TOPICAL REVIEW

A review on deep learning in planetary gearbox health state recognition: methods, applications, and dataset publication

IOP Publishing Measurement Science and Technology
Meas. Sci. Technol. **35** (2024) 012002 (23pp) https://doi.org/10.1088/1361-6501/acf390

Topical Review

A review on deep learning in planetary gearbox health state recognition: methods, applications, and dataset publication

Dongdong Liu[1,*], **Lingli Cui**[1,*] **and Weidong Cheng**[2]

[1] Key Laboratory of Advanced Manufacturing Technology, Beijing University of Technology, Beijing, People's Republic of China
[2] School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing, People's Republic of China

E-mail: liudd@bjut.edu.cn and cuilingli@bjut.edu.cn

Received 15 April 2023, revised 13 August 2023
Accepted for publication 24 August 2023
Published 11 October 2023

Abstract

Planetary gearboxes have various merits in mechanical transmission, but their complex structure and intricate operation modes bring large challenges in terms of fault diagnosis. Deep learning has attracted increasing attention in intelligent fault diagnosis and has been successfully adopted for planetary gearbox fault diagnosis, avoiding the difficulty in manually analyzing complex fault features with signal processing methods. This paper presents a comprehensive review of deep learning-based planetary gearbox health state recognition. First, the challenges caused by the complex vibration characteristics of planetary gearboxes in fault diagnosis are analyzed. Second, according to the popularity of deep learning in planetary gearbox fault diagnosis, we briefly introduce six mainstream algorithms, i.e. autoencoder, deep Boltzmann machine, convolutional neural network, transformer, generative adversarial network, and graph neural network, and some variants of them. Then, the applications of these methods to planetary gearbox fault diagnosis are reviewed. Finally, the research prospects and challenges in this research are discussed. According to the challenges, a dataset is introduced in this paper to facilitate future investigations. We expect that this paper can provide new graduate students, institutions and companies with a preliminary understanding of methods used in this field. The dataset can be downloaded from https://github.com/Liudd-BJUT/WT-planetary-gearbox-dataset.

Keywords: planetary gearbox, deep learning, fault diagnosis, vibration characteristic

1. Introduction

Planetary gearboxes have been widely used in various industrial applications, such as those in aerospace, wind turbines, helicopters and heavy trucks, owing to their unique merits, including their high transmission ratio, compact structure and heavy load ability. Due to their use in harsh environments, planetary gearboxes are usually prone to various faults that can result in the shutdown of an entire system and corresponding economic loss [1–3]. Effective fault diagnosis can reduce downtime, prevent catastrophic

* Authors to whom any correspondence should be addressed.

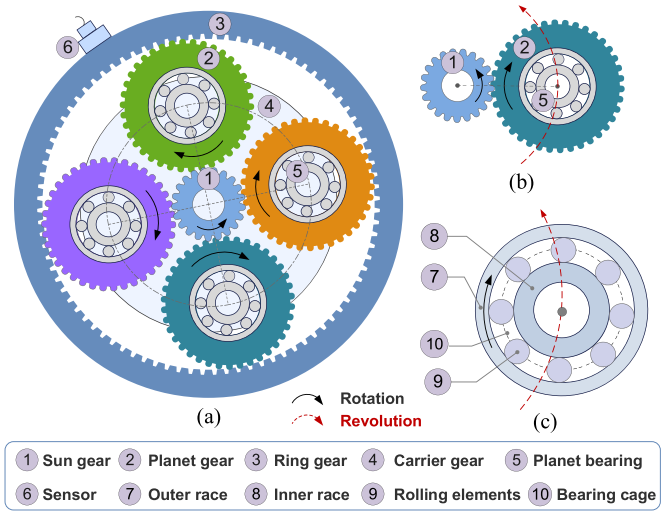

Figure 1. Illustration of (a) planetary gear set, (b) meshing pair of sun gear and planet gear, and (c) planet rolling bearing.

failure, and ensure reliable operation [4]. Therefore, planetary gearbox fault diagnosis is crucial and has attracted much attention.

The effective identification of frequency content is one common and useful method of rotating machinery fault diagnosis. For example, if a rolling bearing has a localized fault, the fault characteristic frequency will exist in the envelope spectra of the bearing [5, 6], and the frequency values will be different when faults are on a different part, such as the outer race, inner race and rolling elements, of a rolling bearing. If a gear has a fault, sidebands will appear around the gear meshing frequency (GMF), and the fault can be localized by analyzing the interval values between the sidebands and the GMF [7, 8]. However, it is challenging to achieve fault diagnosis of planetary gearboxes by analyzing frequency spectra owing to the complicated modulation characteristics caused by their intricate structure and operation modes. As shown in figure 1(a), planetary gearboxes usually consist of a sun gear, multiple planet gears, a ring gear, a planet carrier, and planet bearings. There are multiple meshing pairs, and planet gears and planet bearings, as shown in figures 1(b) and (c), respectively, not only rotate around their own shafts but also revolve around the shaft of the sun gear. In addition, the transmission paths between the fault component and sensor are always time-varying [9, 10]. Thus, the common methods for fixed-shaft gearbox and support bearing fault diagnosis are not applicable for planetary gearboxes owing to their unique vibration characteristics. Researchers have developed many effective methods to implement planetary gearbox fault diagnosis, such as developing special filtering methods to remove the interference components [11] and proposing advanced time-frequency analysis (TFA) methods to better identify the sidebands [12, 13]. These methods have achieved great success, but frequency identification-based methods rely heavily on expert knowledge, and it is time-consuming to generate spectra and analyze the frequency content.

Intelligent fault diagnosis has attracted increasing attention in recent years [14, 15]. Conventional intelligent diagnosis methods mainly consist of feature extraction and pattern recognition [16, 17]. The features, such as time-domain features [18], frequency-domain features [19] and time-frequency-domain features [20], are first calculated based on the collected data, and then they are fed into shallow machine learning classifiers such as support vector machines [21], artificial neural networks [22], *k*-nearest neighbors (*k*NNs) [23] and decision trees (DTs) [24]. To avoid the interference of fault-irrelevant features, feature selection and feature dimension reduction are usually performed in some publications. However, manual feature extraction still requires a great deal of human labor, and feature extraction and pattern recognition cannot be optimized jointly to obtain the best recognition performance [17]. In addition, these methods can only extract shallow features and have a limited ability to mine hidden fault information.

In recent years, deep learning methods have been widely investigated in various fields owing to increases in computing power and the advent of deep learning methods. Deep learning methods include several approaches, such as autoencoders (AEs) [25], deep Boltzmann machines (DBMs) [26] and convolutional neural networks (CNNs) [27]. They extract hierarchical representations from collected data through a deeper neural network structure with multiple layers so that internal representations can be learned directly from raw data, and an end-to-end diagnosis scheme is realized as expected. Compared to conventional machine learning, deep learning does not require extensive human labor and expertise, and health information can be more fully mined by a deeper neural network. In recent years, deep learning has achieved great success in planetary gearbox fault diagnosis owing to its great feature extraction ability.

To summarize the investigations of intelligent fault diagnosis methods, several review papers have been published. However, in [28], traditional machine learning-based methods were mainly reviewed. In [14, 17], the methods mainly focused on bearing and fixed-shaft gearbox fault diagnosis were reviewed. Few publications have been introduced on planetary gearbox fault diagnosis. In [29], only intelligent bearing fault diagnosis methods were reviewed. Compared with rolling bearings and fixed-shaft gearboxes, it is more difficult to realize accurate health state recognition of planetary gearboxes because of the several factors such as complex kinematic mechanism, complex time-varying transfer paths and the interference of multiple meshing vibrations [10]. Various works have discussed the challenges and proposed methods to solve them, such as constructing dynamic models to reveal the special vibration nature [30, 31], installing internal sensors to remove the effect of transfer paths [32, 33], and developing new specialized filters to separate meshing vibrations [11, 34]. Deep learning provides a new solution for planetary gearbox health state recognition, and fruitful and meaningful works have been published. To the best of our knowledge, there are no reviews on deep learning applications in planetary gearbox fault diagnosis. In this paper, we analyze the

vibration characteristics of planetary gearboxes as well as the challenges in fault diagnosis, introduce the mainstream deep learning methods used in planetary gearbox fault diagnosis, and then conduct a comprehensive review on deep learning-enabled planetary gearbox fault diagnosis. Finally, we analyze the challenges in this research and suggest several prospects. Considering the high price of planetary gearbox test rigs, to facilitate investigations on the analyzed challenges, a dataset is introduced. We hope that this paper will provide graduate students, institutions and companies with a preliminary understanding of planetary gearbox vibration characteristics, the principle of deep learning methods and, more importantly, the applications of deep learning methods to planetary gearbox fault diagnosis.

The rest of this review is organized as follows. In section 2, we analyze the vibration characteristics of planetary gearboxes. In section 3, the six mainstream deep learning algorithms are introduced. In section 4, the applications of deep learning methods in planetary gearboxes are reviewed. In section 5, a planetary gearbox dataset is published. In section 6, the research prospects are provided. The conclusions are provided in section 7.

2. Planetary gearbox vibration characteristics and challenges

The transmission diagrams of a fixed-shaft gear set and a planetary gear set with a standstill ring gear are illustrated in figures 2(a) and (b), respectively. The fixed-shaft gearbox is composed of a gear meshing pair, and the two gears rotate around their fixed shafts. Therefore, the vibration of a fixed-shaft gearbox is very simple. A normal gearbox only has the GMF and its harmonics. If a gear fault occurs, sidebands will appear around the GMF and its harmonics with the intervals of the shaft frequency of the faulty gear.

For a planetary gearbox, there are always multiple planet gears that rotate around their unfixed shafts. As shown in figure 1(a), the planetary gear set is composed of a sun gear, a ring gear and four planet gears. The sun gear rotates around its own shaft, the ring gear is stationary, and the four planet gears rotate around their respective shafts and revolve around the fixed sun gear shaft. It can also be seen that these planet gears operate between the sun gear and the ring gear, so they mesh with both of them simultaneously. In addition, there is one rolling bearing in each planet gear, and they also rotate around their own shafts and revolve around the fixed sun gear. Therefore, the vibration of planetary gearboxes exhibits the following characteristics:

(1) The unique structure of a planetary gearbox and its complex operation modes induce more intricate complex frequency components. If a fault occurs in a component of a planetary gearbox, except for the GMF, complex sidebands will exist around the GMF and its harmonics. The intervals of these sidebands with the GMF or its harmonics are various combinations of the shaft rotating frequency of the fault gear and its corresponding characteristic frequency [35]. In addition, the faults on different parts will cause different fault characteristic frequencies [36]. Therefore, there are more complex frequency components in vibration signals compared with that of fixed-shaft gearboxes. In addition, the intervals between the intricate sidebands and GMF harmonics are very close, which makes it difficult to identify them and then localize the fault.

(2) The existence of multiple and time-variant transmission paths adds difficulties in revealing the vibration characteristics of a planetary gearbox. Because there are always several planet gears, except for the ring gear fault, there are multiple transmission paths between a fault component and the sensor mounted on the housing of the planetary gearbox [37]. In addition, with revolving planet gears, the transmission paths are time-variant [10]. Therefore, the vibrations of a faulty gear will be attenuated due to the interference and dissipation effects caused by the paths, making it difficult to reveal the fault characteristics of collected signals.

(3) The vibrations of planet bearings are hard to detect compared to that of the bearings in fixed-shaft gearboxes. Even for fixed-shaft gearboxes, the vibration characteristics of rolling bearings are weaker than those of strong meshing vibrations [38]. For planet bearings, due to the complex meshing vibrations among the planet gears, ring gear and sun gear and the complex transmission paths, the health monitoring of planet bearings is more challenging [9, 39, 40].

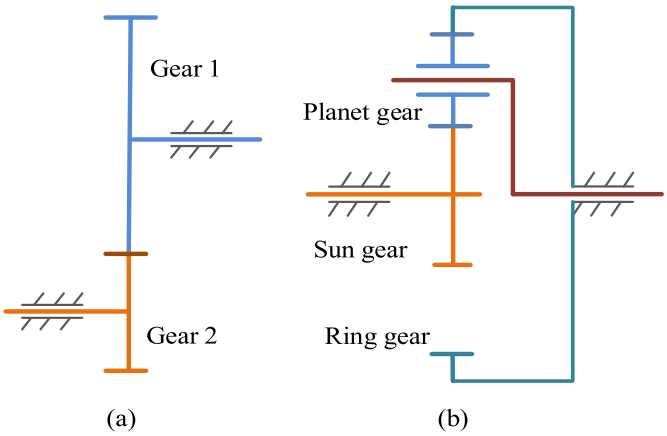

Figure 2. Transmission diagrams of (a) a fixed-shaft gear system and (b) a planetary gear set with a standstill ring gear.

Through the analysis above, it can be found that, compared to that of fixed-shaft gearboxes, planetary gearbox fault diagnosis is more difficult. Researchers have paid much attention to revealing the fault features of planetary gearboxes. With the rapid development of deep learning, it has become a promising approach to recognizing mechanical health states, and it is expected to automatically learn fault information

from collected data. Therefore, various advanced deep learning models have been developed for the fault diagnosis of planetary gearboxes. In the next two sections, the deep learning methods that are commonly used in the fault diagnosis of planetary gearboxes are briefly introduced, and the applications of deep learning to planetary gearbox fault diagnosis are reviewed.

3. Overview of deep learning

To date, various deep learning structures have been investigated, and research on this topic is still rapidly growing. In this section, only some major deep learning methods that have been applied in planetary gearbox fault diagnosis are briefly introduced to make the advanced methods involved in section 4 more understandable.

3.1. AE and its deep models

3.1.1. AE. An AE is composed of two operations, i.e. encoding and decoding [41], as presented in figure 3. These two operations are conducted to form a new representation to reconstruct input data. An encoder is adopted to transform the input $\mathbf{x}$ into a hidden representation $\mathbf{h}$:

$$\mathbf{h}_i = f(\mathbf{W}^{\mathrm{e}}\mathbf{x}_i + \mathbf{b}^{\mathrm{e}}) \tag{1}$$

and the decoder is expressed as follows:

$$\widehat{\mathbf{x}}_i = f(\mathbf{W}^{\mathrm{d}}\mathbf{h}_i + \mathbf{b}^{\mathrm{d}}) \tag{2}$$

where $\mathbf{W}^{\mathrm{d}}$and $\mathbf{W}^{\mathrm{e}}$denote the weight matrices of the decoder and encoder, respectively, $\mathbf{b}^{\mathrm{e}}$ and $\mathbf{b}^{\mathrm{d}}$ denote the corresponding bias terms, and f is an activation function. The model parameters are trained by minimizing the distance between $\mathbf{x}_i$ and $\widehat{\mathbf{x}}_i$ according to

$$\min \frac{1}{N}\sum_{i}^{N}\left\|\mathbf{x}_i - \widehat{\mathbf{x}}_i\right\|_2^2 \tag{3}$$

where N denotes the sample number. An AE is trained through unsupervised learning, and the representation $\mathbf{h}$ is a new and more meaningful representation of the input data.

