Simulation Modelling Practice and Theory 102 (2020) 101981

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

Mechanical fault diagnosis and prediction in IoT based on multi-source sensing data fusion

Min Huang[a], Zhen Liu*,[a], Yang Tao[b]

[a] Department of Software Engineering, South China University of Technology (SCUT), Guangzhou 510006, China
[b] Guangzhou Robustel LTD, Guangzhou 510663, China

ARTICLE INFO

Keywords:
Internet of things
Fault diagnosis and prediction
Big data analysis
Deep learning
Data fusion

ABSTRACT

Using multi-source sensing data based on the Internet of Things (IoT) with artificial intelligence and big data processing technology to achieve predictive maintenance of mechanical equipment can remarkably improve the service life of the machine and reduce labor costs when diagnosing mechanical faults, and it has become a highly relevant research topic. In this paper, the multi-source sensing data fusion models and fusion algorithms are studied and discussed. First, the Joint Directors of Laboratories (JDL) fusion model and the Hierarchical fusion model are compared and analyzed. Then, various types of fusion algorithms based on Neural Networks and Deep Learning, including Dempster-Shafer (D-S) evidence theory and their applications in mechanical fault diagnosis and fault prediction, are studied and compared. The findings reveal that exploring and designing a more intelligent fusion model incorporating the beneficial characteristics of different fusion algorithms are challenging and have a certain value for promoting the development of mechanical fault diagnosis and prediction.

1. Introduction

With the rapid enhancement of computing capacity and the upcoming application of 5G communication infrastructure [1–3], the application of combined IoT [4–6], cloud computing [7–9] and big data processing technologies to predictive maintenance of mechanical equipment [10–15], is the focus of the next stage of development. Wind turbines [16–20], gearboxes [21–24], rolling bearings [25–28], diesel engines [29,30], and other mechanical equipment generally require manual maintenance and troubleshooting. In addition, due to the inability to diagnose the failure of mechanical equipment in time, corresponding safety accidents occur frequently. In an IoT environment, a large number of mechanical devices working data can be collected in a short period of time. How to use the collected data efficiently and to improve the accuracy of fault prediction are a challenging problem. In recent years, multiple sensors, such as those related to vibration, acceleration, temperature, and air pressure have been used, or multiple sensors of the same type have been combined to collect real-time operational status data related to different parts of mechanical equipment. Based on the IoT and combined with the cloud platform [31,32], and using multi-source sensing data fusion technology for big data analysis can effectively improve the accuracy of prediction, this has become a highly relevant research topic. Fig. 1 shows a topology diagram of predictive maintenance of mechanical equipment based on the IoT.

Since the development of data fusion technology, a variety of fusion models have been proposed, such as the JDL model [33], Boyd's OODA (Observe-Orient-Decide-Act) Loop model [34], Intelligence Cycle model [35], the Omnibus model [35,36], State Transition Data Fusion (STDF) model [37], Dasarathy model [38], Waterfall model [39], and Hybrid model [36]. However, the JDL

* Corresponding author.
E-mail address: sezhenliu@mail.scut.edu.cn (Z. Liu).

https://doi.org/10.1016/j.simpat.2019.101981
Received 19 July 2019; Received in revised form 27 August 2019; Accepted 3 September 2019
Available online 03 September 2019

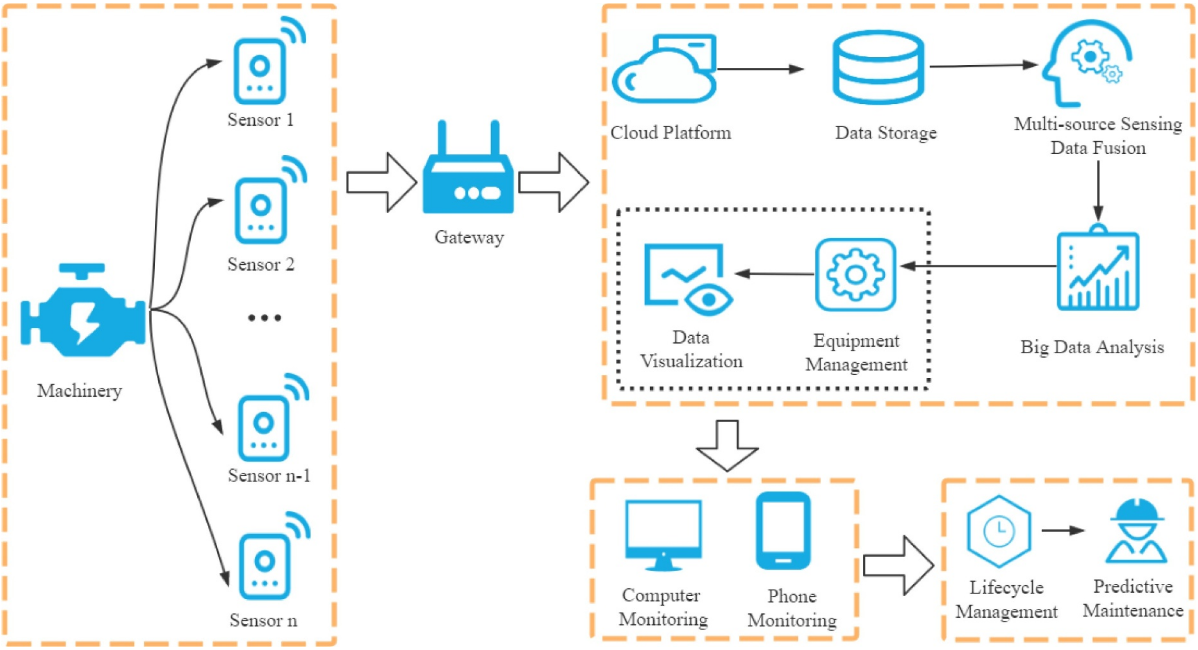


Fig. 1. Topology diagram of predictive maintenance of mechanical equipment based on IoT.

model remains the most frequently applied and studied, and has served as the foundation when developing other models. Among the researchers exploring mechanical fault diagnosis and prediction models, the Hierarchical fusion model developed by using Dasarathy model and Waterfall model is widely used.

Existing research on multi-source sensing data fusion algorithms can be divided into studies focusing on simple processing algorithms, probability-based algorithms, fuzzy logic-based algorithms, artificial intelligence algorithms, and other algorithms. The simple processing algorithms generally incorporate the weighted average method [40–43], and are based on weighting the average data acquired by the same type of sensors in continuous time series according to the chronological order, yielding the final fusion value. Probability-based algorithms include those based on Bayesian theory and D-S evidence theory. Available studies on mechanical fault diagnosis based on Bayesian theory can be generally divided into four theoretical research streams, focusing on Bayesian networks [44–49], naive Bayes [50,51], non-naive Bayes [52] and Bayesian inference [53]. Similarly, extant research on the D-S evidence theory is generally divided into studies on preprocessing the original evidence [54,55], improvement of combination rules [56], and a combination of other algorithms [57,58]. Accordingly, research based on fuzzy logic algorithm, including fuzzy set and rough set, and the application of fuzzy set in mechanical fault diagnostics mainly focuses on fuzzy logic reasoning, fuzzy clustering algorithm [59,60], intuitionistic fuzzy set, and a combination of other algorithms [61,62]. Rough set theory is mainly used for attribute reduction and noise removal, whereby the implementation method typically involves data analysis and decision making [63–67]. Artificial intelligence algorithms include neural networks, Support Vector Machine (SVM), and deep learning. Neural networks generally include Back-Propagation Neural Network (BPNN) [68], Radial Basis Function Neural Network (RBFNN) [69], and Elman Neural Network (ENN) [70] approaches. SVM can effectively solve the problem of dataset classification with a small number of samples [71–74]. There are also many studies on mechanical fault diagnosis using deep learning approach, specifically Convolutional Neural Network (CNN) [75], Deep Belief Network (DBN) [76], or Deep Neural Network (DNN) [77] algorithms. Other algorithms, such as ant colony algorithm [78,79], genetic algorithm [80,81] and Kalman filtering [82,83], are rarely used in the field of mechanical fault diagnosis.

In this work, taking the multi-source sensing data fusion process in Fig. 1 as the research content, we analyze and explain its development and application in predictive maintenance of mechanical equipment. The remainder of the paper is organized as follows: In Section 2, the basic concepts and development status of the JDL fusion model and the Hierarchical fusion model are introduced. This is followed by the discussion on the basic concepts of three fusion algorithms, namely neural networks, deep learning, and D-S evidence theory, presented in Section 3, along with the extent of their use in the field of mechanical fault diagnosis and prediction. In Section 4, the bearing dataset provided by Case Western Reserve University (CWRU) is adopted to conduct fault diagnosis and prediction experiments with the correlation fusion algorithm based on neural networks. The problems and challenges related to adopting multi-source sensing data fusion based on IoT in mechanical fault diagnosis and prediction are discussed in Section 5, before closing the paper with a brief summary provided in Section 6.

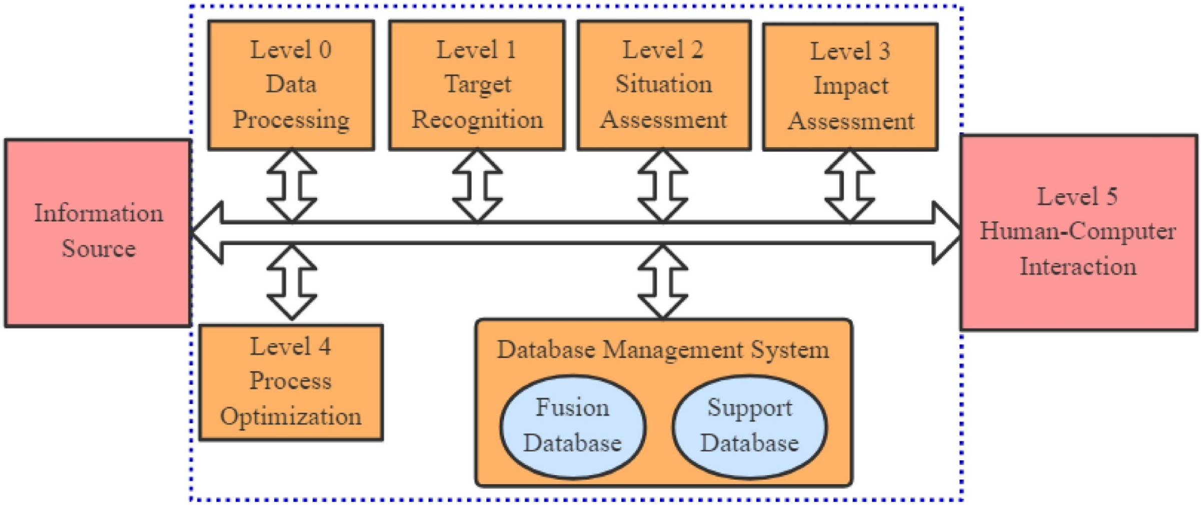


Fig. 2. JDL fusion model.

2. Analysis and comparison of commonly used multi-source sensing data fusion models

2.1. JDL Fusion model

The JDL model is the first fusion model to be proposed in the field of data fusion, and it is also the most well-developed model with the most in-depth application of theory [84]. The model structure is shown in Fig. 2.

In the mid-1980s, the US Department of Defense Data Fusion Joint Command Lab proposed the initial JDL model. At this stage, the model was divided into four levels, namely target recognition, situation assessment, threat assessment, and process optimization. By the end of the 1990s, Steinberg et al. [85] added a data preprocessing level as the first level. Owing to its widespread application, the JDL model gradually shifted from the original military sector to the civilian field, whereby the original threat assessment level was replaced by impact assessment. In addition, five levels of processing models were found not to be detailed enough to effectively reflect the operational process of multi-source information fusion, due to which the results of the corresponding fusion levels are not well applied. Therefore, Blasch et al. [86] proposed a user optimization level based on the original model, also known as human-computer interaction level. It is believed that the user optimization process can reflect the control effect of human subjectivity in the information fusion process, and thus effectively expand the applicability of the JDL model. So far, the JDL model has evolved from the initial 4-level to the 6-level fusion structure [87,88].

In the development process, the JDL model has been applied to different degrees in practical civil life. In their research on automatic driving or vehicle-mounted perception system technology, most scholars apply JDL model to carry out corresponding fault detection and diagnostics [89–91]. Bernardo et al. [92] relied on the JDL model, in combination with relevant information, to propose knowledge level juxtaposed structure, thus effectively improving the accuracy of aircraft operating state diagnostics. Mykich et al. [93] proposed the idea of information fusion based on the JDL model to explore the establishment of formal model of situational awareness system. Razzaq et al. [94] combined the JDL model with a context-aware dialog manager to propose a model for the healthcare dialog system. Other aspects, such as national cybersecurity construction [95], financial markets stock price forecast [96], heartbeat detection [97] and so on, also benefit from the JDL model.

Over the years, the JDL model has been developed substantially, and has been widely used in the military sector. Because the related fusion level processing targets are more diversified, each level of processing takes a long time, and the human participation has a greater impact on the final result. Consequently, it has not become the mainstream method in mechanical fault diagnosis and prediction based on IoT.

2.2. Hierarchical fusion model

The Hierarchical fusion model is developed on the basis of the Dasarathy model and the Waterfall model, and has emerged as the mainstream model for multi-source perceptual information fusion modeling. The model structure is shown in Fig. 3.

