Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 220 (2023) 535–544

Procedia Computer Science

www.elsevier.com/locate/procedia

The 14th International Conference on Ambient Systems, Networks and Technologies (ANT)
March 15-17, 2023, Leuven, Belgium

The effect of augmentation and filtration on noisy environment's acoustic signals to detect abnormalities in industrial machines based on artificial neural networks

Ahmad Qurthobi[a], Rytis Maskeliūnas[a,*]

[a]*Department of Multimedia Engineering, Faculty of Informatics, Kaunas University of Technology, Studentų st. 50, Kaunas 51368, Lithuania*

Abstract

The use of acoustic methods for failure detection is increasing in recent times. By using the acoustic method, some features related to the condition of the equipment under test can also be extracted. The main objective of this study was to compare the effect of using filters and augmentation on audio signals in the data training process on the accuracy and loss generated during training and testing. The model is developed using an artificial neural network. The MIMII dataset, which can be accessed openly, is being used for training and testing purposes. Based on the four scenarios designed, it can be seen that the movement of changes in accuracy values in tests involving the use of original data and augmented data is faster (close to 100% after the twentieth epoch) than the other methods. This is because the augmentation process in the form of a time shift does not change the signal in the frequency domain. Moreover, the test signal used is a periodic signal.

Keywords: Acoustic; mechanical failure; MIMII dataset; artificial neural network;

1. Introduction

The use of acoustic methods for failure detection is increasing in the current state[31]. This is due to the ease of data collection and installation of the necessary instruments, especially if applied in an industrial environment[26]. In addition, by using the acoustic method, some features related to the condition of the equipment under test can also be extracted. In fact, by using the acoustic method, it is still possible to detect failures of the device in a very noisy environment[23, 15].

The appearance of a failure in a device can be recognized from the appearance of an abnormal sound when the machine is operating[4, 11]. If a person has an observant enough ear, he will easily detect which device is operating under

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000.
E-mail address: rytis.maskeliunas@ktu.lt

10.1016/j.procs.2023.03.068

normal and abnormal conditions. However, in noisy conditions, the human ear will generally have difficulty detecting it because of noise interference. Moreover, to some extent, working in a high noise level area can be the cause of serious health issues[3, 14]. Therefore, various classic and artificial intelligence-based failure recognition methods have been developed to handle this condition[29, 26, 23, 33].

Several researchers are trying to develop a failure detection system based on artificial intelligence. Firmino et. al[8], Karabacak et. al.[12], and Taha and Widiyati[27] developed a mechanical failure detection method based on an artificial neural network. Fezari et. al.[7] and Pandya et. al.[20] employed k-nearest neighbors for detecting the abnormalities. Support vector machines were also involved by Al-Obaidi et. al[1], Heydarzadeh et. al.[10], and Kothuru et. al.[13] to categorize the condition of running equipment. Other researchers also used the convolutional neural networks[23] and genetic algorithms[19] to approach the failure condition. However, most of these studies were conducted with audio signals obtained on laboratory scale testing. Therefore, this study tries to design a failure detection based on audio signals obtained in noisy conditions.

The main objective of this study is to compare the effect of audio signal filtration and augmentation in the data training process. The level of influence will be reviewed based on the value of accuracy and loss. The model is developed using an artificial neural network. The dataset used in the training is a MIMII dataset that can be accessed openly[22].

The structure of this paper is as follows. The structure of this paper is as follows. Section 2 discusses the previous studies related to the application of artificial intelligence to failure detection. In Section 3, the configuration of the artificial neural network which will be the reference in this research will be presented. The training scenarios will be explained in Section 4. Section 5 discusses the test results and discusses each scenario. Finally, Section 6 contains conclusions of the conducted study.

2. Related Works

This section discusses several studies related to the application of artificial intelligence for failure detection in industrial machines. In simple terms, failure in industrial machines can be defined as a condition where findings are obtained in the form of the appearance of signal patterns that are not similar to normal conditions.

As mentioned in Section 1, one of the methods used to detect artificial intelligence-based failures is artificial neural networks(ANN). ANN, briefly called neural network (NN), is a computing system inspired by the working principle of biological neural networks found in the brains of living things[24, 32, 21]. Firmino et. al.[8] designed an artificial neural network to detect misfires in internal combustion engines by combining analysis on vibration and acoustic signals and yielding an accuracy value of 98.7 to 99.3%. In another study, Taha and Widiati[27] designed a bearing crack detection system using an artificial neural network based on acoustic emission signals. Their research succeeded in producing a failure detection system using artificial intelligence with an accuracy in the range of 80-100%.

Another artificial intelligence method used by researchers to detect machine failures using artificial intelligence is using a support vector machine. This method was first introduced by Cortes and Vapnik in 1995[5]. The use of SVM can be applied to detect failures in valves[1], gearboxes[10], bearings[19], rotating machines[2], gear reducers[18], mills[25], and others.

