Health-based Fault Generative Adversarial Network for Fault Diagnosis in Machine Tools

Chia-Yu Lin, Tzu-Ting Chen, Li-Chun Wang, Hong-Han Shuai
Department of Electrical and Computer Engineering, National Chiao Tung University, Taiwan
sallylin1987@nctu.edu.tw,{is2586537.eed01,lichun,hhshuai}@g2.nctu.edu.tw

Abstract

Neural network (NN)-based anomaly-detection model has achieved unprecedented performance in internet of things such as fault diagnosis for machine tools. To train a promising neural network-based fault diagnosis model, sufficient fault data is necessary. In reality, machines are rarely operating in fault and cause an unbalance between health data and fault data. Various Generative adversarial networks (GANs) such as auxiliary classifier GAN (ACGAN) were proposed to generate artificial data that are similar to real data for data augmentation. However, applying the current GAN in fault diagnosis model, the fake data makes decision boundary unstable and results in low classification accuracy. Besides, a certain amount of fault data is required and only one-to-one domain generation can be achieved, rather than one-to-many. In this paper, we propose an end-to-end fault diagnosis framework with health-based fault generative adversarial network (HFGAN) to solve data imbalance issue, augmentation uncertainty issue and improve fault diagnosis for machine tools. Based on the proposed HFGAN, various high quality target domain fault data is generated based on health data and one-to-many data generation is achieved. From the experiments, mixing the generated fake data with 7.5% real data as the input of the fault diagnosis model, the classification accuracy achieves 99.8%. Moreover, the cost of data collection in factories can be reduced since only 7.5% of real machine data is required for fault diagnosis.

Introduction

With the advance of deep learning and internet of things technologies, data-driven prognostics and health management (PHM) for machine tools have become promising, which not only reduces the maintenance cost but also enhances the yield rates (Lin et al. 2019). Take high-speed rolling components as an example. High-speed rolling components play an important role in many fields of machine tool industries, such as bearings in a drilling machine, turbines of a power generator, or gears inside an engine (Hasani, Wang, and Grosu 2019). When some rolling components are broken, an ill-functioned machine will be shut down. The repairing cost and the downtime results in lower productivity and increasing productivity cost. Furthermore, the deficient products will be dumped. A fault diagnosis (FD) model to notify fault in advance is necessary.

Building a reliable FD model requires a large amount of fault data for training. However, training data among different machine states is usually unbalanced. For example, a mechanical machine works under health condition most of the time and generates a large amount of health data, and fault data is generated until it crashes. Compared to health data, fault data in training dataset is relatively rare, which is a "data imbalance" problem. Suffered from unbalanced training data, it is relatively difficult to train a deep FD model to achieve accurate prediction of machine conditions.

Generative adversarial network (GAN) was proposed to generate realistic looking image data to solve data insufficient issue (Goodfellow et al. 2014). Since the original GAN had problems with unstable training, variant GAN such as deep convolutional generative adversarial networks (DCGANs), Wasserstein GAN (WGAN) (Arjovsky, Chintala, and Bottou 2017) (Radford, Metz, and Chintala 2015) and semi-supervised learning using GAN (Odena 2016) were developed to improve the quality of generated images. Afterwards, auxiliary classifier GAN (ACGAN) (Odena, Olah, and Shlens 2017) were designed to generate high resolution images and improve the performance in classification tasks. Since limited GAN has been utilized in generating raw sensor data, (Shao and Yan 2019) was the first attempt to generate mechanical sensor signals with subsequent fault classification based on ACGAN architecture.

However, ACGAN has augmentation uncertainty problem, i.e., the augmented data may change the classification decision boundary and result in low accuracy. Besides, (Shao and Yan 2019) requires about 0.8 million data for training data generation model. The insufficient fault data problem is not really solved since certain amount of fault data is necessary before generating data. In addition, current GAN and their variants for smart manufacturing only apply one-to-one domain data generation instead of one-to-many. One type of fault data can not generate more than two other types of fault data at the same time.

