Deep Learning Techniques in Fault Diagnosis and Prognosis

Subjects: Computer Science, Artificial Intelligence

Contributor: Shaohua Qiu , Xiaopeng Cui , Zuowei Ping , Nanliang Shan , Zhong Li , Xianqiang Bao , Xinghua Xu

The fault diagnosis and prognosis (FDP) technique based on data-driven machine learning (ML) methods recognizes or learns the health features of the system from historical data, and tries to discover and mine the information hidden in the data, so that it can accurately analyze and predict future system behavior without precisely knowing the forward physical model.

fault diagnosis   fault prognosis   machine learning   deep learning

---

1. Background

Industrial systems are typical complex systems with various subsystems and device types of mechanical system, power system, information system, electronic system, or their combinations. They are playing an increasingly important role in the economy, such as manufacturing industry, energy industry and chemical industry, which are now developed with more functions, more sophisticated structures, and larger scales [1]. Reliability issues have gradually become the key of whether many modern industrial systems can be truly practical. Once a failure occurs, it may affect the safe and stable operation of the entire system, i.e., reducing the efficiency of the system, and causing system breakdown or damage in severe cases [2]. It may also endanger personnel safety, and cause other catastrophic consequences. Therefore, the early identification of faults in advance can greatly help to take appropriate actions of maintenance to avoid the undesired consequences.

Driven by demand, prognostics and health management (PHM) [3] technology, firstly originated from engine health monitoring systems [4], has gained increasingly more attention. PHM is an expansion of the traditional reliability or predictive maintenance concept oriented for complex industrial systems. It realizes the development from the initial condition monitoring and fault diagnosis that aims to estimate health status, to health management that aims at formulating the countermeasures based on the results of monitoring, diagnosis, and prognosis.

In practical scenes, it is often difficult or even impossible to establish mathematical models of complex components or systems [5], in order to trace and analyze faults. Therefore, a large amount of historical data that were collected in the process of system operation and maintenance have become the major method by which to evaluate the system's health status. As the core part of PHM technology, the fault diagnosis and prognosis (FDP) technique based on data-driven machine learning (ML) methods recognizes or learns the health features of the system from historical data, and tries to discover and mine the information hidden in the data, so that it can accurately analyze and predict future system behavior without precisely knowing the forward physical model. ML methods generally

have a more powerful capacity for FDP without the assumption of data distribution, smoother and more intelligent FDP processes with fewer processing stages and less human intervention, and, moreover, less prior-knowledge requirements for more complex components or systems to be modeled [6].

Consequently, data-driven ML methods have long been applied in various industrial FDP applications. A typical ML pipeline generally consists of three steps [7], i.e., data preprocessing, feature extraction and classification or regression. The performance of ML heavily depends on the manually predefined feature extraction rules. In the past decade, with the great development of mega-scale open datasets [8], evolutional computing capacity of new GPU architectures [9] and innovative neural network training methods [10], deep learning [11] can hierarchically extract highly-abstract features in an end-to-end way from the labeled training dataset. Due to its superior performance over ML methods, deep learning (DL) has gained remarkable success in the tasks of computer vision, natural-language processing, etc. In the community of industrial FDP, researchers have also made great efforts to introduce DL techniques into different and unique industrial FDP scenarios, and tremendous progress has been witnessed.

2. Definitions of Faults

In general, the condition monitoring results of certain object in industrial systems experiences changes all the time, and not all changes in sensory data are failures or faults. Here are some common senses:

- Changes caused by random noise are not necessarily faults, but when the variance of the noise changes, it is generally considered to be a fault.

- Fluctuation within a stable range in a certain operation condition is not a malfunction. In different operating conditions, this fluctuation may be different.

- A change that breaks the current pattern is a fault.