In addition to these two types of algorithms, several studies also involve other algorithms. There are several artificial intelligence algorithms used in failure detection including adaptive neuro-fuzzy inference system[12], k-nearest neighbors[7], genetic algorithms[19], convolutional neural networks[23], random forests[6], decision trees[16], regression tree[9], and long short-term memory[17].

3. Artificial Neural Network Design

3.1. Network Architecture

Fig. 1 shows the developed neural network architecture for this study. It shows the design of an artificial neural network which consists of one input layer, two hidden layers, and an output layer. The output is expected to decide whether the tested signal is from a normal or abnormal operation.

The training carried out on the designed system runs on 4 layers that are between input and output. In the first training layer, a dense layer of 64 fully connected neurons is used. This layer uses the activation function of the rectified linear

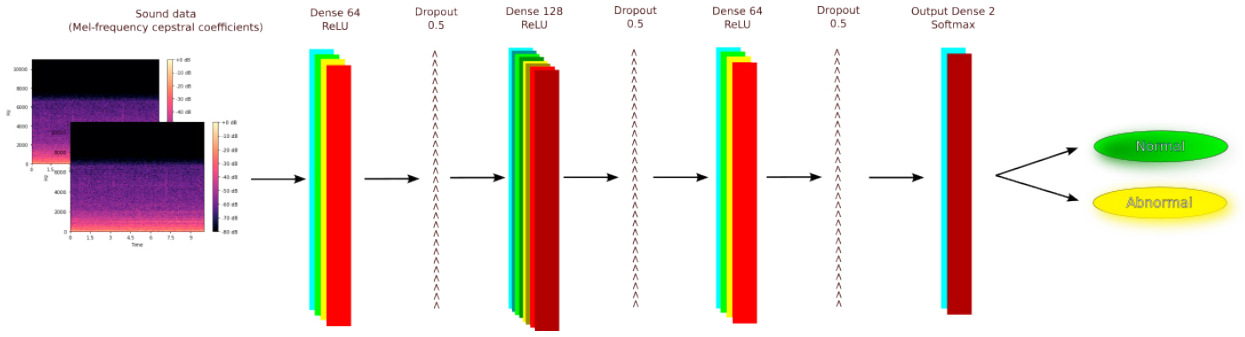


Fig. 1. Developed neural-network architecture

unit (ReLU) which satisfies Eq. 1. This layer also uses a dropout value $p = 0.5$ and is given to avoid overfitting during the training process.

$$f(x) = x^+ = \max(0, x) \tag{1}$$

$$\mathbf{y} = \mathbf{wx} + \mathbf{b} \tag{2}$$
$$= \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

In the second training layer, a dense layer of 128 fully connected neurons is used. The relationship between the output signal of the previous layer and the incoming signal to the next layer can be expressed as eq. 2, where $\mathbf{y}$, $\mathbf{w}$, $\mathbf{x}$, and $\mathbf{b}$ represent inputs to the next layer, the weighting matrix, outputs of the previous layer, and biases. As in the first layer, the ReLU activation function and dropout value with $p = 0.5$ were also involved. In the third training layer, a dense layer of 64 fully connected neurons was used. This layer also employed ReLU as its activation function and $p = 0.5$ for the dropout value. Finally, in the fourth training layer, a dense layer is used with two fully connected neurons. Unlike the previous training layers, this layer uses the softmax activation function which satisfies Eq. 3 for $i = 1, 2, \ldots, K$ and $\mathbf{x} = (x_1, x_2, \ldots, x_K) \in \mathbb{R}^K$. This last layer is the determinant of the resulting output.

$$\sigma(\mathbf{x})_i = \frac{e^{x_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{3}$$

3.2. Implementation and Dataset

The implementation process is carried out using the python programming language with the help of the Tensorflow 2.8.0 library. Code is written and run on Google Colab for easy debugging. In addition, to facilitate the accuracy and loss analysis process, a tensorboard is used to display the movement of changes in the accuracy value and loss of each epoch in each test scenario.
As mentioned in Section 1, the dataset used in the training and testing process is a collection of audio signals consisting of a collection of sound recordings of machines running under normal and abnormal conditions collected by Purohit et. al.[22](also known as MIMII dataset) and can be accessed openly[28, 30]. The collection of fan sound recordings with id_00 for noise level with signal to noise of 6 dB is used in this research. The number of recordings used consisted of 407 machine voice recording files under abnormal conditions and 1011 recording files under normal conditions.