3.1.2. Denoising AE (DAE). The DAE was proposed by Vincent *et al* [25] and has the same structure as an AE, but the input data are corrupted with noise. Dropout noise/binary masking noise is the most common noise. The addition of noise is realized by randomly setting a portion of the features to zero. Compared with an AE, through the DAE, more robust representation can be learned, and identity transformation learning is prevented.

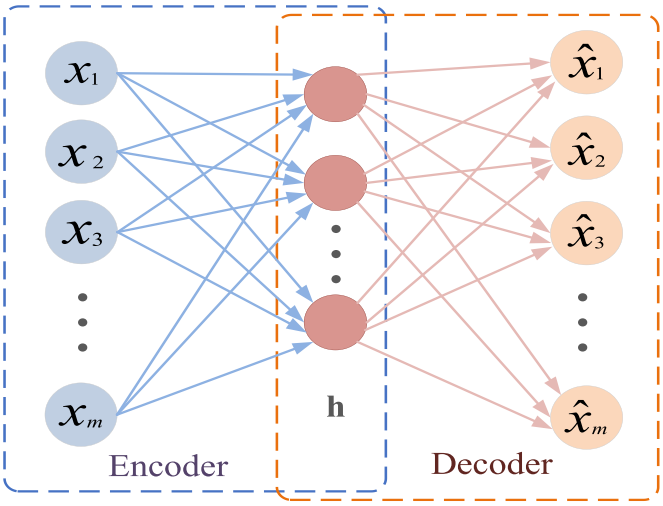

Figure 3. Architecture of AE.

3.1.3. Stacked AE (SAE). An SAE is a deep network constructed by stacking several AEs and is trained by a greedy layerwise method developed by Hinton *et al* [42]. The encoding operation of an SAE with n AEs is expressed as follows:

$$\mathbf{a}_k = f(\mathbf{W}_k^{\mathrm{e}}\mathbf{a}_{k-1} + \mathbf{b}_k^{\mathrm{e}}) \tag{4}$$

where $\mathbf{a}_k$ represents the encoder result of the kth AE and $\mathbf{W}_k^{\mathrm{e}}$ and $\mathbf{b}_k^{\mathrm{e}}$ are the weight matrix and the bias of the kth AE, respectively. If $k = 1$, $\mathbf{a}^0$ denotes the raw data ($\mathbf{a}^0 = \mathbf{x}$). The decoder is calculated by

$$\mathbf{c}_k = f\left(\mathbf{W}_{n-(k-1)}^{\mathrm{d}}\mathbf{c}_{k-1} + \mathbf{b}_{n-(k-1)}^{\mathrm{d}}\right) \tag{5}$$

where if $k = 1$, $\mathbf{c}^0 = \mathbf{a}^n$, and if $k = n$, $\mathbf{c}^n = \widehat{\mathbf{x}}$, which is the final reconstructed data.

3.2. Restricted Boltzmann machine (RBM)

3.2.1. RBM. The RBM is a type of generative stochastic neural network including two types of units, i.e. visible units and hidden units [43]. All units in the two groups are binary. A constraint is applied to make symmetric connections exist between hidden units $\mathbf{h}$ and visible units $\mathbf{v}$, and for each type of unit, no connections exist. The variables of the two groups follow a joint configuration as follows:

$$E(\mathbf{v},\mathbf{h}) = -\sum_i\sum_j w_{ij}v_ih_j - \sum_i b_iv_i - \sum_j a_jh_j \tag{6}$$

where w_{ij} is the weight between hidden unit h_j and visible unit v_i and a_j and b_i are the bias terms of the hidden and visible units, respectively. The joint distribution of the units is obtained according to

$$p(\mathbf{v},\mathbf{h}) = \frac{\exp(-E(\mathbf{v},\mathbf{h}))}{Z} \tag{7}$$

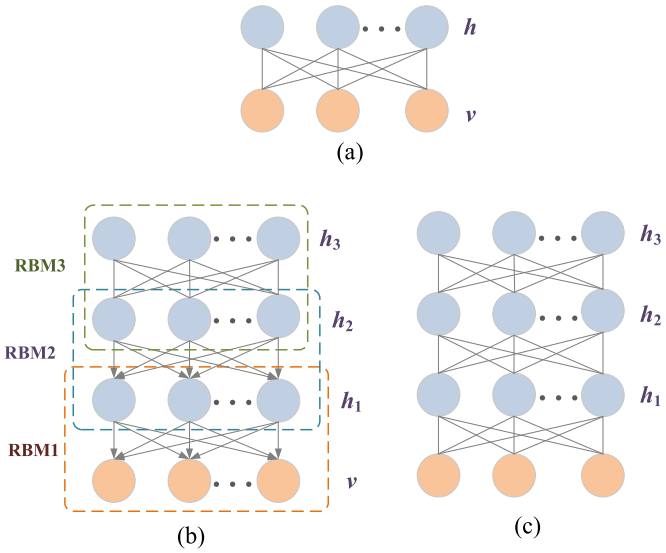

Figure 4. Architectures of (a) RBM, (b) DBN and (c) DBM.

where $Z = \sum_{\mathbf{v},\mathbf{h}} \exp(-E(\mathbf{v},\mathbf{h}))$ is the partition function. The conditional probabilities of $\mathbf{v}$ and $\mathbf{h}$ are computed as:

$$p(v_i = 1 \,|\mathbf{h}) = \sigma\left(\sum\nolimits_i w_{ij} h_j + b_i\right) \tag{8}$$

$$p(h_j = 1 \,|\mathbf{v}) = \sigma\left(\sum\nolimits_i w_{ij} v_i + a_j\right) \tag{9}$$

where $\sigma(x)$ is a logistic function, i.e. $\sigma(x) = 1/[1 + \exp(x)]$. The RBM is optimized by maximizing the joint probability. The weight matrix is learned by contrastive divergence.

3.2.2. Deep models of RBM. A deep belief network (DBN) [44] is a deep learning model formed by stacking several RBM modules, in which the input of the $(l+1)$th layer (visible units) is the output of the lth layer (hidden units), as demonstrated in figure 4(b). A DBN can also be optimized in a greedy layer-wise manner. Fine-tuning is further conducted on the parameters of the model according to the label information by adding a softmax layer.

A DBM [26]is also a deep model. However, rather than a single layer, several hidden layers are adopted to form a hierarchy of layers, as shown in figure 4(c). There are only connections between subsequent layers, i.e. there are no connections between nonadjacent layers or within layers. In addition, different from a DBM model, a DBN is a directed model.

3.3. CNN and its variant

3.3.1. CNN. CNNs were first proposed by LeCun *et al* [27] and their excellent feature learning ability has been demonstrated in computer vision applications [45, 46] and have been widely investigated in fault diagnosis [47–49]. A typical CNN model is usually composed of convolutional layers, pooling layers and fully connected layers, as presented in figure 5. Feature learning is achieved with a sequential convolution

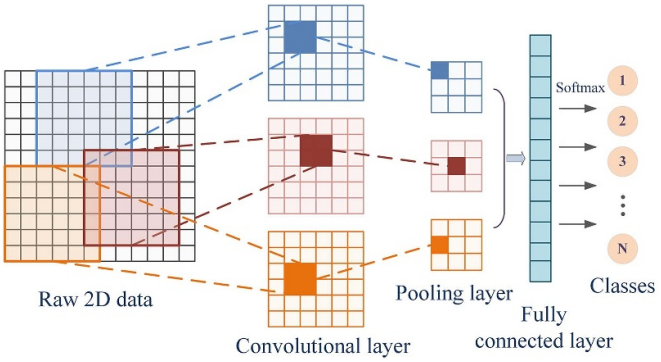

Figure 5. Architecture of 2D CNN.

operation between kernels in each convolutional layer. For the lth layer, which has N kernels and M feature maps, the jth output feature map x_j^l is expressed as:

$$\mathbf{x}_j^l = f\left(\sum_{i=1}^{M} \mathbf{x}_i^{l-1} * \mathbf{k}_{ij}^l + \mathbf{b}_j^l\right), j = 1, \ldots, N \tag{10}$$

where $\mathbf{x}_j^l$ is the ith input feature map, f is the activation function $\mathbf{k}_{ij}^l$ is the jth kernel connected to the ith input feature map, $\mathbf{b}_j^l$ is the bias term, and $*$ is the convolution operation.

After a convolution layer, a pooling layer is usually added to reduce the dimension of the feature map. Max pooling and average pooling are two common pooling operations. Max pooling selects the largest value in a local region, while average pooling computes the average value. By using a pooling operation, the computation time cost is reduced, and the robustness to small variations is improved. Finally, fully connected layers are added after several convolutional layers and pooling layers to make a prediction.

3.3.2. Residual network (ResNet). ResNet is a variant of CNN that is constructed by stacking a series of residual building blocks [50]. A residual building block usually consists of convolutional layers, rectified linear unit (ReLU) activation functions, batch normalization operations and one identity shortcut [51]. Different from common CNN models, identity shortcut layers are added in ResNet so that the gradient can easily flow back to the shallow layers in the training process. A feature is handled by two convolutional layers as follows:

$$y\left(\mathbf{x}_i^{l-1} \,|\theta^l\right) = f\left(\mathbf{x}_i^{l-1} * \mathbf{k}^{l_1} + \mathbf{b}^{l_1}\right) * \mathbf{k}^{l_2} + \mathbf{b}^{l_2} \tag{11}$$

where $\theta^l = \left\{\mathbf{k}^{l_1}, \mathbf{b}^{l_1}, \mathbf{k}^{l_2}, \mathbf{b}^{l_2}\right\}$ are the parameters that need to be trained in the lth residual block. The output of the residual block is computed as:

$$\mathbf{x}_i^l = f\left[y(\mathbf{x}_i^{l-1}|\theta^l) + \mathbf{x}_i^{l-1}\right],\ l = 2, 3, \ldots, L. \tag{12}$$

A deep neural network structure is obtained by stacking several such residual blocks, in which the input of the following layer is the output of the former block.

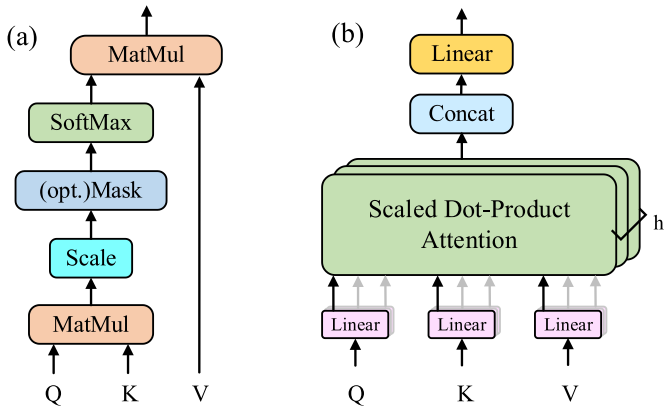

Figure 6. (a) Scaled dot-product attention, and (b) MSA.

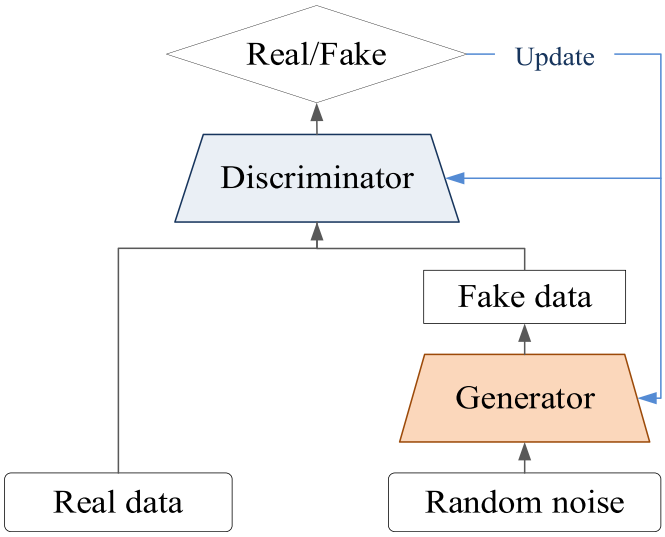

Figure 7. Flowchart of GAN.

3.4. Transformer

The typical transformer mainly consists of multi-head self-attention (MSA), layer normalization (LN), feed-forward neural network (FNN), and residual connection (RC) [52–54]. Among them, as presented in figure 6, MSA aims to conduct multiple attention relation calculations and enables it to focus on features of multiple positions simultaneously.

The self-attention can be calculated as follows:

$$\text{Attention}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{n}}\right)V \tag{13}$$

where $Q \in R^{m\times n}$, $K \in R^{m\times n}$ and $V \in R^{m\times n}$ represent query, key and value respectively; $\sqrt{n}$ represents a scale factor. MSA allows parallel calculation of self-attention, and then the attention relations are converted into a set of outputs by the transformation matrix W^{o}.

$$\text{MultiHead}(Q,K,V) = \text{Concat}(\text{head}_1,\ldots,\text{head}_i)\,W^{\text{o}} \tag{14}$$

$$\text{head}_i = \text{Attention}\left(QW_i{}^{\text{Q}}, KW_i^{\text{K}}, VW_i^{\text{V}}\right) \tag{15}$$

where W_i^{Q}, W_i^{K} and W_i^{V} are the ith liner projection of input embedding X to achieve the query, key and value. W^{o} is the liner projection on concatenated multi-head.

Then, the output of MSA is fed into FNN, and the FNN operates on each position independently. A transformer is composed of several transformer blocks. A block includes a MSA module and a feed-forward module, and both of them utilize RC and LN to obtain the final output of the model.

3.5. Generative adversarial network (GAN)

A GAN is usually composed of a generator $G(\cdot)$ and a discriminator $D(\cdot)$ [55] as shown in figure 7. The former converts a random noise vector into a space with distribution $p_{\mathbf{z}}(\mathbf{z})$. The latter attempts to distinguish the actual data with distribution $p_{\text{data}}(\mathbf{x})$ and the data generated based on $G(\cdot)$. The two models play a game with each other following the value function [56]:

$$\begin{aligned}&\min_G \max_D V(D,G)\\ &= E_{\mathbf{x}\sim p_{\text{data}}(\mathbf{x})}[\log D(\mathbf{x})] + \text{E}_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z})}[\log(1-D(G(\mathbf{x})))]\end{aligned} \tag{16}$$

where E represents the empirical estimate of the expected probability value. Here, the min and max functions are applied to minimize the term $\log(1-D(G(\mathbf{x})))$ and maximize the term $\log D(\mathbf{x}) + \log(1-D(G(\mathbf{x})))$, respectively. In detail, the discriminator attempts to maximize the differentiation of the generated data $\mathbf{z}$ and the actual data $\mathbf{x}$, while the generator attempts to minimize the odds of the generated data being detected by the discriminator.

3.6. Graph neural network (GNN)

GNN is first proposed by Scarselli *et al* [57], which aims to exploit graph theory to construct neural networks. A graph can be expressed as:

$$G = G(V,E) \tag{17}$$

where V and E represent the sets of nodes and edges, respectively. If $v_i \in V$ is a node and $e_{i,j} = (v_i, v_j) \in E$ is an edge between v_i and v_j, the neighborhood of the node is defined as $M(v) = \{u \in V | (v,u) \in E\}$. An adjacency matrix $A \in R^{M\times M}$ can be used to describe a graph, where M is the number of nodes. In particular, if $\{v_i, v_j\} \in E$ and $i \neq j$, $A_{i,j} = 1$. Otherwise, $A_{i,j} \neq 1$.

