Deep Learning based Time-Frequency Image Enhancement Method for Machinery Health Monitoring

Madhurjya D. Choudhury, *Member, IEEE*, Kelly Blincoe, *Member, IEEE*, and Jaspreet S. Dhupia, *Senior Member, IEEE*

Abstract— Reliable machinery health monitoring using measured vibration signals requires a good readability of time-frequency (TF) images. However, conventional TF methods suffer from a limited time–frequency resolution and cross-term interferences, which limit their practical applicability in health monitoring. To address this issue, a TF image improvement method using deep learning is proposed in this paper. The proposed method employs a deep learning technique known as conditional generative adversarial network (cGAN) to convert a noisy low-resolution TF image of a bearing vibration signal into a noise-free high-resolution image such that the true frequency characteristics of measured signals may be revealed. In this paper, the cGAN model is trained using a simulation-based dataset generated from a bearing analytical model. The trained cGAN model is then utilized to improve TF images generated from real bearings under different fault and operating conditions. The results reveal that the proposed image improvement method generates high-resolution TF representations which are better than both the traditional TF images and those generated using TF reassignment methods.

I. INTRODUCTION

Time-frequency analysis of machine vibration signals reveal valuable health information that can be utilised for effective health monitoring. Generally, the measured vibration signal is transformed to a TF representation that can reveal the time-varying frequency characteristics of the signal. The classical tool for TF representation is the Short Time Fourier Transform (STFT) [1]. However, STFT is subject to the Heisenberg uncertainty principle and suffers from limited time-frequency resolution [1]. As a remedy to the time-frequency resolution problem, the reassignment method was developed, which obtains a better TF localization but it still deviates substantially from the ideal noise free TF image [1, 2].

Recently, deep learning has been applied to achieve improvement in image resolution based on supervised training schemes [3]. Convolutional auto-encoders and Generative Adversarial Networks (GANs) have been successfully used in image prediction and generation applications [4, 5]. Especially, a variant of GAN known as the conditional GAN (cGAN) has shown outstanding results in various tasks such as image generation and translation, super-resolution imaging, and face image synthesis. These advances in deep learning techniques provide an opportunity for its use

in improving machine vibration TF images, which can then be effectively used for health monitoring. However, as already established, deep learning models require a large volume of training datasets to minimize their variance during model generalization [5, 6]. These learning models also require knowledge of the ideal characteristics of the images for effective translation of the trained knowledge.

Over the years, researchers have observed that it is relatively expensive to generate such large and diverse data in industrial fault diagnosis and health monitoring applications, and the ideal signal frequency characteristics may be rarely known beforehand for training. Several attempts have been initiated to address the issue with data augmentation methods such as using a deep generative adversarial network approach [6, 7]. However, the dependency on quality data is still an ongoing issue. Simulation-based data generation has been adopted recently as a solution to supplement the datasets for deep learning model training, gaining much interest in different fields [6, 8]. Several studies have been conducted to integrate sophisticated simulation models with deep neural networks to create sufficient variations for training in machine fault diagnosis and health monitoring research [7-12]. Gryllias et al. [9] proposed a simulation-driven bearing fault classification method using support vector machine. In this study, simulated-based model training is utilised to improve TF images of measured vibration signals. The use of simulated data provides an opportunity to fulfil the requirement of ideal image generation, as the accurate simulated frequencies used to generate the training images may also be utilised to generate the ideal TF images required for image-to-image translation in cGAN.

In this study, the issue of low-resolution TF image is solved using a deep learning based framework, where a high-resolution TF representation of a measured bearing vibration signal is obtained by adopting a conditional implementation of GAN called *Pix2Pix*, developed for general image-to-image transformation applications [4]. The training of the deep learning model is based on simulated data obtained from a characteristic bearing fault analytical model under different machine frequency scenarios. The aim of this study is to demonstrate that GAN based models trained on simulated images can be utilized for improving TF image obtained from

Madhurjya D. Choudhury is with the School of Engineering and Computer Science, Victoria University of Wellington, Wellington 6012, New Zealand (email: madhurjyadev.choudhury@vuw.ac.nz).