Figure 1 gives a comparison of the normal three-phase current waveform and the current waveform of interturn short-circuit fault under the same working condition. At no point does the current amplitude exceeds the working condition mode range, but the (blue) curve pattern of $t > 125$ ms changes and it is a fault.
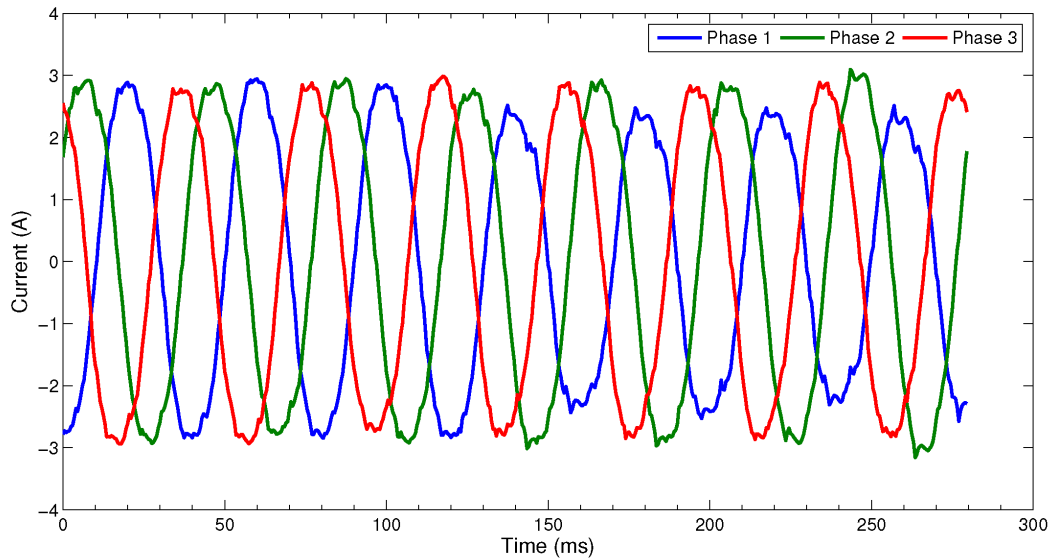
Figure 1. An example of three-phase current waveform.

3. Modern Deep Learning Techniques for Intelligent Industrial FDP

3.1. Modern Deep Learning Techniques

As a young and developing field of AI, ML techniques try to discover knowledge from a large amount of historical data for prediction or classification on new data. More specifically, it is designed to find a projection to fit the input data for desired results, which is often too complex to be explicitly formulated. In terms of application purposes, supervised machine learning is mainly divided into two categories [12]: classification and regression. The former learns the boundaries between categories to achieve classification of new data [13]. The latter fits regularities to the data to predict the properties of new data points. Correspondingly, fault diagnosis is actually a classification problem, and fault prognosis is a regression problem.

As a subset of ML, the emerging DL is currently the hottest topic in AI. It is originated from the paper [10] published in 2006 by Hinton et al. There reveals two characteristics of deep learning. The first is that the neural network with multiple hidden layers has excellent potential for learning more representative features from raw data which are generally designed manually in traditional ML methods. The second is that the difficulty of training deep neural networks can be overcome by layer-by-layer pre-training using the method of unsupervised learning in the Restricted Boltzmann Machine (RBM).

The concept "deep" in deep learning is compared to traditional machine learning algorithms, such as SVM, ANN, and other shallow learning methods, in which there are more layers of non-linear functions in deep learning methods. In traditional shallow neural learning methods, data sample features need to be manually extracted. Conversely, DL automatically learns to obtain feature representations by performing layer-by-layer feature transformation on original data via back-propagation, and these hierarchical feature representations are highly

abstract and task-oriented. One of its major merits is that it can complete the learning in an end-to-end way directly from raw data to results of classification and regression tasks.

Typical DL architectures include deep belief network (DBN) [14], autoencoder (AE) [15], convolutional neural network (CNN) [16], and RNN [17]. With the rapid development of DL techniques in these years, many new architectures have been proposed and introduced into the tasks of intelligent industrial FDP. Examples are generative adversarial network (GAN) [18], transformer [19], and graph neural network (GNN) [20]. Similarly, CNN is prospering again, due to the progress made in the fields of computer vision in recent years.