In practical, the notes are known as features, represented as feature matrix $X \in R^{M\times c}$ where c is the dimension of feature matrix. The degree matrix $D \in R^{M\times M}$ is a diagonal matrix, which is calculated as $D_{ii} \in \sum_{j=1}^{N} A_{ij}$. The graph can

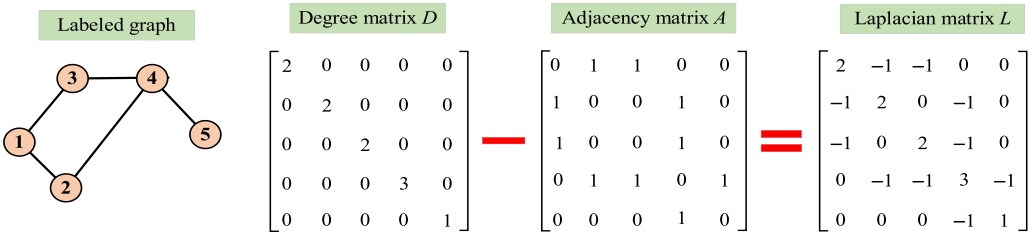

Figure 8. The relationship between the degree matrix, the adjacency matrix and the Laplacian matrix.

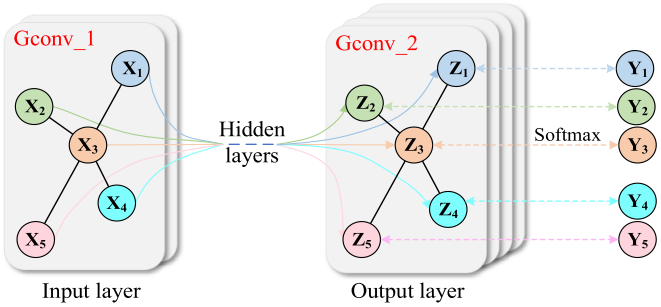

Figure 9. Two-layer GCN.

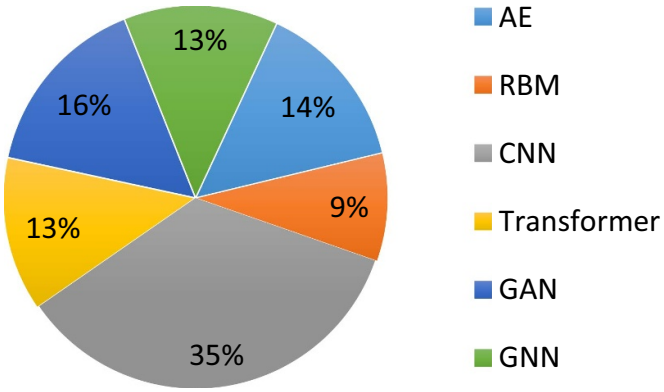

Figure 10. Proportions of selected deep learning methods in planetary gearbox fault recognition.

be described as the Laplacian matrix LM. The relationship between the degree matrix, the adjacency matrix, and the Laplacian matrix is shown in figure 8. The Laplacian matrix is defined as

$$\mathrm{LM} = D - A. \tag{18}$$

The input of GNN includes the feature matrix $\mathbf{X}$ and the adjacency matrix $\mathbf{A}$. The output of GNN is achieved by the general forward propagation as defined by

$$\mathbf{Z}_G = \mathrm{softmax}(G_k(\ldots, G_2(\mathbf{A}, G_1(\mathbf{A}, \mathbf{X})))) \tag{19}$$

where G is the GNN operation, and k is the layer number.

The graph structure provides the node values and the relationships of nodes, so GNN can capture more information than regular data. Owing to this merit, graph CNN (GCN) [58, 59], graph recurrent neural network [60], graph auto-encoder [61] and so on have been developed. Among them, GCN is the most common and effective method in fault diagnosis, and a two-layer GCN is shown in figure 9.

4. Applications of deep learning in planetary gearbox fault diagnosis

In this section, the applications of six deep learning methods in planetary gearbox fault recognition are reviewed. These papers are searched by two databases, i.e. Web of Science and Google Scholar, and are then refined by checking on the topic and the quality of publication journals. The contribution of each technique to planetary gearbox fault recognition is listed in figure 10. It is seen that the most popular method in planetary gearbox fault recognition in recent five years is CNN, i.e. 35%, followed by GAN. It should be noted that although the proportions of transformer and GNN methods are all only 13%, most of papers are published between 2022 and 2023.

4.1. AE and its deep models

An AE and its deep models can extract the hidden representations from collected data automatically, and they have been successfully utilized in the fault diagnosis of planetary gearboxes. Applications of AE-based methods are listed in table 1. Jia *et al* [62] applied an SAE to learn the features from the frequency spectra of vibration signals and implemented planetary gearbox fault classification. This is one of the earliest applications of an SAE to machinery fault diagnosis. In this method, three AEs were stacked to automatically learn the distinguished representations from the frequency spectra rather than manually extracting features. To implement an end-to-end diagnosis, they further developed a normalized sparse AE (NSAE)-based local connection network (LCN), named NSAE-LCN [63]. First, the NSAE was applied to extract the effective features from the raw collected data. The difference between an NSAE and a sparse AE lies in the fact that ReLUs were adopted instead of the common sigmoid function, the bias term was avoided, considering that it was nonessential in feature learning, the L1 norm was used to replace the Kullback–Leibler (KL) divergence function, and in the cost function, a soft orthonormality constraint was added. Second, a LCN was proposed with the NSAE, in which a feature layer was added to average the extracted features. The NSAE-LCN

Table 1. List of references on the AE-based fault diagnosis methods.

References	Method	Input data	Fault components
Jia *et al* [62]	SAE	Frequency spectra of vibration signals	4 sun gears, a planet gear, and a planet bearing
Jia *et al* [63]	NSAE-LCN	Raw vibration signals	3 sun gears, 2 planet bearings, and a planet gear
Yang *et al* [65]	ISAE	Raw vibration signals and current signals	5 sun gears
Zhang *et al* [64]	DSAE	Raw vibration signals	4 sun gears, and 4 rolling bearings
Wang *et al* [66]	BNAE	Raw vibration signals	3 sun gears, 3 pinions, and 3 coupled faults on sun gears and pinions, respectively
Yu *et al* [67]	SDAE	Raw vibration signals	5 gears (the types were not mentioned)
Saufi *et al* [68]	SSAE	Kurtogram images of vibration signals	4 gears (the types were not mentioned)
Shao *et al* [70]	MSAE	Raw vibration signals	8 sun gears
Chen *et al* [71]	FSAE	Raw vibration signals	5 sun gears
Shao *et al* [72]	SWAE	Raw vibration signals collected from three directions	4 sun gears, and 4 rolling bearings

overcomes the limitation of an SAE in feature extraction from raw signals caused by the time-variant property of vibration signals.

To better learn features in a supervised way, Yang *et al* [65] proposed an improved supervision AE (ISAE), in which a new supervision penalty term based on linear discriminant analysis was introduced and a *k*NN classifier was combined with an AE. In addition, the vibration and current signals of the planetary gearbox were fused at the data end, and the fused data were fed into the ISAE for feature learning and classification. They found that the performance was better than that of the conventional SAE, and the recognition rate was higher than that produced by a single vibration or current signal. Zhang *et al* [64] developed a discriminative sparse AE (DSAE), as presented in figure 11, in which a discriminator was constructed to incorporate labels and an AE was introduced to mine the intrinsic correlations of the data. Wang *et al* [66] developed a batch-normalized AE (BNAE) model, in which batch normalization was introduced into each layer of the AE model to extract more robust features, and then a softmax layer was applied for fault classification. Yu *et al* [67] proposed a hybrid model by combining a stacked denoised AE (SDAE) and gated recurrent unit neural network for the fault diagnosis of planetary gearboxes. A softmax layer was applied to identify the features obtained by the hybrid model in a supervised manner, which was proven to be more robust to speed fluctuations.

Although the methods mentioned above have gained good performance in planetary gearbox health state recognition, the hyperparameter values which highly affect the effectiveness of AE models are all determined by manual selection. In the following papers, meta-learning methods have been introduced to adaptively determine the hyperparameters. Saufi *et al* [68] utilized the particle swarm optimization (PSO) algorithm [69] to optimize the hyperparameters of sparse AEs, including the hidden note number, weight decay, sparsity, and sparsity penalty term weight. In addition, different from the publications mentioned above, fast kurtogram images of the planetary

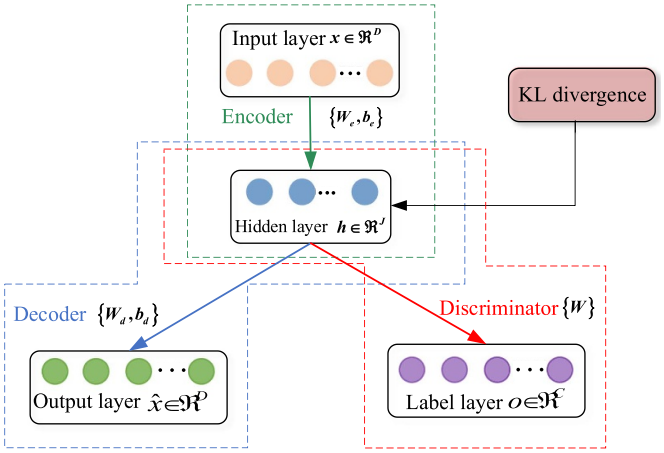

Figure 11. Architecture of DSAE. © [2022] IEEE. Reprinted, with permission, from [64].

gearbox vibration signals were generated as input for a stacked sparse AE (SSAE). It was proven that this method performed better than that of the conventional SSAE and can produce a promising result under limited data. Considering the merit of the Morlet wavelet function in nonlinear mapping, which is more suitable for nonstationary signals, it has been applied to modify an SAE [73]. However, the parameters of the wavelet function were all selected manually. Shao *et al* [70] proposed a modified SAE (MSAE), in which the parameters of the wavelet activation function were determined by the fruit fly optimization algorithm (FFOA) to better match the characteristics of vibration data. The introduction of meta-learning methods makes the AE-based methods more adaptive, and releases burden of manual selection of the hyperparameters.

Each variant of AE has its own merits, so combining those methods is an effective approach to further improve the recognition rate. A fused SAE (FSAE) method was proposed by Chen *et al* [71], which was implemented by combining a sparse AE and a contractive AE (CAE), in which the key

parameters and specific location of the two AE models were optimized by the quantum ant colony algorithm (QACA). Li *et al* [74] constructed a hybrid deep AE model by combining a DAE, a DAE with a linear decoder (DLAE), a sparse AE and a sparse AE with a linear decoder (SLAE), in which the combination strategy was crucial to the decision process. An enhanced weighted voting method was designed to fuse the information of multiple channels to give an appropriate decision. In the voting method, the key class-specific thresholds were determined by the beetle antennae search (BAS) algorithm. Different from the fusion strategy of combining multiple AE models, Shao *et al* [72] proposed a multisensory data fusion strategy, in which the signals of multiple sensors were fed into a stacked wavelet AE (SWAE), and decisions were made by considering the information from multiple channels. The Morlet wavelet function was adopted as the activation function for multisensory data fusion, and an improved voting decision fusion method was developed to obtain a better collaborative result. The experiments proved that the fusion method performed better than the conventional data-level (the different data are put into a matrix/vector according to a rule, and then the results are fed into the model) and feature-level (the data are fed into the model, and the learned representations are fused according to a rule) fusion methods. Compared to a standard AE model, fusing the results of several different variants of AE can take advantage of the merits of different variants, and fusing the multisensory data by a unique AE can obtain more abundant fault information, which results in more accurate recognition results.

Although AE and its variants have achieved good results in planetary gearbox fault recognition, they also have the disadvantages compared to other deep learning methods. Their inherent property results in that the encoded data is too similar to the input data, i.e. the learned features are highly correlated to the input training data, which makes them hardly extract meaningful features and thus poor generalization ability.

4.2. RBM

AE-based deep models provide an end-to-end diagnosis of planetary gearboxes, i.e. the effective representations are learned from raw vibration signals or spectra and recognized by these models. The deep models constructed by an RBM always work together with feature extraction methods.

Lu *et al* [75] applied compressed sensing to compress the vibration signals and then fed the result features into a DBN in which the node number in hidden layers and learning rate were determined via chaotic quantum PSO. In [76], the vibration signals of planetary gearboxes were decomposed by ensemble empirical mode decomposition (EEMD), and then the instantaneous frequency and amplitude of the intrinsic mode functions (IMFs) were tracked via the Teager–Kaiser energy operator. The statistical features of six group signals, such as raw signals, IMFs and frequency spectra, were calculated and refined for a DBN in which the neural number of hidden layers was determined by PSO. Han *et al* [77] extracted the multiscale permutation entropy and wavelet packet energy entropy features and then fed them into a DBN in which an iterative optimization method was applied to avoid overfitting. Zhang *et al* [78] constructed a hybrid data fusion DBN for planetary gearbox fault recognition, in which multiple signals from different sensors were preprocessed using signal processing methods such as low-pass filtering, the time synchronous average [79] and fast Fourier transform (FFT), and then the features were calculated with the resulting signals. A DBN was utilized to identify the health states with the fused features. Qin *et al* [80] applied an optimized Morlet wavelet transform (WT), soft thresholding and the kurtosis index to obtain the impulse components of planetary gearbox data. The statistical features of raw signals and the impulse components were calculated and fed into a DBN in which an improved logistic sigmoid was adopted to avoid the vanishing gradient problem. The above methods are all designed for the fault information from the features obtained by signal processing methods, which makes them highly rely on the quality of extracted features and expert knowledge.

Researchers have also attempted to design optimized DBN models to enable the models to extract effective features automatically. Yang *et al* [81] proposed a joint pairwise graph embedded sparse DBN (J-PBN) in which the pairwise graph and sparse representations were added before the DBN to better mine the distinguished features, and the parameters were trained with partial least squares in the supervision learning process. To implement fault recognition under different working conditions, Xing *et al* [82] proposed a distribution-invariant DBN. The model consists of a fully connected RBM layer (FCRBM), local connected RBM (LCRBM), and an RBM layer with mean discrepancy maximum (MDM-RBM). The LCRBM was utilized to extract the shift-invariant features directly from the time-domain signals, the FCRBM further made the features more sensitive to the health states, and the MDM-RBM layer was used to make the model better learn distribution-invariant features from the signals under different operation conditions.

Compared to AE-based models, artificial features are usually extracted from raw signals through various signal processing methods and then fed into RBM-based models to realize fault recognition. Researchers have mainly focused on how to design more appreciated features through signal processing methods, adaptively determining the hyperparameter values, and designing more advanced models to attempt to learn the representations from raw signals. Therefore, the disadvantage of RBM in planetary gearbox fault recognition is that end-to-end fault diagnosis cannot be realized and thus the results depend on the signal processing methods and expert knowledge.

4.3. CNN

The applications of CNNs in planetary gearbox fault diagnosis are listed in table 2. CNN models can learn complex and useful representation features from raw collected signals or two-dimensional (2D) representations of raw signals. They have been widely investigated in planetary gearbox fault recognition. According to their architecture, they are divided into 1D and 2D CNN models.