The Hierarchical fusion model is broadly divided into four stages, namely data preprocessing, data level fusion, feature level fusion, and decision level fusion. The output of each phase is used as the input to the next stage. The ideas behind data level fusion, feature level fusion, and decision level fusion are borrowed from the Dasarathy model, whereas the concept of merging data by hierarchy is based on the Waterfall model. The Dasarathy model comprises of five stages, based on the fusion function: data level fusion, feature selection and feature extraction, feature level fusion, pattern recognition and pattern processing, and decision level fusion. Through the step-by-step fusion at each stage, the accuracy of the fusion result can be effectively improved; however, the processing speed is slow. The Waterfall model pays more attention to the processing of the underlying data. According to the fusion

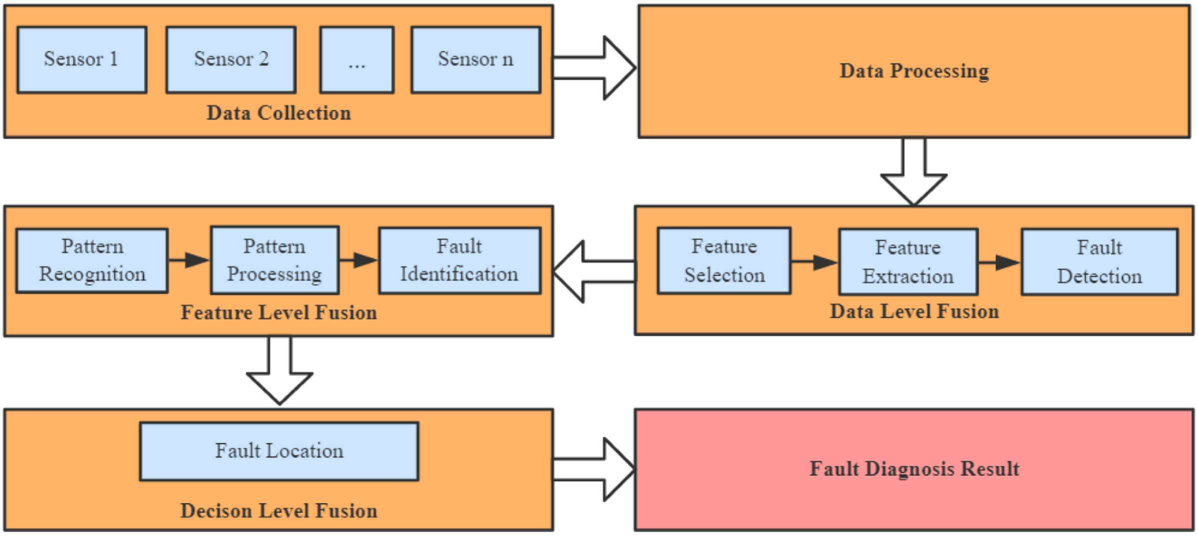


Fig. 3. Hierarchical fusion model.

process, it is divided into five progressive processes, namely signal acquisition, signal processing, feature extraction, pattern processing, and decision making. Although the Waterfall model divides the fusion process in more detail, its main disadvantage is that there is no clear feedback in each process. By extending the idea of the Dasarathy model and using the Waterfall model to process the underlying data, a data preprocessing module is added to form a hierarchical fusion model.

At present, the idea behind the hierarchical fusion model is widely used in mechanical fault diagnostics, whereby the final diagnosis results are obtained by applying data preprocessing, feature level fusion, and decision level fusion in two or three stages. For example, Wang et al. [98] extracted the time domain, frequency domain, and time-frequency domain characteristics of the vibration signal of the hydraulic system, and subsequently used the SVM for feature level fusion. Ma et al. [99] extracted the performance degradation characteristics of aircraft engines using stacked sparse autoencoders, performed feature level fusion through automatic learning, and applied logistic regression for decision level fusion. Zeng et al. [100] used SVM to characterize individual sensor data and then used D-S evidence theory for decision level fusion to achieve engine fault prediction. Hang et al. [61] proposed using empirical mode decomposition for data preprocessing, combined with fuzzy SVM for feature level fusion to achieve fault detection in wind turbines. Guan et al. [76] proposed reconstructing the original data using empirical mode decomposition and sample entropy, and then using DBN for feature level fusion to achieve fault prediction in rotating machinery. In terms of data level fusion, as the amount of original data is generally large and the amount of noise data is also significant, very few studies have been conducted on algorithms related to data level fusion.

Based on the aforementioned analysis, it can be concluded that the data level fusion, feature level fusion, and decision level fusion proposed in the hierarchical fusion model have different advantages and disadvantages. Data level fusion needs to deal with a large amount of information, poor fault tolerance, low level of fusion, and difficult implementation of the fusion algorithm, but the amount of information loss is minimal. The feature level fusion needs to process medium amount of information, fault tolerance, and fusion algorithm implementation, with medium information loss, but generally requires data preprocessing or feature extraction process. The decision level fusion needs to process a small amount of information, but it has good fault tolerance and it is easy to implement the fusion algorithm, despite the greatest information loss. Therefore, in various application scenarios, reasonable selection of the fusion hierarchy can effectively improve the performance of the final fault prediction model.

Although the hierarchical fusion model comprises of too many levels and suffers from slow processing speed, the amount of information will be greatly reduced after the fusion processing within each layer. In the process of mechanical fault diagnosis and prediction, it is generally necessary to transfer the collected data to the cloud server by means of the IoT. If the original data is transferred directly, however, the communication transmission cost will be prohibitive. Therefore, the hierarchical fusion model has certain reference value for the construction of the mechanical fault diagnosis and prediction model based on the IoT.

3. Analysis and comparison of typical multi-source sensing data fusion algorithms

Multi-source sensing data based on IoT generally includes multiple types of data, such as fault data collected at different locations, or fault data collected using different types of sensors, and the amount of data is very large. Therefore, it is difficult to directly use the original data for fault diagnosis and prediction. Mechanical processing can be effectively diagnosed by inputting raw data into a correlation algorithm for fusion processing. In this section, various fusion algorithms based on neural networks are introduced, before discussing deep learning and D-S evidence theory, and finally analyzing and contrasting them.

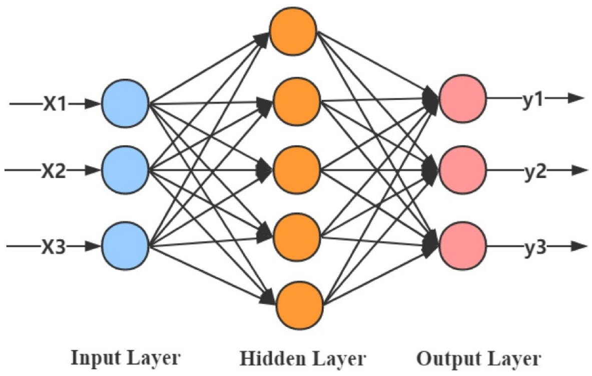


Fig. 4. The basic architecture of a three-layer neural network.

3.1. Fusion algorithms based on neural networks

3.1.1. Basic concepts

Because of their simple structure and the intelligence of processing data, neural networks have been widely used in the field of mechanical fault diagnosis and prediction in recent years. In the 1940s, the concept of neural network was proposed by McCulloch et al. [101], opening up the research direction of neural network models based on various processing algorithms. Neural networks are developed on the basis of multi-layer perceptron [102]. However, as multi-layer perceptron can only be used to perform forward propagation learning, it cannot be effectively used to solve nonlinear classification problems. The back-propagation algorithm eliminates this problem and enables neural networks to perform multi-layer learning. The term “neural networks,” also known as “artificial neural networks,” has a broad meaning. In most scenarios, neural network model is based on multi-layer perceptron and back-propagation algorithm [103–106].

At present, Back-Propagation Neural Network (BPNN), Radial Basis Function Neural Network (RBFNN), Elman Neural Network (ENN), Probabilistic Neural Network (PNN), Fuzzy Neural Network (FNN) and Wavelet Neural Network (WNN) are widely applied in the field of mechanical fault diagnosis research. The aforementioned six neural networks all include an input layer, a hidden layer, and an output layer, with the basic architecture shown in Fig. 4.

BPNN contains multiple hidden layers, and the transfer parameters from the output layer to the input layer are learned by the back-propagation algorithm. This process is iterated repeatedly to obtain the final prediction model. Based on the BPNN, ENN adds a feedback layer behind the hidden layer to provide the network with a memory function, which is a dynamic recurrent neural network. The structure of RBFNN is similar to that of BPNN, but with only three fixed layers, using a radial basis function as the neuron activation function. PNN adds a summation layer based on the RBFNN model. The summation layer determines the number of neurons based on the total number of categories of input data. FNN is developed through the combination of fuzzy logic theory and neural network model, and the input and output data, as well as the weight coefficients in the learning process, are all fuzzy sets. WNN is also based on BPNN, using wavelet basis functions as the transfer function of the hidden layer.

3.1.2. Comparative analysis of different fusion algorithms based on neural networks in mechanical fault diagnosis and prediction

Table 1 provides a comparison of recent studies in the field of mechanical fault diagnosis and prediction according to different fusion algorithms based on neural networks, along with their respective application characteristics.

It can be seen from Table 1 that, when using fusion algorithms based on neural networks for mechanical fault diagnosis and prediction based on IoT, it is essential to select (1) a reasonable data preprocessing method and (2) a reasonable neural network model, and use a heuristic algorithm such as particle swarm or other strategies to optimize the network model.

Fusion algorithms based on neural networks have lower requirements on the configuration of the cloud server, and can quickly complete the construction of the multi-source sensing data fusion model. However, it is necessary to combine appropriate data pre-processing methods or feature extraction methods. In the case of a clear understanding of the various fault data characteristics of mechanical equipment, the choice of fusion algorithms based on neural networks can achieve better results.

3.2. Fusion algorithms based on deep learning

3.2.1. Basic concepts

Algorithms based on Deep Learning (DL) can automatically combine simple features into more complex and abstract features, and can solve practical problems by combining new features. As a result, DL has attracted considerable research interest in various fields. Since the advent of deep learning technology in 2014, deep learning algorithms such as DNN and DBN have been used for mechanical fault diagnosis. Based on the investigation and analysis of relevant literature published in recent years, we conclude that the five fusion algorithms based on deep learning Convolutional Neural Network (CNN), Deep Belief Network (DBN), Deep Neural Network (DNN), Automatic Encoder (AE), and Long and Short-Term Memory network (LSTM)âare widely used in the field of mechanical equipment fault diagnosis and prediction. The corresponding network model structures are shown in Fig. 5.

Table 1
Comparison of application processes incorporated into different neural network models in mechanical fault diagnosis and prediction.

Model	Related Research Work	Application Characteristics Analysis
BPNN	Bin et al. [107] used Wavelet Packet Decomposition (WPD) and Empirical Mode Decomposition (EMD) for data preprocessing, combined with BPNN to achieve fault prediction of rotating machinery. Zhao et al. [108] used an improved shuffled frog leaping algorithm to optimize the weight parameters in the BPNN training process to achieve fault diagnosis in rolling bearings. Fengrong et al. [109] and Liu et al. [110] respectively used Wavelet Packet Transform (WPT) and gray model for data preprocessing, and combined with genetic algorithm and BPNN to build the fault prediction model. Dong et al. [111] used the bat algorithm to optimize the initial parameters of the BPNN to achieve fault diagnosis in the power transformer.	The use of back-propagation algorithm to update the weight parameters can effectively solve most nonlinear classification problems. If the initialization weight parameter setting is unreasonable, it will cause the network model to easily fall into local extreme points or would lead to slow convergence. In order to improve the generalization ability of the model, heuristic algorithms, such as shuffled frog leaping algorithm, bat algorithm, and genetic algorithm, can be used to optimize the weight parameters of the network model. In addition, data preprocessing methods such as EMD, WPD, and WPT can be reasonably selected, which can effectively improve the accuracy of fault diagnosis prediction using BPNN.
ENN	Sun et al. [112] proposed an improved ENN for fault diagnosis in underwater vehicles. Wang et al. [113] used the ENN for fault diagnosis in diesel engines. Jun et al. [114] used an improved particle swarm optimization algorithm combined with the ENN for fault detection in aeroengines. Fu et al. [115] used Ensemble Empirical Mode Decomposition (EEMD) combined with AdaBoost optimized ENN to achieve the diagnosis in bearing faults. Chemseddine et al. [116] combined with Hilbert empirical wavelet transform, Singular Value Decomposition (SVD) and ENN to achieve fault prediction in a gearbox.	With dynamic memory capability, it has good predictive ability for time series data, but it suffers from great instability when used for classification. In order to reduce the instability of the ENN classification, it can be optimized using a similar Adaboost classifier. In the case of a large data dimension, the convergence speed of the ENN is generally slow, and it can be optimized by improving the correlation between the feedback layer and the output layer. ENN uses the back-propagation algorithm to update the weight parameters. It can be combined with a heuristic algorithm, such as particle swarm optimization algorithm, to optimize the weight parameters.
RBFNN	Qi et al. [117] proposed solving the fault diagnosis of rotating machinery by calculating the angular domain adaptive frequency and combining the results with the RBFNN. Zhang et al. [118] combined WPD, with the differential evolution algorithm and RBFNN to achieve fault prediction in rolling bearings. Zhou et al. [119] combined the Kalman filter with the RBFNN to achieve fault diagnosis in a pumping unit.	Similar to BPNN, it can easily lead into local optimization or over-fitting. Heuristic algorithms such as genetic algorithm and differential evolution can be considered to optimize the weight parameters. In addition, effective data preprocessing can improve the generalization ability of RBFNN, for example, by calculating the angular velocity adaptive frequency during the operation of a rotating machine to resample the original data.
PNN	Yu et al. [120] used an improved EEMD and PNN to achieve fault diagnosis in thermal power generation equipment. Yi et al. [121] proposed a PNN based on adaptive strategy for fault diagnosis in a rotor. Liu et al. [122] used EMD and PNN to achieve fault diagnosis in a rotor. Reyes-Archundia et al. [123] used discrete wavelet transform (DWT) and PNN to diagnosis faults in power systems. Jiang et al. [124] proposed a feature fusion model based on information entropy and PNN for fault prediction in rotating machinery.	By using a probabilistic model to diagnose faults, it is not necessary to perform multiple iterations and there is no need to set the initial network weight. However, its commonly used parameter Spread has a great impact on performance. If it is not properly selected, it will produce incorrect prediction results. In order to solve this problem, a PNN optimization method based on adaptive strategy can be adopted. In addition, for feature extraction, it is more important to correctly extract the eigenvectors from the original data.
FNN	Gao et al. [125] used FNN to achieve fault diagnosis in substations. Gai et al. [126] used SVD and FNN to achieve fault diagnosis in diesel engines. Gai et al. [127] used EMD, SVD, and FNN to evaluate the degradation in bearing performance.	When the uncertainty in the original data is significant and the dataset is large, the use of FNN has a big advantage. In addition, if combined with SVD, EMD, and other data preprocessing methods, it can effectively improve the predictive ability of FNN.
WNN	Wu et al. [128] used an improved Laplace Eigen-Mapping (LEM) method and WNN for fault prediction in rolling bearings. Huitao et al. [129] used WNN to achieve fault diagnosis in wind turbines. Jin et al. [130] combined Wavelet Transform (WT) with WNN for monitoring hydraulic cylinder leakage accidents. Guo et al. [131] combined particle swarm optimization with WNN to detect the operating state of the cooling tower.	WNN can easily result in over-fitting or local optimization. Heuristic algorithms such as particle swarm optimization algorithm can be used to optimize the weight parameters. However, WNN can process time series data better. In addition, data preprocessing algorithms, such as WT and LEM, can be combined to improve model prediction accuracy.

