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

# Fault Identification in Rotating Systems using Convolutional Neural Networks

**Carlos A. A. Viana [1], Diogo S. Alves [2], Tiago H. Machado [3]**

[1] Laboratory of Rotating Machinery, School of Mechanical Engineering, UNICAMP, c228839@dac.unicamp.br

[2] Department of Mechanical Engineering, University of Bath, dsa41@bath.ac.uk

[3] Laboratory of Rotating Machinery, School of Mechanical Engineering, UNICAMP, tiagomh@fem.unicamp.br

*Abstract: The application of Machine Learning methods and models in vibration analysis of rotating machines has become an important milestone in the era known as Industry 4.0. Several models are used to identify faults and monitor assets. In this paper, a CNN (Convolutional Neural Network) is tested for classification of ten machine faulty condition classes. The experimental acceleration data – in the time and frequency domains – are converted to images through the vibration images technique. The results pointed to the conclusion that the vibration images in the frequency domain present better contrasts between the types of faults due to its inherent feature extraction characteristics, resulting in an accuracy of 99.4% compared to 97.1% in the domain of time.*

***Keywords: Mechanical vibration, Deep learning, Fault classification, Machine condition monitoring, Predictive maintenance***

## INTRODUCTION

Machine condition monitoring plays a fundamental role in *rotordynamics* and in the daily industrial and production processes. The increasing competitiveness between industries creates a demand for accurate and fast diagnosis of problems in machinery. In this way, the identification of different types of faults, without dismantling and inspecting equipment, and consequently avoiding unwanted stops that impair availability of industrial plants, has become mandatory.

For this, new approaches in the area emerged with the development of Machine Learning and pattern identification and classification techniques. Through vibration signals of the machine, it is possible to classify its health status and faults, such as misalignments, unbalances, component defects, etc., both in the time and frequency domains, using analysis tools such as FFT (Fast Fourier Transform).

At the same time, some researches carried out investigations and developed models based on artificial neural networks that are capable of classifying the signals coming from "in loco" measurement of machinery, such as Souza et al. (2021), who built the PdM-CNN model through training with the dataset of the MaFaulDa (RIBEIRO, 2018) and CWRU (CWRU, 2022) databases, and Brito et al. (2022), with an unsupervised approach for fault classification based on feature importance ranking.

The use of convolutional neural networks in the context of machine health monitoring has been explored in Hoang and Kang (2018), who propose the diagnosis of bearing faults using the raw vibration signal, without extracting parameters from the data, using the vibration images technique. For this, the signal is arranged in sequences of time windows, where the amplitude of vibration of each time step is sequentially inserted in within a matrix that is plotted in a pixel intensity format. All generated images compose the training and test database of the network, which can achieve high accuracy in classifying the intensities of faults.

Jha and Swami (2021) developed a model for classifying the severity of faults in rolling element bearings, using multiclass support vector machines (MSVMs). The one-dimensional vibration signal was converted into two-dimensional grayscale, resulting in a textural pattern mapping, being enhanced using the wave atom transform. The classifier was trained in two phases: first categorizing the fault location and then diagnosing the fault size at that particular location.

Still in relation to vibration monitoring of rotating machines, Fu et al. (2017) explored the possibility of using graphical representations of signals as input to build a fault identification model using convolutional neural networks, also by converting the vibration signal into two dimensions, for exhibit and fuse information. This monitoring has been tested in order to be applied in real time.

Recurrent neural networks were also used to estimate vibration frequencies, as in Liu et al. (2019), which brings an approach with lower computational cost, proposing a long short-term memory-recurrent neural networks (LSTM-RNNs) and muti-target learning, being used for direct prediction of the components of vibration frequencies. The methodology was applied in experiments of the nature of forced and free vibration in structures.

As presented, the combined use of vibration images technique and convolutional neural networks has a great potential to identify abnormal behavior of rotating systems. However, past studies only considered the analysis of vibration in the

time domain, creating the opportunity of a brand-new study about the performance of such methods when pre-processing feature extraction algorithms – as the FFT (Fast Fourier Transform) – are employed. In this context, the objective of the present work is to build a CNN - convolutional artificial neural network - for training vibration measurements from the MaFauldDa database, applying the vibration images technique, and to investigate different types of data input, namely - vibration signals in the time and frequency domains.

## MATERIALS AND METHODS

This section will present succinct details of Machine Learning theory, as well as the sub-area known as Deep Learning and Convolutional Neural Networks (CNN), used in the present paper.

### *General Machine Learning Theory*

The concept of Deep Learning gained strength in the last decade, within the Machine Learning scenario, mainly due to the constant increase in the amount of all types of available data and the significant development of hardware capable of performing training of massive structures. A prominent aspect of deep learning is that the models typically deal with the raw data of the problem, i.e., without any pre-processing or feature extraction. In short, the idea is to "let the data speak for itself". Therefore, the model must learn different levels of representation that allow the problem to be solved. In fact, many deep learning models build a hierarchy of representations, which permits the machine to learn more complicated aspects present in the data by building it from more elementary characteristics. At the same time, the model also learns how to process these representations to perform the desired task (for example, classifying an image).

