Huan-gang Wang, Xin Li, Tao Zhang, 2018. Generative adversarial network based novelty detection using minimized reconstruction error. *Frontiers of Information Technology & Electronic Engineering*, 19(1): 116-125.
https://doi.org/10.1631/FITEE.1700786

Generative adversarial network based novelty detection using minimized reconstruction error

Key words: Generative adversarial networks; Novelty detection; Tennessee Eastman process

Contact: Huan-gang WANG
E-mail: hgwang@tsinghua.edu.cn
ORCID: http://orcid.org/0000-0002-7322-3446

Introduction

- Novelty detection has aroused great attention in the areas such as industry process fault detection, medical diagnosis, drug discovery, and fraud detection in the finance field.
- Generative adversarial network (GAN) is the most exciting machine learning breakthrough in recent years, and it is introduced and investigated for novelty detection.
- GAN is trained on only ordinary data to learn the ordinary data description, and separate novelty from ordinary patterns on previously unknown data.

Structure of the GAN

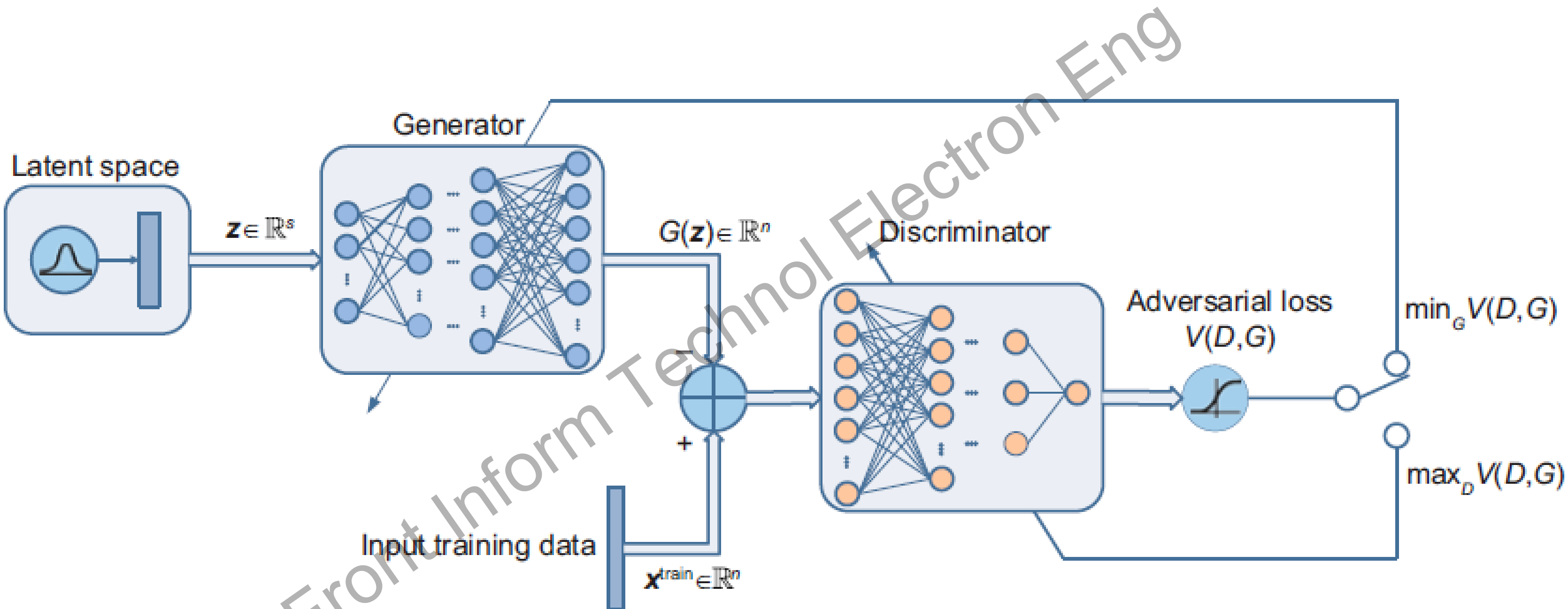

Fig. 1 Structure of the generative adversarial networks (GAN)