Kelly Blincoe is with the Department of Electrical, Computer, and Software Engineering, University of Auckland, Auckland 1010, New Zealand (e-mail: k.blincoe@auckland.ac.nz).

Jaspreet S. Dhupia is with the Department of Mechanical Engineering and Mechatronics Engineering, University of Auckland, Auckland 1010, New Zealand (e-mail: j.dhupia@auckland.ac.nz).

real machine vibration signals. The improved TF images can then be used as an evidence demonstrating the existence of faults for effective machinery health monitoring.

The remainder of this article is organized as follows. Section II summarizes the basic theory of the bearing signal model, TF representation, and GANs. Section III describes the proposed TF image improvement method. Section IV presents the validation of the proposed method using measurements from a bearing test-rig. Finally, Section V concludes the presented work.

II. BACKGROUND

A. Problem Statement

STFT cannot achieve proper time localization and frequency localization simultaneously due to the Heisenberg uncertainty principle leading to low resolution issues. Reassignment offers better resolution but suffers from noise interference. Figure 1(a) shows the STFT of a sinusoidal signal and Fig. 1(b) shows the corresponding reassigned time-frequency plot. It can be observed that due to better localization using reassignment, Fig. 1(b) achieves better resolution as compared to the STFT based TF image of Fig. 1(a). However, even after the reassignment the resulting TF image deviates substantially from the ideal noise free TF image shown in Fig. 1(c). This can lead to misleading decisions, especially in case of machine health monitoring, which is heavily dependent on understanding the accurate time-varying frequency characteristics of a measured signal.
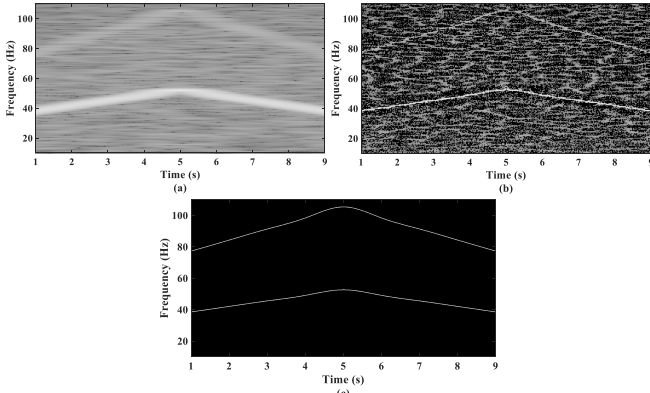

Fig. 1. (a) STFT based TF image of a simulated sinusoidal signal. (b) Reassigned STFT based TF of the simulated sinusoidal signal. (c) Ideal TF image of the simulated sinusoidal signal.

B. Bearing Analytical Model

In this paper, the vibration response of a defective bearing operating under time-varying speed is generated as a chain of impulse responses of a one degree-of-freedom mass-spring damper system [13, 14],

$$x(t) = A(t) \sum_{k=1}^{K} L_k e^{-\beta_d(t-t_k)} \sin[2\pi f_r(t-t_k)] u(t-t_k) + n(t), \tag{1}$$

where $n(t)$ is measurement noise, $A(t) = 1 + \chi \cos\left(\int_0^t f_{sh}(t)\mathrm{d}t\right)$ is the amplitude modulation term caused by the shaft rotational frequency variations ($\chi > 0$ implies inner race faults and $\chi = 0$ implies outer race faults) and $f_{sh}(t)$ is the time-varying shaft frequency, which is also the modulating frequency of the signal. $L_i = L_0 + \zeta f_{sh}(t_i)$, L_0 and ζ are constant terms, t_k is the time of occurrence of the kth impulse, K is the number of impulse responses, β_d is the value of the damping characteristics, f_r is the resonating frequency of the bearing system, and $u(t)$ is the unit step function. The time of occurrence of the impulses t_k can be calculated as,

$$\begin{cases} t_1 = (1+\xi)\left[\frac{1}{f_f(t_0)}\right] \\ t_k = (1+\xi)\left[\frac{1}{f_f(t_0)} + \cdots + \frac{1}{f_f(t_{k-1})}\right], i = 2, \cdots K \end{cases} \tag{2}$$

where ξ is the slip coefficient having typical value between 0.01 and 0.02 [15], $f_f(t)$ is the fault characteristic frequency and t_0 is the time when the first impulse occurs. With specific bearing parameters and operating conditions, the detailed numerical implementation of the analytical model will be demonstrated in Section III.