3.2. Categorization and Literature Trends of DL Techniques for Industrial FDP

Figure 2 shows the categorization of major DL-based approaches used in intelligent FDP. According to the supervision type, they can be divided into unsupervised methods and supervised methods. The former tries to find the inherent common pattern within data which are unlabeled, while the latter refers to methods that learn highly non-linear relationship between the input data and its paired labeled output. More specifically, the supervised methods can be further divided into processing of specific data types or extraction of distinctive features, depending on their objectives. Their detailed introductions will be expanded in the following sections.
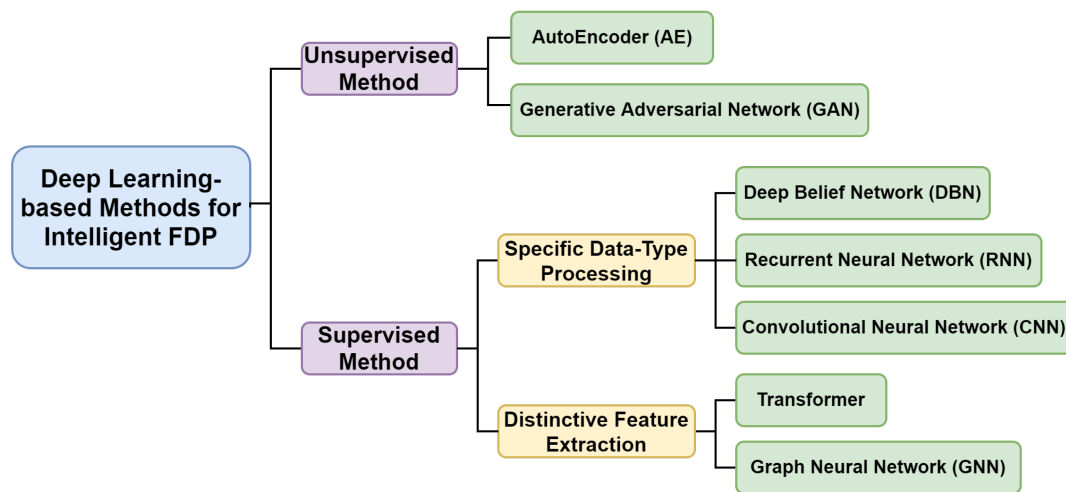

Figure 2. The categorization of deep learning techniques in intelligent FDP.

Figure 3 illustrates the number of journal publications of deep learning methods in intelligent FDP from January 2013 to September 2022 on Web of Knowledge. As can be seen, the number of papers published is increasing year by year, and CNN-based FDP methods account for the majority of all methods. The publication number of typical DL architectures, such as DBN and AE, are stable or growing with relatively slower speed. Note that emerging network architectures are also gradually attracting the attention of researchers.
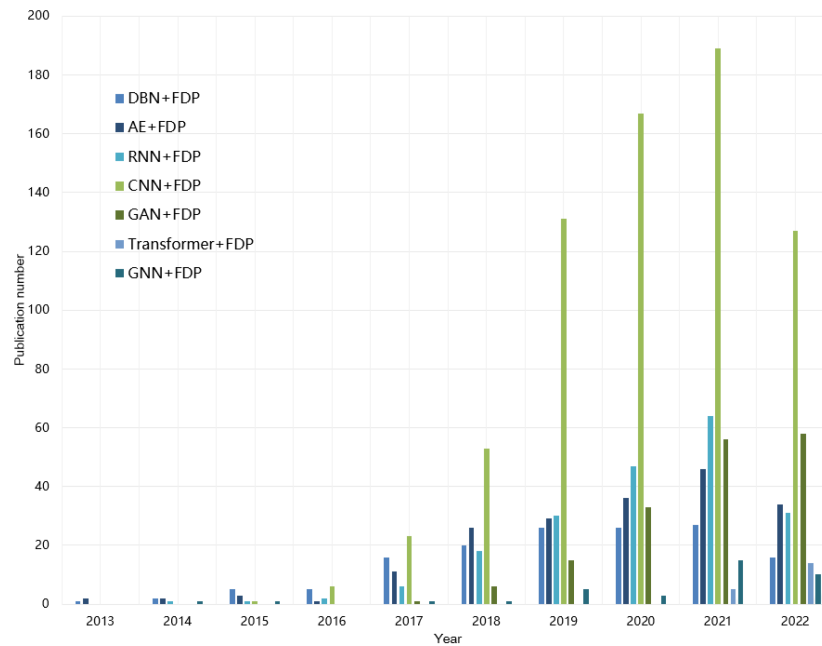
Figure 3. Publication trends of deep learning methods in intelligent industrial FDP.