CNN neural networks are widely used in the field of computer graphics, being models composed of attributes of adjustable weights and subsampling operations (pooling), treating two-dimensional (2D) or three-dimensional (3D) data sequences, being applied in various types of problems, such as classification of audios, texts, and time series. Convolutional networks employ the convolution operation in place of the affine transformation inherent in a perceptron-like layer (Goodfellow et al., 2016). This operation, which is linear, allows exploring information in structures organized in time, such as time series, or in space, such as images. The operation typically presented in clustering (or pooling) CNNs, which performs a sub-sampling of the input, aims to summarize the information. The convolution function can be defined as shown in Eq. (1).

$$s(t) = \int_{a=-\infty}^{\infty} x(a)w(t-a)da \tag{1}$$

*x* being the input value, *w* the weight, *t* the time and *s(t)* the function known as feature map.

The architecture of a convolutional network generally consists of a first part of linear activation, followed by a second part in which each linear activation is performed using nonlinear functions, such as ReLU (Rectified Linear Unit). A third part consists of a pooling function, used to modify the output summarizing the data and, finally, the fourth part with the softmax activation function, in the multi-class scenario, which performs the classification of the data set in the fully connected layer.

The ReLU activation function performs the input processing by canceling all negative values, being linear for positive values, as shown in Eq. (2), where $a^{l(i,j)}$ is the activation function.

$$a^{l(i,j)} = f\left(z^{l(i,j)}\right) = max\left\{0, z^{l(i,j)}\right\} \tag{2}$$

### *Vibration Images*

This method aims to transform the vibration signals into grayscale images, called vibration images. Fig. (1) shows the process of building vibration images. In this calculation, the amplitude of each sample in the vibration signal is normalized from -1 to +1. After a normalization of the amplitudes of each sample, the new value becomes the intensity of the corresponding pixel in the corresponding matrix image.

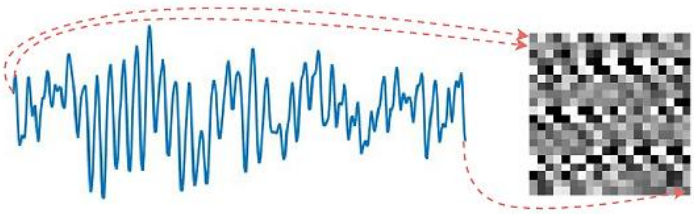

**Figure 1 – Vibration image construction process.**

The conversion between the normalized sample amplitude and corresponding pixel is described by Eq. (3) (Nguyen et al., 2013).

$$P[i,j] = A[(i-1).M + j] \tag{3}$$

where $i$ varies from 1 to $N$, and $j$ varies from 1 to $M$, being $N$ is the number of lines in the vibration image and $M$ is the number of columns.

In this way, the entire vibration signal is mapped to vibration images, respecting the stacking of vibration amplitudes in the matrix components.

### *MaFaulDa Database*

This database is composed of 1951 multivariate time-series, that comprises ten different simulated states: normal function, imbalance fault, horizontal and vertical misalignment faults and, inner, outer and ball bearing faults for underhang bearing and overhang bearing, as shown in Table (1).

| Sequence | Measurements |
|---|---|
| Normal | 49 |
| Horizontal misalignment | 197 |
| Vertical misalignment | 301 |
| Imbalance | 333 |
| **Underhang bearing** | |
| Cage fault | 188 |
| Outer race | 184 |
| Ball fault | 186 |
| **Overhang bearing** | |
| Cage fault | 188 |
| Outer race | 188 |
| Ball fault | 137 |
| **Total** | **1951** |

**Table 1 – Multivariate time-series of MaFaulDa Database. (Machinery Fault Database, 2022).**

Figure (2) shows the test rig used to extract all data from MaFaulDa.

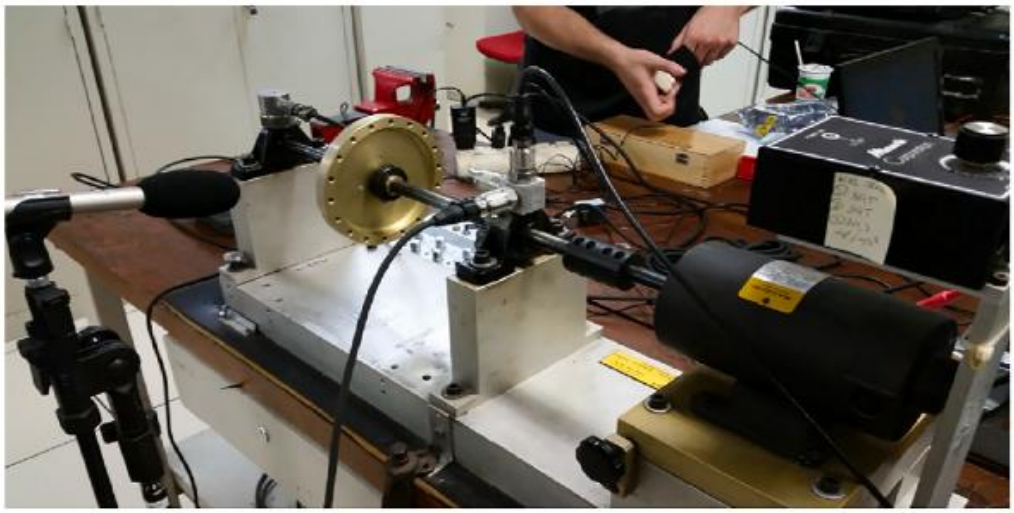

**Figure 2 – Test bench configuration for MaFaulDa database. (Ribeiro, 2018).**

To better represent the test rig components, Figure (3) shows the positioning of the sensors, couplings, motor, disk and shaft of the rotating system.

