DINAME 2023 - Proceedings of the XIX International Symposium on Dynamic Problems of Mechanics,
A.S. De Paula, A.T. Fabro and M.R. Machado (Editors), ABCM.
Pirenópolis, Go, Brazil, February 26th to March 3rd, 2023

Data-driven Gear Faults Diagnostics using Bayesian Neural Networks

Matheus de Moraes[1], **João Paulo Dias**[2], **and Hélio Fiori de Castro**[1]

[1] School of Mechanical Engineering, State University of Campinas, 200 Mendeleyev Street, 13083-860, Campinas, São Paulo, Brazil

[2] Dept. of Civil and Mechanical Engineering, Shippensburg University of Pennsylvania, 1871 Old Main Drive, 17257, Shippensburg, Pennsylvania, United States

Abstract: Condition monitoring of gearboxes is a powerful technique in machinery maintenance. Artificial intelligence is nowadays supporting fault detection of several kinds of components. The objective of this research is to propose a data-driven methodology of gear faults diagnostics that accomplishes uncertainty quantification. Gear fault vibration signals obtained from an open database were converted into images that were used to train a Bayesian neural network. Test results for the Bayesian neural network fault detection capability showed high accuracy and allowed the uncertainties in the fault diagnostics to be properly quantified.

Keywords: gearboxes, condition monitoring, Bayesian neural networks

INTRODUCTION

Condition monitoring of industrial equipment is a powerful method to avoid unexpected failures. With a wide range of industrial applications, transmissions and gearboxes are considered to be critical components due to uncertain environmental conditions and unpredictable overloads. Therefore, condition monitoring is essential to mitigate catastrophic failures of transmission. Furthermore, in the context of the fourth Industrial Revolution, condition monitoring has been supported by artificial intelligence algorithms.

Application of artificial intelligence algorithms for condition monitoring of gearboxes has been extensively studied by researchers in the past decade. Wang *et al* (2019), presented a review of the state-of-art of condition monitoring and fault detection techniques of planetary gearboxes of wind turbines. Jing *et al* (2017) proposed a convolutional neural network (CNN) model to learn features from frequency data directly and detect faults of gearboxes. The results showed that CNN is more powerful than manual feature extraction in deep learning methods. Cao, Zhang, and Tang (2018) proposed a method based on transfer learning for fault detection in a two-stage gearbox and achieved high performance in this task. The work of de Moraes, Dias, and Castro (2021) performed damage detection in ball bearings using vibration images and Bayesian neural networks. The results showed high levels of accuracy in the diagnostic of several damage conditions, furthermore performed uncertainties quantification of the diagnostic.

Although relevant contributions have been made so far, it is possible to identify a gap in the literature regarding the application of Bayesian neural networks (BNN) in gearbox diagnostics. Hence, the general goal of this work is to answer the following research question: are BNNs a viable technique in gear faults diagnostic? To answer this question following specific aims were proposed: (i) to present some theoretical aspects of the BNN; (ii) to construct a data set of vibration images; (iii) to train, to test and to evaluate a BNN in a diagnostic task of gear faults and quantify the uncertainty in the fault detection.

The rest of the paper is organized as follows: in the Methodology section proposes the theoretical and practical aspects of BNN and the construction of vibration mages dataset is proposed; in the Results section the main results are presented and discussed; in the Conclusion section the main conclusions of the work are highlighted.

METHODOLOGY

Bayesian neural networks

Artificial neural networks (ANN) are a type of supervised machine learning algorithm. In other words, during the training, the parameters of the network θ, also called weights and biases, are fitted to input data X_{tr} to the known outputs Y_{tr} that belong to the data set D. The goal of creating neural networks is to construct a non-linear function with parameters θ that maps the output Y given the inputs X, $f^{\theta}: X \rightarrow Y$. For a unique layer of a neural network:

$$f(X) = \sum_{m=1}^{M} g_m(\theta_m^T X) \tag{1}$$

where, g_m is the activation function, θ_m is the vector of parameters (weights and biases).

In the Bayesian approach for neural networks learning, the goal is to obtain a distribution able to predict the outputs of future experiments given the inputs X and the training data X_{tr} and Y_{tr}. Key advantages of Bayesian neural networks are the reduction of overfitting and the inclusion of epistemic and data uncertainties in the analysis.

