

Denoising diffusion implicit model for bearing fault diagnosis under different working loads

Toong Yang Wong^{1*}, Meng Hee Lim^{1,2}, Wai Keng Ngui³, and Mohd Salman Leong²

¹ Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur 54100, Malaysia

² Institute of Noise and Vibration, Universiti Teknologi Malaysia, Kuala Lumpur 4100, Malaysia

³ Faculty of Mechanical & Automotive Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Pahang 26600, Malaysia

Abstract. Rotating machineries always operating under different loads and suffer from various types of bearing fault. Thus, bearing fault diagnosis is essential to prevent further loss or damage. Deep learning has been favoured over machine learning recently due to data explosion and its higher performance. In deep learning-based bearing fault diagnosis, vibration signals are usually transformed into images using time frequency analysis methods such as short-time Fourier transform, wavelet transform, and Hilbert-Huang transform. Convolutional neural network (CNN) is widely used for fault classification method. However, the training dataset and testing dataset usually have different load domains due to different working conditions. Obtaining training data of wide range of loadings are impractical and exhausting. Thus, this study is proposed to solve load domain adaptation using denoising diffusion implicit model (DDIM). In this study, synthetic images are generated using DDIM model while only convolutional neural network (CNN) is used as fault classification model. The classification accuracy of testing dataset is obtained using CNN models trained with original training dataset and augmented training dataset. The results showed that the synthetic scalograms could improve the performance of CNN model by 3.3% under different load domains.

1 Introduction

Rotating machineries such as steam turbines, compressor, gearboxes, aircraft engines, and generators always suffer from various bearing faults due to long working hour under different loads. Serious safety issues and financial losses might as well happen. Thus, highly efficient fault diagnosis is important to detect fault in advance and prevent greater consequences. Commonly found bearing faults are ball fault, inner race fault, and outer race fault.

Deep learning method is widely used for bearing fault diagnosis. Bearing fault diagnosis requires input data which can be captured by sensors in the form of vibration signals, thermal-imaging, acoustic noise, and motor current [1]. Examples of popular vibration signals dataset

* Corresponding author: wongty27@gmail.com

used by researchers are Case Western Reserve University (CWRU) dataset, Paderborn University dataset, PRONOSTIA dataset, and Intelligence Maintenance Systems (IMS) dataset [1]. Vibration signals can be used directly as time series data or two-dimensional images [2]. Common time-frequency analysis used to transform time series vibration signal into two-dimensional images are short-time Fourier Transform (STFT), continuous wavelet transform (CWT), and Hilbert Huang Transform (HHT) [3]. CNN model is usually used for image classification.

In industrial environment, machines always operate under different loads and suffer from various types of bearing fault. Gathering of new data and training of new classification model are too time consuming and computationally expensive.

Generative model such as generative adversarial network (GAN) has been widely studied to diagnose bearing fault. GAN is able to generate high quality images in a short time and has been used to generate synthetic images to oversample imbalanced dataset [4]–[7]. However, many studies of diffusion models are focused on imbalanced datasets instead of domain adaptation.

In this study, denoising diffusion implicit model (DDIM) is utilized for load domain adaptation. Continuous wavelet transform (CWT) is used to convert vibration signals into scalograms. Synthetic images are generated by DDIM for dataset augmentation purpose. Two CNN models are trained and tested with non-augmented dataset and augmented dataset respectively for results' comparison.

2 Bearing Fault Classification

2.1 Vibration Signal Collection

Vibration signals of ball bearing are collected from Bearing Data Center by Case Western Reserve University (CWRU) [2]. CWRU dataset is chosen due to its accessibility and popularity among researchers which enables comparison of results. The specifications of datasets are shown in table 1.

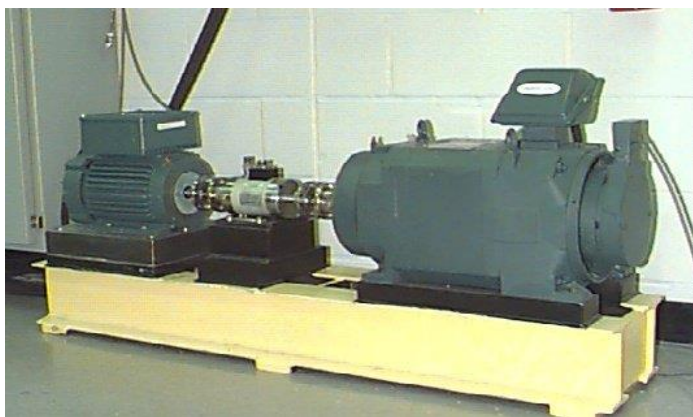


Fig. 1. Experimental setup of CWRU bearing dataset [2].

Table 1. Specifications of bearing dataset.