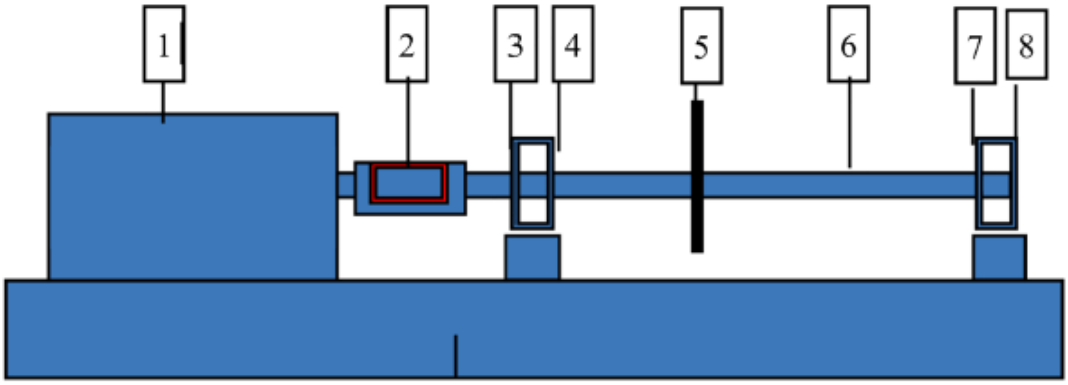


**Figure 3 – Schematic drawing showing each sensor position: 1. Servo motor; 2. Coupling; 3, 7. Bearing house; 4, 8. Accelerometers; 5. Disk; 6. Shaft. (Ribeiro, 2018).**

Data acquisition was performed using an IMI Sensors triaxial accelerometer, Model 604B31, sensitivity (±20%) 100 mV per g (10.2 mV per m/s2), frequency range (±3 dB) 30-300000 CPM (0.5- 5,000 Hz) and measuring range ±50 g (±490 m/s2), returning data over the radial, axial and tangential directions. For this work, only data from tangential direction were used, as they better outline the system's behavior between the classes studied.

### *CNN Model*

The convolution neural network architecture used for the model consists of a two-dimensional convolutional layer followed by max pooling layers, which performs a synthesis of the data coming from the previous layer, passing the information to a flatten that then passes to three-hidden dense layers and for a fully connected softmax layer to perform the classification.

Figure (4) shows the representation of the architecture used in the CNN model.

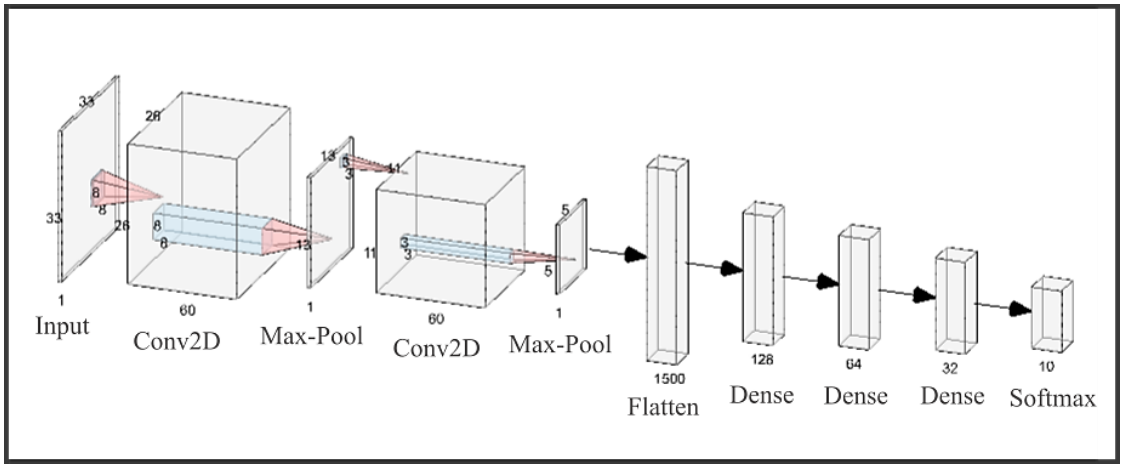


**Figure 4 –Architecture of the CNN model used.**

## RESULTS AND DISCUSSION

Data at rotating speeds from 12 Hz to 60 Hz were adopted from MaFaulDa database. In total 1951 time sequences, with a sampling frequency of 50kHz, were used for each class of rotor state, namely - Normal (0); Horizontal misalignment (1); Vertical misalignment (2); Imbalance (3); Underhang bearing cage fault (4); Underhang bearing outer

race fault (5); Underhang bearing ball fault (6); Overhang bearing cage fault (7); Overhang bearing outer race fault (8); Overhang bearing ball fault (9).

After normalizing the data, they were processed in two distinct ways for each class. In the first situation, the time domain data were directly converted to vibration images, while in the second way, the data were transformed to the frequency domain using the FFT and, afterwards, the amplitudes values were used to generate the images. In both cases, the converted vibration images are a 33x33 pixel map.