In this paper, we propose an end-to-end fault diagnosis framework with data preprocessing module, health-based data generation module and convolutaional neural network (CNN)-based fault diagnosis module to solve data imbalance issue and diagnose faults for machine tools. In the data generation module, we modify ACGAN to design a health-based fault generative adversarial network (HFGAN) by utilizing abundant amount of health data as input instead of random noise. We also modify the loss functions of discriminator and generator to solve augmentation uncertainty issue. The loss function of discriminator in ACGAN considers the classification performance of real data and fake data. Before the generator is well-trained, the fake data is not similar to real data and may change the classification boundary, which results in low accuracy. Therefore, we modify loss functions to make discriminator focus on learning classification boundary based on real data and generator pay attention to generate fake data. We also add the content loss to keep more content of real data. As a result, HFGAN can generate two types of fault data that similar real data based on health data and achieve one-to-many domain generation.

The main contributions of this paper are summarized as follows:

- An end-to-end fault diagnosis framework with data preprocessing module, health-based data generation module and CNN-based fault diagnosis module was proposed to solve data imbalance issue and intensively increase the accuracy of fault diagnosis.
- The proposed HFGAN solves the augmentation uncertainty problem in ACGAN and generates high quality fake data by modifying the loss functions of generator and discriminator and adding content loss.
- Experiments show that mixing the generated fake data with 7.5% real data as the input of the fault diagnosis model, the classification accuracy achieves about 99.8%.
- The proposed HFGAN reduces the cost of data collection in factories since HFGAN only needs 7.5% of real machine data to generate data.

Related Work

Generative adversarial network (GAN) was proposed in (Goodfellow et al. 2014) to solve data insufficient issue by generating artificial images, but the primitive GAN was extremely difficult to train and was unstable. (Radford, Metz, and Chintala 2015) developed deep convolutional generative adversarial networks (DCGANs) to provide a more stable set of architectures for training GAN and learn good representations of images for unsupervised learning and generative modeling. (Arjovsky, Chintala, and Bottou 2017; Gulrajani et al. 2017) designed a new loss function based on Wasserstein distance to solve the collapse mode problem and increase the model stability. To improve the quality of generated images, (Odena 2016) demonstrated the semi-supervised learning using GAN by producing class labels in discriminator network. (Odena, Olah, and Shlens 2017) proposed auxiliary classifier GAN (ACGAN) to generate high resolution images and achieve accurate classification tasks by adding labeling information.

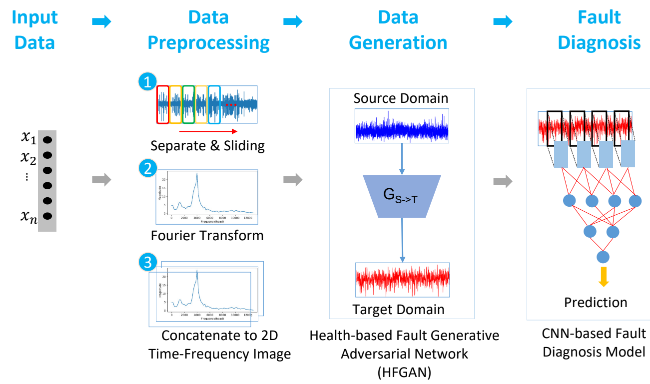


Figure 1: End-to-end data generation and fault diagnosis framework.

Comparing to conventional GANs, ACGAN can generate more discriminant samples by providing label information. The objective function of ACGAN is to maximize log-likelihood of correctly assigning to the source and class they belong to, which can be represented by generator loss L_G and discriminator loss L_D:

$$\begin{aligned} L_D =&E_{y\sim p^y_{data}, c\sim p_c}[\log P(C=c|y)]+ \\ &E_{z\sim p_z, c\sim p_c}[\log P(C=c|G(z,c))]+ \\ &E_{y\sim p^y_{data}}[\log P(S=real|y)]+ \\ &E_{z\sim p_z, c\sim p_c}[\log P(S=fake|G(z,c)]. \end{aligned} \tag{1}$$

$$\begin{aligned} L_G =&E_{y\sim p^y_{data}, c\sim p_c}[\log P(C=c|y)]+ \\ &E_{z\sim p_z, c\sim p_c}[\log P(C=c|G(z,c))]- \\ &E_{y\sim p^y_{data}}[\log P(S=real|y)]- \\ &E_{z\sim p_z, c\sim p_c}[\log P(S=fake|G(z,c)]. \end{aligned} \tag{2}$$

In L_G and L_D, the random noise $z \sim p_z$ and class $c \sim p_c$ is the input of generator G. The output of G is generated samples $G(z,c)$. S denotes the source of input data and y denotes the real fault data from dataset p^y_{data}.