The CNN consists of one or more convolutional and pooling layers, combined with a fully connected layer. The advantage of CNN is that it can better process image and natural language data sets, which can greatly reduce the number of parameters utilized in the learning process. The disadvantage is that the learning process is complicated and the training time is long. The DNN has at least one layer of hidden layer network structure trained by back-propagation and gradient descent algorithms. While DNN can yield more abstract features by adding multiple hidden layers, it can easily cause over-fitting problems. The DBN consists of multiple stacked Restricted Boltzmann Machines (RBM) combined with a layer of BPNN. The advantage of DBN is that, through the unsupervised learning process of RBM, it avoids the need for large amounts of labeled data required by supervised learning models such as BPNN. The disadvantage is that it is easy to ignore local information when characterizing and analyzing the data. AE is a rare unsupervised learning algorithm whose core function is to compress data, and thus it is often used for data denoising and visual dimension reduction. The main advantage of AE is that it can choose the appropriate extended model structure based on the actual processing data. Its disadvantage stems from the fact that an automatic encoder can only process one data type. LSTM is developed based on recursive neural network, and has the ability to learn long distance dependent data. Its structure consists of multiple cells connected in series, and each ”cell” is composed of input gate, output gate, forgetting gate, and a state unit. The advantage of LSTM stems from strong sequence modeling ability. Moreover, the gradient disappearance problem can be solved by the forgetting gate and output gate. The main disadvantage is that insufficient training data will lead to the problem of overfitting.

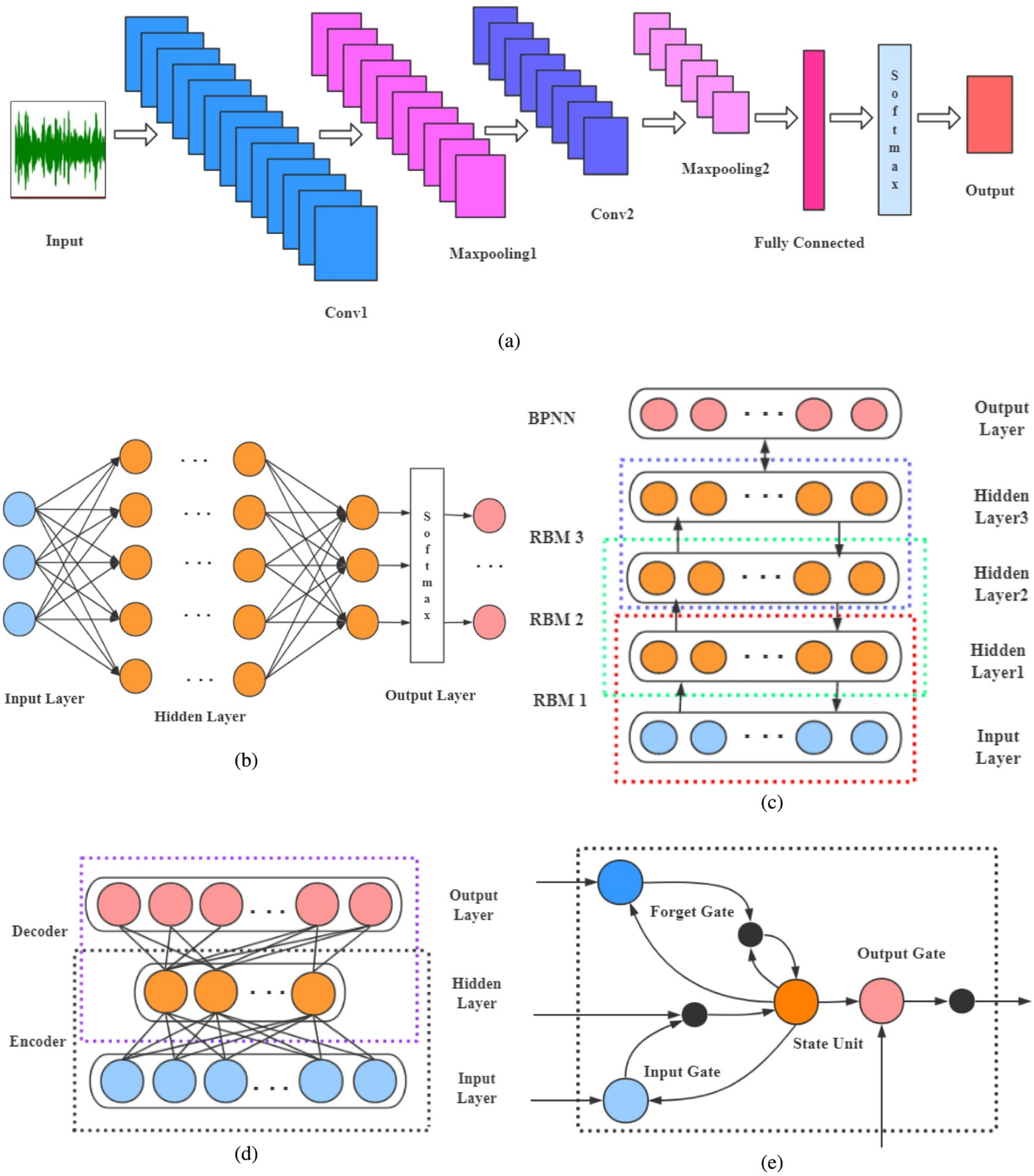


Fig. 5. The Architecture of (a) CNN, (b) DNN, (c) DBN, (d) AE, and (e) LSTM "Cell".

3.2.2. Application characteristics and effect analysis of different fusion algorithms based on deep learning in mechanical fault diagnosis and prediction

(a) Comparative Analysis of Application Characteristics

Table 2 provides a comparison of pertinent recent studies according to different fusion algorithms based on deep learning, along with the analysis of the application characteristics of various algorithms.

It can be seen from Table 2 that fusion algorithms based on deep learning can effectively identify high dimensional input data and can eliminate the need for data preprocessing and the artificial feature extraction process. In addition, migration learning can be used to achieve higher prediction accuracy when the raw dataset is smaller. However, the training time of fusion algorithms based on deep learning is too long and it is easy to fall into a local extremum. It is, however, possible to optimize the network model and weight parameters by using genetic algorithm and particle swarm optimization.

Fusion algorithms based on deep learning have high requirements for the configuration of cloud servers. However, in the case of a lack of understanding of the mechanical fault data characteristics, a fusion algorithm based on deep learning can automatically

Table 2
Comparison of application process of different fusion algorithms in our study.

Model	Related Research Work	Application Characteristics Analysis
CNN	Li et al. [132] used the Fast Fourier Transform (FFT) to extract the root mean square value of the vibration information, combined with the CNN and D-S evidence theory to achieve bearing fault prediction. Guo et al. [133] used continuous wavelet transform to obtain the scale map of the vibration signal, and combined them with CNN to achieve rotating machinery fault diagnosis. Hasan et al. [134] used the discrete orthogonal Stockwell transform vibration imaging method, combined them with CNN to achieve bearing fault diagnosis. Huang et al. [135] and Cao et al. [136] used the acceleration-time series image of the original vibration signal, combined them with CNN to achieve fault prediction in bearings and gears. Jian et al. [137] used the acceleration-time series image of the original vibration signal, combined CNN with D-S evidence theory to achieve motor fault diagnosis.	It is generally necessary to convert raw vibration information into image data, such as applying a root mean square value image, a continuous wavelet transform scale map, and an acceleration-time series image as training input data. The initialization process of the weight parameters will have a certain impact on the training time and the effect of CNN. It can be beneficial to combine CNN with the D-S evidence theory to achieve fault diagnosis, and the feature values extracted by CNN can be effectively used. In addition, the structural design of CNN is flexible, and combining different numbers of convolutional and pooling layers has a greater impact on the final prediction accuracy of the model.
DNN	Guo et al. [138] extracted multiple features from the raw data and combined them with DNN to identify the running state of a bearing. Zhang et al. [139] used DNN to directly learn the raw data to achieve bearing fault diagnosis. Yang et al. [140] conducted two learning trainings on the raw data through DNN to achieve bearing fault prediction.	Raw data can be directly applied for training and learning, and the time-frequency feature value of the vibration signal can also be used as the training data. The training time is long, and it is easy to fall into the local optimum or over-fit the data, but DNN can extract more abstract features from the raw data.
DBN	Tao et al. [141] used a variety of vibration signals as input data, which they combined with DBN to achieve fault diagnosis in bearings. Shao et al. [142] used FFT and DBN to achieve fault diagnosis in a motor. Xinqing et al. [143] combined sliding window spectral features with DBN to achieve fault diagnosis in a hydraulic system. He et al. [144] combined genetic algorithm with DBN to achieve fault prediction in gears. Xie et al. [145] proposed an end-to-end fault diagnosis model based on nesterov matrix optimization for adaptive DBN to achieve fault diagnosis in bearings. Tang et al. [146] proposed a learning model based on Fisher discriminant sparse representation and DBN to achieve fault diagnosis in complex systems about the chemical industry. Liang et al. [147] combined an improved integrated learning method with DBN to achieve fault diagnosis in bearings.	Unsupervised learning through multiple stacked RBMs, and then supervised learning through BP neural network, can adaptively fuse multi-feature data. However, the training time is too long and over-fitting and local convergence are commonly encountered issues. In order to solve this problem, genetic algorithm can be used to realize weight parameter optimization, and Nestrov matrix combined with DBN can be adopted to self-optimize weight parameters. DBN generally has better prediction and classification performance than BPNN, and can directly learn raw data without the need for artificial feature extraction. In addition, data preprocessing algorithms such as FFT and integrated learning can effectively improve the accuracy of DBN fault diagnosis prediction.
AE	Jiang et al. [148] used a Stacked Multilevel Denoising Autoencoders (SMDAE) to achieve fault prediction in wind turbines. Liu et al. [149] combined stacked autoencoders with BPNN to achieve gearbox fault diagnosis. Zhao et al. [150] combined the gray wolf optimizer algorithm with the enhanced autoencoder to achieve bearing fault diagnosis. Li et al. [151] used a fully connected Winner-Take-All autoencoder to achieve bearing fault prediction. Ahmed et al. [152] discussed the effects of Sparse Autoencoder (SAE) on the processing and classification performance in bearing vibration signals. Sohaib et al. [153] combined Complex Envelope Spectra (CES), stacked sparse autoencoder with Softmax classifier to achieve fault diagnosis in bearings. Shi et al. [154] used multiple Stacked Sparse Autoencoders (SSAE) to monitor tool status.	Its core structure consists of encoders and decoders. The encoder can automatically extract the feature vector in the input data, the decoder can convert the feature vector into input data, and the multi-dimensional feature fusion can be realized through the process of encoding and decoding. Application of data preprocessing methods such as CES and WPD can markedly improve the training efficiency of AE. Reasonable selection and design of the AE model structure (such as SMDAE, SAE, or SSAE) can significantly improve the fault diagnostic accuracy. While AE is suitable for feature compression fusion processing, it is necessary to combine the softmax and logistic classifiers to achieve the final fault diagnosis and prediction.
LSTM	Yu et al. [155] used a novel hierarchical algorithm based on stacked LSTM to achieve fault diagnosis in bearings. Zhang et al. [156] used a bi-directional LSTM model to predict the remaining useful life of an engine. Zhang et al. [157] used waveform entropy as a new feature extraction method, combined with LSTM and particle swarm optimization algorithm to achieve fault prediction in bearings. Xiao et al. [158] combined empirical statistical parameters, reproducible quantitative analysis with LSTM to achieve motor fault diagnosis. Lei et al. [159] proposed an end-to-end LSTM model for fault diagnosis in wind turbines.	The time information in the vibration data can be well utilized, and valuable short-term or long-term information can be selected for preservation and memory. Due to the adjustment and optimization of the model weight parameters, the training time is longer. It is considered that heuristic algorithms, such as particle swarm optimization algorithm, can be used to optimize the weight parameters. The use of reasonable time series data feature extraction methods, such as waveform entropy, can effectively improve the learning efficiency of the model.

extract features. As the dataset continues to increase, the prediction accuracy of the model will become higher.

(b) Comparative Analysis of Application Effect

In the field of mechanical fault diagnosis and prediction models and algorithms, most scholars used the bearing fault dataset provided by Case Western Reserve University (CWRU) for experiments. Table 3 is the experimental results of the fault diagnosis of the CWRU dataset using the diagnostic models based on CNN, DNN, DBN, AE and LSTM in the literature [134,139,147,153,155].

As can be seen from Table 3, comparing the experimental results published in [134] and [153], it is evident that the experimental results yielded by AE in [134] are higher than the highest accuracy of AE in [153], but the accuracy of ANN recognition in [134] is below the highest accuracy of BPNN in [153]. It can be further noted that the accuracy of using CNN in [155] is much lower than in [134]. In addition, when the ANN/BPNN method was applied in the control group, [139] had the highest recognition accuracy.

It can be seen from the analysis presented above that, even when the same dataset and the same type of fusion algorithm is used, the final recognition accuracy may be markedly different. The main reason for the difference in accuracy is that the researchers

Table 3
Classification accuracy of CNN, DNN, DBN, AE and LSTM when applied to the CWRU dataset. The Proposed method indicates the diagnostic method proposed in the paper, and Other methods indicates the method used in the control group in the paper, and the underline indicates that the corresponding method was not used in the control group.

Model	Proposed method	Other methods (Control group experiment)				
		ANN/BPNN	SVM	RBFNN	AE	CNN
CNN [134]	99%	66%	82%	-	97%	-
DNN [139]	100%	96.67%	100%	-	-	-
DBN [147]	96.95%	58.05%	47.1%	52.5%	-	-
AE [153]	90%(96.33%)	53%(87%)	-	-	-	-
LSTM [155]	98.8%	70%	90%	-	-	80%

generally only use a part of the data in the CWRU dataset, as well as rely on different application processes and methods. In addition, for the same fusion algorithm, different methods are applied in different studies for data preprocessing and feature extraction, which also has a great impact on the final accuracy.