Variational inference

The central purpose of parametric Bayesian methods is obtaining the posterior distribution of the non-observable parameters θ that is probability density function (pdf), which represents the uncertainties over θ, with the observable data set D. From the Bayes theorem, the posterior is:

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{\int p(D|\theta')p(\theta')d\theta'}, \tag{2}$$

where, $p(D|\theta)$ are the likelihood function, $p(\theta)$ is the prior distribuitions of the parameters θ, and $\int p(D|\theta')p(\theta')d\theta'$ is the evidence, which is a likelihood function in which some parameters have been marginalized.

In general, the integration of the evidence of eq. (2) is untreatable. Therefore, variational inference is employed to approximate the posterior analytically. A family of treatable distributions $\mathscr{Q}$ is chosen with the objective of to search, between the probability density functions $q(\theta)$ that belongs to the family, a probability density function $q^*(\theta)$ in the way of:

$$q^*(\theta) = \text{argmin} D_{KL}(q(\theta)||p(D|\theta)), \tag{3}$$

where, $D_{KL}(q(\theta)||p(D|\theta)$ is the Kullback-Leibler divergence between $q(\theta)$ and $p(D|\theta)$ that is given by:

$$D_{KL}(q(\theta)||p(D|\theta)) = \int \log \frac{q(\theta)}{p(D|\theta)} q(\theta)d\theta. \tag{4}$$

Thus, the goal is to obtain the density $q(\theta) \in \mathscr{Q}$ that minimizes the Kullback-Leibler divergence related to the posterior $p(D|\theta)$. Rewriting eq. (4):

$$D_{KL}(q(\theta)||p(D|\theta)) = \mathbb{E}_q[\log q(\theta)] - \mathbb{E}_q[\log p(D|\theta)], \tag{5}$$

note that the expectations are calculated in relation of the density $q(\theta)$. It is possible to expand the eq. (5):

$$D_{KL}(q(\theta)||p(D|\theta)) = \mathbb{E}_q[\log q(\theta)] - \mathbb{E}_q[\log p(D|\theta)] + \log p(D), \tag{6}$$

when $p(D)$ is unavailable, D_{KL} is untreatable once it depends on the log-evidence $\log p(D|\theta)$. On the other hand, the optimization is given by $q(\theta)$.

We can define the Evidence Lower Bound function (ELBO) $\mathscr{L}$ is given by:

$$\mathscr{L}(q(\theta)) = \mathbb{E}_q[\log p(D|\theta)] - \mathbb{E}_q[\log q(\theta)], \tag{7}$$

we can demonstrate that when the D_{KL} is minimized, the evidence lower bound objective (ELBO) function is maximized with respect of $q(\theta)$. Finally:

$$\log p(D) = D_{KL}(q(\theta)||p(\theta|D) + \mathscr{L}(q(\theta))), \tag{8}$$

thus, the ELBO function is the lower boundary of the evidence and it is treatable (Blundell *et al.*, 2015; Vega and Todd, 2020).

Experimental data set, vibration images, and BNN achitecture

The experimental data set was available in the public Gear Fault Data set (Cao *et al.*, 2018). This data set consists of vibration signals of a two-stage gearbox with replaceable gears. The work of Cao *et al* (2018), describes the detailed experimental procedure. The classes are: gear operating on normal or healthy condition; gear with missing thoot; gear with a thoot with root crack; gear operating with spalling; and, five levels of chip, in which 5 is the less severe and 1 the more severe condition.

For the conversion of vibration signals into vibration images, the following steps are necessary: first, a computational routine normalizes the acceleration's time series into a range of [-1, 1]; after that converts this time series into a 20 by 20 matrix. The works of Hoang *et al.* (2018) and de Moraes, Dias, and Castro (2021) describe the details of this methodology.

The model used in the classification of several classes of gear faults is the same one used in previous work (de Moraes, Dias, de Castro, 2021). Then, the model consists of a sequential model called a convolutional Bayesian neural network with seven layers: convolutional 2D reparametrization; maxpooling 2D; convolutional 2D; maxpooling 2D; flatten layer; dense reparameterization layer. The illustration of Figure 1 shows an example of BNN used in this work.