GAN and its variant have been proved be a useful tool for data augmentation in various image generation tasks but limited GAN has been adopted in raw sensor data generation. (Shao and Yan 2019) utilized ACGAN to generate mechanical sensor signals with subsequent fault classification. Authors built the block of generator and discriminator based on one-dimensional CNN to enhance the ability of capturing representations from raw data. Besides, class labels are added in both generator and discriminator to accelerate model training. However, the augmentation uncertainty problem still existed in ACGAN and cannot achieve one-to-many domain generation. Besides, the model training in (Shao and Yan 2019) still required a certain amount of fault data. Data imbalanced problem in machine tools has not completely solved. Therefore, we propose a health-based fault generative adversarial network based on ACGAN to utilize the abundant health data and generate various types of fault data at the same time.

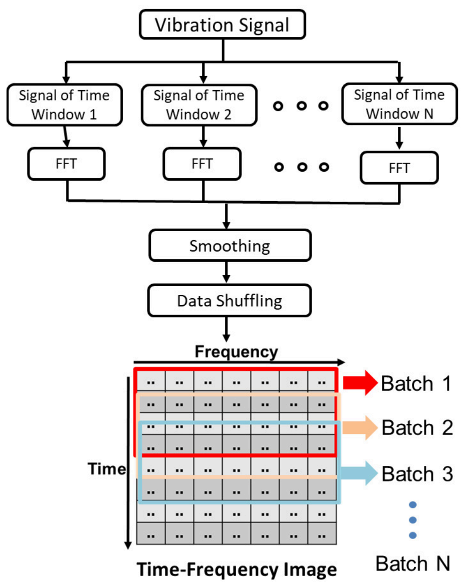


Figure 2: Data preprocessing flow.

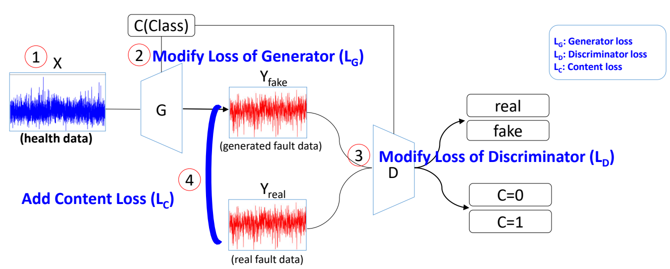


Figure 3: HFGAN.

End-to-end Fault Diagnosis framework

Figure 1 is the proposed end-to-end fault diagnosis framework with health-based fault generative adversarial network (HFGAN). There are three modules: data preprocessing module, health-based data generation module, and CNN-based fault diagnosis model.

Data Preprocessing Module

The input of the proposed framework is vibration data. The data preprocessing for input data is shown in Figure 2. First, we divide vibration data into consecutive non-overlapping windows. Since vibration data in frequency domain are more distinct than in time domain, we execute Fast Fourier Transform (FFT) in every window. To filter the noise and increase the differentiation of peak value, the moving average filter is adopted. After the transformation, we concatenate the window data to time-frequency image to retain both spectrum and time-series information. Therefore, we can utilize CNN to increase the performance of classification.

Health-based Data Generation Module

Taking advantage of plenty of health data in factories, we modify the input and the loss functions of ACGAN (Odena, Olah, and Shlens 2017) to propose a health-based fault GAN (HFGAN), as shown in Figure 3. The input of HFGAN is changed from random noise z in ACGAN to health data x since health data and fault data are all vibration data. Besides, the loss of generator (L_G) and the loss of discriminator (L_D) of ACGAN are modified in HFGAN. The discriminator in ACGAN considers the performance of classifying real data and fake data into different fault conditions. The fake data generated by the model, which is not well-trained, becomes noise and makes classification boundary as well as discriminator unstable. Thus, we remove $E_{x\sim p^x_{data}, c\sim p_c}[\log P(C = c|G(x,c))]$ in discriminator to focus on the classification boundary based on real data, as shown in Eq. (3). Since the discriminator has concentrated on the classification loss of real data, the generator only needs to care the performance of fake data. We remove $E_{y\sim p^y_{data}, c\sim p_c}[\log P(C = c|y)]$ and $E_{y\sim p^y_{data}}[\log P(S = real|y)]$ in generator, as shown in Eq. (4).