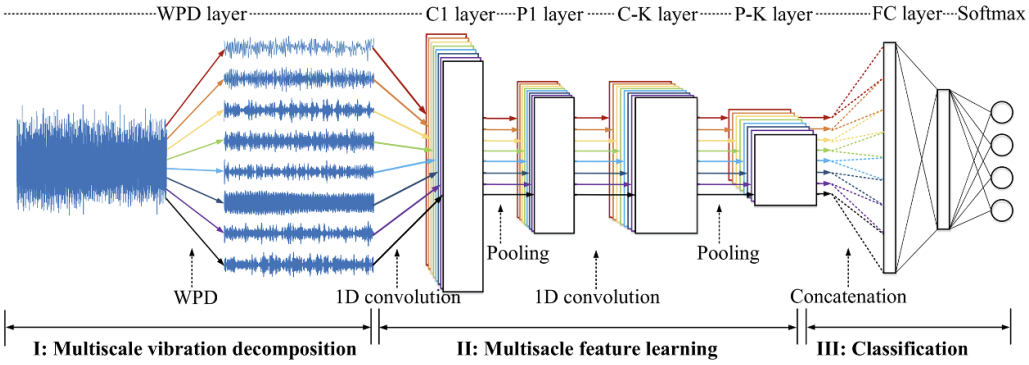

Figure 12. Flowchart of the WPT-HSCNN model. © [2023] IEEE. Reprinted, with permission, from [84].

Jing *et al* [83] applied a 1D CNN to analyze the vibration signals of a planetary gearbox, which was one of the earliest publications on planetary gearbox fault diagnosis. The model includes one convolutional layer, one pooling layer, one fully connected layer, and a softmax layer. Raw time-domain signals, frequency spectra and time-frequency representations (TFRs) were all analyzed by the CNN model, and the results validated that based on the frequency spectra, the recognition accuracy was the highest and it performed better than that obtained by shallow classifiers combined with traditional statistical features. Huang *et al* [84] proposed a novel wavelet packet decomposition (WPT) multiscale CNN (WPT-MSCNN) model as shown in figure 12. In the WPT layer, the vibration signals of a wind turbine gearbox were decomposed into several components, and a hierarchical CNN was used to handle those components. The CNN learned multiscale features from the decomposed components. Han *et al* [85] proposed an enhanced CNN (ECNN) for planetary gearbox fault diagnosis, in which a 1D CNN was used to learn health information from 1D time-domain signals, and the receptive field can be enlarged by fused dilated convolutional layers to further obtain long distance dependencies. The raw collected time-domain data were fed into the ECNN model directly, and the recognition rate was higher than that obtained by the standard CNN model. To realize the fault recognition of planetary gearboxes under variable speeds, Wang *et al* [86] proposed an adaptive normalized CNN (ANCNN) model that was implemented with the Teager calculated order method. For the ANCNN, batch normalization was designed to avoid the feature distribution difference caused by speed fluctuations, and PSO was utilized to determine the hyperparameters. In addition, Teager calculated order spectra were applied to avoid spectrum smearing caused by variable conditions. This method was proven to be capable of analyzing the signals of a planetary gearbox under variable speeds. Chang *et al* [87] applied decoupled operators to the convolutional layers and fully connected layers of a 1D CNN (termed 1D-FDCNET) to improve the robustness to complex conditions and noise by using the merits of the decoupled operators in separating the interclass variation and semantic difference.

1D CNN can learn features directly from raw collected signals, and thus provide an end-to-end fault diagnosis. Compared to AE-based models which are also able to deal with raw data as mentioned in section 4.1, 1D CNN is not subject to the time-shift property of time-domain data, and performs better when the 1D time-domain data is fed as input of models. However, some information, such as the time-frequency domain which directly reflects the instantaneous resonance frequency or fault characteristic frequency, cannot be extracted directly by the 1D CNN via time-domain data.

Inspired by the large advantages of CNN models in image recognition, researchers have also proposed appropriate 2D CNN models for planetary gearbox fault recognition. TFA methods, such as short-time Fourier transform (STFT) [88, 89], WT [90], and Wigner–Ville distribution [91], are common techniques to convert 1D time-domain signals into 2D TFRs. Chen *et al* [92] constructed a 2D CNN for planetary gearbox fault diagnosis, which consisted of three convolutional layers and their corresponding pooling layers, one fully connected layer, and a softmax classifier. They applied this model to analyze the TFRs generated by the discrete WT (DWT), STFT and S transform (ST), and the results showed that the recognition rate based on the DWT was the highest. Different from TFA methods, Wang *et al* [93] used a recurrence plot to generate 2D representations of angular data, and then the resulting 2D representations were fed into a 2D CNN for feature extraction and classification. This preprocessing method was claimed to be more suitable for nonstationary conditions and thus was beneficial for CNNs to extract effective features. Emmanuel *et al* [94] applied a Blackman function to transform the frequency spectra of planetary gearbox signals to 2D images and then fed them into a CNN for health state recognition. Cyclic spectra have been investigated recently in bearing fault diagnosis and were proven to be more effective in identifying the fault characteristic frequency harmonics of rolling bearings [95–97]. Yang *et al* [98] adopted cyclic wavelet spectrum analysis to produce 2D images of planet bearing signals and fed them into an improved CNN model in which an extreme learning machine was adopted instead of a softmax classifier. Jiang *et al* [99] fed 2D bispectrum images into a CNN model to realize the fault diagnosis of a planetary gearbox. The 2D images generated by TFA methods or signal processing methods are obtained before they are fed into CNN models. Kim *et al* [100] proposed an improved CNN model, an

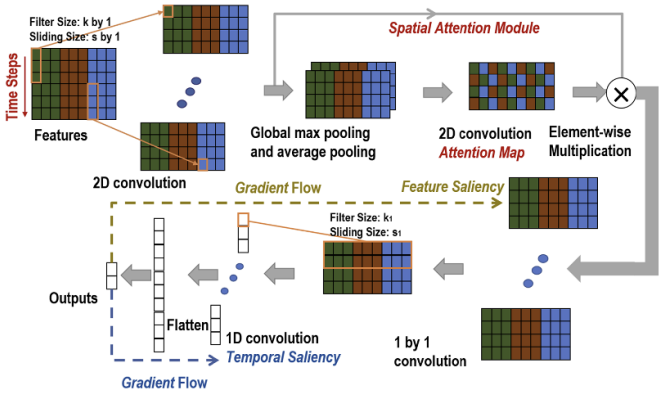

Figure 13. Flowchart of the DEFT-CNN model. © [2021] IEEE. Reprinted, with permission, from [101].

embedded health-adaptive time-scale representation (HTST), in which the HTST module was added before the convolutional layer, and rather than fixed basis functions, the parameters of HTST were learned adaptively. This model was validated to produce better results than those obtained by a CNN combined with STFT- and WT-based TFRs. Zhu *et al* [101] developed a decouplEd feature-temporal CNN (DEFT-CNN) as shown in figure 13. In the model, the feature information and temporal information were captured by convolutional layers, and the spatial attention module was added to make the model focus on important features. In addition, the Grad-CAM was used to obtain the features and temporal saliency maps to visualize the important features in the input 2D feature matrix.

Compared to 1D CNN, 2D CNN can learn richer features such as time-frequency information and cyclic spectrum correlation as listed in table 2. However, the performance of these 2D models rely on the quality of the input images. In addition, when the time-domain data is mapped into the images, it is inevitable that some information will be lost during the conversion process.

The 1D models or 2D models mentioned above can only learn the fault information from one type of representation of one type of signal, such as raw signals, frequency spectra and TFRs of the collected vibration signals. Theoretically, more health information will be learned if the model can learn the features from more representations and more types of data, such as the vibration signals measured by accelerometers installed in different positions or directions. Chen *et al* [102] fused the vibration signals collected by sensors installed in the horizontal and vertical directions at the data level and then fed them into a CNN model to realize planetary gearbox fault recognition, which was validated to produce higher accuracy due to the fusion operation. Guo *et al* [103] proposed a multitask parallel CNN with reinforced input (RI-MPCNN) for wind turbine planetary gearbox fault recognition. One wavelet packet transform matrix (WPTM) was obtained by applying WPT to the signals collected by accelerometers mounted at different positions, and a domain knowledge map (DKM) was constructed with the fault characteristic frequencies and the sidebands around them and operation condition information. Finally, the WPTM and the DKM were fused into one overall information array and then fed into sub-CNNs for feature extraction. Multiple vibration signals, the calculated characteristic frequencies and the operation conditions were all considered in the model, which was expected to generate more fault information by combining the signals with expert knowledge and was validated to be more effective than that of the conventional CNN model. Jiao *et al* [104] proposed a deep coupled dense convolutional network to realize information fusion and feature learning. For this model, a dense connection was used for the feature extraction operation, which was expected to alleviate the loss of fault information and avoid gradient vanishing. In addition, two parallel streams with the same structure were constructed and fused before the softmax layer, and they were adopted to learn the health information from vibration and encoder data. The encoder data are the rotating frequency information, in which pulses are produced equi-angularly with the rotation of a planetary gearbox. If a fault occurs, the stiffness variation will lead to a fluctuation of the rotation speed, so the fault information is also carried by the encoder signal. Encoder signal-based frequency identification methods have also recently been adopted for planetary gearbox fault diagnosis [113–116]. To evaluate the importance of each signal, Li *et al* [105] proposed an adaptive channel weighted CNN (ACW-CNN), as shown in figure 14. Here, ACW layers can adaptively learn the importance of sensors by assigning different weights to each sensor. In the experiment, the acoustic emission signals, vibration signals and sound signals were fed into the model for feature extraction, and it was confirmed that the ACW can improve the fusion result. Xu *et al* [117] proposed a multiscale CNN coupled with multiattention capability (HMS-MACNN), in which raw vibration signals and their corresponding TFRs were fed into parallel CNN for feature extraction and fused in an improved multiattention module through a new weighted soft voting method.

Fusion CNN can take advantage of multisensory data or the data from different domains, and thus more fault information can be extracted for fault recognition. For example, as mentioned in section 2, the time-variant transmission paths will alleviate fault features. However, the signals collected from different locations on the planetary will be complementary due to the time-varying relative locations of fault and the sensors and the different phase differences. In addition, fusion CNN makes it possible to extract features from 1D time-domain data and the conversion images by designing different channels with 1D and 2D CNN structures.

ResNet has the advantage of optimizing the parameters in gradient back-propagation due to the unique design of identity shortcuts, which performs better in avoiding gradient vanishing or exploding and is more suitable for deeper networks. Therefore, ResNet has been investigated in planetary gearbox fault diagnosis [106–108], and also in data fusion. For example, Zhao *et al* [109] proposed a multiple wavelet regularized deep residual network. The signals were transformed into 2D wavelet coefficient matrices by WPT, and the 2D matrices were divided into a primary group generated by the wavelet basis function and an auxiliary group produced by

Table 2. List of references on the CNN-based fault diagnosis methods.

Method	References	Input data	Fault components
1D CNN	Jing *et al* [83]	Frequency spectra of vibration signals	6 planet gears
	Huang *et al* [84]	Multiscale components by WPT	2 shaft, and 1 planet gear
	Han *et al* [85]	Raw vibration signals	4 sun gears, and 3 planet bearings
	Wang *et al* [86]	Raw vibration signals	3 sun gears, 3 planet gears, 3 ring gears, and 2 conditions with fault sun gear and ring gear
	Chang *et al* [87]	Frequency spectra of vibration signals	1 gear, and 4 rolling bearings
2D CNN	Chen *et al* [92]	TFRs of vibration signals obtained by DWT	4 planet gears and 1 sun gear
	Wang *et al* [93]	2D images of vibration signals obtained by RP	1 sun gear, 1 planet gear, and 1 planet carrier
	Emmanuel *et al* [94]	2D images of vibration signals obtained by Blackman	1 sun gear, 1 planet gear, and 1 condition with fault sun gear and planet gear (Note: the data is simulated)
	Yang *et al* [98]	2D Wavelet cyclic spectra of vibration signals	3 planet rolling bearings
	Jiang *et al* [99]	2D bispectra of vibration signals	2 sun gears, 2 planet gears, and 2 ring gears
	Kim *et al* [100]	TFRs of vibration signals obtained by WT	5 planet gears
	Zhu *et al* [101]	2D matrix of time-domain, frequency-domain and time-frequency-domain features	5 planet gears
Fusion CNN	Chen *et al* [102]	Raw vibration signals from two sensors	4 sun gears
	Guo *et al* [103]	TFRs of vibration signals from different sensors obtained by WPT	4 gears, and 3 rolling bearings
	Jiao *et al* [104]	Raw vibration signals and encoder signals	8 planet gears
	Li *et al* [105]	Raw vibration signals, acoustic emission signals and sound signals	4 gears, and 3 rolling bearings
ResNet	Zhang *et al* [106]	TFRs of vibration signals by WPT	4 sun gears, and 4 planet bearings
	Zhao *et al* [107]	TFRs of vibration signals by WPT	4 sun gears, and 4 planet bearings
	Zhang *et al* [108]	Raw vibration signals	4 sun gears, and 4 planet gears
	Zhao *et al* [109]	TFRs of vibration signals by WPT	4 sun gears, and 4 planet bearings
	Huang *et al* [110]	TFRs of vibration signals by WT	4 gears
	Zhang *et al* [111]	TFRs of vibration signals by WPT	4 sun gears, and 4 rolling bearings
	Xie *et al* [112]	RGB images of multisensory data by PCA	4 gears, and 4 rolling bearings

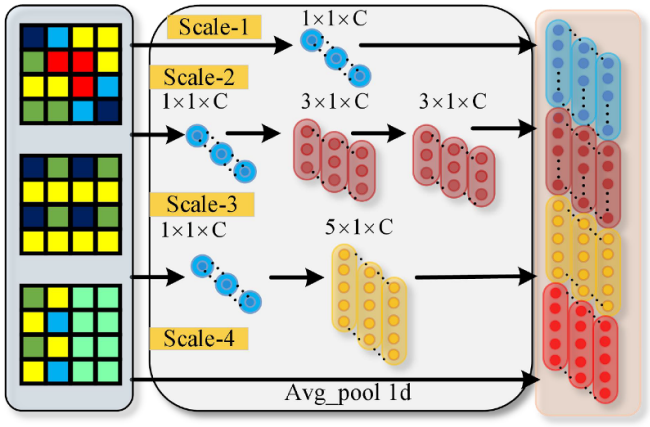

Figure 14. Architecture of ACW-CNN model. © [2020] IEEE. Reprinted, with permission, from [105].

the other type of wavelet basis functions. The primary group was adopted to train a deep network, and the other group was used to optimize a deep network that shared the same weights with the former one. Finally, cross-entropy loss functions were applied to optimize the two deep networks. Huang *et al* [110] developed a ResNet-based planetary gearbox fault diagnosis method in which the wavelet TFRs were generated as input and a channel attention module was adopted to enhance the important features. It should be noted that different from that in the above publications, this method was implemented in a cloud environment to reduce the computational burden. Zhang *et al* [111] also calculated WPT-based wavelet coefficients as input, but a band attention module was designed to highlight the low but crucial frequency band, which was realized by frequency band division and channel attention. ResNet has also been investigated in multisensory fusion-based fault diagnosis. Xie *et al* [112] utilized principal component analysis (PCA) to deal with multiple signals to obtain three-channel RGB images and then fed them into a ResNet model for feature learning and classification.