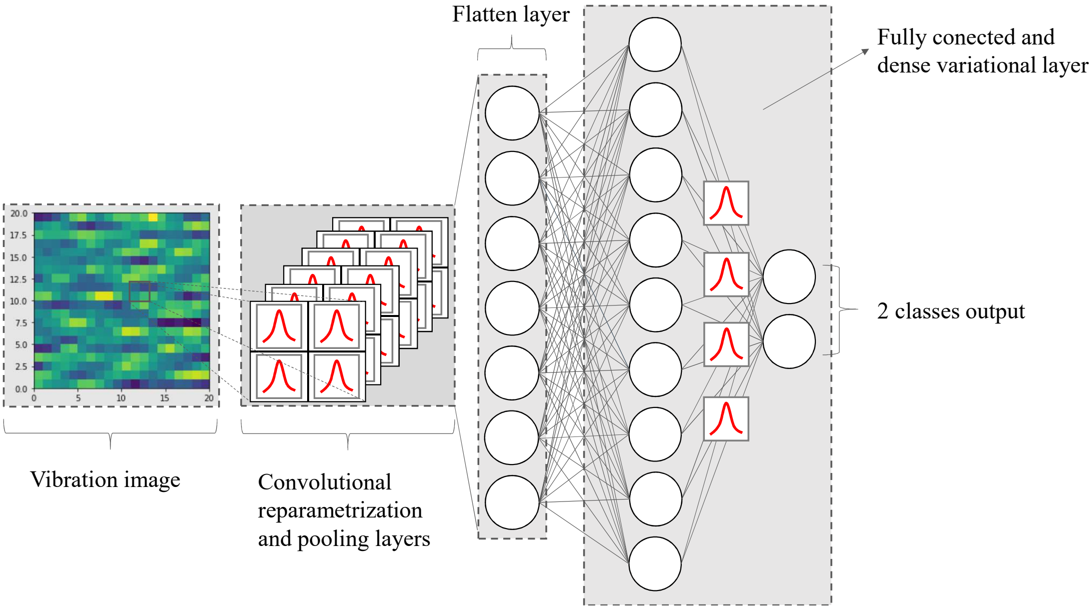


Figure 1: Illustration of an example of the BNN studied.

Figure 2 show the flowchart of the methodology employed in this work.

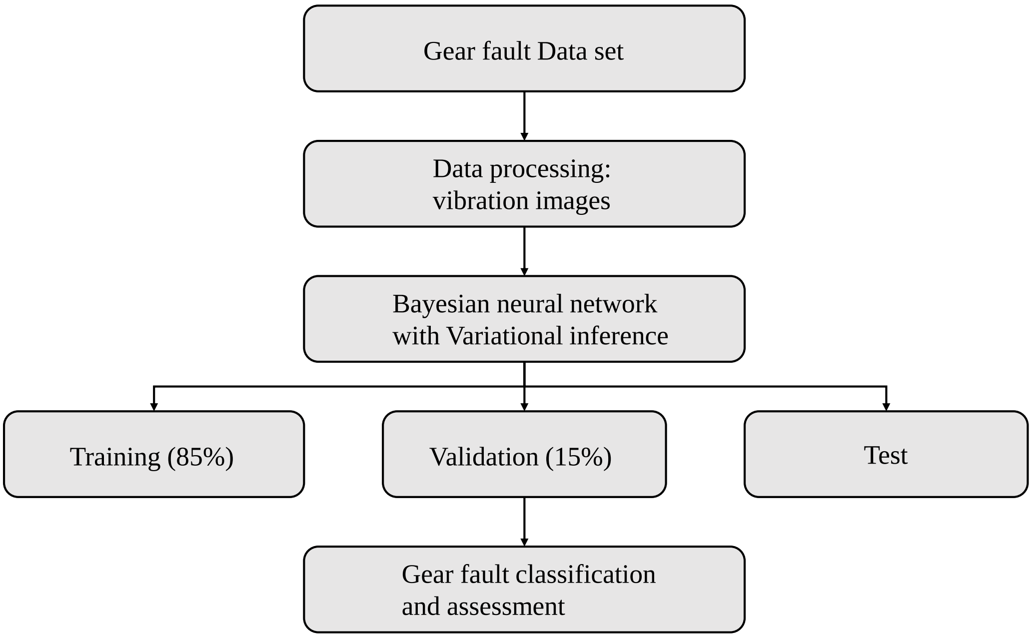
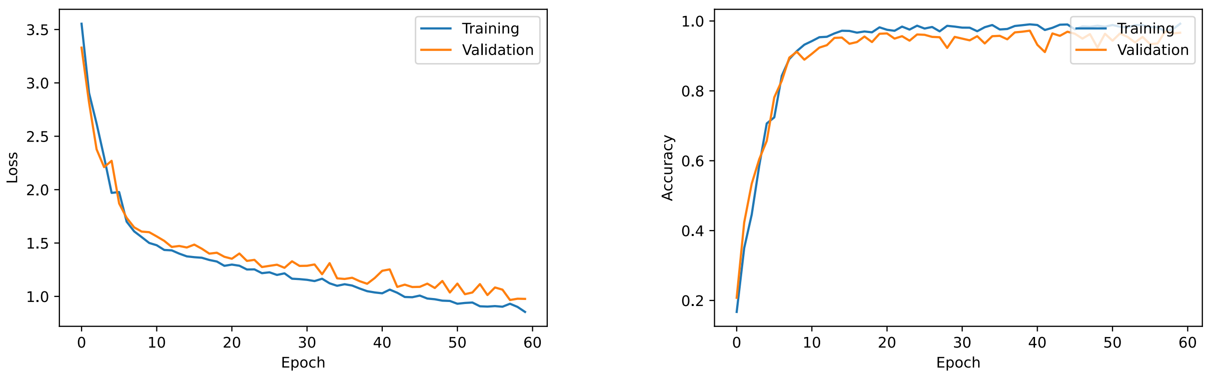