$$\begin{aligned} L_D =& E_{y\sim p^y_{data}, c\sim p_c}[\log P(C = c|y)] + \\ & E_{y\sim p^y_{data}}[\log P(S = real|y)] + \\ & E_{x\sim p^x_{data}, c\sim p_c}[\log P(S = fake|G(x,c)]. \end{aligned} \tag{3}$$

$$\begin{aligned} L_G =& E_{x\sim p^x_{data}, c\sim p_c}[\log P(C = c|G(x,c))] - \\ & E_{x\sim p^x_{data}, c\sim p_c}[\log P(S = fake|G(x,c)]. \end{aligned} \tag{4}$$

Moreover, we add the content loss L_C in generator to keep as more as the content of real data and improve the quality of generated data. The generator loss L_G (4) becomes:

$$\begin{aligned} L_G =& E_{x\sim p^x_{data}, c\sim p_c}[\log P(C = c|G(x,c))] - \\ & E_{x\sim p^x_{data}, c\sim p_c}[\log P(S = fake|G(x,c)] + \mathbf{L_C}. \end{aligned} \tag{5}$$

We can choose maximum absolute error, Huber loss and maximum mean discrepancy as content loss (Gretton et al. 2012).

- Maximum absolute error (MAE)

$$L_C = E_{x\sim p^x_{data}, y\sim p^y_{data}, c\sim p_c} \left| G(x,c) - y \right|. \tag{6}$$

  As shown in the Eq. (6), MAE is the absolute value between generated samples $G(x,c)$ and real samples y. While the gradient descent is executed, the large gradient of MAE increases the training speed.

- Huber loss

$$L_\delta(y, G(x,c)) = \begin{cases} \frac{1}{2}(y - G(x,c))^2, |y - G(x,c)| \le \delta, \\ \delta \left| y - G(x,c) \right| - \frac{1}{2}\delta^2, otherwise. \end{cases} \tag{7}$$

  Huber loss combines the advantages of MAE and MSE by the error parameter δ. When the difference between generated samples $G(x,c)$ and real samples y is less

Table 1: CWRU bearing dataset

Fault width (inch)	Motor load (hp)	Shaft speed (rpm)	Fault Condition
0.007	0,1,2,3	1797,1772, 1750,1730	Inner race Ball, Outer race, Combo
0.014	0,1,2,3	1797,1772, 1750,1730	Inner race Ball, Outer race
0.021	0,1,2,3	1797,1772, 1750,1730	Inner race Ball, Outer race, Combo
0.028	0,1,2,3	1797,1772, 1750,1730	Inner race Ball

than δ, the loss becomes MSE. Otherwise, it is MAE. The large gradient of MAE increases the speed of training model. MSE makes optimization achieve near the minimum point as the gradient decreases.

- Maximum mean discrepancy (MMD)

$$MMD[\mathfrak{F}, p^x_{data}, p^y_{data}] = \sup_{f \in \mathfrak{F}}(E_{x\sim p^x_{data}}[f(G(x,c))] - E_{y\sim p^y_{data}}[f(y)]) \tag{8}$$

Let $\mathfrak{F}$ be a class of function f, the main idea of MMD is to calculate the discrepancy of two samples from two different probability distributions p^x_{data} and p^y_{data} over function f.

Assuming that $[x, y]$ are obtained from two datasets $[p^x_{data}, p^y_{data}]$ and their sizes are $[m, n]$. MMD is the mean of difference between all points in the two samples:

$$MMD[\mathfrak{F}, x, y] = \sup_{f \in \mathfrak{F}}(\frac{1}{m}\sum_{i=1}^{m} f(G(x_i, c)) - \frac{1}{n}\sum_{i=1}^{n} f(y_i)) \tag{9}$$

CNN-based Fault Diagnosis Model

We adopt VGGNet (Simonyan and Zisserman 2014) as CNN-based FD model, since the convolution/pooling layer of CNN can automatically learn important features (Gao et al. 2018; Ding and He 2017). We modify the architecture of VGGNet according to the characteristics of spectrum data and decrease the number of layers. Since the characteristic frequency appears in every length of the spectrum, the kernel size of the first layer is modified to $[8 * 8]$. Besides, the time-frequency diagram is not as complicated as the general graph. We can achieve accurate classification performance with only three convolution and pooling layer. According to the evaluation, the accuracy of the modified VGGNet is 98.8%, which is twice higher than the original VGG-16.