Only the tangential component of the signal captured by the accelerometers was used, as it provides the most evident behavior of the rotor vibration dynamics. In addition, 75% of the data were used for the training stage, while the remaining 25% for the test stage.

As parameters for training the CNN, in each of the two tests, searches were performed for the optimal parameters of quantity and size of kernels for convolutional operations. The model compilation was performed with the "adam" optimizer, and the success metric was calculated with accuracy and the cost function with the categorical cross entropy sparce. For model fit, 60 epochs were used for passing training data to calibrate CNN parameters.

Figure (5) shows samples of the time domain vibration images in each of the classes.

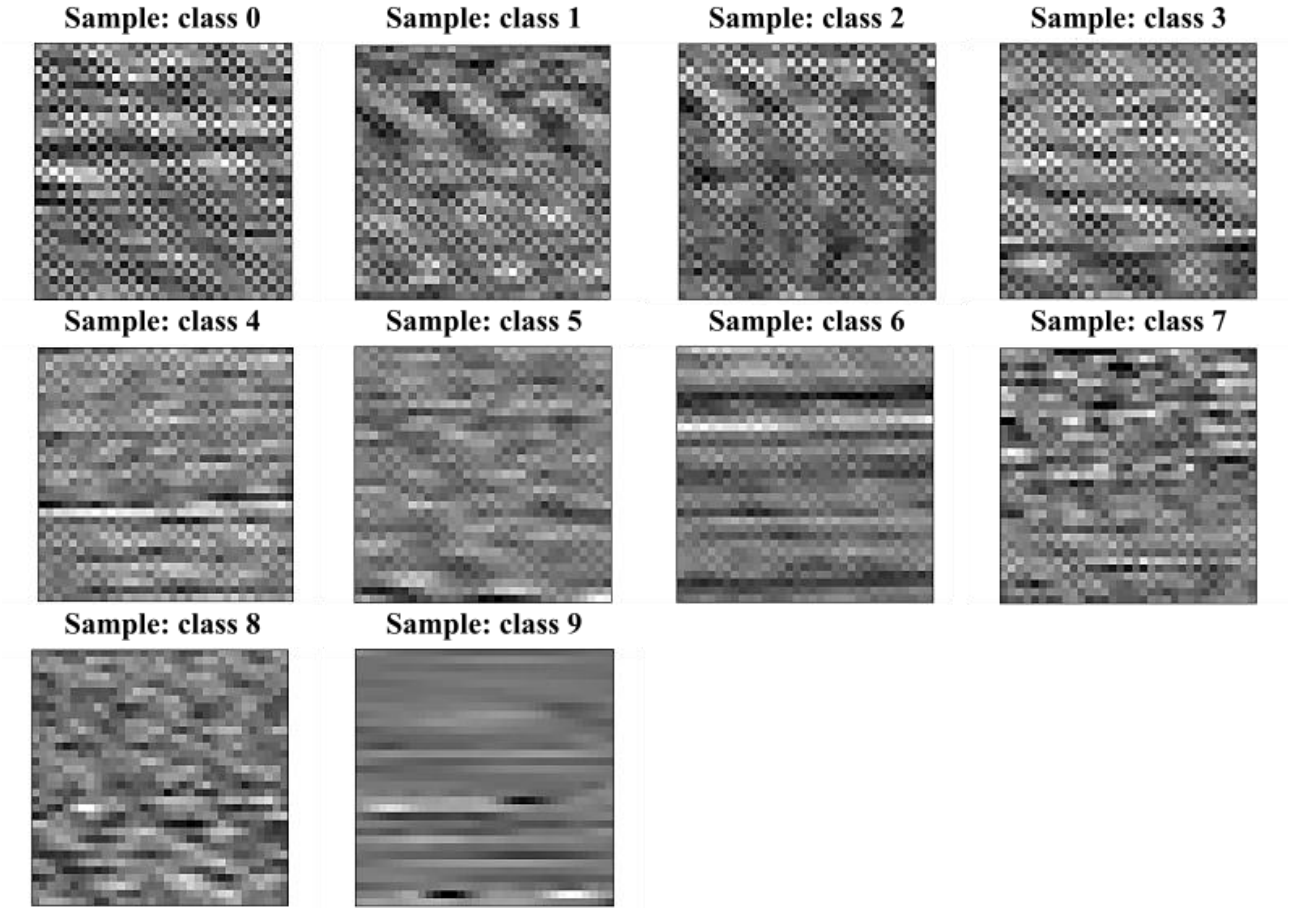


**Figure 5 – Vibration images samples in time domain.**

It is possible to visualize an abstract differentiation between the classes, due to the stacking of vibration amplitudes within the vibration images matrix.

The model's training and validation history is shown in Figure (6), with the graph of accuracy (Fig. 6-a) and loss (Fig. 6-b). Figure (7) presents the confusion matrix for the model, in the validation data.

An average accuracy of 97.1% was found for the database used, which indicates a good ability to generalize classification from model.