To the best of our knowledge, CNNs are the most popular deep learning method for the fault diagnosis of planetary gearboxes. This can also be concluded from the number of CNN models reviewed above. Most of the publications above designed CNN models to learn hidden representations from raw signals or 2D representations obtained from raw data owing to the merit that CNNs can capture the shift-variant properties of raw collected signals. In addition, by weight sharing, the number of parameters to be optimized is reduced, convergence is effectively accelerated, and overfitting is avoided.

However, various CNN models are designed by 2D images obtained by TFA methods or other 1D to 2D conversion methods, but the parameters of their bases are determined before being learned by CNN. The independence between CNN model and these parameters will limit the learning ability. Integrating the 2D image generation techniques to the CNN optimization may be a good solution to improving the accuracy of planetary gearbox fault recognition. In addition, CNN models are subject to low global information learning ability that is one huge advantage of transformer discussed in the next section. Last, the interpretation of CNN models has been studied in bearing fault diagnosis [119, 120], but how to extend these methods to planetary gearbox fault recognition remains challenging. It is expected to be exploited to make these models more transparent by combining the modulation characteristics planetary gearbox vibration data to identify whether the information learned from the models is caused by faults or interference factors such as installation.

4.4. Transformer

Although CNN has exhibited good feature extraction ability in machinery fault diagnosis, it has limited local receptive field which results in poor global feature mining ability. On the contrary, transformer can capture the association relationships from raw data by the global attention mechanism. Owing to the merit of transformer, it has gradually attracted attention in

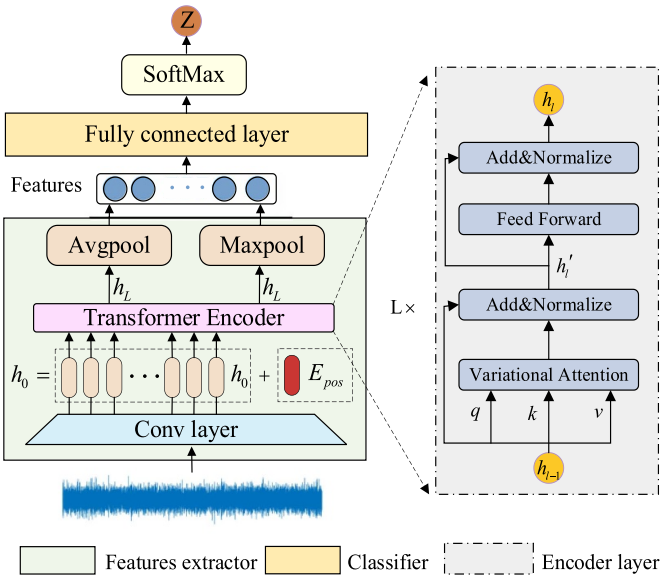

Figure 15. Structure of the VATN model. © [2022] IEEE. Reprinted, with permission, from [118].

the planetary gearbox fault recognition [121–123], and most papers are published in 2023.

To exploit the merits of both CNN and transformer, Han *et al* [124] proposed a Convformer-NSE to extract the local and global information from raw vibration signals, in which two convolutional layers were added to process input data and a novel Senet (NSE) [125]was integrated with transformer to make full use of learning of channel and spatial adaptivity. Sun *et al* [126] designed a multi-stage hierarchical structure via convolutional tokenization for transformer to learn both local and global information from raw signals, and meantime, spatial-reduction attention and linear dimension reduction projections were introduced to transformer to reduce the resource consumption. Li *et al* [118] proposed a variational attention-based transformer network (VATN) as shown in figure 15, in which the basic transformer model was improved by a sparse constrain to embed prior knowledge into the model. In addition, variational inference was added to the sparse constrain to force the attention to focus on the fault impulses, and heat map of attention weights was presented to associate the fault features with fault types.

Researchers also used signal processing methods to process the raw collected data and then adopted transformer to further recognize the planetary gearbox health conditions. Wu *et al* [127] firstly used Gramian angular field and Markov transition field to deal with the frequency-domain signals obtained by FFT, which was expected to remove the effect of time-shift property of time-domain vibration data. Then the features were fed into transformer to achieve further feature extraction and classification, which was validated to be more robust to noise. He *et al* [128] applied empirical mode decomposition (EMD) to process the acoustic emission data, and then the time-domain features of IMFs were calculated to feed into transformer for planetary gearbox fault diagnosis. These

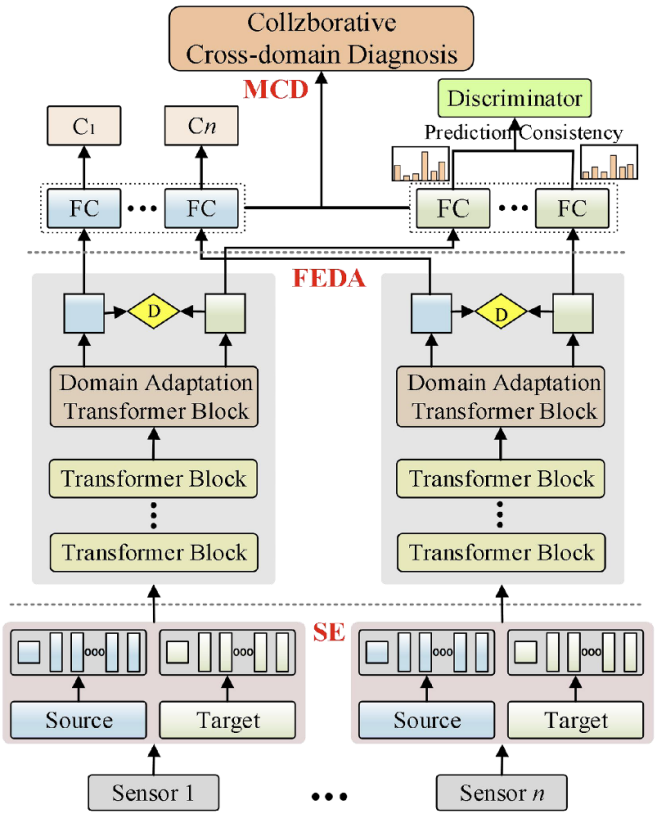

Figure 16. Structure of the MCDT model. © [2023] IEEE. Reprinted, with permission, from [129].

methods also achieve promising results in planetary gearbox fault recognition, but the end-to-end fault diagnosis that is expected in deep learning-enabled diagnosis is not implemented, i.e. prior knowledge is involved in these methods.

Transformer has also been investigated in cross-condition planetary gearbox fault recognition. For these relevant methods, compared to common transformer models, domain adaptation operation usually is used to remove the distribution discrepancy of the data under different operation conditions. Zhang *et al* [129] developed a multisensor cross-domain transformer (MCDT) framework included signal embedding module, a feature extraction and domain adaptation module and a multisensory collaborative diagnosis module (MCD) as shown in figure 16. In detail, an adversarial training via transformer was firstly designed for domain-invariant feature extraction, and a weighted voting method was developed to obtain a fusion result at the decision level. In [130], a multi-scale CNN and a multiscale feature extraction module via pyramid principle were firstly constructed to fuse and extract features from input data and then the features were fed into transformer for further feature learning. Finally, the domain adaptation was adopted to enable the transformer model to learn domain-invariant features, which was implemented by patch-level and class token domain adaptation loss terms.

Compared to CNN, the multihead self-attention enables transformer to extract global information well, which makes it rise abruptly in the last year and become popular. Much attention has been drawn to introduce and improve it in the fault diagnosis of planetary gearboxes. The works mainly focus on optimizing the input data by signal processing methods, improving transformer by combining convolution operations, and constructing cross-domain diagnosis framework by inserting adversarial training modules or domain-adaptation loss terms.

It is predicted that the number of deep learning methods based on transformer will increase rapidly. However, CNN can capture inductive biases, such as locality and translation equivariance, while transformer surpasses inductive biases by large-scale training. In addition, transformer models are usually large and computational expensive, so how to develop efficient transformer models remains an open problem. Maybe applying CNN pruning principle to prune transformer, i.e. only preserving important features, provides a solution to improve its efficiency, and makes it available for real-time monitoring.

4.5. GAN

A GAN provides an effective solution for machine fault diagnosis when the volume of collected data is small. From a literature review, the GAN models utilized for planetary gearbox fault diagnosis are usually constructed via AE [131] or CNN models [132–135].

Wang *et al* [136] designed a generative diagnosis method via a GAN and a stacked DAE, which was one of the earliest applications of a GAN in planetary gearbox fault diagnosis. In the model, the generator of the GAN was used to obtain the data following a similar distribution to that of the actual data to expand the volume of data, and the stacked DAE was adopted as the discriminator to extract features and discriminate their authenticity with fault classifications. This method was validated to be more effective than that of a stacked DAE and SAE. Zhao *et al* [137] proposed a parallel adversarial learning inference model in which 2D TFRs were used as input, a uniform distribution was applied to produce a multidimensional uniform distribution, and the encoder and decoder were simultaneously trained by a parallel adversarial game to subject the learned features to an explicit distribution. It was proven that this method performed well in fault diagnosis under non-stationary conditions, and the results based on TFRs obtained by STFR were better than those obtained by CWT and WVD. Ma *et al* [139] proposed a sparsity-constrained GAN to realize the generation of raw signals and thus improve the recognition accuracy. In this model, a sparse AE with an input layer consisting of constrained units and an output layer was used to reconstruct the input data. The decoder was separated to form the generator of the GAN, and the encoder was also separated to act as the discriminator by adding one output unit.

Considering their excellent feature extraction ability, CNNs have also been widely adapted to construct GAN models in the fault recognition of planetary gearboxes. Liang *et al* [138] developed a method with a WT, a CNN and a GAN as shown in figure 17, in which the WT was used to generate the TFRs, the GAN was applied as a generator to produce similar TFRs, and the CNN acted as a discriminator to extract features and

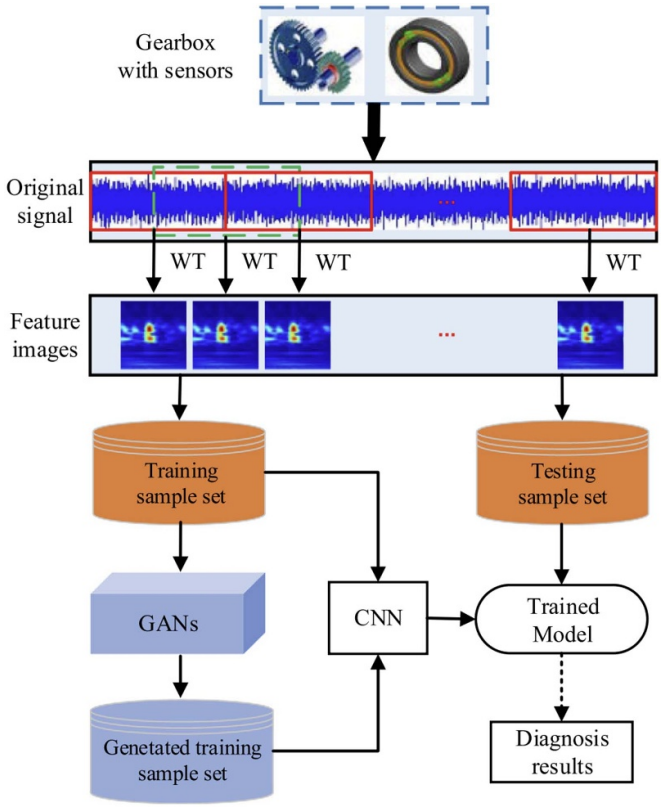

Figure 17. Flowchart of the method by WT, CNN and GAN. Reprinted from [138], Copyright (2020), with permission from Elsevier.

subject the real data and generated data to the same distribution. Luo *et al* [140] developed a conditional-deep convolutional GAN model in which the generator and discriminator were all implemented by CNN models and the conditions, i.e. label and fault type, were added to the generator. The input of the model was raw vibration signals, and the category error and the authenticity label error were utilized as the loss to train the generator and the discriminator. In addition to the label information, Wen *et al* [141] added the condition information to the CNN-based generator, and the generator was trained by minimizing the log-likelihoods of the correct source, the fault class and the working condition. The model was optimized using the planetary gearbox signals of different health states under several rotating speeds, and one dataset including all health states under another condition was adopted to test the model. Compared to AE-based GAN models, CNN-based methods inherit the merits of CNN in dealing with the time-shift property of time-domain, so they are more applicable to implement an end-to-end fault diagnosis and release the dependence on signal processing methods.

In addition to AE and CNN models, Liang *et al* [142] proposed a deep capsule neural network (CapNet) with a GAN model for wind turbine planetary gearbox fault diagnosis. This model was implemented by combining the Stockwell transform, which was adopted to transform data into 2D images, CapNet, which was developed from a CNN to better capture the deformation and viewing location information, and a GAN. It was validated that with the help of CapNet, a GAN can produce higher accuracy than that of a CNN.

In real applications, the data corresponding to a machine fault are usually insufficient. It is an effective solution to compensate for the limitation induced by the lack of data by applying GAN models to generate appropriate data through a generator and a discriminator. As in the literature review above, most GAN models in planetary gearbox fault diagnosis are constructed via AE and CNN models. Considering the data generation ability of GAN models, they are expected to achieve the fault diagnosis of planetary gearboxes under nonstationary conditions. However, the challenges of GAN in planetary gearbox fault diagnosis should also draw our attention. There is not a good solution to quantitatively evaluating the synthetic data, and the complex modulation characteristics of planetary gearbox vibration make it more difficult. In addition, the ideal result of GAN is that the synthetic data is almost the same as the input data, but the lack of data variety will lead to overfitting problem.

4.6. GNN

DNN performs well when handling the regular data, while GNN also considers the interdependence between the collected data, which makes it potentially improve the fault recognition accuracy. For example, the interaction among the data from multiple sensors can be well embedded by GNN. Therefore, GNN, as an emerging network, has been investigated in the planetary gearbox fault recognition recently.

Yu *et al* [144] used WPT to transform the vibration data into TFRs and constructed the graph data by these TFRs. Then, the graph data were fed into GCN for fault recognition, in which the fast graph convolutional kernel and particular pooling operation were applied to release the computational burden. Cao *et al* [145] developed a spiking graph attention network for planetary gearbox fault recognition, which included a graph attention convolutional layer formed by integrate and fire (IF) [146], and fully connected layer formed by leaky IF [147]. The input graph data was formed by applying a chaos theory to the frequency domain data of collected vibration data. Li *et al* [143] proposed a multireceptive field graph convolutional network (MRF-GCN) for fault recognition, in which the input data were converted into weighted graphs to enable the network to learn more stable features and a multireceptive field graph convolutional layer (MRF-GConv) was designed as shown in figure 18, to aggregate information from different receptive fields. Zhang *et al* [148] used a time-shift method to preprocess current data to remove power-line interference, and a hypergraph model as well as hyperedge convolution operation [149] was designed for high-order data learning.