3.3. Fusion algorithm based on dempster-Shafer evidence theory

3.3.1. Basic concepts

In the 1960s, American mathematician Dempster [160] proposed the predecessor of D-S evidence theory, and his student Shafer [161] subsequently formed a complete Dempster-Shafer evidence theory by introducing a trust function. The basic concept of D-S evidence theory is relatively simple, and it is also suitable in practical scenarios; thus, it is widely used in multi-source information fusion research. The basic concepts of D-S evidence theory is as follows:

Set $X = \{x_1, x_2, x_3, \ldots, x_n\}$, where X denotes the frame of discernment (FOD). Then, the meaning of 2^x as follows:

$$2^X = \{\varnothing, \{x_1\}, \{x_2\}, \{x_3\}, \ldots, \{x_n\}, \{x_1, x_2\}, \ldots, \{x_1, x_2, \ldots, x_i\}, X\} \tag{1}$$

In the D-S evidence theory, the BPA of Proposition A indicates the degree of support for A. The larger the BPA, the greater the degree of support for A, and vice versa. Combining different propositions of BPA can achieve comprehensive decision making, and the combination rules are as follows:

$$m(X) = m_1 \oplus m_2 \oplus \cdots \oplus m_n = \begin{cases} \frac{1}{1-K} \sum_{A_1 \cap A_2 \cdots \cap A_n = X} m_1(A_1) \cdots m_n(A_n),\ X \neq \varnothing \\ 0,\ X = \varnothing \end{cases} \tag{2}$$

where $K = \sum_{A_1 \cap A_2 \cdots \cap A_n} = \varnothing m_1(A_1) \cdots m_n(A_n)$ represents the conflict coefficient.

The belief function of Proposition A represents the sum of all BPAs belonging to the subset of Proposition A. The formula is as follows:

$$Bel(A) = \sum_{B \in A} m(B) \tag{3}$$

The plausibility function of Proposition A represents the sum of all BPAs that contain a subset of Proposition A. The formula is as follows:

$$Pl(A) = \sum_{B \cap A \neq \varnothing} m(B) \tag{4}$$

The interval [Bel(A), Pl(A)] represents the belief interval of Proposition A, and the relationship between Bel(A) and Pl(A) can be expressed as shown in Fig. 6 [162].

3.3.2. Application analysis of D-S evidence theory in mechanical fault diagnosis and prediction

Application of D-S evidence theory to achieve multi-source information fusion involves four steps: construct a FOD, define BPA,

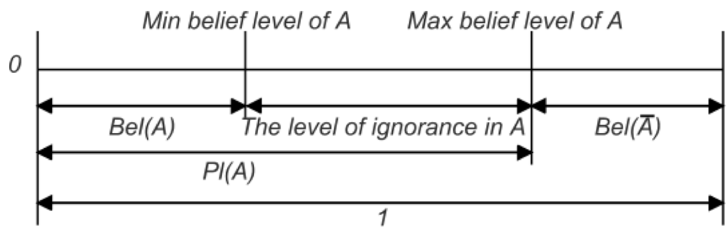


Fig. 6. Relationship between the belief function Bel(A) and the plausibility function Pl(A).

set evidence body combination rules, and set decision criteria. The decision criteria generally set multiple thresholds, such as ε_1, ε_2 and so on. When the original evidence is independent or less conflicting, the use of D-S evidence theory for information fusion can achieve satisfactory results. In the case of high conflicts in the original evidence, the shortcomings of D-S evidence theory emerge, as (1) some results are contrary to the actual situation; (2) some evidence can directly negate the final decision result; (3) the calculation of BPA is sensitive; and (4) if the number of evidence is large, it may lead to an exponential increase in the amount of calculation.

Careful investigation and analysis of pertinent literature has revealed that three approaches can be adopted to solve these problems, namely (a) other algorithms can be combined to achieve information fusion and decision analysis; (b) original evidence can be preprocessed to improve the recognition of different evidence bodies; and (c) evidence combination rules can be improved.

(a) Combining Other Algorithms

By combining with other algorithms, Kari et al. [57] presented a novel method for power transformer incipient fault prediction based on integrated adaptive neuro fuzzy inference system and D-S evidence theory. Wang et al. [58] proposed a fault diagnosis method for roller bearings based on a hybrid classifier ensemble approach and the improved D-S evidence theory. An et al. [163] effectively solved the problem of counter-intuition in the fusion of high-conflict evidence by combining fuzzy set theory and D-S evidence theory. Hu et al. [164] proposed using BPNN and SVM to perform data level and feature level fusion on the fault data of on-board diagnostic system, and finally make decision-level fusion based on D-S evidence theory. Li et al. [132] and Jian et al. [137] used CNN to extract fault information features before calculating the BPA of relevant fault features, and finally used D-S evidence theory for decision level fusion.

(b) Pre-processing the Original Evidence

Many scholars who opted for the preprocessing of original evidence used information entropy to obtain more information from the original evidence body, such as the proposition base of the evidence body or the uncertainty of the mass function. Given that the uncertainty of BPA is difficult to quantify, Pan et al. [165] proposed a new information entropy based on [166] combined with BPA and weighted Hartley entropy of a single event. On the other hand, L. Pan and Y. Deng [167] used the plausibility function and the belief function to analyze the uncertainty of BPA, and combined the findings with Deng entropy [168] to propose a new information entropy. Jousselme el al[169]. proposed an information entropy based on the ambiguity of the quantified belief function for the uncertainty of the original evidence. Because Deng entropy does not consider the FOD scale, which cannot effectively quantify the difference between different single evidence bodies of the same quality value distribution, several research groups [162,170–172] have utilized an improved information entropy based on Deng entropy to quantify the uncertainty of the body of uncertain evidence.

In addition to improving the existing information entropy, some scholars have used existing information entropy to combine evidence distance [173], cosine similarity, Jensen-Shannon divergence or other methods, such as Gaussian fault model for evidence preprocessing. The application process is shown in Table 4.

(c) Improving Evidence Combination Rules

In terms of improving the rules of evidence combination, Lin et al. [184] proposed a new method of evidence combination based on Mahalanobis distance, which can yield the same results as the traditional combination rules, as well as solve the anti-direct problem of high conflict information fusion. Ye et al. [185] used the gun distance function and the spectral angle cosine function to correct the original evidence, and then proposed an improved combination rule. When attempting fusion of high-conflict evidence bodies, Ao et al. [186] regard conflicts between evidences as local conflicts, thus proposing an improved evidence combination rule. Liu et al. [187] proposed an improved D-S evidence combination rule for bearing fault diagnosis.

3.3.3. Comparative analysis of different improved information entropies

D-S evidence theory has been extensively studied. The findings yielded, however, indicate that, faced with different types of original evidence, even when using the same information entropy for evidence preprocessing, the effects may be markedly different. Therefore, for the preprocessing of original evidence based on belief entropy, it is necessary to make a detailed comparison and

Table 4

Evidence preprocessing using existing information entropy or other methods.

Solution	Related Research Work
Information entropy combined with evidence distance	Wang et al. [174] calculated the distance between different pieces of evidence by the number of focal points of the original evidence, combined with the information entropy, to calculate the corresponding credibility, and finally used the combination rule of D-S evidence theory to achieve multi-source data fusion. The authors of [175–177] proposed combining evidence distance and belief entropy, and used weighted average method to correct conflict evidence.
Information entropy combined with cosine similarity	The authors of [178,179] proposed combining cosine similarity and information entropy to measure the similarity between conflicting bodies of evidence.
Information entropy combined with Jensen-Shannon divergence	Xiao F[180]. used Jensen-Shannon divergence to measure the difference and degree of conflict between available evidence bodies, combined with information entropy, to measure the uncertainty and credibility in these evidence bodies, and achieve fusion of multiple evidence bodies.
Using other methods	Jiang et al. [181] used the perceptual data to establish a Gaussian fault model and generated a BPA for the corresponding event, combined with the weighted average method for the fusion between multiple BPAs. Qin et al. [182] combined the interval mathematics with k-means++ clustering methods to calculate BPA. Sun et al. [183] proposed a method based on minimum spanning tree to identify FOD and optimize the D-S evidence theory.

Table 5
Some of existing entropies to measure uncertainty degree. These entropies citation information can be referred to Section 3.3.2, simply using log for $\log_2$ in the table.

Proposer	Formula
Pan et al. [165]	$H_{PQ}(m) = \sum_{A \in 2^X} m(A) \log\left[\frac{\lvert A\rvert}{Pl\_Pm(A)}\right]$
Jirousek et al. [166]	$H_{JS}(m) = \sum_{x \in X} Pl\_Pm(x) \log\left[\frac{1}{Pl\_Pm(x)}\right] + \sum_{A \in 2^X} m(A) \log(\lvert A\rvert)$
L. Pan and Y. Deng [167]	$H_{PD}(m) = -\sum_{A \in 2^X} \frac{Bel(A)+Pl(A)}{2} \log\left[\frac{Bel(A)+Pl(A)}{2(2^{\lvert A\rvert}-1)}\right]$
Y. Deng [168]	$H_D(m) = -\sum_{A \in 2^X} m(A) \log\left[\frac{m(A)}{2^{\lvert A\rvert}-1}\right]$
Jousselme et al. [169]	$H_J(m) = \sum_{x \in X} BetP(x) \log\left[\frac{1}{BetP(x)}\right]$
Chen et al. [170]	$H_{CD}(m) = -\sum_{A \in 2^X} m(A) \log\left[\frac{m(A)}{2^{\lvert A\rvert}-1} \frac{\lvert A\rvert}{\lvert X\rvert}\right]$
Tang et al. [171]	$H_{TD}(m) = -\sum_{A \in 2^X} \frac{\lvert A \mid m(A)}{\lvert X\rvert} \log\left[\frac{m(A)}{2^{\lvert A\rvert}-1}\right]$
Zhou et al. [172]	$H_{ZT}(m) = -\sum_{A \in 2^X} m(A) \log\left[\frac{m(A)}{2^{\lvert A\rvert}-1} e^{\frac{\lvert A\rvert-1}{\lvert x\rvert}}\right]$
Cui et al. [162]	$H_{ZT}(m) = -\sum_{A \in 2^X} m(A) \log\left[\frac{m(A)}{2^{\lvert A\rvert}-1} e^{\sum_{B \subseteq 2^{\lvert X\rvert} \& B \neq A} \frac{\lvert A \cap B\rvert}{2^{\lvert X\rvert}-1}}\right]$

analysis of pertinent preconditions. Table 5 provides the core formula for evidence preprocessing using improved information entropy.

According to the example 4.5 in [168], set m({3, 4, 5}) = 0.05, m({6}) = 0.05, m(A) = 0.8, m(X) = 0.1, where X = {1,2,3, …, 14,15}, A is a continuous sequence starting from 1, and A is set to {1}, {1,2}, {1,2,3}, …, {1, …, 10}, …, {1, …, 14}. Using the nine belief entropies proposed in Table 5, the calculated uncertainty results in simulation are shown in Table 6. Fig. 4 provides graphical visualization of the data analyses reported in Table 6.

It can be seen from Table 6 and Fig. 7 that the information entropy proposed by Jousselme et al. cannot distinguish between evidence bodies A containing different lengths. Other belief entropies can better distinguish evidence body A of different lengths, and the main difference is that the degree of differentiation is different. The information entropy calculation results proposed by Cui et al. are almost identical to those yielded by the Deng entropy, indicating that they are equivalent in some scenarios. In addition, the information entropy calculation results proposed by Chen et al. are larger than Deng entropy, Tang et al. are the opposite, but the results are relatively close.

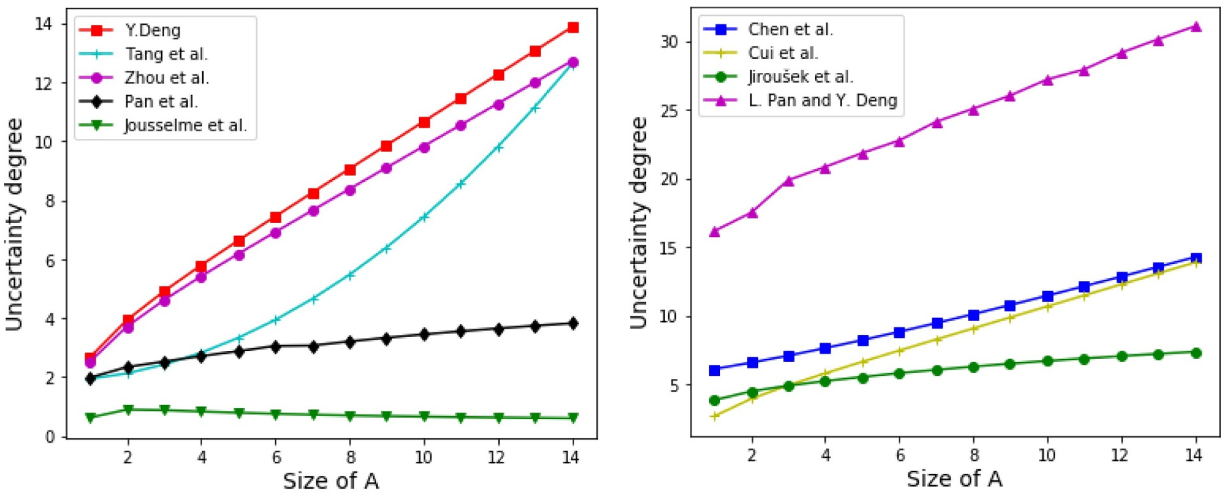
Therefore, based on a mechanical equipment or practical application scenario, an appropriate information entropy can be selected or redesigned for evidence preprocessing, it is helpful to further improve the performance of decision analysis based on big data.

4. Fault diagnosis and prediction experiment involving fusion algorithms based on neural networks

As the six different fusion algorithm models based on neural networks are similar in structure, they have been widely used in mechanical fault diagnosis and prediction research in recent years. Therefore, in the current investigation, CWRU dataset design-related experiments are adopted, as they are widely used in the field of mechanical fault diagnosis and prediction. The aim is to explore the differences in the practical application of six different fusion algorithms based on neural networks.

