Health Index Extraction based on Generative Adversarial Networks

H. J. Park[1], S. Kim[1] and J. H. Choi[2*]

[1]Department of Aerospace and Mechanical Engineering, Korea Aerospace University, Go-yang, Republic of Korea
[2]School of Aerospace and Mechanical Engineering, Korea Aerospace University, Go-yang, Republic of Korea

[*]Corresponding author: jhchoi@kau.ac.kr

1. Introduction

Rapid development of smart factory has significantly increased the importance of the amount and the quality of data. With the acquired data, the diagnostic and prognostic algorithms have been studied to monitor the health of manufacturing systems. Recently, deep learning algorithm has drawn a great attention due to its high diagnosis and predictive performance with increased amount of data. However, in many engineering systems, there are plenty of data from normal condition and few or even no data from fault condition. This causes data imbalance and lowers the performance of deep learning algorithms.

Recently, many literatures propose the Generative Adversarial Networks (GAN) to generate virtual dataset and use it to solve the problem of data imbalance [1–5]. Many literatures use the generative model of GAN to generate virtual datasets. In this paper, authors uses GAN to extract health index of the system and predict its remaining useful life (RUL) by using its discriminative model.

2. Generative Adversarial Network

The GAN is contained with two models, a generative model G and a discriminative model D. The model G generates the virtual data while model D discriminates the real data from the virtual data. As the two models compete against each other, the generative model creates data which is closer to the real.

First, prior on input noise variables $p_z(z)$ is defined to learn the generator's distribution p_g over data x, then represent a mapping to data space as $V(\mathrm{D},\mathrm{G})$, where G is a differentiable function represented by a multilayer perceptron with parameters θ_g. Also define a second multilayer perceptron $D(z;\theta_g)$ that outputs a single scalar. $D(x)$ represents the probability that x came from the data rather than p_g. We train D to maximize the probability of assigning the correct label to both training and samples from G. We simultaneously train G to minimize $\log(1-D(G(z)))$. In other words, D and G play the following two-player minimax game with value function $V(\mathrm{D},\mathrm{G})$ [6]:

$$\min_G \max_D V(D,G) = \mathrm{E}_{x\sim p_{data}(x)}\left[\log D(x)\right] + E_{z\sim p_z(x)}\left[1-\log D(G(z))\right]$$

3. Proposed approach

In many engineering systems, there are considerable of data from normal condition compared to fault condition. Thus, the number of data is not a concern if GAN is constructed to generate normal data. In this approach, authors trained GAN with data from only normal condition and extracted health index using discriminator.

The process is illustrated with numerical simulation. In the Fig. 1, a normal condition data that can be expressed with two features is create which is illustrated with blue dots. Then the degradation begins from normal condition as black dots. The red dots are the fake or virtual data. As GAN models are trained, the red dots move to the true datasets. This means the generator generated virtual data close to the true ones and discriminator have been trained whether input data is normal condition or not.

Then, the degradation feature data are used as input to the trained discriminator model. In the Fig. 2, the output of discriminator (health index) degrades as the input data deviates from the normal condition. To generate virtual fault condition data by GAN, it still needs amount of true fault data to train the models. Therefore, this research used the normal data to train the model which do not have to consider about the number of data in many cases, and extract health index from its model while the system degrades.

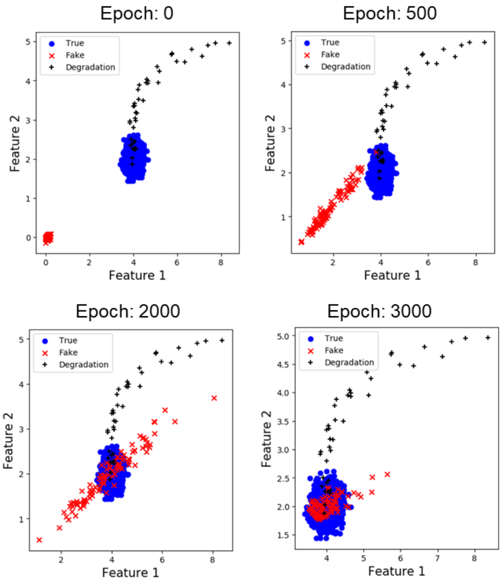

Fig.1 Numerical simulation of GAN

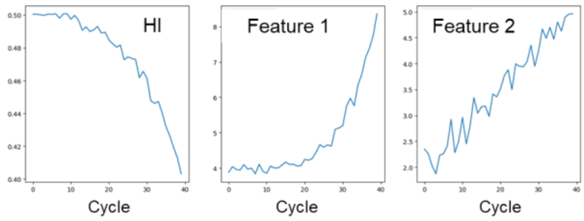

Fig.2 Health index extraction from the trained discriminator

4. Conclusion

The generative adversarial networks have been applied in many studies to solve data imbalance problem in the industrial field and improve deep learning performance. In this paper, authors have used the normal data and trained discriminator model of GANs to extract health index of system degradation.

Acknowledgment

This research was supported by a grant from R&D Program of the Korea Railroad Research Institute, Republic of Korea.

References

[1] S. A. Khan and A. E. Prosvirin, “Towards Bearing Health Prognosis using Generative Adversarial Networks : Modeling Bearing Degradation,” *2018 Int. Conf. Adv. Comput. Sci.*, pp. 1–6.

[2] W. Mao, Y. Liu, L. Ding, and Y. Li, “Imbalanced Fault Diagnosis of Rolling Bearing Based on Generative Adversarial Network : A Comparative Study,” *IEEE Access*, vol. 7, pp. 9515–9530, 2019.

[3] S. Cao, L. Wen, X. Li, and L. Gao, “Application of Generative Adversarial Networks for Intelligent Fault Diagnosis,” *2018 IEEE 14th Int. Conf. Autom. Sci. Eng.*, pp. 711–715, 2018.

[4] Y. Huang, Y. Tang, J. Vanzwieten, J. Liu, and X. Xiao, “An Adversarial Learning Approach for Machine Prognostic Health Management,” *2019 Int. Conf. High Perform. Big Data Intell. Syst.*, pp. 163–168, 2019.

[5] S. Suh, H. Lee, J. Jo, P. Lukowicz, and Y. O. Lee, “applied sciences Generative Oversampling Method for Imbalanced Data on Bearing Fault Detection and Diagnosis,” 2019.

[6] I. J. Goodfellow, J. Pouget-abadie, M. Mirza, B. Xu, and D. Warde-farley, “Generative Adversarial Nets,” pp. 1–9.