C. Time-frequency Representation

Time-frequency analysis is a powerful tool to understand the complex time-varying characteristics of machine vibration signals. In order to recognize the intricate time-variant sidebands and the frequency characteristics under variable speeds, the quality of the time-frequency analysis method is a necessity. Short-time Fourier transform (STFT) based spectrograms can be employed to achieve time–frequency (TF) imaging. For a signal $x(t)$, its short-time Fourier transform is defined as [16],

$$S(t,f) = \int x(\tau)h^*(t-\tau)e^{-j2\pi f\tau}d\tau, \tag{3}$$

where $h(t)$ is a window function of unit energy and * denotes conjugate operation. A spectrogram is defined as the square modulus of the STFT, i.e., $|S(t,f)|^2$.

It is, however, worthwhile to mention that while a spectrogram has low computational complexity, but its time-frequency resolution is limited due to the Heisenberg uncertainty constraint [1, 16]. Depending on the length of the window function, the spectrogram can attenuate the cross-terms between frequencies, however, there is a trade-off between the cross-term suppression and the TF resolution. Therefore, in order to achieve successful fault diagnosis and conduct efficient health monitoring using measured vibration signals, there is a need to generate high-resolution time-frequency images that can correctly reflect the underlying frequency characteristics of the measured signal. In this paper, we have tackled the issue of generating high-resolution TF images using a generative adversarial network trained on a simulation-based dataset to carry out image translation.

D. Generative Adversarial Network (GAN)

A generative adversarial network is a class of deep learning frameworks designed by Goodfellow et al. [4], where two neural networks, namely the generator (G) and the discriminator (D), contest with each other in a game of zero-sum, where one network's gain is another network's loss. Given a training set, a GAN learns to generate new data with the same statistics as the training set. However, GAN generates images based on random noise z, so the generated output image is not controllable. In order to tackle this issue a conditional implementation of GAN known as the cGAN was proposed [4]. In cGAN an image is generated by random noise z and an input conditional variable. In this paper, we have utilized a general algorithm based on cGAN named *Pix2Pix* [4]. In the *Pix2Pix* model, shown in Fig. 1, the generator network G uses U-Net, a network structure that is widely used in the field of image segmentation to fully integrate features and improve details. The discriminator network D utilizes a PatchGAN to output a predicted probability value for each area or patch of the input image, which is equivalent to performing a simple task of classifying whether the input is true or false. The overall objective function of the *Pix2Pix* is,

$$G^* = \arg \min_{G} \max_{D} L_{cGAN}(G, D) + \lambda L_{L1}(G), \qquad (4)$$

where $L_{cGAN}(G, D)$ is the objective of the condition GAN, $L_{L1}(G)$ is the L1 loss of the generator network and λ is the L1 loss regularization parameter.

Therefore, TF images of machine vibration signals can be fed as the input vector to the generator (G) that generates an initial estimate of the input. The discriminator (D) then tries to classify the generated image as a fake or a real image based on the information of the corresponding ideal TF image (high resolution and without noise) fed to it as the ground truth. The architecture of the proposed image enhancement method discussed in this paper is shown in detail in Fig. 2.
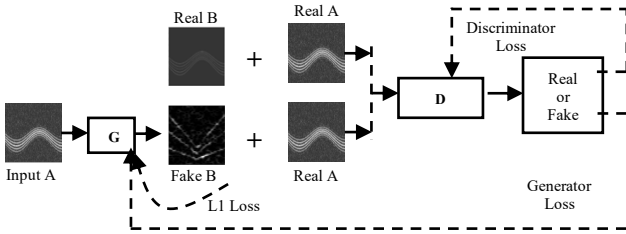

Fig. 2. Architecture of *Pix2Pix* for image improvement.