4. Training Scenario

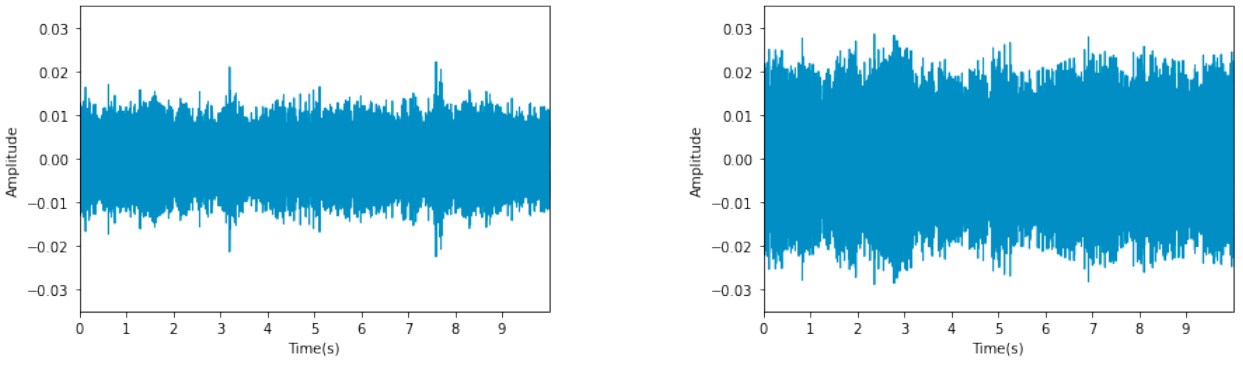
Four scenarios have been prepared to investigate the effect of using augmentation and filters on audio signal training and testing. These scenarios involve the use of original audio signals, filtered audio signals, and augmented audio

signals. For each scenario, the training process is run in 100 epochs with the number of batch sizes used being 64. Meanwhile, 82 and 203 records under conditions, respectively, abnormal and normal were involved in the validation process.

1. First Scenario
   In the first scenario, the dataset used for the training process is a recording of the original normal and abnormal audio signal from the machine with ID 00 with a signal-to-noise of 6 dB. In this scenario, 325 of the 407 recordings and 808 of 1011 recordings for, respectively, abnormal and normal conditions are used for the training process and the rest are used for testing.
2. Second Scenario
   As in the first scenario, the dataset used for the training and validation process is the original recording of normal and abnormal audio signals. However, apart from using the original audio signal, the training process also involves audio signal augmentation. Therefore, in this second scenario, 325 original recordings under abnormal conditions, 808 original recordings under normal conditions, 325 augmented abnormal original recordings, and 808 augmented normal recordings were used for the training process.
3. Third Scenario
   The third scenario is basically similar to the second scenario. However, instead of using augmented data, this scenario involves filtering the audio signal before it is used as an additional signal for the training process. The filtering process is carried out by allowing each training audio signal to pass through a low-pass filter with a critical frequency of 6000 Hz. This critical frequency value was chosen randomly because the purpose of this test was to determine the effect of data loss with a certain frequency on the training and validation process. As in the second scenario, 80% of the original recorded signal of normal and abnormal conditions was involved in the training process. However, in addition, each original recorded signal will undergo a filtering process before being entered as data for training.
4. Fourth Scenario
   The fourth scenario is a combination of the three previous scenarios. The original audio recording signal, the augmented signal, and the filtered signal were involved in the training process.

5. Results and Discussion

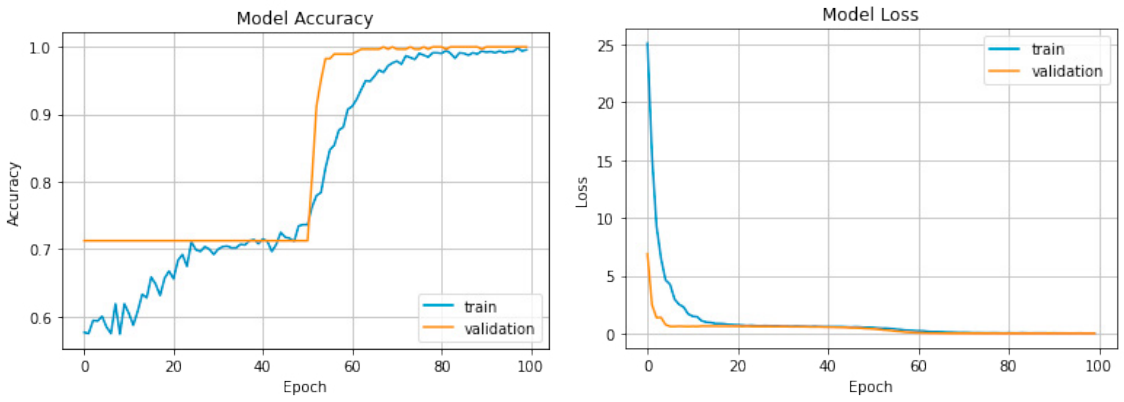
5.1. First Scenario



(a) normal (b) abnormal

Fig. 2. Original waveforms

Fig. 2 shows the waveform of the machine running under normal and abnormal conditions, with the horizontal and vertical axes representing the time(in seconds) and the amplitude, respectively. The amplitude value is dimensionless because it is a comparison of the measured value and the maximum value of the sample measurement results. The