In addition to the convolution optimization, other networks have also been investigated. Li *et al* [150] introduced a sparse constrain tern to random vector functional link network, and a discriminative adjacency graph was designed for feature learning. Sun *et al* [151] introduced neighborhood graph and sparse

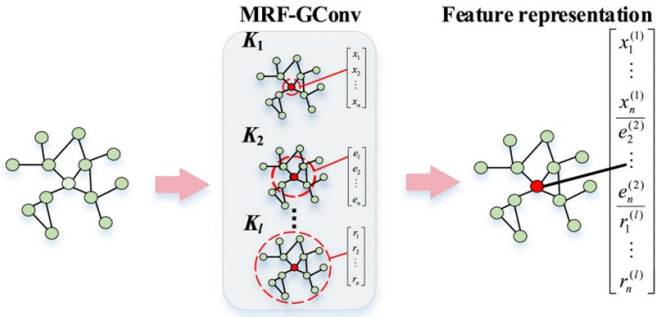

Figure 18. Structure of MRF-GConv. © [2021] IEEE. Reprinted, with permission, from [143].

theory to enhance AE for fault recognition. In [152], a local and non-local neighborhood information graph was embedded into objective function of AE to smooth the data structure. Shan *et al* [153] combined weighted graph-based label propagation and adversarial training to leverage the unlabeled data in the training process. Yu *et al* [154] developed a graph-weighted reinforcement network, in which the adjacency matrix was determined by measuring the Euclidean distance of time- and frequency-domain features and a graph-weighting enhanced mechanism was used to aggregate the node features in graph.

Although GNN has exhibited strong feature learning ability, especially for relationship mode characteristic mine, it also has some drawbacks to be solved. As [155] reported, when the number of layers of the mainstream GCN rise, the recognition rates for planetary gearbox health conditions increase firstly and then drop dramatically, and the best results correspond to the numbers of 2–5, which means that it does not work common deep learning methods. In addition, the performance of GNN will decrease a lot as the noise levels increase. Therefore, the anti-noise ability of GNN should be further improved.

5. Research prospect

Although deep learning has achieved great success in planetary gearbox fault diagnosis, there is still large potential for improvement, especially in actual applications. Accordingly, some research challenges and prospects are provided as follows.

(1) Model interpretation: Deep learning methods can learn hidden representations automatically from raw vibration signals of planetary gearboxes. However, the computation mechanisms of these models are unexplainable, so they are regarded as 'black boxes'. Some limited work has been performed to explain bearing fault diagnosis models. To the best of our knowledge, there are no reports explaining planetary gearbox fault diagnosis models. It seems that the models used for planetary gearbox fault recognition must be explained because decisions are made based on the fault information are more doubtful because the fault information is more difficult to learn due to the complexity of planetary gearboxes. It is suggested that the modules embedded amplitude- and frequency-modulation characteristics of planetary gearbox data can be added to deep learning models to enable the models to mainly focus on the fault-related information.

(2) Deep multisensory models: There are multiple and time-variant transmission paths between a fault component and a sensor, which may attenuate the vibration of the fault component due to the effects of dissipation and interference. Theoretically, vibration signals from different locations will be complementary to each other, so constructing deep models to fuse the data collected from multiple sensors with an appropriate layout will help to improve the accuracy. In addition, the encoder data has been proved to be simpler compared to vibration signals in frequency-demodulation methods, which may be helpful for multisensory intelligent fault diagnosis. In terms of the basic model, transformer and GNN models will be very suitable for multisensory data fusion considering the high feature extraction ability of transformer and the high interaction measurement ability of graph models for multiple sensors.

(3) Feature-based transferable deep learning (FTDL) models: FTDL models have been investigated and achieved success in bearing fault diagnosis through domain adaptation, which helps to achieve the fault recognition of target machines without labeled data. However, different from bearing vibration signals whose amplitude modulations caused by different fault parts are obviously different, the planetary gearbox signals caused by different fault parts have no obvious difference, and there are more components in planetary gearboxes. Therefore, it is more difficult to implement cross-planetary gearbox fault recognition with no labeled data from the target planetary gearbox. One promising way is to utilize multisensory fusion methods to implement domain adaptation, which makes it easier for models to learn domain-invariant features from various source data. In addition, compared to the popular CNN, transformer exhibits good global feature capture ability. In this view, theoretically, transformer will provide better results in domain adaptation, because the operation conditions mainly cause different magnitudes of fault impulses.

(4) Deep models for nonstationary conditions: Planetary gearboxes always operate under nonstationary conditions. For example, the rotation speeds of planetary gearboxes in wind turbines often fluctuate due to variations in wind power and direction. Nonstationary conditions will cause more complex modulation characteristics, such as variations in the amplitudes of vibration signals, the intervals of adjacent impulses, and time-varying characteristic frequencies in the TFR [16, 156], which adds to the difficulty of feature extraction by deep models. However, there are few relevant publications. Therefore, the investigation of deep models that are robust to speed fluctuations and load variations is practical and meaningful to push intelligent diagnosis methods to real industrial

applications. Although some techniques such as maximum mean discrepancy [157] and adversarial training [158], have been utilized to improve the robustness to operation conditions, they are only applicable for several different fixed speeds. The appropriate combination of advanced signal processing methods such as order tracking, time-frequency methods and sparse theory, and deep learning models will be beneficial to improving the accuracy. Perhaps, it is effective to incorporate the encoder data into models to improve the speed-adaptive ability.

(5) Open-source data and investigation on installation effect: In practical applications, compared to the machine providing the data used for training, the monitored machine must be reinstalled, so the installation effect is one practical and important factor in real applications. Because the structure of planetary gearboxes is complicated and the fault characteristics of some components, such as planet bearings, are very weak, if models are trained and tested using the same data collected under the same installation, whether the representations learned from the training data are the fault information or the installation information cannot be determined. Therefore, it is meaningful and valuable to share more data that are collected in accordance with real applications to attract more attention.

(6) Model generalization: Although the deep models that have been validated can produce high recognition results, they have only been tested with a lab experimental test rig. Most test rigs are simple and operate smoothly, but in real applications, the monitored machines are usually more complex, and the loads of planetary gearboxes change with the real-time variations of their loads. In addition, the fault types are more complex than those of the manufactured faults on the gears or the bearings. It is expected that more data collected in real industrial applications will be available to further test these models and to further push researchers to develop more useful deep models. To improve generalization, in addition to developing new model structures, it is suggested to add physical principles or domain knowledge to models to make them learn more fault information, which obviously could improve the generalization to operation conditions or new machines.

6. Dataset publication to facilitate future investigation

6.1. Existing public datasets

Deep leaning-enabled fault diagnosis models are usually optimized by a large number of data, so high-quality data is necessary for model construction and test. However, the collection of data in actual industrial applications is high-cost, and the experimental test rigs are usually expensive. Fortunately, there are several public datasets for our research. The common datasets include the bearing datasets of CWRU [159], IMS [160], MFPT [161] and XJTU-SY [162], PHM2009 fixed-shaft gearbox dataset [163], and SEU planetary gearbox dataset [164, 165]. It is seen that most public datasets are related to bearing faults, and, to the best knowledge, there is only one available public planetary gearbox dataset which is published in [164] and can be downloaded from the link [165].

The SEU planetary gearbox dataset is collected from a drive train dynamic simulator. There are four fault conditions, and two operation conditions are considered for each health condition. The vibration signals in the three directions, i.e. x, y and z, are collected. This dataset is commonly used in intelligent planetary gearbox fault diagnosis.

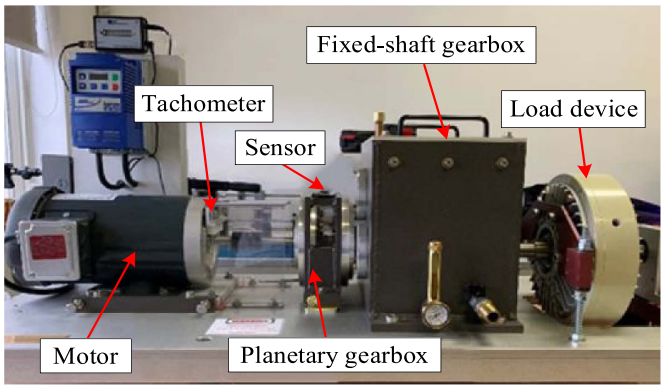

Figure 19. Wind turbine drive train test rig.

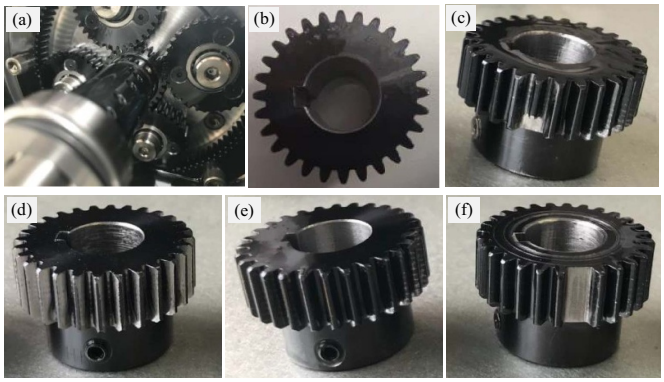

Figure 20. (a) Internal structure of planetary gearbox. (b) Healthy condition, (c) gear with a broken tooth, (d) wear gear, (e) crack occurs in the root, and (f) missing one tooth.

6.2. Dataset publication

In this section, a dataset (named 'WT-Planetary gearbox dataset') collected from the planetary gearbox in the wind turbine drive train test rig presented in figure 19 has been introduced. The test rig consists of a motor, a planetary gearbox, a fixed-shaft gearbox and a load device. There are four planet gears rotating around a sun gear. Five health conditions of the sun gear are considered as shown in figures 20(b)–(f). The vibration data are collected by Sinocera CA-YD-1181 accelerometer, and an encoder is used to capture the speed pulses. The sampling rate of all channels is set to be 48 kHz. Detailed parameters of the planetary gearbox and the relationships between

Table 3. Planetary gearbox parameters.

Tooth number	Sun gear	28
	Ring gear	100
	Planet gear (number)	36 (4)
Meshing frequency		(175/8) f_r
Fault frequency of sun gear		(25/8) f_r

Note: f_r represents sun gear rotating frequency.

the two crucial fault-related frequencies and the input shaft frequency are listed in table 3. The advantages of the dataset are listed as follows.

(1) Large volume. The data for each health condition are collected for a duration of more than 5 min. For the most popular CWRU dataset, the signal length is about 10 s. In previous papers, the length of samples is usually set to be 1000–2000 points. If we set the length of the samples as 2000, more than 7000 samples for each health condition can be obtained without overlapping operation, which is enough for deep network optimization.
(2) Multisensory data. As mentioned in section 4.3, encoder signals are simpler and contain abundant fault information, and various studies have validated they have advantages in signal demodulation-based planetary gearbox fault diagnosis [113–116]. For the dataset, two vibration signals in the x and y directions and the encoder data of the input shaft of the planetary gearbox are collected simultaneously (the first three channels in the 'Data' of each condition), which are available for multisensory deep learning model investigation.
(3) Operation conditions. Cross-domain fault diagnosis is one important subfield of deep learning-based fault diagnosis. In this dataset, eight speed conditions, i.e. 20 Hz, 25 Hz, 30 Hz, 35 Hz, 40 Hz, 45 Hz, 50 Hz, and 55 Hz, for each health condition are considered. This makes it very suitable for the research on model generalization and transfer learning tasks in intelligent planetary gearbox fault recognition.
(4) Installation effect. As shown in figure 21, in actual applications, the historical data used for training and the real-time data must be collected from different installations, so the installation factor should be considered in model test. In this dataset, all health conditions of all eight operation conditions are collected twice under disassembly and installation.

The equipment including the vibration sensors, encoder, and data acquisition card is validated to be effective in a simple bearing test rig [16]. We have analyzed some data collected from the test rig by the same data acquisition equipment via demodulation techniques [166], and the fault characteristic frequency of sun gear can be found. These data are high-quality, and are expected to facilitate relevant researchers' investigations on this topic.

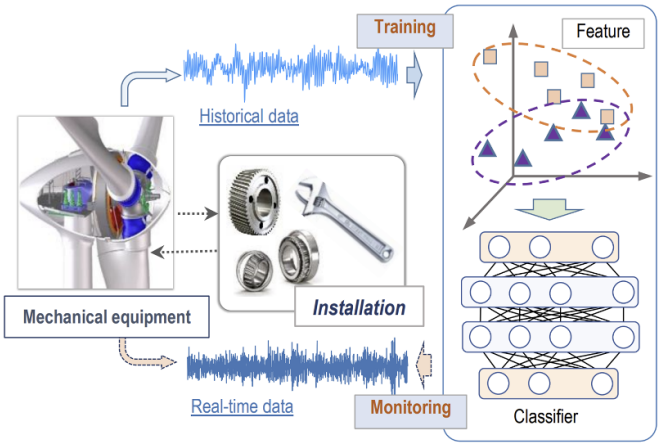

Figure 21. Illustration of installation factor in intelligent fault recognition.

The WT-Planetary gearbox dataset can be found from the link https://github.com/Liudd-BJUT/WT-planetary-gearbox-dataset.

7. Conclusion

In this paper, a comprehensive review of deep learning-based planetary gearbox fault diagnosis is conducted. Deep learning models can learn hidden features from collected data automatically and provide an end-to-end fault diagnosis, which avoids the challenges caused by the complex structure and intricate operation modes of planetary gearboxes. To make the applications of advanced deep learning methods more understandable, the principles of several deep learning methods that have been widely adopted in planetary gearbox fault diagnosis are introduced. Then, the applications of deep learning methods in the recent 5 year timeframe are reviewed. Finally, we provide several open topics for future research. Although promising results have been achieved, there are still some challenges to extending the above mentioned methods to actual applications. It is strongly expected that more actual conditions, such as the effect of installation, will be investigated to ensure that deep models make decisions based on fault information rather than interference factors. Accordingly, we will make our best attempt to solve these challenges.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

The work is supported by the National Natural Science Foundation of China (Grant No. 52305086 and 52075008). The authors thank the postgraduate students for helping them collect the dataset. The valuable comments and suggestions from the reviewers are very much appreciated.