III. PROPOSED METHODOLOGY

In this paper, the cGAN known as *Pix2Pix* is used to improve the resolution of a blurry and noisy TF image in order to reveal accurate instantaneous frequency contents. The proposed idea is to convert bearing vibration data into high-resolution noise-free STFT based spectrogram images, which contain the constituent frequencies that are indicators of the bearing health status.

A. Source Domain Image Generation

The source domain images for training the *Pix2Pix* model were generated using a combination of simulated signals and real-world bearing vibration signals. The simulated signals were generated according to the bearing analytical model of Eq. (1). The values of the various parameters of Eq. (1) used in the simulation are given in Table I.

TABLE I.   VALUES OF PARAMETERS USED IN THE SIMULATION.

Resonance frequency, f_r	Damping, β_d	Slippage coefficient, ξ	Constant terms, L_0 and ζ
2000 Hz	250	0.01	1 and 0.5

Different fault scenarios under various speed variations were simulated to generate the source domain TF images. However, one of the challenges during source domain image generation for training the *Pix2Pix* model was to make the simulated faults adaptable to the given target domain images generated from real fault samples [6, 8]. In order to tackle this issue we utilised some real healthy data samples of the target domain. The noise term in the analytical model of Eq. (1) was simulated by adding these real healthy samples as base signals. It has been observed in previous studies that as the base signal encodes information about the operating and environmental conditions of the bearing under observation, the generated simulated training data becomes more adaptable to the target domain [6, 8]. It is to be mentioned that to generate the source domain training images an assumption was made that the early measurements of the bearing, when the faults have not yet emerged, is available. This is generally a practical assumption as most modern health monitoring systems have access to measurements when a machine is considered to be operating in a healthy condition. In order to test the effectiveness of the proposed method, a training set of 300 simulated signals and corresponding spectrograms $TF_r(t,f)$ and ideal spectrograms $TF_i(t,f)$ were generated. The source domain signals for training were simulated using the bearing vibration model of Eq. (1). Different fault situations were simulated to generate the training images under various speed variation conditions and noise levels. The time-domain signals have 1024 samples. The overall framework of the proposed methodology is presented in Fig. 3.
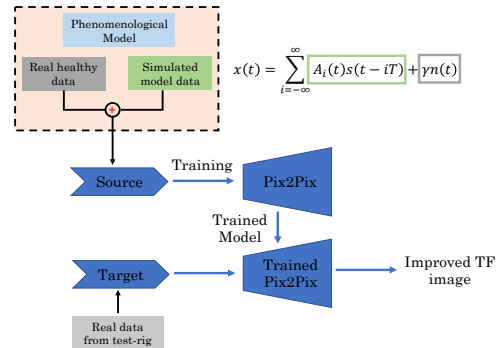

Fig. 3. Methodology of the *Pix2Pix* based TF image enhancement for bearing health monitoring.

B. *Model Training*

The *Pix2Pix* network was implemented to include 8 generator hidden layers and 4 discriminator hidden layers. The L1 regularization parameter was set to $\lambda = 100$ as proposed in [4]. Similarly, the Adam optimizer was selected with learning rates $\alpha = 0.0001$, $\beta_1 = 0.5$, and $\beta_2 = 0.999$. The training was carried out for 100 epochs. The TF images were converted to a resolution of 256×256 pixels before being fed to the model. The structure of the *Pix2Pix* model used in this study is shown in Fig. 4. The detailed architecture of *Pix2Pix* is provided in [4].

Fig. 4. Structure of the proposed model for image improvement.