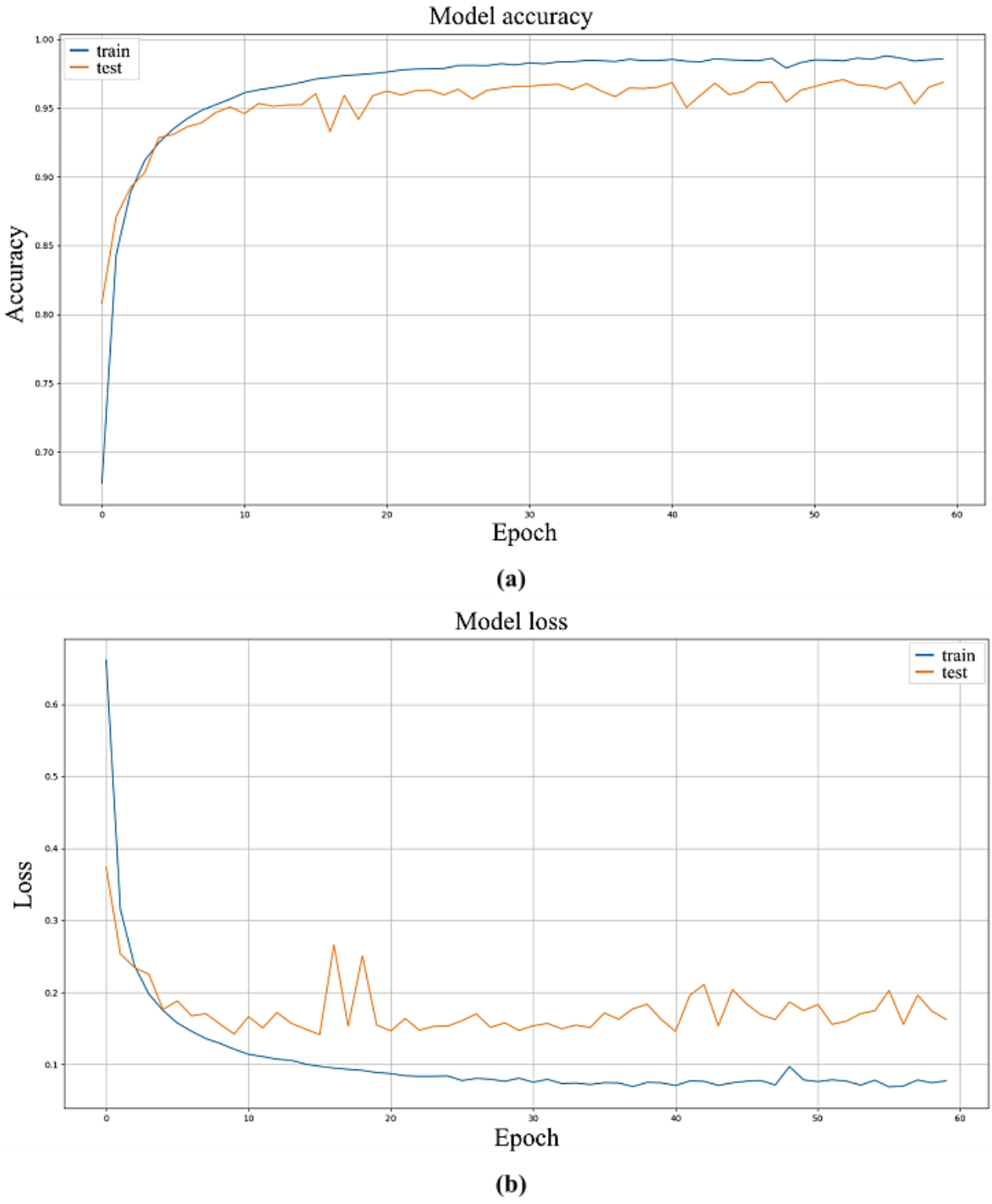


**Figure 6 – Training and validation history in time domain. (a) accuracy; (b) loss.**

From the graphs of Figure (6), it is possible to infer that the model has no overfitting for the number of epochs adopted, with epoch equal to 30 being the ideal stop for the model, since greater values present no considerable increase in accuracy, nor a decrease in the cost function.

In general, a satisfactory fault classification accuracy was obtained. Observing the confusion matrix, it is possible to verify which cases the model can easily predict (classes 0, 2, 3, 6, 7, 8 and 9) and which classes are more difficulty to identify (classes 1, 4 and 5).

