PAPER • OPEN ACCESS

Gas sensor fault diagnosis for imbalanced data based on generative adversarial networks

To cite this article: Yongyi Sun *et al* 2021 *J. Phys.: Conf. Ser.* **1732** 012033

View the article online for updates and enhancements.

Gas sensor fault diagnosis for imbalanced data based on generative adversarial networks

Yongyi Sun[1,2], Shuxia Liu[2], Ying Yu[2], Tingting Zhao[3], Zhihui Zou[3], Jinghua Zhang[4], Shuang Zhang[2] and Hongquan Zhang[1,3,*]

[1]Key Laboratory of Electronics Engineering, College of Heilongjiang Province, Heilongjiang University, Harbin, China
[2]Liaoning Police Academy, Dalian, China
[3]School of Automation, Harbin Engineering University, Harbin, China
[4]Tieling Public Security Bureau of Liaoning Province, Tieling, China

*Corresponding author e-mail: xpsy2150@163.com

Abstract. Gas sensors play a key role in gas detection. Gas sensors are prone to failure, so they are of great significance to classify gas sensor fault. Traditional gas sensor fault diagnosis methods do not consider the condition of the imbalance of fault data. In this paper, the generative adversarial networks (GAN) method is used to generate small sample sensor fault data, balance various types of sensor fault data, and then the random forest is used for fault classification. Finally, the sensor fault data is obtained by self-made experimental system to verify the proposed method. The experimental results show that the proposed method could effectively improve the accuracy of gas sensor fault diagnosis.

1. Introduction

The problem of air pollution is becoming more and more serious, so the problem of air pollution has become an urgent problem for many countries and researchers. The gas sensors could monitor the air quality of environment. If the concentration of polluted gas exceeds a certain threshold, the gas sensors will give an alarm signal immediately. However, the gas sensors will be affected by the external environment, such as temperature, humidity and other factors, prone to fault problems, so the fault diagnosis of gas sensors is of great significance [1][2].

At present, traditional machine learning (ML) algorithm is mainly used in fault diagnosis of gas sensor. Chen et al. used ensemble empirical mode decomposition for gas sensor fault diagnosis [3]. Jan et al. used support vector machine (SVM) for sensor fault classification [4]. Pardo et al. used neural networks for monitoring reliability of sensors [5]. Yang et al. used KNN for fault detection of gas sensor arrays [6]. Tao et al. used PCA for sensor fault diagnosis [7]. Lu et al. used extreme learning machine for sensor fault diagnosis [8][9]. Wan et al. used learning vector quantization neural network for fault identification [10-11]. Chen et al. used gray forecasting model for sensor fault diagnosis [12]. Liu et al. used bp neural network for sensor fault diagnosis [13]. Sun et al. used convolutional neural network for sensor fault diagnosis [14].

However, the premise of these methods is that the sensors have enough fault data and all kinds of fault data are balanced. In the actual collection process of gas sensor fault data, some fault data is difficult to collect, so fault data imbalance phenomenon will occur, which will seriously affect the gas sensor fault diagnosis results.

Recent years, since generative adversarial networks (GAN) could generate data samples with similar statistical characteristics, it has been gradually introduced into fault diagnosis. Wang et al. used GAN for fault classification of machinery [15]. Gao et al. used GAN for fault diagnosis of bearing [16]. Zhang et al. used deep GAN for machinery fault diagnosis [17]. However, it is rarely used in gas sensor fault diagnosis.

This paper presents a gas sensor fault diagnosis method based on a GAN and a random forest (RF). Using GAN to generate small samples fault data to make the fault data samples balanced. Finally, fault diagnosis is carried out by the RF to verify the effectiveness of the method.