(a) Epoch accuracy (b) Epoch loss

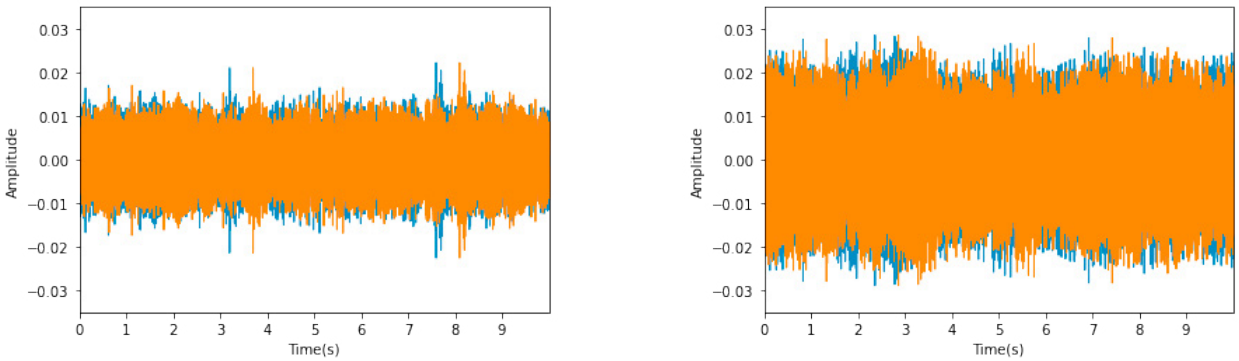
Fig. 3. Epoch accuracy and epoch loss on the first scenario

figure shows a fairly clear difference between the normal and abnormal signal waveforms in the time domain.

The change in the accuracy value for each epoch is shown by Fig. 3a where the x-axis represents the epoch and the y-axis represents the accuracy value. From the figure, it can be seen that there is a significant increase in the accuracy value after the fiftieth epoch. This can be read from the steep gradient in the orange line, where the orange line represents the movement of accuracy based on the validation results, while the blue line indicates the movement of the accuracy of the training results. The figure also shows that the accuracy value of the validation results tends to be higher than the value of the training results, except for the 48th, 49th, and 50th epochs.

On the other hand, the change in the loss value for each epoch is shown in Fig. 3b. The figure shows that the loss value of the validation results tends to be smaller than the training results. Even the loss value of the validation results after the sixtieth epoch tends to be close to zero. This is because the accuracy value of the epoch has a value close to 100%, as shown in Fig. 3a.

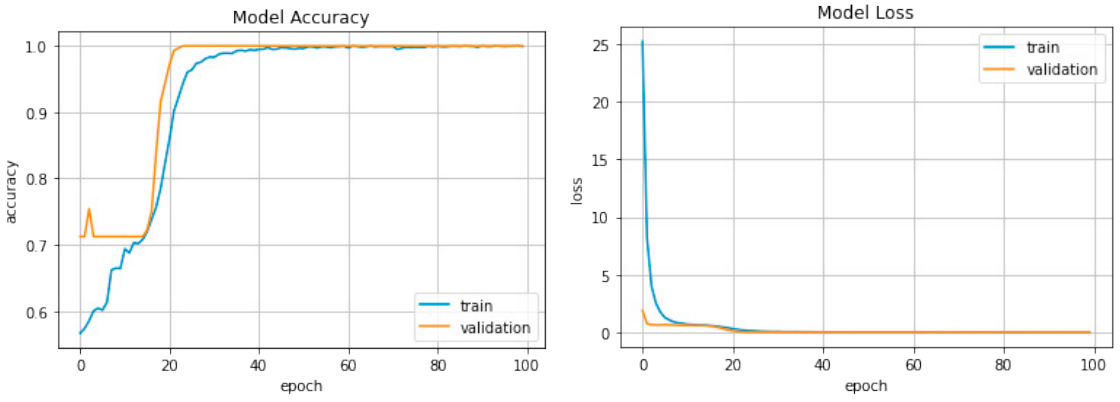
5.2. *Second Scenario*



(a) normal (b) abnormal

Fig. 4. Comparison between original signal(blue) and augmented signal(orange) waveforms

Fig. 4 shows the comparison between the original signal(blue) and the augmented signal(orange), with the horizontal and vertical axes representing the time(in seconds) and the amplitude, respectively. As in Fig. 2, the amplitude value has no dimensional value. Augmentation applied in this study is to time-shift each original audio signal used in the training process, where each signal will be shifted half a second from its original position and the last half-second will appear in the initial position. The purpose of this test is to determine the effect of signal retrieval time on the