Bearing Type	Drive end
Fault Size (inch)	0.007, 0.014, 0.021
Working Load	0 hp, 1 hp, 2 hp, 3 hp
Health Conditions	Normal, inner race fault, outer race fault, ball fault
Sampling Frequency (Hz)	12000

2.2 Image Transformation of Vibration Signals

Time series vibration signal has too many data points and cannot fit in one image. Thus, data segmentation is applied to create segments with equal size. Overlapping sliding window segmentation method [8], [9] is used for segmentation without any loss of features. The length of segment must be enough to cover a bearing’s full rotation. Given the sampling frequency, F_s and motor speed, ω are provided, the segment length can be obtained with equation 1 [10]. The step size is 25% of the segment length. Next, continuous wavelet transform (CWT) are applied to the segments of vibration signal to obtain scalograms. Complex Morlet wavelet is used as mother wavelet.

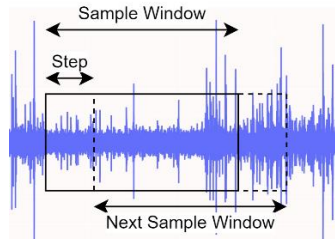


Fig. 2. Overlapping sliding window method [11].

$$L_{segment} = \frac{F_s \times 60}{\omega} \tag{1}$$

2.3 Experimental Conditions

Table 2 shows the distribution of dataset for each domain. Each dataset represents different operating loads of 0 hp, 1 hp, 2 hp, and 3 hp. The dataset is divided into two domains: first is the source domain, which is labelled dataset used for training; the second is target domain, which is unlabelled dataset used for testing. Studies of bearing fault diagnosis under different working loads usually involve fault classification using multiple classification models [11], [12]. In this study, the performance of classification model is studied using original training dataset and augmented dataset.

Source Domain (Load)	Target Domain (Load)
----------------------	----------------------

0 hp	1 hp, 2 hp, 3 hp
1 hp	0 hp, 2 hp, 3 hp
2 hp	0 hp, 1 hp, 3 hp
3 hp	0 hp, 1 hp, 2 hp

Table 2. Dataset for source domain and target domain.

2.4 Generative Model for Synthetic Scalogram

Denoising diffusion implicit model (DDIM) is used in this study to generate synthetic images. DDIM can generate high quality images without adversarial learning and outperforms denoising diffusion probabilistic model (DDPM) [13]. DDIM uses U-Net as its backbone for continuous downsampling and upsampling process of images. Skip connections are employed in the U-Net to decrease information loss during down-sampling and up-sampling of images. Deterministic sampling procedure is implemented in DDIM instead of stochastic sampling for faster images generation [13].

Kernel Inception Distance (KID) is used as loss function to validate the training progress by measuring the similarity between inference samples and training samples. KID is chosen over Frechet Inception Distance (FID) because it is computationally lighter and easier to implement. KID is applied only on evaluation step to use lesser computational resource. During the calculation, the dimensions of images are resized to the minimal resolution of Inception network, which is (75x75) for faster calculation. Model checkpoint callback is used to save the trained DDIM with best weight after every epoch. The best weight is determined by the lowest KID loss.

For this research, two datasets are prepared for training of two models. First model is trained with original dataset. Synthetic scalograms will be added to the original dataset to create a new dataset (augmented dataset) to train second model.

2.5 Fault Classification Model

CNN architecture is used as fault classification model. The input layer of CNN takes RGB image of size 128x128 as input data. There are three convolutional layers with filter size 32, 64, 128 respectively. Each convolutional layer has a 3x3 kernel, 'same' padding, and 'relu' activation. Max pooling layer are applied to first two convolutional layers. Global average pooling layer is added to the last convolutional layer instead of fully connected layer because global average pooling is more robust to spatial translations of the input and more native to the convolution structure. Dropout with rate 0.2 is added to prevent overfitting. 'softmax' activation function is used as output layer since multiclass dataset is used. Adam optimizer with static learning rate of 0.001 was used. The CNN model was trained for 20 epochs. Model checkpoint callback was used to save the trained model with the best weight after every epoch. Table 3 shows the parameters for CNN model. The trainable parameters are relatively small as flatten layer is not used.

Table 3. Parameters of Sequential CNN model.

Layer (Type)	Output Shape	Param #
conv2d (Conv2D)	(None, 128, 128, 32)	896
max_pooling2d (MaxPooling2D)	(None, 64, 64, 32)	0
conv2d_1 (Conv2D)	(None, 64, 64, 64)	18496

max_pooling2d_1 (MaxPooling2D)	(None, 32, 32, 64)	0
conv2d_2 (Conv2D)	(None, 32, 32, 128)	73856
global_average_pooling2d (GlobalAveragePooling2D)	(None, 128)	0
dropout (Dropout)	(None, 128)	0
dense (Dense)	(None, 4)	516
Total parameters: 94276		
Trainable parameters: 94020		
Non-trainable parameters: 256		

3 Results

3.1 Generations of Synthetic Scalograms

DDIM model is successfully trained and synthetic samples for each class of bearing fault types are generated. Figure 3 show the real CWRU scalograms while Figure 4 show generated synthetic scalograms which will be used for data augmentation.