Table 6
The value of uncertainty degree by using different entropies. The length of evidence body A is constantly changing.

Cases	H_{PQ}	H_{JS}	H_{PD}	H_D	H_J	H_{CD}	H_{TD}	H_{ZT}	H_{CL}
A = {1}	1.9757	3.8322	16.1443	2.6623	0.6203	6.0992	1.9351	2.5180	2.6622
A = {1,2}	2.3362	4.4789	17.4916	3.9303	0.8915	6.5672	2.1213	3.7090	3.9302
A = {1,2,3}	2.5232	4.8870	19.8608	4.9082	0.8712	7.0772	2.4186	4.6100	4.9080
A = {1,…,4}	2.7085	5.2250	20.8229	5.7878	0.8271	7.6248	2.8200	5.4127	5.7875
A = {1,…,5}	2.8749	5.5200	21.8314	6.6256	0.7858	8.2051	3.3249	6.1736	6.6253
A = {1,…,6}	3.0516	5.8059	22.7521	7.4441	0.7503	8.8131	3.9336	6.9151	7.4437
A = {1,…,7}	3.0647	6.0429	24.1131	8.2532	0.7204	9.4443	4.6472	7.6473	8.2528
A = {1,…,8}	3.2042	6.2772	25.0685	9.0578	0.6949	10.0947	5.4662	8.3749	9.0573
A = {1,…,9}	3.3300	6.4921	26.0212	9.8600	0.6731	10.7610	6.3911	9.1002	9.8595
A = {1,…,10}	3.4445	6.6903	27.1947	10.6612	0.6542	11.4406	7.4222	9.8244	10.6606
A = {1,…,11}	3.5497	6.8743	27.9232	11.4617	0.6377	12.1311	8.5597	10.5480	11.4611
A = {1,…,12}	3.6469	7.0461	29.1370	12.2620	0.6232	12.8310	9.8037	11.2714	12.2614
A = {1,…,13}	3.7374	7.2071	30.1231	13.0622	0.6103	13.5388	11.1543	11.9946	13.0615
A = {1,…,14}	3.8219	7.3587	31.0732	13.8622	0.5987	14.2533	12.6115	12.7177	13.8615



(a) (b)

Fig. 7. Comparison of the uncertainty values calculated by nine different information entropies. (a) Comparison of uncertainty values of five different belief entropies; (b) Comparison of uncertainty values of other four different belief entropies.

4.1. Experimental simulation environment and dataset introduction

4.1.1. Simulation environment

Our experimental computer was equipped with Ubuntu 18.04 operating system, 32G memory and Intel i7-8700K CPU model. Python programming language were used as the basic development platform, whereby the simulation tool we used in the six different fusion algorithms based on neural networks can be seen in Table 7.

4.1.2. CWRU Dataset introduction

In different operating conditions of the bearing motor, Case Western Reserve University Laboratory used multiple accelerometers to collect vibration data at different locations of the bearing motor, thus obtaining the CWRU dataset. The bearing motor used in the CWRU dataset was artificially damaged by the electric spark, simulating actual failures of the bearing. The fault locations were separately at the inner raceway, rolling element and outer raceway of the bearing driving end or fan end. The bearing rolling element was available in four different types of diameters, each representing a different type of fault. There were 4 different load types and 4 different speeds of bearing motor, respectively representing the fault of the bearing motor under different conditions. The bearing motor used in the CWRU dataset is shown in Fig. 8 [132].

4.2. Experimental content and simulation results

In this section, four eigenvalues of Root Mean Square (RMS), Kurtosis, Crest factor and Skewness were selected for fault diagnosis and prediction, which calculation formulas [115,188,189] can be seen in Table 8.

Select part of the data in the CWRU dataset served as the source of experimental data. The conditions for collecting normal state data was that the Motor Load value was 1 hp and the speed value was 1772 rpm. The fault data acquisition conditions were that the frequency value was 12000sample/s and the fault diameter value was 0.007 in. The fault locations included three different parts in the drive end: the inner raceway, the outer raceway and the rolling raceway. Fig. 9 shows the change of vibration acceleration of the bearing driving end with time under normal and inner raceway fault conditions.

Feature extraction was performed using 480 consecutive sampling points. Each feature value was taken as 250 groups, and the order of feature values was randomly assigned (Using the random function in Python). The corresponding simulation results are shown in Fig. 10.

It can be seen from Fig. 10 that, in the normal state, the inner raceway fault, the outer raceway fault, and the rolling raceway fault corresponding to the bearing have different characteristic values, and the degree of discrimination also differs. Therefore, fault

Table 7
The simulation tool of six different fusion algorithms based on neural networks.

Model	Simulation Tool	Model	Simulation Tool
BPNN	Sklearn	ENN	Neurolab
RBFNN	Keras	FNN	Min Max Neural Network Library
PNN	On the basis of RBFNN	WNN	Tensorflow

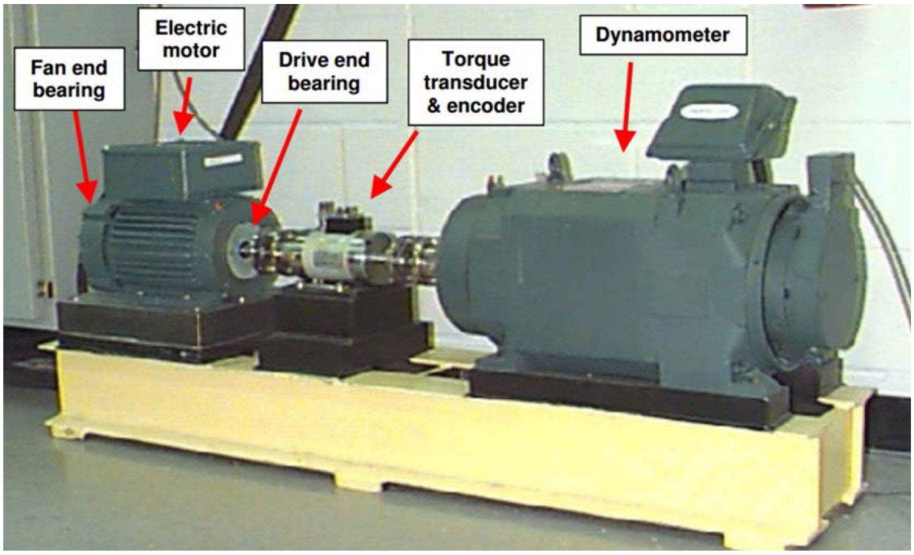


Fig. 8. Bearing motor used in the CWRU dataset.

Table 8
Feature names and calculation formulas.

Feature Name	Calculation Formula
RMS	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i)^2}$
Standard deviation (STD)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$
Kurtosis	$\frac{1}{n*(STD)^4}\sum_{i=1}^{n}(x_i-\bar{x})^4$
Crest factor	$\frac{\max(x)}{RMS}$
Skewness	$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{STD}\right)^3$

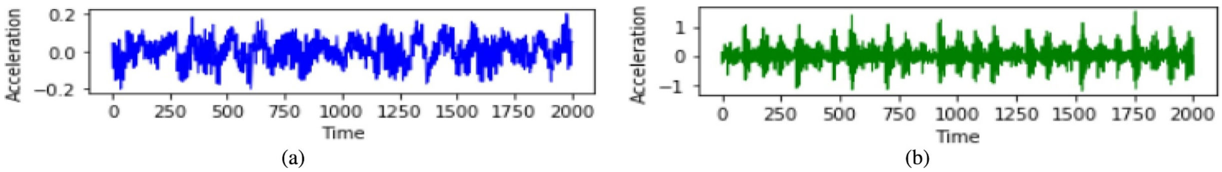


Fig. 9. The change of the bearing drive end acceleration with time under (a) normal conditions and (b) inner raceway fault conditions. The data used 2000 consecutive sampling points.

classification diagnosis combined with four different eigenvalues is more accurate than using a single eigenvalue for fault diagnosis. Thus, the four eigenvalues of RMS, Kurtosis, Crest factor and Skewness were classified and identified using BPNN, ENN, RBFNN, PNN, FNN, and WNN, and the prediction accuracy of different neural network models was compared and analyzed. The experimental results are shown in Table 9.

It can be seen from Table 9 that the classification accuracy of the training set and the test set ratio of 75% and 25% is higher than 50%, because the amount of training data is small. In terms of BPNN, the classification accuracy of the 4-layer structure is higher than that of the 3-layer structure. Based on the overall experimental results, at 91.6%, BPNN has the highest classification accuracy rate, followed by RBFNN and WNN, both of which exceeded 80%, PNN (66%) and finally FNN (65.2%). The classification accuracy of ENN is only 27.2% mainly due to the fact that ENN is a kind of memory neural network, which is suitable for training with continuous time series data. The data used in the experiments conducted as a part of this work are calculated by relevant eigenvectors and randomly distorted the order.

5. Problems and challenges

Application of multi-source sensing data fusion technology based on IoT can be divided into three stages, namely data acquisition,

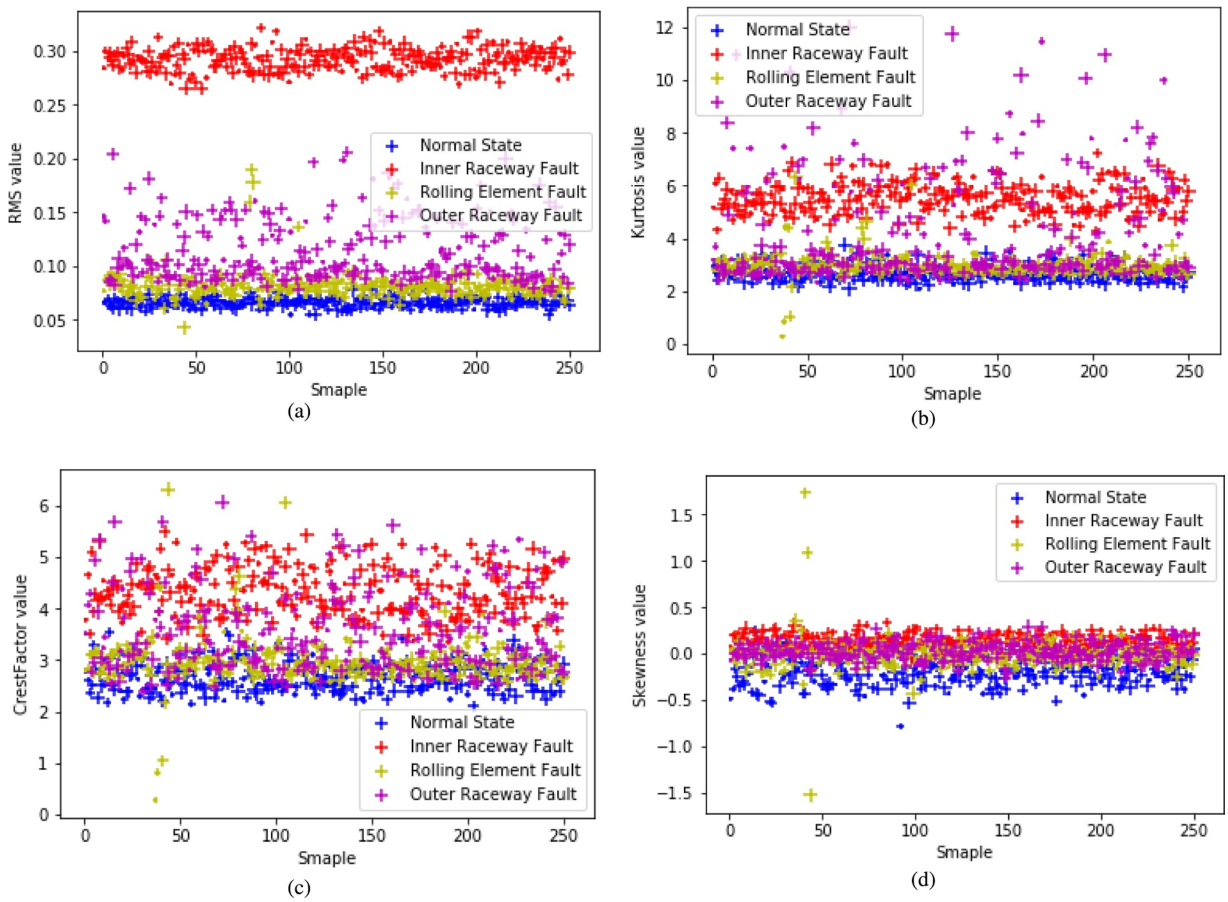


Fig. 10. Distribution of RMS, Kurtosis, Crest factor, and Skewness eigenvalue data extracted using CWRU dataset. (a) Root-mean-square value of all bearing signals; (b) Kurtosis value of all bearing signals; (c) Crest factor value of all bearing signals; (d) Skewness value of all bearing signals.

Table 9
Classification accuracy of BPNN, ENN, RBFNN, PNN, FNN and WNN when applied to the CWRU dataset.

Neural Network	Proportion		Number of Network Layers	Accuracy
	Training Set	Test Set		
BPNN	50%	50%	3	89%
	75%	25%	3	90%
	50%	50%	4	90%
	75%	25%	4	91.6%
ENN	50%	50%	4	25%
	75%	25%	4	27.2%
RBFNN	50%	50%	3	77%
	75%	25%	3	83.6%
PNN	50%	50%	4	62%
	75%	25%	4	66%
FNN	50%	50%	4	60.4%
	75%	25%	4	65.2%
WNN	50%	50%	3	83%
	75%	25%	3	84.8%

fusion model creation, and fusion algorithm application. The data acquisition stage has a significant impact on the industrial application of mechanical fault diagnosis and prediction technology. The problems and challenges faced are summarized below.

(1) Problems related to sensor selection: When presented with actual mechanical equipment, it may be necessary to select different sensor groups for data collection in different environments. For the same type of sensor, the industry standards corresponding to the sensors generated by different manufacturers may also be different, resulting in different degrees of error. In addition, the sensor group may be different when applied to different mechanical equipment. In the future development process, it is hoped that a set of sensors covering all mechanical equipment fault diagnosis and prediction scenarios can be developed, and that these sensors can be reasonably matched for data collection according to actual conditions.