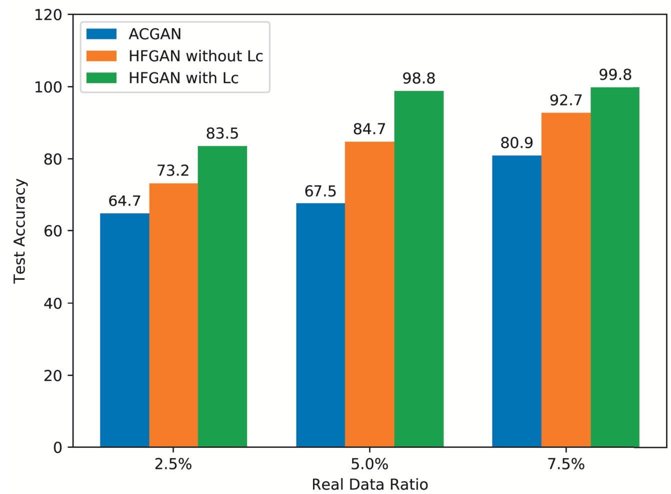


Figure 4: Comparison of classification accuracy.

Experiments

In this section, we give numerical results to do the ablation study for the proposed framework and discuss the data generation performance of HFGAN.

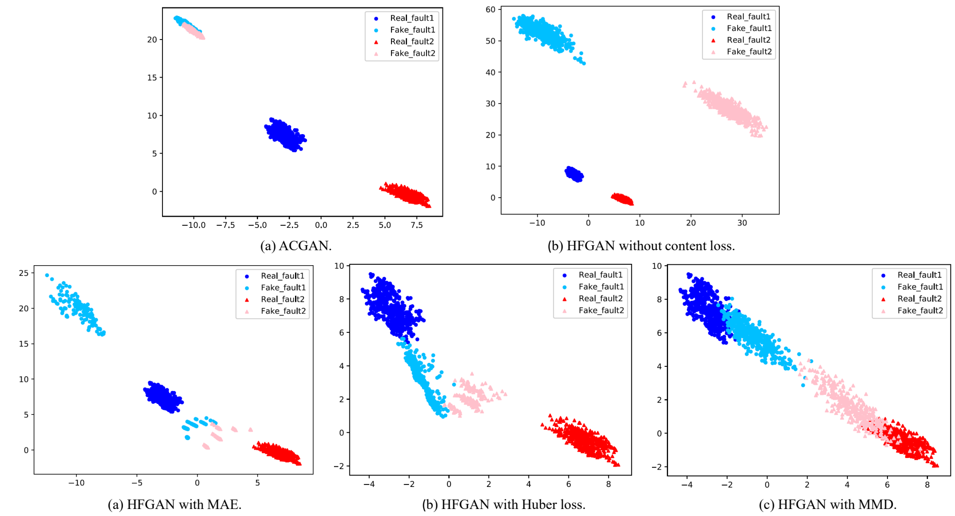
Data Preparation

The dataset is provided by the Case Western Reserve University (CWRU) Bearing Data Center (CWR ; Smith and Randall 2015). CWRU data contains several working setting, as shown in Table 1. The different working setting results in five conditions: 1)health 2)inner ring fault 3)outer ring fault 4)ball fault 5)combination fault on both inner ring and outer ring, which will be called "combo fault." In the following experiments, the 50% health data will be the input of HFGAN to generate different fault conditions. To simulate insufficient fault data, we only utilize 2.5%, 5% and 7.5% fault data for training. The rest of 90% data is for testing.