(a) Epoch accuracy (b) Epoch loss

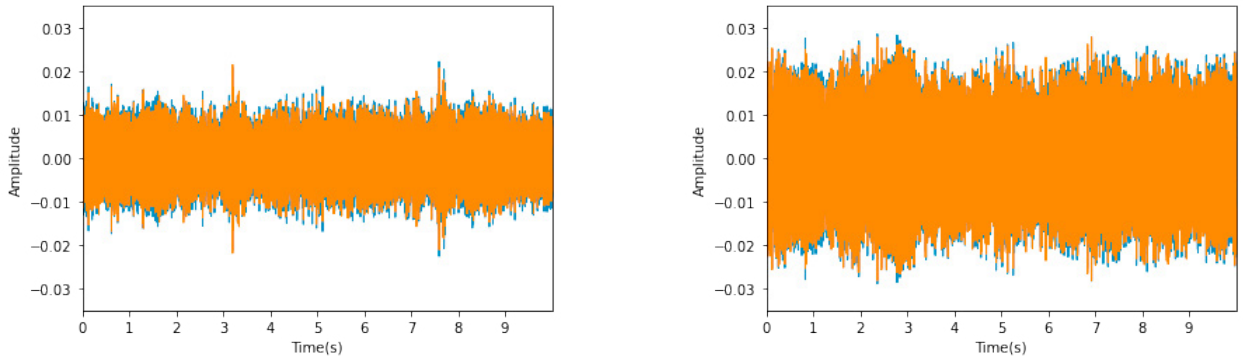
Fig. 5. Epoch accuracy and epoch loss for the second scenario

recognition process.

Fig. 5a displays the movement of the accuracy value of training and validation results for every epoch for the second scenario. The figure shows a significant change toward higher accuracy occurs at an earlier epoch when compared to the first scenario. In this scenario, a significant increase occurred after the fourteenth epoch, whereas the same thing happened to the fiftieth epoch in the first scenario. The figure also shows the tendency of the accuracy value of the validation results which is always higher when compared to the value of the accuracy of the training process.

The change in the loss value for each epoch is shown in Fig. 5b. The figure shows the process of writing down losses occurs faster than in the first scenario. This provides information that additional signals that experience time shifts in the training process actually provide support for a higher level of accuracy and a faster reduction in loss values. In addition, the training and testing accuracy values are close to 100% after the twentieth epoch.

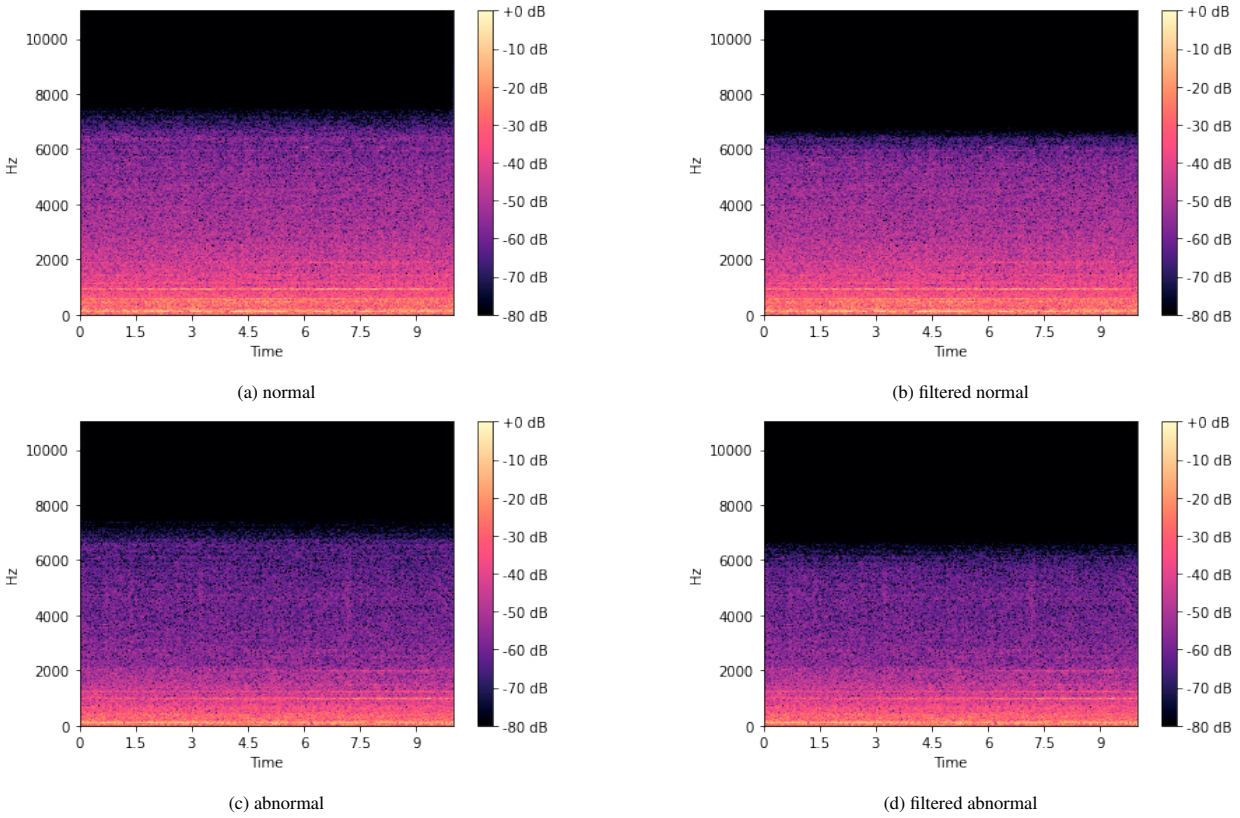
5.3. Third Scenario



(a) normal (b) abnormal

Fig. 6. Comparison between original signal(blue) and filtered signal(orange) waveforms