Figure 2: Flowchart of the methodology employed in this work.

RESULTS AND DISCUSSION

This section provides the previous and highlighted results of the research. For each operating condition, 936 vibration images were generated, totaling 8424 vibration images. We use 6740 (80%) vibration images for training operation and 1684 (20%) vibration images for test.

First of all, Figure 3 shows the graphics of values of loss function *versus* epochs and the values of accuracy in function of epochs. It is possible to notice that in the values loss function (Fig. 4a) even in the validation data set is not observable the overfitting. Also, it is possible to notice that the accuracy (Fig. 4b) shows stability in high levels.



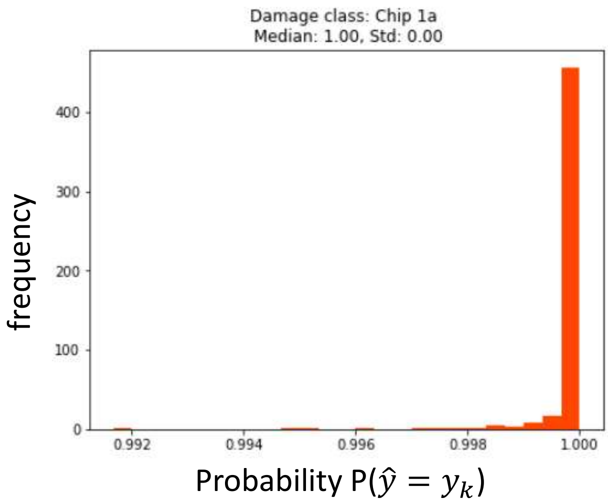
(a) Loss.

(b) Accuracy.

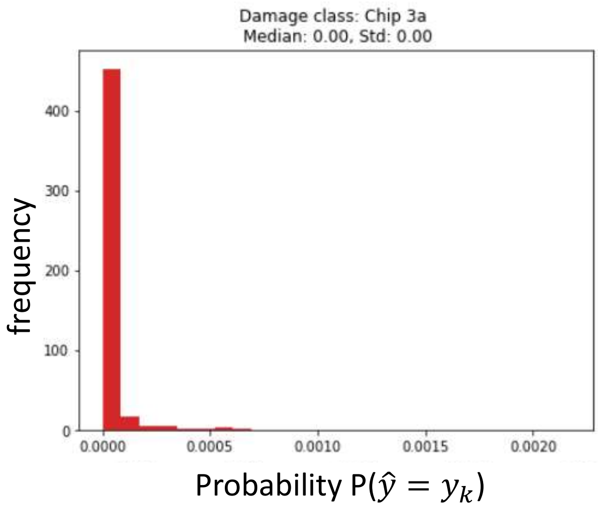
Figure 3: Performance metrics.

Figure 4 shows the histogram of a classification performed by the BNN. The histogram was constructed with 500 trials of classification with a vibration image that belongs to the crack class. In Fig. 4-a the histogram is shown for a classification of a chip 3a damage class and it is possible to notice that the classification achieves high accuracy in the same time that is possible to observe low level of uncertainty. Fig. 4-a show, for the same sample tested, that exist low probability of the sample to belong to class chip 1a, with low levels of uncertainty.

The confusion matrix of mean values for 500 trials is presented in Figure 5. As demonstrated in Figure 5, it has high performance (over than 90%) in classification of several fault classes, the class named chip 3a presented the worst value in which confusion occurs between the class named chip 4a, *i. e.*, the same type of fault with different levels of severity.



(a) Damage class: chip 3a.



(b) Damage class: chip 1a.

Figure 4: Histograms.