Ablation Study of Fault Diagnosis Model

The ablation study is comparing the effect of ACGAN (Shao and Yan 2019), HFGAN without content loss and HFGAN with content loss. The discriminator of HFGAN consists of four convolution layers and two fully-connected layers with Leaky ReLU and softmax. The kernel size of convolution layer is $[4 * 4]$, and there are 64, 128, 256 and 512 kernels in four covolution layers. The generator of HFGAN consists of three convolution layers, one resnet block, three deconvolution layer and output with hyperbolic tanget function. ACGAN generates ball fault data and outer ring fault separately based on random noise. Otherwise, HFGAN generates ball fault data and outer ring fault data based on 50% health data of CWRU. We randomly select 200 fake ball fault data and 200 fake outer ring flat data to mix with 2.5%, 5% and 7.5% real data as the training input of the CNN-based FD model. Eq. 10 is adopted to evaluate the classification accuracy:

$$Testing\ Accuracy = \frac{N_{correct}}{N_{correct} + N_{wrong}}, \tag{10}$$



(a) ACGAN. (b) HFGAN without content loss.

(a) HFGAN with MAE. (b) HFGAN with Huber loss. (c) HFGAN with MMD.

Figure 5: The performance of different loss functions.

Table 2: Classification accuracy under various training setting

Proportion of Real Data	Additional Fake Data	Testing Accuracy
2.5%	0	78.7%
2.5%	200 ball fault+ 200 outer ring fault	83.5%
5%	0	79.6%
5%	200 ball fault+ 200 outer ring fault	98.8%
7.5%	0	82.3%
7.5%	200 ball fault+ 200 outer ring fault	99.8%

where $N_{correct}$ is the number of correct classification results and N_{wrong} is the number of wrong classification result.

In Figure 4, we can see that HFGAN achieves higher accuracy than ACGAN by modifying the loss functions of discriminator and generator and adding the content loss makes the result more accurate.

In addition, we want to discuss whether the generated data improves the accuracy of CNN-based FD model, Table 2 shows that adding fake data improves model accuracy 4.8%, 19% and 17.5% with 2.5%, 5% and 7.5% of real data, respectively. We can see that mixing fake data with real data not only solves data imbalance problem, but also intensively increases the classification accuracy.