IV. EXPERIMENTAL VALIDATION

This section utilises the experimental data provided by the University of Ottawa [17] to validate the effectiveness of the proposed TF image improvement method under different bearing faults in time-varying speed conditions. A Spectra Quest machine fault simulator (MFS-PK5M) was used to collect the experimental measurements. The test-rig consists of a motor that drives a shaft, and an AC drive for rotational speed adjustment, as shown in Fig. 5. The driven shaft was supported by two ball bearings with a healthy bearing placed as the left-end support and the test bearing was placed as the right-end support of the shaft. The right-end support was replaced by bearings of different fault conditions in order to simulate an inner race defect and an outer race defect.
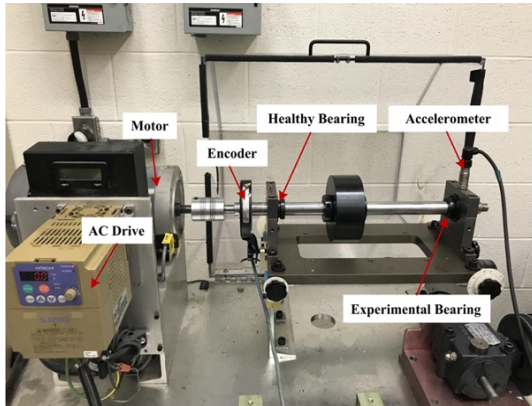

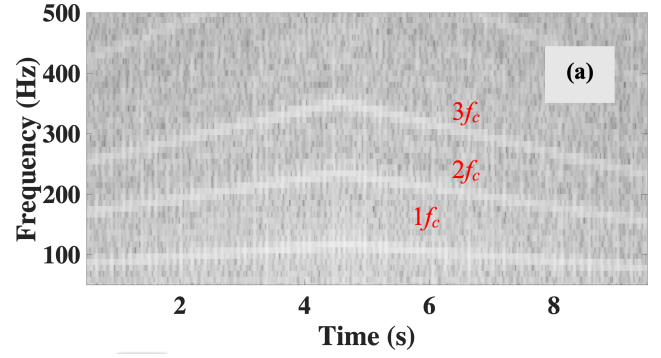**Fig 5.** Experimental set-up used to collect the bearing dataset [19].

Fig. 6. STFT based TF image of bearing inner race fault.