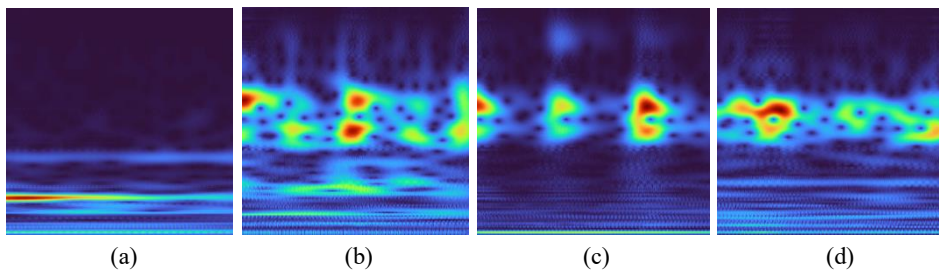


Fig. 3. Scalograms of: (a) real healthy bearing, (b) real inner race fault, (c) real outer race fault, (d) real ball fault.

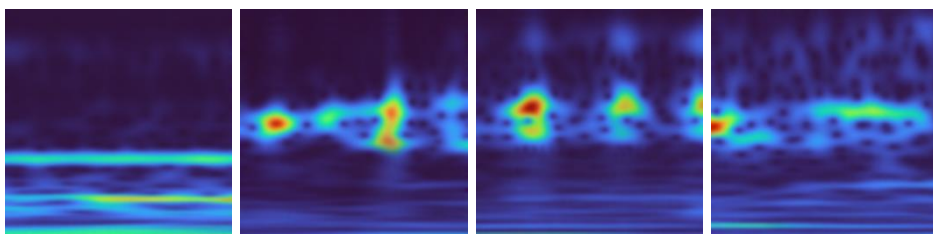


Fig. 4. Scalograms of: (a) synthetic healthy bearing, (b) synthetic inner race fault, (c) synthetic outer race fault, (d) synthetic ball fault.

3.2 Classification Results

Two CNN models are successfully trained with original dataset and augmented dataset respectively. Table 4 shows the testing accuracies of both models. The average testing accuracy of CNN model trained with augmented dataset is 3.3 % higher than the average testing accuracy of CNN model trained with original dataset. Figure 5 shows that there are

increments in testing accuracies for each domain shift by adding synthetic scalograms to the original datasets.

Table 4. Comparison of testing accuracy between original dataset and augmented dataset.

Domain Shift	0→1hp	0→2hp	0→3hp	1→0hp	1→2hp	1→3hp	2→0hp	2→1hp	2→3hp	3→0hp	3→1hp	3→2hp	average
Original Dataset	94.8 %	93.8 %	83.5 %	92.8 %	93.8 %	82.2 %	87.1 %	90.7 %	88.7 %	75.6 %	81.5 %	89.7 %	87.8 %
Augmented Dataset	95.3 %	96.3 %	88.1 %	92.7 %	98.2 %	88.8 %	88.2 %	93.5 %	91.2 %	82.5 %	86.6 %	94.7 %	91.1 %

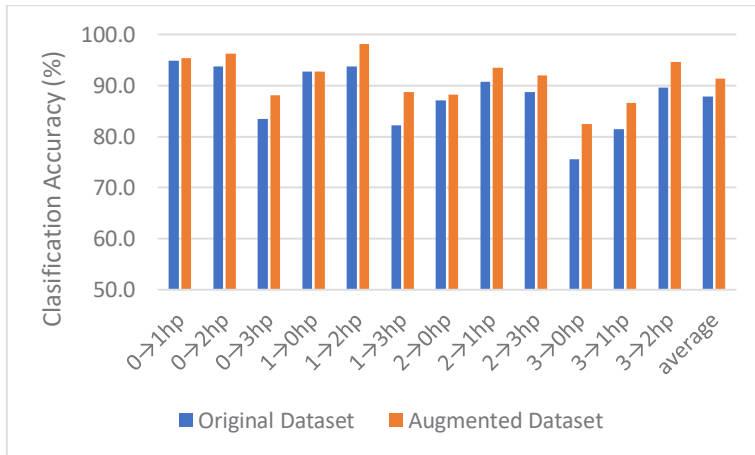


Fig. 5. Comparison of testing accuracy of CNN trained with original dataset and augmented dataset.

4 Conclusion

In this study, DDIM is applied and synthetic scalograms similar to CWRU scalograms can be generated. The difference in testing accuracy of augmented dataset and original dataset shows that synthetic images generated by DDIM could be used to improve performance of CNN model by 3.3% under different working loads.