Results of a GAN model on 2D synthetic datasets

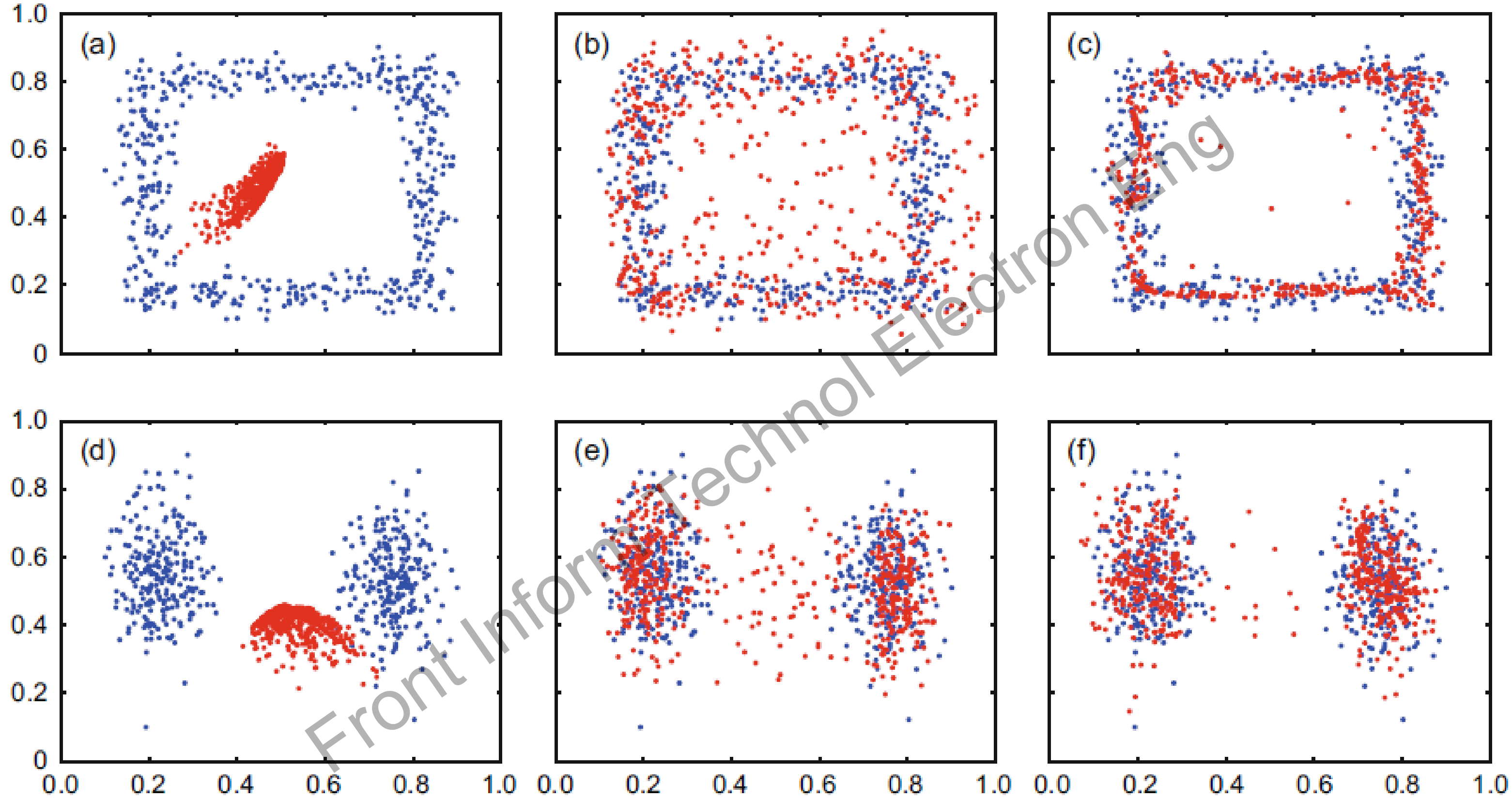

Fig. 2 **Results of a GAN model using multilayer perceptrons on 2D synthetic datasets: (a)–(c) are the results of the 1st, the 300th, and the 3000th iteration, respectively, on a square distribution; (d)–(f) are the results of the 1st, the 300th, and the 3000th iteration, respectively, on a two-model distribution**

Blue points represent real data from the synthetic dataset, and red points represent the points generated by the GAN model. References to color refer to the online version of this figure

Illustration of G-score

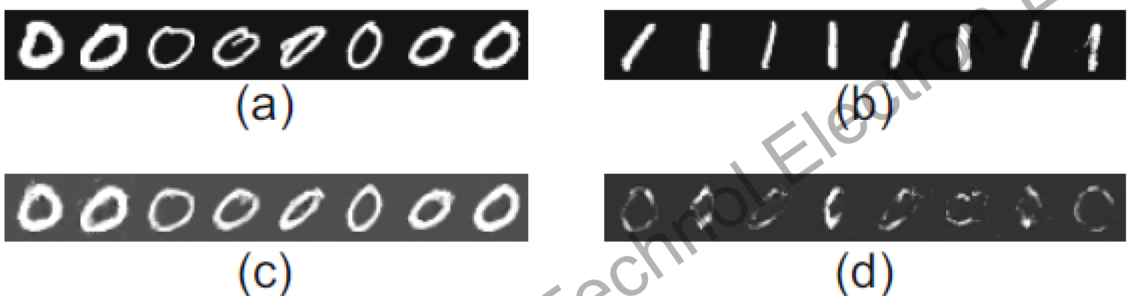

Fig. 4 Illustration of G-score: (a) and (b) are real '0' and '1' digits from the MNIST database not contained in the training dataset, respectively; (c) and (d) are '0' and '1' digits reconstructed by the generator, respectively