An accelerometer mounted on the test bearing housing was used to measure the vibration signals, whereas an incremental encoder measured the shaft rotational speed. The accelerometer signal and encoder signal were collected under various time-varying speed conditions. A detailed account of the data measurement steps is provided in [17]. The signals collected for inner race defect under increasing-decreasing speed and increasing only speed conditions and signals collected for outer race defect under increasing-decreasing speed and decreasing-increasing speed conditions are considered in this paper. Every signal file contained two channels, one having the accelerometer data and the other having the encoder data, respectively. All the signals were collected at a sampling frequency of 200 kHz for 10 seconds. The theoretical fault frequency for a bearing was calculated as FCC*f_{sh}, where FCC is the fault characteristic coefficient. For this test-rig, the FCC due to an inner race fault is 5.43 and an outer race fault is 3.57 [14]. Figure 6 shows the STFT of the bearing with inner race defect where the first three harmonics of the fault characteristic frequency is within a frequency range of 80Hz - 353 Hz as the shaft speed increases from 14.8 Hz to 21.7 Hz and then decreases to 13.6 Hz.. The STFT based spectrogram of the envelop signals are found to be of low resolution and the fault induced frequencies are not clearly visible in these plots. Therefore, for accurate estimation of the envelop signal characteristics, the generated spectrograms of the original envelop signals were fed to the trained *Pix2Pix* model. Figure 7(a) and Fig. 8(a) show the spectrograms of the measured envelop signals of the test bearing under inner race defect under increasing-decreasing speed and increasing only speed conditions. The improved spectrograms after applying the image improvement algorithm were found to clearly reveal the constituent frequencies of the envelop signals, as shown in Fig. 7(b) and Fig. 8(b). The frequencies observed in Fig 7(b) and Fig 8(b) correspond to the theoretical fault characteristic frequency (f_c) and its harmonics due to inner race fault of the bearing under consideration. The improved spectrograms were also found to exhibit better and clearer TF images in comparison to the reassigned spectrograms of Fig. 7(c) and Fig. 8(c). The reassignment procedure is known to achieve better localization than STFT.
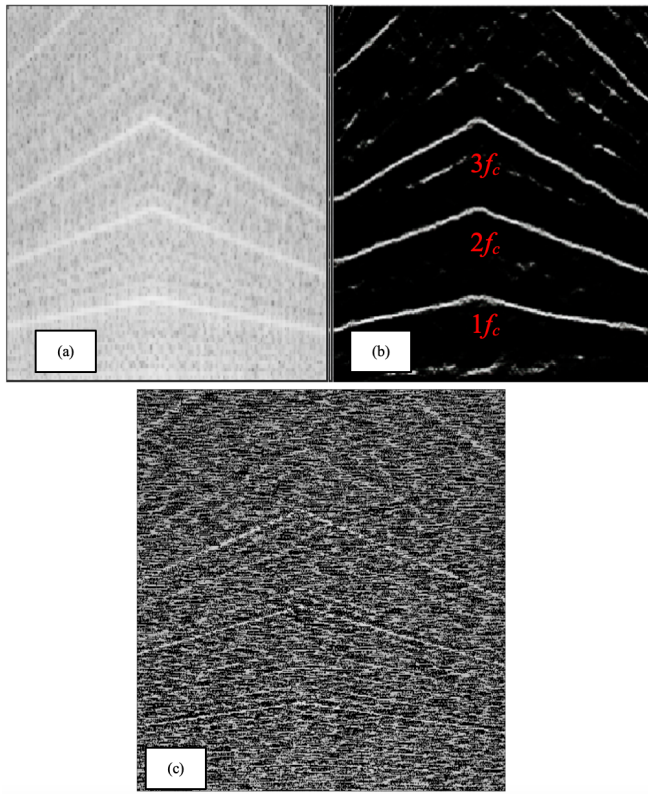
Fig. 7. Inner race fault observed under increasing-decreasing speed condition. (a) Original spectrogram, (b) cGAN generated improved spectrogram, and (c) Reassigned spectrogram.
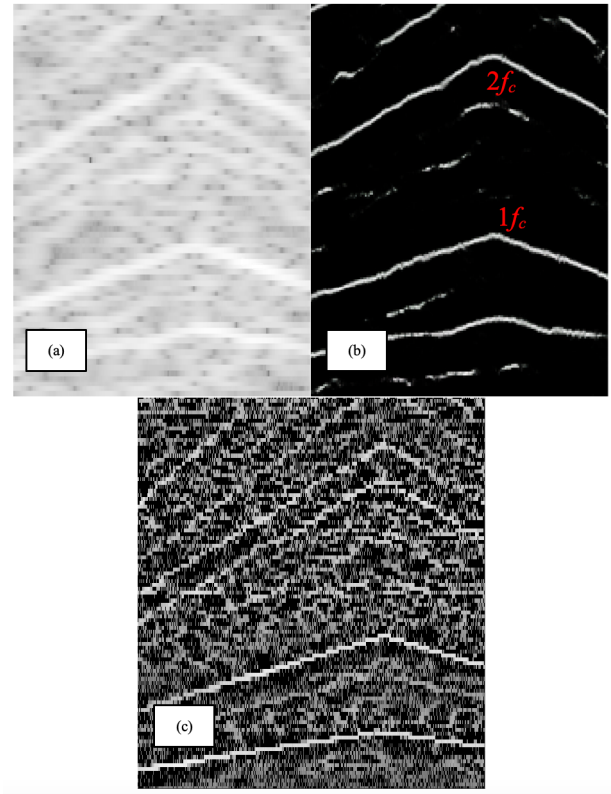

Fig. 8. Inner race fault observed under increasing only speed condition. (a) Original spectrogram, (b) Improved spectrogram, and (c) Reassigned spectrogram.