ORCID iD

Dongdong Liu https://orcid.org/0000-0003-2638-3014

References

[1] Wang T, Han Q, Chu F and Feng Z 2019 Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: a review *Mech. Syst. Signal Process.* **126** 662–85

[2] Wang J, Liang Y, Zheng Y, Gao R X and Zhang F 2020 An integrated fault diagnosis and prognosis approach for predictive maintenance of wind turbine bearing with limited samples *Renew. Energy* **145** 642–50

[3] Chen X, Yan R and Liu Y 2016 Wind turbine condition monitoring and fault diagnosis in China *IEEE Instrum. Meas. Mag.* **19** 22–28

[4] Zhang W, Yang D and Wang H 2019 Data-driven methods for predictive maintenance of industrial equipment: a survey *IEEE Syst. J.* **13** 2213–27

[5] Liu D, Cui L and Cheng W 2023 Flexible iterative generalized demodulation filtering for the fault diagnosis of rotating machinery under nonstationary conditions *Struct. Health Monit.* **22** 1421–36

[6] Liu D, Cui L and Cheng W 2023 Flexible generalized demodulation for intelligent bearing fault diagnosis under nonstationary conditions *IEEE Trans. Ind. Inform.* **19** 2717–28

[7] Lei Y, Lin J, Zuo M J and He Z 2014 Condition monitoring and fault diagnosis of planetary gearboxes: a review *Measurement* **48** 292–305

[8] Li M, Wang T, Kong Y and Chu F 2022 Synchro-reassigning transform for instantaneous frequency estimation and signal reconstruction *IEEE Trans. Ind. Electron.* **69** 7263–74

[9] Kong Y, Qin Z, Wang T, Rao M, Feng Z and Chu F 2022 Data-driven dictionary design–based sparse classification method for intelligent fault diagnosis of planet bearings *Struct. Health Monit.* **21** 1313–28

[10] Kong Y, Qin Z, Wang T, Han Q and Chu F 2021 An enhanced sparse representation-based intelligent recognition method for planet bearing fault diagnosis in wind turbines *Renew. Energy* **173** 987–1004

[11] Wang T, Han Q, Chu F and Feng Z 2016 A new SKRgram based demodulation technique for planet bearing fault detection *J. Sound Vib.* **385** 330–49

[12] Zhang D and Feng Z 2023 Proportion-extracting chirplet transform for nonstationary signal analysis of rotating machinery *IEEE Trans. Ind. Inform.* **19** 2674–83

[13] Li M, Wang T, Chu F and Feng Z 2021 Component matching chirplet transform via frequency-dependent chirp rate for wind turbine planetary gearbox fault diagnostics under variable speed condition *Mech. Syst. Signal Process.* **161** 107997

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

[15] Zhu Z, Lei Y, Qi G, Chai Y, Mazur N, An Y and Huang X 2023 A review of the application of deep learning in intelligent fault diagnosis of rotating machinery *Measurement* **206** 112346

[16] Liu D, Cheng W and Wen W 2021 Intelligent cross-condition fault recognition of rolling bearings based on normalized resampled characteristic power and self-organizing, map *Mech. Syst. Signal Process.* **153** 107462

[17] Zhao R, Yan R, Chen Z, Mao K, Wang P and Gao R X 2019 Deep learning and its applications to machine health monitoring *Mech. Syst. Signal Process.* **115** 213–37

[18] Qiu H, Lee J, Lin J and Yu G 2003 Robust performance degradation assessment methods for enhanced rolling element bearing prognostics *Adv. Eng. Inform.* **17** 127–40

[19] Helmi H and Forouzantabar A 2019 Rolling bearing fault detection of electric motor using time domain and frequency domain features extraction and ANFIS *IET Electr. Power Appl.* **13** 662–9

[20] Li Z, Fang H, Huang M, Wei Y and Zhang L 2018 Data-driven bearing fault identification using improved hidden Markov model and self-organizing map *Comput. Ind. Eng.* **116** 37–46

[21] Li Y, Wang S and Deng Z 2021 Intelligent fault identification of rotary machinery using refined composite multi-scale Lempel–Ziv complexity *J. Manuf. Syst.* **61** 725–35

[22] Lei Y, He Z and Zi Y 2011 EEMD method and WNN for fault diagnosis of locomotive roller bearings *Expert Syst. Appl.* **38** 7334–41

[23] Liu Y, Cheng Y, Zhang Z and Wu J 2022 Multi-information fusion fault diagnosis based on KNN and improved evidence theory *J. Vib. Eng. Technol.* **10** 841–52

[24] Wang X, Si S and Li Y 2022 Variational embedding multiscale diversity entropy for fault diagnosis of large-scale machinery *IEEE Trans. Ind. Electron.* **69** 3109–19

[25] Vincent P, Larochelle H, Bengio Y and Manzagol P-A 2008 Extracting and composing robust features with denoising autoencoders *Proc. 25th Int. Conf. on Machine learning—ICML'08* (ACM Press) pp 1096–103

[26] Salakhutdinov R and G H 2009 Deep Boltzmann machines *J. Mach. Learn. Res.* **5** 448–55

[27] LeCun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W and Jackel L 1989 Handwritten digit recognition with a back-propagation network *Advances in Neural Information Processing Systems*

[28] Liu R, Yang B, Zio E and Chen X 2018 Artificial intelligence for fault diagnosis of rotating machinery: a review *Mech. Syst. Signal Process.* **108** 33–47

[29] Hoang D-T and Kang H-J 2019 A survey on deep learning based bearing fault diagnosis *Neurocomputing* **335** 327–35

[30] Luo Y, Cui L, Zhang J and Ma J 2021 Vibration mechanism and improved phenomenological model of planetary gearbox with broken sun gear fault *Measurement* **178** 109356

[31] Luo Y, Cui L and Ma J 2021 Effect of bolt constraint of ring gear on the vibration response of the planetary gearbox *Mech. Mach. Theory* **159** 104260

[32] Elasha F, Greaves M, Mba D and Fang D 2017 A comparative study of the effectiveness of vibration and acoustic emission in diagnosing a defective bearing in a planetry gearbox *Appl. Acoust.* **115** 181–95

[33] Elasha F, Greaves M and Mba D 2018 Planetary bearing defect detection in a commercial helicopter main gearbox with vibration and acoustic emission *Struct. Health Monit.* **17** 1192–212

[34] Kong Y, Wang T and Chu F 2019 Meshing frequency modulation assisted empirical wavelet transform for fault diagnosis of wind turbine planetary ring gear *Renew. Energy* **132** 1373–88

[35] Feng Z, Zhu W and Zhang D 2019 Time-frequency demodulation analysis via Vold-Kalman filter for wind turbine planetary gearbox fault diagnosis under nonstationary speeds *Mech. Syst. Signal Process.* **128** 93–109

[36] Feng Z, Yu X, Zhang D and Liang M 2020 Generalized adaptive mode decomposition for nonstationary signal

analysis of rotating machinery: principle and applications *Mech. Syst. Signal Process.* **136** 106530
[37] Lei Y, Han D, Lin J and He Z 2013 Planetary gearbox fault diagnosis using an adaptive stochastic resonance method *Mech. Syst. Signal Process.* **38** 113–24
[38] Wang T, Liang M, Li J, Cheng W and Li C 2015 Bearing fault diagnosis under unknown variable speed via gear noise cancellation and rotational order sideband identification *Mech. Syst. Signal Process.* **62–63** 30–53
[39] Kong Y, Wang T, Chu F, Feng Z and Selesnick I 2021 Discriminative dictionary learning-based sparse classification framework for data- driven machinery fault diagnosis *IEEE Sens. J.* **21** 8117–29
[40] Feng Z, Ma H and Zuo M J 2016 Amplitude and frequency demodulation analysis for fault diagnosis of planet bearings *J. Sound Vib.* **382** 395–412
[41] Rumelhart D E, Hinton G E and Williams R J 1986 Learning representations by back-propagating errors *Nature* **323** 533–6
[42] Hinton G E, Osindero S and Teh Y-W 2006 A fast learning algorithm for deep belief nets *Neural Comput.* **18** 1527–54
[43] Hinton G E 2002 Training products of experts by minimizing contrastive divergence *Neural Comput.* **14** 1771–800
[44] Hinton G E and Salakhutdinov R R 2006 Reducing the dimensionality of data with neural networks *Science* **313** 504–7
[45] Krizhevsky A, Sutskever I and Hinton G E 2017 ImageNet classification with deep convolutional neural networks *Commun. ACM* **60** 84–90
[46] LeCun Y, Bengio Y and Hinton G 2015 Deep learning *Nature* **521** 436–44
[47] Zhang J, Y S, Guo L, H G, Hong X and Song H 2020 A new bearing fault diagnosis method based on modified convolutional neural networks *Chin. J. Aeronaut.* **33** 439–47
[48] Jiao J, Zhao M, Lin J and Liang K 2020 A comprehensive review on convolutional neural network in machine fault diagnosis *Neurocomputing* **417** 36–63
[49] Liu D, Cui L, Cheng W, Zhao D and Wen W 2022 Rolling bearing fault severity recognition via data mining integrated with convolutional neural network *IEEE Sens. J.* **22** 5768–77
[50] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* pp 770–8
[51] He K, Zhang X, Ren S and Sun J 2016 Identity mappings in deep residual networks *Computer Vision—ECCV 2016* pp 630–45
[52] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser Ł and Polosukhin I 2017 Attention is all you need *Advances in Neural Information Processing Systems* pp 5998–6008
[53] Tang J, Zheng G, Wei C, Huang W and Ding X 2022 Signal-transformer: a robust and interpretable method for rotating machinery intelligent fault diagnosis under variable operating conditions *IEEE Trans. Instrum. Meas.* **71** 1–11
[54] Han K *et al* 2023 A survey on vision transformer *IEEE Trans. Pattern Anal. Mach. Intell.* **45** 87–110
[55] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y 2020 Generative adversarial networks *Commun. ACM* **63** 139–44
[56] Goodfellow I 2016 NIPS 2016 tutorial: generative adversarial networks (arXiv:1701.00160v4)
[57] Scarselli F, Gori M, Chung Tsoi A, Hagenbuchner M and Monfardini G 2009 The graph neural network model *IEEE Trans. Neural Netw.* **20** 61–80
[58] Kipf T N and Welling M 2017 Semi-supervised classification with graph convolutional networks *ICLR 2017* (arXiv:1609.02907)
[59] Zhang S, Tong H, Xu J and Maciejewski R 2019 Graph convolutional networks: a comprehensive review *Comput. Soc. Netw.* **6** 11
[60] Ruiz L, Gama F and Ribeiro A 2020 Gated graph recurrent neural networks *IEEE Trans. Signal Process.* **68** 6303–18
[61] Kollias G, Kalantzis V, Ide T, Lozano A and Abe N 2022 Directed graph auto-encoders *Proc. AAAI Conf. on Artificial Intelligence* vol 36 pp 7211–9
[62] Jia F, Lei Y, Lin J, Zhou X and Lu N 2016 Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data *Mech. Syst. Signal Process.* **72–73** 303–15
[63] Jia F, Lei Y, Guo L, Lin J and Xing S 2018 A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines *Neurocomputing* **272** 619–28
[64] Zhang Z, Yang Q, Zi Y and Wu Z 2022 Discriminative sparse autoencoder for gearbox fault diagnosis toward complex vibration signals *IEEE Trans. Instrum. Meas.* **71** 1–11
[65] Yang S, Wang Y and Li C 2021 Wind turbine gearbox fault diagnosis based on an improved supervised autoencoder using vibration and motor current signals *Meas. Sci. Technol.* **32** 114003
[66] Wang J, Li S, Han B, An Z, Xin Y, Qian W and Wu Q 2019 Construction of a batch-normalized autoencoder network and its application in mechanical intelligent fault diagnosis *Meas. Sci. Technol.* **30** 015106
[67] Yu J, Xu Y and Liu K 2019 Planetary gear fault diagnosis using stacked denoising autoencoder and gated recurrent unit neural network under noisy environment and time-varying rotational speed conditions *Meas. Sci. Technol.* **30** 095003
[68] Saufi S R, Ahmad Z A B, Leong M S and Lim M H 2020 Gearbox fault diagnosis using a deep learning model with limited data sample *IEEE Trans. Ind. Inform.* **16** 6263–71
[69] Ab Wahab M N, Nefti-Meziani S and Atyabi A 2015 A comprehensive review of swarm optimization algorithms *PLoS One* **10** e0122827
[70] Shao H, Xia M, Wan J and de Silva C W 2022 Modified stacked autoencoder using adaptive Morlet wavelet for intelligent fault diagnosis of rotating machinery *IEEE/ASME Trans. Mechatronics* **27** 24–33
[71] Chen X, Ji A and Cheng G 2019 A novel deep feature learning method based on the fused-stacked AEs for planetary gear fault diagnosis *Energies* **12** 4522
[72] Shao H, Lin J, Zhang L, Galar D and Kumar U 2021 A novel approach of multisensory fusion to collaborative fault diagnosis in maintenance *Inform. Fusion* **74** 65–76
[73] Hassairi S, Ejbali R and Zaied M 2018 A deep stacked wavelet auto-encoders to supervised feature extraction to pattern classification *Multimedia Tools Appl.* **77** 5443–59
[74] Li X, Jiang H, Niu M and Wang R 2020 An enhanced selective ensemble deep learning method for rolling bearing fault diagnosis with beetle antennae search algorithm *Mech. Syst. Signal Process.* **142** 106752
[75] Lu L, He Y, Ruan Y and Yuan W 2021 An optimized stacked diagnosis structure for fault diagnosis of wind turbine planetary gearbox *Meas. Sci. Technol.* **32** 075102
[76] Chen H, Wang J, Tang B, Xiao K and Li J 2017 An integrated approach to planetary gearbox fault diagnosis using deep belief networks *Meas. Sci. Technol.* **28** 025010
[77] Han D, Guo X and Shi P 2020 An intelligent fault diagnosis method of variable condition gearbox based on improved DBN combined with WPEE and MPE *IEEE Access* **8** 131299–309