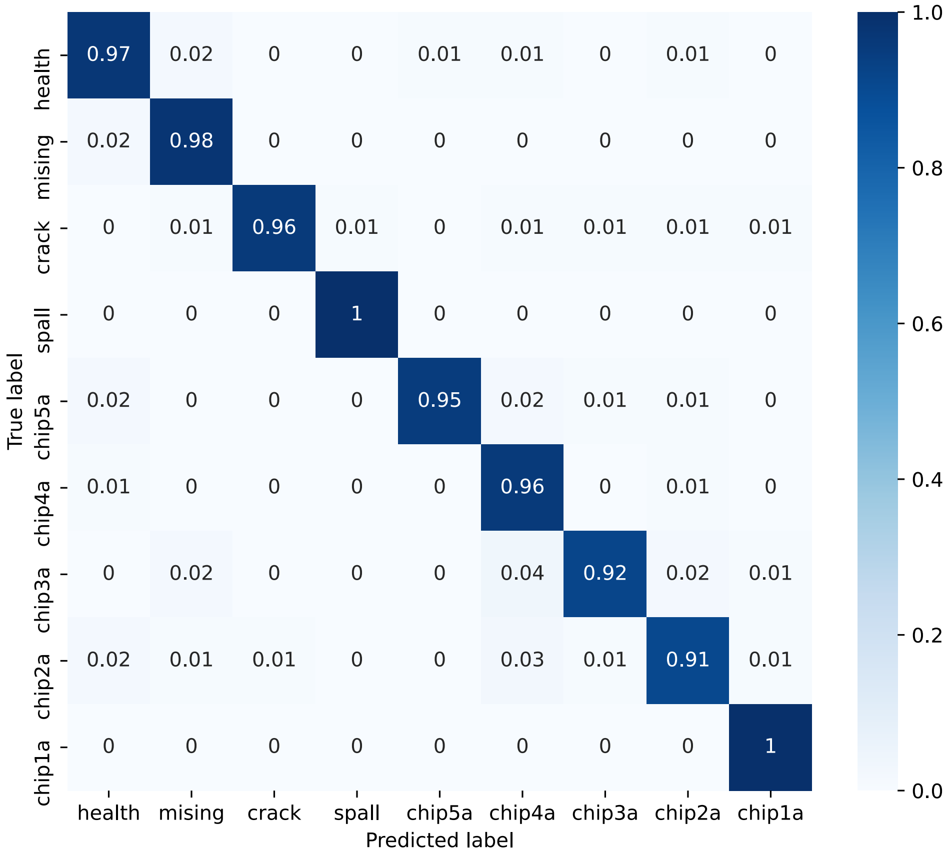


Figure 5: Confusion matrix of the mean values.

CONCLUSIONS

This paper's main conclusions are:

- it was possible to generate vibration images from acceleration signal of a public data set of a two-stage gearbox operating under different type of conditions, including faults with different levels of severity;
- the construction, training, testing, and evaluation of a Bayesian neural network to perform classification of the gear faults, including uncertainty quantification;
- the diagnostic achieves a performance of 96% over the testing data.

All of these conclusions allows that the methodology proposed in this work presented high performance in diagnostic of gear faults, therefore this methodology indicates viability for condition monitoring routines.

ACKNOWLEDGMENTS

The authors acknowledge the financial support given by Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES) and the Sao Paulo Research Foundation (FAPESP) [grant number #2019/00974-1].

REFERENCES

Blundell, C., Cornebise, J., Kavukcuoglu, K. and Wierstra, D., 2015. "Weight uncertainty in neural network". In International Conference on Machine Learning. PMLR, pp. 1613–1622.

de Moraes, M., Dias, J.P. and de Castro, H.F., 2021, "Condition monitoring of ball bearings using bayesian neural networks." In Proceedings of the XLII Ibero-Latin-American Congress on Computational Methods in Engineering and III Pan-American Congress on Computational Mechanics, ABMEC-IACM. Rio de Janeiro, pp. 1–7.

Cao, Pei, Zhang, Shengli , and Tang, Jiong, 2018 "Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning." Ieee Access, Vol. 6, pp. 26241-26253.

Hoang, Duy-Tang, and Hee-Jun Kang, 2019, "Rolling element bearing fault diagnosis using convolutional neural network and vibration image." Cognitive Systems Research, Vol. 53, pp. 42-50.

Jing, Luyang, et al., 2017, "A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox.", Measurement, Vol. 111, pp. 1-10.

Vega, M.A. and Todd, M.D., 2020. "A variational bayesian neural network for structural health monitoring and costinformed decision-making in miter gates". Structural Health Monitoring, Vol. 00, pp. 1–15.

Wang, Tianyang, *et al.*, 2019, "Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review." Mechanical Systems and Signal Processing, Vol. 126, pp. 662-685.

RESPONSIBILITY NOTICE