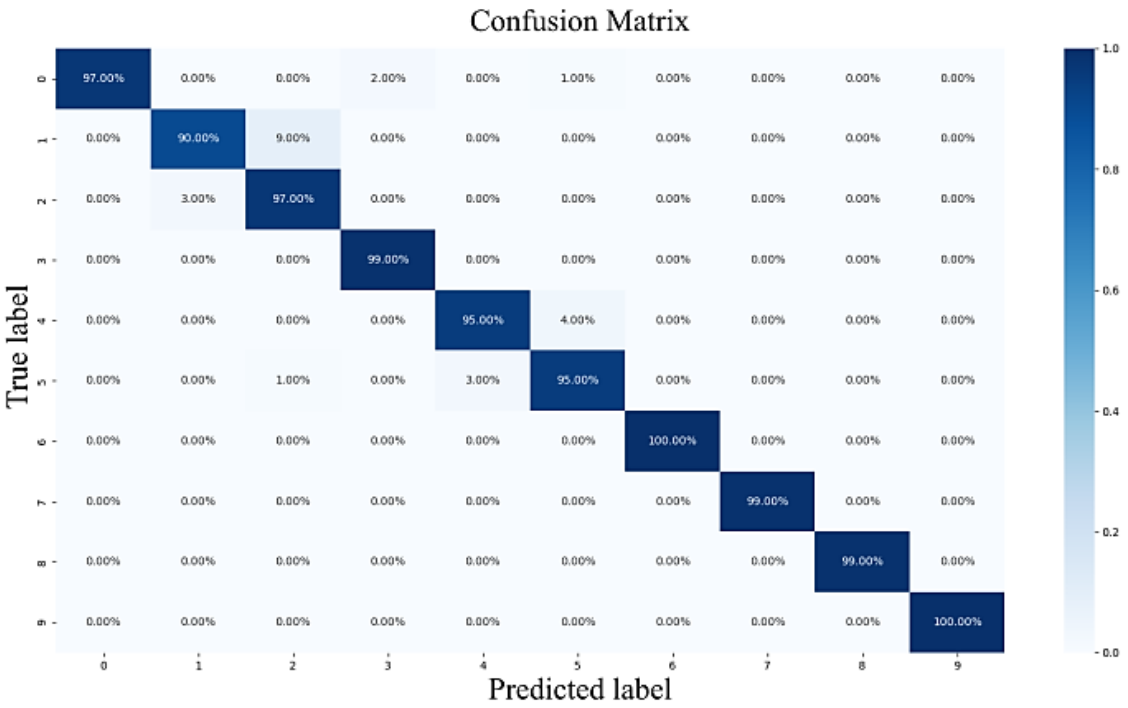


**Figure 7 – Confusion Matrix for validation in time domain.**

The second part of the tests concerns the results in the frequency domain. Figure (8) shows samples of the frequency domain vibration images in each of the classes.

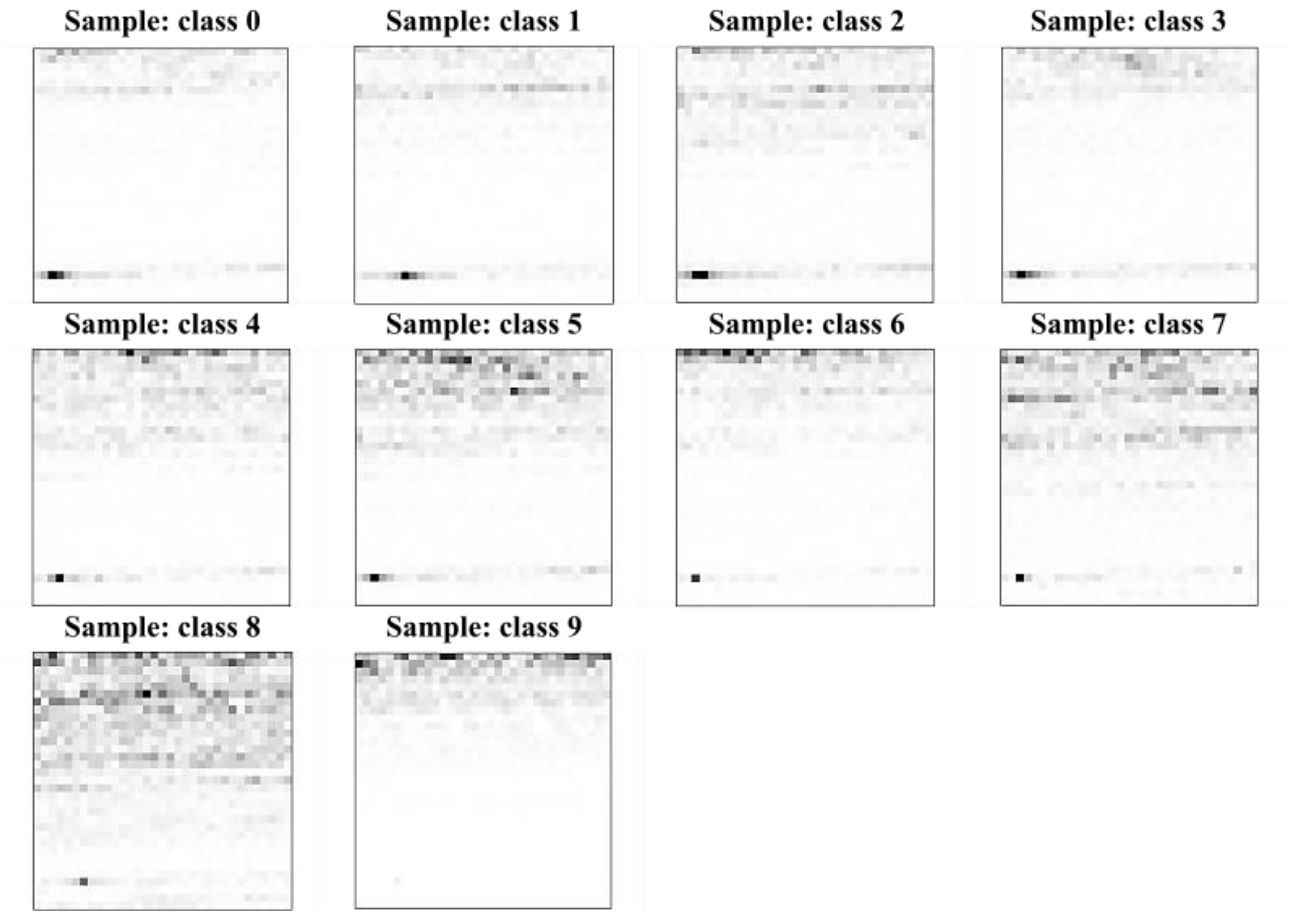


**Figure 8 – Vibration images samples in frequency domain.**

Figure (9) shows the graph of accuracy (Fig. 9-a) and loss (Fig. 9-b) for the training stage in frequency domain, and Figure (10) contains the confusion matrix for this validation.