Fig. 6 shows the comparison between the original signal(blue) and the filtered signal(orange). As in Fig. 2 and 4, the horizontal and vertical axes represent the time(in seconds) and the amplitude, respectively. Based on the figure, the original signal and the filtered signal are almost indistinguishable in the time domain. However, when observed in the frequency domain, the effect of this filtering process causes the loss of signals having frequencies above the critical frequency, as shown in Fig. 7. Additionally, Fig. 7a and 7c also show the shape of the spectrogram signal from normal and abnormal signals which are quite difficult to distinguish by eye. However, the values on the spectrogram



(a) normal

(b) filtered normal

(c) abnormal

(d) filtered abnormal

Fig. 7. Comparison between the original signal and filtered signal spectrograms

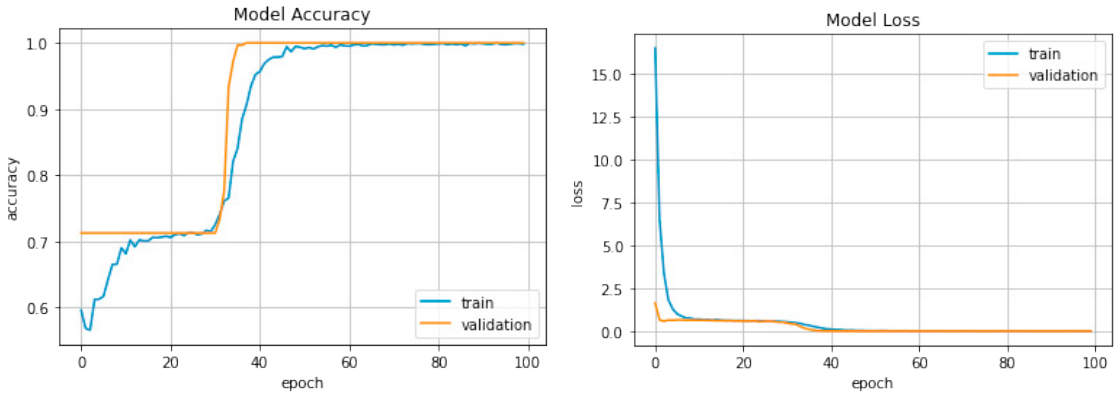
will be easily recognized by computational methods.

Fig. 8a displays the movement of the accuracy value of training and validation results for every epoch for the third scenario. This figure provides information about a significant change towards higher accuracy that occurs in the thirtieth epoch, which indicates an earlier condition when compared to the first scenario (fiftieth epoch), but slower than the second scenario (thirteenth epoch). The figure also provides information on the accuracy value of the validation results which tend to be higher than the training results, except for the twenty-eighth to thirty-first epochs. The movement of the loss value from the training and validation results for each epoch is shown in Fig. 8b. The loss value in the training process moves down faster than the conditions in the first scenario. In fact, the loss value from the training and validation results, respectively, has reached 0.1304 and 0.0077 in the fortieth epoch and continues to fall in the following epoch. This shows that, as in the second scenario, the addition of the screening data can help improve the quality of the training process.

In this study, the critical frequency values were randomly selected to add to the training data. Even though the training results were closer to 100% at an earlier epoch (after the thirtieth epoch) than the first experiment, it can be seen that there is missing information in the spectrogram in Fig. 7.

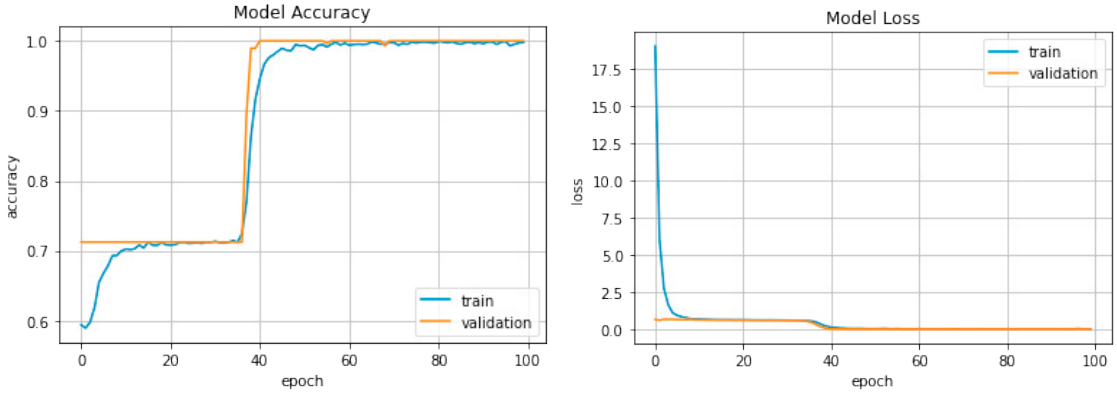
5.4. Fourth Scenario

Fig. 9a shows the change in the value of the accuracy of the training and validation results for each epoch for the fourth scenario. When compared with the previous training, the change in accuracy in the fourth scenario was slower than in the second and third scenarios, even though more signals were used for the training process. This is caused by the high variation of the training signal and slows down the training and validation process. On the other hand, Fig. 9a presents the change of epoch loss for every epoch during the training and validation process. In the figure, the