(2) Data acquisition frequency and communication: In the data acquisition process, if the acquisition frequency is too high, data redundancy will occur. Conversely, if the sampling frequency is too low, the obtained data cannot be reliably used in fault diagnosis. Mechanical equipment is generally continuously operated, and the generated dataset is relatively large. If the initial fusion processing is not performed at the collection end, the communication cost related to data transfer will be greatly increased. In the future development process, setting a data acquisition frequency standard for different sensors and initial integration of the original data is one of the main objectives.

According to the development status of fusion models and fusion algorithms in the context of multi-source sensing data fusion, the problems and challenges faced can be summarized as follows:

(1) The fusion model is not uniform: In the field of mechanical fault diagnosis and prediction research, there is no unified modeling framework to deal with various mechanical fault diagnosis scenarios. Most of the existing fusion models are specific to a particular type of device. For example, the famous JDL model is mainly used in the military field. In the future research process, it would be beneficial to design and build a unified fusion framework for diagnosing faults in mechanical equipment.

(2) Uncertainty of the original data: In the process of data collection, the actual data obtained contains a lot of noise due to uncontrollable environmental factors. If the original data is used directly for feature extraction and data fusion, the resulting decision results are often wrong. Therefore, when presented with raw data, it is necessary to combine the fusion model and the fusion algorithm to select an appropriate data preprocessing method. In the future development process, it would be advantageous to establish a set of data preprocessing methods for different sensors used in fault diagnosis and prediction specifically aimed at mechanical equipment.

(3) Long running time: When fusion algorithms based on deep learning are used for training, the running time required to find suitable hyperparameters is generally long, and sometimes over-fitting occurs. Fusion algorithms based on neural networks generally require manual feature extraction, resulting in longer runtimes. Most fusion algorithms are aimed at feature level fusion and decision level fusion research. There are few algorithms at the data level fusion level. In the future development process, further development of data level fusion algorithms is thus needed.

6. Conclusions

The rapid development of artificial intelligence technology and the advent of 5G, make it possible for predicting mechanical faults based on the IoT. Big data processing and analysis is the key to multi-source sensing data fusion technology. This paper studied the development and application of multi-source sensing data fusion models and algorithms in mechanical equipment fault diagnosis and prediction. In terms of fusion models, the JDL fusion model was relatively mature, but it was not best suited for mechanical fault diagnosis applications, because the ideas related to hierarchical fusion model were more widely used in mechanical fault diagnosis. In terms of fusion algorithms, those algorithms that were based on neural network have been widely used. However, they need to be combined with data preprocessing algorithms such as WPD, EMD, and SVD to achieve higher accuracy. The related fusion algorithms based on deep learning could be directly applied to train and extract the original data. Generally, data preprocessing is not required, but the training time was long, and the performance requirements of the machine equipment are high. D-S evidence theory was prone to paradoxes when presented with high-conflict evidence. Consequently, most scholars preprocessed the original evidence by adopting different belief entropy models to optimize the D-S evidence theory.

Mechanical intelligence has been gaining prominence in the mechanical fault diagnosis and prediction research. It was shown that, we could consider choosing the neural network and D-S evidence theory combined with the appropriate data preprocessing algorithm for fault diagnosis and prediction in some special scenarios. In the case of mechanical fault diagnosis related to more complex scenes, the correlation fusion algorithm based on deep learning could effectively avoid the process of artificial feature extraction. In addition, a good set of hyperparameters was significant for fusion algorithms based on deep learning. Combining the heuristic algorithms, such as evolutionary algorithm, particle swarm optimization, and genetic algorithm, to optimize the network structure and parameters could be beneficial in some cases. Combined with the multi-source sensing data based on the IoT and the cloud server, rationally adopting fusion algorithms based on neural network and deep learning, and the advantages of D-S evidence theory in decision-level fusion, as well as designing and creating more intelligent big data analysis fusion models and fusion algorithms, could be an important research trends for the future.

References

[1] J. Cheng, W. Chen, F. Tao, C.-L. Lin, Industrial iot in 5g environment towards smart manufacturing, J. Industr. Inf. Integr. 10 (2018) 10–19.

[2] M.R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, L. Ladid, Internet of things in the 5g era: enablers, architecture, and business models, IEEE J. Sel. Areas Commun. 34 (3) (2016) 510–527.

[3] S. Li, L. Da Xu, S. Zhao, 5g internet of things: A survey, J. Industr. Inf. Integr. 10 (2018) 1–9.
[4] M. Huang, Y. Chen, Guest editorial: internet of things and intelligent devices and services, CAAI Trans. Intell. Technol. 3 (2) (2018) 73–74.
[5] Y. Chen, Analyzing and visual programming internet of things and autonomous decentralized systems, Simul. Modell. Pract. Theory 65 (2016) 1–10.
[6] G.L. Stavrinides, H.D. Karatza, A hybrid approach to scheduling real-time iot workflows in fog and cloud environments, Multimed. Tools Appl. (2018) 1–17.
[7] W.-T. Tsai, G. Qi, Y. Chen, Choosing cost-effective configuration in cloud storage, 2013 IEEE Eleventh International Symposium on Autonomous Decentralized Systems (ISADS), IEEE, 2013, pp. 1–8.
[8] I. Mavridis, H. Karatza, Combining containers and virtual machines to enhance isolation and extend functionality on cloud computing, Future Gener. Comput. Syst. 94 (2019) 674–696.
[9] I.A. Moschakis, H.D. Karatza, Towards scheduling for internet-of-things applications on clouds: a simulated annealing approach, Concurrency Comput. 27 (8) (2015) 1886–1899.
[10] R. Duwairi, H. Karatza, Advances on information and communication systems, Simul. Modell. Pract. Theory 100 (64) (2016) 1–2.
[11] J. Wang, L. Zhang, L. Duan, R.X. Gao, A new paradigm of cloud-based predictive maintenance for intelligent manufacturing, J. Intell. Manuf. 28 (5) (2017) 1125–1137.
[12] Y. Chen, H. Hu, Internet of intelligent things and robot as a service, Simul. Modell. Pract. Theory 34 (2013) 159–171.
[13] M. Baptista, S. Sankararaman, I.P. de Medeiros, C. Nascimento Jr, H. Prendinger, E.M. Henriques, Forecasting fault events for predictive maintenance using data-driven techniques and arma modeling, Comput. Industr. Eng. 115 (2018) 41–53.
[14] J. Yan, Y. Meng, L. Lu, L. Li, Industrial big data in an industry 4.0 environment: challenges, schemes, and applications for predictive maintenance, IEEE Access 5 (2017) 23484–23491.
[15] F. Civerchia, S. Bocchino, C. Salvadori, E. Rossi, L. Maggiani, M. Petracca, Industrial internet of things monitoring solution for advanced predictive maintenance applications, J. Industr. Inf. Integr. 7 (2017) 4–12.
[16] Y. Pan, R. Hong, J. Chen, J. Singh, X. Jia, Performance degradation assessment of a wind turbine gearbox based on multi-sensor data fusion, Mech. Mach. Theory 137 (2019) 509–526.
[17] W. Teng, X. Ding, H. Cheng, C. Han, Y. Liu, H. Mu, Compound faults diagnosis and analysis for a wind turbine gearbox via a novel vibration model and empirical wavelet transform, Renew. Energy 136 (2019) 393–402.
[18] H. Habibi, I. Howard, S. Simani, Reliability improvement of wind turbine power generation using model-based fault detection and fault tolerant control: a review, Renew. Energy (2018) 877–896.
[19] R. Chen, X. Huang, L. Yang, X. Xu, X. Zhang, Y. Zhang, Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform, Comput. Ind. 106 (2019) 48–59.
[20] A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, J. Keane, G. Nenadic, Machine learning methods for wind turbine condition monitoring: a review, Renew. Energy (2018) 620–635.
[21] D. Han, N. Zhao, P. Shi, Gear fault feature extraction and diagnosis method under different load excitation based on emd, pso-svm and fractal box dimension, J. Mech. Sci. Technol. 33 (2) (2019) 487–494.
[22] W. Du, J. Zhou, Z. Wang, R. Li, J. Wang, Application of improved singular spectrum decomposition method for composite fault diagnosis of gear boxes, Sensors 18 (11) (2018) 3804.
[23] Z. Wang, J. Wang, W. Du, Research on fault diagnosis of gearbox with improved variational mode decomposition, Sensors 18 (10) (2018) 3510.
[24] X. Chen, G. Cheng, H. Li, Y. Li, Research of planetary gear fault diagnosis based on multi-scale fractal box dimension of ceemd and elm. Strojniski Vestnik/ Journal of Mechanical Engineering 63 (1) (2017) 45–55.
[25] X. Li, W. Zhang, Q. Ding, Cross-domain fault diagnosis of rolling element bearings using deep generative neural networks, IEEE Trans. Ind. Electron. 66 (7) (2018) 5525–5534.
[26] Z. Xiang, X. Zhang, W. Zhang, X. Xia, Fault diagnosis of rolling bearing under fluctuating speed and variable load based on tco spectrum and stacking auto-encoder, Measurement 138 (2019) 162–174.
[27] M. Zair, C. Rahmoune, D. Benazzouz, Multi-fault diagnosis of rolling bearing using fuzzy entropy of empirical mode decomposition, principal component analysis, and som neural network, Proc. Instit. Mech. Eng., Part C: J. Mech. Eng. Sci. 233 (9) (2019) 3317–3328.
[28] H. Li, T. Liu, X. Wu, Q. Chen, Application of eemd and improved frequency band entropy in bearing fault feature extraction, ISA Trans. 88 (2019) 170–185.
[29] X. Bi, S. Cao, D. Zhang, Diesel engine valve clearance fault diagnosis based on improved variational mode decomposition and bispectrum, Energies 12 (4) (2019) 661.
[30] J. Zhang, C.-W. Liu, F.-R. Bi, X.-B. Bi, X. Yang, Fault feature extraction of diesel engine based on bispectrum image fractal dimension, Chin. J. Mech. Eng. 31 (1) (2018) 40.
[31] H.D. Karatza, G.L. Stavrinides, Modeling and simulation of cloud computing and big data, Simul. Modell. Pract. Theory (2019) 1–2.
[32] Y. He, X. Wang, Y. Chen, Z. Du, W. Huang, X. Chai, A simulation cloud monitoring framework and its evaluation model, Simul. Modell. Pract. Theory 38 (2013) 20–37.
[33] E. Blasch, A. Steinberg, S. Das, J. Llinas, C. Chong, O. Kessler, E. Waltz, F. White, Revisiting the jdl model for information exploitation, Proceedings of the 16th International Conference on Information Fusion, IEEE, 2013, pp. 129–136.
[34] B. Brehmer, The dynamic ooda loop: Amalgamating boyds ooda loop and the cybernetic approach to command and control, Proceedings of the 10th international command and control research technology symposium, (2005), pp. 365–368.
[35] M. Bedworth, J. O'Brien, The omnibus model: a new model of data fusion? IEEE Aerosp. Electron. Syst. Mag. 15 (4) (2000) 30–36.
[36] J. Esteban, A. Starr, R. Willetts, P. Hannah, P. Bryanston-Cross, A review of data fusion models and architectures: towards engineering guidelines, Neural Comput. Appl. 14 (4) (2005) 273–281.
[37] D.A. Lambert, Stdf model based maritime situation assessments, 2007 10th International Conference on Information Fusion, IEEE, 2007, pp. 1–8.
[38] B.V. Dasarathy, Sensor fusion potential exploitation-innovative architectures and illustrative applications, Proc. IEEE 85 (1) (1997) 24–38.
[39] E. Blasch, P. Hanselman, Information fusion for information superiority, Proceedings of the IEEE 2000 National Aerospace and Electronics Conference. NAECON 2000. Engineering Tomorrow (Cat. No. 00CH37093), IEEE, 2000, pp. 290–297.
[40] S. Li, X. Kang, L. Fang, J. Hu, H. Yin, Pixel-level image fusion: a survey of the state of the art, Inf. Fusion 33 (2017) 100–112.
[41] H. Lidong, Z. Wei, W. Jun, S. Zebin, Combination of contrast limited adaptive histogram equalisation and discrete wavelet transform for image enhancement, IET Image Proc. 9 (10) (2015) 908–915.
[42] B.S. Kumar, Image fusion based on pixel significance using cross bilateral filter, Signal Image Video Process. 9 (5) (2015) 1193–1204.
[43] R. Nayak, D. Patra, Super resolution image reconstruction using weighted combined pseudo-zernike moment invariants, AEU-Int. J. Electron. Commun. 70 (11) (2016) 1496–1505.
[44] H.-B. Jun, D. Kim, A bayesian network-based approach for fault analysis, Expert Syst. Appl. 81 (2017) 332–348.
[45] Q. Jiang, B. Huang, S.X. Ding, X. Yan, Bayesian fault diagnosis with asynchronous measurements and its application in networked distributed monitoring, IEEE Trans. Ind. Electron. 63 (10) (2016) 6316–6324.
[46] Z. Wang, Z. Wang, X. Gu, S. He, Z. Yan, Feature selection based on bayesian network for chiller fault diagnosis from the perspective of field applications, Appl. Therm. Eng. 129 (2018) 674–683.
[47] B. Cai, Y. Zhao, H. Liu, M. Xie, A data-driven fault diagnosis methodology in three-phase inverters for pmsm drive systems, IEEE Trans. Power Electron. 32 (7) (2016) 5590–5600.
[48] K. Li, Q. Zhang, K. Wang, P. Chen, H. Wang, Intelligent condition diagnosis method based on adaptive statistic test filter and diagnostic bayesian network, Sensors 16 (1) (2016) 76.
[49] B. Cai, L. Huang, M. Xie, Bayesian networks in fault diagnosis, IEEE Trans. Ind. Inf. 13 (5) (2017) 2227–2240.
[50] N. Zhang, L. Wu, J. Yang, Y. Guan, Naive bayes bearing fault diagnosis based on enhanced independence of data, Sensors 18 (2) (2018) 463.