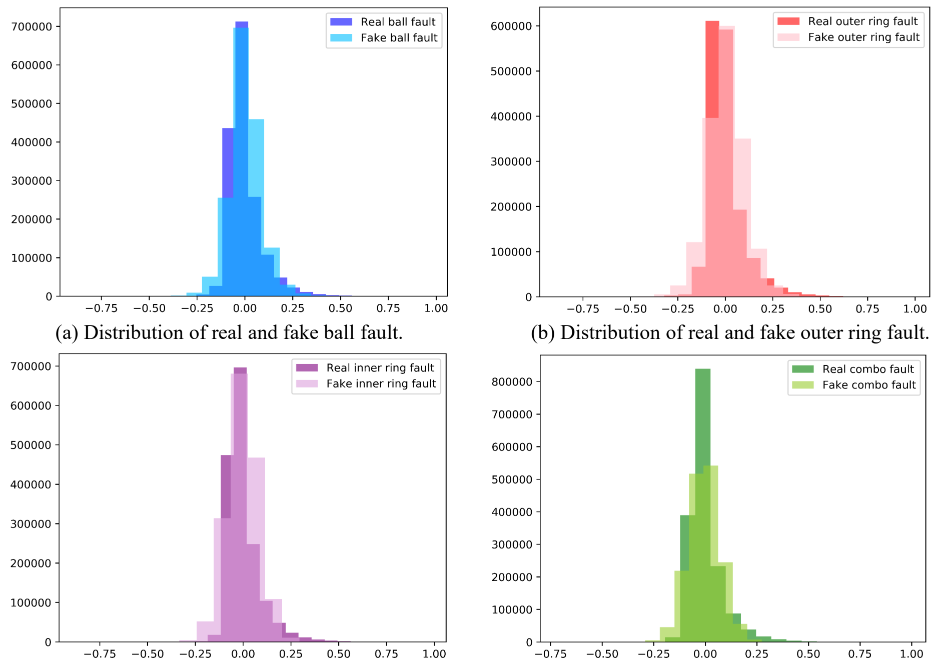
Quality of Data Generation

Figure 5 shows the distribution of generated data and real data in classification latent space of ACGAN, HFGAN without content loss and HFGAN with different content losses. We pretrain a fault classification model and output the last full-connected layer with two neurons. The fake data of fault 1 and fault 2 generated by ACGAN are too close to each other, as shown in Figure 5(a). These data will result in poor classification results. After modifying the generator loss and discriminator loss, the fake data of fault 1 and fault 2 is separated, as shown in Figure 5(b). The classification result is better. However, the fake data is still far from the real data in latent space. We add different content loss in generator. After adding content loss, fake data and real data becomes closer. Since the invariable large gradient makes MAE unable to reach the minimum point and domain shift between two dataset is not considered in Huber loss, some of the fake data of fault 1 and fault 2 still overlaps in Figure 5(c) and Figure 5(d). With MMD as content loss, the distribution of different fake data and real data becomes perfect. Besides, the test accuracy of FD model with MAE, Huber loss and MMD are 68.25%, 86.39% and 99.8%, respectively. Therefore, according to the distribution of generated data and test accuracy, MMD outperforms MAE and Huber loss.

By the modification of loss functions and adding content loss, Figure 6 shows the distribution of different fake fault data generated by HFGAN based on health data. Regardless of generating which type of fault data, the distribution of fake fault data is very similar to real fault data.

Conclusion

In this paper, we proposed an end-to-end fault diagnosis framework to generate different fault data and diagnose faults for machine tools. In the framework, data preprocessing module transformed fault diagnosis to pattern recognition task by converting time-series vibration to time-frequency image. In data generation module, we replaced the input noise with health data and proposed a health-based fault generative adversarial network (HFGAN). In HFGAN, we modified the loss functions of generator and discriminator and added maximum mean discrepancy (MMD) as content loss to solve augmentation uncertainty issue in ACGAN and improve the quality of data generation. The proposed HFGAN can generate different types of fault data at



(a) Distribution of real and fake ball fault.

(b) Distribution of real and fake outer ring fault.

(c) Distribution of real and fake inner ring fault.

(d) Distribution of real and fake combo fault.

Figure 6: Distribution of real and fake fault.

the same time based on health data to achieve one-to-many domain data generation. According to the experiments, factories can only collect less than 7.5% real machine data and mix with the generated fake data to achieve 99.8% classification accuracy. In conclusion, the proposed end-to-end fault diagnosis framework with HFGAN not only generated high quality fault data based on health data to solve data imbalance issue as well as augmentation uncertainty issue but also can intensively reduce the data collection cost in factories.

References

Arjovsky, M.; Chintala, S.; and Bottou, L. 2017. Wasserstein gan. *arXiv preprint arXiv:1701.07875*.

Case Western Reserve University Bearing Data Center. http://csegroups.case.edu/bearingdatacenter/home.

Ding, X., and He, Q. 2017. Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis. *IEEE Transactions on Instrumentation and Measurement* 66(8):1926–1935.

Gao, D.; Zhu, Y.; Wang, X.; Yan, K.; and Hong, J. 2018. A fault diagnosis method of rolling bearing based on complex morlet cwt and cnn. In *IEEE Prognostics and System Health Management Conference (PHM-Chongqing)*.

Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial nets. In *Advances in neural information processing systems*.

Gretton, A.; Borgwardt, K. M.; Rasch, M. J.; Schölkopf, B.; and Smola, A. 2012. A kernel two-sample test. *Journal of Machine Learning Research* 13(Mar):723–773.

Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; and Courville, A. C. 2017. Improved training of wasserstein gans. In *Advances in neural information processing systems*.

Hasani, R.; Wang, G.; and Grosu, R. 2019. A machine learning suite for machine components' health-monitoring. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 9472–9477.

Lin, C.-Y.; Weng, C.-P.; Wang, L.-C.; Shuai, H.-H.; and Tseng, W.-P. 2019. Edge-based rnn anomaly detection platform in machine tools. *Smart Science* 7(2):139–146.

Odena, A.; Olah, C.; and Shlens, J. 2017. Conditional image synthesis with auxiliary classifier gans. In *International Conference on Machine Learning-Volume 70*. JMLR. org.

Odena, A. 2016. Semi-supervised learning with generative adversarial networks. *arXiv preprint arXiv:1606.01583*.

Radford, A.; Metz, L.; and Chintala, S. 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

Shao, S.; Wang, P., and Yan, R. 2019. Generative adversarial networks for data augmentation in machine fault diagnosis. *Computers in Industry* 106:85–93.

Simonyan, K., and Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Smith, W. A., and Randall, R. B. 2015. Rolling element bearing diagnostics using the case western reserve university data: A benchmark study. *Mechanical Systems and Signal Processing* 64:100–131.