ROC and AUC on MNIST digit database

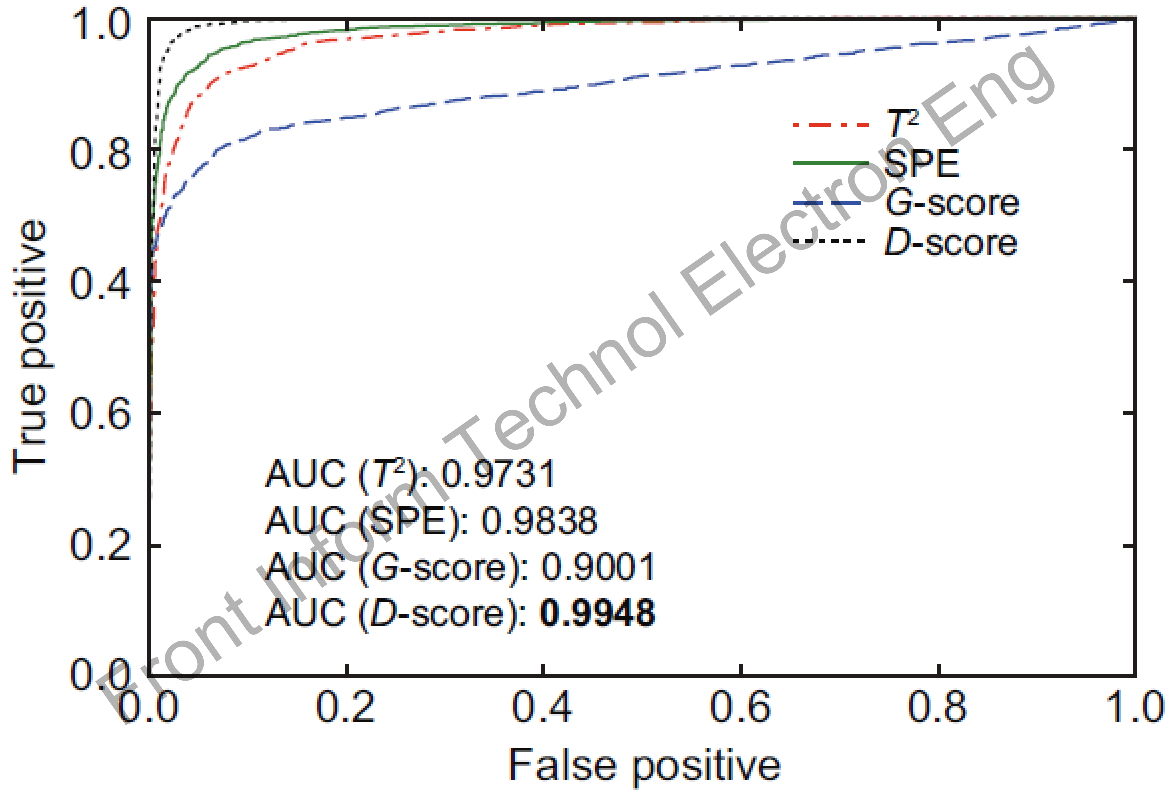

Fig. 6 Receiver operating characteristic (ROC) curves on the test dataset and the area under curve (AUC) value of each score