References

1. Xu, Y.; Sun, Y.; Wan, J.; Liu, X.; Song, Z. Industrial Big Data for Fault Diagnosis: Taxonomy, Review, and Applications. IEEE Access 2017, 5, 17368–17380.

2. Dash, S.; Venkatasubramanian, V. Challenges in the industrial applications of fault diagnostic systems. Comput. Chem. Eng. 2000, 24, 785–791.

3. Zio, E. Prognostics and health management of industrial equipment. In Diagnostics and Prognostics of Engineering Systems: Methods and Techniques; IGI-Global: Hershey, PA, USA, 2013; pp. 333–356.

4. Tumer, I.; Bajwa, A. A survey of aircraft engine health monitoring systems. In Proceedings of the 35th Joint Propulsion Conference and Exhibit, Los Angeles, CA, USA, 20–24 June 1999; p. 2528.

5. Zhong, K.; Han, M.; Han, B. Data-driven based fault prognosis for industrial systems: A concise overview. IEEE/CAA J. Autom. Sin. 2019, 7, 330–345.

6. Tsui, K.L.; Chen, N.; Zhou, Q.; Hai, Y.; Wang, W. Prognostics and health management: A review on data driven approaches. Math. Probl. Eng. 2015, 2015, 793161.

7. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. Mech. Syst. Signal Process. 2020, 138, 106587.

8. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern

Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.

9. NVIDIA. NVIDIA Technologies and GPU Architectures. Available online: https://www.nvidia.com/en-us/technologies/ (accessed on 4 September 2022).

10. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. Science 2006, 313, 504–507.

11. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444.

12. Jain, A.K.; Duin, R.P.W.; Mao, J. Statistical pattern recognition: A review. IEEE Trans. Pattern Anal. Mach. Intell. 2000, 22, 4–37.

13. Zhao, B.; Zhang, X.; Zhan, Z.; Wu, Q. Deep multi-scale adversarial network with attention: A novel domain adaptation method for intelligent fault diagnosis. J. Manuf. Syst. 2021, 59, 565–576.

14. Xie, J.; Du, G.; Shen, C.; Chen, N.; Chen, L.; Zhu, Z. An end-to-end model based on improved adaptive deep belief network and its application to bearing fault diagnosis. IEEE Access 2018, 6, 63584–63596.

15. Mao, W.; Feng, W.; Liu, Y.; Zhang, D.; Liang, X. A new deep auto-encoder method with fusing discriminant information for bearing fault diagnosis. Mech. Syst. Signal Process. 2021, 150, 107233.

16. Huang, Y.C.; Wang, P.J. Infrared Air Turbine Dental Handpiece Rotor Fault Diagnosis with Convolutional Neural Network. Sens. Mater. 2020, 32, 3545–3558.

17. Wu, Q.; Ding, K.; Huang, B. Approach for fault prognosis using recurrent neural network. J. Intell. Manuf. 2020, 31, 1621–1633.

18. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014.

19. Tang, X.; Xu, Z.; Wang, Z. A Novel Fault Diagnosis Method of Rolling Bearing Based on Integrated Vision Transformer Model. Sensors 2022, 22, 3878.

20. Li, T.; Zhao, Z.; Sun, C.; Yan, R.; Chen, X. Multireceptive field graph convolutional networks for machine fault diagnosis. IEEE Trans. Ind. Electron. 2020, 68, 12739–12749.