Fig. 9. Outer race fault observed under increasing-decreasing speed condition. (a) Original spectrogram, (b) cGAN generated improved spectrogram, and (c) Reassigned spectrogram.
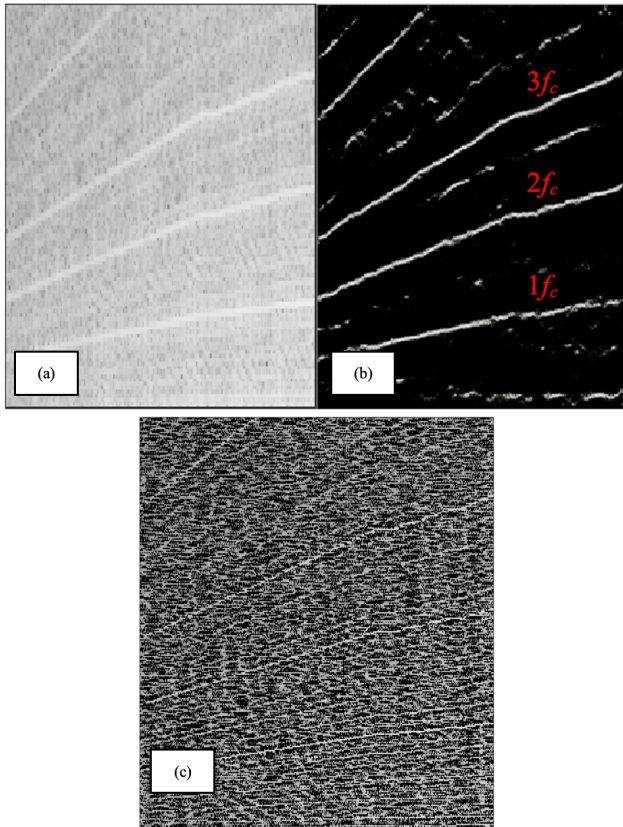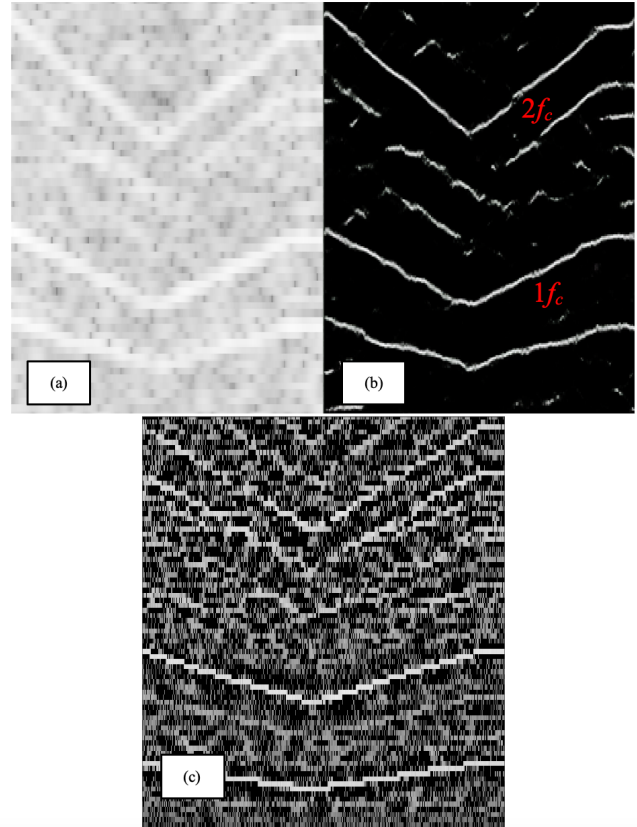

Fig. 10. Outer race fault observed under decreasing-increasing speed condition. (a) Original spectrogram, (b) cGAN generated improved spectrogram, and (c) Reassigned spectrogram.

However, it unavoidably introduces noise to the eventual TF result [18], which is not observed in the proposed cGAN based method. The improved TF images for outer race fault under increasing-decreasing speed and decreasing-increasing speed conditions are shown in Fig. 9(b) and Fig. 10(b). The corresponding STFT-based spectrograms are shown in Fig. 9(a) and Fig. 10(a) and the reassigned spectrograms are shown in Fig. 9(c) and Fig. 10(c). Thus, the TF images obtained using the proposed conditional GAN based image improvement strategy show promise in monitoring the presence of faults in bearings operating under variable speed conditions.