Structure of the Tennessee Eastman process

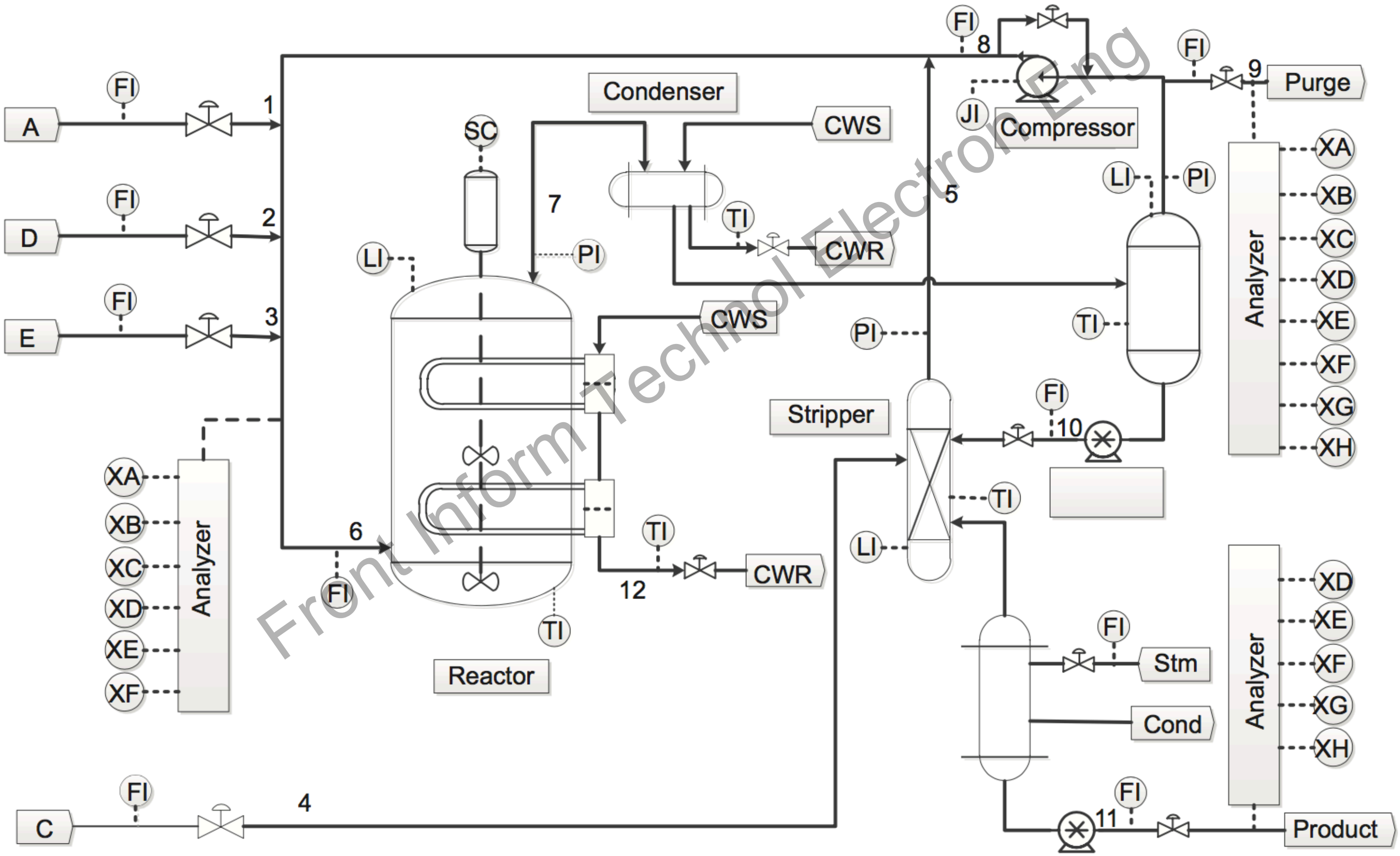

Fig. 7 Structure diagram of the Tennessee Eastman (TE) process

AUC on 21 faults of the TE process

Table 2 The area under curve (AUC) values on 21 faults of the Tennessee Eastman (TE) process

Fault	AUC			
	T^2	SPE	G-score	D-score
IDV(1)	0.9967	**0.9999**	0.9993	0.9962
IDV(2)	**0.9962**	0.9957	0.9956	0.9888
IDV(3)	0.5704	0.5234	0.4980	**0.6388**
IDV(4)	0.8314	**1.0000**	0.9012	0.5086
IDV(5)	**0.7540**	0.7301	0.7065	0.6812
IDV(6)	0.9994	**1.0000**	**1.0000**	**1.0000**
IDV(7)	0.9111	**1.0000**	**1.0000**	0.9843
IDV(8)	**0.9920**	0.9919	0.9918	0.9771
IDV(9)	0.4524	0.4907	**0.5111**	0.4535
IDV(10)	0.8238	0.8466	**0.8829**	0.7749
IDV(11)	0.7865	**0.9399**	0.8636	0.6730
IDV(12)	0.9951	0.9943	**0.9960**	0.9826
IDV(13)	0.9847	0.9764	**0.9918**	0.9646
IDV(14)	0.9981	**1.0000**	0.9999	0.8067
IDV(15)	0.6741	0.5645	**0.6864**	0.6131
IDV(16)	0.6047	**0.7465**	0.6983	0.6671
IDV(17)	0.9489	**0.9819**	0.9429	0.8878
IDV(18)	**0.9672**	0.9595	0.9652	0.9179
IDV(19)	0.6062	**0.8958**	0.6762	0.5449
IDV(20)	0.8601	**0.8868**	0.8545	0.8057
IDV(21)	0.7328	0.7418	**0.7938**	0.7086

Conclusions

- The GAN can generate new samples similar to the training data, and such implicit data description of normal data is transformed to a novelty score for novelty detection by formulating the G-score and D-score.
- A new structure to directly map the data to latent space in order to shorten this time needs further study.
- The scores can also be used in other areas like medical diagnosis when trained on medical measurements or images and used in drug discovery when molecules are represented in a feature space.