It is observed an average accuracy of 99.4% in the frequency domain, which indicates a very accurate classification performance of the model with this approach.

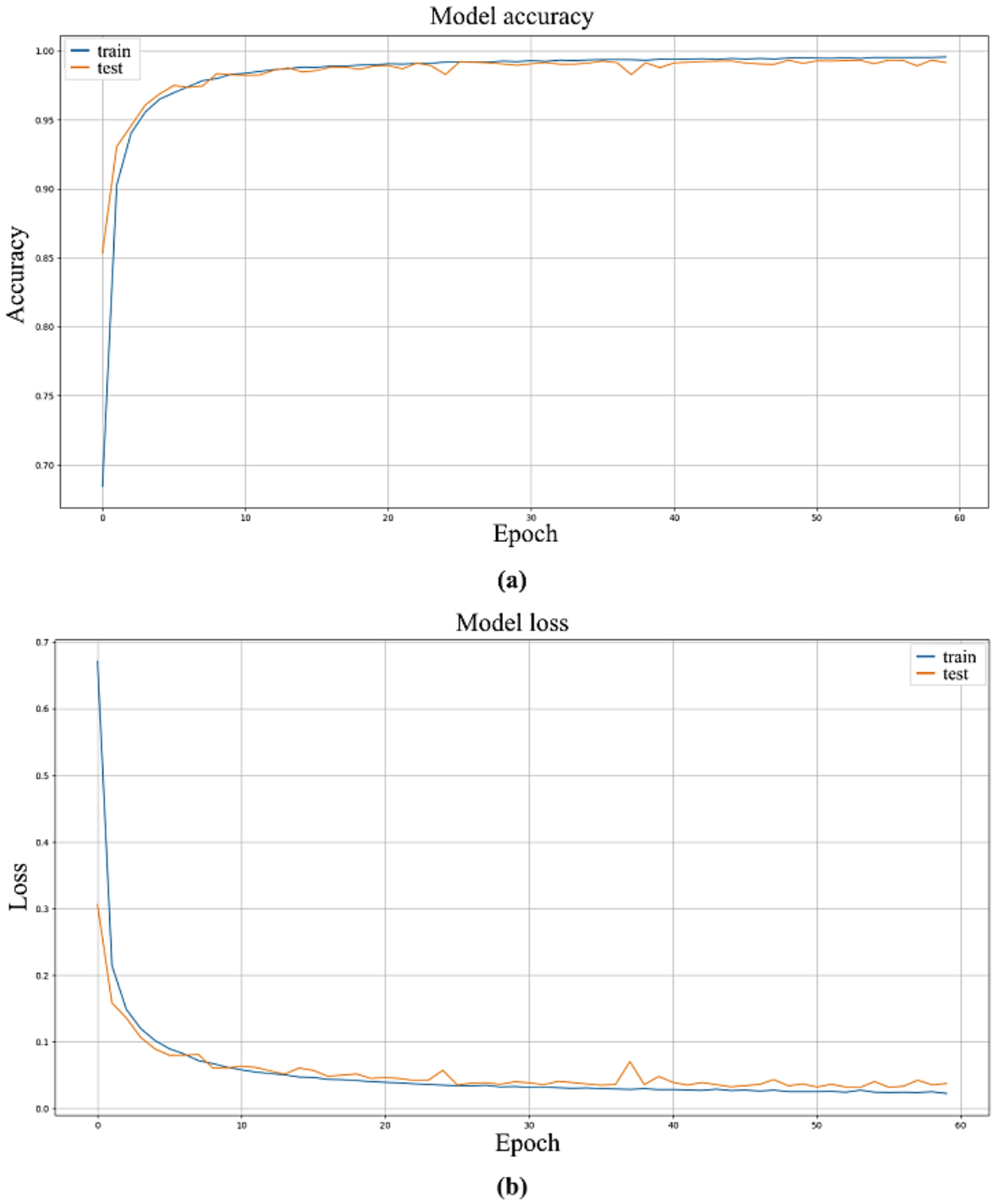


**Figure 9 – Training and validation history in frequency domain. (a) accuracy; (b) loss.**

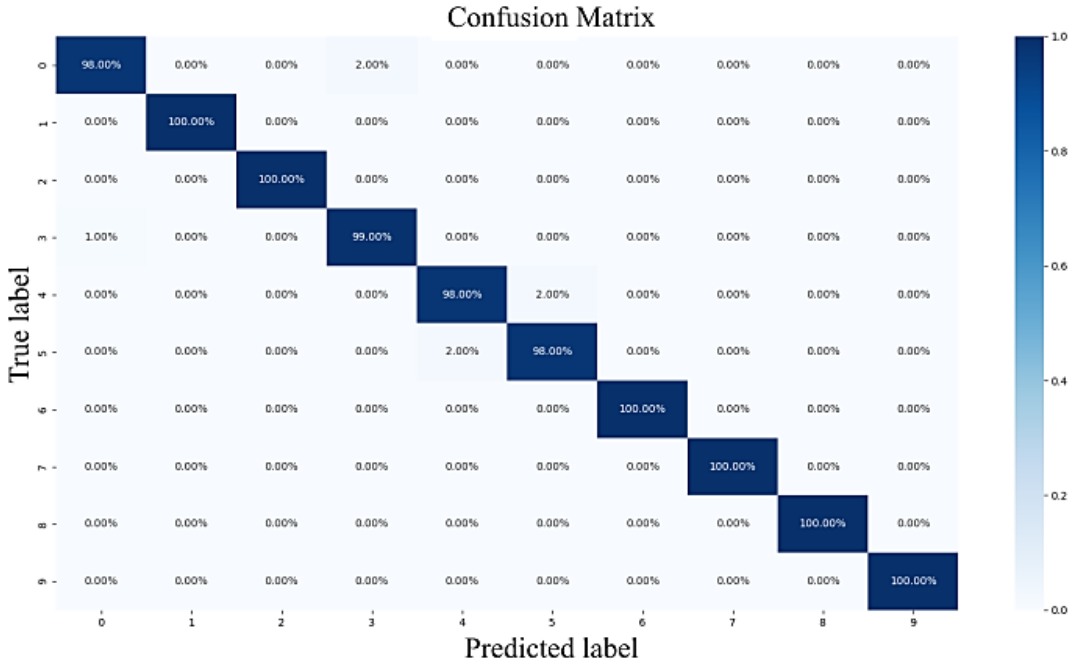


**Figure 10 – Confusion Matrix for validation in frequency domain.**

Again, there is no model overfitting for the number of epochs tested when analyzing the frequency domain results. Epoch 52 presented the best accuracy result with the lowest cost function value.

From the confusion matrix obtained for the validation, one has that the classes with high classification capacity (1, 2, 6, 7, 8 and 9) obtained 100% discriminatory precision. The classes with lower prediction accuracy (0, 3, 4 and 5), despite the lower nominal value, obtained precision above 98%, which indicates a very high performance in the classification of normal and fault cases.

## CONCLUSION AND FUTURE WORK

The proposed method for verifying the best input condition of the database for the constructed CNN model allowed the conclusion that the vibration images in frequency domain shows better contrasts between fault types, due to the feature extraction characteristics of the FFT. Comparing the accuracy results with the PdM-CNN model (Souza et al. 2021), the prediction accuracy of the CNN model with input from vibration images in the frequency domain was higher. Moreover, considering that the entire database was used in the present work, without elimination of any sequence to favor the classification (as done in Souza et al. 2021), nor characteristic pre-treatment of the vibration signal (even though the pre-treatment performed on the data in Souza et al. 2021 is not specified, a simple comparison of a FFT presented in the paper in question shows that the data was pre-treated), the results presented in the present work are more robust and easier to apply in practice in industrial plant machines.