V. CONCLUSION

This paper proposes a conditional generative adversarial network (cGAN) based time-frequency (TF) image improvement method for bearing health monitoring. The proposed method leverages a bearing phenomenological model to simulate spectrogram images that are utilised to train a cGAN variant called *Pix2Pix* developed for image-to-image translation. The results show that a *Pix2Pix* model trained on a simulation-based bearing vibration TF image dataset can reliably denoise and improve the resolution of TF images generated from real bearing vibration signals. The proposed method was validated using experimental datasets from a mechanical fault simulator test-rig. Compared to the conventional TF images like spectrograms and reassigned spectrograms, the images generated using the proposed method significantly reduced noise and improved the readability of the bearing vibration TF images to pinpoint the frequency contents and time variability of such signals necessary for reliable health monitoring. The experimental results indicate that the cGAN based method is able to generate physically meaningful TF images. The proposed method was found to achieve satisfactory results using simulation-based model training and availability of only a small amount of real data, which reinforces the method's applicability in situations where collecting large datasets for model training is expensive.

ACKNOWLEDGMENT

VI. REFERENCES

[1] Feng, Z., et al., *Generalized adaptive mode decomposition for nonstationary signal analysis of rotating machinery: Principle and applications.* Mechanical Systems and Signal Processing, 2020. **136**.

[2] Auger, F., and P. Flandrin, *Improving the readability of time–frequency and time–scale representations by the reassignment method.* IEEE Transactions on Signal Processing, 1995. **43** (5), pp. 1068-1089.

[3] Liang, P., et al., *Intelligent Fault Diagnosis via Semisupervised Generative Adversarial Nets and Wavelet Transform.* IEEE Transactions on Instrumentation and Measurement. 2020. **69** (7).

[4] Isola, P., et al., *Image-to-Image Translation with Conditional Adversarial Networks*, in *CVPR*. 2017.

[5] Shao, S., et al., *Generative adversarial networks for data augmentation in machine fault diagnosis.* Computers in Industry. 2019. **106**, pp. 85-93.

[6] Liu, C., and K. Gryllias, *Simulation-driven domain adaptation for rolling element bearing fault diagnosis.* IEEE Transactions on Industrial Informatics. 2021 (in press).

[7] Zhang, W., et al., *Machinery fault diagnosis with imbalanced data using deep generative adversarial networks.* Measurement. 2020. **152**, 107377.

[8] Wang, Q., et al., *Integrating expert knowledge with domain adaptation for unsupervised fault diagnosis.* IEEE Transactions on Instrumentation and Measurement. 2022. **71**.

[9] Gryllias, K., and I. Antoniadis, *A support vector machine approach based on physical model training for rolling element bearing fault detection in industrial environments.* Engineering Applications of Artificial Intelligence. 2012. **25**(2), pp. 326–344.

[10] Sobie, C., et al., *Simulation-driven machine learning: Bearing fault classification.* Mechanical Systems and Signal Processing. 2018. **99**, pp. 403–419.

[11] Gao, Y., et al., *FEM simulation-based generative adversarial networks to detect bearing faults.* IEEE Transactions on Industrial Informatics. 2020. **16**(7), pp. 4961–4971.

[12] Booyse, W., et al., *Deep digital twins for detection, diagnostics and prognostics.* Mechanical Systems and Signal Processing. 2020. **140**.

[13] Antoni, J., and R.B. Randall, *A Stochastic Model for Simulation and Diagnostics of Rolling Element Bearings with Localized Faults.* Journal of Vibration and Acoustics, 2003. 125(3): p. 282-289.

[14] Choudhury, M.D., et al., *A novel tacholess order analysis method for bearings operating under time-varying speed conditions.* Measurement. 2021. **186**.

[15] Randall, R.B., *Vibration-based Condition Monitoring: Industrial, Aerospace and Automotive Applications.* 2011: John Wiley & Sons, Ltd.

[16] Feng, Z., et al., *Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples.* Mechanical Systems and Signal Processing. 2013. **38** (1), pp. 165-205.

[17] Huang, H., and N. Baddour, *Bearing vibration data collected under time-varying rotational speed conditions.* Data in Brief. 2018. **21**, pp. 1745–1749.

[18] Ahrabian, A., and Mandic, D. P., *Selective Time-Frequency Reassignment Based on Synchrosqueezing.* IEEE Signal Processing Letters. 2015. **22**(11), pp. 2039-2043.