(a) Epoch accuracy (b) Epoch loss

Fig. 8. Epoch accuracy and epoch loss for the third scenario



(a) Epoch accuracy (b) Epoch loss

Fig. 9. Epoch accuracy and epoch loss for the forth scenario

training and validation loss values at the fortieth epoch are 0.1899 and 0.0447, respectively. This value is certainly greater when compared to the second and third scenarios.

6. Conclusion

This article discusses the effect of using augmentation and screening on the training and validation process using artificial neural networks. The artificial neural network built consists of one input layer, four fully connected training layers, and one output layer. Based on the four scenarios designed, it can be seen that the movement of the increase in accuracy value in the second scenario, which involves the use of original data and augmented data, is faster than the movement in the other scenarios, where accuracy approaching 100% is achieved after the twentieth epoch. This is due to the augmentation process, which is a time shift, almost does not change the shape of the original signal which is a periodic signal. The second scenario also produces a faster movement in accuracy than the third scenario (original and filtered data, close to 100) after the fiftieth epoch) and the fourth scenario (original data, filtered data, and augmented data, close to 100 after the thirtieth epoch). However, further research is needed to determine the filter parameter values so as not to eliminate the required signal characteristics. In addition, the use of dense and fully connected artificial neural networks will also be a burden on the computing process.

References

[1] Al-Obaidi, S., Hui, K.H., Hee, L., Leong, M., 2017. Automated valve fault detection based on acoustic emission parameters and support vector machine. Alexandria Engineering Journal 57. doi:10.1016/j.aej.2016.12.010.

[2] Altaf, M., Khan, M., Ahmad, A., Badshah, S., Shah, J., Anjum, M.A., 2019. Automatic and efficient fault detection in rotating machinery using sound signals. Acoustics Australia 47. doi:10.1007/s40857-019-00153-6.

[3] Araújo Alves, J., Paiva, F., Silva, L., Remoaldo, P., 2020. Low-frequency noise and its main effects on human health—a review of the literature between 2016 and 2019. Applied Sciences 10, 39. doi:10.3390/app10155205.

[4] Cooper, C., Wang, P., Zhang, J., Gao, R., Roney, T., Ragai, I., Shaffer, D., 2020. Convolutional neural network-based tool condition monitoring in vertical milling operations using acoustic signals. Procedia Manufacturing 49, 105–111. doi:10.1016/j.promfg.2020.07.004.

[5] Cortes, C., Vapnik, V., 1995. Support vector network. Machine Learning 20, 273–297. doi:10.1007/BF00994018.

[6] Cruz, R., Silva, F., Fileti, A., 2020. Machine learning and acoustic method applied to leak detection and location in low-pressure gas pipelines. Clean Technologies and Environmental Policy 22. doi:10.1007/s10098-019-01805-x.

[7] Fezari, M., Taif, F., Lafifi, M.M., 2014. Noise emission analysis a way for early detection and classification faults in rotating machines. doi:10.1109/EPEPEMC.2014.6980655.

[8] Firmino, J., Neto, J., Oliveira, A., Silva, J., Mishina, K., Rodrigues, M., 2021. Misfire detection of an internal combustion engine based on vibration and acoustic analysis. Journal of the Brazilian Society of Mechanical Sciences and Engineering 43. doi:10.1007/s40430-021-03052-y.

[9] Griffin, J., Shanbhag, V., Pereira, M., Rolfe, B., 2021. Application of machine learning for acoustic emissions waveform to classify galling wear on sheet metal stamping tools. doi:10.21203/rs.3.rs-186756/v1.

[10] Heydarzadeh, M., Nourani, M., Hansen, J., Kia, S., 2017. Non-invasive gearbox fault diagnosis using scattering transform of acoustic emission. doi:10.1109/ICASSP.2017.7952180.

[11] Holford, K., Eaton, M., Hensman, J., Pullin, R., Evans, S., Dervilis, N., Worden, K., 2017. A new methodology for automating acoustic emission detection of metallic fatigue fractures in highly demanding aerospace environments: An overview. Progress in Aerospace Sciences 90. doi:10.1016/j.paerosci.2016.11.003.

[12] Karabacak, Y., Ozmen, N., 2021. Common spatial pattern-based feature extraction and worm gear fault detection through vibration and acoustic measurements. Measurement 187. doi:10.1016/j.measurement.2021.110366.

[13] Kothuru, A., Nooka, S., Liu, R., 2018. Application of audible sound signals for tool wear monitoring using machine learning techniques in end milling. The International Journal of Advanced Manufacturing Technology 95. doi:10.1007/s00170-017-1460-1.