[51] K. Vernekar, H. Kumar, K. Gangadharan, Engine gearbox fault diagnosis using empirical mode decomposition method and naïve bayes algorithm, Sādhanā 42 (7) (2017) 1143–1153.
[52] M.Y. Asr, M.M. Ettefagh, R. Hassannejad, S.N. Razavi, Diagnosis of combined faults in rotary machinery by non-naive bayesian approach, Mech. Syst. Signal Process. 85 (2017) 56–70.
[53] M.M. Islam, J. Kim, S.A. Khan, J.-M. Kim, Reliable bearing fault diagnosis using bayesian inference-based multi-class support vector machines, J. Acoust. Soc. Am. 141 (2) (2017) EL89–EL95.
[54] H. Zhang, Y. Deng, Engine fault diagnosis based on sensor data fusion considering information quality and evidence theory, Adv. Mech. Eng. 10 (11) (2018). 1687814018809184
[55] F. Lu, C. Jiang, J. Huang, Y. Wang, C. You, A novel data hierarchical fusion method for gas turbine engine performance fault diagnosis, Energies 9 (10) (2016) 828.
[56] R.R. Yager, On the dempster-shafer framework and new combination rules, Inf. Sci. (Ny) 41 (2) (1987) 93–137.
[57] T. Kari, W. Gao, D. Zhao, Z. Zhang, W. Mo, Y. Wang, L. Luan, An integrated method of anfis and dempster-shafer theory for fault diagnosis of power transformer, IEEE Trans. Dielectr. Electr. Insul. 25 (1) (2018) 360–371.
[58] Y. Wang, F. Liu, A. Zhu, Bearing fault diagnosis based on a hybrid classifier ensemble approach and the improved dempster-shafer theory, Sensors 19 (9) (2019) 2097.
[59] A.R. Ramos, J.M.B. de Lázaro, A.J. da Silva Neto, C.C. Corona, J.L. Verdegay, O. Llanes-Santiago, An approach to fault diagnosis using fuzzy clustering techniques, Advances in Fuzzy Logic and Technology 2017, Springer, 2017, pp. 232–243.
[60] C. Li, M. Cerrada, D. Cabrera, R.V. Sanchez, F. Pacheco, G. Ulutagay, J. Valente de Oliveira, A comparison of fuzzy clustering algorithms for bearing fault diagnosis, J. Intell. & Fuzzy Syst. (Preprint) (2018) 1–16.
[61] J. Hang, J. Zhang, M. Cheng, Application of multi-class fuzzy support vector machine classifier for fault diagnosis of wind turbine, Fuzzy Sets Syst. 297 (2016) 128–140.
[62] Y. Li, M. Xu, H. Zhao, W. Huang, Hierarchical fuzzy entropy and improved support vector machine based binary tree approach for rolling bearing fault diagnosis, Mech. Mach. Theory 98 (2016) 114–132.
[63] F. Pacheco, M. Cerrada, R.-V. Sánchez, D. Cabrera, C. Li, J.V. de Oliveira, Attribute clustering using rough set theory for feature selection in fault severity classification of rotating machinery, Expert Syst. Appl. 71 (2017) 69–86.
[64] P. Lu, X.-h. Wang, J.-m. Xiao, Method of fault diagnosis in power system based on rough set theory and graph theory, Control Decis. (2013) 511–516 + 524.
[65] W. Huang, F. Kong, X. Zhao, Spur bevel gearbox fault diagnosis using wavelet packet transform and rough set theory, J. Intell. Manuf. 29 (6) (2018) 1257–1271.
[66] Q.-j. Xie, H.-x. Zeng, L. Ruan, X.-m. Chen, H.-l. Zhang, Transformer fault diagnosis based on bayesian network and rough set reduction theory, IEEE 2013 Tencon-Spring, IEEE, 2013, pp. 262–266.
[67] P. Lu, D. Li Wenhuib amd Huang, Transformer fault diagnosis method based on graph theory and rough set, J. Intell. Fuzzy Syst. (2018) 223–230.
[68] H. Mekki, A. Mellit, H. Salhi, Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules, Simul. Modell. Pract. Theory 67 (2016) 1–13.
[69] H.R. Baghaee, M. Mirsalim, G.B. Gharehpetian, H.A. Talebi, Application of rbf neural networks and unscented transformation in probabilistic power-flow of microgrids including correlated wind/pv units and plug-in hybrid electric vehicles, Simul. Modell. Pract. Theory 72 (2017) 51–68.
[70] T. Zheng, Y. Zhang, Y. Li, L. Shi, Real-time combustion torque estimation and dynamic misfire fault diagnosis in gasoline engine, Mech Syst Signal Process 126 (2019) 521–535.
[71] L.-l. Jiang, H.-k. Yin, X.-j. Li, S.-w. Tang, Fault diagnosis of rotating machinery based on multisensor information fusion using svm and time-domain features, Shock Vib. 2014 (2014) 1–8.
[72] L. Pan, D. Zhu, S. She, A. Song, X. Shi, S. Duan, Gear fault diagnosis method based on wavelet-packet independent component analysis and support vector machine with kernel function fusion, Adv. Mech. Eng. 10 (11) (2018). 1687814018811036
[73] F. Jiang, Z. Zhu, W. Li, Y. Ren, G. Zhou, Y. Chang, A fusion feature extraction method using eemd and correlation coefficient analysis for bearing fault diagnosis, Appl. Sci. 8 (9) (2018) 1621.
[74] G. Qi, Z. Zhu, K. Erqinhu, Y. Chen, Y. Chai, J. Sun, Fault-diagnosis for reciprocating compressors using big data and machine learning, Simul. Modell. Pract. Theory 80 (2018) 104–127.
[75] X. Zhu, D. Hou, P. Zhou, Z. Han, Y. Yuan, W. Zhou, Q. Yin, Rotor fault diagnosis using a convolutional neural network with symmetrized dot pattern images, Measurement 138 (2019) 526–535.
[76] Z. Guan, Z. Liao, K. Li, P. Chen, A precise diagnosis method of structural faults of rotating machinery based on combination of empirical mode decomposition, sample entropy, and deep belief network, Sensors 19 (3) (2019) 591.
[77] M. Ou, H. Wei, Y. Zhang, J. Tan, A dynamic adam based deep neural network for fault diagnosis of oil-immersed power transformers, Energies 12 (6) (2019) 995.
[78] Z. Liu, W. Guo, Z. Tang, Y. Chen, Multi-sensor data fusion using a relevance vector machine based on an ant colony for gearbox fault detection, Sensors 15 (9) (2015) 21857–21875.
[79] X. Zhang, W. Chen, B. Wang, X. Chen, Intelligent fault diagnosis of rotating machinery using support vector machine with ant colony algorithm for synchronous feature selection and parameter optimization, Neurocomputing 167 (2015) 260–279.
[80] M. Cerrada, G. Zurita, D. Cabrera, R.-V. Sánchez, M. Artés, C. Li, Fault diagnosis in spur gears based on genetic algorithm and random forest, Mech. Syst. Signal Process. 70 (2016) 87–103.
[81] M. Cerrada, R. Sánchez, D. Cabrera, G. Zurita, C. Li, Multi-stage feature selection by using genetic algorithms for fault diagnosis in gearboxes based on vibration signal, Sensors 15 (9) (2015) 23903–23926.
[82] Z. Liu, H. He, Sensor fault detection and isolation for a lithium-ion battery pack in electric vehicles using adaptive extended kalman filter, Appl. Energy 185 (2017) 2033–2044.
[83] Q. Zhang, Adaptive kalman filter for actuator fault diagnosis, Automatica 93 (2018) 333–342.
[84] D.L. Hall, J.M. Jordan, Human-centered information fusion, Artech House, 2010.
[85] A.N. Steinberg, C.L. Bowman, Revisions to the Jdl Data Fusion Model, Handbook of multisensor data fusion, CRC press, 2008, pp. 65–88.
[86] E. Blasch, S. Plano, Dfig level 5 (user refinement) issues supporting situational assessment reasoning, 2005 7th International Conference on Information Fusion, 1 IEEE, 2005, pp. xxxv–xliii.
[87] S. Schreiber-Ehle, W. Koch, The jdl model of data fusion applied to cyber-defencea review paper, 2012 Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF), IEEE, 2012, pp. 116–119.
[88] L. Snidaro, I. Visentini, J. Llinas, G.L. Foresti, Context in fusion: some considerations in a jdl perspective, Proceedings of the 16th International Conference on Information Fusion, IEEE, 2013, pp. 115–120.
[89] M. Realpe, B.X. Vintimilla, L. Vlacic, A fault tolerant perception system for autonomous vehicles, 2016 35th Chinese Control Conference (CCC), IEEE, 2016, pp. 6531–6536.
[90] M. Realpe, B. Vintimilla, L. Vlacic, Sensor fault detection and diagnosis for autonomous vehicles, MATEC Web of Conferences, 30 EDP Sciences, 2015, p. 04003.
[91] M. Realpe, B. Vintimilla, L. Vlacic, Towards fault tolerant perception for autonomous vehicles: Local fusion, 2015 IEEE 7th International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), IEEE, 2015, pp. 253–258.
[92] J.T. Bernardo, Cognitive and functional frameworks for hard/soft fusion for the condition monitoring of aircraft, 2015 18th International Conference on Information Fusion (Fusion), IEEE, 2015, pp. 287–294.
[93] K. Mykich, Y. Burov, Algebraic model for knowledge representation in situational awareness systems, 2016 XIth International Scientific and Technical Conference Computer Sciences and Information Technologies (CSIT), IEEE, 2016, pp. 165–167.
[94] M.A. Razzaq, W.A. Khan, S. Lee, Intent-context fusioning in healthcare dialogue-based systems using jdl model, International Conference on Smart Homes and

Health Telematics, Springer, 2017, pp. 61–72.
[95] I. Swart, B.V. Irwin, M.M. Grobler, Adaptation of the Jdl Model for Multi-sensor National Cyber Security Data Fusion, National Security: Breakthroughs in Research and Practice, IGI Global, 2019, pp. 92–107.
[96] L. Evans, M. Owda, K. Crockett, A.F. Vilas, Big data fusion model for heterogeneous financial market data (findf), Proceedings of SAI Intelligent Systems Conference, Springer, 2018, pp. 1085–1101.
[97] T. Zia, Z. Arif, Probabilistic data fusion model for heart beat detection from multimodal physiological data, Turkish J. Electr. Eng. Comput. Sci. 25 (1) (2017) 449–460.
[98] L. Wang, X.-q. Wu, C. Zhang, H. Shi, Hydraulic system fault diagnosis method based on a multi-feature fusion support vector machine, J. Eng. 2019 (13) (2019) 215–218.
[99] J. Ma, H. Su, W.-l. Zhao, B. Liu, Predicting the remaining useful life of an aircraft engine using a stacked sparse autoencoder with multilayer self-learning, Complexity 2018 (2018) 1–13.
[100] R. Zeng, L. Zhang, J. Mei, H. Shen, H. Zhao, Fault detection in an engine by fusing information from multivibration sensors, Int. J. Distrib. Sens. Netw. 13 (7) (2017). 1550147717719057
[101] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (4) (1943) 115–133.
[102] J. Tang, C. Deng, G.-B. Huang, Extreme learning machine for multilayer perceptron, IEEE Trans. Neural Netw. Learn. Syst. 27 (4) (2015) 809–821.
[103] R.S. Gunerkar, A.K. Jalan, S.U. Belgamwar, Fault diagnosis of rolling element bearing based on artificial neural network, J. Mech. Sci. Technol. 33 (2) (2019) 505–511.
[104] J.B. Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, F. Fnaiech, Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals, Applied Acoustics 89 (2015) 16–27.
[105] R. Ahmed, M. El Sayed, S.A. Gadsden, J. Tjong, S. Habibi, Automotive internal-combustion-engine fault detection and classification using artificial neural network techniques, IEEE Trans. Veh. Technol. 64 (1) (2014) 21–33.
[106] T. Waqar, M. Demetgul, Thermal analysis mlp neural network based fault diagnosis on worm gears, Measurement 86 (2016) 56–66.
[107] G. Bin, J. Gao, X. Li, B. Dhillon, Early fault diagnosis of rotating machinery based on wavelet packetsempirical mode decomposition feature extraction and neural network, Mech. Syst. Signal Process. 27 (2012) 696–711.
[108] Z. Zhao, Q. Xu, M. Jia, Improved shuffled frog leaping algorithm-based bp neural network and its application in bearing early fault diagnosis, Neural Comput. Appl. 27 (2) (2016) 375–385.
[109] F. Bi, Y. Liu, Fault diagnosis of valve clearance in diesel engine based on bp neural network and support vector machine, Trans. Tianjin Univ. 22 (6) (2016) 536–543.
[110] H.-F. Liu, C. Ren, Z.-T. Zheng, Y.-J. Liang, X.-J. Lu, Study of a gray genetic bp neural network model in fault monitoring and a diagnosis system for dam safety, ISPRS Int. J. Geoinf. 7 (1) (2017) 4.
[111] H. Dong, X. Yang, A. Li, A novel method for power transformer fault diagnosis based on bat-bp algorithm, 2018 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), IEEE, 2018, pp. 566–569.
[112] Y.-s. Sun, Y.-m. Li, G.-c. Zhang, Y.-h. Zhang, H.-b. Wu, Actuator fault diagnosis of autonomous underwater vehicle based on improved elman neural network, J. Central South Univ. 23 (4) (2016) 808–816.
[113] Y. Wang, F. Zhang, T. Cui, J. Zhou, Fault diagnosis for manifold absolute pressure sensor (map) of diesel engine based on elman neural network observer, Chinese J. Mech. Eng. 29 (2) (2016) 386–395.
[114] J. Pi, J. Huang, L. Ma, Aeroengine fault diagnosis using optimized elman neural network, Math. Prob. Eng. 2017 (2017) 1–8.
[115] Q. Fu, B. Jing, P. He, S. Si, Y. Wang, Fault feature selection and diagnosis of rolling bearings based on eemd and optimized elman_adaboost algorithm, IEEE Sens. J. 18 (12) (2018) 5024–5034.
[116] R. Chemseddine, M. Boualem, B. Djamel, F. Semchedine, Gear fault feature extraction and classification of singular value decomposition based on hilbert empirical wavelet transform, J. Vibroengineering 20 (4) (2018) 1603–1618.
[117] X. Qi, Z. Yuan, X. Han, Diagnosis of misalignment faults by tacholess order tracking analysis and rbf networks, Neurocomputing 169 (2015) 439–448.
[118] Q. Zhang, J. Gao, H. Dong, Y. Mao, Wpd and de/bbo-rbfnn for solution of rolling bearing fault diagnosis, Neurocomputing 312 (2018) 27–33.
[119] W. Zhou, X. Li, J. Yi, H. He, A novel ukf-rbf method based on adaptive noise factor for fault diagnosis in pumping unit, IEEE Trans. Ind. Inf. 15 (3) (2018) 1415–1424.
[120] Y. Yu, W. Li, D. Sheng, J. Chen, A novel sensor fault diagnosis method based on modified ensemble empirical mode decomposition and probabilistic neural network, Measurement 68 (2015) 328–336.
[121] J.-H. Yi, J. Wang, G.-G. Wang, Improved probabilistic neural networks with self-adaptive strategies for transformer fault diagnosis problem, Adv. Mech. Eng. 8 (1) (2016). 1687814015624832
[122] D. Liu, H. Zeng, Z. Xiao, L. Peng, O. Malik, Fault diagnosis of rotor using emd thresholdingbased de-noising combined with probabilistic neural network. J. Vibroengineering 19 (8) (2017) 5920–5931.
[123] E. Reyes-Archundia, J. Guardado, J. Gutiérrez-Gnecchi, E. Moreno-Goytia, N. Guerrero-Rodriguez, Fault analysis in tcsc-compensated lines using wavelets and a pnn, Neural Comput. Appl. (2018) 1–14.
[124] Q. Jiang, Y. Shen, H. Li, F. Xu, New fault recognition method for rotary machinery based on information entropy and a probabilistic neural network, Sensors 18 (2) (2018) 337.
[125] J.-P. Gao, C.-B. Xu, L. Zhang, J.-L. Zheng, H. Shu, X. Yuan, A method of information fusion based on fuzzy neural network and its application, ITM Web of Conferences, 11 EDP Sciences, 2017, p. 01015.
[126] J. Gai, Y. Hu, Research on fault diagnosis based on singular value decomposition and fuzzy neural network, Shock Vib. 2018 (2018) 1–7.
[127] J. Gai, Y. Hu, J. Shen, A bearing performance degradation modeling method based on emd-svd and fuzzy neural network, Shock Vib. 2019 (2019) 1–10.
[128] L. Wu, B. Yao, Z. Peng, Y. Guan, Fault diagnosis of roller bearings based on a wavelet neural network and manifold learning, Appl. Sci. 7 (2) (2017) 158.
[129] C. Huitao, J. Shuangxi, W. Xianhui, W. Zhiyang, Fault diagnosis of wind turbine gearbox based on wavelet neural network, J. Low Frequency Noise Vibration Active Control 37 (4) (2018) 977–986.
[130] Y. Jin, C. Shan, Y. Wu, Y. Xia, Y. Zhang, L. Zeng, Fault diagnosis of hydraulic seal wear and internal leakage using wavelets and wavelet neural network, IEEE Trans. Instrum. Meas. (99) (2018) 1–9.
[131] Q. Guo, X. Qi, Z. Wei, Q. Yin, P. Sun, P. Guo, J. Liu, Modeling and characteristic analysis of fouling in a wet cooling tower based on wavelet neural networks, Appl. Therm. Eng. 152 (2019) 907–916.
[132] S. Li, G. Liu, X. Tang, J. Lu, J. Hu, An ensemble deep convolutional neural network model with improved ds evidence fusion for bearing fault diagnosis, Sensors 17 (8) (2017) 1729.
[133] S. Guo, T. Yang, W. Gao, C. Zhang, Y. Zhang, An intelligent fault diagnosis method for bearings with variable rotating speed based on pythagorean spatial pyramid pooling cnn, Sensors 18 (11) (2018) 3857.
[134] M.J. Hasan, J.-M. Kim, Bearing fault diagnosis under variable rotational speeds using stockwell transform-based vibration imaging and transfer learning, Appl. Sci. 8 (12) (2018) 2357.
[135] R. Huang, Y. Liao, S. Zhang, W. Li, Deep decoupling convolutional neural network for intelligent compound fault diagnosis, IEEE Access 7 (2018) 1848–1858.
[136] P. Cao, S. Zhang, J. Tang, Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning, IEEE Access 6 (2018) 26241–26253.
[137] X. Jian, W. Li, X. Guo, R. Wang, Fault diagnosis of motor bearings based on a one-dimensional fusion neural network, Sensors 19 (1) (2019) 122.
[138] L. Guo, H. Gao, H. Huang, X. He, S. Li, Multifeatures fusion and nonlinear dimension reduction for intelligent bearing condition monitoring, Shock Vib. 2016 (2016) 1–10.
[139] R. Zhang, Z. Peng, L. Wu, B. Yao, Y. Guan, Fault diagnosis from raw sensor data using deep neural networks considering temporal coherence, Sensors 17 (3)

