsbeurteilung von Wälzlagern in Maschinen und Anlagen, Beuth Verlag GmbH, 2013.

[7] C. M. Bishop, Pattern Recognition and Machine Learning, Cambridge: Springer Science & Business Media, 2009.

[8] D. Baughman and Y. Liu, “Fundamental and Practical Aspects of Neural Computing,” *Neural Networks in Bioprocessing and Chemical Engineering,* pp. 3-31, January 1995, doi: 10.1016/B978-0-12-083030-5.50008-4.

[9] I. H. Witten, E. Frank and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Elsevier, 2011, doi: 10.1145/2020976.2021004.

[10] J. Davis and M. Goadrich, “The Relationship Between Precision-Recall and ROC Curves,” in *International Conference on Machine Learning*, Pittsburgh, 2006, doi: 10.1145/1143844.1143874.

[11] T. Fawcett, “ROC Graphs: Notes and Practical Considerations for Researchers,” *Pattern Recognition Letters,* pp. 1-38, 2004.

[12] SKF Group, “Railway technical handbook,” Göteborg, 2012, pp. 122-135.

[13] W. Kuttner, Praxiswissen Schwingungsmesstechnik, Wiesbaden: Springer Fachmedien, 2015, doi: 10.1007/978-3-658-04638-5.