2. Theoretical background of generative adversarial networks

A GAN is a deep learning (DL) model, which is one of the most promising methods for unsupervised learning on complex distribution recently. In the year of 2016, the boom of GAN swept the top conferences in AI field. From ICLR to NIPS, a large number of high-quality GAN related papers were published and discussed. Yann LeCun once commented that GAN was "the coolest idea in machine learning for 20 years. GAN models can be divided into two categories: generator and discriminator. The generator needs input variables, which can be determined by a certain model. The discriminator needs to judge whether the received information is true or false. The structure of classical GAN is shown in Figure 1.
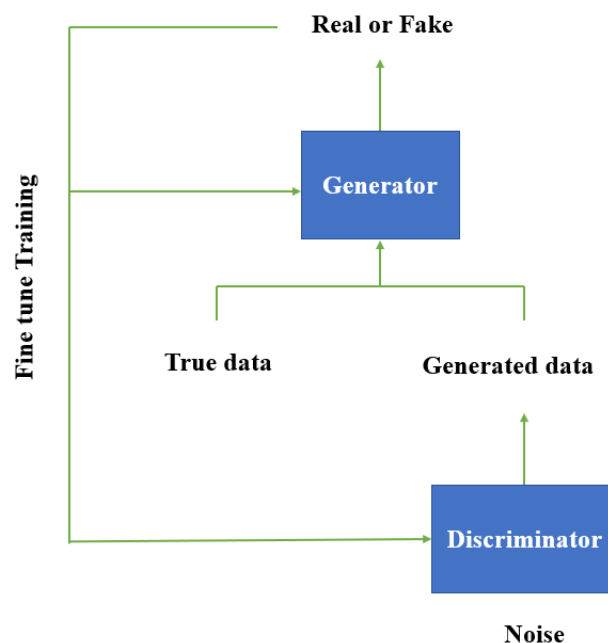

Figure 1. The structure of classical GAN

3. Experiment and results

3.1. Experimental setup

The fault data needed was obtained by our self-made experimental system, and hydrogen was selected as the verification gas. The experimental system included gas molecular flowmeter, stabilized power supply, hydrogen cylinder, gas mixer, etc. SnO_2 sensors were selected, and the sensors model was MQ-8. According to long-term observation and literature, the common fault types of gas sensors are exfoliation of sensitive body (ESB) fault, stuck fault, erratic fault. The training sample data and the testing sample data of the three sensor fault types and one normal signal obtained by the experimental device, as shown in Table 1.

Table 1. Sample number of seven modes of sensor signal.

Models	Models description	Training sample number	Testing sample number
1	Normal signal	100	20
2	ESB fault	100	20
3	Stuck fault	100	20
4	Erratic fault	20	20

3.2. Results

The GAN method was used to increase the volume of Erratic fault data. The number of generated samples was shown in Table 2.

Table 2. Sample dataset after data generation

Modes	Modes description	Newly generated samples	Final training samples
1	Normal signal	0	100
2	ESB fault	0	100
3	Stuck fault	0	100
4	Erratic fault	80	100

Then the RF was used to classify the fault types of gas sensors. In order to verify the effectiveness of the algorithm, two groups of experiments were carried out. In Experiment 1, the fault data was generated by the GAN, and the RF was used for fault classification. In Experiment 2, the RF algorithm was used to classify the faults types using the original fault data in Table 1. The results of Experiment 1 are obviously better than those of Experiment 2, shown in Table 3.

Table 3. Comparison of experimental results

Experiment number	Experiment description	Fault classification results
1	GAN+RF	91.25%
2	RF	88.75%

4. Conclusion

In this paper, the gas sensor fault diagnosis method based on the GAN and the RF is proposed. Firstly, the GAN is used to expand the small sample fault data to make the sensor fault data sample balanced. Secondly, the fault types are classified by the RF algorithm. Finally, the fault data are obtained through the self-made experimental system, and the proposed method is verified. The experimental results show that the proposed method has the advantages of high accuracy.

5. Acknowledgments

References

[1] S. G. Chalk and J. F. Miller, "Key challenges and recent progress in batteries, fuel cells, and hydrogen storage for clean energy systems," J. Power Sour., vol. 159, no. 1, pp. 73_80, Sep. 2006.