[78] Zhang T, Li Z, Deng Z and Hu B 2019 Hybrid data fusion DBN for intelligent fault diagnosis of vehicle reducers *Sensors* **19** 2504
[79] Guo Y, Wu X, Na J and Fung R-F 2016 Envelope synchronous average scheme for multi-axis gear faults detection *J. Sound Vib.* **365** 276–86
[80] Qin Y, Wang X and Zou J 2019 The optimized deep belief networks with improved logistic sigmoid units and their application in fault diagnosis for planetary gearboxes of wind turbines *IEEE Trans. Ind. Electron.* **66** 3814–24
[81] Yang J, Bao W, Liu Y, Li X, Wang J, Niu Y and Li J 2021 Joint pairwise graph embedded sparse deep belief network for fault diagnosis *Eng. Appl. Artif. Intell.* **99** 104149
[82] Xing S, Lei Y, Wang S and Jia F 2021 Distribution-invariant deep belief network for intelligent fault diagnosis of machines under new working conditions *IEEE Trans. Ind. Electron.* **68** 2617–25
[83] Jing L, Zhao M, Li P and Xu X 2017 A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox *Measurement* **111** 1–10
[84] Huang D, Zhang W-A, Guo F, Liu W and Shi X 2023 Wavelet packet decomposition-based multiscale CNN for fault diagnosis of wind turbine gearbox *IEEE Trans. Cybern.* **53** 443–53
[85] Han Y, Tang B and Deng L 2019 An enhanced convolutional neural network with enlarged receptive fields for fault diagnosis of planetary gearboxes *Comput. Ind.* **107** 50–58
[86] Wang C, Li H, Zhang K, Hu S and Sun B 2021 Intelligent fault diagnosis of planetary gearbox based on adaptive normalized CNN under complex variable working conditions and data imbalance *Measurement* **180** 109565
[87] Chang X, Tang B, Tan Q, Deng L and Zhang F 2020 One-dimensional fully decoupled networks for fault diagnosis of planetary gearboxes *Mech. Syst. Signal Process.* **141** 106482
[88] Feng Z, Liang M and Chu F 2013 Recent advances in time–frequency analysis methods for machinery fault diagnosis: a review with application examples *Mech. Syst. Signal Process.* **38** 165–205
[89] Liu D, Cheng W and Wen W 2020 Rolling bearing fault diagnosis via STFT and improved instantaneous frequency estimation method *Proc. Manuf.* **49** 166–72
[90] Yan R, Gao R X and Chen X 2014 Wavelets for fault diagnosis of rotary machines: a review with applications *Signal Process.* **96** 1–15
[91] Cohen L 1989 Time-frequency distributions-a review *Proc. IEEE* **77** 941–81
[92] Chen R, Huang X, Yang L, Xu X, Zhang X and Zhang Y 2019 Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform *Comput. Ind.* **106** 48–59
[93] Wang D-F, Guo Y, Wu X, Na J and Litak G 2020 Planetary-gearbox fault classification by convolutional neural network and recurrence plot *Appl. Sci.* **10** 932
[94] Emmanuel S, Yihun Y, Nili Ahmedabadi Z and Boldsaikhan E 2021 Planetary gear train microcrack detection using vibration data and convolutional neural networks *Neural Comput. Appl.* **33** 17223–43
[95] Chen Z, Mauricio A, Li W and Gryllias K 2020 A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks *Mech. Syst. Signal Process.* **140** 106683
[96] Abboud D, Elbadaoui M, Smith W A and Randall R B 2019 Advanced bearing diagnostics: a comparative study of two powerful approaches *Mech. Syst. Signal Process.* **114** 604–27
[97] Perez-Sanjines F, Peeters C, Verstraeten T, Antoni J, Nowé A and Helsen J 2023 Fleet-based early fault detection of wind turbine gearboxes using physics-informed deep learning based on cyclic spectral coherence *Mech. Syst. Signal Process.* **185** 109760
[98] Yang R, An Z and Song S 2022 Multilayer extreme learning convolutional feature neural network model for the weak feature classification and status identification of planetary bearing *J. Sens.* **2022** 1–11
[99] Lingli J, Shuhui L, Xuejun L, Jiale L and Dalian Y 2022 Fault diagnosis of a planetary gearbox based on a local bi-spectrum and a convolutional neural network *Meas. Sci. Technol.* **33** 045008
[100] Kim Y, Na K and Youn B D 2022 A health-adaptive time-scale representation (HTSR) embedded convolutional neural network for gearbox fault diagnostics *Mech. Syst. Signal Process.* **167** 108575
[101] Zhu C, Chen Z, Zhao R, Wang J and Yan R 2021 Decoupled feature-temporal CNN: explaining deep learning-based machine health monitoring *IEEE Trans. Instrum. Meas.* **70** 1–13
[102] Chen H, Hu N, Cheng Z, Zhang L and Zhang Y 2019 A deep convolutional neural network based fusion method of two-direction vibration signal data for health state identification of planetary gearboxes *Measurement* **146** 268–78
[103] Guo S, Yang T, Hua H and Cao J 2021 Coupling fault diagnosis of wind turbine gearbox based on multitask parallel convolutional neural networks with overall information *Renew. Energy* **178** 639–50
[104] Jiao J, Zhao M, Lin J and Ding C 2019 Deep coupled dense convolutional network with complementary data for intelligent fault diagnosis *IEEE Trans. Ind. Electron.* **66** 9858–67
[105] Li T, Zhao Z, Sun C, Yan R and Chen X 2020 Adaptive channel weighted CNN with multisensor fusion for condition monitoring of helicopter transmission system *IEEE Sens. J.* **20** 8364–73
[106] Zhang K, Tang B, Deng L, Tan Q and Yu H 2021 A fault diagnosis method for wind turbines gearbox based on adaptive loss weighted meta-ResNet under noisy labels *Mech. Syst. Signal Process.* **161** 107963
[107] Zhao M, Zhong S, Fu X, Tang B, Dong S and Pecht M 2021 Deep residual networks with adaptively parametric rectifier linear units for fault diagnosis *IEEE Trans. Ind. Electron.* **68** 2587–97
[108] Zhang S, Liu Z, Chen Y, Jin Y and Bai G 2023 Selective kernel convolution deep residual network based on channel-spatial attention mechanism and feature fusion for mechanical fault diagnosis *ISA Trans.* **133** 369–83
[109] Zhao M, Kang M, Tang B and Pecht M 2019 Multiple wavelet coefficients fusion in deep residual networks for fault diagnosis *IEEE Trans. Ind. Electron.* **66** 4696–706
[110] Huang X, Qi G, Mazur N and Chai Y 2022 Deep residual networks-based intelligent fault diagnosis method of planetary gearboxes in cloud environments *Simul. Model. Pract. Theory* **116** 102469
[111] Zhang K, Tang B, Deng L and Liu X 2021 A hybrid attention improved ResNet based fault diagnosis method of wind turbines gearbox *Measurement* **179** 109491
[112] Xie T, Huang X and Choi S-K 2022 Intelligent mechanical fault diagnosis using multisensor fusion and convolution neural network *IEEE Trans. Ind. Inform.* **18** 3213–23
[113] Feng Z, Gao A, Li K and Ma H 2021 Planetary gearbox fault diagnosis via rotary encoder signal analysis *Mech. Syst. Signal Process.* **149** 107325
[114] Jiao J, Zhao M, Lin J and Zhao J 2018 A multivariate encoder information based convolutional neural network for intelligent fault diagnosis of planetary gearboxes *Knowl.-Based Syst.* **160** 237–50

[115] Miao Y, Zhao M, Yi Y and Lin J 2020 Application of sparsity-oriented VMD for gearbox fault diagnosis based on built-in encoder information *ISA Trans.* **99** 496–504

[116] Liang K, Zhao M, Lin J, Jiao J and Ding C 2022 Toothwise health monitoring of planetary gearbox under time-varying speed condition based on rotating encoder signal *IEEE Trans. Ind. Electron.* **69** 6267–77

[117] Xu Z, Bashir M, Zhang W, Yang Y, Wang X and Li C 2022 An intelligent fault diagnosis for machine maintenance using weighted soft-voting rule based multi-attention module with multi-scale information fusion *Inform. Fusion* **86–87** 17–29

[118] Li Y, Zhou Z, Sun C, Chen X and Yan R 2022 Variational attention-based interpretable transformer network for rotary machine fault diagnosis *IEEE Trans. Neural Netw. Learn. Syst.* pp 1–14

[119] Wang H, Liu Z, Peng D and Qin Y 2020 Understanding and learning discriminant features based on multiattention 1DCNN for wheelset bearing fault diagnosis *IEEE Trans. Ind. Inform.* **16** 5735–45

[120] Wang H, Liu Z, Peng D and Cheng Z 2022 Attention-guided joint learning CNN with noise robustness for bearing fault diagnosis and vibration signal denoising *ISA Trans.* **128** 470–84

[121] Zhang Y, Ji J C, Ren Z, Ni Q and Wen B 2023 Multi-sensor open-set cross-domain intelligent diagnostics for rotating machinery under variable operating conditions *Mech. Syst. Signal Process.* **191** 110172

[122] Zhang Y, Ren Z, Feng K, Yu K, Beer M and Liu Z 2023 Universal source-free domain adaptation method for cross-domain fault diagnosis of machines *Mech. Syst. Signal Process.* **191** 110159

[123] Weng C, Lu B, Gu Q and Zhao X 2023 A novel multisensor fusion transformer and its application into rotating machinery fault diagnosis *IEEE Trans. Instrum. Meas.* **72** 1–12

[124] Han S, Shao H, Cheng J, Yang X and Cai B 2023 Convformer-NSE: a novel end-to-end gearbox fault diagnosis framework under heavy noise using joint global and local information *IEEE/ASME Trans. Mechatronics* **28** 340–9

[125] Hu J, Shen L and Sun G 2018 Squeeze-and-excitation networks *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* pp 7132–41

[126] Sun W, Wang H, Xu J, Yang Y and Yan R 2022 Effective convolutional transformer for highly accurate planetary gearbox fault diagnosis *IEEE Open J. Instrum. Meas.* **1** 1–9

[127] Wu R, Liu C, Han T, Yao J and Jiang D 2023 A planetary gearbox fault diagnosis method based on time-series imaging feature fusion and a transformer model *Meas. Sci. Technol.* **34** 024006

[128] He D, He M and Yoon J 2023 A natural language processing based planetary gearbox fault diagnosis with acoustic emission signals *2023 IEEE Aerospace Conf.* (IEEE) pp 01–6

[129] Zhang Y, Ren Z, Feng K, Yu K, Ma H and Liu Z 2023 Transformer-enabled cross-domain diagnostics for complex rotating machinery with multiple sensors *IEEE/ASME Trans. Mechatronics* **28** 2293–304

[130] Zhang Y, Feng K, Ma H, Yu K, Ren Z and Liu Z 2022 MMFNet: multisensor data and multiscale feature fusion model for intelligent cross-domain machinery fault diagnosis *IEEE Trans. Instrum. Meas.* **71** 1–11

[131] Cui L, Tian X, Shi X, Wang X and Cui Y 2021 A semi-supervised fault diagnosis method based on improved bidirectional generative adversarial network *Appl. Sci.* **11** 9401

[132] Pu Z, Cabrera D, Bai Y and Li C 2022 A one-class generative adversarial detection framework for multifunctional fault diagnoses *IEEE Trans. Ind. Electron.* **69** 8411–9

[133] Su Y, Meng L, Kong X, Xu T, Lan X and Li Y 2022 Small sample fault diagnosis method for wind turbine gearbox based on optimized generative adversarial networks *Eng. Fail. Anal.* **140** 106573

[134] fan Y, Huang D, Li D, Zhao Y, Lin S and M S M 2022 A novel convolution network with self-adaptation high-pass filter for fault diagnosis of wind turbine gearbox *Meas. Sci. Technol.* **34** 025024

[135] Zhou K, Diehl E and Tang J 2023 Deep convolutional generative adversarial network with semi-supervised learning enabled physics elucidation for extended gear fault diagnosis under data limitations *Mech. Syst. Signal Process.* **185** 109772

[136] Wang Z, Wang J and Wang Y 2018 An intelligent diagnosis scheme based on generative adversarial learning deep neural networks and its application to planetary gearbox fault pattern recognition *Neurocomputing* **310** 213–22

[137] Zhao C and Zhang Y 2022 Parallel adversarial feature learning and enhancement of feature discriminability for fault diagnosis of a planetary gearbox under time-varying speed conditions *Meas. Sci. Technol.* **33** 125019

[138] Liang P, Deng C, Wu J and Yang Z 2020 Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network *Measurement* **159** 107768

[139] Ma L, Ding Y, Wang Z, Wang C, Ma J and Lu C 2021 An interpretable data augmentation scheme for machine fault diagnosis based on a sparsity-constrained generative adversarial network *Expert Syst. Appl.* **182** 115234

[140] Luo J, Huang J and Li H 2021 A case study of conditional deep convolutional generative adversarial networks in machine fault diagnosis *J. Intell. Manuf.* **32** 407–25

[141] Wen W, Bai Y and Cheng W 2020 Generative adversarial learning enhanced fault diagnosis for planetary gearbox under varying working conditions *Sensors* **20** 1685

[142] Liang P, Deng C, Yuan X and Zhang L 2023 A deep capsule neural network with data augmentation generative adversarial networks for single and simultaneous fault diagnosis of wind turbine gearbox *ISA Trans.* **135** 462–75

[143] Li T, Zhao Z, Sun C, Yan R and Chen X 2021 Multireceptive field graph convolutional networks for machine fault diagnosis *IEEE Trans. Ind. Electron.* **68** 12739–49

[144] Yu X, Tang B and Zhang K 2021 Fault diagnosis of wind turbine gearbox using a novel method of fast deep graph convolutional networks *IEEE Trans. Instrum. Meas.* **70** 1–14

[145] Cao S, Li H, Zhang K, Yang C, Xiang W and Sun F 2023 A novel spiking graph attention network for intelligent fault diagnosis of planetary gearboxes *IEEE Sens. J.* **23** 13140–54

[146] Cessac B, Paugam-Moisy H and Viéville T 2010 Overview of facts and issues about neural coding by spikes *J. Physiol. Paris* **104** 5–18

[147] Jin D Z 2002 Fast convergence of spike sequences to periodic patterns in recurrent networks *Phys. Rev. Lett.* **89** 208102

[148] Zhang K, Li H, Cao S, Yang C, Sun F and Wang Z 2022 Motor current signal analysis using hypergraph neural networks for fault diagnosis of electromechanical system *Measurement* **201** 111697

[149] Feng Y, You H, Zhang Z, Ji R and Gao Y 2019 Hypergraph neural networks *Proc. AAAI Conf. on Artificial Intelligence* vol 33 pp 3558–65

[150] Li X, Yang Y, Wu Z, Yan K, Shao H and Cheng J 2022 High-accuracy gearbox health state recognition based on graph sparse random vector functional link network *Reliab. Eng. Syst. Saf.* **218** 108187

[151] Sun Z, Wang Y and Gao J 2023 Intelligent fault diagnosis of rotating machinery under varying working conditions with global–local neighborhood and sparse graphs embedding deep regularized autoencoder *Eng. Appl. Artif. Intell.* **124** 106590
[152] Sun Z, Wang Y, Sun G, Gao J and PAN Z 2023 Neighborhood graph embedding interpretable fault diagnosis network based on local and non-local information balanced under imbalanced samples *Struct. Health Monit.* **22** 1721–44
[153] Shan D, Cheng C, Li L, Peng Z and He Q 2023 Semisupervised fault diagnosis of gearbox using weighted graph-based label propagation and virtual adversarial training *IEEE Trans. Instrum. Meas.* **72** 1–11
[154] Yu X, Tang B and Deng L 2023 Fault diagnosis of rotating machinery based on graph weighted reinforcement networks under small samples and strong noise *Mech. Syst. Signal Process.* **186** 109848
[155] Li T, Zhou Z, Li S, Sun C, Yan R and Chen X 2022 The emerging graph neural networks for intelligent fault diagnostics and prognostics: a guideline and a benchmark study *Mech. Syst. Signal Process.* **168** 108653
[156] Liu D, Cheng W and Wen W 2020 An online bearing fault diagnosis technique via improved demodulation spectrum analysis under variable speed conditions *IEEE Syst. J.* **14** 2323–34
[157] Chen P, Zhao R, He T, Wei K and Yuan J 2023 A novel bearing fault diagnosis method based joint attention adversarial domain adaptation *Reliab. Eng. Syst. Saf.* **237** 109345
[158] Wang X, Jiang H, Liu Y and Yang Q 2023 Data-augmented patch variational autoencoding generative adversarial networks for rolling bearing fault diagnosis *Meas. Sci. Technol.* **34** 055102
[159] (Available at: https://engineering.case.edu/bearingdatacenter)
[160] Qiu H, Lee J, Lin J and Yu G 2006 Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics *J. Sound Vib.* **289** 1066–90
[161] Bechhoefer E 2012 (available at: www.mfpt.org/fault-data-sets/)
[162] Wang B (available at: https://biaowang.tech/xjtu-sy-bearing-datasets/)
[163] (Available at: https://phmsociety.org/public-data-sets/)
[164] Shao S, McAleer S, Yan R and Baldi P 2019 Highly accurate machine fault diagnosis using deep transfer learning *IEEE Trans. Ind. Inform.* **15** 2446–55
[165] (Available at: https://github.com/cathysiyu/Mechanical-datasets)
[166] Liu D, Cui L and Cheng W 2023 Fault diagnosis of wind turbines under nonstationary conditions based on a novel tacho-less generalized demodulation *Renew. Energy* **206** 645–57