[14] Lai, A.J., Huang, C.Y., 2019. Effect of occupational exposure to noise on the health of factory workers. Procedia Manufacturing 39, 942–946. doi:10.1016/j.promfg.2020.01.395.

[15] Li, G., Wu, J., Deng, C., Chen, Z., 2021. Parallel multi-fusion convolutional neural networks based fault diagnosis of rotating machinery under noisy environments. ISA Transactions 128. doi:10.1016/j.isatra.2021.10.023.

[16] Liu, X., Pei, D., Lodewijks, G., Zhao, Z., Mei, J., 2020. Acoustic signal based fault detection on belt conveyor idlers using machine learning. Advanced Powder Technology 31. doi:10.1016/j.apt.2020.04.034.

[17] Medina, R., Cerrada, M., Cabrera, D., Sánchez, R., Li, C., de Oliveira, J., 2019. Deep learning-based gear pitting severity assessment using acoustic emission, vibration and currents signals, pp. 210–216. doi:10.1109/PHM-Paris.2019.00042.

[18] Oh, S.W., Lee, C., You, W., 2019. Gear reducer fault diagnosis using learning model for spectral density of acoustic signal, pp. 1027–1029. doi:10.1109/ICTC46691.2019.8939913.

[19] Omoregbee, H., Heyns, S., 2019. Fault classification of low-speed bearings based on support vector machine for regression and genetic algorithms using acoustic emission. Journal of Vibration Engineering & Technologies 7. doi:10.1007/s42417-019-00143-y.

[20] Pandya, D., Upadhyay, S., HARSHA, S., 2013. Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using apf-knn. Expert Systems with Applications 40, 4137–4145. doi:10.1016/j.eswa.2013.01.033.

[21] Park, K.C., Motai, Y., Yoon, J., 2017. Acoustic fault detection technique for high power insulators. IEEE Transactions on Industrial Electronics PP, 1–1. doi:10.1109/TIE.2017.2716862.

[22] Purohit, H., Tanabe, R., Ichige, K., Endo, T., Nikaido, Y., Suefusa, K., Kawaguchi, Y., 2019. Mimii dataset: Sound dataset for malfunctioning industrial machine investigation and inspection. URL: https://arxiv.org/abs/1909.09347, doi:10.48550/ARXIV.1909.09347.

[23] Qiao, M., Yan, S., Tang, X., Xu, C., 2020. Deep convolutional and lstm recurrent neural networks for rolling bearing fault diagnosis under strong noises and variable loads. IEEE Access PP, 1–1. doi:10.1109/ACCESS.2020.2985617.

[24] Simonovic, M., Kovandzic, M., Ciric, I., Nikolic, V., 2021. Acoustic recozgnition of noise-like environmental sounds by using artificial neural network. Expert Systems with Applications 184, 115484. doi:10.1016/j.eswa.2021.115484.

[25] Sun, S., Hu, X., Zhang, W., 2020. Detection of tool breakage during milling process through acoustic emission. The International Journal of Advanced Manufacturing Technology 109, 1–10. doi:10.1007/s00170-020-05751-7.

[26] Tagawa, Y., Maskeliunas, R., Damasevicius, R., 2021. Acoustic anomaly detection of mechanical failures in noisy real-life factory environments. Electronics 10, 2329. doi:10.3390/electronics10192329.

[27] Taha, Z., Widiyati, K., 2010. Artificial neural network for bearing defect detection based on acoustic emission. The International Journal of Advanced Manufacturing Technology 50, 289–296. doi:10.1007/s00170-009-2476-y.

[28] Tan, J., Oyekan, J., 2021. Attention augmented convolutional neural network for acoustics based machine state estimation. Applied Soft Computing 110, 107630. doi:10.1016/j.asoc.2021.107630.

[29] Vicuna, C., Howeler, C., 2017. A method for reduction of acoustic emission (ae) data with application in machine failure detection and diagnosis. Mechanical Systems and Signal Processing 97. doi:10.1016/j.ymssp.2017.04.040.

[30] Wang, L., Sun, G., Wang, Y., Ma, J., Zhao, X., Liang, R., 2022. Afexplorer: Visual analysis and interactive selection of audio features. Visual

Informatics 6. doi:10.1016/j.visinf.2022.02.003.
[31] Weintroub, S., 1960. Noise in factories and its control. Nature 187, 1085. doi:10.1038/1871085a0.
[32] Yun, H., Kim, H., Jeong, Y., Jun, M., 2021. Autoencoder-based anomaly detection of industrial robot arm using stethoscope based internal sound sensor. Journal of Intelligent Manufacturing , 1–18doi:10.1007/s10845-021-01862-4.
[33] Zhang, J., Yang, J., Litak, G., Hu, E., 2019. Feature extraction under bounded noise background and its application in low speed bearing fault diagnosis. Journal of Mechanical Science and Technology 33. doi:10.1007/s12206-019-0614-5.