In the future, conditional generative models should be considered to control type of images generated. Since unconditional DDIM is used in this study, the amount for each class of samples is generated randomly. This would be a problem if large amounts of synthetic samples of a specific class are required. Different evaluation methods for generative model should be considered since existing evaluation methods such as KID and FID evaluate the images in human perspectives which might not works for scalograms.

Transfer learning using generated scalograms also can be applied to an existing model instead of training a new model. Transfer learning is more effective in terms of computational power and time.

Training datasets and testing datasets from different test rigs also could be obtained to verify the effectiveness of DDIM in load domain adaptation. Also, alternative analysis methods such as current analysis, ultrasound analysis and acoustic analysis should be considered since vibration analysis could only detect mechanical faults.

Acknowledgements: The authors would like to extend their greatest gratitude to the Institute of Noise and vibration UTM for funding the current study under the Higher Institution Centre of Excellence (HICoE) Grant Scheme of Advanced Health Monitoring for Turbomachinery

(R.K130000.7843.4J227) and Health Monitoring and Integrity Assessment of Ageing Assets (R.K130000.7843.4J228).

References

- [1] O. Alshorman *et al.*, “A Review of Artificial Intelligence Methods for Condition Monitoring and Fault Diagnosis of Rolling Element Bearings for Induction Motor,” *Shock and Vibration*, vol. **2020**. Hindawi Limited, (2020). doi: 10.1155/2020/8843759.
- [2] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, “Deep Learning Algorithms for Bearing Fault Diagnostics - A Comprehensive Review,” *IEEE Access*, vol. **8**. Institute of Electrical and Electronics Engineers Inc., pp. 29857–29881, (2020). doi: 10.1109/ACCESS.2020.2972859.
- [3] K. P. Soman, S. Sachin Kumar, N. Mohan, and P. Poornachandran, “Modern methods for signal analysis and its applications,” in *Studies in Computational Intelligence*, vol. **823**, Springer Verlag, (2019), pp. 263–290. doi: 10.1007/978-3-030-12500-4_17.
- [4] S. Suh, H. Lee, J. Jo, P. Lukowicz, and Y. O. Lee, “Generative oversampling method for imbalanced data on bearing fault detection and diagnosis,” *Applied Sciences (Switzerland)*, vol. **9**, no. 4, (Feb. 2019), doi: 10.3390/app9040746.
- [5] K. Zhou, E. Diehl, and J. Tang, “Deep convolutional generative adversarial network with semi-supervised learning enabled physics elucidation for extended gear fault diagnosis under data limitations,” *Mech Syst Signal Process*, vol. **185**, (Feb. 2023), doi: 10.1016/j.ymssp.2022.109772.
- [6] L. Tan, T. Huang, J. Liu, Q. Li, and X. Wu, “Deep adversarial learning system for fault diagnosis in fused deposition modeling with imbalanced data,” *Comput Ind Eng*, vol. **176**, (Feb. 2023), doi: 10.1016/j.cie.2022.108887.
- [7] M. Buda, A. Maki, and M. A. Mazurowski, “A systematic study of the class imbalance problem in convolutional neural networks,” *Neural Networks*, vol. **106**, pp. 249–259, (Oct. 2018), doi: 10.1016/j.neunet.2018.07.011.
- [8] W. Jiang, C. Wang, J. Zou, and S. Zhang, “Application of deep learning in fault diagnosis of rotating machinery,” *Processes*, vol. **9**, no. 6, (Jun. 2021), doi: 10.3390/pr9060919.
- [9] F. Zhou, S. Yang, H. Fujita, D. Chen, and C. Wen, “Deep learning fault diagnosis method based on global optimization GAN for unbalanced data ☆,” vol. **187**, p. 104837, (2020), doi: 10.1016/j.knosys.
- [10] J. Luo, L. Zhu, Q. Li, D. Liu, and M. Chen, “Imbalanced fault diagnosis of rotating machinery based on deep generative adversarial networks with gradient penalty,” *Processes*, vol. **9**, no. 10, (Oct. 2021), doi: 10.3390/pr9101751.
- [11] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, “A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy

- environment and different working load,” *Mech Syst Signal Process*, vol. **100**, pp. 439–453, (Feb. 2018), doi: 10.1016/j.ymsp.2017.06.022.
- [12] S. Liu, H. Jiang, Y. Wang, K. Zhu, and C. Liu, “A deep feature alignment adaptation network for rolling bearing intelligent fault diagnosis,” *Advanced Engineering Informatics*, vol. **52**, (Apr. 2022), doi: 10.1016/j.aei.2022.101598.
- [13] J. Song, C. Meng, and S. Ermon, “Denoising Diffusion Implicit Models,” (Oct. 2020), [Online]. Available: <http://arxiv.org/abs/2010.02502>