Other attempts were made with the input of 2D amplitude *versus* time and amplitude *versus* frequency graphs, but they did not generate an accuracy greater than 80% of classification, which is understandable, given that the input of graphical images to the network constitutes a more challenging task than the analysis of the input with raw data. Other architectures and parameters of the CNN models will be tested, as well as checks for more incipient fault conditions.

## REFERENCES

Brito, L. C., Susto, G. A., Brito, J. N., Duarte, M. A. V., “An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery”, Mechanical Systems and Signal Processing, 163–108105, 21 p.

CWRU, “Case Western Reserve University Bearing Data Center”. <https://cwru-db-group.github.io>. Last access 01/07/2022.

Goodfellow, I., Bengio. Y., Courville, A., 2016, “Deep Learning”, MIT Press.

Hoang, D. T., Kang, H. J, "Rolling element bearing fault diagnosis using convolutional neural network and vibration image",Cognitive Systems Research, Article in Press, 2018.

Jha, R. K., Swami, P. D., "Fault diagnosis and severity analysis of rolling bearings using vibration image texture enhancement and multiclass support vector machines", Applied Acoustics, 182 – 108243, 2021.

Krizhevsky, A., Sutskever, I., Hinton, G. E., 2012, "ImageNet Classification with Deep Convolutional Neural Networks", Advances in Neural Information Processing Systems 25, 9 p.

Liu, J., Yang, X., Li, L., "VibroNet: Recurrent neural networks with multi-target learning for image-based vibration frequency measurement", Journal of Sound and Vibration, 457 – 51-66, 2019.

Ribeiro, F. M. L., 2018, "Machinery Fault Database (MaFaulDa) - multivariate time-series acquired by sensors on a SpectraQuest's Machinery Fault Simulator (MFS) Alignment-Balance-Vibration -ABVT". <http://www02.smt.ufrj.br/~offshore/mfs/page_01.html>. Last access 01/07/2022.

Souza, R. M., Nascimento, E. G. S., Miranda, U. A., Silva, W. J. D., Lepikson, H. A., 2021, "Deep learning for diagnosis and classification of faults in industrial rotating machinery", Computers & Industrial Engineering, 153-107060, 12 p.

Yang, F., Yun, Z., Gao, Y., Gao, H., Mao, T., Zhou, H., Li, D., "Machining vibration states monitoring based on image representation using convolutional neural networks", Engineering Applications of Artificial Intelligence, 65 – 240-251, 2017.

## RESPONSIBILITY NOTICE

The author(s) is (are) the only responsible for the printed material included in this paper.