(2017) 549.
[140] Y. Yang, P. Fu, Y. He, Bearing fault automatic classification based on deep learning, IEEE Access 6 (2018) 71540–71554.
[141] J. Tao, Y. Liu, D. Yang, Bearing fault diagnosis based on deep belief network and multisensor information fusion, Shock Vib. 2016 (2016) 1–9.
[142] S.-Y. Shao, W.-J. Sun, R.-Q. Yan, P. Wang, R.X. Gao, A deep learning approach for fault diagnosis of induction motors in manufacturing, Chin. J. Mech. Eng. 30 (6) (2017) 1347–1356.
[143] X. Wang, J. Huang, G. Ren, D. Wang, A hydraulic fault diagnosis method based on sliding-window spectrum feature and deep belief network. J. Vibroengineering 19 (6) (2017) 4272–4284.
[144] J. He, S. Yang, C. Gan, Unsupervised fault diagnosis of a gear transmission chain using a deep belief network, Sensors 17 (7) (2017) 1564.
[145] J. Xie, G. Du, C. Shen, N. Chen, L. Chen, Z. Zhu, An end-to-end model based on improved adaptive deep belief network and its application to bearing fault diagnosis, IEEE Access 6 (2018) 63584–63596.
[146] Q. Tang, Y. Chai, J. Qu, H. Ren, Fisher discriminative sparse representation based on dbn for fault diagnosis of complex system, Appl. Sci. 8 (5) (2018) 795.
[147] T. Liang, S. Wu, W. Duan, R. Zhang, Bearing fault diagnosis based on improved ensemble learning and deep belief network, Journal of Physics: Conference Series, 1074 IOP Publishing, 2018, p. 012154.
[148] G. Jiang, H. He, P. Xie, Y. Tang, Stacked multilevel-denoising autoencoders: a new representation learning approach for wind turbine gearbox fault diagnosis, IEEE Trans. Instrum. Meas. 66 (9) (2017) 2391–2402.
[149] G. Liu, H. Bao, B. Han, A stacked autoencoder-based deep neural network for achieving gearbox fault diagnosis, Math. Problems Eng. 2018 (2018) 1–10.
[150] L. Zhao, X. Wang, A deep feature optimization fusion method for extracting bearing degradation features, IEEE Access 6 (2018) 19640–19653.
[151] C. Li, W. Zhang, G. Peng, S. Liu, Bearing fault diagnosis using fully-connected winner-take-all autoencoder, IEEE Access 6 (2017) 6103–6115.
[152] H. Ahmed, M.D. Wong, A.K. Nandi, Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features, Mech. Syst. Signal Process. 99 (2018) 459–477.
[153] M. Sohaib, J.-M. Kim, Reliable fault diagnosis of rotary machine bearings using a stacked sparse autoencoder-based deep neural network, Shock Vib. 2018 (2018) 1–11.
[154] C. Shi, G. Panoutsos, B. Luo, H. Liu, B. Li, X. Lin, Using multiple-feature-spaces-based deep learning for tool condition monitoring in ultraprecision manufacturing, IEEE Trans. Ind. Electron. 66 (5) (2018) 3794–3803.
[155] L. Yu, J. Qu, F. Gao, Y. Tian, A novel hierarchical algorithm for bearing fault diagnosis based on stacked lstm, Shock Vib. 2019 (2019) 1–10.
[156] A. Zhang, H. Wang, S. Li, Y. Cui, Z. Liu, G. Yang, J. Hu, Transfer learning with deep recurrent neural networks for remaining useful life estimation, Appl. Sci. 8 (12) (2018) 2416.
[157] B. Zhang, S. Zhang, W. Li, Bearing performance degradation assessment using long short-term memory recurrent network, Comput. Ind. 106 (2019) 14–29.
[158] D. Xiao, Y. Huang, C. Qin, H. Shi, Y. Li, Fault diagnosis of induction motors using recurrence quantification analysis and lstm with weighted bn, Shock Vib. 2019 (2019) 1–14.
[159] J. Lei, C. Liu, D. Jiang, Fault diagnosis of wind turbine based on long short-term memory networks, Renew. Energy 133 (2019) 422–432.
[160] A.P. Dempster, Upper and Lower Probabilities Induced by a Multivalued Mapping, Classic Works of the Dempster-Shafer Theory of Belief Functions, Springer, 2008, pp. 57–72.
[161] G. Shafer, A mathematical theory of evidence, 42 Princeton university press, 1976.
[162] H. Cui, Q. Liu, J. Zhang, B. Kang, An improved deng entropy and its application in pattern recognition, IEEE Access 7 (2019) 18284–18292.
[163] J. An, M. Hu, L. Fu, J. Zhan, A novel fuzzy approach for combining uncertain conflict evidences in the dempster-shafer theory, IEEE Access 7 (2019) 7481–7501.
[164] J. Hu, T. Huang, J. Zhou, J. Zeng, Electronic systems diagnosis fault in gasoline engines based on multi-information fusion, Sensors 18 (9) (2018) 2917.
[165] Q. Pan, D. Zhou, Y. Tang, X. Li, J. Huang, A novel belief entropy for measuring uncertainty in dempster-shafer evidence theory framework based on plausibility transformation and weighted hartley entropy, Entropy 21 (2) (2019) 163.
[166] R. Jiroušek, P.P. Shenoy, A new definition of entropy of belief functions in the dempster–shafer theory, Int. J. Approximate Reasoning 92 (2018) 49–65.
[167] L. Pan, Y. Deng, A new belief entropy to measure uncertainty of basic probability assignments based on belief function and plausibility function, Entropy 20 (11) (2018) 842.
[168] Y. Deng, Deng entropy, Chaos, Solitons & Fractals 91 (2016) 549–553.
[169] A.-L. Jousselme, C. Liu, D. Grenier, É. Bossé, Measuring ambiguity in the evidence theory, IEEE Trans. Syst., Man, and Cybern. 36 (5) (2006) 890–903.
[170] L. Chen, L. Diao, J. Sang, A novel weighted evidence combination rule based on improved entropy function with a diagnosis application, Int. J. Distrib. Sens. Netw. 15 (1) (2019). 1550147718823990
[171] Y. Tang, D. Zhou, S. Xu, Z. He, A weighted belief entropy-based uncertainty measure for multi-sensor data fusion, Sensors 17 (4) (2017) 928.
[172] D. Zhou, Y. Tang, W. Jiang, An improved belief entropy and its application in decision-making, Complexity 2017 (2017) 1–15.
[173] A.-L. Jousselme, D. Grenier, É. Bossé, A new distance between two bodies of evidence, Inf. fusion 2 (2) (2001) 91–101.
[174] Z. Wang, F. Xiao, An improved multisensor data fusion method and its application in fault diagnosis, IEEE Access 7 (2018) 3928–3937.
[175] Y. Tang, D. Zhou, M. Zhuang, X. Fang, C. Xie, An improved evidential-iowa sensor data fusion approach in fault diagnosis, Sensors 17 (9) (2017) 2143.
[176] K. Yuan, F. Xiao, L. Fei, B. Kang, Y. Deng, Conflict management based on belief function entropy in sensor fusion, SpringerPlus 5 (1) (2016) 638.
[177] K. Yuan, F. Xiao, L. Fei, B. Kang, Y. Deng, Modeling sensor reliability in fault diagnosis based on evidence theory, Sensors 16 (1) (2016) 113.
[178] F. Xiao, B. Qin, A weighted combination method for conflicting evidence in multi-sensor data fusion, Sensors 18 (5) (2018) 1487.
[179] W. Jiang, B. Wei, X. Qin, J. Zhan, Y. Tang, Sensor data fusion based on a new conflict measure, Math. Probl. Eng. 2016 (2016) 1–11.
[180] F. Xiao, Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy, Inf. Fusion 46 (2019) 23–32.
[181] W. Jiang, W. Hu, C. Xie, A new engine fault diagnosis method based on multi-sensor data fusion, Appl. Sciences 7 (3) (2017) 280.
[182] B. Qin, F. Xiao, An improved method to determine basic probability assignment with interval number and its application in classification, Int. J. Distrib. Sens. Netw. 15 (1) (2019). 1550147718820524
[183] R. Sun, Y. Deng, A new method to identify incomplete frame of discernment in evidence theory, IEEE Access 7 (2019) 15547–15555.
[184] Y. Lin, C. Wang, C. Ma, Z. Dou, X. Ma, A new combination method for multisensor conflict information, J. Supercomput. 72 (7) (2016) 2874–2890.
[185] F. Ye, J. Chen, Y. Li, Improvement of ds evidence theory for multi-sensor conflicting information, Symmetry (Basel) 9 (5) (2017) 69.
[186] S. Ao, J. He, Z. Peng, A health diagnosis model for sluices based on the improved evidence combination algorithm, IOP Conference Series: Earth and Environmental Science, 189 IOP Publishing, 2018, p. 022069.
[187] J. Liu, A. Chen, N. Zhao, An intelligent fault diagnosis method for bogie bearings of metro vehicles based on weighted improved ds evidence theory, Energies 11 (1) (2018) 232.
[188] A. Boudiaf, A. Moussaoui, A. Dahane, I. Atoui, A comparative study of various methods of bearing faults diagnosis using the case western reserve university data, J. Fail. Anal. Prev. 16 (2) (2016) 271–284.
[189] K.H. Hui, C.S. Ooi, M.H. Lim, M.S. Leong, A hybrid artificial neural network with dempster-shafer theory for automated bearing fault diagnosis. J. Vibroengineering 18 (7) (2016) 4409–4418.