[2] H. Lee, J.-R. Sohn, S.-H. Byeon, S. Yoon, and K. Moon, "Alternative risk assessment for dangerous chemicals in South Korea regulation: Comparing three modeling programs," Int. J. Environ. Res. Public Health, vol. 15, no. 8, p. 1600, 2018.

[3] Y.-S. Chen, Y.-H. Xu, J.-L. Yang, Z. Shi, S.-D. Jiang, and Q.Wang, "Fault detection, isolation, and diagnosis of status self-validating gas sensor arrays," Rev. Sci. Instrum., vol. 87, no. 4, Apr. 2016, Art. no. 045001.

[4] S. U. Jan, Y.-D. Lee, J. Shin, and I. Koo, "Sensor fault classi_cation based on support vector machine and statistical time-domain features," IEEE Access, vol. 5, pp. 8682_8690, 2017.

[5] M. Pardo, G. Faglia, G. Sberveglieri, M. Corte, F. Masulli, and M. Riani, "Monitoring reliability of sensors in an array by neural networks," Sens. Actuators B, Chem., vol. 67, nos. 1_2, pp. 128_133, Aug. 2000.

[6] J. Yang, Z. Sun, and Y. Chen, "Fault detection using the clustering-kNN rule for gas sensor arrays," Sensors, vol. 16, no. 12, p. 2069, 2016.

[7] E. P. Tao, W. H. Shen, T. L. Liu, and X. Q. Chen, "Fault diagnosis based on PCA for sensors of laboratorial wastewater treatment process," Chemometric Intell. Lab. Syst., vol. 128, pp. 49_55, Oct. 2013.

[8] J. Lu, J. Huang, and F. Lu, "Sensor fault diagnosis for aero engine based on online sequential extreme learning machine with memory principle," Energies, vol. 10, no. 1, p. 39, 2017.

[9] K. Song, P. Xu, Y. Chen, T. Zhang, G. Wei, and Q. Wang, " A fault diagnosis and recon_guration strategy for self-validating hydrogen sensor array based on mwpca and elm," IEEE Access, vol. 7, pp. 115 075-115 092, 2019

[10] N. Wan, W. Zhang, and K. Li, "Fault identi_cation of high frequency current ripple of fuel cell based on learning vector quantization neural network," DEStech Trans. Eng. Technol. Res., 2017.

[11] A. M. Bassiuny, X. Li, and R. Du, "Fault diagnosis of stamping process based on empirical mode decomposition and learning vector quantization," Int. J. Mach. Tools Manuf., vol. 47, no. 15, pp. 2298_2306, Dec. 2007.

[12] Y. Chen, J. Yang, Y. Xu, S. Jiang, X. Liu, and Q. Wang, "Status selfvalidation of sensor arrays using gray forecasting model and bootstrap method," IEEE Trans. Instrum. Meas., vol. 65, no. 7, pp. 1626_1640, Jul. 2016.

[13] H. Liu, B. Li, T. Liu, M. Wang, H. Fu, and R. Guo, "Angular acceleration sensor fault diagnosis based on lm-bp neural network," in 2018 37th Chinese Control Conference (CCC). IEEE, 2018, pp. 6028-6032.

[14] Y. Sun, H. Zhang, T. Zhao, Z. Zou, B. Shen, and L. Yang, " A new convolutional neural network with random forest method for hydrogen sensor fault diagnosis," IEEE Access, vol. 8, pp. 85 421-85 430, 2020.

[15] J. Wang, S. Li, B. Han, Z. An, H. Bao, and S. Ji, "Generalization of deep neural networks for imbalanced fault classification of machinery using generative adversarial networks," Ieee Access, vol. 7, pp. 111 168-111 180, 2019.

[16] Y. Gao, X. Liu, and J. Xiang, "Fem simulation-based generative adversarial networks to detect bearing faults," IEEE Transactions on Industrial Informatics, vol. 16, no. 7, pp. 4961-4971, 2020.

[17] W. Zhang, X. Li, X.-D. Jia, H. Ma, Z. Luo, and X. Li, "Machinery fault diagnosis with imbalanced data using deep generative adversarial networks," Measurement, vol. 152, p. 107377, 2020.